Qiuzhen Lectures on Functional Analysis

BIN GUI

YAU MATHEMATICAL SCIENCES CENTER, TSINGHUA UNIVERSITY.

binguimath@gmail.com   bingui@tsinghua.edu.cn

December 29, 2025

Last page of TOC

Contents

0 Preface

0.1 前言

> 欧几里得方法论已形成一种特定的强制性表述风格，我将其称为“演绎主义风格”。它从一系列精心陈述的公理、引理和/或定义开始，这些公理和定义往往显得人为且晦涩复杂，却无人解释这些复杂性从何而来……遵循欧几里得仪式，数学学习者被迫目睹这场魔术表演，既不得追问背景，也不得探究手法奥秘。倘若学生偶然发现某些生硬的定义实为证明过程所催生，倘若他仅仅疑惑这些定义、引理与定理如何可能先于证明存在，魔术师便会以“数学不成熟”为由将他排斥在外。[1]
>
> — Imre Lakatos [Lak63, Appendix 2]

一直以来，我都有写作泛函分析教材的心愿。正逢 2025 年秋季学期，我有幸接手清华大学求真书院的 64 课时泛函分析本科课程的教授任务，本讲义（教材）便是在这个契机下完成的。得益于求真书院给予/默许的教学自由度，以及书院学生极佳的数学资质，本讲义和课程能够大致遵照我自己的意愿得以设计。与过往的讲义写作经历（例如 [Gui-A]）相比，本讲义中所展现的教学理念更加完整，主题也更为统一。

按照 Zeidler 在“应用泛函分析”一书 [Zei] 前言中的提法，有两种教授数学的方式：体系化的与应用导向的。具体到泛函分析，这两种类型都有许多优秀的教材。本讲义尝试的是第三种方式：**历史的**教学方式。采纳这种方式并非意在放弃体系化构建泛函分析理论，也并非要放弃展示泛函分析的应用。恰恰相反，我试图通过历史方法来沟通抽象理论和应用之间的鸿沟。

这种理论和应用之间的鸿沟长期为人们所忽略。人们似乎以为：通过把抽象定义和定理运用在足够多的场合中，给予足够多的应用语境，人们便能更好地理解定义和定理。这么做或许在别的课程中有用，但在泛函分析中，知道定理的应用往往无益于理解定理的证明为何如此这般布局。作为这一失败的后果，哪怕是注重应用导向的教学方式，在不可避免地涉及理论构建的地方，引导读者的动机仍然和偏向体系化的教材一样，以“**点集拓扑化的几何线性代数**”为主导。

所谓几何线性代数，便是以线性映射和线性子空间（而非以行列式和矩阵等有具体表达式的对象）为基本语言的线性代数。在历史上，线性代数的几何化

[1] 原文：Euclidean methodology has developed a certain obligatory style of presentation. I shall refer to this as ‘deductivist style’. This style starts with a painstakingly stated list of axioms, lemmas and/or definitions. The axioms and definitions frequently look artificial and mystifyingly complicated. One is never told how these complications arose...The student of mathematics is obliged, according to the Euclidean ritual, to attend this conjuring act without asking questions either about the background or about how this sleight-of-hand is performed. If the student by chance discovers that some of the unseemly definitions are proofgenerated, if he simply wonders how these definitions, lemmas and the theorem can possibly precede the proof, the conjuror will ostracize him for this display of mathematical immaturity

本身就是泛函分析的产物，是泛函分析反哺线性代数的成果。而当几何化的线性代数融入了点集拓扑（从而融入了分析学），呈现在人们面前的便是"闭线性子空间"、"连续/有界线性映射"，"闭算子"，"有界线性扩张"，"弱紧性/弱*紧性"，"关于闭线性子空间的商空间"，"有界线性泛函构成的对偶空间"等一系列概念。

从某种意义上讲，"点集拓扑化的几何线性代数"是泛函分析有别于其它数学方向的最鲜明的特征。然而，以这种范式为动机去解释理论构建，而非以促成这种范式的具体数学现象[2]为动机去解释，其实只是转移了困惑而非消解了困惑。以下是一些常见的此类解释方式：

- 研究一个数学对象的好方式通常是研究它上面的函数。因此，我们来研究赋范空间上的有界线性泛函。
- 既然引入了赋范空间的概念，那么让我们来研究一下，有限维的赋范空间如何刻画。
- 欧氏空间上有 Heine-Borel 定理和 Bolzano-Weierstrass 定理等关于紧性的定理。它们在无限维空间上是否成立?
- 既然通常无限维赋范空间的单位闭球不是紧的，我们能否找到更弱一点的拓扑，使得紧性成立?
- 有限维线性代数中，kernel 的维数等于 range 的余维数。无限维情况下什么样的有界线性算子也满足这个条件?
- 无限维情况下，哪种线性算子的性质和有限维情形下的算子比较接近?
- 有限维情形下的实对称矩阵和复 Hermite 矩阵的对角化理论是否可以推广到无限维?
- 我们需要一种连续版本的对角化理论。因此，我们需要在适当的地方引入测度论技术。
- 有限维线性代数中，矩阵求逆是基本问题。因此，对于无限维赋范空间上的有界算子 T，我们关心使得 $T - \lambda$ 可逆的标量 λ。
- 如果要放弃要求线性算子是有界/连续的，那么"闭算子"看起来像是一个合适的较弱的分析条件。
- 一个有界线性映射是双射，那么自然的问题是：它的逆映射是否有界/连续?

[2] 关于"数学现象"的更多讨论见附录第 A.6 节。

这一类解释方式看似给定义和定理提供了动机，实际上却难以为相关定理的证明思想提供良好的辩护。更麻烦的问题在于：只要人们稍加细想便会意识到，在泛函分析中有无数种类似上面的提问方式，但这些提问方式和动机解释都不引向真正有意义的研究。真正有效的定义和真正深刻的提问方式都是依赖于对数学现象的深切关怀的。

很遗憾，如今我们常见的"泛函分析之应用"也并不完全是促使泛函分析之研究范式形成的数学现象。许多人都知道，偏微分方程对泛函分析的起源和成长起到了核心作用。但是现代偏微分方程的理论立足于分布理论和 Sobolev 空间理论，它和 19 世纪末 20 世纪初通过积分方程来研究微分方程的方法有极大差异。然而， Hilbert 空间与紧算子理论的基本形态和方法论却是在积分方程的语境下形成的。倘若我们不知道偏微分方程问题是如何转化为积分方程问题的，那么积分算子便只不过是紧算子的一个例子——哪怕学了这个例子，学生也并不知道为何要关心积分算子和积分方程。倘若不知道研究偏微分方程的经典方法（即积分方程法），那么任何试图用积分方程来启发学生学习紧算子理论的尝试都会遇到如下困境：紧算子的部分意义作为"例子"被转移到积分方程，而积分方程本身的意义被悬置了。可是难道仅仅因为积分算子看起来比较具体，它就有资格成为"动机"吗？它和我们随手写下来的算子相比重要在哪里？在现代教科书中，无论应用还是例子，都不必然是理论背后的数学现象。

开头说过，本讲义以历史为核心教学理念，试图以历史方法来为泛函分析理论提供更好的动机解释。我对"**更好的解释**"是以科研为标准的：我尝试提供一种历史语境，在这种语境中，读者对为何引入这一定义、为何期待这一定理、为何如此布局定理证明、为何走这一条路而非另一条路能够有更好的共情。我对何为"自然的解释"的标准是：能够启发读者发现这一定义、定理、以及证明的解释，而非给它们一个事后合理化的解释。甚至，借助历史方法，我希望触及更加底层的东西：我希望"自然的解释"能够使读者对观念和范式的转换感同身受——因为深刻的数学研究不只是在一个观念下发现了定义、定理、或是证明，而是自身成为观念和范式转化的承载者。

这也便意味着：通常教材中以之为天经地义的范式和观念——例如完备性，例如线性扩张——在本讲义中会接受历史批判。我们不曾质疑这些观念，我们甚至用这些观念来解释别的概念，而不是将这些观念当作待解释的对象，这是因为我们的本科数学教育一开始便在为它们做准备、在学习泛函分析之前我们便和它们打过许多照面。甚至高中数学便会学的"映射"概念，都是在为学生熟悉这些观念铺路。然而，光凭范式本身所许诺的东西是无法准确评估范式的生命力的。当相同的数学物理现象能够同时被泛函分析，Lie 理论，微分几何，代数几何等不同的方法研究的时候，只有通过泛函分析的历史——通过它曾经成功解决过的问题，通过范例性的解决模式——才能评估泛函分析自身的特质、潜力、未来、以及与其它方法相比的优势。

本讲义 5.4 节提到的泛函分析中的三个范式转换，串起了讲义所涉及的各个题

材。因此，这三个范式转换构成了讲义的一条核心主线：

(a) 从有限逼近到线性扩张的范式转换；

(b) 从双线性型到线性映射/线性算子的范式转换；

(c) 从对偶性到完备性的范式转换。

关于这三条范式转换，正文多处有详细讨论，此处主要对 (a) 作大致介绍。

我以矩问题为例。矩问题是贯穿本讲义的数学现象，因此也是本讲义的另一条核心主线。大致来说，**矩问题**问的是给定一列函数 (ξ_n) 和一个数列，存在函数（或测度）f 满足

$$\int \xi_n f = c_n \qquad \text{或} \qquad \int \xi_n df = c_n \qquad \forall n$$

的充分必要条件是什么。最重要的两类矩问题是 ξ_n 是三角函数的情况和多项式的情况。前者即是 Fourier 级数理论。求解三角矩问题，便是寻找 Fourier 系数为给定数列的函数。而在 ξ_n 是（不一定有界的）区间 I 上的多项式 x^n 的情况下，矩问题的解 f 作为 I 上的递增函数/Borel 测度给出的全纯函数

$$z \in \mathbb{C}\backslash I \mapsto \int_I \frac{df(x)}{z-x} \tag{0.1}$$

便是发散级数 $\sum_n c_n/z^{n+1}$ 的一个解析表达式；因此，解多项式矩问题便是找出级数表示为 $\sum_n c_n/z^{n+1}$ 的全纯函数。

现代泛函分析经常把矩问题翻译为如下形式：令 V 为一个包含 (ξ_n) 的函数空间，且 (ξ_n) 在 V 中稠密，从而有唯一的有界线性泛函/正线性泛函满足

$$\Lambda : V \to \mathbb{C} \qquad \xi_n \mapsto c_n$$

只要我们能把 Λ 实现为关于某个函数/测度 f 的积分，那么 f 就给出了原矩问题的解。泛函分析中对线性泛函和对偶空间的刻画皆来源于矩问题。按照**线性扩张范式**，矩问题的解决方式为：找到一个比 V 更大的函数空间 W，其上的线性泛函更容易刻画为关于函数或测度的积分；因此，刻画 V 上的线性泛函的问题便转化为了证明 V 上的线性泛函能够扩张到 W 上的问题。

经典分析学对矩问题的处理采取**有限逼近范式**。其核心思想为：找到一列函数 (f_k) 作为矩问题的逼近解，即

$$\int \xi_n f_k = c_n \qquad \text{或} \qquad \int \xi_n df_k = c_n \qquad \forall |n| \leqslant k$$

然后对 (f_k) 或者它的 某个子列取极限，收敛到的函数 f 给出了原矩问题的解。按照现代表述，这种收敛性实为弱*收敛。但在更偏向经典分析学的视角下，解矩问

题涉及到的收敛性常常可以表现为逐点收敛或几乎处处收敛。因此，在有限逼近范式下，函数逼近论处于核心地位——在三角矩问题中，相关的逼近论问题是一个函数可否由它的 **Fourier 级数**逼近；在多项式矩问题中，相应的逼近论问题则是发散级数 $\sum_n c_n/z^{n+1}$ 及其解析式 (0.1) 被对应的**连分数**逼近。

倘若没有见识过这种范式比较，现代读者在看到 Hahn-Banach 定理的证明时、在看到 $C[a,b]$ 上正线性泛函的 Riesz 表示定理证明时，或许会以为线性扩张法是最自然的理解方式。但阅读了本讲义的读者应该能明白：线性扩张绝不是唯一自然的理解方式，它甚至在某些地方还不如有限逼近范式——在线性扩张的范式中，函数逼近这一主题没有容身之处。

那么，从有限逼近到线性扩张的范式转换为何得以发生？当时的数学家、以及之后的教科书作者、现代的科研人员，为何采纳后一范式作为泛函分析的核心？是什么让人们宁愿隔断与经典分析的联系，也要拥抱线性扩张范式？线性扩张范式和有限逼近范式相比，究竟厉害在哪？对这些问题的探讨占据了本讲义相当多的篇幅；甚至在纯粹构建数学理论的地方，讲义对理论体系的呈现方式也是以这些问题为关怀的。并且读者会发现，这一范式转换 (a) 与前面提到的另外两个范式转换 (b) 和 (c) 紧密相连。

例如，读者会读到：通过所谓 GNS 构造[3]（其原型为多项式矩问题中的 Hankel 矩阵构造），多项式矩问题、以及相应的正线性泛函刻画问题，能够转化为对称算子/对称二次型的谱分解问题。而多项式矩问题涉及的有限逼近——即连分数逼近——对应了相应算子/二次型作为无穷矩阵，被它的有限截断给逼近。[4]借此，有限逼近范式成为了谱理论的早期范式。

在谱理论中，范式转换 (a) 最惊人的例子，便是 von Neumann 的无界算子扩张理论替代了继承自 Hilbert 传统的有限矩阵逼近无限矩阵/二次型——借助 Cayley 变换，von Neumann 把无界算子扩张的问题约化为（有界）等距映射的扩张问题。这一方法毫无疑问立足于线性映射视角而非双线性型视角，因为只有借助于前者，Cayley 变换与等距映射的扩张才能得到自然和精简的表述。与此相对，Hilbert 学派的有限逼近方法依赖于双线性型范式：双线性型无法直接定义乘法和预解式（resolvent），因为没有映射复合和逆映射的直接对应物。因此，要定义双线性型的乘法和预解式，只能借助于对其有限截断的乘法和预解式（表达为有限矩阵的乘法和行列式）取极限。由此可见，von Neumann 的无界算子扩张理论同时是范式转换 (a) 和 (b) 的典型案例。

随着上述三大范式转换得以完成，“点集拓扑化的几何线性代数”这一整体风格得以定型，泛函分析作为一门学科也趋于成熟。从这个意义上讲，本讲义涉及的许多数学题材属于“成熟视角下的未成熟泛函分析”。而我在整个讲义里所做的，无非是在讲述一个传奇故事，讲述“未成熟泛函分析的成长史”。因此，本讲义虽然常常不以真正的历史顺序为教学顺序，但明白本文所述之历史取向的读者不

[3]Gelfand-Naimark-Segal 构造

[4]多项式矩问题中的有限逼近（同时也体现为连分数的收敛方式）对应于谱理论中无穷矩阵的有限截断逼近，这是我在备课期间最为惊喜的发现之一。

难理解：为什么绝大多数泛函分析教材在开头便会介绍的内容——那些被普遍当作“泛函分析基本定理”的结论——本讲义会留到最后再做讨论。

实际上，随着本讲义趋于末尾——在第九章后半和第十章——读者会发现：“点集拓扑化的几何线性代数”越来越成为动机本身，而不是成为待解释的东西。这是所处理之对象的本性使然：讲义最后呈现的，本就是以 Banach、Mazur、Saks、Schauder 等人为代表的 Lwów 学派对 Banach 空间进行系统化和抽象化研究的结果。或许我们很难推测出：如果没有 Lwów 学派，泛函分析的风格会和如今所见相差多远。但毫无疑问，即使这些数学家在研究中更为显著地以“点集拓扑化的几何线性代数”的风格提问，这种风格对他们而言也是有历史厚度的。毕竟，Banach 发表他的著名专著 [Ban32] 的 1932 年，距离 Hilbert 在 [Hil06] 中为了给积分方程提供更加概念化的理解方式而提出 Hilbert space 的 1906 年，也就差了不到 30 年。

作为一门极致依赖于演绎推理的学科，数学给人的感觉通常与“不变”和“永恒”挂钩。作为一门追求永垂不朽之真理的学科，数学的静态特质也深深影响了数学教学和数学学习。借由历史方法，本讲义尝试让读者感受到数学的另一面：感受到数学理论的动态、竞争、不确定性，如同任何一种富有创造力的人类活动一样。我期待这本讲义能够给后人的教学和教材写作提供新的方向。我期待不同的写作风格在未来可以有更多的对话。我也期待，未来人们能够更好地理解：在不同的数学方向和数学主题下，历史风格的写作和教学有哪些属于这个主题自身的特质。[5]

> Some working mathematicians who do not like logicians, philosophers and other cranks interfering in their work, usually say that the introduction of heuristic style would require the rewriting of textbooks, and would make them so long that one could never read them to the end. Papers would become much longer too. The answer to this pedestrian argument is: let us try.
>
> — Imre Lakatos [Lak63, Appendix 2]

2025 年 12 月于北京

[5] 关于数学教育理念的更多讨论见附录 A。

1 Preliminaries

1.1 Notation

Most of the problems in this monograph are not used in the main text. Those that are used are marked with a †.

In this monograph, unless otherwise stated, we understand the field $\mathbb{F}$ as either $\mathbb{R}$ or $\mathbb{C}$.

We use frequently the abbreviations:

iff=if and only if

LHS=left hand side RHS=right hand side

$\exists$=there exists $\forall$=for all

i.e.=id est=that is=namely e.g.=for example

cf.=compare/check/see/you are referred to

resp.=respectively WLOG=without loss of generality

LCH=locally compact Hausdorff

MCT=monotone convergence theorem

DCT=dominated convergence theorem

When we write $A := B$ or $A \xlongequal{\text{def}} B$, we mean that A is defined by the expression B. When we write $A \equiv B$, we mean that A are B are different symbols of the same object.

Unless otherwise stated, an inner product space V denotes a complex inner product space, and its sesquilinear form $\langle\cdot|\cdot\rangle$ is linear on the right argument $|\cdot\rangle$ and antilinear on the left argument $\langle\cdot|$. Note that this convention is different from that of [Gui-A], where the right variable is antilinear.

If V is an $\mathbb{F}$-vector space, then for each $v \in V$ and each linear map $\varphi : V \to \mathbb{F}$, we write

$$\langle v, \varphi\rangle = \langle \varphi, v\rangle := \varphi(v)$$

We assume $a\cdot(+\infty) = (+\infty)\cdot a = +\infty$ if $a \in (0, +\infty]$, and $0\cdot(+\infty) = (+\infty)\cdot 0 = 0$.

An increasing function/sequence/net means a non-decreasing one, that is, $x \leqslant y \Rightarrow f(x) \leqslant f(y)$.

Unless otherwise stated, when mentioning a function space, i.e., a linear subspace of $\mathbb{F}^X$ where X is a set, we assume that X is non-empty.

- Unless otherwise specified, completeness of a metric space or normed vector space refers to Cauchy completeness.
- $\mathbb{N} = \{0, 1, 2, \dots\}$, $\mathbb{Z}_+ = \{1, 2, \dots\}$.

- $\mathbb{R}_{\geqslant 0} = [0, +\infty)$, $\overline{\mathbb{R}}_{\geqslant 0} = [0, +\infty]$, $\overline{\mathbb{R}} = [-\infty, +\infty]$. We equip $\overline{\mathbb{R}}$ with the topology generated by all $(a, b), (a, +\infty], [-\infty, b)$ where $a, b \in \mathbb{R}$. The space $\overline{\mathbb{R}}$ is a compact and Hausdorff.
- An **interval** denotes a connected subset of $\overline{\mathbb{R}}$. A **proper interval** denotes an interval with non-zero Lebesgue measure.
- Y^X is the set of functions with domain X and codomain Y.
- 2^X is the set of subsets of X.
- $\mathrm{fin}(2^X)$ is the set of finite subsets of X.
- For each vector space V, we let $V[x_1, \ldots, x_k]$ be the space of polynomials of the (mutually commuting) abstract variables $x_1, \ldots, x_k$ with coefficients in V. Therefore, its elements are of the form

$$\sum_{n_1,\ldots,n_k \in E} v_{n_1,\ldots,n_k} x_1^{n_1} \cdots x_k^{n_k} \qquad \text{where } E \in \mathrm{fin}(2^{\mathbb{N}}) \text{ and } v_{n_1,\ldots,n_k} \in V$$

- If $f : X \to Y$ is a map, then

$$\mathrm{Rng}(f) = f(X)$$

If X, Y are vector spaces and f is linear, then

$$\mathrm{Ker}(f) = f^{-1}(0)$$

- If V is a vector space and X is a set, then V^X is viewed as a vector space whose linear structure is defined by

$$(af + bg)(x) = af(x) + bg(x) \qquad \text{for all } f, g \in V^X \text{ and } a, b \in \mathbb{F}$$

- If X is a set and $A \subset X$, the **characteristic function** is

$$\chi_A : X \to \{0, 1\} \qquad x \mapsto \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \in X \backslash A \end{cases}$$

- If X is a metric space and $p \in X, r \in [0, +\infty]$, we let

$$B_X(p, r) = \{x \in X : d(x, p) < r\} \qquad \overline{B}_X(p, r) = \{x \in X : d(x, p) \leqslant r\}$$

For each $E \subset X$, we define the **diameter**

$$\mathrm{diam}(E) = \sup\{d(x, y) : x, y \in E\}$$

For each $E, F \subset X$, define the **distance**

$$\mathrm{dist}(E, F) = \inf_{x \in E, y \in F} d(x, y) \tag{1.1}$$

- If X is a topological space, then $\mathcal{T}_X$ denotes the topology of X, i.e.,

$$\mathcal{T}_X = \{\text{open subsets of } X\}$$

  If $x \in X$, a **neighborhood** of x denotes an *open* subset of X containing x. We let

$$\mathrm{Nbh}_X(x) \equiv \mathrm{Nbh}(x) := \{\text{neighborhoods of } x \text{ in } X\}$$

- $\mathrm{Cl}_X(A)$, also denoted by $\mathrm{Cl}(A)$ or $\overline{A}$, is the closure of $A \subset X$ with respect to the topological space X.

- $\mathrm{Int}_X(A)$, also denoted by $\mathrm{Int}(A)$, is the interior of $A \subset X$ with respect to the topological space X. In other words, $\mathrm{Int}_X(A)$ consists of all $x \in A$ such that A contains U for some $U \in \mathrm{Nbh}_X(x)$.

- If X, Y are topological spaces, then

$$\begin{gathered} C(X,Y) = \{f \in Y^X : f \text{ is continuous}\} \\ \mathfrak{B}_X = \text{the Borel } \sigma\text{-algebra of } X \\ \mathscr{Bor}(X,Y) = \{f \in Y^X : f \text{ is Borel}\} \end{gathered}$$

- m^n, as a measure, denotes the Lebesgue measure on $\mathbb{R}^n$, and is abbreviated to m when no confusion arises.

- $\mathbb{T} = \mathbb{S}^1 = \{z \in \mathbb{C} : |z| = 1\} \simeq \mathbb{R}/2\pi\mathbb{Z}$. If f is a function on $\mathbb{S}^1$, equivalently, a 2π-periodic function on $\mathbb{R}$, then

$$\widehat{f}(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-\mathbf{i}nx}dm(x)$$

  is its n-th Fourier coefficient (whenever the integral can be defined).

- $(X, \mathfrak{M}, \mu)$, often abbreviated to (X, μ), denotes a measure space where $\mathfrak{M}$ is the σ-algebra and $\mu : \mathfrak{M} \to \overline{\mathbb{R}}_{\geqslant 0}$ is the measure.

- Let V be a normed vector space. Let X be either a set or a topological space, depending on the context. Let $1 \leqslant p < +\infty$. For each $f \in V^X$,

$$\begin{gathered} \mathrm{Supp}_X(f) \equiv \mathrm{Supp}(f) = \mathrm{Cl}_X(\{x \in X : f(x) \neq 0\}) \\ \|f\|_{l^\infty(X,V)} = \|f\|_{l^\infty} = \sup_{x\in X} \|f(x)\| \\ \|f\|_{l^p(X,V)} = \|f\|_{l^p} = \Big(\sum_{x\in X} \|f(x)\|^p\Big)^{\frac{1}{p}} \\ |f| \text{ is the function } X \to \mathbb{R}_{\geqslant 0} \text{ such that } |f|(x) = \|f(x)\| \end{gathered}$$

We call $|f|$ the **absolute value function** of f. For each $E \subset V$, we let

$$
\begin{aligned}
C_c(X,E) &= \{f \in C(X,E) : \mathrm{Supp}(f) \text{ is compact in } X\} \\
l^\infty(X,V) &= \{f \in V^X : \|f\|_\infty < +\infty\} \\
l^p(X,V) &= \{f \in V^X : \|f\|_p < +\infty\} \\
\mathscr{Bor}_{\mathrm{b}}(X,V) &= \mathscr{Bor}(X,V) \cap l^\infty(X,V) = \{f \in V^X : f \text{ is Borel and bounded}\} \\
C_{\mathrm{b}}(X,V) &= C(X,V) \cap l^\infty(X,V) = \{\text{bounded continuous } f : X \to V\}
\end{aligned}
$$

We are particularly interested in the case that $E = V$, $E = [0,1]$, and $E = \mathbb{R}_{\geqslant 0}$.

- Let $k \in \mathbb{N} \cup \{\infty\}$. Suppose that X is a subset of $\mathbb{R}^m$ (or more generally, a subset of a C^k-manifold M with or without boundary). Let $Y \subset \mathbb{R}^n$. We let

$$
\begin{aligned}
C^k(X,\mathbb{R}^n) &= \{C^k\text{-functions } X \to \mathbb{R}^n\} \\
C^k(X,Y) &= \{f \in C^k(X,\mathbb{R}^n) : f(X) \subset Y\} \\
C_c^k(X,Y) &= \{f \in C^k(X,Y) : \mathrm{Supp}_X(f) \text{ is compact}\}
\end{aligned}
$$

  Here, a function $f : X \to \mathbb{R}^m$ is called a ***C^k*-function** if for each $x \in X$ there exists $U \in \mathrm{Nbh}_M(x)$ such that $f|_{U \cap X}$ can be extended to a C^k-function $U \to \mathbb{R}^n$. A C^∞-function is called a **smooth function**.

- Unless otherwise stated, if $f : X \to \mathbb{C}$ where X is a set, we let

$$
f^* : X \to \mathbb{C} \qquad f^*(x) = \overline{f(x)}
$$

  Thus $f = f^*$ iff f is real-valued.

- Let V be a normed vector space. Let X be a set. We say that a family $(f_\alpha)_{\alpha \in \mathscr{A}}$ in V^X is **uniformly bounded** if $\sup_{\alpha \in \mathscr{A}} \|f_\alpha\|_{l^\infty(X,V)} < +\infty$.

- If X is LCH and V is a normed $\mathbb{F}$-vector space, we understand $C_c(X,V)$ as a normed $\mathbb{F}$-vector space whose linear structure inherits from that of V^X, and whose norm is chosen to be the l^∞-norm.

- If $(X,\mathfrak{M})$ and $(Y,\mathfrak{N})$ are measurable spaces, then

$$
\mathcal{L}(X,Y) = \{\text{measurable functions } X \to Y\}
$$

  If V is a normed vector space, for each $f \in \mathcal{L}(X,V)$ and $1 \leqslant p < +\infty$, we let

$$
\|f\|_{L^p(X,\mu)} = \|f\|_{L^p} = \Big(\int_X |f|^p d\mu\Big)^{\frac{1}{p}}
$$
$$
\|f\|_{L^\infty(X,\mu)} = \|f\|_{L^\infty} = \inf\{\lambda \in \overline{\mathbb{R}}_{\geqslant 0} : \mu\{x \in X : \|f(x)\| > a\} = 0\}
$$

  which are potentially infinite.

- If V, W are $\mathbb{F}$-linear maps, we let

$$\mathrm{Lin}(V,W) = \{\mathbb{F}\text{-linear maps } V \to W\} \qquad \mathrm{Lin}(V) = \mathrm{Lin}(V,V)$$

If $A, B \in \mathrm{Lin}(V)$, we let

$$[A,B] = AB - BA$$

- If $(V_\alpha)_{\alpha\in\mathscr{A}}$ is a family of linear subspaces of a vector space V, we let

$$\sum_{\alpha\in\mathscr{A}} V_\alpha = \mathrm{Span}\{\xi \in V_\alpha \text{ where } \alpha \in \mathscr{A}\}$$

- In the notation of function spaces, the codomain is understood to be $\mathbb{C}$ when it is suppressed. For example,

$$C_c(X) = C_c(X,\mathbb{C}) \qquad \mathscr{Bor}(X) = \mathscr{Bor}(X,\mathbb{C}) \qquad L^p(X,\mu) = L^p(X,\mu,\mathbb{C})$$

However, this convention does not apply to $\mathfrak{L}(V)$ and $\mathrm{Lin}(V)$: If V is a normed vector space, then $\mathfrak{L}(V)$ denotes $\mathfrak{L}(V,V)$, the space of bounded linear operators on V. Likewise, if V is a vector space, then $\mathrm{Lin}(V)$ denotes $\mathrm{Lin}(V,V)$.

- Given an abstract set X without prescribed topology, we understand 2^X as the topology of X. In this context, for instance, $C_c(X)$ is the set of functions $f : X \to \mathbb{C}$ such that

$$\mathrm{Supp}_X(f) = \{x \in X : f(x) \neq 0\} \tag{1.2}$$

is a finite set.

1.2 Nets

In this section, we present results on nets that are essential for the topics covered in this course. For more details, see [Gui-A].

1.2.1 Basic definitions

Definition 1.2.1. A relation $\leqslant$ on a set I is called a **preorder** if for all $\alpha, \beta, \gamma \in I$, the following are satisfied:

- (Reflexivity) $\alpha \leqslant \alpha$.

- (Transitivity) If $\alpha \leqslant \beta$ and $\beta \leqslant \gamma$ then $a \leqslant \gamma$.

The pair $(I, \leqslant)$ (or simply I) is called a **preordered set**. For each $\beta \in I$, we write

$$I_{\geqslant\beta} = \{\alpha \in I : \alpha \geqslant \beta\} \tag{1.3}$$

Definition 1.2.2. A preordered set $(I, \leqslant)$ is called a **directed set** if

$$\forall \alpha, \beta \in I \quad \exists \gamma \in I \quad \text{such that } \alpha \leqslant \gamma, \beta \leqslant \gamma \tag{1.4}$$

If I is a directed set and X is a set, then a function $x : I \to X$ is called a **net** with directed set/index set I. We often write $x(\alpha)$ as x_α if $\alpha \in I$, and write x as $(x_\alpha)_{\alpha \in I}$.

Unless otherwise stated, for any net $(x_\alpha)_{\alpha \in I}$ we assume that $I \neq \varnothing$. □

Example 1.2.3. $(\mathbb{Z}_+, \leqslant)$ is a directed set. A net with index set $\mathbb{Z}_+$ in a set X is precisely a sequence in X.

Example 1.2.4. Let X be a topological space and $x \in X$. Then $\mathrm{Nbh}_X(x)$, together with the preorder $\supset$ (that is, $U \leqslant V$ iff $U \supset V$), is a directed set. Unless otherwise stated, the preorder on $\mathrm{Nbh}_X(x)$ is always chosen to be $\supset$.

Definition 1.2.5. Suppose that $(I, \leqslant_I)$ and $(J, \leqslant_J)$ are preordered sets (resp. directed sets), then the **product** $(I \times J, \leqslant)$ is a preordered set (resp. directed set) if for every $\alpha, \alpha' \in I, \beta, \beta' \in J$ we define

$$(\alpha, \beta) \leqslant (\alpha', \beta') \qquad \Longleftrightarrow \qquad \alpha \leqslant_I \alpha' \text{ and } \beta \leqslant_J \beta' \tag{1.5}$$

Unless otherwise stated, the preorder on $I \times J$ is assumed to be defined by (1.5).

Definition 1.2.6. If X is a set, then $(2^X, \subset)$ and $(\mathrm{fin}(2^X), \subset)$ are directed sets where

$$\mathrm{fin}(2^X) = \{A \subset X : A \text{ is a finite set}\} \tag{1.6}$$

Definition 1.2.7. Let P be a property about elements of a set X, i.e., P is a function $X \to \{\text{true, false}\}$. Let $(x_\alpha)_{\alpha \in I}$ be a net in X.

We say that x_α **eventually** satisfies P (equivalently, we say that x_α satisfies P for **sufficiently large** α) if:

- There exists $\beta \in I$ such that for every $\alpha \in I_{\geqslant\beta}$, the element x_α satisfies P.

"Sufficiently large" is also called "**large enough**".

We say that x_α **frequently** satisfies P if:

- For every $\beta \in I$ there exists $\alpha \in I_{\geqslant\beta}$ such that x_α satisfies P.

□

Remark 1.2.8. Let P and Q be two properties about elements of X. Then

$$\neg(x_\alpha \text{ eventually satisfies } P) \quad = \quad (x_\alpha \text{ frequently satisfies } \neg P)$$

By the crucial condition (1.4) for directed sets, we have

$$\begin{gathered}(x_\alpha \text{ eventually satisfies } P) \wedge (x_\alpha \text{ eventually satisfies } Q)\\ \Downarrow \\ x_\alpha \text{ eventually satisfies } P \wedge Q\end{gathered} \tag{1.7a}$$

We will use (1.7a) very frequently without explicitly mentioning it. Clearly, we also have

$$\begin{gathered}(x_\alpha \text{ eventually satisfies } P) \wedge (x_\alpha \text{ frequently satisfies } Q)\\ \Downarrow \\ x_\alpha \text{ frequently satisfies } P \wedge Q\end{gathered} \tag{1.7b}$$

1.2.2 Nets and topological spaces

Let $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ be topological spaces.

Definition 1.2.9. Let $(x_\alpha)_{\alpha\in I}$ be a net in X. Let $x \in X$. We say that (x_α) **converges to** x and write

$$\lim_{\alpha\in I} x_\alpha \equiv \lim_\alpha x_\alpha = x$$

or simply $x_\alpha \to x$ if the following statement holds:

- For every $U \in \mathrm{Nbh}_X(x)$, the net (x_α) is eventually in U.

Convention 1.2.10. If I is a directed set, and if $(x_{\alpha_1,\dots,\alpha_n})_{(\alpha_1,\dots,\alpha_n)\in I^n}$ is a net in X (where I^n is equipped with the product preorder, cf. Def. 1.2.5), we

$$\text{abbreviate} \qquad \lim_{(\alpha_1,\dots,\alpha_n)\in I^n} x_{\alpha_1,\dots,\alpha_n} \qquad \text{to} \qquad \lim_{\alpha_1,\dots,\alpha_n\in I} x_{\alpha_1,\dots,\alpha_n}$$

Example 1.2.11. Let $(x_n)_{n\in\mathbb{Z}_+}$ be a sequence in a metric space. Then $(x_n)_{n\in\mathbb{Z}_+}$ is a Cauchy sequence iff $\lim_{m,n\in\mathbb{Z}_+} d(x_m, x_n) = 0$.

Proposition 1.2.12. *Let $(x_\alpha)_{\alpha\in I}$ be an increasing (resp. decreasing) net in $\overline{\mathbb{R}}$. Let x be the supremum (resp. infimum) of $\{x_\alpha : \alpha \in I\}$. Then $\lim_\alpha x_\alpha = x$.*

Proof. We address the case that (x_α) is increasing; the other case is similar. Let $S = \{x_\alpha : \alpha \in I\}$. Since $x = \sup S$, for each interval U open in $\overline{\mathbb{R}}$ and containing x, we have $S \cap U \neq \emptyset$, and hence there exists $\alpha \in I$ such that $x_\alpha \in U$. Since the net is increasing, we have $x_\alpha \leqslant x_\beta \leqslant x$ for each $\beta \geqslant \alpha$. Since U is an interval containing x_α, x, it must contain x_β for all $\beta \geqslant \alpha$. This proves that $\lim_\alpha x_\alpha = x$.. □

Definition 1.2.13. A net $(x_\alpha)_{\alpha\in I}$ in a metric space X is called a **Cauchy net** if $\lim_{\alpha,\beta\in I} d(x_\alpha, x_\beta) = 0$.

Proposition 1.2.14. *Let (x_α) be a convergent net in a metric space X. Then (x_α) is a Cauchy net.*

Proof. Let x be the limit of (x_α). Then for each $\varepsilon > 0$, there exists $\gamma \in I$ such that $d(x, x_\alpha) < \varepsilon$ for all $\alpha \geqslant \gamma$. Therefore, for all $\alpha, \beta \geqslant \gamma$ we have $d(x_\alpha, x_\beta) \leqslant d(x_\alpha, x) + d(x, x_\beta) < 2\varepsilon$. $\square$

Recall that a metric space X is called **(Cauchy) complete** if each Cauchy sequence in X converges.

Theorem 1.2.15. *Suppose that X is a complete metric space. Then every Cauchy net in X converges.*

Proof. Let $(x_\alpha)_{\alpha\in I}$ be a Cauchy net in X. Then for each $n \in \mathbb{Z}_+$ there exists $\gamma_n \in I$ such that $d(x_\alpha, x_\beta) < 1/n$ for all $\alpha, \beta \geqslant \gamma_n$. Since I is directed, by successively replacing $\gamma_1, \gamma_2, \ldots$ with larger ones, we may assume that $(\gamma_n)_{n\in\mathbb{Z}_+}$ is increasing. Thus $d(x_{\gamma_m}, x_{\gamma_n}) < 1/m$ whenever $m \leqslant n$, and hence $(x_{\gamma_n})_{n\in\mathbb{Z}_+}$ is a Cauchy sequence in X. Since X is complete, $(x_{\gamma_n})_{n\in\mathbb{Z}_+}$ converges to some $x \in X$.

Now, for each n, and for each $\alpha \geqslant \gamma_n$ and $k \geqslant n$, we have $d(x_\alpha, x_{\gamma_k}) < 1/n$. Since $\lim_k d(x_{\gamma_k}, x) = 0$, we may find $k \geqslant n$ such that $d(x_{\gamma_k}, x) < 1/n$. Therefore $d(x_\alpha, x) < 2/n$ for all $\alpha \geqslant \gamma_n$. This proves $\lim_\alpha x_\alpha = x$. $\square$

Proposition 1.2.16. *X is Hausdorff iff every net in X has at most one limit.*

Proof. First, assume that X is not Hausdorff. Then there exist $x \neq y$ in X such that every neighborhood of x intersects every neighborhood of y. Let $I = \mathrm{Nbh}_X(x) \times \mathrm{Nbh}_X(y)$. For each $\alpha = (U, V) \in I$, by assumption, there exists $x_\alpha \in U \cap V$. Then $(x_\alpha)_{\alpha\in I}$ is a net in X converging to both x and y.

Conversely, assume that X has a net (x_α) converging to distinct points $x, y \in X$. Then for each $U \in \mathrm{Nbh}_X(x)$, (x_α) is eventually in X. Similarly, for each $V \in \mathrm{Nbh}_X(y)$, (x_α) is eventually in Y. By Rem. 1.2.8, (x_α) is eventually in $U \cap V$. In particular, there exists α such that $x_\alpha \in U \cap V$. So $U \cap V \neq \emptyset$. Thus, X is not Hausdorff. $\square$

Proposition 1.2.17. *Let $A \subset X$ and $x \in X$. Then $x \in \mathrm{Cl}_X(A)$ iff there exists a net (x_α) in A such that x is a limit of (x_α).*

Proof. Suppose that $x \in \overline{A}$. Then each $U \in \mathrm{Nbh}_X(x)$ intersects A, and hence there exists $x_U \in U \cap A$. Then $(x_U)_{U\in\mathrm{Nbh}_X(x)}$ is a net in A converging to x.

Conversely, assume that $x \notin \overline{A}$. Then there exists $U \in \mathrm{Nbh}_X(x)$ disjoint from A. Therefore, any net (x_α) in A is never in U, and hence does not converge to x. $\square$

Theorem 1.2.18. *Let $f : X \to Y$ be a map. Let $x \in X$. Then the following are equivalent:*

(1) *f is continuous at x, that is, for each $V \in \mathrm{Nbh}_Y(f(x))$, there exists $U \in \mathrm{Nbh}_X(x)$ that is contained in $f^{-1}(V)$.*

(2) *For each net (x_α) in X converging to x, the net $f(x_\alpha)$ converges to $f(x)$.*

(3) *For each net (x_α) in $X\backslash\{x\}$ converging to x, the net $f(x_\alpha)$ converges to $f(x)$.*

Proof. (1)⇒(2): Assume (1). Let (x_α) be a net in X converging to x. Then for any $V \in \mathrm{Nbh}(f(x))$, by choosing $U \in \mathrm{Nbh}(x)$ contained in $f^{-1}(V)$, we have that (x_α) is eventually in U, and hence $(f(x_\alpha))$ is eventually in V. Therefore $f(x_\alpha) \to f(x)$. This proves (2).

(2)⇒(3): Obvious.

¬(1)⇒¬(3): Since (1) is not true, there exists $V \in \mathrm{Nbh}(f(x))$ such that any $U \in \mathrm{Nbh}(x)$ is not contained in $f^{-1}(V)$. Since $x \in f^{-1}(V)$, the set $U\backslash\{x\}$ is not contained in $f^{-1}(V)$. (In particular, $U\backslash\{x\}$ is not empty.) Therefore, we can choose $x_U \in U\backslash\{x\}$ such that $f(x_U) \notin V$. It follows that $(x_U)_{U\in\mathrm{Nbh}_X(x)}$ is a net in $X\backslash\{x\}$ converging to x that is always outside V. Thus $f(x)$ is not a limit of $(f(x_U))_{U\in\mathrm{Nbh}_X(x)}$, because $f(x_U)$ is always outside V. Therefore, (3) is not true. □

Corollary 1.2.19. *Let $f : X \to Y$ be continuous. Let $A \subset X$. Then*

$$f(\overline{A}) \subset \overline{f(A)}$$

Proof. For each $x \in \overline{A}$, by Prop. 1.2.17, there exists a net (x_α) in A converging to x. By Thm. 1.2.18, $(f(x_\alpha))$ converges to $f(x)$. By Prop. 1.2.17 again, we conclude that $f(x) \in \overline{f(A)}$. □

Corollary 1.2.20. *Let $f : X \to Y$ be a bijection. Then the following are equivalent.*

(1) *f is a homeomorphism.*

(2) *For each $x \in X$ and each net (x_α) in X, we have $\lim_\alpha x_\alpha = x$ iff $\lim_\alpha f(x_\alpha) = f(x)$.*

Proof. This follows immediately from Thm. 1.2.18. □

Corollary 1.2.21. *Let $\mathcal{T}_X$ and $\mathcal{T}'_X$ be two topologies on X. Then the following are equivalent.*

(1) *$\mathcal{T}_X = \mathcal{T}'_X$.*

(2) *For each $x \in X$ and each net (x_α) in X, the net (x_α) converges to x in $\mathcal{T}_X$ iff (x_α) converges to x in $\mathcal{T}'_X$.*

In other words, topologies are determined by net convergence.

Proof. Apply Cor. 1.2.20 to the identity map of X. □

Definition 1.2.22. Recall that a topological space X is called **first-countable** if for each $x \in X$, there exists a sequence $(U_n)_{n\in\mathbb{Z}_+}$ in $\mathrm{Nbh}_X(x)$ such that $\{U_n : n \in \mathbb{Z}_+\}$ is cofinal in $\mathrm{Nbh}_X(x)$ (i.e., for each $V \in \mathrm{Nbh}_X(x)$ there exists n such that $U_n \subset V$). Moreover, once we have such $(U_n)_{n\in\mathbb{Z}_+}$, by replacing U_n with $U_1 \cap \cdots \cap U_n$, we may assume that $U_1 \supset U_2 \supset \cdots$. Therefore, X being first-countable means that for each $x \in X$, the net $(U)_{U\in\mathrm{Nbh}_X(x)}$ has a subnet that is also a sequence.

Remark 1.2.23. In Prop. 1.2.16, Prop. 1.2.17, Thm. 1.2.18, Cor. 1.2.20, and Cor. 1.2.21, if all the topologies involved are assumed to be first-countable, then the statements remain valid when nets are replaced by sequences. We leave the proof to the readers.

For example, Cor. 1.2.21 can be modified as follows: If $\mathcal{T}_X$ and $\mathcal{T}'_X$ are two first-countable topologies on X, then $\mathcal{T}_X = \mathcal{T}'_X$ iff any sequence (x_n) converges to x under $\mathcal{T}_X$ iff (x_n) converges to x under $\mathcal{T}'_X$.

Note that in Thm. 1.2.18, only the domain X needs to be first-countable; the codomain Y does not. □

1.2.3 Subnets

Definition 1.2.24. A subset E of a directed set I is called **cofinal** if:

$$\forall \alpha \in I \quad \exists \beta \in E \quad \text{such that } \alpha \leqslant \beta$$

By the transitivity in Def. 1.2.1 and property (1.3), we clearly have

$$\forall \alpha_1, \dots, \alpha_n \in I \quad \exists \beta \in E \quad \text{such that } \alpha_1 \leqslant \beta, \dots, \alpha_n \leqslant \beta$$

Definition 1.2.25. Let $(x_\alpha)_{\alpha\in I}$ be a net in a set X. A **subnet** of $(x_\alpha)_{\alpha\in I}$ is, by definition, of the form $(x_{\alpha_s})_{s\in S}$ where S is a directed set, and

$$(\alpha_s)_{s\in S} : S \to I \qquad s \mapsto \alpha_s$$

is an increasing function (i.e. $s \leqslant s' \Rightarrow \alpha_s \leqslant \alpha_{s'}$) satisfying one of the following (clearly) equivalent conditions:

(a) The range $\{\alpha_s : s \in S\}$ is cofinal in I.

(b) For each $\beta \in I$, the net $(\alpha_s)_{s\in S}$ is eventually $\geqslant \beta$.

Example 1.2.26. A subsequence of a sequence is a subnet of that sequence.

Proposition 1.2.27. *Let $(x_\alpha)_{\alpha\in I}$ be a net in a topological space X converging to $x \in X$. Then every subnet $(x_{\alpha_s})_{s\in S}$ converges to x.*

Proof. Choose any $U \in \mathrm{Nbh}(x)$. Since $x_\alpha \to x$, there exists $\beta \in I$ such that for all $\alpha \geqslant \beta$ we have $x_\alpha \in U$. Since α_t is eventually $\geqslant \beta$, we see that x_{α_t} is eventually in U. □

Note that a subnet does not necessarily have the same index set as the original net. This definition of subnets is motivated largely by the following important property, which will play a crucial role in Subsec. 1.3.1.

Theorem 1.2.28. *Let $(x_\alpha)_{\alpha\in I}$ be a net in a topological space X. Let $x \in X$. Then the following are equivalent.*

(1) $(x_\alpha)_{\alpha\in I}$ has a subnet converging to x.

(2) For every $U \in \mathrm{Nbh}_X(x)$, the net (x_α) is frequently in U.

(3) x belongs to $\bigcap_{\alpha\in I}\overline{\{x_\beta : \beta \geqslant \alpha\}}$.

Any $x \in X$ satisfying one of these three conditions is called a ***cluster point*** *of $(x_\alpha)_{\alpha\in I}$.*

Proof. (2)⇔(3): (3) holds iff x belongs to $\overline{\{x_\beta : \beta \geqslant \alpha\}}$ for each α, iff each $U \in \mathrm{Nbh}(x)$ intersects $\{x_\beta : \beta \geqslant \alpha\}$ for each α, iff for each $U \in \mathrm{Nbh}(x)$ and each α, there exists $\beta \geqslant \alpha$ such that $x_\beta \in U$, iff (2) holds.

(1)⇒(2): Let (x_{α_s}) be a subnet converging to x. Then for each $U \in \mathrm{Nbh}_X(x)$ and $\beta \in I$, the net (x_{α_s}) is eventually in U, and α_s is eventually $\geqslant \beta$. Therefore, by Rem. 1.2.8, it is eventually true that x_{α_s} is in U and, simultaneously, $\alpha_s \geqslant \beta$. In particular, there exists s such that $x_{\alpha_s} \in U$ and $\alpha_s \geqslant \beta$. This proves that (x_α) is frequently in U.

(2)⇒(1): Assume (2). Define a preordered set $(J, \leqslant)$ by

$$\begin{aligned} J &= \{(\alpha, U) \in I \times \mathrm{Nbh}_X(x) : x_\alpha \in U\} \\ (\alpha, U) \leqslant (\alpha', U') \qquad &\Longleftrightarrow \qquad \alpha \leqslant \alpha' \text{ and } U \supset U' \end{aligned} \tag{1.8}$$

Let us prove that J is directed: Suppose that (α_1, U_1) and (α_2, U_2) belong to J. Since $(\alpha)_{\alpha\in I}$ is eventually $\geqslant \alpha_1$ and eventually $\geqslant \alpha_2$, and since (by (2)) (x_α) is frequently in $U_1 \cap U_2$, by Rem. 1.2.8, it is frequently true that $\alpha \geqslant \alpha_1, \alpha_2$ and $x_\alpha \in U_1 \cap U_2$. Choose $\alpha \in I_{\geqslant\alpha_1} \cap I_{\geqslant\alpha_2}$ such that $x_\alpha \in U_1 \cap U_2$. Then $(\alpha, U_1 \cap U_2)$ belongs to J and is $\geqslant (\alpha_1, U_1)$ and $\geqslant (\alpha_2, U_2)$. This proves that J is directed.

The map $(\alpha, U) \in J \mapsto \alpha \in I$ is clearly increasing; its range is cofinal, since $(\alpha, X) \in J$ for each $\alpha \in I$. Therefore, $(x_{\alpha,U})_{(\alpha,U)\in J}$ is a subnet of (x_α). To prove (1), it remains to show that $(x_{\alpha,U})_{(\alpha,U)\in J}$ converges to x. For each $U \in \mathrm{Nbh}_X(x)$, by (2) there exists $\alpha \in I$ such that $x_\alpha \in U$, and hence $(\alpha, U) \in J$. For each $(\beta, V) \in J_{\geqslant(\alpha,U)}$, we have $x_\beta \in V \subset U$. This verifies $\lim_{(\alpha,U)\in J} x_{\alpha,U} = x$. □

Theorem 1.2.29. *Assume that X is first-countable, and let (x_n) be a sequence in X. Let $x \in X$. Then x is a cluster point of (x_n) iff (x_n) has a subsequence converging to x.*

Proof. We leave the proof to the reader as an exercise. □

1.2.4 First-countable nets

Definition 1.2.30. Let $(I, \leqslant)$ be a directed set. Let ∞_I (often abbreviated to ∞) be a new symbol not in I. Then

$$I^* = I \cup \{\infty_I\}$$

is also a directed set if we extend the preorder $\leqslant$ of I to I^* by setting

$$\alpha \leqslant \infty_I \qquad (\forall \alpha \in I^*)$$

For each $\alpha \in I$, let

$$I^*_{\geqslant \alpha} = \{\beta \in I^* : \beta \geqslant \alpha\}$$

The **standard topology** on I^* is defined to be the one induced by the basis

$$\mathcal{B} = \big\{\{\alpha\} : \alpha \in I\big\} \cup \big\{I^*_{\geqslant \alpha} : \alpha \in I\big\} \tag{1.9}$$

Remark 1.2.31. Suppose that $(x_\alpha)_{\alpha \in I}$ is a net in a topology space X. Let $x_\infty \in X$. Extend (x_α) to a function

$$x : I^* \to X \qquad \alpha \mapsto x_\alpha, \quad \infty \mapsto x_\infty$$

The following facts are easy to check:

1. x is continuous at every point of I.

2. x is continuous at ∞ iff the net $(x_\alpha)_{\alpha \in I}$ converges to x_∞.

In particular, $(x_\alpha)_{\alpha \in I}$ converges to x_∞ iff the function $x : I^* \to X$ is continuous.

Definition 1.2.32. A directed set I is called **first countable** if one of the following equivalent conditions holds:

(1) The standard topology on I^* is first-countable.

(2) I has a countable cofinal subset.

(3) The net $(\alpha)_{\alpha \in I}$ has a subnet which is also a sequence. (In other words, there is an increasing sequence in I converging to ∞.)

If $(x_\alpha)_{\alpha \in I}$ is a net in a set X such that the index set I is first-countable, we also say that $(x_\alpha)_{\alpha \in I}$ is a **first-countable net**.

Proof of equivalence. (1)⇒(2): Assume (1). Then $\mathrm{Nbh}_{I^*}(\infty)$ has a countable cofinal subset, which (due to (1.9)) can be chosen to be of the form $I^*_{\geqslant\alpha_1}, I^*_{\geqslant\alpha_2}, \dots$ where $(\alpha_n)_{n\in\mathbb{Z}_+}$ is a sequence in I. Since $(I^*_{\geqslant\alpha_n})_{n\in\mathbb{Z}_+}$ is cofinal in $\mathrm{Nbh}_{I^*}(\infty)$, for each $\beta\in I$ there exists n such that $I^*_{\geqslant\alpha_n}\subset I^*_{\geqslant\beta}$, and hence $\alpha_n\geqslant\beta$. This proves that $\{\alpha_n : n\in\mathbb{Z}_+\}$ is a countable cofinal subset of I. Hence (2) is proved.

(2)⇒(1): Assume (2). For every $\alpha\in I$, the direct set $\mathrm{Nbh}_{I^*}(\alpha)$ clearly has a countable cofinal subset, namely $\{\{\alpha\}\}$. Let (α_n) be a countable cofinal sequence in I. One checks easily that $(I^*_{\geqslant\alpha_n})_{n\in\mathbb{Z}_+}$ is a cofinal sequence in $\mathrm{Nbh}_{I^*}(\infty)$. This proves (1).

(3)⇒(2): Obvious.

(2)⇒(3): Let $(\alpha_n)_{n\in\mathbb{Z}_+}$ be a cofinal sequence in I. Since I is directed, by increasing $\alpha_2,\alpha_3,\dots$ successively, we can assume that $(\alpha_n)_{n\in\mathbb{Z}_+}$ is increasing. Clearly (α_n) converges to ∞. □

Example 1.2.33. If X is a topological space, then X is first-countable iff $\mathrm{Nbh}_X(x)$ is a first-countable directed set for each $x\in X$.

Example 1.2.34. If I,J are first-countable directed sets, then the directed set $I\times J$ is also first-countable. In particular, double sequences (i.e., nets of the form $(x_{m,n})_{m,n\in\mathbb{Z}_+}$) are first-countable nets.

Theorem 1.2.35. *Let $(x_\alpha)_{\alpha\in I}$ be a first-countable net in a topological space X. Let $x_\infty\in X$. Then the following are equivalent.*

(1) (x_α) converges to x_∞.

(2) For each sequence $(\alpha_n)_{n\in\mathbb{Z}_+}$ in I converging to ∞, the sequence $(x_{\alpha_n})_{n\in\mathbb{Z}_+}$ converges to x_∞.

Proof. This follows immediately from Thm. 1.2.18 and Rem. 1.2.23. □

With the help of Thm. 1.2.35, one can easily generalize the MCT, the DCT, and Fatou's lemma from sequences to first-countable nets of functions. We leave the details of the proofs to the reader.

Theorem 1.2.36 (Monotone convergence theorem (MCT)). *Let (X,μ) be a measure space. Let (f_α) be an increasing first-countable net of measurable functions $X\rightarrow\overline{\mathbb{R}}_{\geqslant 0}$. Let $f : X\rightarrow\overline{\mathbb{R}}_{\geqslant 0}$ be the pointwise limit of (f_α). Then f is measurable, and*

$$\lim_\alpha\int_X f_\alpha d\mu=\int_X f d\mu$$

Theorem 1.2.37 (Dominated convergence theorem (DCT)). *Let (X,μ) be a measure space. Let $(f_\alpha)_{\alpha\in\mathscr{I}}$ be a first-countable net of measurable functions $X\rightarrow\mathbb{C}$. Assume that (f_α) converges pointwise to a function $f : X\rightarrow\mathbb{C}$. Suppose that there exists an integrable function $g : X\rightarrow\overline{\mathbb{R}}_{\geqslant 0}$ such that $|f_\alpha|\leqslant g$ for all α. Then f_α, f are integrable, and*

$$\lim_\alpha\int_X f_\alpha d\mu=\int_X f d\mu$$

Note that the measurability of f follows from the fact that one can choose an increasing sequence $(\alpha_n)_{n\in\mathbb{Z}_+}$ in $\mathscr{I}$ with cofinal range, and hence f is the pointwise limit of the sequence of measurable functions $(f_{\alpha_n})_{n\in\mathbb{Z}_+}$.

Theorem 1.2.38 (Fatou's lemma). *Let (X,μ) be a measure space. Let (f_α) be a first-countable net of measurable functions $X\rightarrow\overline{\mathbb{R}}_{\geqslant 0}$. Assume that (f_α) converges pointwise to a function $f:X\rightarrow\overline{\mathbb{R}}_{\geqslant 0}$. Then f is measurable and*

$$\liminf_{\alpha}\int_X f_\alpha d\mu\geqslant\int_X fd\mu$$

See Def. 1.3.10 for the definition of $\liminf$.

1.2.5 Unordered sum

In this subsection, we fix a normed vector space V over $\mathbb{F}\in\{\mathbb{R},\mathbb{C}\}$. Fix a set X. Recall that $\mathrm{fin}(2^X)$ is a directed set, ordered by the inclusion ($A\leqslant B$ meaning $A\subset B$).

Definition 1.2.39. Let $f:X\rightarrow V$ be a map. The expression

$$\sum_{x\in X}f(x)$$

(or simply $\sum_X f$) is called an **unordered sum**. If $v\in V$, we say that $\sum_{x\in X}f(x)$ equals (or converges to) v, if

$$\lim_{A\in\mathrm{fin}(2^X)}\sum_{x\in A}f(x)=v \tag{1.10}$$

In this case, we write

$$\sum_{x\in X}f(x)=v \tag{1.11}$$

Unwinding the definition of net convergence, (1.10) says that for every $\varepsilon>0$, there exists a finite set $B\subset X$ such that for every finite set A satisfying $B\subset A\subset X$, we have $\|v-\sum_{x\in A}f(x)\|<\varepsilon$.

Remark 1.2.40. If V is complete, then $\sum_X f$ converges precisely when the associated net $(\sum_A f)_{A\in\mathrm{fin}(2^X)}$ satisfies the Cauchy condition. Let us spell out what this Cauchy condition means:

(1) For every $\varepsilon>0$, there exists a finite set $B\subset X$ such that for any finite sets A_1,A_2 satisfying $B\subset A_1\subset X, B\subset A_2\subset X$, we have

$$\Big\|\sum_{A_1\backslash A_2}f-\sum_{A_2\backslash A_1}f\Big\|<\varepsilon$$

Note that the term inside the norm is $\sum_{A_1} f - \sum_{A_2} f$. This is also equivalent to:

(2) For every $\varepsilon > 0$, there exists a finite set $B \subset X$ such that for any finite set $E \subset X \backslash B$, we have

$$\Big\| \sum_E f \Big\| < \varepsilon$$

In practice, we will mainly use (2) as the Cauchy criterion for the convergence of $\sum_X f$.

Proof of the equivalence. (2) follows from (1) by taking $A_1 = B$ and $A_2 = B \cup E$. (1) follows from (2) by taking $E_1 = A_1 \backslash A_2$ and $E_2 = A_2 \backslash A_1$ and then concluding $\| \sum_{E_1} f - \sum_{E_2} f \| < 2\varepsilon$. □

Definition 1.2.41. Let $g : X \to \overline{\mathbb{R}}_{\geqslant 0}$ be a map. Note that the net $(\sum_A g)_{A \in \mathrm{fin}(2^X)}$ is increasing. Hence, its limit exists in $\overline{\mathbb{R}}$ and equals $\sup_{A \in \mathrm{fin}(2^X)} \sum_A g$ (by Prop. 1.2.12). We write this as $\sum_X g$, or more precisely:

$$\sum_X g \equiv \sum_{x \in X} g(x) \xlongequal{\text{def}} \lim_{A \in \mathrm{fin}(2^X)} \sum_A g = \sup_{A \in \mathrm{fin}(2^X)} \sum_A g \tag{1.12}$$

We say that $\sum_X g$ **converges** or **converges absolutely**, if $\sum_X g < +\infty$.

It is clear that $\sum_X g < +\infty$ iff there exists $C \in \mathbb{R}_{\geqslant 0}$ such that $\sum_A g \leqslant C$ for all $A \in \mathrm{fin}(2^X)$.

Remark 1.2.42. Note that when $g : X \to \mathbb{R}_{\geqslant 0}$, the convergence in Def. 1.2.41 agrees with that in Def. 1.2.39. Therefore, Rem. 1.2.40 still gives a Cauchy criterion for convergence.

Definition 1.2.43. Let $f : X \to V$. We say that $\sum_X f$ **converges absolutely** if

$$\sum_{x \in X} \|f(x)\| < +\infty$$

Proposition 1.2.44. *Let $f : X \to V$, and assume that $\sum_X f$ converges absolutely. Then $\mathrm{Supp}(f) := \{x \in X : f(x) \neq 0\}$ is a countable set.*

Proof. For each $\varepsilon > 0$, let $A_\varepsilon = \{x \in X : |f(x)| \geqslant \varepsilon\}$. Then

$$\sum_X |f| \geqslant \sum_{A_\varepsilon} |f| \geqslant \varepsilon \sum_{A_\varepsilon} 1$$

So $\sum_{A_\varepsilon} 1 < +\infty$, and hence A_ε is a finite set. Since $\mathrm{Supp}(f) = \bigcup_{n \in \mathbb{Z}_+} A_{1/n}$, the set $\mathrm{Supp}(f)$ must be countable. □

Proposition 1.2.45. *Assume that V is complete. Let $f : X \to V$. If $\sum_X f$ converges absolutely, then it converges, and*

$$\Big\|\sum_{x\in X} f(x)\Big\| \leqslant \sum_{x\in X} \|f(x)\| \tag{1.13}$$

We write this simply as $\|\sum_X f\| \leqslant \sum_X |f|$.

Proof. (1.13) clearly holds when X is finite. In the general case, assume that $\sum_X f$ converges absolutely. Then by the Cauchy criterion, for every $\varepsilon > 0$ there is $A \in \mathrm{fin}(2^X)$ such that for each finite $E \subset X\backslash A$ we have $\sum_E |f| < \varepsilon$, and hence $\|\sum_E f\| < \varepsilon$. Therefore $\sum_X f$ converges by Cauchy criterion again.

By the continuity of the norm function $v \in V \mapsto \|v\| \in \mathbb{R}_{\geqslant 0}$ (cf. Rem. 2.3.3),

$$\Big\|\sum_X f\Big\| = \Big\|\lim_A \sum_A f\Big\| = \lim_A \Big\|\sum_A f\Big\|$$

Since $\|\sum_A f\| \leqslant \sum_A |f|$, by Prop. 1.2.12, the above expression is no greater than

$$\lim_A \sum_A |f| = \sum_X |f|$$

$\square$

Example 1.2.46. Let X be a set, equipped with the **counting measure** $\mu : 2^X \to [0, +\infty]$. That is, for each $A \subset X$, we have $\mu(A) = |A|$ if A is a finite set, and $\mu(A) = +\infty$ if A is infinite. Then for each $f : X \to \overline{\mathbb{R}}_{\geqslant 0}$, we have

$$\int_X f d\mu = \sum_{x\in X} f(x) \tag{1.14}$$

If $f : X \to \mathbb{C}$ satisfies $\sum_X |f| < +\infty$, then (1.14) still holds. The details are left to the reader.

Example 1.2.47. Let $f : \mathbb{Z}_+ \to V$, and assume that $\sum_{\mathbb{Z}_+} f$ converges. Then

$$\sum_{\mathbb{Z}_+} f = \lim_{n\to\infty} (f(1) + \cdots + f(n))$$

Proof. By assumption, the net $(\sum_A f)_{A\in \mathrm{fin}(2^{\mathbb{Z}_+})}$ converges to $v := \sum_{\mathbb{Z}_+} f$. By Prop. 1.2.27, the subnet $(\sum_{\{1,\dots,n\}} f)_{n\in\mathbb{Z}_+}$ also converges to v. $\square$

1.3 Nets and compactness

1.3.1 Compactness and cluster points

Proposition 1.3.1. *Let X be a topological space. Then the following are equivalent.*

(1) X is compact.

*(2) **(Increasing chain property)** If $(U_\mu)_{\mu\in I}$ is an increasing net of open subsets of X satisfying $\bigcup_{\mu\in I} U_\mu = X$, then $U_\mu = X$ for some μ.*

*(3) **(Decreasing chain property)** If $(E_\mu)_{\mu\in I}$ is a decreasing net of nonempty closed subsets of X, then $\bigcap_{\mu\in I} E_\mu \neq \varnothing$.*

Here, "increasing net" means $U_\mu \subset U_\nu$ if $\mu \leqslant \nu$, and "decreasing net" means the opposite.

Proof. (1)⇒(2): Assume (1). Then $X = \bigcup_\mu U_\mu$ is an open cover of X. So, by the compactness of X, we have $X = U_{\mu_1} \cup \cdots \cup U_{\mu_n}$ for some $\mu_1, \dots, \mu_n \in I$. Choose $\mu \in I$ which is $\geqslant \mu_1, \dots, \mu_n$. Then $X = U_\mu$.

(2)⇒(1): Assume (2). Let $X = \bigcup_{\alpha\in\mathscr{A}} W_\alpha$ be an open cover of X. Let $I = \mathrm{fin}(2^{\mathscr{A}})$. For each $\mu = \{\alpha_1, \dots, \alpha_n\} \in I$, let $U_\mu = W_{\alpha_1} \cup \cdots \cup W_{\alpha_n}$. Then $(U_\mu)_{\mu\in I}$ is an increasing net of open sets covering X. Thus, by (2), we have $U_\mu = X$ for some μ. This proves (1).

(2)⇔(3): If we let $E_\mu = X\backslash U_\mu$, then (2) says that if (E_μ) is a decreasing net of closed sets whose intersection is $\varnothing$, then $E_\mu = \varnothing$ for some μ. This is the contraposition of (3). □

Theorem 1.3.2. *Let X be a topological space. Then X is compact iff every net in X has at least one cluster point.*

Proof. Assume that X is compact. Let $(x_\alpha)_{\alpha\in I}$ be a net in X. Define F_α by

$$F_\alpha = \{x_\beta : \beta \in I, \beta \geqslant \alpha\} \tag{1.15}$$

Then $(\overline{F}_\alpha)_{\alpha\in I}$ is a decreasing net of nonempty closed subsets. So $\bigcap_\alpha \overline{F}_\alpha$ is nonempty by the decreasing chain property (cf. Prop. 1.3.1). By Thm. 1.2.28, $\bigcap_\alpha \overline{F}_\alpha$ is the set of cluster points of (x_α). Therefore, (x_α) has a cluster point.

Conversely, assume that every net of X has a cluster point. By Prop. 1.3.1, to prove that X is compact, it suffices to prove that X satisfies the decreasing chain property. Let $(E_\alpha)_{\alpha\in I}$ be a decreasing net of nonempty closed subsets of X. For each α we choose $x_\alpha \in E_\alpha$, which gives a net $(x_\alpha)_{\alpha\in I}$ in X. The fact that (E_α) is decreasing implies that $F_\alpha \subset E_\alpha$ if we let $F_\alpha =$ (1.15). Thus, the closure $\overline{F}_\alpha$ is a subset of E_α since E_α is closed.

By Thm. 1.2.28, $\bigcap_{\alpha\in I} \overline{F}_\alpha$ is the set of cluster points of (x_α), which is non-empty by assumption. Therefore, $\bigcap_\alpha E_\alpha$ is nonempty. □

Definition 1.3.3. Recall that X is called

- **separable** if X has a countable dense subset.
- **second countable** if the topology $\mathcal{T}_X$ has a countable basis.

- **Lindelöf** if every open cover of X has a countable subcover.

Example 1.3.4. Any subspace of a second countable space is second countable.

Proposition 1.3.5. *We have*

$$\text{separable} \quad \Longleftarrow \quad \text{second countable} \quad \Longrightarrow \quad \text{Lindelöf}$$

Moreover, for metrizable spaces, we have

$$\text{separable} \quad \Longleftrightarrow \quad \text{second countable}$$

In fact, Lindelöf metrizable spaces are also second countable. We will not need this fact.

Proof. Step 1. Suppose that X is second countable. Let $\{U_1, U_2, \dots\}$ be a countable basis of $\mathcal{T}_X$. We may assume WLOG that each U_n is nonempty. Choose $x_n \in U_n$. Then $\{x_n : n \in \mathbb{Z}_+\}$ is a countable dense subset of X. Therefore, X is separable.

Let $\mathcal{W}$ be an open cover of X. For each n, let W_n be any member of $\mathcal{W}$ containing U_n; if no such W_n exists then we set $W_n = \emptyset$. By the fact that $\{U_1, U_2, \dots\}$ is a basis of $\mathcal{T}_X$, one easily shows that $\bigcup_n W_n = X$. Therefore, (W_n) is a countable subcover of $\mathcal{W}$. This proves that X is Lindelöf.

Step 2. Suppose that X is a separable metric space. Let $\{x_1, x_2, \dots\}$ be a dense subset of X. Then $(B(x_n, 1/k))_{n,k\in\mathbb{Z}_+}$ is a countable basis of $\mathcal{T}_X$. This proves that X is second countable. $\square$

Definition 1.3.6. Recall that a topological space X is called **sequentially compact** if every sequence in X has a convergent subsequence. By Thm. 1.2.29, if X is first-countable (in particular, if X metrizable or second-countable), then X is sequentially compact iff every sequence in X has a cluster point.

Theorem 1.3.7. *Let X be a second-countable topological space. Then X is compact iff X is sequentially compact.*

Proof. Similar to the proof of Thm. 1.3.2, one shows that every sequence (x_n) has a cluster point (i.e., $\bigcap_n \overline{\{x_k : k \geqslant n\}}$ is nonempty, cf. Thm. 1.2.28) iff the intersection of a decreasing *sequence* of non-empty closed subsets of X is nonempty. The latter condition is equivalent to that X is **countably compact**, that is, every countable open cover of X has a finite subcover. Since X is Lindelöf (Prop. 1.3.5), compactness and countable compactness are equivalent. $\square$

Corollary 1.3.8. *Let X be a metric space. Then X is compact iff X is sequentially compact.*

Proof. By Thm. 1.3.7 and Prop. 1.3.5, it suffices to prove that X is separable. Note that for each $\varepsilon > 0$, there exists a finite set $E_\varepsilon \subset X$ such that the distance from any point of X to E_ε is $\leqslant \varepsilon$. (Otherwise, one can find a sequence $(x_n)_{n\in\mathbb{Z}_+}$ such that for each n, the distance from x_{n+1} to $\{x_1, \dots, x_n\}$ is $> 1/\varepsilon$. Then any subsequence of (x_n) is not a Cauchy sequence, and hence does not converge.) Then $\bigcup_{n\in\mathbb{Z}_+} E_{1/n}$ is a countable dense subset of X. $\square$

1.3.2 lim inf and lim sup

Theorem 1.3.9. *Let (x_α) be a net in a compact Hausdorff space X, and let $x \in X$. The following are equivalent.*

(1) (x_α) converges to x.

(2) x is the only cluster point of (x_α).

In other words, (x_α) converges to x iff any convergent subnet converges to x.

Proof. (1)⇒(2): This is obvious even without assuming that the Hausdorff space X is compact.

¬(1)⇒¬(2): Assume that (1) is not true. Since (x_α) does not converge to x, there exists $U \in \mathrm{Nbh}_X(x)$ such that (x_α) is frequently not in U. Therefore, $J = \{\alpha \in I : x_\alpha \notin U\}$ is a cofinal subset of I. Note that J is also directed: if $\alpha, \beta \in J$, choose $\gamma \in I$ such that $\alpha, \beta \leqslant \gamma$, then there exists $\delta \in I_{\geqslant \gamma}$ such that $x_\delta \notin U$, and hence $\delta \in J$. It follows that $(x_\alpha)_{\alpha \in J}$ is a subnet of $(x_\alpha)_{\alpha \in I}$ that is always outside U.

Since X is compact, by Thm. 1.3.2, $(x_\alpha)_{\alpha \in J}$ has a subnet $(x_\mu)_{\mu \in K}$ converging to some $y \in X$. Then $y \notin U$, since (x_μ) is always not in U. Therefore, $x \neq y$, and y is a cluster point of (x_α). Hence (2) is not true. □

Definition 1.3.10. Let $(x_\alpha)_{\alpha \in I}$ be a net in the compact Hausdorff space $\overline{\mathbb{R}}$. Let S be the set of cluster points of (x_α) in $\overline{\mathbb{R}}$. Recall that S is a closed subset by Thm. 1.2.28, and is non-empty by Thm. 1.3.2. Define

$$\liminf_{\alpha \in I} x_\alpha = \inf S \qquad \limsup_{\alpha \in I} x_\beta = \sup S \tag{1.16}$$

Since S is closed, $\liminf_{\alpha \in I} x_\alpha$ and $\limsup_{\alpha \in I} x_\beta$ are both cluster points of (x_α). Therefore, they are the smallest and the largest cluster points of (x_α), respectively.

Remark 1.3.11. Let (x_α) be a net in $\overline{\mathbb{R}}$. Clearly $\liminf_\alpha x_\alpha \leqslant \limsup_\alpha x_\alpha$. By Thm. 1.3.9, if $x \in \overline{\mathbb{R}}$, then

$$\lim_\alpha x_\alpha = x \qquad \Longleftrightarrow \qquad \liminf_\alpha x_\alpha = x = \limsup_\alpha x_\alpha \tag{1.17}$$

Proposition 1.3.12. *Let (x_α) be a net in $\overline{\mathbb{R}}$. For each $\alpha \in I$, define*

$$A_\alpha = \inf\{x_\beta : \beta \geqslant \alpha\} \qquad B_\alpha = \sup\{x_\beta : \beta \geqslant \alpha\} \tag{1.18}$$

Then (A_α) is increasing and (B_α) is decreasing; in particular, they converge in $\overline{\mathbb{R}}$. Moreover, we have

$$\lim_{\alpha \in I} A_\alpha = \liminf_{\alpha \in I} x_\alpha \qquad \lim_{\alpha \in I} B_\alpha = \limsup_{\alpha \in I} x_\alpha \tag{1.19}$$

Proof. The monotonicities of (A_α) and (B_α) are obvious. Let $E_\alpha = \{x_\beta : \beta \geqslant \alpha\}$. Then $B_\alpha = \sup E_\alpha = \sup \overline{E_\alpha}$. By Thm. 1.2.28, we have $\limsup_\alpha x_\alpha = \sup \bigcap_\alpha \overline{E_\alpha}$. Since $\bigcap_\alpha \overline{E_\alpha} \subset \overline{E_\beta}$, we have $\limsup_\alpha x_\alpha \leqslant \sup \overline{E_\beta} = B_\beta$. Therefore $\limsup x_\alpha \leqslant \lim B_\alpha$.

For each open interval U in $\overline{\mathbb{R}}$ containing $\lim B_\alpha$, the net $(B_\alpha)_{\alpha\in I} = (\sup E_\alpha)_{\alpha\in I}$ must be eventually in U. From the definition of E_α, we see that (x_α) is frequently in U. It follows from Thm. 1.2.28 that $\lim B_\alpha$ is a cluster point of (x_α). Therefore $\limsup x_\alpha \geqslant \lim B_\alpha$. This proves one of the two relations in (1.19); the other one can be proved in the same way. □

1.4 Review of important facts in point-set topology

Fix a normed vector space $\mathcal{V}$.

1.4.1 Miscellaneous definitions and properties

Definition 1.4.1. If X, Y are metric spaces and $f : X \to Y$ is map, we say that $C \in \mathbb{R}_{\geqslant 0}$ is a **Lipschitz constant** of f if

$$d(f(x_1), f(x_2)) \leqslant C d(x_1, x_2) \qquad \text{for all } x_1, x_2 \in X$$

If f has a Lipschitz constant, we say that f is **Lipschitz continuous**.

Definition 1.4.2. If d and d' are two metrics on a set X, we say that d and d' are **equivalent** if there exist $\alpha, \beta \in \mathbb{R}_{>0}$ such that

$$d(x,y) \leqslant \alpha d'(x,y) \qquad d'(x,y) \leqslant \beta d(x,y) \qquad \text{for all } x, y \in X$$

Definition 1.4.3. Let $X_1, \dots, X_N$ be metric spaces. For each $1 \leqslant p < +\infty$, the $\boldsymbol{l^\infty}$**-product metric** d_∞ and the $\boldsymbol{l^p}$**-product metric** d_p are the metrics on $X_1 \times \cdots \times X_N$ defined by

$$\begin{aligned} d_\infty((x_1, \dots, x_N), (y_1, \dots, y_N)) &:= \max\{d(x_1, y_1), \dots, d(x_N, y_N)\} \\ d_p((x_1, \dots, x_N), (y_1, \dots, y_N)) &:= \sqrt[p]{d(x_1, y_1)^p + \cdots + d(x_N, y_N)^p} \end{aligned}$$

for all $x_i, y_i \in X_i$. These metrics are equivalent. We equip $X_1 \times \cdots \times X_N$ with any metric equivalent to l^∞ and l^p.

Remark 1.4.4. Recall that if $f : X \to Y$ is a map of topological spaces, and $X = \bigcup_{i\in I} U_i$ is an open cover of X, then f is continuous iff $f|_{U_i} : U_i \to Y$ is continuous for any $i \in I$.

Definition 1.4.5. Let $f : X \to Y$ be a map where $(Y, \mathcal{T}_Y)$ is a topological space. The **pullback topology** on X is defined to be

$$f^*\mathcal{T}_Y := f^{-1}(\mathcal{T}_Y) = \{f^{-1}(V) : V \in \mathcal{T}_Y\}$$

Then, a net (x_α) in X converges under $f^*\mathcal{T}_Y$ to x iff

$$\lim_\alpha f(x_\alpha) = f(x)$$

The following facts are well-known.

Proposition 1.4.6. *Let X be a metric space. Let $A \subset X$, whose metric inherits from that of X. If A is complete, then A is closed. Conversely, if A is closed and X is complete, then A is complete.*

Proposition 1.4.7. *Let X be a Hausdorff space. Let $A \subset X$. If A is compact, then A is closed. Conversely, if A is closed and X is compact, then A is compact.*

1.4.2 Semicontinuous functions

Let X be a topological space.

Definition 1.4.8. We say that $f : X \to \overline{\mathbb{R}}$ is **lower semicontinuous** if $f^{-1}(a, +\infty]$ is open for each $a \in \overline{\mathbb{R}}$. We say that f is **upper semicontinuous** if $f^{-1}[-\infty, a)$ is open for each $a \in \overline{\mathbb{R}}$.

Example 1.4.9. Let $A \subset X$. Then χ_A is lower semicontinuous iff A is open.

Proposition 1.4.10. *Let $(f_i)_{i\in I}$ be a family of lower semicontinuous functions $X \to \overline{\mathbb{R}}$. Let $f(x) = \sup_{i\in I} f_i(x)$. Then $f : X \to \overline{\mathbb{R}}$ is lower semicontinuous.*

Proof. Let $x \in f^{-1}(a, +\infty]$. Since $f(x) > a$, there exists i such that $f_i(x) > a$. Since f_i is lower semi-continuous, there exists $U \in \mathrm{Nbh}(x)$ such that $f_i|_U > a$ (e.g. $U = f_i^{-1}(a, +\infty]$), and hence $f|_U > a$. So $U \subset f^{-1}(a, +\infty]$. We have thus proved that any $x \in f^{-1}(a, +\infty]$ is an interior point of $f^{-1}(a, +\infty]$. $\square$

Proposition 1.4.11. *Let $f : X \to \overline{\mathbb{R}}$. Then the following are equivalent.*

(1) f is lower semicontinuous.

(2) For each $x \in X$ and each net (x_α) in X converging to x, we have $\liminf_\alpha f(x_\alpha) \geqslant f(x)$.

(3) For each $x \in X$ and each net (x_α) in X converging to x, we have $\limsup_\alpha f(x_\alpha) \geqslant f(x)$.

When X is first-countable, the same equivalence holds with nets replaced by sequences.

Proof. (1)⇒(2): Assume (1). Let $x \in X$, and let (x_α) be a net converging to x. If $\lambda \in \mathbb{R}$ satisfies $\lambda < f(x)$, then by (1), $f^{-1}(\lambda, +\infty]$ is a neighborhood of x. Therefore, x_α is eventually in $f^{-1}(\lambda, +\infty]$, and hence $f(x_\alpha)$ is eventually $> \lambda$. Therefore $\liminf_\alpha f(x_\alpha) \geqslant \lambda$. Since λ is arbitrary, we must have $\liminf_\alpha f(x_\alpha) \geqslant f(x)$. This proves (2).

(2)⇒(3): This is obvious.

¬(1)⇒¬(3): Assume that f is not lower semicontinuous. Then there exists $\lambda \in [-\infty, +\infty)$ such that $f^{-1}(\lambda, +\infty]$ is not open, and hence has non-interior point x. Therefore, $f(x) > \lambda$, and for any $U \in \mathrm{Nbh}_X(x)$ there exists $x_U \in U$ such that $f(x_U) \leqslant \lambda$. So $(x_U)_{U \in \mathrm{Nbh}_X(x)}$ is a net in X converging to x, and $\limsup_U f(x_U) \leqslant \lambda < f(x)$. Hence (3) is false. □

1.4.3 Product topology and pointwise convergence

Let $(X_\alpha)_{\alpha \in \mathscr{A}}$ be a family of topological spaces. Elements of the product space

$$S = \prod_{\alpha \in \mathscr{A}} X_\alpha$$

are denoted by $x = (x_\alpha)_{\alpha \in \mathscr{A}}$. Let

$$\pi_\alpha : S \to X_\alpha \qquad x \mapsto x(\alpha)$$

It is easy to check that

$$\begin{aligned}\mathcal{B} = \Big\{ \prod_{\alpha \in \mathscr{A}} U_\alpha &: \text{each } U_\alpha \text{ is open in } X_\alpha, \\ &U_\alpha = X_\alpha \text{ for all but finitely many } \alpha \Big\} \\ =& \Big\{ \bigcap_{\alpha \in E} \pi_\alpha^{-1}(U_\alpha) : E \in \mathrm{fin}(2^{\mathscr{A}}),\ U_\alpha \text{ is open in } X_\alpha \text{ for each } \alpha \in E \Big\}\end{aligned}$$

is a base for a topology, namely, for each $W_1, W_2 \in \mathcal{B}$ and $x \in W_1 \cap W_2$, there exists $W_3 \in \mathcal{B}$ such that $W_3 \subset W_1 \cap W_2$. Therefore, $\mathcal{B}$ generates a topology.

Definition 1.4.12. The topology of S generated by $\mathcal{B}$ is called the **product topology** or **pointwise convergence topology** of S. Unless otherwise stated, the product of a family of topological spaces is equipped with the product topology.

Remark 1.4.13. If each X_α is Hausdorff, then S is clearly Hausdorff.

Theorem 1.4.14. *Let $(x_\mu)_{\mu \in I}$ be a net in S, and let $x \in S$. Then the following conditions are equivalent:*

(a) $\lim\limits_{\mu \in I} x_\mu = x$ *in the product topology.*

(b) $(x_\mu)_{\mu\in I}$ *converges pointwise to* x*, namely, for each* $\alpha\in\mathscr{A}$ *we have* $\lim_{\mu\in I} x_\mu(\alpha) = x(\alpha)$ *in* X_α*.*

Proof. (a)⇒(b): Fix $\alpha\in\mathscr{A}$. For each open $U_\alpha\subset X_\alpha$, we have $\pi^{-1}(U_\alpha)\in\mathcal{B}$. Therefore,

$$\pi_\alpha : S\to X_\alpha \qquad \text{is continuous} \tag{1.20}$$

Thus, if $\lim_\mu x_\mu = x$, then $\lim_\mu \pi_\alpha(x_\mu)=\pi_\alpha(x)$. This proves (b).

(b)⇒(a): Assume (b). Choose any $W\in\mathcal{B}$ containing x. Then there exists $E\in\mathrm{fin}(2^{\mathscr{A}})$ such that $W=\bigcap_{\alpha\in E}\pi_\alpha^{-1}(U_\alpha)$, where each $U_\alpha\subset X_\alpha$ is open and containing x_α. For such $\alpha\in E$, since $\lim_\mu x_\mu(\alpha)=x(\alpha)$, we know that $(x_\mu(\alpha))$ is μ-eventually in U_α. Therefore, since E is finite, we conclude that (x_μ) is eventually in W. This proves (a). □

Corollary 1.4.15. *Let* Z *be a topological space. Suppose that for each* $\alpha\in\mathscr{A}$*, a map* $f_\alpha : Z\to X_\alpha$ *is chosen. Then*

$$\bigvee_{\alpha\in\mathscr{A}} f_\alpha : Z\to\prod_{\alpha\in\mathscr{A}} X_\alpha \qquad z\mapsto (f_\alpha(z))_{\alpha\in\mathscr{A}} \tag{1.21}$$

is continuous iff f_α *is continuous for each* $\alpha\in\mathscr{A}$*.*

Proof. If $F:=\bigvee_{\alpha\in\mathscr{A}} f_\alpha$ is continuous, then since π_α is continuous, $f_\alpha=\pi_\alpha\circ f_\alpha$ is also continuous. Conversely, suppose that each f_α is continuous. Let (z_i) be a net in Z converging to $z\in Z$. For each α, since f_α is continuous, we see that $\lim_i f_\alpha(z_i)=f_\alpha(z)$. By Thm. 1.4.14, $F(z_i)$ converges to $F(z)$. This proves that F is continuous. □

Proposition 1.4.16. *Suppose that* $\mathscr{A}$ *is countable. If each* X_α *is second countable, then* S *is second countable. If each* X_α *is metrizable, then* S *is metrizable.*

Proof. If $\mathcal{U}_\alpha$ is a base of the topology of X_α, then

$$\mathcal{U}:=\Big\{\bigcap_{\alpha\in E}\pi_\alpha^{-1}(U_\alpha) : E\in\mathrm{fin}(2^{\mathscr{A}}), U_\alpha\in\mathcal{U}_\alpha\Big\}$$

is a base of the the product topology, which is countable if each $\mathcal{U}_\alpha$ is countable.

Now assume that each X_α is equipped with a metric d_α. Fix any $R\in\mathbb{R}_{>0}$, let $\widetilde{d}_\alpha$ be metric on X_α inducing the same topology as d_α and satisfying $\widetilde{d}_\alpha\leqslant R$. For example,

$$\widetilde{d}_\alpha(x_\alpha,y_\alpha)=\min\{d_\alpha(x_\alpha,y_\alpha),R\} \qquad \text{for each } x_\alpha,y_\alpha\in X_\alpha \tag{1.22a}$$

Let $\nu:\mathscr{A}\to\mathbb{Z}_+$ be an injective map, and define a metric d on S by

$$d(x,y)=\sum_{\alpha\in\mathscr{A}} 2^{-\nu(\alpha)}\widetilde{d}_\alpha(x(\alpha),y(\alpha)) \qquad \text{for each } x,y\in S \tag{1.22b}$$

One shows easily that a net (x_μ) in S converges in the metric d to $x \in S$ iff $\lim_\mu \widetilde{d}_\alpha(x_\mu(\alpha), x(\alpha)) = 0$ for all $\alpha \in \mathscr{A}$. Thus, by Thm. 1.4.14, d induces the product topology. $\square$

Theorem 1.4.17 (Tychonoff theorem). *Assume that X_α is compact for each $\alpha \in \mathscr{A}$. Then S is compact.*

$\star$ *Proof.* Assume WLOG that $\mathscr{A}$ is non-empty, and that each X_α is non-empty. Let $(x_\mu)_{\mu\in I}$ be a net in S. We want to show that $(x_\mu)_{\mu\in I}$ has a cluster point.

For each $\mathscr{E} \subset \mathscr{A}$, let $S_{\mathscr{E}} = \prod_{\alpha\in\mathscr{E}} X_\alpha$. For each $x \in S_{\mathscr{E}}$, we write $\mathrm{Dom}(x) = \mathscr{E}$. For each $\mathscr{E} \subset \mathscr{F} \subset \mathscr{A}$ and $y \in S_{\mathscr{F}}$, let $y|_{\mathscr{E}} = (y(\alpha))_{\alpha\in\mathscr{E}}$. Let

$$P = \bigcup_{\mathscr{E}\subset\mathscr{A}} \{x \in S_{\mathscr{E}} : x \text{ is a cluster point of } (x_\mu|_{\mathscr{E}})_{\mu\in I} \text{ in } S_{\mathscr{E}}\}$$

equipped with the partial order "$\subset$". In other words, if $x, y \in P$, then $x \leqslant y$ means that $\mathrm{Dom}(x) \subset \mathrm{Dom}(y)$ and $x = y|_{\mathscr{E}}$.

Since each X_α is compact, P is clearly non-empty. Let us show that every totally ordered non-empty subset $Q \subset P$ has an upper bound in P, so that Zorn's lemma can be applied. Let x be the union of all elements of Q. Thus $x \in S_{\mathscr{E}}$ where $\mathscr{E} = \bigcup_{y\in Q} \mathrm{Dom}(y)$, and we have $x|_{\mathrm{Dom}(y)} = y$ for each $y \in Q$.

To show that x is a cluster point of $(x_\mu|_{\mathscr{E}})_{\mu\in I}$ in $S_{\mathscr{E}}$, we pick any neighborhood of x in $S_{\mathscr{E}}$, which, after shrinking if necessary, is of the form $W = \prod_{\alpha\in\mathscr{E}} U_\alpha$ where each $U_\alpha \subset X_\alpha$ is open, and there exists $K \in \mathrm{fin}(2^{\mathscr{E}})$ such that $U_\alpha = X_\alpha$ whenever $\alpha \notin K$. Since $\mathscr{E} = \bigcup_{y\in Q} \mathrm{Dom}(y)$, there exists $y \in Q$ such that $K \subset \mathrm{Dom}(y)$. Namely, $(x_\mu|_{\mathrm{Dom}(y)})_{\mu\in I}$ has cluster point y, and $K \subset \mathrm{Dom}(y)$. Therefore $(x_\mu|_K)_{\mu\in I}$ has cluster point $y|_K$ (which equals $x|_K$ because $x|_{\mathrm{Dom}(y)} = y$), and hence is frequently in $\prod_{\alpha\in K} U_\alpha$. Thus $(x_\mu|_{\mathscr{E}})_{\mu\in I}$ is frequently in W. This finishes the proof that $x \in P$. Clearly x is an upper bound of Q.

Now we can apply Zorn's lemma, which claims that P has a maximal element $x \in P$. The proof of the Tychonoff theorem will be finished by showing that $\mathscr{E} := \mathrm{Dom}(x)$ equals $\mathscr{A}$. Suppose not. Choose $\beta \in \mathscr{A}\backslash\mathscr{E}$. Since $x \in P$, there is a subnet $(x_{\mu_\nu}|_{\mathscr{E}})_{\nu\in J}$ of $(x_\mu|_{\mathscr{E}})_{\mu\in I}$ converging pointwise to x. Since X_β is compact, $(x_{\mu_\nu}(\beta))_{\nu\in J}$ has a converging subnet $(x_{\mu_{\nu_\upsilon}}(\beta))_{\upsilon\in L}$. Define $\widetilde{x} \in S_{\mathscr{E}\cup\{\beta\}}$ to be x when restricted to $\mathscr{E}$, and $\widetilde{x}(\beta) := \lim_\upsilon x_{\mu_{\nu_\upsilon}}(\beta)$. Then $\widetilde{x} \in P$, and $\widetilde{x}$ is strictly larger than x, contradicting the maximality of x. $\square$

Remark 1.4.18. If $\mathscr{A}$ is a countable set, and if each X_α is compact and second-countable, the **diagonal method** can be used in place of Zorn's lemma to prove that S (which is second countable by Prop. 1.4.16) is compact.

We consider the case that $\mathscr{A} = \mathbb{Z}_+$. (The case that $\mathscr{A}$ is finite is even simpler.) Let $(x_n)_{n\in\mathbb{Z}_+}$ be a sequence in S. We construct inductively a double sequence $(x_{m,n})_{m,n\in\mathbb{Z}_+}$ in S as follows. Since (by Thm. 1.3.7) X_1 is sequentially compact, (x_n) has subseqeunce $(x_{1,n})_{n\in\mathbb{Z}_+}$ whose first component $(x_{1,n}(1))_{n\in\mathbb{Z}_+}$ converges to some $x(1) \in X_1$. Suppose that $(x_{m-1,n})_{n\in\mathbb{Z}_+}$ has been constructed (where

$m - 1 \geqslant 1$). Since X_m is sequentially compact, $(x_{m-1,n})_{n\in\mathbb{Z}_+}$ has a subsequence $(x_{m,n})_{n\in\mathbb{Z}_+}$ whose m-th component $(x_{m,n}(m))_{n\in\mathbb{Z}_+}$ to some $x(m) \in X_m$. In this way, the double sequence $(x_{m,n})$ in S and the element $x \in S$ are constructed. One checks easily that $(x_{n,n})_{n\in\mathbb{Z}_+}$ is a subsequence of (x_n) converging to x.

We have thus proved that S is sequentially compact. By Thm. 1.3.7, S is compact. □

1.4.4 Precompact sets

Let X be a Hausdorff space.

Definition 1.4.19. Let $A \subset X$. We say that A is **precompact** relative to X and write

$$A \Subset X$$

if $\mathrm{Cl}_X(A)$ is compact, equivalently, if A is contained in a compact subset of X.

Recall that a subset of a compact Hausdorff space is closed iff it is compact (cf. Prop. 1.4.7).

Proof of equivalence. "$\Rightarrow$": Obvious. "$\Leftarrow$": Let $B \subset X$ be compact and containing A. Then B is closed in X. So $\mathrm{Cl}_X(A) \subset B$. Since $\mathrm{Cl}_X(A)$ is closed in X and hence closed in B, it is compact. □

Remark 1.4.20. Let $W \subset X$. Then for each $A \subset W$, we have

$$A \Subset W \qquad \Longleftrightarrow \qquad A \Subset X \text{ and } \mathrm{Cl}_X(A) \subset W$$

When either side is true, we have $\mathrm{Cl}_W(A) = \mathrm{Cl}_X(A)$. Thus, both $\mathrm{Cl}_W(A)$ and $\mathrm{Cl}_X(A)$ can be denoted unambiguously by $\overline{A}$.

In practice, we often choose W to be an open subset of X.

Proof. "$\Leftarrow$": $\mathrm{Cl}_X(A)$ is a compact set inside W and contains A. So $A \Subset W$.

"$\Rightarrow$": We have a compact set B such that $A \subset B \subset W$. So $A \Subset X$. Since B is closed in X, we have $\mathrm{Cl}_X(A) \subset B$ and hence $\mathrm{Cl}_X(A) \subset W$.

It is obvious that $\mathrm{Cl}_W(A) \subset \mathrm{Cl}_X(A)$. Assume $A \Subset W$. Then $\mathrm{Cl}_W(A)$ is compact. In the above paragraph, if we choose $B = \mathrm{Cl}_W(A)$. then we have $\mathrm{Cl}_X(A) \subset B = \mathrm{Cl}_W(A)$. This proves $\mathrm{Cl}_W(A) = \mathrm{Cl}_X(A)$. □

Remark 1.4.21. Let U be an open subset of X. Let $f \in C_c(U, \mathcal{V})$. Then by zero-extension, f can be viewed as an element of $C_c(X, \mathcal{V})$ supported in U. Briefly speaking, we have

$$C_c(U, \mathcal{V}) \subset C_c(X, \mathcal{V})$$

Moreover, for each $f \in C_c(U, \mathcal{V})$, we have

$$\mathrm{Supp}_U(f) = \mathrm{Supp}_X(f)$$

Proof. Let f take value 0 outside U. Let $K = \mathrm{Supp}_U(f)$, which is compact by assumption. Since $f|_U$ and $f|_{K^c} = 0$ are continuous, and since $X = U \cup K^c$ is an open cover on X, f is continuous. By the Rem. 1.4.20, we have $\mathrm{Supp}_U(f) = \mathrm{Supp}_X(f)$. Therefore $f \in C_c(X, \mathcal{V})$. $\square$

Under the setting of Rem. 1.4.21, it is clear that

$$C_c(U, \mathcal{V}) = \{f \in C_c(X, \mathcal{V}) : \mathrm{Supp}_X(f) \subset U\} \tag{1.23}$$

Proposition 1.4.22. *Assume that X is a metric space, and let $A \subset X$. The following are equivalent.*

(1) A is precompact.

(2) Every net in A has a cluster point in X.

(3) Every sequence in A has a cluster point in X.

Note that since X is a metric space, Prop. 1.2.29 implies that any cluster point of a sequence is not only the limit of a convergent subnet of that sequence, but also the limit of a convergent subsequence.

Proof. (1)⇒(2): Since $\overline{A}$ is compact, every net in A has a cluster point in $\overline{A}$ and hence in X.

(2)⇒(3): This is obvious.

(3)⇒(1): Assume (3). Let (x_n) be a sequence in $\overline{A}$. For each n, choose $y_n \in A$ such that $d(x_n, y_n) < 1/n$. By assumption, (y_n) has a subsequence (y_{n_k}) converging to some $x \in X$. One easily checks that (x_{n_k}) converges to x. This proves that $\overline{A}$ is sequentially compact, and hence compact by Cor. 1.3.8. $\square$

1.4.5 LCH spaces

Let X be LCH.

Proposition 1.4.23. *Any closed or open subset of X is LCH.*

Proof. See [Gui-A, Subsec. 8.6.2]. $\square$

Corollary 1.4.24. *Let $W \subset X$ be an open subset. Let $K \subset W$ be compact. Then there exists an open subset U of X such that $K \subset U \Subset W$.*

Proof. The case that K is a single point follows from the fact that W is LCH, cf. Prop. 1.4.23. The general case follows from the compactness of K. $\square$

Corollary 1.4.25. *Let K_1, K_2 be mutually disjoint compact subsets of X. Then there exist open subsets U_1, U_2 of X such that $K_1 \subset U_1$ and $K_2 \subset U_2$.*

Proof. This corollary in fact holds even without the assumption that X is locally compact, and its proof is a straightforward exercise in point-set topology. However, it also follows directly from the results established above. Indeed, by Prop. 1.4.23, $X\backslash K_2$ is LCH. Therefore, by Cor. 1.4.24, there exists an open set U_1 such that $K_1 \subset U_1 \Subset X\backslash K_2$. Let $U_2 = X\backslash\overline{U}_1$. □

Theorem 1.4.26 (Urysohn's lemma). *Let $K \subset X$ be compact. Then there exists a (continuous)* ***Urysohn function*** *f with respect to K and X, i.e., $f \in C_c(X, [0, 1])$ and $f|_K = 1$.*

Proof. See [Gui-A, Sec. 15.4]. □

Remark 1.4.27. Urysohn's lemma can be used in the following way. Suppose that $K \subset U \subset X$ where K is compact and U is open in X. By Prop. 1.4.23, U is LCH. Therefore, by Thm. 1.4.26, there exists $f \in C_c(U, [0, 1])$ such that $f|_K = 1$. By Rem. 1.4.21, f can be viewed as an element of $C_c(X, [0, 1])$ satisfying $f|_K = 1$ and $\mathrm{Supp}_X(f) \subset U$.

Theorem 1.4.28. *Let K be a compact subset of X. Let $\mathfrak{U} = (U_1, \dots, U_n)$ be a finite collection of open subsets of X covering K (i.e. $K \subset U_1 \cup \cdots \cup U_n$). Then there exist $h_i \in C_c(U_i, \mathbb{R}_{\geqslant 0})$ (for all $1 \leqslant i \leqslant n$) satisfying the following conditions:*

(1) $0 \leqslant \sum_{i=1}^{n} h_i \leqslant 1$ *on* X.

(2) $\sum_{i=1}^{n} h_i\big|_K = 1$.

Such $h_1, \dots, h_n$ are called a ***partition of unity of K subordinate to $\mathfrak{U}$****.*

In fact, $h_1, \dots, h_n$ should be viewed as a partition of the Urysohn function $h := h_1 + \cdots + h_n$.

Proof. See [Gui-A, Sec. 15.4]. Note that condition (1) is not stated in some textbooks on partitions of unity. However, even if (1) is not initially satisfied, one can enforce it by setting $g(x) = \max\{\sum_i h_i(x), 1\}$ and replacing each h_i with h_i/g. □

Theorem 1.4.29 (Tietze extension theorem). *Let K be a compact subset of X. Let $f \in C(K, \mathbb{F})$. Then there exists $\widetilde{f} \in C_c(X, \mathbb{F})$ such that $\widetilde{f}|_K = f$, and that $\|\widetilde{f}\|_{l^\infty(X)} = \|f\|_{l^\infty(K)}$.*

Proof. See [Gui-A, Sec. 15.4]. □

Remark 1.4.30. Let $k \in \mathbb{N} \cup \{\infty\}$. Suppose that X is an open subset of $\mathbb{R}^n$ (or more generally, a C^k-manifold with or without boundary, e.g. $[a, b]$). Then Urysohn's lemma (Thm. 1.4.26), the partition of unity (Thm. 1.4.28), and the Tietze extension theorem (Thm. 1.4.29) hold verbatim with C_c replaced by C_c^k. See [Gui-A, Sec. 30.6] for detailed explanations.

As an application of the C^k-Urysohn lemma, we see that elements of $C_c^k(X,\mathbb{F})$ vanishes nowhere and separates the points of X. It follows from the Stone-Weierstrass Theorem (see Thm. 1.5.13) that $C_c^k(X,\mathbb{F})$ is l^∞-dense in $C_0(X,\mathbb{F})$. $\square$

Definition 1.4.31. We let $\boldsymbol{C_0(X,\mathcal{V})}$ be the set of all $f\in C(X,\mathcal{V})$ whose zero-extension to the one-point compactification of X is continuous at ∞. Equivalently, $C_0(X,\mathcal{V})$ is the set of all $f\in C(X,\mathcal{V})$ such that for any $\varepsilon>0$ there exists a compact $K\subset X$ such that $\|f\|_{l^\infty(X\backslash K)}<\varepsilon$. See [Gui-A, Subsec. 15.8.1] for more discussions. For each $E\subset\mathcal{V}$, we let

$$C_0(X,E)=C_0(X,V)\cap E^X$$

Unless otherwise stated, we equip $C_0(X,\mathcal{V})$ with the l^∞-norm.

Remark 1.4.32. $C_c(X,\mathcal{V})$ is dense in $C_0(X,\mathcal{V})$ under the l^∞-norm.

Proof. Choose any $f\in C_0(X,\mathcal{V})$. Then for each $\varepsilon>0$ there exists a compact $K\subset X$ such that $\|f\|_{l^\infty(K^c)}<\varepsilon$. By Urysohn's lemma, there exists $h\in C_c(X,[0,1])$ such that $h|_K=1$. Then $\|hf\|_{l^\infty(K^c)}<\varepsilon$, and hence $\|f-hf\|_{l^\infty(X)}<2\varepsilon$. This finishes the proof, since $hf\in C_c(X,\mathcal{V})$. $\square$

Exercise 1.4.33. Suppose that $\mathcal{V}$ is complete. Show that $C_0(X,\mathcal{V})$ is also complete.

Remark 1.4.34. Suppose that X is second countable. Then X is Lindelöf by Prop. 1.3.5. Therefore, X has a countable open cover $\mathfrak{U}=(U_n)_{n\in\mathbb{Z}_+}$ whose members U_n are precompact open subsets of X. In particular, X is σ-compact, since $X=\bigcup_{n\in\mathbb{Z}_+}\overline{U_n}$ where each $\overline{U_n}$ is compact.

Since any open subset $W\subset X$ is second-countable and is LCH (by Prop. 1.4.23), it follows that W is a countable union of precompact open subsets of W; in particular, W is σ-compact. $\square$

1.4.6 Equicontinuity and the Arzelà-Ascoli theorem

Let X be a topological space. Let V be a normed vector space.

Definition 1.4.35. Let $(f_\alpha)_{\alpha\in\mathscr{A}}$ be a family of functions $X\to V$. We say that $(f_\alpha)_{\alpha\in\mathscr{A}}$ is **equicontinuous at $\boldsymbol{x\in X}$** if for each $\varepsilon>0$ there exists $U_x\in\mathrm{Nbh}_X(x)$ such that

$$\|f_\alpha(y)-f_\alpha(x)\|\leqslant\varepsilon\qquad\text{for all }\alpha\in\mathscr{A},y\in U_x\tag{1.24}$$

This is equivalent to saying that

$$\lim_{y\to x}\sup_\alpha\|f_\alpha(y)-f_\alpha(x)\|=0$$

We say that $(f_\alpha)_{\alpha\in\mathscr{A}}$ is an **equicontinuous family of functions** if it is equicontinuous at every point of X.

Theorem 1.4.36. *Let $(f_\alpha)_{\alpha\mathscr{A}}$ be an equicontinuous net of functions $X \to V$ converging pointwise to some $f : X \to V$. Then f is continuous. Moreover, if X is compact, then $(f_\alpha)_{\alpha\mathscr{A}}$ converges uniformly to f.*

Proof. For each $x \in X$ and $\varepsilon > 0$, choose $U_x \in \mathrm{Nbh}_X(x)$ such that (1.24) holds. Applying $\lim_\alpha$, we see that $\|f(y) - f(x)\| \leqslant \varepsilon$ for all $y \in U_x$. This proves that f is continuous at every $x \in X$.

Next, assume that X is compact. Then there exist $x_1, \dots, x_n \in X$ such that $X = U_1 \cup \cdots \cup U_n$ if we set $U_i = U_{x_i}$. For each $y \in X$, choose i such that $y \in U_i$. Then

$$\begin{aligned}&\|f_\alpha(y) - f(y)\| \leqslant \|f_\alpha(y) - f_\alpha(x_i)\| + \|f_\alpha(x_i) - f(x_i)\| + \|f(x_i) - f(y)\|\\ \leqslant&2\varepsilon + \|f_\alpha(x_i) - f(x_i)\|\end{aligned}$$

and hence

$$\sup_{y\in X}\|f_\alpha(y) - f(y)\| \leqslant 2\varepsilon + \sum_{i=1}^n \|f_\alpha(x_i) - f(x_i)\|$$

Since f_α converges to f at $x_1, \dots, x_n$, we conclude that

$$\limsup_\alpha \sup_{y\in X}\|f_\alpha(y) - f(y)\| \leqslant 2\varepsilon$$

for all $\varepsilon > 0$. Hence $\lim_\alpha \sup_{y\in X}\|f_\alpha(y) - f(y)\| = 0$. $\square$

Theorem 1.4.37 (Arzelà-Ascoli theorem). *Assume that X is LCH. Let $\mathscr{F}$ be an equicontinuous set of functions $X \to \mathbb{F}$. If X is non-compact, we assume that*

$$\lim_{x\to\infty} \sup_{f\in\mathscr{F}} |f(x)| = 0 \tag{1.25}$$

which means that for each $\varepsilon > 0$, there exists a compact $K \subset X$ such that

$$|f(x)| \leqslant \varepsilon \qquad \text{for all } f \in \mathscr{F}, x \in X\backslash K$$

Assume that $\mathscr{F}$ is ***pointwise bounded****, i.e., $\sup_{f\in\mathscr{F}}|f(x)| < +\infty$ for each $x \in X$. Then $\mathscr{F}$ is precompact in $C_0(X,\mathbb{F})$ (in the l^∞-norm).*

Proof. It is easy to check that $\overline{\mathscr{F}}$, the closure of $\mathscr{F}$, satisfies the same properties as $\mathscr{F}$. Therefore, replacing $\mathscr{F}$ with its closure, we assume WLOG that $\mathscr{F}$ is closed.

We first consider the case that X is compact. Since $\mathscr{F}$ is pointwise bounded, it can be viewed as a subset of $\prod_{x\in X} D_x$ where D_x is a compact set in $\mathbb{F}$. Therefore, by the Tychonoff Thm. 1.4.17, every net (f_α) in $\mathscr{F}$ has a pointwise convergent subnet (f_β). Since (f_β) is equicontinuous and X is compact, by Thm. 1.4.36, (f_β) converges uniformly to some $f \in C(X,\mathbb{F})$. Since $\mathscr{F}$ is closed, we have $f \in \mathscr{F}$. This proves that $\mathscr{F}$ is compact.

Next, assume that X is not compact. Let $\widehat{X} = X \cup \{\infty\}$ be the one-point compactification of X. Extend each $f \in \mathscr{F}$ to $\widehat{X} \to \mathbb{F}$ by setting $f(0) = 0$. Then $\mathscr{F}$ is a pointwise bounded and equicontiuous family of functions $\widehat{X} \to \mathbb{C}$. By the previous paragraph, every net in $\mathscr{F}$ has a uniformly convergent subnet. Thus, $\mathscr{F}$ is compact. $\square$

1.5 *-algebras and the Stone-Weierstrass theorem

Recall that $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. In this section, we let $\mathbb{K}$ be any subfield of $\mathbb{C}$ closed under complex conjugation, such as $\mathbb{R}, \mathbb{C}, \mathbb{Q}, \mathbb{Q} + \mathbf{i}\mathbb{Q}$.

Definition 1.5.1. A $\mathbb{K}$**-algebra** is defined to be a ring $\mathscr{A}$ (not necessarily having 1) that is also a $\mathbb{K}$-vector space, together with a bilinear map

$$\mathscr{A} \times \mathscr{A} \to \mathscr{A} \qquad (x, y) \mapsto xy$$

satisfying the associacitity rule

$$(xy)z = x(yz) \qquad \text{for all } x, y, z \in \mathscr{A}$$

A $\mathbb{K}$-algebra is called **unital** if $\mathscr{A}$, as a ring, has a multiplicative identity $1_{\mathscr{A}} = 1$ (i.e. $1x = x1 = 1$ for all $x \in \mathscr{A}$). In this case, we write $\lambda \cdot 1$ as λ if $\lambda \in \mathbb{K}$.

A $\mathbb{K}$-algebra is called **commutative** or **abelian** if $xy = yx$ for all $x, y \in \mathscr{A}$.

If $\mathscr{A}$ is a $\mathbb{K}$-algebra, then a **($\mathbb{K}$-)subalgebra** is a subset $\mathscr{B}$ which is closed under the ring addition, ring multiplication, and scalar multiplication. (Namely, $\mathscr{B}$ is a subring and also a linear subspace of $\mathscr{A}$.) If $\mathscr{A}$ is unital, then a **unital ($\mathbb{K}$-)subalgebra** of $\mathscr{A}$ is a $\mathbb{K}$-subalgebra containing the identity of $\mathscr{A}$. $\square$

Example 1.5.2. If V is a $\mathbb{F}$-vector space, then $\mathrm{End}(V)$, the set of $\mathbb{F}$ linear maps $V \to V$, is naturally an $\mathbb{F}$-algebra. If V is a normed vector space, then $\mathfrak{L}(V)$ is an $\mathbb{F}$-algebra.

Definition 1.5.3. A ***-$\mathbb{K}$-algebra** is defined to be a $\mathbb{K}$-algebra together with an **antilinear map** $* : \mathscr{A} \to \mathscr{A}$ sending x to x^* (where "antilinear" means that for every $a, b \in \mathbb{C}$ and $x, y \in \mathscr{A}$ we have $(ax + by)^* = \overline{a}x^* + \overline{b}y^*$) such that for every $x, y \in \mathscr{A}$, we have

$$(x^*)^* = x \qquad (xy)^* = y^*x^*$$

Note that $*$ must be bijective. We call $*$ an **involution**. A ***-$\mathbb{K}$-subalgebra** $\mathscr{B}$ is defined to be a subalgebra satisfying $x \in \mathscr{B}$ iff $x^* \in \mathscr{B}$. If $\mathscr{A}$ is a unital algebra with unit $1_{\mathscr{A}}$, we say that $\mathscr{A}$ is a **unital *-$\mathbb{K}$-algebra** if $\mathscr{A}$ is equipped with an involution $* : \mathscr{A} \to \mathscr{A}$ such that $\mathscr{A}$ is a *-algebra, and that

$$1_{\mathscr{A}}^* = 1_{\mathscr{A}}$$

A **unital *-$\mathbb{K}$-subalgebra** of $\mathscr{A}$ is defined to be a $*$-$\mathbb{K}$-subalgebra of $\mathscr{A}$ containing $1_{\mathscr{A}}$.

Convention 1.5.4. We omit "$\mathbb{K}$-" when $\mathbb{K}$ is $\mathbb{C}$. For example, a **unital *-algebra** means a unital *-$\mathbb{C}$-algebra.

Example 1.5.5. The set $\mathbb{R}[x_\bullet] \equiv \mathbb{R}[x_1, \dots, x_N]$ of real polynomials of $x_1, \dots, x_N$ is a unital *-$\mathbb{R}$-algebra, whose involution $*$ is the identity map.

Example 1.5.6. The set

$$\mathbb{C}[z_\bullet, \overline{z_\bullet}] \equiv \mathbb{C}[z_1, \overline{z_1}, \dots, z_N, \overline{z_N}]$$

of complex polynomials of the $2N$ mutually-independent (and mutually commuting) variables $z_1, \dots, z_N$ and $\overline{z_1}, \dots, \overline{z_N}$ is a unital *-algebra. The involution is described by

$$(z_i)^* = \overline{z_i} \qquad (\overline{z_i})^* = z_i$$

for each $1 \leqslant i \leqslant N$.

Example 1.5.7. The set of complex $n \times n$ matrices $\mathbb{C}^{n\times n}$ is naturally a unital *-algebra if for every $A \in \mathbb{C}^{n\times n}$ we define $A^* = \overline{A}^{\mathrm{t}}$, the complex conjugate of the transpose of A.

Example 1.5.8. Let X be a set. Then $\mathbb{K}^X$ is a unital *-algebra if for every $f \in \mathbb{K}^X$ we define

$$f^* : X \to \mathbb{K} \qquad \boldsymbol{f^*(x)} = \overline{f(x)} \tag{1.26}$$

and $l^\infty(X, \mathbb{K})$ is a unital *-$\mathbb{K}$-subalgebras of $\mathbb{K}^X$.

If $f_1, \dots, f_n \in \mathbb{K}^X$, then $\boldsymbol{\mathbb{K}[f_1, \dots, f_n]}$, the set of polynomials of $f_1, \dots, f_n$ with coefficients in $\mathbb{K}$, is a unital $\mathbb{K}$-subalgebra of $C(X, \mathbb{F})$. And $\mathbb{K}[f_1, f_1^*, \dots, f_n, f_n^*]$ is a unital *-$\mathbb{K}$-subalgebra of $C(X, \mathbb{K})$. □

More generally, we have:

Example 1.5.9. Let $\mathscr{A}$ be a unital $\mathbb{K}$-algebra. Let $\mathfrak{S} \subset \mathscr{A}$. Then

$$\boldsymbol{\mathbb{K}\langle\mathfrak{S}\rangle} = \mathrm{Span}_{\mathbb{K}}\{x_1^{n_1} \cdots x_k^{n_k} : k \in \mathbb{Z}_+, x_i \in \mathfrak{S}, n_i \in \mathbb{N}\} \tag{1.27}$$

the set of (possibly non-commutative) polynomials of elements in $\mathfrak{S}$, is the smallest unital $\mathbb{K}$-subalgebra containing $\mathfrak{S}$, called the **unital $\mathbb{K}$-subalgebra generated by** $\mathfrak{S}$. (Here, we understand $x^0 = 1$ if $x \in \mathscr{A}$.) Thus, if $\mathscr{A}$ is an abelian unital *-algebra, then $\mathbb{C}\langle\mathfrak{S} \cup \mathfrak{S}^*\rangle$ (where $\mathfrak{S}^* = \{x^* : x \in \mathfrak{S}\}$) is the smallest unital *-algebra containing $\mathfrak{S}$, called the **unital *-$\mathbb{K}$-subalgebra generated by** $\mathfrak{S}$.

If $\mathscr{A}$ is commutative, we write $\mathbb{K}\langle\mathfrak{S}\rangle$ as $\boldsymbol{\mathbb{K}[\mathfrak{S}]}$. □

Definition 1.5.10. Let X be a set. Let $(f_\alpha)_{\alpha\in\mathfrak{A}}$ be a family of maps where $f_\alpha : X \to Y_\alpha$ and each Y_α is a set. We say that $(f_\alpha)_{\alpha\in\mathfrak{A}}$ **separates the points of X** if for any distinct $x_1, x_2 \in X$ there exists $\alpha \in \mathfrak{A}$ such that $f_\alpha(x_1) \neq f_\alpha(x_2)$. This is equivalent to saying that the map

$$\bigvee_{\alpha\in\mathfrak{A}} f_\alpha : X \to \prod_{\alpha\in\mathfrak{A}} Y_\alpha \qquad x \mapsto (f_\alpha(x))_{\alpha\in\mathfrak{A}} \tag{1.28}$$

is injective.

Example 1.5.11. Let X be an LCH space. Then $C_c(X, [0,1])$ separates the points of X.

Proof. Choose any distinct points $x, y \in X$. By Urysohn's lemma (Rem. 1.4.27), there exists $f \in C_c(X, [0,1])$ such that $f(x) = 1$ and $\mathrm{Supp}(f) \subset X\backslash\{y\}$. So f separates x and y. $\square$

Theorem 1.5.12 (Stone-Weierstrass theorem). *Let X be a compact Hausdorff space. Suppose that $\mathfrak{S} \subset C(X, \mathbb{F})$ contains 1_X. Suppose that $\mathfrak{S}$ separates the points of X. Then the (unital) *-$\mathbb{F}$-subalgebra $\mathbb{F}[\mathfrak{S} \cup \mathfrak{S}^*]$ generated by $\mathfrak{S}$ is dense in $C(X, \mathbb{F})$ in the l^∞-norm.*

Note that if $\mathbb{F} = \mathbb{R}$, then $\mathfrak{S}^* = \mathfrak{S}$ by (1.26).

If $\mathbb{F} = \mathbb{C}$, then since $(\mathbb{Q} + \mathrm{i}\mathbb{Q})[\mathfrak{S} \cup \mathfrak{S}^*]$ is l^∞-dense in $\mathbb{C}[\mathfrak{S} \cup \mathfrak{S}^*]$, it follows that $(\mathbb{Q} + \mathrm{i}\mathbb{Q})[\mathfrak{S} \cup \mathfrak{S}^*]$ is l^∞-dense in $C(X)$. Similarly, if $\mathbb{F} = \mathbb{R}$, then $\mathbb{Q}[\mathfrak{S}]$ is l^∞-dense in $C(X, \mathbb{R})$.

Proof. See [Gui-A, Ch. 15]. $\square$

Theorem 1.5.13 (Stone-Weierstrass theorem). *Let X be an LCH space. Let $\mathfrak{S} \subset C_0(X, \mathbb{F})$. Suppose that the following conditions are satisfied:*

(1) $\mathfrak{S}$ separates the points of X.

(2) $\mathfrak{S}$ vanishes nowhere, that is, for each $x \in X$ there exists $f \in \mathfrak{S}$ such that $f(x) \neq 0$.

*Then the *-$\mathbb{F}$-subalgebra $\mathbb{F}[\mathfrak{S} \cup \mathfrak{S}^*]$ generated by $\mathfrak{S}$ is dense in $C_c(X, \mathbb{F})$ in the l^∞-norm.*

Proof. Let $\widehat{X} = X \cup \{\infty\}$ be the one-point compactification of X. Extend each $f \in C_c(X, \mathbb{F})$ to $\widetilde{f} \in C(\widehat{X}, \mathbb{F})$ by setting $f(\infty) = 0$. By (1) and (2), the set $\widehat{\mathfrak{S}} = \{\widetilde{f} : f \in C_c(X, \mathbb{F})\} \sqcup \{1_{\widehat{X}}\}$ separates the points of $\widehat{X}$. (In particular, condition (2) guarantees that $\widehat{\mathfrak{S}}$ separates ∞ from any point of X.) By Thm. 1.5.12, $\mathbb{F}[\widehat{\mathfrak{S}} \cup \widehat{\mathfrak{S}}^*]$ is dense in $C(\widehat{X}, \mathbb{F})$. This implies that $\mathbb{F}[\mathfrak{S} \cup \mathfrak{S}^*]$ is dense in $C_c(X, \mathbb{F})$. $\square$

The following application of the Stone-Weierstrass theorem will only be used in Thm. 1.7.14 on the separability of L^p-spaces. (The easy direction (a)$\Rightarrow$(b) will also be used in Exp. 7.A.2.)

Theorem 1.5.14. *Let X be a compact Hausdorff space. Then the following are equivalent:*

(a) X is metrizable.

(b) X is second countable.

(c) There is a sequence $(f_n)_{n\in\mathbb{Z}_+}$ in $C(X,\mathbb{F})$ separating the points of X.

(d) $C(X,\mathbb{F})$ is separable.

The Stone-Weierstrass theorem will be used in the direction (c)⇒(d). The equivalence of (a), (b), and (c) does not rely on the Stone-Weierstrass theorem. In this course, we will mainly use the equivalence of (b), (c), and (d).

Proof. (a)⇒(b): Fix a metric on X. By the compactness, for each $n\in\mathbb{Z}_+$, X is covered by finitely many open balls with radius $1/n$. One checks easily that the collection of all these open balls for all $n\in\mathbb{Z}_+$ is a countable basis of the topology of X.

(b)⇒(c): Since X is second countable, we can choose an infinite countable base $(U_n)_{n\in\mathbb{Z}}$ of the topology. For each $m,n\in\mathbb{Z}_+$, if $U_n\Subset U_m$, we choose $f_{m,n}\in C_c(U_m,[0,1])\subset C_c(X,[0,1])$ such that $f|_{\overline{U_n}}=1$ (which exists by Urysohn's lemma); otherwise, we let $f_{m,n}=0$.

Let us prove that $\{f_{m,n}:m,n\in\mathbb{Z}_+\}$ separates the points of X: Choose distinct $x,y\in X$. Since $X\backslash\{y\}\in\mathrm{Nbh}_X(x)$, there exists U_m containing x and is contained in $X\backslash\{y\}$. By Cor. 1.4.24, there exists n such that $\{x\}\subset U_n\Subset U_m$. Then $f_{m,n}(x)=1$ and $f_{m,n}(y)=0$.

(c)⇒(a,b): Since (f_n) separates points, the map

$$\Phi=\bigvee_{n\in\mathbb{Z}_+}f_n:X\rightarrow\mathbb{F}^{\mathbb{Z}_+}\qquad x\mapsto(f_n(x))_{n\in\mathbb{Z}_+}$$

is injective. By Cor. 1.4.15, Φ is continuous. Since X is compact, the map Φ restricts to a homeomorphism $\Phi:X\rightarrow\Phi(X)$, where $\Phi(X)$ is equipped with the subspace topology of the product topology of $\mathbb{F}^{\mathbb{Z}_+}$. By Prop. 1.4.16, $\mathbb{F}^{\mathbb{Z}_+}$ is metrizable and second countable, so $\Phi(X)$, and hence X, is metrizable and second countable. This proves (a) and (b).

(c)⇒(d): Let $\mathbb{K}=\mathbb{F}\cap(\mathbb{Q}+\mathbf{i}\mathbb{Q})$. By the Stone-Weierstrass Thm. 1.5.13, the countable set $\mathbb{K}[\{f_n:n\in\mathbb{Z}_+\}]$ is dense in $C(X,\mathbb{F})$. Thus $C(X,\mathbb{F})$ is separable.

(d)⇒(c): By Exp. 1.5.11, $C(X,\mathbb{F})$ separates the points of X. Therefore, any dense subset of $C(X,\mathbb{F})$ separates the points of X. Since $C(X,\mathbb{F})$ is separable, it has a countable dense subset separating the points of X. $\square$

1.6 Review of measure theory: general facts

Recall the following basic property:

Proposition 1.6.1. *Let (X, μ) be a measure space, and let $f : X \to [0, +\infty]$ be measurable. Then $\int_X f d\mu = 0$ iff $f = 0$ μ-a.e..*

1.6.1 Some useful definitions and their basic properties

Definition 1.6.2. Let X be a set. Let $\mathbb{K} \in \{\mathbb{R}_{\geqslant 0}, \mathbb{R}, \mathbb{C}\}$. Suppose that $\mathscr{C}$ is an $\mathbb{K}$-linear subspace of $\mathbb{K}^X$. A **positive linear functional** on $\mathscr{C}$ denotes an $\mathbb{K}$-linear map $\Lambda : \mathscr{C} \to \mathbb{K}$ such that $\Lambda(f) \geqslant 0$ for all $f \in \mathscr{C} \cap \mathbb{R}^X_{\geqslant 0}$. Note that this condition is redundant when $\mathbb{K} = \mathbb{R}_{\geqslant 0}$.

Recall that if $(X, \mathfrak{M})$ is a measurable space, an **$\mathbb{F}$-valued simple function** on X is an $\mathbb{F}$-linear combination of characteristic functions over measurable sets; that is, an element of $\mathrm{Span}_{\mathbb{F}}\{\chi_E : E \in \mathfrak{M}\}$.

Definition 1.6.3. Let X be a set. Let $x \in X$. The **Dirac measure $\boldsymbol{\delta_x}$** of x is defined to be the measure $\delta_x : 2^X \to \overline{\mathbb{R}}_{\geqslant 0}$ satisfying $\delta_x(A) = 1$ if $x \in A$, and $\delta_x(A) = 0$ if $x \notin A$.

Definition 1.6.4. Let $(X, \mathcal{T}_X)$ be a topological space. Let $\mathfrak{M} \subset 2^X$ be a σ-algebra containing the Borel σ-algebra $\mathfrak{B}_X$. Let $\mu : \mathfrak{M} \to \overline{\mathbb{R}}_{\geqslant 0}$ be a measure. Assume that one of the following conditions holds:

(1) X is second countable.

(2) X is LCH, and $\mu|_{\mathfrak{B}_X}$ is a Radon measure.

The **support $\mathrm{Supp}(\boldsymbol{\mu})$** is defined to be

$$\mathrm{Supp}(\mu) = \{x \in X : \mu(U) > 0 \text{ for each } U \in \mathrm{Nbh}_X(x)\}$$

Then $\mathrm{Supp}(\mu)$ is a closed subset of X, because we clearly have

$$X \backslash \mathrm{Supp}(\mu) = \bigcup_{U \in \mathcal{T}_X, \mu(U) = 0} U$$

Moreover, we have $\mu(X \backslash \mathrm{Supp}(\mu)) = 0$. Thus, $\mathrm{Supp}(\mu)$ is the smallest closed subset whose complement is μ-null.

Proof that $X \backslash \mathrm{Supp}(\mu)$ is null. It suffices to show that if a family of open subsets $(U_\alpha)_{\alpha \in \mathscr{A}}$ is null, then the union $U := \bigcup_\alpha U_\alpha$ is null.

Assume that condition (1) holds. Since any subset of a second countable space is second countable and hence Lindelöf, the set U is Lindelöf. So (U_α) has a countable subfamily covering U. Therefore, by the countable additivity, U is null.

Assume that condition (2) holds. Since Radon measures are inner regular on open sets (cf. Def. 1.7.3), $\mu(U)$ is the supremum of $\mu(K)$ where K runs through all compact subsets of U. Since K is compact, (U_α) has a finite subfamily covering K. Therefore K is null, and hence U is null. $\square$

Lemma 1.6.5. *Let $\mu : \mathfrak{M} \to \overline{\mathbb{R}}_{\geqslant 0}$ be as in Def. 1.6.4, and assume that Condition (1) or (2) of Def. 1.6.4 holds. The following are equivalent:*

(a) $\mathrm{Supp}(\mu)$ *is a finite set.*

(b) μ *is a linear combination of Dirac measures (restricted to $\mathfrak{M}$).*

Proof. (b)$\Rightarrow$(a): This is obvious.

(a)$\Rightarrow$(b): Write $E = \mathrm{Supp}(\mu)$. Choose any measurable $f : X \to \overline{\mathbb{R}}_{\geqslant 0}$. Then, since $\mu|_{X\setminus E} = 0$, the integral of any measurable function $g : X \to \overline{\mathbb{R}}_{\geqslant 0}$ vanishing ourside E is zero. In particular, we can choose g to be the unique one such that $g + \sum_{x\in E} f(x)\chi_{\{x\}} = f$. Therefore

$$\int_X f d\mu = \int_X \sum_{x\in E} f(x)\chi_{\{x\}} d\mu = \sum_{x\in E} f(x) \cdot \mu(\{x\})$$

This shows that $\mu = \sum_{x\in E} \mu(\{x\})\delta_x$. $\square$

Definition 1.6.6. Let $\Phi : (X, \mathfrak{M}) \to (Y, \mathfrak{N})$ be a measurable map between two measurable spaces. Let μ be a measure on $\mathfrak{M}$. The **pushforward measure** of μ by Φ is defined to be

$$\Phi_*\mu : \mathfrak{N} \to [0, +\infty] \qquad B \mapsto \mu(\Phi^{-1}(B)) \tag{1.29}$$

Then

$$\int_Y f d\Phi_*\mu = \int_X (f \circ \Phi) d\mu \tag{1.30}$$

holds for any simple function $f : X \to \mathbb{R}_{\geqslant 0}$, and hence (by MCT) for any measurable $f : X \to \overline{\mathbb{R}}_{\geqslant 0}$. If Y is a topological space and $\mathfrak{N}$ contains $\mathfrak{B}_Y$, we call

$$\mathbf{Rng^{ess}(\Phi, \mu)} := \mathrm{Supp}(\Phi_*\mu)$$

the **essential range** of Φ with respect to μ and abbreviate it to $\mathrm{Rng}^{\mathrm{ess}}(\Phi)$ when no confusion arises.

Remark 1.6.7. Let (X, μ) be a measure space, and let $f : X \to \mathbb{C}$ be measurable. It is a standard fact that

$$\|f\|_{l^\infty(X,\mu)} = \sup\{|z| : z \in \mathrm{Rng}^{\mathrm{ess}}(f, \mu)\}$$

Definition 1.6.8. Let $(X,\mathfrak{M})$ and $(Y,\mathfrak{N})$ be measurable spaces. A bijection $f : X \to Y$ such that both f and f^{-1} are measurable is called a **measurable isomorphism**. If X, Y are topological spaces and $\mathfrak{M} = \mathfrak{B}_X, \mathfrak{N} = \mathfrak{B}_Y$, a measurable isomorphism is called a **Borel isomorphism**.

If $f : X \to Y$ is a measurable bijection and $\nu : Y \to [0,+\infty]$ is a measure, then

$$\boldsymbol{\Phi^*\nu} = (\Phi^{-1})_*\nu$$

is called the **pullback measure** of ν by Φ. It can be equivalently described by

$$\int_Y f d\nu = \int_X (f \circ \Phi) d\Phi^*\mu \tag{1.31}$$

for each $f : X \to \overline{\mathbb{R}}_{\geqslant 0}$. □

1.6.2 Radon-Nikodym derivatives

Fix a measurable space $(X,\mathfrak{M})$.

Definition 1.6.9. Let $\mu, \nu : \mathfrak{M} \to [0,+\infty]$ be measures. We say that ν is **absolutely continuous** with respect to μ and write $\boldsymbol{\nu \ll \mu}$ if any μ-null set is ν-null. We say that $h \in \mathcal{L}(X, \overline{\mathbb{R}}_{\geqslant 0})$ is a **Radon-Nikodym derivative** of ν with respect to μ if

$$\int_X f d\nu = \int_X f h d\mu \qquad \text{for all } f \in \mathcal{L}(X, \overline{\mathbb{R}}_{\geqslant 0})$$

By MCT, the above condition is equivalent to

$$\nu(E) = \int_E h d\mu \qquad \text{for all } E \in \mathfrak{M}$$

We write $\boldsymbol{d\nu = h d\mu}$.

Remark 1.6.10. If μ is σ-finite, and if h_1, h_2 are both Radon-Nikodym derivatives of ν with respect to μ, then $h_1(x) = h_2(x)$ for μ-a.e. $x \in X$.

Proof. It suffices to assume that $\mu(X) < +\infty$. For each $k \in \mathbb{N}$, let

$$A_k = \{x \in X : h_1(x) < h_2(x) \text{ and } h_1(x) \leqslant k\}$$

Then $\int_{A_k} h_1 d\mu \leqslant k\mu(X) < +\infty$, and

$$\int_{A_k} h_1 d\mu = \int_{A_k} d\nu = \int_{A_k} h_2 d\mu$$

Taking subtraction, we get $\int_{A_k} (h_2 - h_1) d\mu = 0$. Let $A = \bigcup_k A_k = \{x \in X : h_1(x) < h_2(x)\}$. By MCT, $\int_A (h_2 - h_1) d\mu = 0$. Since $h_2 - h_1 \geqslant 0$ on A, we conclude from Prop. 1.6.1 that $h_2 - h_1 = 0$ μ-a.e. on A, and hence $\mu(A) = 0$. Similarly, $\mu(B) = 0$ where $B = \{x \in X : h_1(x) > h_2(x)\}$. □

Remark 1.6.11. If ν is σ-finite, and if $d\nu = hd\mu$, then $h(x) < +\infty$ for μ-a.e. $x \in X$.

Proof. Let $A = \{x \in A : h(x) = +\infty\}$. Since ν is σ-finite, we can write $A = \bigcup_{k \in \mathbb{N}} A_k$ where $A_k \in \mathfrak{M}$ and $\nu(A_k) < +\infty$. Since $\nu(A_k) = \int_{A_k} hd\mu = +\infty\mu(A_k)$, we have $\mu(A_k) = 0$, and hence $\mu(A) = 0$. □

Theorem 1.6.12 (Radon-Nikodym theorem). *Assume that $\mu, \nu : \mathfrak{M} \to [0, +\infty]$ are σ-finite measures. Then $\nu \ll \mu$ iff ν has a Radon-Nikodym derivative with respect to μ.*

Proof. "$\Leftarrow$" is obvious. Let us prove "$\Rightarrow$". It is easy to reduce to the case that $\mu(X), \nu(X) < +\infty$. Let $d\psi = d\mu + d\nu$. So $\mu, \nu \leqslant \psi$. Therefore, the linear functional

$$\Lambda : L^2(X, \psi) \to \mathbb{C} \qquad \xi \mapsto \int_X \xi d\mu$$

is bounded (with operator norm $\leqslant \psi(X)$). Since $L^2(X, \psi)$ is a Hilbert space (Thm. 1.6.14), by the Riesz-Fréchet Thm. 3.5.3, there exists $f \in L^2(X, \psi)$ such that $\int_X \xi d\mu = \int_X \xi f d\psi$ for all $\xi \in L^2(X, \psi)$. Since Λ sends positive functions to $\mathbb{R}_{\geqslant 0}$, after replacing f with $\mathrm{Re}(f)$ and adding a measurable ψ-a.e. zero function to f, we have $f \geqslant 0$ everywhere.

We have found $f \in \mathcal{L}(X, \mathbb{R}_{\geqslant 0})$ such that $d\mu = fd\psi$. Similarly, we have $g \in \mathcal{L}(X, \mathbb{R}_{\geqslant 0})$ such that $d\nu = gd\psi$. Since $\mu \leqslant \psi \ll \mu$, we have $f > 0$ outside a ψ-null set Δ. Let $h = g/f$ outside Δ, and $h = 0$ on Δ. Then $d\nu = hd\mu$. □

1.6.3 L^p-spaces

Let $(X, \mathfrak{M}, \mu)$ be a measure space. Let $1 \leqslant p, q \leqslant +\infty$ such that $p^{-1} + q^{-1} = 1$.

Theorem 1.6.13. *Let $1 \leqslant p < +\infty$. Then the set of integrable $\mathbb{F}$-valued simple functions is dense in $L^p(X, \mu, \mathbb{F})$. In other words,*

$$\{\chi_E : E \subset \mathfrak{M}, \mu(E) < +\infty\}$$

spans a dense subspace of $L^p(X, \mu, \mathbb{F})$.

Proof. See [Gui-A, Sec. 27.2]. □

Theorem 1.6.14 (Riesz-Fischer theorem, the modern form**).** *The normed vector space $L^p(X, \mu, \mathbb{F})$ is (Cauchy) complete. Moreover, if (f_n) is a sequence in $L^p(X, \mu, \mathbb{F})$ converging (in L^p) to $f \in L^p(X, \mu, \mathbb{F})$, then (f_n) has a subsequence converging μ-a.e. to f.*

Proof. See [Gui-A, Sec. 27.3]. □

Lemma 1.6.15. *Assume that (X, μ) is σ-finite. Let $\mathcal{S}_+$ be the set of simple functions $X \to \mathbb{R}_{\geqslant 0}$. Then for each $f \in \mathcal{L}(X, \overline{\mathbb{R}}_{\geqslant 0})$ we have*

$$\|f\|_{L^p(X,\mu)} = \sup\left\{\int_X fgd\mu : g \in \mathcal{S}_+, \|g\|_{L^q(X,\mu)} \leqslant 1\right\} \tag{1.32}$$

Consequently, for each $f \in \mathcal{L}(X, \mathbb{C})$ we have

$$\|f\|_{L^p(X,\mu)} = \sup\left\{\int_X |fg| : g \in L^q(X, \mu), \|g\|_q \leqslant 1\right\} \tag{1.33}$$

Proof. By Hölder's inequality, we have "$\geqslant$". To prove "$\leqslant$", we note that (1.33) follows immediately from (1.32) by writing $f = u|f|$ where $u \in \mathcal{L}(X, \mathbb{S}^1)$ and applying (1.32) to $|f|$. Thus, in the following, we assume $f \in \mathcal{L}(X, \overline{\mathbb{R}}_{\geqslant 0})$. Moreover, we assume $\|f\|_{L^p} > 0$; otherwise, the inequality is trivial.

Case $1 < p < +\infty$: Choose an increasing sequence (f_n) (i.e. $f_1 \leqslant f_2 \leqslant \cdots$) in $\mathcal{S}_+$ converging pointwise to f such that each f_n vanishes outside a measurable μ-finite set. Let $g_n = (f_n)^{p-1}$. After removing the first several terms, we assume $\|g_n\|_{L^q} > 0$ for all n. Then

$$0 < \|g_n\|_q = \|f_n\|_p^{p/q} < +\infty$$

By MCT, we have $\lim_n \|g_n\|_p = \|f\|_p^{p/q}$ and $\lim_n \int_X fg_n = \|f\|_p^p$. Thus, if $\|f\|_p < +\infty$, then

$$\lim_n \|g_n\|_q^{-1} \int_X fg_n = \|f\|_p^{-p/q} \cdot \|f\|_p^p = \|f\|_p$$

This proves (1.32) when $\|f\|_p < +\infty$. If $\|f\|_p = +\infty$, then, by MCT, $\|f_n\|_p < +\infty$ can be sufficiently large. Applying (1.32) to f_n instead of f, we obtain $g \in \mathcal{S}_+$ such that $\|g\|_q \leqslant 1$ and $\int f_n g$ is sufficiently large, and hence $\int fg$ is sufficiently large. Thus (1.32) holds again.

Case $p = 1$: Let $g = 1$.

Case $p = +\infty$: Write $X = \bigcup_{n \in \mathbb{N}} \Omega_n$ where $\Omega_n \in \mathfrak{M}$ and $\mu(\Omega_n) < +\infty$. Choose any $0 \leqslant \lambda < \|f\|_\infty$. Then $A := \{f > \lambda\}$ satisfies $\mu(A) > 0$. Thus, there exists n such that $0 < \mu(A \cap \Omega_n) < +\infty$. Let $g = \chi_{A \cap \Omega_n}/\mu(A \cap \Omega_n)$. Then $g \in \mathcal{S}_+$, $\|g\|_1 = 1$, and $\int fg \geqslant \lambda$. This proves (1.32). □

Theorem 1.6.16. *Assume that (X, μ) is σ-finite. Assume $1 < p \leqslant +\infty$. Then we have an isomorphism of normed vector spaces*

$$\Psi : L^p(X, \mu, \mathbb{F}) \to L^q(X, \mu, \mathbb{F})^* \qquad f \mapsto \left(g \in L^q(X, \mu, \mathbb{F}) \mapsto \int_X fgd\mu\right) \tag{1.34}$$

When $p < +\infty$, the assumption on σ-finiteness can be removed. See [Fol-R, Sec. 6.2]. When $p = 2$, this is simply due to the completeness of $L^2(X, \mu, \mathbb{F})$ and the Riesz-Fréchet theorem.

Proof. By Hölder's inequality and Lem. 1.6.15, Ψ is an isometry. Let us show that any $\Lambda \in L^q(X,\mu,\mathbb{F})^*$ belongs to the range of Ψ.

Step 1. By considering the real and imaginary parts

$$(\mathrm{Re}\Lambda)(f) = \frac{\Lambda(f) + \overline{\Lambda(\overline{f})}}{2} \qquad (\mathrm{Im}\Lambda)(f) = \frac{\Lambda(f) - \overline{\Lambda(\overline{f})}}{2\mathbf{i}}$$

separately, we can assume that Λ is real, i.e., $\Lambda(f) \in \mathbb{R}$ for any $f \in L^q(X,\mu,\mathbb{R}_{\geqslant 0})$.

Let us define $\mathbb{R}_{\geqslant 0}$-linear maps $\Lambda^+,\Lambda^- : L^q(X,\mu,\mathbb{R}_{\geqslant 0}) \to \mathbb{R}_{\geqslant 0}$ with operator norms $\leqslant \|\Lambda\|$, i.e.,

$$\|\Lambda^\pm(g)\| \leqslant \|\Lambda\| \cdot \|g\|_q \qquad \text{for all } g \in L^q(X,\mu,\mathbb{R}_{\geqslant 0}) \tag{1.35}$$

and let us check that

$$\Lambda(g) = \Lambda^+(g) - \Lambda^-(g) \qquad \text{for all } g \in L^q(X,\mu,\mathbb{R}_{\geqslant 0}) \tag{1.36}$$

Eq. (1.36) is called the **Jordan decomposition** of Λ.

Define the $\Lambda^\pm : L^q(X,\mu,\mathbb{R}_{\geqslant 0}) \to \overline{\mathbb{R}}$ by sending each $g \in L^q(X,\mu,\mathbb{R}_{\geqslant 0})$ to

$$\Lambda^\pm(g) = \sup\{\pm\Lambda(h) : h \in L^q(X,\mu,\mathbb{R}_{\geqslant 0}), h \leqslant g\} \tag{1.37}$$

Since $0 \leqslant g$, we clearly have $\Lambda^+(g) \geqslant 0$. Since Λ is bounded and $\|h\|_q \leqslant \|g\|_q$, we clearly have $\|\Lambda^+(g)\| \leqslant \|\Lambda\| \cdot \|g\|_q$. In particular, Λ^+ has range in $\mathbb{R}_{\geqslant 0}$. Since $\Lambda^\pm = (-\Lambda)^\mp$, a similar property holds for Λ^-. Thus, we have checked (1.35).

Clearly, for each $f,g \in L^q(X,\mu,\mathbb{R}_{\geqslant 0})$, we have $\Lambda^+(f+g) \geqslant \Lambda^+(f) + \Lambda^+(g)$. To prove the other direction, choose any $h \in L^q(X,\mu,\mathbb{R}_{\geqslant 0})$ such that $h \leqslant f+g$. Let $h_1 = fh/(f+g)$ and $h_2 = gh/(f+g)$, understood to be zero where the denominator vanishes. Then $h_1,h_2 \in L^q(X,\mu,\mathbb{R}_{\geqslant 0})$ and $h_1 \leqslant f$ and $h_2 \leqslant g$. This proves $\Lambda^+(f+g) \leqslant \Lambda^+(f) + \Lambda^+(g)$. Thus Λ^+ (and similarly Λ^-) is $\mathbb{R}_{\geqslant 0}$-linear.

From (1.37), one easily checks $\Lambda(g) + \Lambda^-(g) \leqslant \Lambda^+(g)$ for each $g \in L^q(X,\mu,\mathbb{R}_{\geqslant 0})$. Replacing Λ with $-\Lambda$, we get $-\Lambda(g) + \Lambda^+(g) \leqslant \Lambda^-(g)$. Thus (1.36) holds.

Step 2. Let us prove that Λ^+ is represented by some $f^+ \in L^p(X,\mu,\mathbb{R}_{\geqslant 0})$, namely,

$$\Lambda^+(g) = \int_X f^+ g d\mu \qquad \text{for all } g \in L^q(X,\mu,\mathbb{R}_{\geqslant 0}) \tag{1.38}$$

Then, similarly, Λ^- is represented by some $f^- \in L^p(X,\mu,\mathbb{R}_{\geqslant 0})$. Thus Λ is represented by $f^+ - f^-$, finishing the proof.

Write $X = \bigsqcup_n X_n$ where $\mu(X_n) < +\infty$. Suppose that we can find $f_n^+ \in L^p(X_n,\mu)$ representing $\Lambda^+|_{L^q(X_n,\mu)}$, then we can define $f^+ : X \to \mathbb{R}_{\geqslant 0}$ such that $f^+|_{X_n} = f_n$ for all n. Clearly f^+ represents Λ^+. In particular, by Lem. 1.6.15 and (1.35), $\|f^+\|_p \leqslant \|\Lambda\| < +\infty$. Thus $f \in L^p(X,\mu)$.

Therefore, according to the previous paragraph, we may assume at the beginning that $\mu(X) < +\infty$. Define

$$\nu : \mathfrak{M} \to [0, +\infty] \qquad E \mapsto \Lambda(\chi_E)$$

Then one checks easily that ν is a measure,[1] and that $\nu \ll \mu$. Therefore, by the Radon-Nikodym Thm. 1.6.12, there exists $f^+ \in \mathcal{L}(X, \mathbb{R}_{\geqslant 0})$ such that $d\nu = f^+ d\mu$. Thus

$$\Lambda^+(g) = \int_X g d\nu = \int_X f^+ g d\mu \quad \text{for each simple function } g \in L^q(X, \mu, \mathbb{R}_{\geqslant 0}) \tag{1.39}$$

Lem. 1.6.15 and (1.35) then imply $\|f^+\|_p \leqslant \|\Lambda\| < +\infty$, and hence $f \in L^p(X, \mu, \mathbb{R}_{\geqslant 0})$.

Finally, for $g \in L^q(X, \mu, \mathbb{R}_{\geqslant 0})$, find an increasing sequence of simple functions $g_n \in L^q(X, \mu, \mathbb{R}_{\geqslant 0})$ converging pointwise to g. By (1.35), $\Lambda^+(g - g_n) \leqslant \|\Lambda\| \cdot \|g - g_n\|_q$ where the RHS converges to zero by DCT. By MCT, $\int_X f^+ g_n d\mu \to \int_X f^+ g d\mu$. Thus, by (1.39), we conclude (1.38). $\square$

1.7 Review of measure theory: Radon measures

1.7.1 Radon measures and the Riesz-Markov representation theorem

Let X be LCH. The reference for this subsection is [Gui-A, Ch. 25].

Definition 1.7.1. Let $\mathfrak{M} \subset 2^X$ be a σ algebra containing $\mathfrak{B}_X$, and let $\mu : \mathfrak{M} \to \overline{\mathbb{R}}_{\geqslant 0}$ be a measure. Let $E \in \mathfrak{M}$. We say that μ is **outer regular** on E if

$$\mu(E) = \inf\{\mu(U) : U \supset E, U \text{ is open}\}$$

We say that μ is **inner regular** on E if

$$\mu(E) = \sup\{\mu(K) : K \subset E, K \text{ is compact}\}$$

We say that μ is **regular** on E if μ is both outer and inner regular on E.

Lemma 1.7.2. *Let $\mu : \mathfrak{B}_X \to \overline{\mathbb{R}}_{\geqslant 0}$ be a Borel measure. Let $U \subset X$ be open. Then*

$$\sup\left\{\mu(K) : K \subset U, K \textit{ is compact}\right\} = \sup\left\{\int_X f d\mu : f \in C_c(U, [0,1])\right\}$$

Therefore, μ is inner regular on U iff

$$\mu(U) = \sup\left\{\int_X f d\mu : f \in C_c(U, [0,1])\right\}$$

[1] To check the countable additivity, we let $E_1 \subset E_2 \subset \cdots$ be measurable and $E = \bigcup_n E_n$. Let $F_n = E \backslash E_n$. By (1.35), $\nu(F_n) \leqslant \|\Lambda\| \mu(F_n)^{\frac{1}{q}} \to 0$. Thus $\nu(E_n) \to \nu(E)$.

Proof. Let A, B denote the LHS and the RHS. If $f \in C_c(U, [0,1])$, then setting $K = \mathrm{Supp}(f)$, we have $\mu(K) = \int_X \chi_K d\mu \geqslant \int_X f d\mu$. This proves $A \geqslant B$.

Conversely, let $K \subset U$. By Urysohn's lemma, there exists $f \in C_c(U, [0,1])$ such that $f|_K = 1$. So $\mu(K) = \int_X \chi_K d\mu \leqslant \int_X f d\mu$. This proves $A \leqslant B$. $\square$

Definition 1.7.3. A Borel measure $\mu : \mathfrak{B}_X \to \overline{\mathbb{R}}_{\geqslant 0}$ is called a **Radon measure** if the following conditions are satisfied:

(a) μ is outer regular on Borel sets.

(b) μ is inner regular on open sets. Equivalently, for each open $U \subset X$, we have

$$\mu(U) = \sup\Big\{\int_X f d\mu : f \in C_c(U, [0,1])\Big\} \tag{1.40}$$

(c) $\mu(K) < +\infty$ if K is a compact subset of X. Equivalently, for each $f \in C_c(X, \mathbb{R}_{\geqslant 0})$ we have

$$\int_X f d\mu < +\infty \tag{1.41}$$

Proof of equivalence. The equivalence in (b) is due to Lem. 1.7.2. The equivalence in (c) can be proved in a similar way to Lem. 1.7.2. $\square$

Note that Radon measures are determined by their integrals against functions in $C_c(X, [0,1])$. Indeed, by (a), μ is determined by its values on open sets. (b) shows that those values on open sets are determined by the integrals $\int_X f d\mu$ where $f \in C_c(X, [0,1])$.

Remark 1.7.4. There exist canonical bijections among:

- $\mathbb{R}_{\geqslant 0}$-linear maps $C_c(X, \mathbb{R}_{\geqslant 0}) \to \mathbb{R}_{\geqslant 0}$.
- Positive linear functionals on $C_c(X, \mathbb{R})$.
- Positive linear functionals on $C_c(X) = C_c(X, \mathbb{C})$.

Proof. An $\mathbb{R}_{\geqslant 0}$-linear map $\Lambda : C_c(X, \mathbb{R}_{\geqslant 0}) \to \mathbb{R}_{\geqslant 0}$ can be extended uniquely to a linear map $\Lambda : C_c(X, \mathbb{R}) \to \mathbb{R}$ due to the following Lem. 1.7.5. The latter can be extended to a linear functional on $C_c(X)$ by setting

$$\Lambda(f) = \Lambda(\mathrm{Re} f) + \mathbf{i}\Lambda(\mathrm{Im} f) \tag{1.42}$$

for all $C_c(X)$. $\square$

Lemma 1.7.5. *Let K be an $\mathbb{R}_{\geqslant 0}$-linear subspace of an $\mathbb{R}$-vector space V. Let W be an $\mathbb{R}$-linear space. Let $\Gamma : K \to W$ be an $\mathbb{R}_{\geqslant 0}$-linear map. Suppose that $V = \mathrm{Span}_{\mathbb{R}} K$. Then Γ can be extended uniquely to an $\mathbb{R}$-linear map $\Lambda : V \to W$.*

Proof. The uniqueness is obvious. To prove the existence, note that any $v \in V$ can be written as

$$v = v^+ - v^-$$

where $v^+, v^- \in K$. (Proof: Since $V = \mathrm{Span}_{\mathbb{R}} K$, we have $v = a_1u_1 + \cdots + a_m u_m - b_1 w_1 - \cdots - b_n w_n$ where each u_i, w_j are in K, and each a_i, b_j are in $\mathbb{R}_{\geqslant 0}$. One sets $v^+ = \sum_i a_i u_i$ and $v^- = \sum_j b_j w_j$.) We then define $\Lambda(v) = \Gamma(v^+) - \Gamma(v^-)$.

Let us show that this gives a well-defined map $\Lambda : V \to W$. Assume that $v = v^+ - v^- = w^+ - w^-$ where $v^\pm, w^\pm \in K$. Then $\Gamma(v^+) - \Gamma(v^-) = \Gamma(w^+) - \Gamma(w^-)$ iff $\Gamma(v^+) + \Gamma(w^-) = \Gamma(v^-) + \Gamma(w^+)$, iff (by the additivity of Γ) $\Gamma(v^+ + w^-) = \Gamma(v^- + w^+)$. The last statement is true because $v^+ - v^- = w^+ - w^-$ implies $v^+ + w^- = v^- + w^+$.

It is easy to see that Λ is additive. If $c \geqslant 0$, then $cv = cv^+ - cv^-$ where $cv^+, cv^- \in K$. So $\Lambda(cv) = \Gamma(cv^+) - \Gamma(cv^-)$, which (by the $\mathbb{R}_{\geqslant 0}$-linearity of Γ) equals $c\Gamma(v^+) - c\Gamma(v^-) = c\Lambda(v)$. Since $-v = v^- - v^+$, we have $\Lambda(-v) = \Gamma(v^-) - \Gamma(v^+) = -\Lambda(v)$. Hence $\Lambda(-cv) = c\Lambda(-v) = -c\Lambda(v)$. This proves that Λ commutes with the $\mathbb{R}$-multiplication. □

Theorem 1.7.6 (Riesz-Markov representation theorem). *For every positive linear $\Lambda : C_c(X, \mathbb{F}) \to \mathbb{F}$ there exists a unique Radon measure $\mu : \mathfrak{B}_X \to \overline{\mathbb{R}}_{\geqslant 0}$ such that*

$$\Lambda(f) = \int_X f d\mu \tag{1.43}$$

for all $f \in C_c(X, \mathbb{F})$. Moreover, every Radon measure on X arises from some Λ in this way.

In addition, the operator norm $\|\Lambda\|$ equals $\mu(X)$. Therefore, Λ is bounded iff μ is a finite measure.

Proof. See [Gui-A, Sec. 25.3] for the first paragraph. The second paragraph asserts that

$$\sup_{f \in \overline{B}_{C_c(X)}(0,1)} |\Lambda(f)| = \mu(X)$$

The inequality "$\leqslant$" is obvious. The reverse inequality "$\geqslant$" follows from (1.40). □

1.7.2 Basic properties of Radon measures

Theorem 1.7.7. *Let μ be a Radon measure (or its completion) on X. Then μ is regular on any measurable set E satisfying $\mu(E) < +\infty$.*

Proof. See [Gui-A, Sec. 25.4]. A sketch of the proof (different from that in [Gui-A]) is as follows.

Assume WLOG that E is Borel. Since Radon measures are outer regular on Borel sets, it remains to prove that μ is inner regular on E. Pick an open set U

such that $\mu(U\backslash E)$ is small. Since μ is inner regular on U, there is a compact $K \subset U$ such that $\mu(U\backslash K)$ is small. However, K is not necessarily contained in E.

To fix this issue, we note that since μ is outer regular on $U\backslash E$, we can find an open set $V \subset U$ containing $U\backslash E$ whose measure is close to $\mu(U\backslash E)$. In particular, $\mu(V)$ is small. Then $K\backslash V$ is a compact subset of E whose measure is close to $\mu(E)$. $\square$

The following criterion for Radon measures is of fundamental importance and will be used repeatedly throughout this course.

Theorem 1.7.8. *Assume that X is second countable. Let μ be a Borel measure on X. Then μ is Radon iff $\mu(K) < +\infty$ for any compact $K \subset X$.*

In particular, a finite Borel measure on $\mathbb{R}^n$ (where $n \in \mathbb{N}$) is Radon.

Proof. See [Gui-A, Sec. 25.5]. The rough idea is as follows. By Thm. 1.7.6, there exists a Radon measure ν such that $\int f d\mu = \int f d\nu$ for all $f \in C_c(X)$. It suffices to show $\mu = \nu$.

If $U \subset X$ is open, then U is σ-compact by Rem. 1.4.34. Using Urysohn's lemma, we can find an increasing sequence (f_n) in $C_c(I, [0,1])$ converging pointwise to χ_U. By MCT, we have $\mu(U) = \lim_n \int f_n d\mu = \lim_n \int f_n d\nu = \nu(U)$.

Next, let $E \in \mathfrak{B}_X$ such that $\nu(E) < +\infty$. Since ν is Radon, by Thm. 1.7.7, there exist a compact $K \subset E$ and an open $U \supset E$ such that $\nu(U\backslash K)$ is small. By the above paragraph, $\mu(U\backslash K) = \nu(U\backslash K)$ is also small, and $\mu(U) = \nu(U)$. Thus $\mu(E) \approx \mu(U) = \nu(U) \approx \nu(E)$. Hence $\mu(E) = \nu(E)$.

Finally, choose any $E \in \mathfrak{B}_X$. Since X is σ-compact, it is covered by an increasing sequence (K_n) of compact subsets. Since ν is Radon, we have $\nu(K_n) < +\infty$ and hence $\nu(E \cap K_n) < +\infty$. By the above paragraph, we have $\mu(E \cap K_n) = \nu(E \cap K_n)$. Hence $\mu(E) = \nu(E)$. $\square$

1.7.3 Approximation and density

The main reference for this subsection is [Gui-A, Sec. 27.2].

Theorem 1.7.9 (Lusin's theorem). *Let X be LCH. Let μ be a Radon measure (or its completion) on X with σ-algebra $\mathfrak{M}$. Let $f : X \to \mathbb{F}$ be measurable. Let $A \in \mathfrak{M}$ such that $\mu(A) < +\infty$. Then for each $\varepsilon > 0$ there exists a compact $K \subset A$ such that $\mu(A\backslash K) < \varepsilon$ and that $f|_K : K \to \mathbb{F}$ is continuous.*

With the help of the Tietze extension Thm. 1.4.29, Lusin's theorem implies that for each $\varepsilon > 0$ there exist a compact $K \subset A$ and some $\widetilde{f} \in C_c(X, \mathbb{F})$ such that $\widetilde{f}|_K = f|_K$ and $\mu(A\backslash K) < \varepsilon$.

Proof. See [Gui-A, Sec. 25.4]. The rough idea is that one first uses Thm. 1.7.7 to prove the case where $\widetilde{f} := f\chi_A$ is a simple function. The general case follows by choosing an increasing sequence of simple functions converging uniformly to $\widetilde{f}\chi_{F_n}$, where $F_n = |\widetilde{f}|^{-1}([0,n])$ and n is sufficiently large. □

Theorem 1.7.10. *Let $1 \leqslant p < +\infty$. Let μ be a Radon measure (or its completion) on an LCH space X. Then, in the L^p-norm, the space $C_c(X,\mathbb{F})$ is dense in $L^p(X,\mu,\mathbb{F})$. More precisely, the map $f \in C_c(X,\mathbb{F}) \mapsto f \in L^p(X,\mu,\mathbb{F})$ has dense range.*

Proof. See [Gui-A, Sec. 27.2]. Rough idea: Choose $f \in L^p$. Use Lusin's theorem to approximate $f\chi_{E_n}$ by functions of $C_c(X,\mathbb{F})$, where $E_n = \{1/n < |f| < n\}$ and n is sufficiently large. Alternatively, first approximate f by a simple function g (cf. Thm. 1.6.13). Then use Thm. 1.7.7 and Urysohn's lemma to approximate g by functions of $C_c(X,\mathbb{F})$. □

Theorem 1.7.11. *Let $1 \leqslant p < +\infty$. Let $k \in \mathbb{N} \cup \{\infty\}$. Assume that X is an open subset of $\mathbb{R}^n$ (or more generally, a C^k-manifold with or without boundary). Let μ be a Radon measure (or its completion) on X. Then, in the L^p-norm, the space $C_c^k(X,\mathbb{F})$ is dense in $L^p(X,\mu,\mathbb{F})$.*

Proof. See [Gui-A, Sec. 30.7]. Rough idea: Due to Thm. 1.7.10, a function $f \in L^p(X,\mu,\mathbb{F})$ can be approximated in L^p by some $g \in C_c(X,\mathbb{F})$. Choose a precompact open set $U \subset X$ containing $\mathrm{Supp}(g)$. By Rem. 1.4.30, g can be approximated uniformly by functions of $C_c^k(U,\mathbb{F})$. This is also an L^p-approximation, because $\mu(U) < +\infty$. □

Remark 1.7.12. One easily checks that

$$
\begin{aligned}
&\mathrm{Span}_{\mathbb{F}}\{\chi_I : I \subset \mathbb{R} \text{ is a bounded interval}\}\\
=&\mathrm{Span}_{\mathbb{F}}\{\chi_I : I \subset \mathbb{R} \text{ is a compact interval}\}\\
=&\mathrm{Span}_{\mathbb{F}}\{\chi_I : I \subset \mathbb{R} \text{ is a bounded open interval}\}
\end{aligned}
$$

An element in these sets is called an $\mathbb{F}$-valued **step function**. Moreover, one checks that

$$
\begin{aligned}
\{\text{right-continuous } \mathbb{F}\text{-valued step functions}\} &= \mathrm{Span}_{\mathbb{F}}\{\chi_{[a,b)} : a,b \in \mathbb{R}\}\\
\{\text{left-continuous } \mathbb{F}\text{-valued step functions}\} &= \mathrm{Span}_{\mathbb{F}}\{\chi_{(a,b]} : a,b \in \mathbb{R}\}
\end{aligned}
$$

Theorem 1.7.13. *Let $1 \leqslant p < +\infty$. Let μ be a Radon measure (or its completion) on $\mathbb{R}$. Then each of the following classes of functions form a dense subset of $L^p(\mathbb{R},\mu,\mathbb{F})$:*

(a) Right-continuous $\mathbb{F}$-valued step functions.

(b) Left-continuous $\mathbb{F}$-valued step functions.

(c) Elements of $\mathrm{Span}_{\mathbb{F}}\{\chi_{(-\infty,b]} : b \in \mathbb{R}\}$.

(d) *Elements of* $\mathrm{Span}_{\mathbb{F}}\{\chi_{(-\infty,b)} : b \in \mathbb{R}\}$.

Proof. With the help of Thm. 1.7.10, the density of (a) and (b) can be proved by approximating a function $f \in C_c(X, \mathbb{F})$ with left/right-continuous step functions. See [Gui-A, Sec. 27.2] for details.

Since (a)⊂(c) and (b)⊂(d), the density of (c) and (d) follows. □

Theorem 1.7.14. *Let* $1 \leqslant p < +\infty$. *Let* μ *be a Radon measure (or its completion) on a second countable LCH space* X. *Then* $L^p(X, \mu, \mathbb{F})$ *is separable.*

Proof. See [Gui-A, Sec. 27.2]. The rough idea is as follows. First assume that X is compact. Then $\mu(X) < +\infty$. By Thm. 1.5.14, $C(X, \mathbb{F})$ is l^∞-separable, and hence L^p-separable. Therefore, by Thm. 1.7.10, $L^p(X, \mu, \mathbb{F})$ is separable.

In the general case, one writes $X = \bigcup_n K_n$ where $K_1 \subset K_2 \subset \cdots$ are compact. Clearly $\sum_n L^p(K_n, \mu|_{K_n}, \mathbb{F})$ is dense in $L^p(X, \mu, \mathbb{F})$. By Thm. 1.7.8, each $\mu|_{K_n}$ is Radon. By the above paragraph, each $L^p(K_n, \mu|_{K_n}, \mathbb{F})$ is separable. Thus $L^p(X, \mu, \mathbb{F})$ is separable. □

1.7.4 Complex Radon measures

Definition 1.7.15. If X is a set and $\mathfrak{M} \subset 2^X$ is a σ-algebra, a **complex measure** (resp. **signed measure**) is a function $\mathfrak{M} \to \mathbb{C}$ (resp. $\mathfrak{M} \to \mathbb{R}$) that can be written as a $\mathbb{C}$-linear (resp. $\mathbb{R}$-linear) combination of finite measures on $\mathfrak{M}$.

We now assume that X is LCH.

Definition 1.7.16. A complex (resp. signed) measure on $\mathfrak{B}_X$ is called **Radon** if it is a $\mathbb{C}$-linear (resp. $\mathbb{R}$-linear) combination of finite Radon measures.

Suppose that μ is a complex Radon measure on X. Then similar to the proof of Rem. 1.7.4, for each $f \in C_0(X)$, we can extend the $\mathbb{R}_{\geqslant 0}$-linear functional $\mu \mapsto \int_X f d\mu$ (for all finite Radon measures μ) to $\mu \mapsto \int_X f d\mu$ (for all complex Radon measures μ). This gives a $\mathbb{C}$-bilinear map

$$(f, \mu) \mapsto \int_X f d\mu \qquad \in \mathbb{C}$$

for $f \in C_0(X)$ and complex Radon measures μ.

Theorem 1.7.17 (Riesz-Markov representation theorem). *Let* $\mathbb{F} = \mathbb{C}$ *(resp.* $\mathbb{F} = \mathbb{R}$*.) Then the elements of* $C_c(X, \mathbb{F})^*$ *are precisely linear functionals*

$$\Lambda : C_c(X, \mathbb{F}) \to \mathbb{F} \qquad f \mapsto \int_X f d\mu$$

where μ *is complex (resp. signed) Radon measure on* X.

Proof. It suffices to assume that Λ is real, i.e., sending $C_c(X,\mathbb{R})$ into $\mathbb{R}$. Similar to the proof of Thm. 1.6.16, one writes $\Lambda = \Lambda^+ - \Lambda^-$ where $\Lambda^\pm$ are positive. Then apply Thm. 1.7.6 to $\Lambda^\pm$. See [Gui-A, Subsec. 25.10.2] for details. □

Remark 1.7.18. By Rem. 1.4.32, $C_c(X,\mathbb{F})$ is l^∞-dense in $C_0(X,\mathbb{F})$. Therefore, by Cor. 2.4.3, the dual spaces $C_c(X,\mathbb{F})^*$ and $C_0(X,\mathbb{F})^*$ are canonically identified. Therefore, Thm. 1.7.17 holds verbatim if $C_c(X,\mathbb{F})$ is replaced by $C_0(X,\mathbb{F})$.

1.7.5 Products of Radon measures

Let μ and ν be Radon measures on LCH spaces X and Y, respectively. In this subsection, we recall Tonelli's and Fubini's theorems in the setting of Radon measures. Their more general forms beyond this setting (cf. [Fol-R, Rud-R]) will not be used in this course.

Definition 1.7.19. By the Riesz-Markov Thm. 1.7.6, there is a unique Radon measure $\boldsymbol{\mu\times\nu}$ on $X\times Y$ associated to the linear functional

$$C_c(X\times Y)\to\mathbb{C}\qquad f\mapsto\int_X\int_Y fd\nu d\mu=\int_Y\int_X fd\mu d\nu \tag{1.44}$$

This Radon measure (or rather its completion) is called the **product** of μ and ν.

Remark 1.7.20. By the Stone-Weierstrass Thm. 1.5.13, functions of the form $g(x)h(y)$ (where $g\in C_c(X),h\in C_c(Y)$) span a dense linear subspace of $C_c(X\times Y)$. Therefore, since the relation $\int_X\int_Y fd\nu d\mu=\int_Y\int_X fd\mu d\nu$ holds for all f of this form, it holds for all $f\in C_c(X,Y)$. See [Gui-A, Sec. 26.1] for details.

Theorem 1.7.21 (Tonelli theorem). *Assume that μ and ν are σ-finite. Let $f: X\times Y\to\overline{\mathbb{R}}_{\geqslant 0}$ be $\mu\times\nu$-measurable. Then the following are true.*

(a) For almost every $x\in X$, the function $y\in Y\mapsto f(x,y)$ is ν-measurable.

(b) Set $\int_Y f(x,y)d\nu(y):=0$ if $y\in Y\mapsto f(x,y)$ is not ν-measurable. Then the function $x\in X\mapsto\int_Y f(x,y)d\nu(y)$ is μ-integrable.

(c) We have

$$\int_{X\times Y} fd(\mu\times\nu)=\int_X\int_Y fd\nu d\mu$$

Exchanging the roles of X and Y, we see that $\int_X\int_Y fd\nu d\mu=\int_Y\int_X fd\mu d\nu$.

Proof. See [Gui-A, Sec. 26.2]. □

Theorem 1.7.22 (Fubini theorem). *Assume that μ and ν are σ-finite. Let $f: X\times Y\to\mathbb{C}$ be $\mu\times\nu$-integrable. Then the following are true.*

(a) *For almost every $x \in X$, the function $y \in Y \mapsto f(x,y)$ is ν-integrable.*

(b) *Set $\int_Y f(x,y)d\nu(y) := 0$ if $y \in Y \mapsto f(x,y)$ is not ν-measurable. Then the function $x \in X \mapsto \int_Y f(x,y)d\nu(y)$ is μ-integrable.*

(c) *We have*

$$\int_{X\times Y} fd(\mu\times\nu) = \int_X \int_Y fd\nu d\mu$$

Proof. This is an easy consequence of Tonelli's theorem. See [Gui-A, Sec. 26.2] for details. $\square$

1.8 Basic facts about increasing functions

1.8.1 Notation

If $I \subset \mathbb{R}$ is a proper interval, a function $\rho : I \to \mathbb{R}$ is called **increasing** if it is non-decreasing, i.e., $\rho(x) \leqslant \rho(y)$ whenever $x, y \in I$ and $x \leqslant y$. For each $t \in \mathbb{R}$, let

$$I_{\leqslant t} = I \cap (-\infty, t] \quad I_{<t} = I \cap (-\infty, t) \quad I_{\geqslant t} = I \cap [t, +\infty) \quad I_{>t} = I \cap (t, +\infty)$$

Suppose that $a = \inf I$ and $b = \sup I$. Let $\rho : I \to \mathbb{R}$ be increasing. If $x \in (a,b)$, then the left and right limits[2]

$$\rho(x^-) = \lim_{y\to x^-} \rho(y) \qquad \rho(x^+) = \lim_{y\to x^+} \rho(y) \tag{1.45}$$

exist, and

$$\rho(x^-) \leqslant \rho(x) \leqslant \rho(x^+)$$

If $a \in I$, then $\rho(a^+)$ exists, and $\rho(a) \leqslant \rho(a^+)$. If $b \in I$, then $\rho(b^-)$ exists, and $\rho(b^-) \leqslant \rho(b)$. Let

$$\Omega_\rho = \{x \in (a,b) : \rho|_{(a,b)} \text{ is continuous at } x\}$$

Then for each $x \in (a,b)$, we have

$$x \in \Omega_\rho \quad \Leftrightarrow \quad \rho(x^-) = \rho(x^+) \quad \Leftrightarrow \quad \rho(x^-) = \rho(x) = \rho(x^+) \tag{1.46}$$

[2] When taking the limit $\lim_{y\to x^\pm}$, we do not allow y to be equal to x.

1.8.2 Basic properties of increasing functions

Let $I\subset\mathbb{R}$ be a proper interval with $a=\inf I, b=\sup I$.

Proposition 1.8.1. *If $\rho:I\to\mathbb{R}$ is increasing, then $I\backslash\Omega_\rho$ is countable.*

Proof. Replacing ρ with $\arctan\circ\rho$, we may assume that ρ is bounded. Let $C=\mathrm{diam}(\rho(I))=\sup_{x,y\in I}|\rho(x)-\rho(y)|$. Let $A=(a,b)\backslash\Omega_\rho$. Then for each $B\in\mathrm{fin}(2^A)$, we have

$$\sum_{x\in B}(\rho(x^+)-\rho(x^-))\leqslant C$$

Applying $\lim_B$, we get $\sum_{x\in A}(\rho(x^+)-\rho(x^-))\leqslant C<+\infty$. It follows from Prop. 1.2.44 that A is countable. □

Definition 1.8.2. Let $\rho:I\to\mathbb{R}$ be increasing. The **right-continuous normalization** of ρ is the function $\widetilde{\rho}:I\to\mathbb{R}$ defined by

$$\widetilde{\rho}(x)=\begin{cases}\rho(x^+) & \text{if } x<b\\ \rho(b) & \text{if } x=b\end{cases}$$

The function $\widetilde{\rho}$ is clearly increasing and right-continuous. Moreover, $\widetilde{\rho}$ clearly agrees with ρ on Ω_ρ. Therefore, $\widetilde{\rho}$ and ρ are almost equal, as defined by the following proposition.

Proposition 1.8.3. *Let $\rho_1,\rho_2:I\to\mathbb{R}$ be increasing. Then the following are equivalent:*

(a) There exists a dense subset $E\subset I$ such that $\rho_1|_E=\rho_2|_E$.

(b) $\Omega_{\rho_1}=\Omega_{\rho_2}$, and $\rho_1|_{\Omega_{\rho_1}}=\rho_2|_{\Omega_{\rho_2}}$.

(c) The right-continuous normalizations of ρ_1 and ρ_2 agree on $I_{<b}$.

*If any of these statements are true, we say that ρ_1,ρ_2 are **almost equal**.*

Proof. (a)⇒(b): Assume (a). Choose any $x\in I$. If $x>a$ then

$$\rho_1(x^-)=\lim_{E\ni y\to x^-}\rho_1(y)=\lim_{E\ni y\to x^-}\rho_2(y)=\rho_2(x^-)\tag{1.47a}$$

Similarly, if $x<b$ then

$$\rho_1(x^+)=\rho_2(x^+)\tag{1.47b}$$

Thus (b) follows from (1.46).

(b)⇒(a): By Prop. 1.8.1, $E:=(a,b)\cap\Omega_{\rho_1}$ is a dense subset of (a,b).

(b)⇔(c): Let $\widetilde{\rho}_i$ be the right continuous normalization of ρ_i. Then by (a)⇒(b), we have $\Omega_{\rho_i} = \Omega_{\widetilde{\rho}_i}$ and $\rho_i|_{\Omega_{\rho_i}} = \widetilde{\rho}_i|_{\Omega_{\widetilde{\rho}_i}}$. Therefore, (b) holds iff

$$\Omega_{\widetilde{\rho}_1} = \Omega_{\widetilde{\rho}_2} \qquad \text{and} \qquad \widetilde{\rho}_1|_{\Omega_{\widetilde{\rho}_1}} = \widetilde{\rho}_2|_{\Omega_{\widetilde{\rho}_2}} \tag{1.48}$$

Clearly (c) implies (1.48). Suppose that (1.48) is true. Then for each $x \in I_{<b}$ we have

$$\widetilde{\rho}_1(x) = \widetilde{\rho}_1(x^+) \xlongequal{(1.47b)} \widetilde{\rho}_2(x^+) = \widetilde{\rho}_2(x)$$

Thus (1.48) implies (c). Therefore (b) and (c) are equivalent. □

1.9 The Stieltjes integral

1.9.1 Definitions and basic properties

In this subsection, we fix a proper interval $I \subset \mathbb{R}$, and let $\rho : I \to \mathbb{R}_{\geqslant 0}$ be an increasing function.

Definition 1.9.1. Let J be any proper bounded interval. Let $a = \inf J, b = \sup J$. A **partition** of the interval J is defined to be an element of the form

$$\sigma = \{a_0, a_1, \ldots, a_n \in [a, b] : a_0 = a < a_1 < a_2 < \cdots < a_n = b, n \in \mathbb{Z}_+\} \tag{1.49}$$

The **mesh** of σ is defined to be

$$\max\{a_i - a_{i=1} : i = 1, \ldots, n\}$$

If $\sigma, \sigma' \in \mathrm{fin}(2^J)$ are partitions of J, we say that σ' is a **refinement** of σ (or that σ' is **finer than** σ), if $\sigma \subset \sigma'$. In this case, we also write

$$\sigma \prec \sigma'$$

We define $\mathcal{P}(J)$ to be

$$\mathcal{P}(J) = \{\text{partitions of } J\}$$

Remark 1.9.2. If $\sigma, \sigma' \in \mathcal{P}(J)$, then clearly $\sigma \cup \sigma' \in \mathcal{P}(J)$ and $\sigma, \sigma' \prec \sigma \cup \sigma$. Therefore, $\prec$ is a partial order on $\mathcal{P}(J)$. We call $\sigma \cup \sigma'$ the **common refinement** of σ and σ'.

Definition 1.9.3. A **tagged partition** of I is an ordered pair

$$(\sigma, \xi_\bullet) = \big(\{a_0 = a < a_1 < \cdots < a_n = b\}, (\xi_1, \ldots, \xi_n)\big) \tag{1.50}$$

where $\sigma \in \mathcal{P}(J)$ and

$$\xi_i \in (a_{j-1}, a_j]$$

for all $1 \leqslant j \leqslant n$. The set

$$\mathcal{Q}(J) = \{\text{tagged partitions of } J\}$$

equipped with the preorder $\prec$ defined by

$$(\sigma, \xi_\bullet) \prec (\sigma', \xi'_\bullet) \qquad \Longleftrightarrow \qquad \sigma \subset \sigma' \tag{1.51}$$

is a directed set.

Definition 1.9.4. Let V be a complete normed vector space. Assume $[a,b] \subset I$ and $a < b$. Let $f \in C([a,b],V)$. For each $(\sigma,\xi_\bullet) \in \mathcal{Q}(I)$, define the **Stieltjes sum**

$$S_\rho(f,\sigma,\xi_\bullet) = \sum_{j \geqslant 1} f(\xi_j)\big(\rho(a_j) - \rho(a_{j-1})\big)$$

abbreviated to $S(f,\sigma,\xi_\bullet)$ when no confusion arises. The **Stieltjes integral** on $(a,b]$ is defined to be the limit of the net $(S_\rho(f,\sigma,\xi_\bullet))_{(\sigma,\xi_\bullet)\in\mathcal{Q}([a,b])}$:

$$\int_{(a,b]} f d\rho = \lim_{(\sigma,\xi_\bullet)\in\mathcal{Q}(I)} S_\rho(f,\sigma,\xi_\bullet) \tag{1.52}$$

The **Stieltjes integral** on $[a,b]$ is defined to be

$$\int_{[a,b]} f d\rho = f(a)\rho(a) + \int_{(a,b]} f d\rho \tag{1.53}$$

Note that when $f(a) \neq 0$, the integral $\int_{(a,b]} f d\rho$ depends not only on $\rho|_{(a,b]}$ but also on the value $\rho(a)$. On the other hand, it is clear that

$$\int_{(a,b]} f d\rho = \int_{(a,b]} f d\rho|_{[a,b]} \qquad \int_{[a,b]} f d\rho = \int_{[a,b]} f d\rho|_{[a,b]} \tag{1.54}$$

Proof of the convergence of (1.52). Since f is uniformly continuous, for each $\varepsilon > 0$, there exists $\delta > 0$ such that $\|f(x) - f(y)\| \leqslant \varepsilon$ for all $x,y \in [a,b]$ and $|x-y| \leqslant \delta$. Choose any tagged partition $(\sigma,\xi_\bullet)$ of $[a,b]$ with mesh $\leqslant \delta$. Then one easily sees that for any $(\sigma',\xi'_\bullet) \succ (\sigma,\xi)$ we have

$$\|S_\rho(f,\sigma',\xi'_\bullet) - S_\rho(f,\sigma,\xi_\bullet)\| \leqslant \varepsilon(\rho(b) - \rho(a)) \tag{1.55}$$

Therefore, the net $(S(f,\sigma,\xi_\bullet))_{(\sigma,\xi_\bullet)\in\mathcal{Q}(I)}$ is Cauchy. So it must converge due to Thm. 1.2.15 and the completeness of V. □

Remark 1.9.5. The above proof implies the following useful fact: Let $f \in [a,b]$. Let $\varepsilon, \delta > 0$ such that $\|f(x) - f(y)\| \leqslant \varepsilon$ for all $x,y \in [a,b]$ satisfying $|x-y| \leqslant \delta$. Then for each tagged partition $(\sigma,\xi_\bullet)$ of $[a,b]$ with mesh $\leqslant \delta$, we have

$$\Big\| \int_{(a,b]} f d\rho - S_\rho(f,\sigma,\xi_\bullet) \Big\| \leqslant \varepsilon(\rho(b) - \rho(a)) \tag{1.56}$$

and hence

$$\Big\| \int_{[a,b]} f d\rho - f(a)\rho(a) - S_\rho(f,\sigma,\xi_\bullet) \Big\| \leqslant \varepsilon(\rho(b) - \rho(a)) \tag{1.57}$$

Example 1.9.6. The integrals of the constant function 1 are

$$\int_{(a,b]} d\rho = \rho(b) - \rho(a) \qquad \int_{[a,b]} d\rho = \rho(b)$$

Remark 1.9.7. It is easy to see that

$$\Lambda : C([a,b],V) \to V \qquad f \mapsto \int_{[a,b]} f d\rho$$

is linear. Moreover, since $\|S(f,\sigma,\xi_\bullet)\| \leqslant (\rho(b)-\rho(a))\|f\|_{l^\infty}$ and hence $\|f(a)\rho(a) + S(f,\sigma,\xi_\bullet)\| \leqslant \rho(b)\|f\|_{l^\infty}$, the operator norm $\|\Lambda\|$ satisfies $\|\Lambda\| \leqslant \rho(b)$, that is

$$\Big\|\int_{[a,b]} f d\rho\Big\| \leqslant \rho(b)\|f\|_{l^\infty} \qquad \text{for all } f \in C([a,b],V)$$

In particular, Λ is bounded. Similarly, the linear functional

$$C([a,b],V) \to V \qquad f \mapsto \int_{(a,b]} f d\rho$$

has operator norm $\leqslant \rho(b) - \rho(a)$.

Remark 1.9.8. It is easy to check that $\rho \mapsto \int_{(a,b]} f d\rho$ and $\rho \mapsto \int_{[a,b]} f d\rho$ are $\mathbb{R}_{\geqslant 0}$-linear over increasing functions $\rho : [a,b] \to \mathbb{R}_{\geqslant 0}$. Moreover, if $c \in (a,b)$, one easily shows

$$\int_{(a,b]} f d\rho = \int_{(a,c]} f d\rho + \int_{(c,b]} f d\rho \tag{1.58}$$

by considering tagged partitions finer than $\{a,c,b\}$.

Lemma 1.9.9. *Suppose that $\rho_1, \rho_2 : [a,b] \to \mathbb{R}_{\geqslant 0}$ are increasing and satisfies $\rho_1|_{(a,b]} = \rho_2|_{(a,b]}$. Then for each $f \in C([a,b],V)$ we have $\int_{[a,b]} f d\rho_1 = \int_{[a,b]} f d\rho_2$.*

See Thm. 1.9.11 for a generalization of this lemma.

Proof. Since f is continuous, for each increasing $\rho : [a,b] \to \mathbb{R}_{\geqslant 0}$, the integral $\int_{[a,b]} f d\rho$ can be approximated by $f(a)\rho(a) + S_\rho(f,\sigma,\xi_\bullet)$ where $(\sigma,\xi_\bullet) = $ (1.50) and $\xi_1 = a$. Then

$$f(a)\rho(a) + S_\rho(f,\sigma,\xi_\bullet) = f(a)\rho(a_1) + \sum_{j\geqslant 2} f(\xi_j)(\rho(a_j) - \rho(a_{j-1}))$$

is independent of $\rho(a)$. Hence $\int_{[a,b]} f d\rho$ is independent of $\rho(a)$. $\square$

1.9.2 Dependence of the Stieltjes integral on ρ

Let $I \subset \mathbb{R}$ be a proper interval, and let $a = \inf I$ and $b = \sup I$. Note that I is not assumed to be bounded.

Definition 1.9.10. For each $f \in C_c(I, V)$ and each increasing $\rho : I \to \mathbb{R}_{\geqslant 0}$, we can still define the **Stieltjes integral**

$$\int_I f d\rho := \int_J f d\rho$$

where J is any compact sub-interval of I containing $\mathrm{Supp}_I(f)$. The value of the integral is clearly independent of the choice of such J. Moreover, this definition is compatible with the definitions of $\int_{[a,b]} f d\rho$ and $\int_{(a,b]} f d\rho$ in Def. 1.9.4.

Theorem 1.9.11. *Let $\rho_1, \rho_2 : I \to \mathbb{R}_{\geqslant 0}$ be increasing functions satisfying the following condition:*

- *ρ_1 and ρ_2 are almost equal, and $\rho_1(b) = \rho_2(b)$ if $b \in I$. (By Prop. 1.8.3, this is equivalent to that ρ_1, ρ_2 have the same right-continuous normalization.)*

Then for each $f \in C_c(I, V)$, we have

$$\int_I f d\rho_1 = \int_I f d\rho_2$$

Proof. By Lem. 1.9.9, we may assume that $\rho_1(a) = \rho_2(a)$ if $a \in I$.

Fix $f \in C_c(I, V)$. Choose $\alpha, \beta \in \mathbb{R}$ satisfying $\mathrm{Supp}_I(f) \subset [\alpha, \beta] \subset I$. Due to the assumption on ρ_1, ρ_2, we may slightly enlarge the compact interval $J := [\alpha, \beta]$ so that

$$\rho_1(\alpha) = \rho_2(\alpha) \qquad \rho_1(\beta) = \rho_2(\beta)$$

(When $a \in I$ resp. $b \in I$, one can even set $\alpha = a$ resp. $\beta = b$.) Then $\int_I f d\rho_i = \int_J f d\rho_i$.

Let $C = \max\{\rho_i(\beta) - \rho_i(\alpha) : i = 1, 2\}$. Choose any $\varepsilon > 0$. Since f is uniformly continuous, there exists $\delta > 0$ such that $|f(x) - f(y)| \leqslant \varepsilon$ whenever $x, y \in I$ and $|x - y| \leqslant \delta$. Choose a tagged partition

$$(\sigma, \xi_\bullet) = \big(\{a_0 = \alpha < a_1 < \cdots < a_n = \beta\}, (\xi_1, \ldots, \xi_n)\big)$$

of J with mesh $< \delta$. Moreover, due to the assumption on ρ_1, ρ_2, by a slight adjustment, we may assume that $\rho_1(a_j) = \rho_2(a_j)$ for each $0 \leqslant j \leqslant n$. This implies

$$S_{\rho_1}(f, \sigma, \xi_\bullet) = S_{\rho_2}(f, \sigma, \xi_\bullet)$$

Therefore, by Rem. 1.9.5, we obtain

$$\Big\| \int_J f d\rho_1 - \int_J f d\rho_2 \Big\| \leqslant 2\varepsilon \cdot C$$

This completes the proof by choosing arbitrary ε. □

Theorem 1.9.12. *Let $\rho_1, \rho_2 : I \to \mathbb{R}_{\geqslant 0}$ be increasing functions satisfying*

$$\lim_{x\to a^+} \rho_1(x) = \lim_{x\to a^+} \rho_2(x) = 0 \qquad \text{if } a \notin I \tag{1.59}$$

Then the following are equivalent:

(1) *ρ_1 and ρ_2 are almost equal, and $\rho_1(b) = \rho_2(b)$ if $b \in I$. (By Prop. 1.8.3, this is equivalent to that ρ_1, ρ_2 have the same right-continuous normalization.)*

(2) *For each $f \in C_c(I, \mathbb{F})$ we have*

$$\int_I f d\rho_1 = \int_I f d\rho_2$$

Proof. By Thm. 1.9.11, we have "(1)⇒(2)". Assume (2). Let us prove (1). Let $\widetilde{\rho}_i$ be the right-normalization of ρ_i. By "(1)⇒(2)", we have $\int_I f d\rho_i = \int_I f d\widetilde{\rho}_i$. Therefore, to prove (1), it suffices to assume that ρ_1 and ρ_2 are right-continuous on I.

We shall prove (1) by choosing an arbitrary bounded increasing right-continuous $\rho : I \to \mathbb{R}_{\geqslant 0}$, and show that for each $x \in I$, the value $\rho(x)$ can be recovered from the integrals $\int_I f d\rho$ where $f \in C_c(I, \mathbb{R})$.

Case 1: Assume $a \notin I$ and $a < x < b$. For each real numbers v, y satisfying

$$a < v < x < y < b$$

choose $\varphi_{v,y} \in C_c(I, [0, 1])$ satisfying

$$\chi_{[v,x]} \leqslant \varphi_{v,y} \leqslant \chi_{(a,y]}$$

Choose $u \in (a, v)$ such that $\varphi_{v,y}$ vanishes outside $[u, y]$. Then by Rem. 1.9.8,

$$\int_I \varphi_{v,y} d\rho = \int_{[u,y]} \varphi_{v,y} d\rho = \int_{[u,v]} \varphi_{v,y} d\rho + \int_{(v,x]} \varphi_{v,y} d\rho + \int_{(x,y]} \varphi_{v,y} d\rho$$
$$= \int_{[u,v]} \varphi_{v,y} d\rho + \rho(x) - \rho(v) + \int_{(x,y]} \varphi_{v,y} d\rho$$

where Exp. 1.9.6 is used in the last equality. By Rem. 1.9.7, we have $\int_{[u,v]} \varphi_{v,y} d\rho \leqslant \rho(v)$ and $\int_{(x,y]} \varphi_{v,y} d\rho \leqslant \rho(y) - \rho(x)$. Since ρ is right-continuous and satisfies (1.59), we have

$$\lim_{v \searrow a^+} \rho(v) = \lim_{y \searrow x^+} (\rho(y) - \rho(x)) = 0$$

Therefore, the above calculation of $\int_I \varphi_{v,y} d\rho$ shows

$$\lim_{\substack{v \searrow a^+ \\ y \searrow x^+}} \int_I \varphi_{v,y} d\rho = \lim_{v \searrow a^+} (\rho(x) - \rho(v)) = \rho(x)$$

Case 2: Assume $a \in I$ and $a \leqslant x < b$. For each $y \in (x, b)$, choose $\varphi_y \in C_c(I, [0,1])$ such that $\chi_{[a,x]} \leqslant \varphi_y \leqslant \chi_{[a,y]}$. Similar to the argument in Case 1, one shows

$$\int_I \varphi_y d\rho = \int_{[a,x]} \varphi_y d\rho + \int_{(x,y]} \varphi_y d\rho = \rho(x) + \int_{(x,y]} \varphi_y d\rho$$

where Exp. 1.9.6 is used. By Rem. 1.9.7, $\int_{(x,y]} \varphi_y d\rho \leqslant \rho(y) - \rho(x)$. Therefore, the right-continuity of ρ implies

$$\lim_{y \searrow x^+} \int_I \varphi_y d\rho = \rho(x)$$

Case 3: Assume $I = (a, b]$ and $x = b$. For each $v \in (a, x)$, choose $\varphi_v \in C_c(I, [0,1])$ such that $\chi_{[v,b]} \leqslant \varphi_v \leqslant \chi_I$. Similar to the argument above,

$$\lim_{v \searrow a^+} \int_I \varphi_v d\rho = \rho(b)$$

Case 4: Assume $I = [a, b]$ and $x = b$. Then $\displaystyle\int_I d\rho = \rho(b)$. $\square$

Remark 1.9.13. The assumption (1.59) imposes little restriction. Indeed, suppose $a \notin I$. Then for each $f \in C_c(I)$, since there exists $v \in \mathbb{R}_{>a}$ such that f vanishes on $(a, v]$, for any constant $\varkappa \in \mathbb{R}$ with $\rho + \varkappa \geqslant 0$, we clearly have

$$\int_I f d\rho = \int_I f d(\rho + \varkappa) \tag{1.60}$$

Therefore, when $a \notin I$, given any two increasing functions $\rho_1, \rho_2 : I \to \mathbb{R}_{\geqslant 0}$, we can freely add constants to ρ_1 and ρ_2 to ensure that (1.59) holds.

1.10 The Riesz representation theorem via the Stieltjes integral

In this section, we fix a proper interval $I \subset \mathbb{R}$, and let $a = \inf I$ and $b = \sup I$.

1.10.1 The positive case

Theorem 1.10.1 (Riesz representation theorem). *We have a bijection between:*

(a) A bounded increasing right-continuous function $\rho : I \to \mathbb{R}_{\geqslant 0}$ satisfying $\lim_{x \to a^+} \rho(a) = 0$ if $a \notin I$.

(b) A bounded positive linear functional $\Lambda : C_c(I, \mathbb{F}) \to \mathbb{F}$.

Λ is determined by ρ by

$$\Lambda : C(I, \mathbb{F}) \to \mathbb{F} \qquad f \mapsto \int_I f d\rho \tag{1.61}$$

ρ is determined by Λ by

$$\rho(x) = \mu(I_{\leqslant x}) \qquad \textit{for all } x \in I \tag{1.62}$$

where μ is the finite Borel measure on I associated to Λ as in the Riesz-Markov representation Thm. 1.7.6.

Note that by Thm. 1.7.8, finite Borel measures on I and finite Radon measures on I coincide.

Proof. Step 1. Thm. 1.7.6 establishes the equivalence between a bounded positive linear functional Λ and a finite Borel measure μ. Let us prove the equivalence between the radon measures μ and the functions ρ satisfying (a).

More precisely, given a Radon measure μ on I, let $\rho_\mu : I \to \mathbb{R}_{\geqslant 0}$ be defined by (1.62), that is, for each $x \in I$ we have

$$\rho_\mu(x) = \mu(I_{\leqslant x}) \tag{1.63}$$

Then ρ_μ is clearly bounded and increasing. By DCT, ρ_μ is right-continuous, and we have $\lim_{x \to a^-} \rho(x) = 0$ when $a \notin I$. Therefore, ρ_μ satisfies (a).

Conversely, given any ρ satisfying (a), let μ_ρ be the unique Radon measure corresponding to ρ via (1.61), i.e., for each $f \in C_c(I, \mathbb{F})$ we have

$$\int_I f d\mu_\rho = \int_I f d\rho \tag{1.64}$$

By Rem. 1.9.7, the linear functional $f \in C_c(I, \mathbb{F}) \mapsto \int_I f d\rho$ is bounded with operator norm $\leqslant \sup_{x \in I} \rho(x)$. Thus, μ_ρ is a finite measure.

We want to show that $\Phi : \rho \mapsto \mu_\rho$ and $\Psi : \mu \mapsto \rho_\mu$ are inverses of each other. By Thm. 1.9.12, the map Φ is injective. Therefore, it suffices to prove that $\Phi \circ \Psi = \mathrm{id}$, i.e., that $\mu_{\rho_\mu} = \mu$. This means proving

$$\int_I f d\mu = \int_I f d\rho_\mu \tag{1.65}$$

for each $f \in C_c(I, \mathbb{F})$.

Step 2. Let us fix $f \in C_c(I, \mathbb{F})$ and prove (1.65). Choose $\alpha, \beta \in \mathbb{R}$ such that $J := [\alpha, \beta]$ is a sub-interval of I containing $\mathrm{Supp}_I(f)$. Choose any $\varepsilon > 0$. Since f is uniformly continuous, there exists $\delta > 0$ such that $|f(x) - f(y)| \leqslant \varepsilon$ whenever $x, y \in I$ and $|x - y| \leqslant \delta$. Choose a tagged partition

$$(\sigma, \xi_\bullet) = \big(\{a_0 = \alpha < a_1 < \cdots < a_n = \beta\}, (\xi_1, \ldots, \xi_n)\big)$$

of J with mesh $\leqslant \delta$. By Rem. 1.9.5, we have

$$\Big|\int_J fd\rho_\mu - f(\alpha)\rho_\mu(\alpha) - S_{\rho_\mu}(f,\sigma,\xi_\bullet)\Big| \leqslant \varepsilon(\rho_\mu(\beta)-\rho_\mu(\alpha)) = \varepsilon\cdot\mu((\alpha,\beta]) \tag{1.66}$$

Also, we have $\|f-g\|_{l^\infty(I)}\leqslant\varepsilon$ where

$$g = f(\alpha)\chi_{\{a\}} + \sum_{i=1}^n f(\xi_i)\chi_{(a_{i-1},a_i]}$$

By (1.63), we have

$$\mu(\{\alpha\}) = \rho_\mu(\alpha) - \mu(I_{<\alpha}) \qquad \mu((a_{i-1},a_i]) = \rho_\mu(a_i) - \rho_\mu(a_{i-1})$$

Note that if $f(\alpha)\neq 0$, then by $\mathrm{Supp}_I(f)\subset[\alpha,\beta]$, we must have $\alpha = a\in I$ and hence $I_{<\alpha}=\varnothing$. Therefore, we must have

$$\int_I gd\mu = f(\alpha)\rho_\mu(\alpha) + S_{\rho_\mu}(f,\sigma,\xi_\bullet)$$

Combining this fact with $\|f-g\|_{l^\infty(I)}\leqslant\varepsilon$, we get

$$\Big|\int_I fd\mu - f(\alpha)\rho_\mu(\alpha) - S_{\rho_\mu}(f,\mu,\xi_\bullet)\Big| \leqslant \varepsilon\cdot\mu(J)$$

This inequality, together with (1.66), implies

$$\Big|\int_I fd\mu - \int_I fd\rho_\mu\Big| \leqslant 2\varepsilon\cdot\mu(J)$$

Since ε is arbitrary, we conclude (1.65). $\square$

1.10.2 The general case

Definition 1.10.2. A real-valued function $I\to\mathbb{F}$ is called of **bounded variation** (or simply **BV**) if it is an $\mathbb{F}$-linear combination of bounded increasing functions $I\to\mathbb{R}_{\geqslant 0}$. The space of BV functions from I to $\mathbb{F}$ is denoted by $BV(I,\mathbb{F})$.

Remark 1.10.3. By Rem. 1.9.7 and 1.9.8, we have an $\mathbb{R}_{\geqslant 0}$-bilinear functional

$$(f,\rho)\mapsto\int_I fd\rho \qquad \in\mathbb{R}_{\geqslant 0}$$

for $f\in C_c(I,\mathbb{R}_{\geqslant 0})$ and bounded increasing $\rho: I\to\mathbb{R}_{\geqslant 0}$. Similar to the proof of Rem. 1.7.4, it can be extended to a positive $\mathbb{C}$-bililinear functional

$$C_c(I,\mathbb{C})\times BV(I,\mathbb{C})\to\mathbb{C} \qquad (f,\rho)\mapsto\int_I fd\rho$$

Theorem 1.10.4 (Riesz representation theorem). *The elements of the dual space* $C_c(I,\mathbb{F})^*$ *are precisely linear functionals of the form*

$$\Lambda : C(I,\mathbb{F}) \to \mathbb{F} \qquad f \mapsto \int_I f d\rho$$

where $\rho \in BV(I,\mathbb{F})$*. Moreover, the BV function* ρ *can be chosen such that it is right-continuous on* I*, and that* $\lim_{x\to a^+} \rho(x) = 0$ *if* $a \notin I$*.*

Proof. This is immediate from Thm. 1.7.17 and 1.10.1. $\square$

2 Normed vector spaces and their dual spaces

2.1 The origin of dual spaces in the calculus of variations

Linear functional analysis treats function spaces as linear spaces with appropriate geometric/topological structures and analytic properties. In the foundational theory of functional analysis, two analytic properties are especially important: (Cauchy) completeness and duality. In this course, our focus is primarily on normed vector spaces V. For such spaces, Cauchy completeness is interpreted in the same way as in any metric space. Duality, on the other hand, refers to the natural identification of V as the dual space U^* of another normed vector space U.

Many early results in functional analysis were related to duality, while the significance of completeness was not immediately recognized. In fact, the history of functional analysis experienced a paradigm shift from the study of (scalar-valued) functionals to linear maps between vector spaces. Specifically, attention moved from continuous bilinear maps of the form $U \times V \to \mathbb{F}$ to the analysis of continuous linear maps $V \to W$, where U, V, W are normed vector spaces. With this shift, completeness became increasingly central to modern analysis. See Sec. 2.5 for further illustrations.

The early part of this course will also focus more on dual spaces. If V is a normed $\mathbb{F}$-vector space, then the **dual space** $V^* = \mathfrak{L}(V, \mathbb{F})$ is defined to be the space of bounded (i.e. continuous) linear maps $V \to \mathbb{F}$. One of the major themes in early functional analysis was the characterization of dual spaces of various function spaces under appropriate norms. Among the most notable results are F. Riesz's characterization of $C([a,b],\mathbb{R})^*$ (cf. Thm. 1.10.4) in [Rie09, Rie11], and his proof that $L^q([a,b],m,\mathbb{R})^* \simeq L^p([a,b],m,\mathbb{R})$ (cf. Thm. 1.6.16) in [Rie10]. These results highlight a profound connection between dual spaces and measure/integration theory. Nevertheless, the study of dual spaces originally arose from a somewhat different field: the calculus of variations in the 19th century.

Consider a nonlinear functional $S : f \mapsto S(f) \in \mathbb{R}$, for example, of the form

$$S(f) = \int_a^b L(f(t), f'(t), \ldots, f^{(r)}(t))dt$$

where L is a "nice" real valued function with r-variables, and f is defined on $[a,b]$. If we perturb f slightly by a variation η, then the corresponding change in S can be approximated by

$$\delta S[f,\eta] := S(f+\eta) - S(f) \approx \int_a^b \beta_f(t) \cdot \eta(t)dt \tag{2.1}$$

where $\beta_f : [a,b] \to \mathbb{R}$ is a function depending on f. This function should be interpreted loosely. In some cases, it may involve delta functions or similar objects that are not functions in the classical sense, but rather distributions:

Example 2.1.1. Consider the case where L is smooth and $r = 1$, i.e.

$$S(f) = \int_a^b L(f(t), f'(t))dt$$

(For example, $L(x, y) = T(y) - V(x)$ where $T(y) = \frac{1}{2}my^2$ the kinetic energy for the mass $m \in \mathbb{R}_{>0}$, and $V(x)$ is the potential energy at x.) Then

$$\begin{aligned}\delta S[f, \eta] = &\int_a^b L(f + \eta, f' + \eta') \approx \int_a^b (\partial_x L(f, f')\eta + \partial_y L(f, \eta)\eta')\\ =&\partial_y L(f, f')\eta\big|_a^b + \int_a^b (\partial_x L(f, f') - \partial_t \partial_y L(f, f'))\eta\end{aligned}$$

If we assume that the function f and its variation η always vanish at the endpoints a, b, then we obtain (2.1) with

$$\beta_f(t) = \partial_x L(f(t), f'(t)) - \partial_t \partial_y L(f(t), f'(t))$$

The equation $\beta_f = 0$ is called the **Euler-Lagrange equation**.

However, if no boundary conditions are imposed on the endpoints, then the term $\partial_y L(f, f')\eta\big|_a^b$ is not necessarily zero. As a result, we have

$$\beta_f = L(f(b), f'(b))\delta_b - L(f(a), f'(a))\delta_a + \partial_x L(f, f') - \partial_t \partial_y (f, f')$$

where, for each $c \in \mathbb{R}$, δ_c is the "**delta function**" at c, namely, the imaginary function $\mathbb{R} \to \mathbb{R}_{\geqslant 0}$ vanishing outside c and satisfying $\int_{\mathbb{R}} \delta_c = 1$. The situation becomes even more singular if we define S by

$$S(f) = \sum_{i=1}^n \lambda_i f(c_i) + \int_a^b L(f(t), f'(t))dt$$

where $\lambda_i \in \mathbb{R}$ and $a < c_i < b$, then

$$\beta_f = \sum_{i=1}^n \lambda_i \delta_{c_i} + L(f(b), f'(b))\delta_b - L(f(a), f'(a))\delta_a + \partial_x L(f, f') - \partial_t \partial_y (f, f')$$

This raises the question: what should the function β_f, alternatively the integral operator $\eta \mapsto \int_a^b \beta_f \eta$, actually look like in the general case? □

It is in this context that the problem of classifying bounded linear functionals on $C([a, b], \mathbb{R})$, originally posed by Hadamard in 1903, should be understood. Recall that if V, W are normed vector spaces, $\Omega \subset V$ is open, and $S : \Omega \to W$ is a map, one says that S is differentiable at $f \in \Omega$ if

$$S(f + \eta) - S(f) = \Lambda(\eta) + o(\eta)$$

where $\Lambda : V \to W$ is a bounded linear operator (called the **differential** of S at f), and $\lim_{\|\eta\|\to 0} o(\eta)/\|\eta\| = 0$. In the calculus of variations, one sets $W = \mathbb{F}$. Then $\Lambda \in V^*$. One can thus understand $\eta \mapsto \delta S[f,\eta]$ as a bounded linear functional on a function space V equipped with a suitable norm.

The problem of expressing $\delta S[f,\eta]$ as an integral involving η is therefore transformed to the problem of characterizing the dual space V^*. More precisely, the space V—and in particular its norm—is not fixed in advance. The situation is not that one starts with a given normed space and is then asked to characterize its dual. Rather, the task is to find an appropriate norm on a suitable function space V such that the bounded linear functionals on V, once studied and classified as integrals, are well-suited to capturing the variation of S. [1] The two perspectives on $\delta S[f,\eta]$—as a bounded linear functional on V, and as an integral involving η—together offer a deeper and more complete understanding of the variation of S.

More discussion of the relationship between dual spaces and the calculus of variations can be found in [Gray84].

2.2 Moment problems: a bridge between integral theory and dual spaces

The theory of dual spaces would not have reached its current depth and sophistication if it were developed solely within the framework of the calculus of variations. For instance, Riesz's classification of the duals of $C([a,b])$ and $L^p([a,b],m)$ would have been impossible without the Lebesgue and Stieltjes integrals. In fact, the very form of Riesz's theorems presents a striking connection between integration theory and dual spaces.

But why should such a connection exist in the first place? The way this relationship appears in Riesz's theorems calls for a deeper explanation. My short answer is this: it is the moment problems that form the bridge between integration theory and the theory of dual spaces. (Readers may jump ahead to Subsection 2.2.5 for the detailed final conclusion.)

To clarify my point, consider the first major example of a duality theorem: the identification $(L^2)^* \simeq L^2$ proved by Riesz and Fréchet in 1907:

Theorem 2.2.1 (Riesz-Fréchet theorem, the classical form**).** *We have a linear isomorphism*

$$\Lambda : L^2\big([-\pi,\pi],\frac{m}{2\pi}\big) \to L^2\big([-\pi,\pi],\frac{m}{2\pi}\big)^*$$

[1] The same function space V, when equipped with different norms, leads to different classifications of bounded linear functionals. For example, let $V = C([a,b])$. If the norm is l^∞, then by Thm. 1.10.4, the bounded linear functionals are the Stieltjes integrals with respect to BV functions. If the norm is L^2, then by Exp. 2.4.4, the bounded linear functionals are those of the form $f \mapsto \int fgdm$ where $g \in L^2([a,b],m)$.

$$\langle \Lambda(f), g \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} fg dm$$

In fact, Riesz studied L^2 spaces several years before introducing the more general L^p spaces. His interest in L^2 spaces was clearly influenced by Hilbert's earlier work on the Hilbert space $l^2(\mathbb{Z})$ and its applications to the theory of integral equations. It was Hilbert's insights that served as the crucial bridge leading to the Riesz-Fréchet Thm. 2.2.1—the first major result linking Lebesgue integration with dual spaces.

As I will explain in the following, Hilbert's role in this development is best understood through the lens of moment problems.

2.2.1 Moment problems and dual spaces

Let me begin by introducing moment problems and explaining how they relate to dual spaces—particularly to the characterization of dual spaces in terms of integral representations.

Problem 2.2.2 (**Moment problem**, original version)**.** Let (ξ_n) be a sequence of scalar-valued functions defined on a space, e.g., an interval $I \subset \mathbb{R}$. Choose a sequence of scalars (c_n) satisfying certain conditions. Find a scalar valued function f on I such that for all n, we have

$$\int \xi_n f = c_n \qquad \text{resp.} \qquad \int \xi_n df = c_n \tag{2.2}$$

The numbers $c_1, c_2, \dots$ are called the **moments** of f resp. df.

There are two typical types of moment problems:

- **Trigonometric moment problem**: Here $I = \mathbb{S}^1 \simeq \mathbb{R}/2\pi\mathbb{Z}$, and $\xi_n(x) = e^{-\mathbf{i}nx}$ for $n \in \mathbb{Z}$. The problem then amounts to finding a function f with prescribed Fourier coefficients $c_1, c_2, \dots$.

- **Polynomial moment problem**: Here $I \subset \mathbb{R}$ is an interval, not necessarily bounded, and $\xi_n(x) = x^n$ for $n \in \mathbb{N}$. One is asked to find an increasing or BV function f such that df has moments $c_1, c_2, \dots$.

Many (but not all) moment problems can be reformulated in the language of bounded linear functionals and dual spaces as follows:

Problem 2.2.3 (**Moment problem**, dual space version)**.** Let (ξ_n) be a sequence in a normed vector space V, and let (c_n) be a sequence of scalars. Suppose that there exists $M \in \mathbb{R}_{\geqslant 0}$ such that

$$\Big| \sum_n a_n c_n \Big| \leqslant M \Big\| \sum_n a_n \xi_n \Big\| \tag{2.3}$$

for each sequence of scalars (a_n) with finitely many nonzero terms. Find $\varphi \in V^*$ such that

$$\langle \xi_n, \varphi \rangle = c_n \qquad \text{for all } n \tag{2.4}$$

Remark 2.2.4. Note that (2.3) is necessary for the existence of φ satisfying (2.4), because

$$\Big|\sum_n a_n c_n\Big| = \Big|\Big\langle \sum_n a_n \xi_n, \varphi \Big\rangle\Big| \leqslant \|\varphi\| \cdot \Big\|\sum_n a_n \xi_n\Big\|$$

where $\|\varphi\|$ is the operator norm. Hence (2.3) holds for any M satisfying $\|\varphi\| \leqslant M$.

Conversely, if we know that $V_0 = \mathrm{Span}\{\xi_n\}$ is dense in V, then (2.3) guarantees that the linear functional

$$\varphi : V_0 \to \mathbb{F} \qquad \sum_n a_n \xi_n \mapsto \sum_n a_n c_n$$

is well-defined and bounded, with operator norm $\|\varphi\| \leqslant M$. By boundedness, φ extends uniquely to a bounded linear functional on all of V, cf. Thm. 2.4.2. Therefore, Problem 2.2.3 can always be solved.

The case where V_0 is not dense in V is more subtle and will be treated in detail in Ch. 8. $\square$

Once Problem 2.2.3 is resolved—for example, when $\mathrm{Span}\{\xi_n\}$ is dense in V—Problem 2.2.2 can be solved by answering the following:

Problem 2.2.5 (Characterization of the dual space). Characterize the elements of V^* as precisely those linear functionals $\varphi : V \to \mathbb{F}$ of the form

$$\langle \xi, \varphi \rangle = \int \xi f \;\text{ resp. }\; \int \xi df$$

(for all $\xi \in V$), where f is a function satisfying suitable regularity or integrability conditions.

Conversely, Problem 2.2.5 can be reduced to the moment Problem 2.2.2 by choosing a densely-spanning (ξ_n) and taking $c_n = \langle \xi_n, \varphi \rangle$. The solution to Problem 2.2.2 then yields a function f such that $\langle \xi_n, f \rangle = \langle \xi_n, \varphi \rangle$. By the density of $\mathrm{Span}\{\xi_n\}$ in V, it follows that φ is represented by f.

Thus, we conclude that when (ξ_n) spans a dense subspace of V, the moment problem (Problem 2.2.2) and the characterization of dual spaces (Problem 2.2.5) are equivalent.

2.2.2 Moment problems and integral theory/function theory

In the remainder of this section, we focus on the case where the sequence of functions (ξ_n) is "sufficiently rich", for example, when it spans a dense subspace of V in Problem 2.2.3. Under this assumption, the function f (resp. df) in Problem 2.2.2 or the functional φ in Problem 2.2.3 is uniquely determined by the moments (c_n). Therefore, (c_n) can be understood as the **coordinates** of f (resp. df) and φ under the **coordinate system** (ξ_n).

We now explain how the moment problems connect to integral theory—in other words, to **function theory**. A central theme in function theory is the approximation of abstract or complicated functions by simpler, more elementary ones. This motivation often arises from practical mathematical problems, particularly those originating in physics, where one seeks to express the solution as a series of elementary functions, such as a power series or a Fourier series. The question of how such series should converge—uniformly, pointwise, or in some other sense—and what kinds of functions they can approximate was a central focus of function theory in the 18th and 19th centuries.

The first step in understanding and solving the approximation problem is to analyze the corresponding moment problem. A typical scenario unfolds as follows. In the setting of Problem 2.2.2, suppose there exists a sequence of elementary functions (f_n) such that

$$\int \xi_k f_n \quad \text{resp.} \quad \int \xi_k df_n \quad = c_k \qquad \text{when } |k| \leqslant |n| \tag{2.5}$$

This situation arises, for instance, in the study of continued fractions and polynomial moments, where $\xi_k(x) = x^k$. In the case of Fourier series, an even stronger condition holds:

$$\int \xi_k f_n \quad \text{resp.} \quad \int \xi_k df_n \quad = \begin{cases} c_k & \text{if } |k| \leqslant |n| \\ 0 & \text{if } |k| > |n| \end{cases} \tag{2.6}$$

where $\xi_k(x) = e^{-\mathbf{i}kx}$ and $f_n(x) = \sum_{|k|\leqslant n} c_k e^{\mathbf{i}kx}$. The approximation problem asks:

Problem 2.2.6. Does the sequence (f_n) converge to some function f? If so, in what sense does it converge?

To approach this problem, observe that if such a function f exists, and if the integral commutes with the convergence of sequence of functions, then

$$\int \xi_k f = \int \lim_{|n|\to\infty} \xi_k f_n = \lim_{|n|\to\infty} \int \xi_k f_n \overset{(2.5)}{=\!=} c_k \tag{2.7a}$$

resp.

$$\int \xi_k df = \int \lim_{|n|\to\infty} \xi_k df_n = \lim_{|n|\to\infty} \int \xi_k df_n \overset{(2.5)}{=\!=} c_k \tag{2.7b}$$

Therefore, the first step in solving Problem 2.2.6 is to find a function f solving the moment Problem 2.2.2. Once such an f is found, the next step is to prove that the sequence (f_n) converges to f, and investigate the mode of convergence.

Historically, the understanding of convergence, the properties of the limiting function f, and the integrals appearing in (2.7) was often insufficient to resolve the approximation problem at the outset. In many cases, addressing the approximation problem required the development of new theories of integration or the extension of the class of integrable functions. Both the Lebesgue and Stieltjes integrals emerged from such needs. For instance, the challenges posed by Fourier series played a central role in motivating the development of the Riemann and later the Lebesgue integral. See [Jah, Ch. 6, 9] and [Haw-L] for a detailed discussion of how Fourier series drove this evolution. The connection between continued fractions and the Stieltjes integral will be explored in Ch. 4.

Function theory	Moment Problems	Dual spaces
Lebesgue integral & Fourier series	Fourier coefficients	$L^2([a,b],m)^*$
Stieltjes integral & Continued fractions	Polynomial moments	$C([a,b])^*$

Table 2.1: The origin of moment problems in function theory

2.2.3 Convergence of functions, moments, and linear functionals

In the previous subsection, we noted that solving moment problems determines the function f that appears in Problem 2.2.6. But can the moment problem perspective also help us understand the convergence of f_n to f? Or conversely, can the convergence behavior of f_n toward f offer deeper insight into the structure of moment problems themselves? Thanks to Hilbert's foundational work on the Hilbert space $l^2(\mathbb{Z})$—especially his groundbreaking 1906 paper [Hil06]—the answer is yes.[2]

A key concept introduced by Hilbert in [Hil06] is **weak convergence**: If (ψ_n) is a sequence in $l^2(\mathbb{Z})$ with uniformly bounded norm, i.e.,

$$\sup_n \|\psi_n\|_2 < +\infty \tag{2.8}$$

we say that (ψ_n) converges weakly to $\psi \in l^2(\mathbb{Z})$ if it converges pointwise $\mathbb{Z}$, i.e.,

$$\lim_n \psi_n(k) = \psi(k) \qquad \text{for all } k \in \mathbb{Z} \tag{2.9}$$

[2]Indeed, Hilbert originally worked with the real Hilbert space $l^2(\mathbb{Z},\mathbb{R})$, rather than the complex one $l^2(\mathbb{Z}) = l^2(\mathbb{Z},\mathbb{C})$. For clarity and simplicity, however, we will work with $l^2(\mathbb{Z})$ in what follows.

Since $l^2(\mathbb{Z})$ is typically interpreted as the space of Fourier series of L^2-integrable functions, Hilbert's notion of weak convergence corresponds to the (pointwise) convergence of Fourier coefficients. That is,

$$\lim_n \widehat{f}_n(k) = \widehat{f}(k) \qquad \text{for all } k \in \mathbb{Z}$$

where f_n and f are L^2-integrable functions on $[-\pi, \pi]$.[3]

The notion of weak convergence—later extended to weak-* convergence—provided a fundamentally new insight into the study of moment problems and their connection to dual spaces and function theory/integral theory. Since Fourier coefficients are simply trigonometric moments, the weak convergence described by (2.9) can be understood as the (pointwise) **convergence of moments**, which means, in the setting of the moment Problem 2.2.2, that

$$\lim_n \int \xi_k f_n \quad \text{resp.} \quad \lim_n \int \xi_k df_n \quad = c_n \qquad \text{for all } k \tag{2.10}$$

The translation of (2.10) into the setting of the dual space version of the moment Problem 2.2.3 is straightforward: One considers a sequence (φ_k) in V^* such that $\lim_n \langle \xi_k, \varphi_n \rangle = \langle \xi_k, \varphi \rangle$ holds for all k. Since we have assumed at the beginning of Subsec. 2.2.2 that (ξ_n) spans a dense subspace of V, it follows from (2.8) that this convergence of moments is equivalent to the **weak-* convergence** of (φ_n) to φ. That is, we say that (φ_n) converges weak-* to φ if

$$\lim_n \langle \xi, \varphi_n \rangle = \langle \xi, \varphi \rangle \qquad \text{for all } \xi \in V \tag{2.11}$$

Thus, the second and third columns of Table 2.2 are equivalent. See Thm. 2.6.2 for the formal statement of this equivalence.

On the other hand, (2.10) generalizes the condition (2.5), which, as previously mentioned, arises naturally in the study of Fourier series and continued fractions. As such, its function-theoretic interpretation—highlighted by the following theorems—provides a general framework for understanding the convergence of the sequence (f_n) to f in Problem 2.2.6.

Theorem 2.2.7. *Let $1 < p \leqslant +\infty$ and $p^{-1} + q^{-1} = 1$. Let (f_n) be a uniformly L^p-norm bounded sequence in $L^p([a,b], m)$. Suppose that (f_n) converges pointwise to f. Then we have $f \in L^p([a,b], m)$. Moreover, (f_n) converges weak-* to f, which means that $\lim_n \int f_n g dm = \int f g dm$ for all $g \in L^q([a,b], m)$.*

Proof. See Thm. 2.7.2. □

[3]Hilbert himself did not initially connect $l^2(\mathbb{Z})$ with the Lebesgue integral. The precise relationship between $l^2(\mathbb{Z})$ and $L^2([-\pi, \pi], \frac{m}{2\pi})$ was later clarified by Riesz and Fischer in 1907.

Theorem 2.2.8. *Let $1 < p \leqslant +\infty$ and $p^{-1} + q^{-1} = 1$. Let (f_n) be a uniformly L^p-norm bounded sequence in $L^p([a,b],m)$. Then (f_n) converges weak-* to some element of $L^p([a,b],m)$ iff the limit*

$$F(x) := \lim_n \int_a^x f_n dm \tag{2.12}$$

exists for every $x \in [a,b]$. Moreover, if $f \in L^p([a,b],m)$, then (f_n) converges weak- to $f \in L^p([a,b],m)$ iff for each $x \in [a,b]$ we have*

$$F(x) = \int_a^x f dm \tag{2.13}$$

Proof. If (f_n) converges weak-* to f, then $\lim_n \int f_n \chi_{[a,x]} = \int f \chi_{[a,x]}$, which implies that $F(x)$ exists and equals $\int_a^x f dm$.

The other direction is more difficult. Indeed, it is almost equivalent to the duality $L^p([a,b],m) \simeq L^q([a,b],m)^*$. See Thm. 2.7.1. A classical proof is as follows. One shows that $F :=$ (2.12) is absolutely-continuous. Thus, its a.e. derivative $f = F'$ is integrable, and $\int f \chi_{[a,x]} = \int_a^x F' = F(x) = \lim_n \int f_n \chi_{[a,x]}$ by the fundamental theorem of calculus for the Lebesgue integral. Thus $\int fg = \lim_n \int f_n g$ for each step function g. From this fact and the density of step functions in $L^p([a,b],m)$ (cf. Thm. 1.7.13), one shows that $f \in L^p$ (by showing $\|f\|_p \leqslant \sup_n \|f_n\|_p$) and that (f_n) converges weak-* to f. □

Theorem 2.2.9. *Let (ρ_n) be a uniformly l^∞-bounded sequence of increasing functions $[a,b] \to \mathbb{R}_{\geqslant 0}$. The following are true.*

1. *Let $\rho : [a,b] \to \mathbb{R}_{\geqslant 0}$ be bounded and increasing. Then $(d\rho_n)$ converges weak-* to $d\rho$ iff (ρ_n) converges pointwise to ρ at b and at any point where $\rho|_{(a,b)}$ is continuous.*

2. *$(d\rho_n)$ converges weak-* to $d\rho$ for some bounded increasing $\rho : [a,b] \to \mathbb{R}_{\geqslant 0}$ iff (ρ_n) converges pointwise at b and on a dense subset of I.*

By saying that $(d\rho_n)$ converges weak-* to $d\rho$, we mean $\lim_n \int g d\rho_n = \int g d\rho$ for all $g \in C([a,b],m)$.

Proof. See Thm. 2.9.6 and Cor. 2.9.7. □

The above theorems establish an intimate connection between the (pointwise) convergence of moments and the pointwise convergence of (the antiderivatives of) a sequence of functions.[4] Our understanding of convergence from various perspectives can thus be summarized in Table 2.2.

[4]We are viewing ρ_n and ρ as the antiderivatives of $d\rho_n$ and $d\rho$.

Function theory	Moment Problems	Dual spaces
Pointwise convergence of (antiderivatives of) a sequence of functions	Pointwise convergence of moments	Weak-* convergence

Table 2.2: Equivalence of convergence notions

2.2.4 Equivalence of the first and second columns of Table 2.2

Thm. 2.2.8 and 2.2.9, which establish the equivalence of the first and second columns of Table 2.2, are not easy to prove. In fact, proving Thm. 2.2.8 typically requires the duality $L^p([a,b]) \simeq L^q([a,b])^*$, or at least techniques closely related to those used in establishing this duality.

Therefore, the solvability of the moment problems (Problems 2.2.2)—equivalently, the solvability of Problem 2.2.5 concerning the characterization of dual spaces—is closely related to the equivalence between the first and second columns of Table 2.2. This close connection rests on the following principle:

Principle 2.2.10. Usually, if V is a normed vector space consisting of functions, any element φ of V^* can be weak-* approximated by elementary functions with uniformly bounded norms. More precisely, there exists a sequence (or a net) of elementary functions (f_n) such that the operator norms of the linear functionals $\xi \in V \mapsto \int \xi f_n$ are uniformly bounded, and

$$\lim_n \int \xi f_n = \langle \xi, \varphi \rangle \qquad \text{for all } \xi \in V$$

Remark 2.2.11. Here is how, with the help of Principle 2.2.10, the characterization of V^* can be derived from the equivalence of the first and second columns of Table 2.2:

By this principle, for each $\varphi \in V^*$, we can select a sequence (f_n) approximating weak-* to φ. Since the second column of Table 2.2 implies the first column, the sequence (f_n) converges to some function f in the sense described in the first column of Table 2.2. Then, by the equivalence of the three modes of convergence in that table, it follows that (f_n) converges weak-* to f. Consequently, φ is represented by integration against f, thereby solving the problem of characterizing the dual space V^*. □

The idea outlined in Rem. 2.2.11 is roughly the approach Riesz employed in [Rie07a] to solve the following trigonometric moment problem. See [Haw-L, Ch. 6] or [Rie07a].

Theorem 2.2.12 (Riesz-Fischer theorem, Riesz's original version). [5] *For each $(c_k)_{k\in\mathbb{Z}}$ in $l^2(\mathbb{Z})$, there is a (necessarily unique) $f \in L^2([-\pi,\pi],\frac{m}{2\pi})$ whose Fourier series is equal to (c_k).*

Riesz's proof. Choose $(c_k)_{k\in\mathbb{Z}}$ in $l^2(\mathbb{Z})$. Riesz aimed to solve the moment problem that there exists $f \in L^2$ such that $\frac{1}{2\pi}\int fe_{-k} = c_k$ for all $k \in \mathbb{Z}$, where $e_k(x) = e^{\mathbf{i}kx}$. For each $0 < r < 1$, let

$$f_r = \sum_{n\in\mathbb{Z}} r^{|n|}c_n e_n$$

which belongs to $C^\infty(\mathbb{S}^1)$. Then $\lim_{r\to 1} f_r$ converges weak-* to the bounded linear functional $\varphi \in (L^2)^*$ satisfying $\langle e_{-k},\varphi\rangle = c_k$ for all k. (This is an instance of Principle 2.2.10.)

Using a result of Fatou, Riesz showed that $\lim_{r\to 1} f_r$ converges a.e. to some Lebesgue-measurable function f, thereby partially illustrating that the second cell of Table 2.2 implies the first. Using Thm. 2.2.7, Riesz deduced that $f \in L^2$, and that $\lim_{r\to 1} f_r$ converges weak-* to f. This can be seen as an example that the first cell of Table 2.2 implies the second one. Since $\lim_{r\to 1} f_r$ also converges weak-* to φ, we conclude that φ is represented by the integral against f, and hence f solves the desired moment problem. [6] □

See [Gui-A, Sec. 27.3] for further discussion on the relationship between the classical and modern proofs of the duality $L^p \simeq (L^q)^*$, and the connection between this duality and the completeness of L^p-spaces. In Ch. 4, we will see that the strategy outlined in Rem. 2.2.11 was also the original approach used to study polynomial moment problems, i.e., the moment problems related to the dual space of $C(I)$.

2.2.5 Conclusion

We now summarize the discussion so far by addressing the question posed at the beginning of this section: Why are dual spaces related to integral theory?

[5] The modern interpretation of the Riesz-Fischer theorem as stating that $L^2(X,\mu)$ (or more generally $L^p(X,\mu)$) is Cauchy-complete for any measure space (X,μ) has led to a significant misunderstanding. In fact, while Fischer formulated the theorem for $L^2([-\pi,\pi],\frac{m}{2\pi})$ in terms of Cauchy sequences, Riesz understood it quite differently—through the lens of moment problems.

Therefore, once Riesz realized that solving moment problems is equivalent to the characterization of dual spaces, he immediately obtained the Riesz-Fréchet Thm. 2.2.1 in [Rie07b]. As we have emphasized at the beginning of Sec. 2.1, completeness and duality are fundamentally distinct properties, each serving distinct purposes and arising from different considerations. The fact that they coincide in the case of inner product spaces is purely a coincidence.

[6] Another proof using the equivalence in Table 2.2: Let $f_n = \sum_{k=-n}^{n} c_k e_k$. Then (f_n) converges weak-* to φ, again illustrating Principle 2.2.10. One easily shows that $\lim_n \int_{-\pi}^{x} f_n$ converges for each x. Therefore, by Thm. 2.2.8, (f_n) converges weak-* to some L^2-function f. Hence f solves the desired moment problem.

More specifically, from the mathematical-historical perspective, why is it possible to characterize the dual spaces of $L^p(X, \mu)$ and $C(X)$?

Function theory	Moment Problems	Dual spaces
	Solving moment problems	Characterizing V^*
Related by $\Updownarrow$ Principle 2.2.10		
Pointwise convergence of (antiderivatives of) a sequence of functions	Pointwise convergence of moments	Weak-* convergence

Table 2.3: The cells in each row are equivalent

The answer, in my view, is captured in Table 2.3: The power of the Lebesgue and Stieltjes integrals lies in their ability to establish the equivalence between the two gray cells in that table. Once this equivalence is established, with the help of Principle 2.2.10, the characterization of dual spaces in terms of integrals becomes straightforward.

But why are these two integrals powerful enough to establish the equivalence between the two gray cells in Table 2.3?—Because both the Lebesgue and Stieltjes integrals arise from the study of moment problems, which in turn are rooted in the corresponding approximation problems, as illustrated in Table 2.1. The emphasis of these integral theories on the commutativity of limits and integration anticipates the equivalence of the two gray cells.

In light of the equivalences in Table 2.3, the Lebesgue integral, as the completion of the Riemann integral, can be interpreted as the weak-* completion of trigonometric functions and continuous functions. Similarly, the Stieltjes integral, as the completion of finite sums, can be viewed as the weak-* completion of discrete spectra—a perspective that will be one of the main themes of Ch. 4 and 5. See Table 2.4.

Completion of Integrals	Extension of classes of functions	**Weak-* completion**
Riemann integral $\cap$ Lebesgue integral	Continuous functions $\cap$ Measurable functions	of continuous functions
Finite sum $\cap$ Stieltjes integral	Discrete spectra $\cap$ Continuous spectra	of discrete spectra

Table 2.4

Side note. A common viewpoint—motivated by the completeness of L^1-spaces—regards the Lebesgue integral and the Lebesgue measurable/integrable functions as the Cauchy completion of Riemann integrals and continuous functions. In my view, this perspective is not only historically inaccurate, but also mathematically misleading.

Historically, the first L^p-space considered is $L^2([a,b],m)$, due to its close relation with $l^2(\mathbb{Z})$, the space of trigonometric moments of L^2-integrable functions. The space $l^2(\mathbb{Z})$ was introduced by Hilbert in [Hil06], where weak convergence (equivalently, pointwise convergence of moments) plays a central role in his proof of the Hilbert-Schmidt theorem. In [Rie10], Riesz studied the space $L^p([a,b],m)$ for $1 < p < +\infty$, and in particular proved the duality $L^p([a,b],m) \simeq L^q([a,b],m)^*$. The completeness of $L^p([a,b],m)$ follows as a corollary. However, $L^1([a,b],m)$ was not considered, likely due to its lack of a satisfactory duality structure. This clearly shows that duality was originally viewed as more fundamental than Cauchy completeness.

Mathematically, to perform a Cauchy completion, one needs a norm, which in this context is defined via an integral. Yet, while integrals are linear functionals, norms only satisfy the subadditivity. As a result, norms and Cauchy completions do not provide the right conceptual framework for understanding the nature of the Lebesgue integral from a functional-analytic perspective.

The more appropriate viewpoint is to regard the Lebesgue integral as arising from weak-* completion, not Cauchy completion.

2.3 Bounded multilinear maps

2.3.1 Seminorms, norms, normed vector spaces, and Banach spaces

Definition 2.3.1. If V is an $\mathbb{F}$-vector space, a function $\|\cdot\| : V \to \mathbb{R}_{\geqslant 0}$ is called a **seminorm** if

$$\|av\| = |a| \cdot \|v\| \qquad \|u+v\| \leqslant \|u\| + \|v\| \qquad \text{for any } u,v \in V \text{ and } a \in \mathbb{F} \tag{2.14}$$

A seminorm is called a **norm** if any $v \in V$ satisfying $\|v\| = 0$ is the zero vector 0. A vector space V, equipped with a norm, is called a **normed vector space**.

If V is a normed vector space, then a **normed vector subspace** of V denotes a linear subspace $U \subset V$ equipped with the norm inherited from V, i.e., the restriction of V's norm to U.

We say that V is **separable** if it is so in the **norm topology**, namely, the topology induced by the metric $d(u,v) = \|u - v\|$. □

Remark 2.3.2. In Def. 2.3.1, the condition $\|av\| = |a| \cdot \|v\|$ can be weakened to

$$\|av\| \leqslant |a| \cdot \|v\| \qquad \text{for any } v \in V \text{ and } a \in \mathbb{F} \tag{2.15}$$

Therefore, (2.14) can be weakened to

$$\|au + bv\| \leqslant |a| \cdot \|u\| + |b| \cdot \|v\| \qquad \text{for any } u,v \in V \text{ and } a,b \in \mathbb{F} \tag{2.16}$$

Proof. Suppose that (2.15) is true. Then we clearly have $\|av\| = |a|\cdot\|v\|$ when $a = 0$. Suppose that $a \neq 0$. Then $\|v\| = \|a^{-1}av\| \leqslant |a|^{-1}\|av\|$, and hence $\|av\| \geqslant |a|\cdot\|v\|$. Therefore $\|av\| = |a|\cdot\|v\|$. □

Remark 2.3.3. The norm function $\|\cdot\| : V \to \mathbb{R}_{\geqslant 0}$ is continuous. This is because

$$\|u\| - \|v\| \leqslant \|u - v\| \tag{2.17}$$

Therefore, if (v_α) is a net in V converging (in norm) to v, then

$$\|v\| = \lim_\alpha \|v_\alpha\|$$

Proposition 2.3.4. *Let $\|\cdot\|_V$ be a seminorm on an $\mathbb{F}$-vector space V. Let $V_0 = \{v \in V : \|v\|_V = 0\}$. Then V_0 is a linear subspace on V, and there is a (clearly unique) norm $\|\cdot\|_{V/V_0}$ on the quotient space V/V_0 such that*

$$\|v + V_0\|_{V/V_0} = \|v\|_V \qquad \text{for all } v \in V \tag{2.18}$$

In the future, unless otherwise stated, we will always equip V/V_0 with this norm $\|\cdot\|_{V/V_0}$.

Proof. We abbreviate $\|\cdot\|_V$ to $\|\cdot\|$. If $u, v \in V_0$ and $a, b \in \mathbb{F}$, then

$$\|au + bv\| \leqslant |a|\|u\| + |b|\|v\| = 0$$

This shows that V_0 is a linear subspace of V. On the other hand, if $u, v \in V$ satisfy $u + V_0 = v + V_0$, then $u - v \in V_0$, and hence

$$\|v\| = \|u + v - u\| \leqslant \|u\| + \|v - u\| = \|u\|$$

Similarly, $\|u\| \leqslant \|v\|$. Therefore $\|u\| = \|v\|$. This implies that we have a well-defined function $\|\cdot\|_{V/V_0} : V/V_0 \to \mathbb{R}_{\geqslant 0}$ satisfying (2.18).

If $u, v \in V$ and $a, b \in \mathbb{F}$, then

$$\begin{aligned}&\|a(u + V_0) + b(v + V_0)\|_{V/V_0} = \|au + bv + V_0\|_{V/V_0} = \|au + bv\|\\ \leqslant&|a|\|u\| + |b|\|v\| = |a|\|u + V_0\|_{V/V_0} + |b|\|v + V_0\|_{V/V_0}\end{aligned}$$

□

Definition 2.3.5. A complete normed $\mathbb{F}$-vector space is called a **Banach $\mathbb{F}$-space** (or **Banach space** over $\mathbb{F}$). A **real Banach space** and a **complex Banach space** mean a Banach space over $\mathbb{R}$ and $\mathbb{C}$, respectively.

2.3.2 Bounded multilinear maps

In the rest of this section, $V_1, V_2, \ldots$ and U, V, W all denote normed $\mathbb{F}$-vector spaces.

Definition 2.3.6. Let $N \in \mathbb{Z}_+$. A map $T : V_1 \times \cdots \times V_N \to W$ is called a **multilinear map** if for each $1 \leqslant i \leqslant N$ and each fixed $v_j \in V_j$ (for all $j \neq i$), the map

$$v_i \in V_i \mapsto T(v_1, \ldots, v_N) \in W$$

is $\mathbb{F}$-linear. We let

$$\mathbf{Lin}(\boldsymbol{V_1} \times \cdots \times \boldsymbol{V_N}, \boldsymbol{W}) = \{\text{multilinear maps } V_1 \times \cdots \times V_N \to W\}$$

For each $T \in \mathrm{Lin}(V_1 \times \cdots \times V_N, W)$, we define the **operator norm**

$$\|T\| := \|T\|_{l^\infty(\overline{B}_{V_1}(0,1) \times \cdots \times \overline{B}_{V_N}(0,1), W)} = \sup_{v_1 \in \overline{B}_{V_1}(0,1), \ldots, v_N \in \overline{B}_{V_N}(0,1)} \|T(v_1, \ldots, v_N)\|$$

We say that T is **bounded** if $\|T\| < +\infty$.

Definition 2.3.7. We let

$$\boldsymbol{\mathfrak{L}}(\boldsymbol{V_1} \times \cdots \times \boldsymbol{V_N}, \boldsymbol{W}) := \{\text{bounded multilinear maps } V_1 \times \cdots \times V_N \to W\} \tag{2.19}$$

viewed as an $\mathbb{F}$-linear subspace of $W^{V_1 \times \cdots \times V_N}$. We let

$$\mathfrak{L}(V) := \mathfrak{L}(V, V) \qquad V^* := \mathfrak{L}(V, \mathbb{F})$$

Elements of $\mathfrak{L}(V)$ are called **bounded linear operators on** V. An element $T \in \mathfrak{L}(V)$ is called **invertible** if there exists $T^{-1} \in \mathfrak{L}(V)$ such that

$$TT^{-1} = T^{-1}T = \mathrm{id}_V$$

The space V^* is called the (analytic) **dual space** of V.

Remark 2.3.8. In this course, the most frequently encountered cases of (2.19) are $\mathfrak{L}(V)$, V^*, and $\mathfrak{L}(U \times V, \mathbb{F})$.

Remark 2.3.9. $\|T\|$ is the smallest element in $\overline{\mathbb{R}}_{\geqslant 0}$ satisfying

$$\|T(v_1, \ldots, v_N)\| \leqslant \|T\| \cdot \|v_1\| \cdots \|v_N\| \tag{2.20}$$

Proof. If one of $v_1, \ldots, v_N$ is zero, then $T(v_1, \ldots, v_N) = 0$ by the multilinearity, and hence (2.20) holds. So we assume that $v_1, \ldots, v_N$ are all non-zero. So their norms are all nonzero. Since $v_i/\|v_i\| \in B_{V_i}(0, 1)$, we have

$$\Big\|T\Big(\frac{v_1}{\|v_1\|}, \cdots, \frac{v_N}{\|v_N\|}\Big)\Big\| \leqslant \|T\|$$

which implies (2.20) by the multilinearity.

We have proved that $\|T\|$ satisfies (2.20). Now, suppose that $C \in \overline{\mathbb{R}}_{\geqslant 0}$ and

$$\|T(v_1, \ldots, v_N)\| \leqslant C \cdot \|v_1\| \cdots \|v_N\|$$

for all $v_i \in V_i$. Taking $v_i \in \overline{B}_V(0, 1)$, we see that $\|T\| \leqslant C$. $\square$

Recall Def. 1.4.3.

Proposition 2.3.10. *Let $T : V_1 \times \cdots \times V_N \to W$ be multilinear. The following are equivalent.*

(a) T is continuous.

(b) T is continuous at $0 \times \cdots \times 0$.

(c) T is bounded.

(d) T is Lipschitz continuous on $\overline{B}_{V_1}(0,R) \times \cdots \times \overline{B}_{V_N}(0,R)$ for every $R \in \mathbb{R}_{>0}$.

(e) T is Lipschitz continuous on $\overline{B}_{V_1}(0,1) \times \cdots \times \overline{B}_{V_N}(0,1)$.

Moreover, if T is bounded, and if $V_1 \times \cdots \times V_N$ is equipped with the l^∞-product metric, then the Lipschitz constant in (d) can be chosen to be $NR^{N-1}\|T\|$.

What matters about the Lipschitz constant above is not its exact formula, but the implication it carries: namely, that any family (T_α) in $\mathfrak{L}(V_1 \times \cdots \times V_N, W)$ satisfying $\sup_\alpha \|T_\alpha\| < +\infty$, when restricted to a bounded subset of $V_1 \times \cdots \times V_N$, admits a uniform Lipschitz constant.

Proof. Clearly (a)⇒(b).

(b)⇒(c): Assume (b). Then $0 \times \cdots \times 0$ is an interior point of $T^{-1}(B_W(0,1))$, and hence contains $B_{V_1}(0,2\delta_1) \times \cdots \times B_{V_N}(0,2\delta_N)$ for some $\delta_1,\dots,\delta_N > 0$. So T sends $\overline{B}_{V_1}(0,\delta_1) \times \cdots \times \overline{B}_{V_N}(0,\delta_N)$ (which equals $\delta_1\overline{B}_{V_1}(0,1) \times \cdots \times \delta_N\overline{B}_{V_N}(0,1)$) into $B_W(0,1)$. By multilinearity, T sends $\overline{B}_{V_1}(0,1) \times \cdots \times \overline{B}_{V_N}(0,1)$ into $B_W(0,\delta_1^{-1}\cdots\delta_N^{-1})$. This proves (c).

(c)⇒(d): Assume (c). Choose $v_i \in \overline{B}_{V_i}(0,R)$. Then, for each $\xi_i \in \overline{B}_{V_i}(0,R)$,

$$\begin{aligned}
&\|T(\xi_1,\dots,\xi_N) - T(v_1,\dots,v_N)\| \\
\leqslant&\|T(\xi_1 - v_1,\xi_2,\xi_3,\dots,\xi_N)\| + \|T(v_1,\xi_2 - v_2,\xi_3,\dots,\xi_N)\| \\
&+ \|T(v_1,v_2,\xi_3 - v_3,\dots,\xi_N)\| + \cdots + \|T(v_1,v_2,v_3,\dots,\xi_N - v_N)\| \\
\leqslant&NR^{N-1}\|T\| \cdot \max\{\|\xi_1 - v_1\|,\dots,\|\xi_N - v_N\|\}
\end{aligned}$$

where (2.20) is used in the last inequality. Thus T has Lipschitz constant $NR^{N-1}\|T\|$.

(e)⇔(d): This is clear by scaling the vectors.

(d)⇒(a): This is clear from Rem. 1.4.4. □

Corollary 2.3.11. *Let $T \in \mathfrak{L}(V,W)$. Then $\mathrm{Ker}(T)$ is a closed linear subspace of V.*

Proof. By Prop. 2.3.10, T is continuous. Since the preimage of any closed set under a continuous map is closed, $\mathrm{Ker}(T) = T^{-1}(0)$ is closed. □

Example 2.3.12. A linear map $T : V \to W$ is called a **linear isometry** if it is an isometry of metric spaces, i.e., $\|Tv_1 - Tv_2\| = \|v_1 - v_2\|$ for all $v_1, v_2 \in V$. This is clearly equivalent to

$$\|Tv\| = \|v\| \qquad \text{for all } v \in V$$

A linear isometry is clearly bounded with operator norm $\|T\| = 1$ (unless when $V = \{0\}$). Moreover, a linear isometry is clearly injective. A linear isometry $T : V \to W$ which is also surjective (and hence bijective) is called an **isomorphism of normed vector spaces**. In that case, we say that the normed vector spaces V, W are **isomorphic**.

Remark 2.3.13. Suppose that $\Phi : V \to W$ is a linear map of vector spaces, and W is a normed vector space. Then V has a seminorm defined by

$$\|v\|_V := \|\Phi(v)\|_V$$

Equip $V/\mathrm{Ker}\Phi$ with the norm defined by Prop. 2.3.4. Then Φ descends to a linear map $\widetilde{\Phi} : V/\mathrm{Ker}\Phi \to W$, which is clearly a linear isometry.

Example 2.3.14. Let $1 \leqslant p \leqslant +\infty$, let X be an LCH space, let μ be a Radon measure (or its completion) on X. Let $\Phi : C_c(X, \mathbb{F}) \to L^p(X, \mu, \mathbb{F})$ be the obvious map. Then Φ descends to a linear isometry of normed vector spaces

$$C_c(X, \mathbb{F})/\{f \in C_c(X, \mathbb{F}) : f = 0\ \mu\text{-a.e.}\} \longrightarrow L^p(X, \mu, \mathbb{F}) \tag{2.21}$$

Now assume $p < +\infty$. Then by Thm. 1.7.10, the map (2.21) has dense range. This is often expressed by saying that $C_c(X, \mathbb{F})/\{f \in C_c(X, \mathbb{F}) : f = 0\ \mu\text{-a.e.}\}$ is dense in $L^p(X, \mu, \mathbb{F})$, or simply that $C_c(X, \mathbb{F})$ is dense in $L^p(X, \mu, \mathbb{F})$.

2.4 Fundamental properties of bounded multilinear maps

Let $V_1, V_2, \dots, U, V, W$ be normed vector spaces. In this section, we establish several fundamental properties of bounded multilinear maps that will be used frequently throughout the course. The importance of these properties will be discussed in Sec. 2.5. We first note the elementary fact:

Remark 2.4.1. Let U be a linear subspace of V. Let $R \in \mathbb{R}_{>0}$. Then U is dense in V iff $\overline{B}_U(0, R)$ is dense in $\overline{B}_V(0, R)$.

Proof. The direction "$\Leftarrow$" is obvious. Let us prove "$\Rightarrow$". Let $\xi \in \overline{B}_V(0, R)$, choose a sequence (ξ_n) in U converging to ξ. Assume WLOG that $\xi \neq 0$ and $R \in \mathbb{R}_{>0}$; otherwise, the approximation is obvious. Since the norm function is continuous, $\|\xi_n\| \to \|\xi\|$. In particular, $\|\xi_n\|$ is eventually nonzero. Thus $\frac{\|\xi\|}{\|\xi_n\|}\xi_n \to \xi$. $\square$

Recall that two sequences $(x_n), (y_n)$ in a metric space X is called **Cauchy equivalent** if $\lim_n d(x_n, y_n) = 0$.

Theorem 2.4.2. *Suppose that W is complete. For each i, let U_i be a dense linear subspace of V_i. Then we have an isomorphism of normed vector spaces*

$$\begin{gathered}\mathfrak{L}(V_1 \times \cdots \times V_N, W) \xrightarrow{\simeq} \mathfrak{L}(U_1 \times \cdots \times U_N, W) \\ T \mapsto T\big|_{U_1 \times \cdots \times U_N}\end{gathered} \tag{2.22}$$

The following proof shows that the map (2.22) is a linear isometry even without assuming that W is complete.

Proof. By Rem. 2.4.1, $\overline{B}_{U_1}(0,1) \times \cdots \times \overline{B}_{U_N}(0,1)$ is dense in $\overline{B}_{U_1}(0,1) \times \cdots \times \overline{B}_{U_N}(0,1)$. Since the l^∞-norm of a continuous function is unchanged when the domain is restricted to a dense subset, it follows that $T\big|_{U_1 \times \cdots \times U_N}$ have the same operator norm. Thus (2.22) is a linear isometry.

We now show that (2.22) is surjective. Here, the completeness of W is needed. Let $T \in \mathfrak{L}(U_1 \times \cdots \times U_N, W)$. It suffices to extend T in the first component to a bounded multilinear $V_1 \times U_2 \times U_3 \times \cdots \times U_N \to W$. Then, a similar argument applies to the second component, extending T to a bounded multilinear $V_1 \times V_2 \times U_3 \times \cdots \times U_N \to W$. By repeating this procedure, we obtain a bounded multilinear map $V_1 \times \cdots \times V_N \to W$ extending T.

Let $\xi \in V_1, u_2 \in U_2, \dots, u_N \in U_N$. Let (ξ_n) be a sequence in U_1 converging to ξ. In particular, (ξ_n) is a Cauchy sequence. By Rem. 2.3.9, $T(\xi_n, v_2, \dots, v_N)$ is a Cauchy sequence in W. Therefore, by the completeness of W, $T(\xi_n, v_2, \dots, v_N)$ converges to some element, which we denote by $T(\xi, v_2, \dots, v_N)$.

Let us show that the definition of $T(\xi, v_2, \dots, v_N)$ is independent of the choice of sequence converging to ξ. Suppose that (ξ_n') is another sequence converging to ξ. Then (ξ_n) and (ξ_n') are Cauchy equivalent. By Rem. 2.3.9, $T(\xi_n, v_2, \dots, v_N)$ and $T(\xi_n', v_2, \dots, v_N)$ are Cauchy equivalent. So they converge to the same element.

Thus, we have defined a map $T : V_1 \times U_2 \times \cdots \times U_N \to W$. We leave it to the reader to check that T is bounded multi-linear map. $\square$

Corollary 2.4.3. *Let U be a dense linear subspace of V. Then we have an isomorphism of normed vector spaces*

$$V^* \xrightarrow{\simeq} U^* \qquad \varphi \mapsto \varphi|_U \tag{2.23}$$

Proof. This follows immediate from Thm. 2.4.2. $\square$

Example 2.4.4. Let $1 \leqslant q < +\infty$ and $p^{-1} + q^{-1} = 1$. Let X be an LCH space. Let μ be a Radon measure (or its completion) on X. By Exp. 2.3.14, the L^q-seminorm on $C_c(X, \mathbb{F})$ descends to the L^q-norm on $V = C_c(X, \mathbb{F})/\{f \in C_c(X, \mathbb{F}) : f = 0 \ \mu\text{-a.e.}\}$, and V is dense in $L^p(X, \mu)$. Therefore, by Thm. 1.6.16 and Cor. 2.4.3, the map (1.34) gives an isomorphism of normed vector spaces $V^* \simeq L^p(X, \mu)$.

The following Prop. 2.4.5 and Thm. 2.4.6 will imply Thm. 2.6.2, which establishes the equivalence of the second and third columns of Table 2.2.

Proposition 2.4.5. *For each i, let E_i be a densely spanning subset of V_i. Let (T_α) be a net in $\mathfrak{L}(V_1 \times \cdots \times V_N, W)$ with* ***uniformly bounded operator norms****, i.e.,* $\sup_\alpha \|T_\alpha\| < +\infty$. *Choose $T \in \mathfrak{L}(V_1 \times \cdots \times V_N, W)$, and assume that (T_α) converges pointwise on $E_1 \times \cdots \times E_N$ to T. Then (T_α) converges pointwise on $V_1 \times \cdots \times V_N$ to T.*

Proof. Let $U_i = \mathrm{Span}(E_i)$, which is dense in V_i. Then (T_α) converges pointwise on $U_1 \times \cdots \times U_N$ to T.

Choose any $\xi_i \in V_i$. Choose $R \in \mathbb{R}_{>0}$ such that $\|\xi_i\| \leqslant R$ for each i. Since $\sup_\alpha \|T_\alpha\| < +\infty$, by Prop. 2.3.10, $\{T_\alpha, T : \alpha \in I\}$ has a uniform Lipschitz constant $C \in \mathbb{R}_{\geqslant 0}$ (with respect to the l^∞-product metric) when restricted to $\overline{B}_{V_1}(0, R) \times \cdots \times \overline{B}_{V_N}(0, R)$. By Rem. 2.4.1, for each $\varepsilon > 0$, there exists $v_i \in \overline{B}_{U_i}(0, R)$ such that $\|\xi_i - v_i\| \leqslant \varepsilon$. Then

$$
\begin{aligned}
&\limsup_\alpha \|T(\xi_1, \ldots, \xi_N) - T_\alpha(\xi_1, \ldots, \xi_N)\| \\
\leqslant &\limsup_\alpha \|T(v_1, \ldots, v_N) - T_\alpha(v_1, \ldots, v_N)\| + 2C\varepsilon = 2C\varepsilon
\end{aligned}
$$

Since ε is arbitrary, we conclude that $T_\alpha(\xi_1, \ldots, \xi_N) \to T(\xi_1, \ldots, \xi_N)$. $\square$

Theorem 2.4.6. *Suppose that W is complete. For each i, let E_i be a densely spanning subset of V_i. Let (T_α) be a net in $\mathfrak{L}(V_1 \times \cdots \times V_N, W)$ satisfying $\sup_\alpha \|T_\alpha\| < +\infty$. Suppose that (T_α) converges pointwise on $E_1 \times \cdots \times E_N$. Then (T_α) converges pointwise on $V_1 \times \cdots \times V_N$ to some $T \in \mathfrak{L}(V_1 \times \cdots \times V_N, W)$, and*

$$
\|T\| \leqslant \liminf_\alpha \|T_\alpha\| \tag{2.24}
$$

Inequality (2.24) is sometimes referred to as **Fatou's lemma**.

Proof. Let $U_i = \mathrm{Span}(E_i)$, which is dense in V_i. Let $T : U_1 \times \cdots \times U_N \to W$ be the pointwise limit of $(T_\alpha)_{\alpha \in I}$ restricted to $U_1 \times \cdots \times U_N$, which is clearly linear. Moreover, for each $v_i \in \overline{B}_{U_i}(0, 1)$ we have

$$
\|T(v_1, \ldots, v_N)\| = \liminf_\alpha \|T_\alpha(v_1, \ldots, v_N)\| \leqslant \liminf_\alpha \|T_\alpha\|
$$

Taking sup over all $v_i \in \overline{B}_{U_i}(0, 1)$, we see that $\|T\| \leqslant \liminf_\alpha \|T_\alpha\| < +\infty$. In particular, $T \in \mathfrak{L}(U_1 \times \cdots \times U_N, W)$. By Thm. 2.4.2, T can be extended to a bounded multilinear map $T : V_1 \times \cdots \times V_N \to W$ with $\|T\|$ unchanged. By Prop. 2.4.5, this extended T is the pointwise limit of (T_α) on the whole domain $V_1 \times \cdots \times V_N$. $\square$

Remark 2.4.7. Recall that if X is a set, then $l^\infty(X, W)$, equipped with the l^∞-norm, is a normed vector space.

By the definition of operator norms, we have a linear isometry of normed vector spaces

$$\begin{aligned}\mathfrak{L}(V_1\times\cdots\times V_N,W)&\rightarrow l^\infty(\overline{B}_{V_1}(0,1)\times\cdots\times\overline{B}_{V_N}(0,1),W)\\ T&\mapsto T|_{\overline{B}_{V_1}(0,1)\times\cdots\times\overline{B}_{V_N}(0,1)}\end{aligned}\tag{2.25}$$

Therefore, by identifying $\mathfrak{L}(V_1\times\cdots\times V_N,W)$ with its image under (2.25), we view $\mathfrak{L}(V_1\times\cdots\times V_N,W)$ as a normed vector subspace of $l^\infty(\overline{B}_{V_1}(0,1)\times\cdots\times\overline{B}_{V_N},W)$.

Consequently, if (T_α) is a net in $\mathfrak{L}(V_1\times\cdots\times V_N,W)$, and if $T\in\mathfrak{L}(V_1\times\cdots\times V_N,W)$, then $\lim_\alpha\|T-T_\alpha\|=0$ is equivalent to that (T_α) converges uniformly to T on $\overline{B}_{V_1}(0,1)\times\cdots\times\overline{B}_{V_N}(0,1)$. □

Theorem 2.4.8. *$\mathfrak{L}(V_1\times\cdots\times V_N,W)$ is a closed linear subspace of $l^\infty(\overline{B}_{V_1}(0,1)\times\cdots\times\overline{B}_{V_N}(0,1),W)$.*

Proof. Let $T\in l^\infty(\overline{B}_{V_1}(0,1)\times\cdots\times\overline{B}_{V_N}(0,1),W)$ be the limit of a sequence (T_n) in $\mathfrak{L}(V_1\times\cdots\times V_N,W)$. Then (T_n) converges uniformly on $\overline{B}_{V_1}(0,1)\times\cdots\times\overline{B}_{V_N}(0,1)$ to T. By scaling the vectors, we see that (T_n) converges uniformly on $\overline{B}_{V_1}(0,R)\times\cdots\times\overline{B}_{V_N}(0,R)$ for any $R>0$. Let $T:V_1\times\cdots\times V_N\rightarrow W$ be the pointwise limit of (T_n), which automatically extends the original T defined on $\overline{B}_{V_1}(0,1)\times\cdots\times\overline{B}_{V_N}(0,1)$.

Since each T_n is multilinear, clearly T is multilinear. Thus $T\in\mathfrak{L}(V_1\times\cdots\times V_N,W)$. This proves that $\mathfrak{L}(V_1\times\cdots\times V_N,W)$ is a closed. □

Corollary 2.4.9. *Suppose that W is complete. Then $\mathfrak{L}(V_1\times\cdots\times V_N,W)$ is complete.*

Proof. Since W is complete, by the following Prop. 2.4.10, $l^\infty(\overline{B}_{V_1}(0,1)\times\cdots\times\overline{B}_{V_N},W)$ is complete. Since any closed subset of a complete space is complete, by Thm. 2.4.8, $\mathfrak{L}(V_1\times\cdots\times V_N,W)$ is complete. □

Proposition 2.4.10. *Suppose that W is complete. Then for each $1\leqslant p\leqslant+\infty$, the normed vector space $l^p(X,W)$ is complete.*

Proof. Let (f_n) be a Cauchy sequence in $l^p(X,W)$. Then for each $x\in X$, $(f_n(x))$ is a Cauchy sequence in W, and hence converges to some $f(x)\in W$. This defines $f:X\rightarrow W$.

Case $p=+\infty$: For each $\varepsilon>0$, choose $N\in\mathbb{Z}_+$ such that for all $m,n\geqslant N$ we have $\|f_n-f_m\|_{l^\infty}\leqslant\varepsilon$, i.e., $\|f_n(x)-f_m(x)\|\leqslant\varepsilon$ for every $x\in X$. Applying $\lim_{m\rightarrow\infty}$, we get $\|f_n(x)-f(x)\|\leqslant\varepsilon$ for all $x\in X$ and $n\geqslant N$. Thus, for all $n\geqslant N$ we have $\|f_n-f\|_{l^\infty}\leqslant\varepsilon$; in particular, we have $f\in l^\infty(X,W)$. Thus $\|f_n-f\|_{l^\infty}\rightarrow 0$.

Case $p<+\infty$: For each $\varepsilon>0$, choose $N\in\mathbb{Z}_+$ such that for all $m,n\geqslant N$ we have $\|f_n-f_m\|_{l^p(X)}\leqslant\varepsilon$, equivalently, $\|f_n-f_m\|_{l^p(A)}\leqslant\varepsilon$ for each $A\in\mathrm{fin}(2^X)$. Applying $\lim_{m\rightarrow\infty}$, we get $\|f_n-f\|_{l^p(A)}\leqslant\varepsilon$ for all $n\geqslant N$ and $A\in\mathrm{fin}(2^X)$. Thus $\|f_n-f\|_{l^p(X)}\leqslant\varepsilon$ for all $n\geqslant N$; in particular, we have $f\in l^p(X,W)$. This proves $\|f_n-f\|_p\rightarrow 0$. □

Corollary 2.4.11. *The dual space V^*, equipped with the operator norm, is complete.*

Proof. This follows immediately from Cor. 2.4.9. □

2.5 The roles of duality and Cauchy completeness

Let $V_1, \dots, V_N$ and V, W be normed vector spaces.

2.5.1 The role of Cauchy completeness

In functional analysis, Cauchy completeness plays two primary roles:

1. Completeness as a domain property, where it is often used in conjunction with the Baire category theorem.

2. Completeness as a codomain property, which ensures that linear operators can be restricted from the whole space to a dense subspace without loss. Thm. 2.4.2 and 2.4.6 are typical examples illustrating this usage.

Among these two, completeness as a codomain is the more widely encountered in practice. This suggests that the recognition and widespread appreciation of **Cauchy completeness** in function spaces developed alongside the study of linear operators—that is, linear maps from V to W—rather than with linear, bilinear, or multilinear functionals, such as $V \times W \to \mathbb{F}$. In the early days of functional analysis, particularly in Hilbert's foundational work [Hil06], the dominant perspective was centered not on linear operators, but on bilinear forms and linear functionals. Within this (bi)linear framework, completeness is not required—indeed, in Thm 2.4.2, 2.4.6, and Corollary 2.4.9, when $W = \mathbb{F}$, none of the remaining vector spaces involved (namely $V_1, \dots, V_N$) are assumed to be complete.

Historically, the focus on bilinear forms gradually gave way to the linear operator viewpoint. As this shift took place, Cauchy completeness came to occupy a central role in functional analysis. The fact that the bilinear/multilinear form viewpoint can be reformulated in terms of linear operators is a consequence of the following elementary observation:

Proposition 2.5.1. *Let $U_1, \dots, U_M$ be normed vector spaces. Then we have an isomorphism of normed vector spaces*

$$\begin{gathered}\mathfrak{L}(U_1 \times \cdots \times U_M \times V_1 \times \cdots \times V_N, W) \xrightarrow{\simeq} \mathfrak{L}(U_1 \times \cdots \times U_M, \mathfrak{L}(V_1 \times \cdots \times V_N, W)) \\ T \mapsto \Big((u_1, \dots, u_M) \mapsto T(u_1, \dots, u_M, -, \dots, -) \Big)\end{gathered} \tag{2.26}$$

where $T(u_1, \dots, u_M, -, \dots, -)$ denotes the multilinear map $V_1 \times \cdots \times V_N \to W$ sending $(v_1, \dots, v_N)$ to $T(u_1, \dots, u_M, v_1, \dots, v_N)$.

Proof. It is easy to verify that the second line of (2.26) defines a linear isomorphism

$$\Psi : \mathrm{Lin}(U_1 \times \cdots \times U_M \times V_1 \times \cdots \times V_N, W)$$

$$\xrightarrow{\simeq} \mathrm{Lin}(U_1 \times \cdots \times U_M, \mathrm{Lin}(V_1 \times \cdots \times V_N, W))$$

To explain the idea of comparing the operator norms, we assume for simplicity that $M = N = 1$, and write $U_1 = U$ and $V_1 = V$.

Choose any $T \in \mathrm{Lin}(U \times V, W)$. Then $\Psi(T) : U \to \mathrm{Lin}(V, W)$ sends each $u \in V$ to the linear map

$$\Psi(T)(u) : v \in \mathrm{Lin}(V, W) \mapsto T(u, v)$$

Thus, for each $u \in U$ and $v \in V$, we have

$$\|T(u, v)\| = \|\Psi(T)(u)(v)\| \leqslant \|\Psi(T)(u)\| \cdot \|v\| \leqslant \|\Psi(T)\| \cdot \|u\| \cdot \|v\|$$

This proves $\|T\| \leqslant \|\Psi(T)\|$. Conversely, for each $u \in U$,

$$\|\Psi(T)(u)\| = \sup_{v \in \overline{B}_V(0,1)} \|\Psi(T)(u)(v)\| = \sup_{v \in \overline{B}_V(0,1)} \|T(u, v)\|$$
$$\leqslant \sup_{v \in \overline{B}_V(0,1)} \|T\| \cdot \|u\| \cdot \|v\| = \|T\| \cdot \|u\|$$

This proves $\|\Psi(T)\| \leqslant \|T\|$.

We have proved that $\|\Psi(T)\| = \|T\|$. In particular, if T is bounded, then $\Psi(T)(u)$ is bounded for each $u \in U$, and $\Psi(T)$ is bounded. Conversely, if $\Psi(T)(u)$ is bounded for each u, and if $\Psi(T)$ is bounded, then T is bounded. This proves that Ψ restricts to the linear isomorphism (2.26), which is an isometry because $\|\Psi(T)\| = \|T\|$. □

2.5.2 The role of duality

The following two corollaries follow immediate from Prop. 2.5.1.

Corollary 2.5.2. *We have an isomorphism of normed vector spaces*

$$\mathfrak{L}(U \times V, \mathbb{F}) \xrightarrow{\simeq} \mathfrak{L}(U, V^*) \qquad T \mapsto \big(u \mapsto T(u, -)\big) \tag{2.27}$$

Corollary 2.5.3. *Suppose that V is the dual space of another normed vector space V_*. Then we have an isomorphism of normed vector spaces*

$$\mathfrak{L}(V \times V_*, \mathbb{F}) \xrightarrow{\simeq} \mathfrak{L}(V) \qquad T \mapsto \big(v \mapsto T(v, -)\big) \tag{2.28}$$

In Sec. 2.1 and 2.2, we explored the motivation for introducing dual spaces from the perspectives of the calculus of variations and moment problems. Cor. 2.5.3 now offers yet another compelling reason for the study of duality: when a space V possesses a **dual structure**—specifically, when V is the dual of some normed space V_*—it allows us to approach problems from both the bilinear form and linear operator perspectives.

What are the respective advantages of these two viewpoints? To address this, I would like to revisit the arguments presented in [Gui-A], particularly in the Introduction and in Ch. 21 and 25 of [Gui-A]:

1. The bilinear form framework allows us to draw upon the full strength of measure theory. In fact, measure theory can be understood as a method of **monotone convergence extension**—a procedure for extending linear functionals in such a way that the monotone convergence theorem (or its variants) holds; see Sec. 5.3 for details. This type of extension aligns naturally with the structure of bilinear forms.

2. The space $\mathfrak{L}(V)$ of bounded linear operators on V is not just a vector space but also an algebra, with multiplication given by composition. This algebraic structure enables the use of **symbolic calculus**, a technique developed in the mid-19th century in the study of linear algebras, and it connects directly to the representation-theoretic perspectives that flourished in the 20th century.

As discussed in [Gui-A, Sec. 25.8, 25.9], and as we will also explore in Ch. 5, Riesz's spectral theorem provides a striking example of how these two advantages can be fruitfully combined.

2.6 Dual spaces and the weak-* topology

Let $V_1, V_2, \dots, U, V, W$ be normed $\mathbb{F}$-vector spaces.

Definition 2.6.1. By viewing V^* as a subset of $\mathbb{F}^V$, the subspace topology on V^* inherited from the product topology of $\mathbb{F}^V$ is called the **weak-* topology** on V^*. By Thm. 1.4.14, this is the unique topology such that for any net (φ_α) in V^* and any $\varphi \in V$, the net (φ_α) **converges weak-*** to φ—that is, converges to φ in the weak-* topology—iff

$$\lim_\alpha \langle \varphi_\alpha, v\rangle = \langle \varphi, v\rangle \qquad \text{for any } v \in V \tag{2.29}$$

Since $\mathbb{F}^V$ is Hausdorff, the weak-* topology is also Hausdorff.

Weak-* topology is mainly considered for closed balls of V^*, rather than the whole dual space V^*, because for such subsets, pointwise convergence of moments is equivalent to weak-* convergence—that is, the second and third columns of Table 2.2 are equivalent. This equivalence is formally stated in the following theorem.

Theorem 2.6.2. *Suppose that E is a densely spanning subset of V. Let (φ_α) be a net in V^* satisfying $\sup_\alpha \|\varphi_\alpha\| < +\infty$. Then (φ_α) converges weak-* in V^* iff the limit $\lim_\alpha \langle \varphi_\alpha, v\rangle$ exists for any $v \in E$.*

Moreover, if $\varphi \in V^*$ satisfies that

$$\lim_\alpha \langle \varphi_\alpha, v\rangle = \langle \varphi, v\rangle \qquad \text{for any } v \in E$$

then (φ_α) converges weak- to φ.*

Proof. This is clear from Prop. 2.4.5 and Thm. 2.4.6. □

Remark 2.6.3. Let U be a dense linear subspace of V. (For example, take $V = C_0(X,\mathbb{F})$ and $U = C_c(X,\mathbb{F})$.) Recall the canonical isomorphism $V^* \simeq U^*$ given in Cor. 2.4.3. Then by Prop. 2.6.2, for each $R \in \mathbb{R}_{\geqslant 0}$, the weak-* topology on $\overline{B}_{V^*}(0,R)$ agrees with the weak-* topology on $\overline{B}_{U^*}(0,R)$. However, the weak-* topology on V^* is in general not equal to the weak-* topology on U^*.

In Prop. 2.6.2, one might further ask whether a net (φ_α) in $\overline{B}_{V^*}(0,R)$ that converges weak-* has its limit also in $\overline{B}_{V^*}(0,R)$. The answer is yes:

Proposition 2.6.4 (Fatou's lemma for weak-* convergence). *Let (φ_α) be a net in V^* converging weak-* to some $\varphi \in V^*$. Then*

$$\|\varphi\| \leqslant \liminf_\alpha \|\varphi_\alpha\| \tag{2.30}$$

In other words, the norm function $\|\cdot\| : V^* \to \mathbb{R}_{\geqslant 0}$ is lower semicontinuous with respect to the weak- topology on V^*.*

In contrast, if (φ_α) converges in the operator norm to φ, then $\|\varphi\| = \lim_\alpha \|\varphi_\alpha\|$. Cf. Rem. 2.3.3.

Proof. For each $v \in \overline{B}_V(0,1)$, we have

$$|\langle\varphi,v\rangle| = \lim_\alpha |\langle\varphi_\alpha,v\rangle| = \liminf_\alpha |\langle\varphi_\alpha,v\rangle| \leqslant \liminf_\alpha \|\varphi_\alpha\|\cdot\|v\| \leqslant \liminf_\alpha \|\varphi_\alpha\|$$

Applying $\sup_{v\in\overline{B}_V(0,1)}$ to the LHS above yields (2.30). (See also Thm. 2.4.6.) □

Theorem 2.6.5 (Banach-Alaoglu theorem). *$\overline{B}_{V^*}(0,1)$ is **weak-* compact**—that is, it is compact in the weak-* topology.*

Thus, $\overline{B}_{V^*}(0,1)$ is a compact Hausdorff space.

First proof. Let (φ_α) be a net $\overline{B}_{V^*}(0,1)$. Since $|\langle\varphi_\alpha,v\rangle| \leqslant \|v\|$ for each $v \in V$, we can view (φ_α) as a net in

$$S = \prod_{v\in V} \overline{B}_{\mathbb{F}}(0,\|v\|)$$

By Tychonoff's Thm. 1.4.17, S is compact. Therefore, (φ_α) has a subnet (φ_{α_μ}) converging pointwise on V to some function $\varphi : V \to \mathbb{F}$. The function φ is clearly linear and satisfies $\|\varphi\| \leqslant \sup_\mu \|\varphi_{\alpha_\mu}\| \leqslant 1$, cf. Thm. 2.4.6. Thus (φ_{α_μ}) converges weak-* to $\varphi \in \overline{B}_{V^*}(0,1)$. This finishes the proof that $\overline{B}_{V^*}(0,1)$ is compact. □

The above proof relies on Tychonoff's theorem, which in turn relies on Zorn's lemma. When V is separable, one can prove the Banach-Alaoglu theorem without using Zorn's lemma:

Second proof assuming that V is separable. Let E be a densely-spanning subset of V. Then

$$\Phi : \overline{B}_{V^*}(0,1) \to \mathbb{F}^E \qquad \varphi \mapsto \varphi|_E$$

is injective. Moreover, if (φ_α) is a net in $\overline{B}_{V^*}(0,1)$ and $\varphi \in \overline{B}_{V^*}(0,1)$, then Prop. 2.4.5 indicates that (φ_α) converges weak-* to φ iff (φ_α) converges pointwise on E to φ. Therefore, Φ restricts to a homeomorphism from $\overline{B}_{V^*(0,1)}$ to its image. Thus, since $\mathbb{F}^E$ is second countable (cf. Prop. 1.4.16), so is any subset—in particular, $\overline{B}_{V^*}(0,1)$.

Therefore, by Thm. 1.3.7, showing that $\overline{B}_{V^*}(0,1)$ is compact is equivalent to showing that it is sequentially compact. Let (φ_n) be a sequence in $\overline{B}_{V^*}(0,1)$. By the diagonal method (cf. Rem. 1.4.18), (φ_n) has a subsequence (φ_{n_k}) converging pointwise on E. Thm. 2.4.6 now implies that (φ_{n_k}) converges weak-* to some $\varphi \in \overline{B}_{V^*}(0,1)$. □

The above proof shows that when V is separable, then $\overline{B}_{V^*}(0,1)$ is second-countable and therefore sequentially compact in the weak-* topology. We record this result below.

Proposition 2.6.6. *Suppose that the normed vector space V is separable. Then, when equipped with the weak-* topology, the compact Hausdorff space $\overline{B}_{V^*}(0,1)$ is second countable (equivalently, metrizable; cf. Thm. 1.5.14).*

2.7 Weak-* convergence in L^p-spaces

Let $(X, \mathfrak{M}, \mu)$ be a σ-finite measure space.[7] Let $1 < p \leqslant +\infty$ and $p^{-1} + q^{-1} = 1$.

We identify $L^p(X, \mu, \mathbb{F})$ with the dual space $L^q(X, \mu, \mathbb{F})^*$ via the isomorphism described in Thm. 1.6.16. This defines the **weak-* topology on $L^p(X, \mu, \mathbb{F})$** . In particular, a net (f_α) in $L^p(X, \mu, \mathbb{F})$ converges weak-* to $f \in L^p(X, \mu, \mathbb{F})$ iff

$$\lim_\alpha \int_X f_\alpha g d\mu = \int_X f g d\mu \qquad \text{for all } g \in L^q(X, \mu, \mathbb{F})$$

2.7.1 Pointwise convergence and weak-* convergence

Let us prove Thm. 2.2.8 in a slightly more general setting. Note that a finite Borel measure μ on an interval $I \subset \mathbb{R}$ can be extended by zero to a finite Borel measure on $\mathbb{R}$, which is Radon by Thm. 1.7.8. Therefore, to generalize Thm. 2.2.8, it suffices to consider finite Borel (equivalently, finite Radon) measures on $\mathbb{R}$.

[7]The condition on σ-finiteness can be removed at least when $p = 2$. See the paragraph after Thm. 1.6.16.

Theorem 2.7.1. *Let μ be a finite Borel measure on $\mathbb{R}$. Let (f_α) be a net in $L^p(\mathbb{R},\mu,\mathbb{F})$ satisfying $\sup_\alpha \|f_\alpha\|_{L^p} < +\infty$. Then (f_α) converges weak-* to some element of $L^p(\mathbb{R},\mu,\mathbb{F})$ iff the following limit exists for every $x \in \mathbb{R}$:*

$$F(x) := \lim_\alpha \int_{(-\infty,x]} f_\alpha d\mu \tag{2.31}$$

Moreover, if $f \in L^p(\mathbb{R},\mu,\mathbb{F})$, then (f_α) converges weak- to f iff for each $x \in \mathbb{R}$ we have*

$$F(x) = \int_{(-\infty,x]} f d\mu \tag{2.32}$$

Note that since μ is finite, the constant function 1 belongs to L^q. Therefore, by Hölder's inequality, any function in $L^p(\mathbb{R},\mu,\mathbb{F})$ is integrable.

Proof. First, assume that (f_α) converges weak-* to f in $L^p(\mathbb{R},\mu,\mathbb{F})$. Then for each $x \in \mathbb{R}$, we have $\lim_\alpha \int f_\alpha \chi_{(-\infty,x]} d\mu = \int f \chi_{(-\infty,x]} d\mu$. This proves that (2.31) exists and (2.32) holds.

Next, we assume that (2.31) exists for every x. In the following, we give two proofs that (f_α) converges weak-* to some $f \in L^p(\mathbb{R},\mu,\mathbb{F})$. Then (2.32) will follow from the first paragraph.

First proof. Let $\varphi_\alpha \in L^q(\mathbb{R},\mu,\mathbb{F})^*$ be the linear functional associated to f_α, i.e., $\langle \varphi_\alpha, g\rangle = \int f_\alpha g d\mu$ for each $g \in L^q$. By assumption, φ_α converges when evaluated with any member of

$$\mathcal{E} = \mathrm{Span}_{\mathbb{F}}\{\chi_{(-\infty,x]} : x \in \mathbb{R}\}$$

By Thm. 1.7.13, $\mathcal{E}$ is dense in L^q. Therefore, since

$$\sup_\alpha \|\varphi_\alpha\| = \sup_\alpha \|f_\alpha\|_p < +\infty$$

by Thm. 2.6.2, (φ_α) converges weak-* to some $\varphi \in (L^q)^*$. By Thm. 1.6.16, φ is represented by some $f \in L^p(\mathbb{R},\mu,\mathbb{F})$. Thus (f_α) converges weak-* to f.

Second proof. In this proof, we use the fact that any bounded closed ball of $L^p(\mathbb{R},\mu,\mathbb{F})$ is weak-* compact, which is due to Thm. 1.6.16 and the Banach-Alaoglu theorem.

Since $\sup_\alpha \|f_\alpha\|_p < +\infty$, the net (f_α) has a subnet (f_{α_ν}) converging weak-* to some $f \in L^p$. By the first paragraph, for each $x \in \mathbb{R}$ we have

$$\lim_\nu \int_{(-\infty,x]} f_{\alpha_\nu} d\mu = \int_{(-\infty,x]} f d\mu$$

Since (2.31) converges, we conclude

$$\lim_\alpha \int_{(-\infty,x]} f_\alpha d\mu = \int_{(-\infty,x]} f d\mu$$

That is, if we let $\varphi_\alpha \in (L^q)^*$ represent f_α and let $\varphi \in (L^q)^*$ represent f, then (φ_α) converges to φ when evaluated on $\mathcal{E}$. By Thm. 1.7.13, $\mathcal{E}$ is dense in L^q. Therefore, by Thm. 2.6.2, (φ_α) converges weak-* to φ. That is, (f_α) converges weak-* to f. □

We now present another connection between pointwise convergence and weak-* convergence.

Theorem 2.7.2. *Let (f_n) be a sequence in $L^p(X, \mu, \mathbb{F})$ satisfying $\sup_n \|f_n\|_{L^p} < +\infty$. Suppose that (f_n) converges pointwise to f. Then $f \in L^p(X, \mu, \mathbb{F})$, and (f_n) converges weak-* to f.*

Proof. By Fatou's lemma, we have $f \in L^p$, since

$$\int |f|^p \leqslant \liminf_n \int |f_n|^p < +\infty$$

Thm. 2.7.1 suggests that when $X = \mathbb{R}$ and μ is a finite Borel measure, to prove that (f_n) converges weak-* to f, it suffices to verify that $\lim_n \int_{(-\infty,x]} f_n = \int_{(-\infty,x]} f$ for each $x \in \mathbb{R}$. Motivated by this, we claim that in the general case, it suffices to prove

$$\lim_n \int_E f_n d\mu = \int_E f d\mu \tag{2.33}$$

for each $E \in \mathfrak{M}$ satisfying $\mu(E) < +\infty$. (Note that any L^p function is integrable in E by Hölder's inequality.) Indeed, suppose (2.33) is true. Then, by the density of integrable simple functions in L^p (Thm. 1.6.13), and by Thm. 2.6.2, the sequence (f_n) converges weak-* to f.

Let us prove (2.33). Since (f_n) converges pointwise to f, and since $\mu(E) < +\infty$, it follows that (f_n) converges in measure to f on E. That is, for each $\varepsilon > 0$,

$$\lim_n \mu(A_n) = 0 \qquad \text{where } A_n = \{x \in E : |f(x) - f_n(x)| \geqslant \varepsilon\}$$

(Proof: Let $g_n = |f - f_n|$. Then (g_n) converges pointwise to 0. Let $h_n(x) = \sup_{k \geqslant n} g_n(x)$. Then (h_n) is decreasing, and $\lim_n h_n(x) = \limsup_n g_n(x) = 0$. Thus, the sequence of sets (B_n) defined by $B_n = \{x \in E : h_n(x) \geqslant \varepsilon\}$ is decreasing and $\bigcap B_n = \varnothing$. Hence $\lim_n \mu(B_n) = 0$. Since $g_n \leqslant h_n$, we have $\mu(A_n) \leqslant \mu(B_n)$, and hence $\lim_n \mu(A_n) = 0$.)

For each $\varepsilon > 0$ and A_n as above, we have

$$\int_{E \backslash A_n} |f_n - f| d\mu \leqslant \varepsilon \mu(E)$$

Let $C = \sup_n \|f_n\|_{L^p}$. By Hölder's inequality, we have

$$\limsup_n \int_{A_n} |f_n - f| d\mu \leqslant \limsup_n \left(\|f_n - f\|_{L^p} \cdot \mu(A_n)^{\frac{1}{q}} \right) \leqslant 2C \limsup_n \mu(A_n)^{\frac{1}{q}} = 0$$

where the last equality is due to $\lim_n \mu(A_n) = 0$. Thus

$$\limsup_n \left| \int_E (f_n - f) d\mu \right| \leqslant \varepsilon \mu(E)$$

Since ε is arbitrary, we conclude (2.33). $\square$

2.7.2 Weak-* approximation by elementary functions

Let X be an LCH space, and let μ be a Radon measure (or its completion) on X with σ-algebra $\mathfrak{M}$. We assume that μ is σ-finite. This condition holds, for example, when X is σ-compact (in particular, when X is second countable; cf. Rem. 1.4.34.)

In this subsection, we examine Principle 2.2.10 in the context of L^p-spaces. We begin with the following observation:

Remark 2.7.3. Let V be a normed vector space, and let U be a linear subspace of V^*. Let $R \in \mathbb{R}_{>0}$. By Rem. 2.4.1, U is norm-dense in V^* iff $\overline{B}_U(0, R)$ is norm-dense in $\overline{B}_{V^*}(0, R)$.

It is clear from linearity that if $\overline{B}_U(0, R)$ is weak-* dense in $\overline{B}_{V^*}(0, R)$, then U is weak-* dense in V^*. However, the weak-* density of U in V^* does not imply the weak-* density of $\overline{B}_U(0, R)$ in $\overline{B}_{V^*}(0, R)$. Therefore, when studying weak-* approximation in V^*, we aim—when possible—to approximate any $\varphi \in V^*$ by a net (φ_α) in U such that $\|\varphi_\alpha\| \leqslant \|\varphi\|$. This ensures not only convergence but also control of norms. $\square$

Theorem 2.7.4. *The closed unit ball of $C_c(X, \mathbb{F})$ is weak-* dense in the closed unit ball of $L^p(X, \mu, \mathbb{F})$. More precisely, the obvious map $C_c(X, \mathbb{F}) \to L^p(X, \mu, \mathbb{F})$ sends $\overline{B}_{C_c(X,\mathbb{F})}(0, 1)$ to a weak-* dense subset of $\overline{B}_{L^p(X,\mu,\mathbb{F})}(0, 1)$.*

Proof. By Thm. 1.7.10, if $p < +\infty$, then $\overline{B}_{C_c(X,\mathbb{F})}(0, 1)$ is norm-dense in $\overline{B}_{L^p(X,\mu,\mathbb{F})}(0, 1)$, and hence also weak-* dense.

Now, we assume $p = +\infty$. let $\mathscr{I}$ be the directed set

$$\mathscr{I} = \{(\mathcal{G}, \varepsilon) : \mathcal{G} \in \mathrm{fin}(2^{C_c(X,\mathbb{F})}), \varepsilon \in \mathbb{R}_{>0}\}$$
$$(\mathcal{G}_1, \varepsilon_1) \leqslant (\mathcal{G}_2, \varepsilon_2) \qquad \text{means} \qquad \mathcal{G}_1 \subset \mathcal{G}_2, \varepsilon_1 \geqslant \varepsilon_2$$

Fix any $f \in \overline{B}_{L^\infty(X,\mu,\mathbb{F})}(0, 1)$. By adding a μ-a.e. zero function to f, we assume that $\|f\|_{l^\infty(X)} = \|f\|_{L^\infty(X,\mu,\mathbb{F})} \leqslant 1$. We claim that for any $(\mathcal{G}, \varepsilon) \in \mathscr{I}$, there exists $f_{\mathcal{G},\varepsilon} \in \overline{B}_{C_c(X,\mathbb{F})}(0, 1)$ such that

$$\left| \int_X (f - f_{\mathcal{G},\varepsilon}) g d\mu \right| \leqslant \varepsilon \qquad \text{for all } g \in \mathcal{G}$$

If this is true, then $(f_{\mathcal{G},\varepsilon})_{(\mathcal{G},\varepsilon)\in\mathscr{I}}$ converges to f when integrated against any element of $C_c(X, \mathbb{F})$. Since $C_c(X, \mathbb{F})$ is dense in $L^1(X, \mu, \mathbb{F})$ (Thm. 1.7.10), it follows from Thm. 2.6.2 that $(f_{\mathcal{G},\varepsilon})_{(\mathcal{G},\varepsilon)\in\mathscr{I}}$ converges weak-* to f, finishing the proof.

Let us prove the claim. We write $\mathcal{G} = \{g_1, \dots, g_n\}$. Let $A_i = \mathrm{Supp}(g_i)$ and $A = A_1 \cup \cdots \cup A_n$. Since A is compact, we have $\mu(A) < +\infty$. Let $M = \|g_1\|_\infty + \cdots + \|g_n\|_\infty$. By Lusin's Thm. 1.7.9 and the Tietze extension Thm. 1.4.29, there exist a compact set $K \subset A$ and a function $f_{\mathcal{G},\varepsilon} \in C_c(X, \mathbb{F})$ satisfying

$$f_{\mathcal{G},\varepsilon}|_K = f|_K \qquad \|f_{\mathcal{G},\varepsilon}\|_{l^\infty} = \|f\|_{l^\infty} \qquad \mu(A\backslash K) \leqslant \varepsilon/2M$$

Recall that $\|f\|_{l^\infty} \leqslant 1$. Thus, for each $1 \leqslant i \leqslant n$, we have

$$\begin{aligned} &\Big|\int_X (f - f_{\mathcal{G},\varepsilon})g_i\Big| = \Big|\int_{A\backslash K} (f - f_{\mathcal{G},\varepsilon})g_i\Big| \leqslant M \int_{A\backslash K} (|f| + |f_{\mathcal{G},\varepsilon}|) \\ \leqslant &2M \cdot \mu(A\backslash K) \leqslant \varepsilon \end{aligned}$$

□

Corollary 2.7.5. *Let $\varphi \in L^1(X, \mu, \mathbb{F})$. Then the linear functional*

$$\Lambda : C_0(X, \mathbb{F}) \to \mathbb{F} \qquad f \mapsto \int_X f\varphi d\mu$$

satisfies $\|\Lambda\| = \int_X |\varphi| d\mu$.

Proof. Clearly $\|\Lambda\| \leqslant \|\varphi\|_{L^1}$. To prove the reverse inequality, assume WLOG that $\|\varphi\|_{L^1} > 0$. Define $g = \overline{\varphi}/|\varphi|$, understood to be zero wherever $\varphi = 0$. Then $\int g\varphi d\mu = \|\varphi\|_{L^1}$. Since $g \in \overline{B}_{L^\infty(X,\mu,\mathbb{F})}(0, 1)$, by Thm. 2.7.4, for each $0 < \gamma < 1$ there exists $f \in \overline{B}_{C_c(X,\mathbb{F})}(0, 1)$ such that $|\int f\varphi d\mu| \geqslant \gamma\|\varphi\|_{L^1}$; in particular, $f \neq 0$. Therefore $\|\Lambda\| \geqslant \gamma\|\varphi\|_{L^1}$. Since γ is arbitrary, we obtain $\|\Lambda\| \geqslant \|\varphi\|_{L^1}$. □

2.8 Weak-* convergence in l^p-spaces

Let X be a set, and let $1 \leqslant p \leqslant +\infty$ and $p^{-1} + q^{-1} = 1$. In this section, we prove the equivalence of the first two columns of Table 2.2 for $V = l^q(X, \mathbb{F})$, cf. Thm. 2.8.5. The most important case is when X is countable and $p = q = 2$. For example, $l^2(\mathbb{Z}^n)$ corresponds to the space of Fourier coefficients of L^2-functions on $\mathbb{T}^n := (\mathbb{S}^1)^n$.

Recall from the text near (1.2) that X is equipped with the discrete topology 2^X. Therefore, the support of each $f : X \to \mathbb{F}$ is $\mathrm{Supp}(f) = \{x \in X : f(x) \neq 0\}$.

2.8.1 The linear isometry $l^p(X, \mathbb{F}) \to l^q(X, \mathbb{F})^*$

Proposition 2.8.1. *Assume that $1 \leqslant p < +\infty$. Then $C_c(X, \mathbb{F})$ is dense in $l^p(X, \mathbb{F})$, where*

$$C_c(X, \mathbb{F}) := \{f \in \mathbb{F}^X : \mathrm{Supp}(f) \textit{ is a finite set}\} \tag{2.34}$$

The notation of $C_c(X,\mathbb{F})$ in (2.34) is compatible with our usual notation for LCH spaces if X is equipped with the discrete topology $\mathcal{T}_X = 2^X$.

Proof. Choose $f \in l^p(X,\mathbb{F})$. Then, since

$$\lim_{A\in\mathrm{fin}(2^X)} \sum_A |f|^p = \sum_X |f|^p$$

we have

$$\lim_{A\in\mathrm{fin}(2^X)} \|f - f\chi_A\|_{l^p}^p = \lim_{A\in\mathrm{fin}(2^X)} \sum_{X\backslash A} |f| = \sum_X |f|^p - \lim_{A\in\mathrm{fin}(2^X)} \sum_A |f|^p = 0$$

Thus, $(f\chi_A)_{A\in\mathrm{fin}(2^X)}$ is a net in $C_c(X,\mathbb{F})$ converging to f. □

Remark 2.8.2. We have a linear map

$$\begin{aligned} \Psi : l^p(X,\mathbb{F}) &\to l^q(X,\mathbb{F})^* \\ f &\mapsto \Big(g \in l^q(X,\mathbb{F}) \mapsto \sum_{x\in X} f(x)g(x)\Big) \end{aligned} \tag{2.35}$$

Indeed, by Hölder's inequality, for each $A \in \mathrm{fin}(2^X)$,

$$\Big|\sum_A fg\Big| \leqslant \sum_A |fg| \leqslant \|f\|_{l^p(A)} \cdot \|g\|_{l^q(X)} \leqslant \|f\|_{l^p(X)} \cdot \|g\|_{l^q(X)}$$

Applying $\lim_A$, we see that $\sum_X fg$ is absolutely convergent (i.e. $\sum_X |fg| < +\infty$), and

$$\Big|\sum_X fg\Big| \leqslant \sum_X |fg| \leqslant \|f\|_{l^p(X)} \cdot \|g\|_{l^q(X)}$$

This justifies the claim that Ψ has range in $l^q(X,\mathbb{F})^*$ (rather than just in $\mathrm{Lin}(l^q(X,\mathbb{F}),\mathbb{F})$), and that $\|\Psi\| \leqslant 1$.

Proposition 2.8.3. *The map Ψ in (2.35) is a linear isometry.*

Proof. We already know $\|\Psi\| \leqslant 1$, and we want to show $\|\Psi\| = 1$.

Case $p < +\infty$: By Prop. 2.8.1 and Thm. 2.4.2, we have $\|\Psi\| = \|\Psi|_{C_c(X,\mathbb{F})}\|$. Therefore, it suffices to show that $\|\Psi(f)\| = \|f\|$ for each $f \in C_c(X,\mathbb{F})$. We assume WLOG that $f \neq 0$. Then

$$\langle \Psi(f), g\rangle = \|f\|_{l^p} \cdot \|g\|_{l^q}$$

if we write $f = u|f|$ (where $u : X \to \mathbb{S}^1$) and let $g = \overline{u} \cdot |f|^{p-1}$. Since $\|\Psi(f)\| \cdot \|g\|_{l^q} \geqslant |\langle\Psi(f), g\rangle|$ and $\|g\|_{l^q} > 0$, we conclude that $\|\Psi(f)\| \geqslant \|f\|_{l^p}$, and hence $\|\Psi(f)\| = \|f\|_{l^p}$.

Case $p = +\infty$: For each $0 \leqslant \lambda < 1$, let $x \in X$ such that $|f(x)| \geqslant \lambda\|f\|_{l^\infty}$. Take $g = \overline{u}\chi_{\{x\}}$ where $u \in \mathbb{S}^1$ is such that $f(x) = u|f(x)|$. Then

$$\langle \Psi(f), g\rangle = |f(x)| \geqslant \lambda\|f\|_{l^\infty} = \lambda\|f\|_{l^\infty} \cdot \|g\|_{l^1}$$

This, together with $\|\Psi(f)\| \cdot \|g\|_{l^1} \geqslant |\langle\Psi(f), g\rangle|$, implies $\|\Psi(f)\| \geqslant \lambda\|f\|_{l^\infty}$. Since λ is arbitrary, we conclude $\|\Psi(f)\| = \|f\|_{l^\infty}$. □

2.8.2 Weak-* convergence in $l^p(X, \mathbb{F})$

Definition 2.8.4. Assume that $1 < p \leqslant +\infty$. The **weak-* topology on $l^p(X, \mathbb{F})$** is defined to be the pullback topology via the (injective) map $\Phi : l^p(X, \mathbb{F}) \to l^q(X, \mathbb{F})^*$ of the weak-* topology of $l^q(X, \mathbb{F})^*$. In other words, a net (f_α) in $l^p(X, \mathbb{F})$ converges weak-* to $f \in l^p(X, \mathbb{F})$ iff for each $g \in l^q(X, \mathbb{F})$ we have

$$\lim_\alpha \sum_X f_\alpha g = \sum_X fg \tag{2.36}$$

Theorem 2.8.5. *Assume $1 < p \leqslant +\infty$. Let (f_α) be a net in $L^p(X, \mathbb{F})$ satisfying $\sup_\alpha \|f_\alpha\|_{l^p} < +\infty$. Then (f_α) converges weak-* to some element of $l^p(X, \mathbb{F})$ iff $\lim_\alpha f_\alpha(x)$ converges for each $x \in X$.*

Moreover, if $f \in l^p(X, \mathbb{F})$, then (f_α) converges weak- to f iff $f(x) = \lim_\alpha f_\alpha(x)$ for each $x \in X$.*

Consequently, if $p > 1$ and (f_α) is a uniformly l^p-bounded net in $L^p(X, \mathbb{F})$ converging pointwise to $f : X \to \mathbb{F}$, then $f \in l^p(X, \mathbb{F})$. (Indeed, by Thm. 2.8.5, (f_α) converges weak-* to some $\widetilde{f} \in l^p(X, \mathbb{F})$, and $\widetilde{f}$ is the pointwise limit of (f_α). Therefore $f = \widetilde{f}$ belongs to $l^p(X, \mathbb{F})$.)

However, as we will see below, this conclusion must in fact be established first in order to complete the proof of Thm. 2.8.5

Proof. First, assume that (f_α) converges weak-* to $f \in l^p(X, \mathbb{F})$. Applying (2.36) to $g = \chi_{\{x\}}$ (for each $x \in X$), we see that (f_α) converges pointwise to f.

Conversely, assume that (f_α) converges pointwise on X. Let $f \in \mathbb{F}^X$ be the pointwise limit of (f_α). Recall that $C = \sup_\alpha \|f_\alpha\|_{l^p}$ is finite. We claim that $f \in l^p(X, \mathbb{F})$. Indeed, if $p = +\infty$, then for each $x \in X$, we have

$$|f(x)| = \lim_\alpha |f_\alpha(x)| \leqslant \sup_\alpha \|f_\alpha\|_{l^\infty} < +\infty$$

If $p < +\infty$, then for each $A \in \mathrm{fin}(2^X)$,

$$\sum_A |f|^p = \lim_\alpha \sum_A |f_\alpha|^p \leqslant \sup_\alpha \|f_\alpha\|_{l^p}^p \leqslant C^p$$

Applying $\lim_A$, we see that $\sum_X |f|^p \leqslant C^p$, and hence $f \in l^p(X, \mathbb{F})$.

Let Ψ be as in (2.35). By Prop. 2.8.1, $C_c(X, \mathbb{F})$ is dense in $L^q(X, \mathbb{F})$. Therefore, to show that (f_α) converges weak-* to f, by Thm. 2.6.2 and the observation that

$$\sup_\alpha \|\Psi(f_\alpha)\| = \sup_\alpha \|f_\alpha\|_{l^p} < +\infty$$

it suffices to show that $\langle \Psi(f_\alpha), g \rangle$ converges to $\langle \Psi(f), g \rangle$ (that is, $\sum f_\alpha g$ converges to $\sum fg$) for each $g \in C_c(X, \mathbb{F})$. But this follows from the fact that (f_α) converges pointwise to f. $\square$

As an application of Thm. 2.8.5, we prove a variant of Prop. 2.8.1.

Proposition 2.8.6. *Let $1 < p \leqslant +\infty$. Then $\overline{B}_{C_c(X,\mathbb{F})}(0,1)$ is weak-* dense in $\overline{B}_{l^\infty(X,\mathbb{F})}$.*

Proof. Let $f \in \overline{B}_{l^\infty(X,\mathbb{F})}$. Then $(f\chi_A)_{A\in\mathrm{fin}(2^X)}$ is a net in $\overline{B}_{C_c(X,\mathbb{F})}(0,1)$ converging pointwise to f. By Thm. 2.8.5, this net converges weak-* to f. □

2.8.3 The isomorphism $l^p(X,\mathbb{F}) \simeq l^q(X,\mathbb{F})^*$

Now that the equivalence of the first two columns of Table 2.2 for $V = l^q(X,\mathbb{F})$ has been established in Thm. 2.8.5 for $p > 1$, we can prove the isomorphism $l^p(X,\mathbb{F}) \simeq l^q(X,\mathbb{F})^*$. Of course, at least when X is countable, this isomorphism is a special case of the duality $L^p(X,\mu,\mathbb{F}) \simeq L^q(X,\mu,\mathbb{F})^*$ from Thm. 1.6.16, by taking $\mu : 2^X \to [0,+\infty]$ to be the counting measure. However, there are good reasons to study the proof of $l^q(X,\mathbb{F})^* \simeq l^p(X,\mathbb{F})$ independently.

First, the proof of Thm. 1.6.16 is significantly more involved than the direct proof in the l^p setting. Whenever a result admits a simpler proof in a special case, it is worthwhile to examine that proof directly. Second, Thm. 1.6.16 depends crucially on the Radon-Nikodym Thm. 1.6.12, which in turn can be derived from the Riesz-Fréchet Thm. 3.5.3. The latter can be proved with the help of the isomorphism $l^2(X,\mathbb{F}) \simeq l^2(X,\mathbb{F})^*$.

Theorem 2.8.7. *Assume that $1 < p \leqslant +\infty$. Then the map $\Psi : l^p(X,\mathbb{F}) \to l^q(X,\mathbb{F})^*$ is an isomorphism of normed vector spaces.*

Proof. By Prop. 2.8.3, it remains to show that Ψ is surjective. Choose $\varphi \in l^q(X,\mathbb{F})^*$. We want to find $f \in l^p(X,\mathbb{F})$ such that $\Psi(f) = \varphi$.

Define $f : X \to \mathbb{F}$ by $f(x) = \varphi(\chi_{\{x\}})$. Then

$$\varphi(g) = \sum_X fg \tag{2.37}$$

holds whenever $g = \chi_{\{x\}}$ for some $x \in X$, and hence for all $g \in C_c(X,\mathbb{F})$. For each finite set $A \subset X$, due to the canonical linear isometry $l^p(A,\mathbb{F}) \to l^q(A,\mathbb{F})^*$ (as in (2.35)), we have

$$\|f|_A\|_{l^p} = \sup_{g\in l^q(A,\mathbb{F}),\|g\|_q\leqslant 1} \Big|\sum_A fg\Big| = \sup_{g\in l^q(A,\mathbb{F}),\|g\|_q\leqslant 1} |\varphi(g)| \leqslant \|\varphi\|$$

Hence, if $p = +\infty$, we clearly have $\|f\|_{l^\infty} \leqslant \|\varphi\|$; if $p < +\infty$, we have

$$\sum_X |f|^p = \lim_{A\in\mathrm{fin}(2^X)} \sum_A |f|^p = \lim_{A\in\mathrm{fin}(2^X)} \|f|_A\|_{l^p}^p \leqslant \|\varphi\|^p$$

and hence $\|f\|_{l^p} \leqslant \|\varphi\|$. In both cases, we have $f \in l^p(X,\mathbb{F})$.

By 4.2.16, the bounded linear functionals φ and $\Psi(f)$ agree on $C_c(X,\mathbb{F})$. Since $C_c(X,\mathbb{F})$ is dense in $l^q(X,\mathbb{F})$ (cf. Prop. 2.8.1), we must have $\varphi = \Psi(f)$. □

2.9 Weak-* convergence of distribution functions

In this section, we fix a proper interval $I \subset \mathbb{R}$, and let $a = \inf I, b = \sup I$. We use freely the notation in Subsec. 1.8.1. In particular, for each function ρ on I, we let

$$\Omega_\rho = \{x \in (a,b) : \rho|_{(a,b)} \text{ is continuous at } x\}$$

A family of functions (ρ_α) from I to $\mathbb{R}$ is called **uniformly bounded** if $\sup_\alpha \|\rho_\alpha\|_{l^\infty(I,\mathbb{R})} < +\infty$.

The goal of this section is to prove Thm. 2.2.9, which characterizes the relationship between pointwise convergence and weak-* convergence for increasing functions. To this end, we begin with several preparatory results concerning the pointwise convergence of such functions.

2.9.1 Almost convergence of increasing functions

Lemma 2.9.1. *Let (ρ_α) be a uniformly bounded net of increasing functions $I \to \mathbb{R}_{\geqslant 0}$. Suppose that (ρ_α) converges pointwise on a dense subset $E \subset I$. Then there exists a bounded increasing function $\rho : I \to \mathbb{R}_{\geqslant 0}$ such that (ρ_α) converges pointwise on E to ρ.*

Proof. Let $\varrho : E \to \mathbb{R}_{\geqslant 0}$ be the pointwise limit of (ρ_α), which is clearly bounded and increasing. Extend $\varrho|_{E\cap I_{<b}}$ to a function $\rho : I_{<b} \to \mathbb{R}_{\geqslant 0}$ by setting

$$\rho(x) = \lim_{E\ni y\to x^+} \varrho(y)$$

for any $x \in I\backslash E$. If $b \notin I$ then we are done. If $b \in E$, set $\rho(b) = \varrho(b)$. If $b \in I\backslash E$, set $\rho(b) = \lim_{x\to b^-} \rho(x)$. Then ρ is bounded and increasing, and (ρ_α) converges pointwise to ρ on E. □

Proposition 2.9.2. *Let (ρ_α) be a uniformly bounded net of increasing functions $I \to \mathbb{R}_{\geqslant 0}$. Let $\rho : I \to \mathbb{R}_{\geqslant 0}$ be increasing. Then the following are equivalent:*

(a) There exists a dense subset $E \subset (a,b)$ such that (ρ_α) converges pointwise on E to ρ.

(b) The net (ρ_α) converges pointwise on Ω_ρ to ρ.

If either of these two statements are true, we say that (ρ_α) ***almost converges*** *to ρ.*

Proof. Since Ω_ρ is dense (Prop. 1.8.1), clearly (b) implies (a).

Now assume (a). Choose any $x \in \Omega_\rho$. We will show that every convergent subnet $(\rho_{\alpha_\nu}(x))$ of $(\rho_\alpha(x))$ converges to $\rho(x)$. Then by Thm. 1.3.9, we have $\lim_\alpha \rho_\alpha(x) = \rho(x)$, proving (b).

By Lem. 2.9.1, there exists an increasing function $\widetilde{\rho} : I \to \mathbb{R}_{\geqslant 0}$ such that (ρ_{α_ν}) converges on $E \cup \{x\}$ to $\widetilde{\rho}$. Since (ρ_{α_ν}) converges pointwise on E to ρ, the functions ρ and $\widetilde{\rho}$ agree on E. Namely, ρ and $\widetilde{\rho}$ are almost equal. Therefore, by Prop. 1.8.3, ρ and $\widetilde{\rho}$ agree on Ω_ρ, and in particular at x. This proves $\lim_\nu \rho_{\alpha_\nu}(x) = \rho(x)$. □

The following theorem can be viewed as a concrete manifestation of the Banach-Alaoglu Thm. 2.6.5 in the setting of $C_c(I)^*$. It will be used in the proof of Thm. 2.9.6.

Theorem 2.9.3 (Helly selection theorem). *Let (ρ_α) be a uniformly bounded net (resp. sequence) of increasing functions $I \to \mathbb{R}_{\geqslant 0}$. Then (ρ_α) admits a pointwise convergent subnet (resp. subsequence).*

Proof. The existence of a pointwise convergent subnet follows directly from the Tychonoff Thm. 1.4.17. Therefore, let us assume that (ρ_α) is a sequence (ρ_n). Let $E = I \cap \mathbb{Q}$. Then, by the diagonal method (cf. Rem. 1.4.18), (ρ_n) has a subsequence (ρ_{n_k}) converging pointwise on E. By Lem. 2.9.1, there exists a bounded increasing $\rho : I \to \mathbb{R}_{\geqslant 0}$ such that (ρ_{n_k}) converges pointwise on E to ρ. Therefore, by Prop. 2.9.2, (ρ_{n_k}) converges pointwise on Ω_ρ to ρ. Since $I \backslash \Omega_\rho$ is countable, by the diagonal method again, (ρ_{n_k}) has a subsequence converging pointwise on $I \backslash \Omega_\rho$, and hence on I. □

2.9.2 Almost convergence and weak-* convergence

Definition 2.9.4. Let (ρ_α) be a net in $BV(I, \mathbb{F})$. Let $\rho \in BV(I, \mathbb{F})$. Let Λ_α and Λ be the elements of $C_c(I, \mathbb{F})^*$ corresponding to ρ_α and ρ, respectively, via the Riesz representation Thm. 1.10.4. We say that the net $(d\rho_\alpha)$ **converges weak-*** to $d\rho$ if (Λ_α) converges weak-* to Λ. Namely, for each $f \in C_c(I, \mathbb{F})$, we have

$$\lim_\alpha \int_I f d\rho_\alpha = \int_I f d\rho \tag{2.38}$$

Remark 2.9.5. Suppose that (ρ_α) is a uniformly bounded net of increasing functions $I \to \mathbb{R}_{\geqslant 0}$. Let $\rho : I \to \mathbb{R}_{\geqslant 0}$ be increasing. Recall from Rem. 1.4.32 that $C_c(I, \mathbb{F})$ is dense in $C_0(I, \mathbb{F})$. Therefore, by Thm. 2.6.2, the net $(d\rho_\alpha)$ converges weak-* to $d\rho$ iff (2.38) holds for any $f \in C_0(I, \mathbb{F})$.

The following Thm. 2.9.6 is parallel to Thm. 2.7.1. However, unlike Thm. 2.7.1 whose proof relies on the isomorphism $L^p \simeq (L^q)^*$, Thm. 2.9.6 does not rely on the Riesz representation theorem. (Note that when the measure in Thm. 2.7.1 is the Lebesgue measure, one can also prove Thm. 2.7.1 from the fundamental theorem of calculus, without invoking the duality $L^p \simeq (L^q)^*$; see the proof of Thm. 2.2.8.)

Theorem 2.9.6. *Let $(\rho_\alpha)_{\alpha \in \mathscr{A}}$ be a uniformly bounded net of bounded increasing functions $I \to \mathbb{R}_{\geqslant 0}$. Let $\rho : I \to \mathbb{R}_{\geqslant 0}$ be bounded and increasing. Then the following are equivalent:*

(a) *There exists a bounded family $(\varkappa_\alpha)_{\alpha \in \mathscr{A}}$ in $\mathbb{R}$ (assumed to be zero if $a \in I$) satisfying the following conditions:*

- *$(\rho_\alpha + \varkappa_\alpha)$ almost converges to ρ.*

- $\lim_\alpha(\rho_\alpha(b)+\varkappa_\alpha)=\rho(b)$ *if* $b\in I$.

(b) The net $(d\rho_\alpha)$ *converges weak-* to* $d\rho$.

The boundedness of $(\varkappa_\alpha)_{\alpha\in\mathscr{A}}$ means that $\sup_\alpha|\varkappa_\alpha|<+\infty$.

Proof. (a)⇒(b): Assume (a). We verify (2.38) for each $f\in C_c(I,\mathbb{F})$, which established (b). Recall from Rem. 1.9.13 that if $a\notin I$, adding constants to ρ_α and ρ does not affect the values of $\int_I fd\rho_\alpha$ and $\int_I fd\rho$.

Since (ρ_α) is uniformly bounded and $(\varkappa_\alpha)$ is bounded, there exists $c\geqslant 0$ such that $\rho_\alpha+\varkappa_\alpha+c\geqslant 0$ for all α. Therefore, replacing ρ_α with $\rho_\alpha+\varkappa_\alpha+c$ and ρ with $\rho+c$, we assume that there exists a dense subset $E\subset I$ such that (ρ_α) converges pointwise on E to ρ, and that $b\in E$ if $b\in I$.

Choose any $f\in C_c(I,\mathbb{F})$. Choose $u,v\in\mathbb{R}$ satisfying $\mathrm{Supp}_I(f)\subset[u,v]\subset I$, and let $J=[u,v]$. By increasing v if possible, we may assume that $v\in E$. (When $b\in I$, one simply choose $v=b$.)

In the case where $a\in I$, by Lem. 1.9.9, the values of $\int_J fd\rho_\alpha$ and $\int_J fd\rho$ remain unchanged if we change the values of $\rho_\alpha(a)$ and $\rho(a)$ to 0. Therefore, we may assume that $\rho_\alpha(a)=\rho(a)=0$ (so that a can be included to E), and we may also choose $u=a$. In the case where $a\notin I$, by the density of E, we can slightly decrease u so that $u\in E$. To summarize, whether a or b belongs to I or not, we can assume

$$u,v\in E$$

Since f is uniformly continuous, for each $\varepsilon>0$ there exists $\delta>0$ such that $|f(x)-f(y)|\leqslant\varepsilon$ for each $x,y\in I$ satisfying $|x-y|\leqslant\delta$. Choose a tagged partition

$$(\sigma,\xi_\bullet)=\big(\{a_0=u<a_1<\cdots<a_n=v\},(\xi_1,\dots,\xi_n)\big)$$

of J with mesh $<\delta$. Since E is dense, by a slight adjustment, we may assume that $a_0,a_1,\dots,a_n\in E$. This implies

$$\lim_\alpha\big(f(u)\rho_\alpha(u)+S_{\rho_\alpha}(f,\sigma,\xi_\bullet)\big)=f(u)\rho(u)+S_\rho(f,\sigma,\xi_\bullet)$$

Therefore, if we let $C=\sup\{\rho_\alpha(v)-\rho_\alpha(u),\rho(v)-\rho(u):\alpha\in\mathscr{A}\}$, then Rem. 1.9.5 implies

$$\limsup_\alpha\Big|\int_J fd\rho_\alpha-\int_J fd\rho\Big|\leqslant 2\varepsilon\cdot C$$

This finishes the proof of (2.38).

(b)⇒(a): Assume (b). We first consider the case where $a\notin I$. Fix $t\in\Omega_\rho$, and let

$$\varkappa_\alpha=\rho(t)-\rho_\alpha(t)$$

Then $(\varkappa_\alpha)$ is bounded. Therefore, $(\rho_\alpha+\varkappa_\alpha)$ is uniformly bounded, and hence there exists $c \geqslant 0$ such that $\rho_\alpha + \varkappa_\alpha + c \geqslant 0$ for all α. Replacing ρ_α with $\rho_\alpha + c$ and ρ with $\rho + c$, we assume that $\rho_\alpha + \varkappa_\alpha \geqslant 0$ for all α. (Of course, we still have $\rho \geqslant 0$.)

Choose any $x \in \Omega_\rho$. To show that $(\rho_\alpha(x) + \varkappa_\alpha)_\alpha$ converges to $\rho(x)$, by Thm. 1.3.9, it suffices to show that every convergent subnet $(\rho_\beta(x) + \varkappa_\beta)_\beta$ converges to $\rho(x)$.

By the Helly selection Thm. 2.9.3, the net of functions $(\rho_\beta + \varkappa_\beta)_\beta$ has a pointwise convergent subnet $(\rho_\gamma + \varkappa_\gamma)_\gamma$. Let $\widetilde{\rho} : I \to \overline{\mathbb{R}}_{\geqslant 0}$ be the pointwise limit of this subnet, which is clearly bounded and increasing. By (a)⇒(b), the net $(d(\rho_\gamma+\varkappa_\gamma))_\gamma$ converges weak-* to $d\widetilde{\rho}$. By assumption, it also converges weak-* to $d\rho$. Therefore, we have $\int_I f d\widetilde{\rho} = \int_I f d\rho$ for each $f \in C_c(I)$.

By Thm. 1.9.12 (and noting Rem. 1.9.13), we have

$$\widetilde{\rho} - \lim_{y\to a^+} \widetilde{\rho}(y) = \rho - \lim_{y\to a^+} \rho(y) \qquad \text{on } \Omega_\rho$$

In other words, there exists a constant $c \in \mathbb{R}$ such that

$$\widetilde{\rho} + c = \rho \qquad \text{on } \Omega_\rho \tag{2.39}$$

Since $\rho_\alpha(t) + \varkappa_\alpha = \rho(t)$ is constant over α, and since its subnet $(\rho_\gamma(t) + \varkappa_\gamma)_\gamma$ converges to $\widetilde{\rho}(t)$, we conclude $\widetilde{\rho}(t) = \rho(t)$. Therefore, since $t \in \Omega_\rho$, by (2.39), we have $c = 0$. Since $x \in \Omega_\rho$, by (2.39), we obtain $\widetilde{\rho}(x) = \rho(x)$. This proves that $(\rho_\gamma(x) + \varkappa_\gamma)_\gamma$ converges to $\rho(x)$, and hence $(\rho_\beta(x) + \varkappa_\beta)_\beta$ converges to $\rho(x)$.

Now consider the case where $a \in I$. We set $\varkappa_\alpha = 0$. Similar to the above argument, we choose any $x \in \Omega_\rho$, choose a subnet ρ_β converging at x, and further choose a subnet ρ_γ converging pointwise on I to $\widetilde{\rho} : I \to \overline{\mathbb{R}}_{\geqslant 0}$. By (a)⇒(b), we have $\int_I f d\widetilde{\rho} = \int_I f d\rho$ for each $f \in C_c(I)$. Consequently, Thm. 1.9.12 implies that $\widetilde{\rho} = \rho$ on Ω_ρ. Since $x \in \Omega_\rho$, we obtain again $\lim_\beta \rho_\beta(x) = \lim_\gamma \rho_\gamma(x) = \widetilde{\rho}(x) = \rho(x)$. Therefore $(\rho_\alpha(x))_\alpha$ converges to $\rho(x)$ for each $x \in \Omega_\rho$. □

Corollary 2.9.7. *Let $(\rho_\alpha)_{\alpha\in\mathscr{A}}$ be a uniformly bounded net of increasing functions $I \to \mathbb{R}_{\geqslant 0}$. Then the following are equivalent:*

(1) There exists a bounded family $(\varkappa_\alpha)_{\alpha\in\mathscr{A}}$ in $\mathbb{R}$ (assumed to be zero if $a \in I$) such that $(\rho_\alpha + \varkappa_\alpha)$ converges pointwise on a dense subset $E \subset I$, and also at b if $b \in I$.

(2) There exists a bounded increasing $\rho : I \to \mathbb{R}_{\geqslant 0}$ such that $(d\rho_\alpha)_{\alpha\in\mathscr{A}}$ converges weak- to $d\rho$.*

Proof. "(2)⇒(1)" follows immediately from Thm. 2.9.6. Conversely, assume (1). By Lem. 2.9.1, there exists a bounded increasing $\rho : I \to \mathbb{R}_{\geqslant 0}$ such that $(\rho_\alpha + \varkappa_\alpha)$ converges pointwise on $E \cup \{I \cap \{b\}\}$ to ρ. Then Thm. 2.9.6 implies (2). □

In Thm. 2.9.6, if $a \notin I$, one might suspect that the appearance of the family $(\varkappa_\alpha)_{\alpha\in\mathscr{A}}$ is needed only because (ρ_α) and ρ are not normalized, that is, because

$\lim_{x\to a^+}\rho_\alpha(x)$ and $\lim_{x\to a^+}\rho(x)$ don't vanish. However, even when both limits are zero, a nontrivial family $(\varkappa_\alpha)_{\alpha\in\mathscr{A}}$ may still be required, as the following example shows:

Example 2.9.8. Let $I=\mathbb{R}$ and $\rho_n=\chi_{[-n,+\infty)}$. Then $\lim_{x\to-\infty}\rho_n(x)=0$, and the sequence $(d\rho_n)_{n\in\mathbb{Z}_+}$ converges weak-* to 0. In this case, the family $(\varkappa_n)$ is chosen to be the constant sequence -1.

2.10 Weak-* approximation of Radon measures by Dirac measures

Fix an LCH space X. Recall that we have assumed throughout the notes that $\mathbb{F}\in\{\mathbb{R},\mathbb{C}\}$. Let

$$\begin{aligned}\mathcal{RM}(X,\overline{\mathbb{R}}_{\geqslant0})&=\{\text{Radon measures on }X\}\\ \mathcal{RM}(X,\mathbb{R}_{\geqslant0})&=\{\text{finite Radon measures on }X\}\\ \mathcal{RM}(X,\mathbb{R})&=\{\text{signed Radon measures on }X\}\\ \mathcal{RM}(X,\mathbb{C})&=\{\text{complex Radon measures on }X\}\end{aligned}\tag{2.40}$$

which are vectors spaces over $\overline{\mathbb{R}}_{\geqslant0},\mathbb{R}_{\geqslant0},\mathbb{R},\mathbb{C}$ respectively. Note the inclusion relation

$$\mathcal{RM}(X,\mathbb{R}_{\geqslant0})\subset\mathcal{RM}(X,\overline{\mathbb{R}}_{\geqslant0})\qquad\mathcal{RM}(X,\mathbb{R}_{\geqslant0})\subset\mathcal{RM}(X,\mathbb{R})\subset\mathcal{RM}(X,\mathbb{C})$$

Recall that for each $x\in X$, the Dirac measure at x is denoted by δ_x.

The goal of this section is to prove Principle 2.2.10 for $V=C_c(X,\mathbb{F})$. In this context, elementary functions are understood as linear combinations of Dirac measures. When X is an interval $I\subset\mathbb{R}$, these elementary functions correspond to bounded increasing functions $I\to\mathbb{R}_{\geqslant0}$ whose ranges are finite sets.

2.10.1 Definitions and basic properties

Definition 2.10.1. Recall the $\mathbb{F}$-linear isomorphism

$$\mathcal{RM}(X,\mathbb{F})\simeq C_c(X,\mathbb{F})^*$$

defined by the Riesz-Markov representation Thm. 1.7.17. The pullback of the operator norm on $C_c(X,\mathbb{F})^*$ to $\mu\in\mathcal{RM}(X,\mathbb{F})$ is called the **total variation** of μ, and is denoted by $\|\mu\|$. In other words,

$$\|\mu\|=\sup\Big\{\Big|\int fd\mu\Big|:f\in C_c(X,\mathbb{F}),|f|\leqslant1\Big\}$$

A family of complex Radon measures $(\mu_\alpha)_{\alpha\in\mathscr{A}}$ is called **uniformly bounded** if

$$\sup_{\alpha\in\mathscr{A}}\|\mu_\alpha\|<+\infty$$

The weak-* topology on $C_c(X,\mathbb{F})^*$ defines the **weak-* topology on $\mathcal{RM}(X,\mathbb{F})$**. Thus, if (μ_α) is a uniformly bounded net in $\mathcal{RM}(X,\mathbb{F})$, and if $\mu\in\mathcal{RM}(X,\mathbb{F})$, then (μ_α) converges weak-* to μ [8] iff for each $f\in C_c(X,\mathbb{F})$ we have

$$\lim_\alpha\int_X fd\mu_\alpha=\int_X fd\mu \tag{2.41}$$

By Rem. 1.4.32 and Thm. 2.6.2, (μ_α) converges weak-* to μ iff (2.41) holds for all $f\in C_0(X,\mathbb{F})$.

Example 2.10.2. By Thm. 1.7.6, if $\mu\in\mathcal{RM}(X,\mathbb{R}_{\geqslant 0})$, then

$$\|\mu\|=\mu(X)$$

Example 2.10.3. Let $E\subset X$ be a finite set, and let $c:E\to\mathbb{F}$ be a function. Then

$$\Big\|\sum_{x\in E}c(x)\delta_x\Big\|=\sum_{x\in E}|c(x)| \tag{2.42}$$

Proof. Let $\mu=\sum_{x\in E}c(x)\delta_x$. By Exp. 2.10.2, we have $\|\delta_x\|=1$. Since norms satisfy the sub-additivity, we have

$$\|\mu\|\leqslant\sum_{x\in E}|c(x)|\cdot\|\delta_x\|=\sum_{x\in E}|c(x)|$$

By Urysohn's lemma, there exists $f\in C_c(X,\mathbb{F})$ such that $\|f\|_{l^\infty}\leqslant 1$, and that for each $x\in E$, we have $|f(x)|=1$ and $f(x)c(x)=|c(x)|$. Then $\int_X fd\mu=\sum_{x\in E}|c(x)|$. This proves $\|\mu\|\geqslant\sum_{x\in E}|c(x)|$. □

Lemma 2.10.4. *Let $\mu\in\mathcal{RM}(X,\mathbb{F})$. Let $A_1,\dots,A_k$ be mutually disjoint Borel subsets of X. Then*

$$\|\mu\|\geqslant\sum_{j=1}^k|\mu(A_j)|$$

Proof. Since μ is a linear combination of finite Radon measures, there exists $\widehat{\mu}\in\mathcal{RM}(X,\mathbb{R}_{\geqslant 0})$ such that $|\int_X gd\mu|\leqslant\int_X|g|d\widehat{\mu}$ for each bounded Borel $g:X\to\mathbb{C}$. [9]

---

[8] We also say that $(d\mu_\alpha)$ converges weak-* to $d\mu$.

[9] For example, if μ is a finite sum $\sum_j\lambda_j\mu_j$ where $\lambda_j\in\mathbb{F}$ and $\mu_j\in\mathcal{RM}(X,\mathbb{R}_{\geqslant 0})$, set $\widehat{\mu}=\sum_j|\lambda_j|\mu_j$.

Since Radon measures are regular on Borel sets with finite measures (Thm. 1.7.7), for each $\varepsilon > 0$ there exists compact $K_j \subset A_j$ such that $\widehat{\mu}(A_j \backslash K_j) \leqslant \varepsilon$.

By Cor. 1.4.25, there exist mutually disjoint open subsets $U_1, \ldots, U_k \subset X$ such that $U_j \supset K_j$. Since $\widehat{\mu}$ is regular on K_j, we may assume that $\widehat{\mu}(U_j \backslash K_j) < \varepsilon$. By Urysohn's lemma, there exists $f_j \in C_c(U_j, \mathbb{F})$ such that $|f_j| \leqslant 1$, that $f_j|_{K_j}$ equals a constant $c_j \in \mathbb{F}$, and that $c_j \mu(K_j) = |\mu(K_j)|$. Let $f = f_1 + \cdots f_k$, which is an element of $C_c(X, \mathbb{F})$ satisfying $|f| \leqslant 1$. Then

$$\int_{\bigcup_j K_j} f d\mu = \sum_j |\mu(K_j)| \qquad \Big| \int_{X \backslash \bigcup_j K_j} f d\mu \Big| \leqslant k\varepsilon$$

Since $|\mu(A_j) - \mu(K_j)| = |\mu(A_j \backslash K_j)| \leqslant \widehat{\mu}(A_j \backslash K_j) \leqslant \varepsilon$, we obtain $|\mu(K_j)| \geqslant |\mu(A_j)| - \varepsilon$, and hence

$$\|\mu\| \geqslant \Big| \int_X f d\mu \Big| \geqslant \Big| \int_{\bigcup_j K_j} f d\mu \Big| - \Big| \int_{X \backslash \bigcup_j K_j} f d\mu \Big| \geqslant \sum_j |\mu(A_j)| - 2k\varepsilon$$

Since ε is arbitrary, we obtain the desired inequality. □

2.10.2 Approximation of Radon measures by Dirac measures

In this section, we let $\mathbb{K} \in \{\mathbb{R}_{\geqslant 0}, \mathbb{R}, \mathbb{C}\}$.

Theorem 2.10.5. *Define*

$$\mathcal{D}(X, \mathbb{K}) = \mathrm{Span}_{\mathbb{K}}\{\delta_x : x \in X\}$$

Then the closed unit ball of $\mathcal{D}(X, \mathbb{K})$ is weak- dense in the closed unit ball of $\mathcal{RM}(X, \mathbb{K})$. In other words, $\overline{B}_{\mathcal{D}(X,\mathbb{K})}(0,1)$ is weak-* dense in $\overline{B}_{\mathcal{RM}(X,\mathbb{K})}(0,1)$.*

The most important case is where $\mathbb{K} = \mathbb{R}_{\geqslant 0}$. In this case, the following proof can be slightly simplified by choosing $\widehat{\mu}$ to be μ.

Proof. Fix $\mu \in \mathcal{RM}(X, \mathbb{K})$ satisfying $\|\mu\| \leqslant 1$. Similar to the proof of Thm. 2.7.4, we let $\mathscr{I}$ be the directed set

$$\begin{gathered} \mathscr{I} = \{(\mathcal{G}, \varepsilon) : \mathcal{G} \in \mathrm{fin}(2^{C_c(X,\mathbb{K})}), \varepsilon \in \mathbb{R}_{\geqslant 0}\} \\ (\mathcal{G}_1, \varepsilon_1) \leqslant (\mathcal{G}_2, \varepsilon_2) \qquad \text{means} \qquad \mathcal{G}_1 \subset \mathcal{G}_2, \varepsilon_1 \geqslant \varepsilon_2 \end{gathered}$$

We claim that for any $(\mathcal{G}, \varepsilon) \in \mathscr{I}$, there exists $\mu_{\mathcal{G},\varepsilon} \in \overline{B}_{\mathcal{D}(X,\mathbb{K})}(0,1)$ such that

$$\Big| \int_X f d\mu - \int_X f d\mu_{\mathcal{G},\varepsilon} \Big| \leqslant \varepsilon \qquad \text{for all } f \in \mathcal{G}$$

If this is true, then $(\mu_{\mathcal{G},\varepsilon})_{(\mathcal{G},\varepsilon) \in \mathscr{I}}$ is a net in $\overline{B}_{\mathcal{D}(X,\mathbb{K})}(0,1)$ converging weak-* to μ. This will finish the proof.

Let us prove the claim. Since μ is a linear combination of finite Radon measures, similar to the proof of Lem. 2.10.4, there exists $\widehat{\mu} \in \mathcal{RM}(X, \mathbb{R}_{\geqslant 0})$ such that

$$\Big|\int_X g d\mu\Big| \leqslant \int_X |g| d\widehat{\mu}$$

for each bounded Borel function $g : X \to \mathbb{C}$.

Let $K \subset X$ be compact and containing $\mathrm{Supp}(f)$ for all $f \in \mathcal{G}$. By the compactness of K, there exist open sets $U_1, \dots, U_k \subset X$ whose union contains K, such that $\mathrm{diam}(f(U_j)) \leqslant \varepsilon/\widehat{\mu}(K)$ for each j and $f \in \mathcal{G}$. Choose a Borel set $A_j \subset U_j$ such that $K = A_1 \sqcup \cdots \sqcup A_k$.[10] Choose any $x_j \in A_j$ if $A_j \neq \emptyset$, and choose any $x_j \in X$ if $A_j = \emptyset$. Let

$$\mu_{\mathcal{G},\varepsilon} = \sum_{j=1}^k \mu_j(A_j)\delta_{x_j} \tag{2.43}$$

Then, for each $f \in \mathcal{G}$,

$$\begin{aligned}
&\Big|\int_X fd(\mu - \mu_{\mathcal{G},\varepsilon})\Big| \leqslant \sum_{j=1}^k \Big|\int_{A_i} fd(\mu - \mu_{\mathcal{G},\varepsilon})\Big| = \sum_{j=1}^k \Big|\int_{A_i} fd\mu - \mu_j(A_j)f(x_i)\Big| \\
&= \sum_{j=1}^k \Big|\int_{A_i} (f - f(x_j))d\mu\Big| \leqslant \sum_{j=1}^k \int_{A_j} |f - f(x_j)|d\widehat{\mu} \leqslant \frac{\varepsilon}{\widehat{\mu}(K)} \sum_{j=1}^k \widehat{\mu}(A_j) = \varepsilon
\end{aligned}$$

This proves the desired inequality. Moreover, by Exp. 2.10.3 and Lem. 2.10.4,

$$\|\mu_{\mathcal{G},\varepsilon}\| = \sum_{j=1}^k |\mu_j(A_j)| \leqslant \|\mu\| \leqslant 1$$

This proves that $\mu_{\mathcal{G},\varepsilon} \in \overline{B}_{\mathcal{D}(X,\mathbb{K})}(0,1)$. □

The proof of Thm. 2.10.5 immediately implies:

Theorem 2.10.6. *For each $\mu \in \mathcal{RM}(X, \mathbb{F})$, we have*

$$\|\mu\| = \sup\Big\{\sum_{j=1}^k |\mu(A_j)| : k \in \mathbb{Z}_+, \text{ and } A_1, \dots, A_k \in \mathfrak{B}_X \text{ are mutually disjoint}\Big\} \tag{2.44}$$

Proof. Lem. 2.10.4 implies "$\geqslant$". Let us prove "$\leqslant$". Let $(\mu_{\mathcal{G},\varepsilon})_{(\mathcal{G},\varepsilon)\in\mathscr{I}}$ be the net in $\mathcal{D}(X, \mathbb{F})$ converging weak-* to μ and satisfying $\|\mu_{\mathcal{G},\varepsilon}\| \leqslant \|\mu\|$, where each $\mu_{\mathcal{G},\varepsilon}$ is of the form (2.43). By Lem. 2.10.3, the RHS of (2.44) is $\geqslant \|\mu_{\mathcal{G},\varepsilon}\|$. Since the net $(\mu_{\mathcal{G},\varepsilon})_{(\mathcal{G},\varepsilon)\in\mathscr{I}}$ converges weak-* to μ, by Fatou's lemma for weak-* convergence (Prop. 2.6.4), the RHS of (2.44) is $\geqslant \|\mu\|$. □

[10]For example, take $A_1 = K \cap U_1$ and $A_j = K \cap U_j \backslash (U_1 \cup \cdots \cup U_{j-1})$ if $j > 1$.

2.11 Problems

Problem 2.1. Let X be an LCH space. Let $(\mu_\alpha)_{\alpha\in\mathscr{I}}$ be a net of finite Radon measures on X satisfying $\sup_\alpha \mu_\alpha(X) < +\infty$ and converging weak-* to some finite Radon measure μ on X (cf. Def. 2.10.1). Let $E \subset X$ be a Borel set.

1. Prove that $\mu(E) \leqslant \liminf\limits_{\alpha} \mu_\alpha(E)$ if E is open, and that $\mu(E) \geqslant \limsup\limits_{\alpha} \mu_\alpha(E)$ if E is compact.
2. Suppose that $\mathrm{Cl}_X(E)$ is compact and $\mathrm{Cl}_X(E)\backslash\mathrm{Int}_X(E)$ is μ-null. Prove that $\mu(E) = \lim_\alpha \mu_\alpha(E)$.
3. Assume that X is a compact interval $[a,b]$ in $\mathbb{R}$ where $a < b$. Find an example of a sequence (μ_n) of Borel measures on $[a,b]$ with $\sup_n \mu_n([a,b]) < +\infty$ converging weak-* to some finite Borel measure μ, such that $\lim_n \mu_n([a,x])$ does not converge to $\mu([a,x])$ for some $x \in [a,b]$.

Note that part 2 is a generalization of the direction (b)⇒(a) in Thm. 2.9.6.

Hint. Part 1. Recall (1.40). If E is compact, use the outer regularity and Urysohn's lemma to show that $\mu(E)$ can be approximated from above by $\int f d\mu$ where $f \in C_c(X,[0,1])$ and $f|_E = 1$.

Part 3. Reduce the problem to one concerning the convergence of distribution functions. □

Problem 2.2. Let X be an LCH space. Let $(\mu_\alpha)_{\alpha\in\mathscr{I}}$ be a net of finite Radon measures on X satisfying $\sup_\alpha \mu_\alpha(X) < +\infty$ and converging weak-* to some finite Radon measure μ on X.

1. Let $f : X \to \mathbb{R}_{\geqslant 0}$. For each $\varepsilon > 0$, show that $0 \leqslant f - f_\varepsilon \leqslant \varepsilon$ where $f_\varepsilon = \sum_{k\in\mathbb{Z}_+} \varepsilon \cdot \chi_{f^{-1}(k\varepsilon,+\infty)}$.
2. Let $f : X \to \mathbb{R}_{\geqslant 0}$. Use part 1 and Pb. 2.1 to prove that if f is lower semicontinuous, then
$$\int_X f d\mu \leqslant \liminf_\alpha \int_X f d\mu_\alpha$$
3. Assume that X is compact. Let $f : X \to \mathbb{C}$ be a bounded Borel function. Suppose that the set $\{x \in X : f \text{ is not continuous at } x\}$ is μ-null. Prove that
$$\lim_\alpha \int_X f d\mu_\alpha = \int_X f d\mu$$

Hint for Part 3. Assume that f is real-valued, and bound f from above and below by semicontinuous functions. □

3 Basics of inner product spaces

3.1 Sesquilinear forms

Let V, W be $\mathbb{C}$-vector spaces.

3.1.1 Sesquilinear forms

Definition 3.1.1. A map of $\mathbb{C}$-vector spaces $T : V \to W$ is called **antilinear** or **conjugate linear** if for every $a, b \in \mathbb{C}$ and $\xi, \eta \in V$ we have

$$T(a\xi + b\eta) = \overline{a}\xi + \overline{b}\eta$$

where $\overline{a}, \overline{b}$ are the complex conjugates of a, b.

Definition 3.1.2. A function $\langle\cdot|\cdot\rangle : V \times V \to \mathbb{C}$ (sending $\xi \times \eta \in V^2$ to $\langle\xi|\eta\rangle$) is called a **sesquilinear form** if it is antilinear on the first variable, and linear on the second one.[1] Namely, for each $a, b \in \mathbb{C}$ and $\xi, \eta, \psi \in V$ we have

$$\langle a\xi + b\eta|\psi\rangle = \overline{a}\langle\xi|\psi\rangle + \overline{b}\langle\eta|\psi\rangle \qquad \langle\psi|a\xi + b\eta\rangle = a\langle\psi|\xi\rangle + b\langle\psi|\eta\rangle$$

More generally, a map $V \times W \to \mathbb{C}$ is also called **sesquilinear** if it is antilinear on the V-component and linear on the W-component. The function

$$V \to \mathbb{C} \qquad \xi \mapsto \langle\xi|\xi\rangle$$

is called the **quadratic form** associated to the sesquilinear form $\langle\cdot|\cdot\rangle$.

Notice the difference between the notations $\langle\xi|\eta\rangle$ and $\langle\xi, \eta\rangle$: the latter always means a bilinear form, i.e., a function which is linear on both variables.

Remark 3.1.3. For each sesquilinear form $\langle\cdot|\cdot\rangle$ on V, we have the **polarization identity**

$$\begin{aligned}\langle\xi|\eta\rangle &= \frac{1}{4}\sum_{t=0,\frac{\pi}{2},\pi,\frac{3\pi}{2}} \langle e^{\mathbf{i}t}\xi + \eta|e^{\mathbf{i}t}\xi + \eta\rangle e^{\mathbf{i}t} \\ &= \frac{1}{4}\Big(\langle\xi + \eta|\xi + \eta\rangle - \langle-\xi + \eta| - \xi + \eta\rangle + \mathbf{i}\langle\mathbf{i}\xi + \eta|\mathbf{i}\xi + \mathbf{i}\eta\rangle - \mathbf{i}\langle-\mathbf{i}\xi + \eta| - \mathbf{i}\xi + \eta\rangle\Big)\end{aligned} \tag{3.1}$$

Therefore, sesquilinear forms are determined by their associated quadratic forms.

Definition 3.1.4. Let $\omega(\cdot|\cdot) : V \times W \to \mathbb{C}$ be a sesquilinear form. The **adjoint sesequilinear form ω^*** is defined to be

$$\omega^* : W \times V \to \mathbb{C} \qquad \omega^*(\eta|\xi) = \overline{\omega(\xi|\eta)}$$

[1] This is different from [Gui-A], where the second variable is assumed to be antilinear

Definition 3.1.5. A sesquilinear form $\langle\cdot|\cdot\rangle : V \times V \to \mathbb{C}$ is called a **Hermitian form** if is equal to its adjoint, namely,

$$\langle\eta|\xi\rangle = \overline{\langle\xi|\eta\rangle} \qquad \text{for each } \xi, \eta \in V$$

Proposition 3.1.6. *Let $\langle\cdot|\cdot\rangle$ be a sesquilinear form on V. The following are equivalent:*

(1) $\langle\cdot|\cdot\rangle$ is a Hermitian form.

(2) The quadratic form associated to $\langle\cdot|\cdot\rangle$ is real-valued, that is, for each $\xi \in V$ we have $\langle\xi|\xi\rangle \in \mathbb{R}$.

Proof. Let $\omega = \langle\cdot|\cdot\rangle$. By the polarization identity, we have $\omega^* = \omega$ iff $\omega^*(\xi|\xi) = \omega(\xi|\xi)$ (i.e. $\overline{\omega(\xi|\xi)} = \omega(\xi|\xi)$) for each $\xi \in V$. $\square$

3.1.2 Positive sesquilinear forms

Definition 3.1.7. A sesquilinear form $\langle\cdot|\cdot\rangle$ on V is called **positive semi-definite** (or simply **positive**) and written as $\langle\cdot|\cdot\rangle \geqslant 0$, if $\langle\xi|\xi\rangle \geqslant 0$ for all $\xi \in V$. If a positive sesquilinear form $\langle\cdot|\cdot\rangle$ on V is fixed, we define

$$\|\xi\| = \sqrt{\langle\xi|\xi\rangle} \qquad \text{for all } \xi \in V \tag{3.2}$$

Then it is clear that $\|\lambda\xi\| = |\lambda| \cdot \|\xi\|$ for each $\xi \in V$ and $\lambda \in \mathbb{C}$. A vector $\xi \in V$ satisfying $\|\xi\| = 1$ is called a **unit vector**.

By Prop. 3.1.6, a positive sesquilinear form is Hermitian. More generally, we have the following definition:

Definition 3.1.8. Let ω_1, ω_2 be Hermitian forms on V. We write

$$\boldsymbol{\omega_1 \leqslant \omega_2}$$

(equivalently, $\omega_2 \geqslant \omega_1$) if the (real-valued) quadratic forms associated to ω_1 and ω_2 satisfy the corresponding inequality, that is,

$$\omega_1(\xi|\xi) \leqslant \omega_2(\xi|\xi) \qquad \text{for each } \xi \in V$$

Thus, "$\leqslant$" defines a partial order on the set of Hermitian forms on V. Moreover, the meaning of $0 \leqslant \omega$ agrees with that in Def. 3.1.7

Theorem 3.1.9 (Cauchy-Schwarz inequality). *Let $\langle\cdot|\cdot\rangle$ be a positive sesquilinear form on V. Then for each $\xi, \eta \in V$ we have*

$$|\langle\xi|\eta\rangle| \leqslant \|\xi\| \cdot \|\eta\|$$

Proof. By linear algebra, if $f : \mathbb{R}^2 \to \mathbb{R}$ is a quadratic form

$$f(x,y) = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = ax^2 + 2bxy + cy^2$$

where $a, b, c \in \mathbb{R}$, then $f \geqslant 0$ iff $a \geqslant 0, b \geqslant 0$ and

$$ac - b^2 \equiv \det \begin{pmatrix} a & b \\ b & c \end{pmatrix} \geqslant 0$$

Indeed, we only need the fact that if $f \geqslant 0$ then $ac-b^2 \geqslant 0$. To see this, note that if f is not always 0, then one of a, c must be nonzero; otherwise, $f(x,y) = 2bxy$ cannot be always $\geqslant 0$. Thus, assume WLOG that $a \neq 0$. Then $f(x,1) = ax^2 + 2bx + c = a(x+b/a)^2 + c - b^2/a$, which implies $a > 0$ and $c - b^2/a \geqslant 0$, and hence $ac - b^2 \geqslant 0$.

Now, we let $f : \mathbb{R}^2 \to \mathbb{R}_{\geqslant 0}$ be the quadratic form defined by pulling back the form $\xi \in V \mapsto \langle \xi | \xi \rangle$ via the map $(x,y) \in \mathbb{R}^2 \mapsto x\xi + y\eta \in V$, that is,

$$f(x,y) = \langle x\xi + y\eta | x\xi + y\eta \rangle = \|\xi\|^2 \cdot x^2 + 2\mathrm{Re}\langle \xi | \eta \rangle \cdot xy + \|\eta\|^2 \cdot y^2$$

Then, the above paragraph shows that $\|\xi\|^2 \cdot \|\eta\|^2 - (\mathrm{Re}\langle \xi|\eta\rangle)^2 \geqslant 0$, equivalently,

$$|\mathrm{Re}\langle \xi | \eta \rangle| \leqslant \|\xi\| \cdot \|\eta\|$$

Choose $\lambda \in \mathbb{S}^1$ such that $\lambda\langle \xi|\eta\rangle \in \mathbb{R}$. Since the above inequality holds when η is replaced by $\lambda\eta$, we get

$$|\langle \xi | \eta \rangle| = |\mathrm{Re}\langle \xi | \lambda\eta \rangle| \leqslant \|\xi\| \cdot \|\lambda\eta\| = \|\xi\| \cdot \|\eta\|$$

□

Corollary 3.1.10. *Let $\langle\cdot|\cdot\rangle$ be a positive sesquilinear form on V. Then we have*

$$\{\xi \in V : \|\xi\| = 0\} = \{\xi \in V : \langle \xi|\psi\rangle = 0 \textit{ for all } \psi \in V\}$$

where the RHS is clearly a linear subspace of V. We call this space the ***null space*** *of $\langle\cdot|\cdot\rangle$.*

Proof. Let $\xi \in V$. If $\langle \xi|V\rangle = 0$, then $\|\xi\|^2 = \langle \xi|\xi\rangle = 0$. Conversely, if $\|\xi\| = 0$, then by the Cauchy-Schwarz inequality, for each $\psi \in V$ we have $|\langle \xi|\psi\rangle| \leqslant \|\xi\| \cdot \|\psi\| = 0$. □

Corollary 3.1.11. *Let $\langle\cdot|\cdot\rangle$ be a positive sesquilinear form on V. Then $\xi \in V \mapsto \|\xi\| \in \mathbb{R}_{\geqslant 0}$ is a seminorm on V.*

Proof. It remains to check the subadditivity: for each $\xi, \eta \in V$, the Cauchy-Schwarz inequality imlies

$$\begin{aligned} &\|\xi + \eta\|^2 = \langle \xi + \eta | \xi + \eta \rangle = \|\xi\|^2 + 2\mathrm{Re}\langle \xi|\eta\rangle + \|\xi\|^2 \\ \leqslant &\|\xi\|^2 + 2\|\xi\| \cdot \|\eta\| + \|\eta\|^2 = (\|\xi\| + \|\eta\|)^2 \end{aligned}$$

□

3.2 Inner product spaces and bounded sesquilinear forms

3.2.1 Inner product spaces

Definition 3.2.1. Let $\langle\cdot|\cdot\rangle$ be a positive sesquilinear form on a $\mathbb{C}$-vector space V. We call $\langle\cdot|\cdot\rangle$ an **inner product** if it is **non-degenerate**, i.e., the null space is 0. We call the pair $(V, \langle\cdot|\cdot\rangle)$ (or simply call V) an **inner product space** or a **pre-Hilbert space** .

Exercise 3.2.2. Let $\langle\cdot|\cdot\rangle$ be a positive sesquilinear form on V with null space $\mathscr{N}$. Prove that there is a (necessarily unique) inner product $\langle\cdot|\cdot\rangle_{V/\mathscr{N}}$ on the quotient space $V/\mathscr{N}$ such that for any $\xi, \eta \in V$, the cosets $\xi + \mathscr{N}$ and $\eta + \mathscr{N}$ satisfy

$$\langle\xi + \mathscr{N}|\eta + \mathscr{N}\rangle_{V/\mathscr{N}} = \langle\xi|\eta\rangle$$

Example 3.2.3. Let X be a set. Then $l^2(X) = l^2(X, \mathbb{C})$ is an inner product space, where

$$\langle f|g\rangle = \sum_{x\in X} \overline{f(x)}g(x) \qquad \text{for any } f, g \in l^2(X)$$

Example 3.2.4. Let (X, μ) be a measure space. Then $L^2(X, \mu)$ is an inner product space, where

$$\langle f|g\rangle = \int_X \overline{f}gd\mu \qquad \text{for any } f, g \in L^2(X, \mu)$$

Remark 3.2.5. By Rem. 3.1.11, an inner product space V is equipped with the norm defined by $\|\xi\| = \sqrt{\langle\xi|\xi\rangle}$. In particular, V is a metric space with metric $d(\xi, \eta) = \|\xi - \eta\|$. The topology on V induced by this metric is called the **norm topology** of V.

Remark 3.2.6. Let V, W be inner product spaces. If $T : V \to V$ is a linear map, then T is an isometry of metric spaces iff T is an isometry of normed vector spaces, i.e.,

$$\langle T\xi|T\xi\rangle = \langle\xi|\xi\rangle \qquad \text{for all } \xi \in V$$

By the polarization identity, this is equivalent to

$$\langle T\xi|T\eta\rangle = \langle\xi|\eta\rangle \qquad \text{for all } \xi, \eta \in V$$

A surjective linear isometry $T : V \to W$ is called a **unitary map**. If $T : V \to W$ is unitary, we say that V, W are **isomorphic inner product spaces** (or that V, W are **unitarily equivalent**).

Similarly, if $T : V \to V$ is antilinear map between inner product spaces, then T is an isometry of metric spaces iff

$$\langle T\xi|T\xi\rangle = \langle\xi|\xi\rangle \qquad \text{for all } \xi \in V$$

By the polarization identity, this is equivalent to

$$\langle T\xi|T\eta\rangle = \langle \eta|\xi\rangle \qquad \text{for all } \xi,\eta\in V$$

A surjective antilinear isometry $T : V \to W$ is called an **antiunitary map**. If $T : V \to W$ is antiunitary, we say that V and W are **antiunitarily equivalent**. □

3.2.2 Bounded sesquilinear forms

Let V, W be inner product spaces.

Definition 3.2.7. The **(complex) conjugate** of V is the inner product space $V^{\complement}$ defined as follows. The elements of $V^{\complement}$ correspond bijectively to those of V by the map

$$\complement : V \to V^{\complement} \qquad \xi \mapsto \xi^{\complement} \equiv \overline{\xi}$$

where $\xi^{\complement} \equiv \overline{\xi}$ is an abstract element, called the **conjugate** of ξ. Moreover, the structure of an inner product space on $V^{\complement}$ is defined in such a way that $\complement$ is antiunitary. In other words, for each $\xi,\eta\in V$ and $a,b\in\mathbb{C}$, we have

$$\begin{gathered}\overline{a}\cdot\overline{\xi}+\overline{b}\cdot\overline{\eta} := \overline{a\xi+b\eta}\\ \langle\overline{\xi}|\overline{\eta}\rangle_{V^{\complement}} := \overline{\langle\xi|\eta\rangle_V} = \langle\eta|\xi\rangle_V\end{gathered}$$

The conjugate of $V^{\complement}$ is defined to be V, that is,

$$(V^{\complement})^{\complement} = V$$

Moreover, the conjugate map $\complement : V^{\complement} \to V$ is defined by

$$\complement : V^{\complement} \to V \qquad \overline{\xi} \mapsto \xi$$

Thus $\overline{\overline{\xi}} = \xi$ for each $\xi\in V$. □

Remark 3.2.8. An antilinear map $T : V \to W$ is equivalent to the linear map

$$V \to W^{\complement} \qquad \xi \mapsto \overline{T\xi} \tag{3.3a}$$

and is also equivalent to the linear map

$$V^{\complement} \to W \qquad \overline{\xi} \mapsto T\xi \tag{3.3b}$$

It is clear that T is an antilinear isometry (resp. antiunitary) iff (3.3a) is a linear isometry (resp. unitary) iff (3.3b) is a linear isometry (resp. unitary).

Remark 3.2.9. A sesquilinear form $\omega : V \times W \to \mathbb{C}$ is equivalent to a bilinear form

$$\widetilde{\omega} : V^{\mathbb{C}} \times W \to \mathbb{C} \qquad (\overline{\xi}, \eta) \mapsto \langle \xi | \eta \rangle$$

Unless otherwise stated, we always view ω and $\widetilde{\omega}$ as the same.

Definition 3.2.10. Let $\omega : V \times W \to \mathbb{C}$ be a sesquilinear form. The **norm** $\|\omega\|$ is defined to be the norm of the associated bilinear form $V^{\mathbb{C}} \times W \to \mathbb{C}$. That is,

$$\|\omega\| = \sup_{\xi \in \overline{B}_V(0,1), \eta \in \overline{B}_W(0,1)} |\omega(\xi|\eta)|$$

Recalling the notation (2.19), we let

$$\mathfrak{Ses}\,(V|W) := \mathfrak{L}(V^{\mathbb{C}} \times W, \mathbb{C})$$

which is the space of bounded sesquilinear forms $V \times W \to \mathbb{C}$. We write

$$\mathfrak{Ses}\,(V) := \mathfrak{Ses}\,(V|V)$$

The elements of $\mathfrak{Ses}\,(V|W)$ (resp. $\mathfrak{Ses}\,(V)$) are called **bounded sesquilinear forms** on $V \times W$ (resp. on V).

Example 3.2.11. The inner product

$$\langle \cdot | \cdot \rangle : V \times V \to \mathbb{C} \qquad (\xi, \eta) \mapsto \langle \xi | \eta \rangle$$

has norm 1, and hence belongs to $\mathfrak{Ses}\,(V)$. Therefore, by Prop. 2.3.10, this map is continuous.

The following useful property says that a sesquilinear form is bounded iff the associated quadratic form is bounded.

Proposition 3.2.12. *Let ω be sesquilinear form on V. Let $M \in \mathbb{R}_{\geqslant 0}$. Assume that*

$$|\omega(\xi|\xi)| \leqslant M\|\xi\|^2$$

for each $\xi \in V$. Then $\|\omega\| \leqslant 4M$.

Proof. Choose any $\xi, \eta \in \overline{B}_V(0,1)$. For each $\lambda \in \mathbb{S}^1$, we have

$$|\omega(\lambda\xi + \eta|\lambda\xi + a\eta)| \leqslant M\|\lambda\xi + \eta\|^2 \leqslant M(\|\xi\| + \|\eta\|)^2 \leqslant 4M$$

Therefore, by the polarization identity (3.1),

$$|\omega(\xi|\eta)| = \frac{1}{4}\Big\| \sum_{t=0,\frac{\pi}{2},\pi,\frac{3\pi}{2}} \omega(e^{\mathbf{i}t}\xi + \eta|e^{\mathbf{i}t}\xi + \eta)e^{\mathbf{i}t}\Big\| \leqslant 4M$$

□

3.3 Orthogonality

Let V be an inner product space.

3.3.1 Orthogonal and orthonormal vectors

Definition 3.3.1. A set $\mathfrak{S}$ of vectors of V are called **orthogonal** if $\langle\xi|\eta\rangle = 0$ for any distinct $\xi, \eta \in V$. An orthogonal set $\mathfrak{S}$ is called **orthonormal** if $\|\xi\| = 1$ for all $\xi \in \mathfrak{S}$.

Remark 3.3.2. We will also talk about an **orthogonal** resp. **orthonormal family of vectors** $(e_i)_{i\in I}$. This means that $\langle e_i|e_j\rangle = 0$ for any distinct $i, j \in I$ (resp. $\langle e_i|e_j\rangle = \delta_{i,j}$ for any $i, j \in I$).

In particular, two vectors $u, v \in V$ are called orthogonal and written as

$$\xi \perp \eta$$

when $\langle\xi|\eta\rangle = 0$. A fundamental fact about orthogonal vectors is

Proposition 3.3.3 (Pythagorean identity). *Suppose that $\xi, \eta \in V$ are orthogonal. Then*

$$\|\xi + \eta\|^2 = \|\xi\|^2 + \|\eta\|^2 \tag{3.4}$$

In particular,

$$\|\xi\| \leqslant \|\xi + \eta\| \tag{3.5}$$

Proof. $\|\xi + \eta\|^2 = \langle\xi + \eta|\xi + \eta\rangle = \langle\xi|\xi\rangle + \langle\eta|\eta\rangle + 2\mathrm{Re}\langle\xi|\eta\rangle = \langle\xi|\xi\rangle + \langle\eta|\eta\rangle$. □

Note that by applying (3.4) repeatedly, we see that if $\xi_1, \dots, \xi_n \in V$ are orthogonal, then

$$\|\xi_1 + \cdots + \xi_n\|^2 = \|\xi_1\|^2 + \cdots + \|\xi_n\|^2 \tag{3.6}$$

Remark 3.3.4. Suppose that $\mathfrak{S}$ is an orthonormal set of vectors of V. Then $\mathfrak{S}$ is clearly linearly independent. (If $e_1, \dots, e_n \in \mathfrak{S}$ and $\sum_i a_i e_i = 0$, then $a_j = \sum_i \langle e_j|a_i e_i\rangle = \langle e_j|0\rangle = 0$.) Thus, by linear algebra, if $\mathfrak{S} = \{e_1, \dots, e_n\}$ is finite, then for each $\xi \in V$, one can find uniquely $a_1, \dots, a_n \in \mathbb{C}$ and $\eta \in V$ such that $\xi = a_1 e_1 + \cdots + a_n e_n + \eta$ and that η is orthogonal to $e_1, \dots, e_n$. The expressions of $a_1, \dots, a_n, \eta$ can be expressed explicitly:

Proposition 3.3.5 (Gram-Schmidt). *Let $e_1, \dots, e_n$ be orthonormal vectors in V. Let $\xi \in V$. Then*

$$\xi - \sum_{i=1}^n e_i \cdot \langle e_i|\xi\rangle \tag{3.7}$$

is orthogonal to $e_1, \dots, e_n$.

Proof. This is a direct calculation and is left to the readers. □

Remark 3.3.6. "Gram-Schmidt" usually refers to the following process. Let $\xi_1, \dots, \xi_n$ be a set of linearly independent vectors of V. Then there is an algorithm of finding an orthonormal basis of $U = \mathrm{Span}\{\xi_1, \dots, \xi_n\}$: Let $e_1 = \xi_1/\|\xi_1\|$. Suppose that a set of orthonormal vectors $e_1, \dots, e_k$ in U have been found. Then e_{k+1} is defined by $\widetilde{\xi}_{k+1}/\|\widetilde{\xi}_{k+1}\|$ where $\widetilde{\xi}_{k+1} = \xi_{k+1} - \sum_{i=1}^k e_i \cdot \langle e_i|\xi_{k+1}\rangle$.

Combining Pythagorean with Gram-Schmidt, we have:

Corollary 3.3.7 (Bessel's inequality). *Let $(e_i)_{i\in I}$ be a family of orthonormal vectors of V. Then for each $\xi \in V$ we have*

$$\sum_{i\in I} |\langle e_i|\xi\rangle|^2 \leqslant \|\xi\|^2 \tag{3.8}$$

In particular, the set $\{i \in I : \langle e_i|\xi\rangle \neq 0\}$ is countable.

Proof. The LHS of (3.8) is $\lim_{J\in\mathrm{fin}(2^I)} \sum_{j\in J} |\langle e_j|\xi\rangle|^2$. Thus, it suffices to show that for each $J \in \mathrm{fin}(2^I)$ we have $\sum_{j\in J} |\langle e_j|\xi\rangle|^2 \leqslant \|\xi\|^2$. Let

$$\eta_1 = \sum_{j\in J} e_j \cdot \langle e_j|\xi\rangle \qquad \eta_2 = \xi - \eta_1$$

(Namely, $\xi = \eta_1 + \eta_2$ is the orthogonal decomposition of ξ with respect to $\mathrm{Span}\{e_j : j \in J\}$.) By Gram-Schmidt, we have $\langle \eta_1|\eta_2\rangle = 0$. By Pythagorean, we have $\|\eta_1\|^2 \leqslant \|\xi\|^2$. But Pythagorean (3.6) also implies

$$\|\eta_1\|^2 = \sum_{j\in J} |\langle e_j|\xi\rangle|^2$$

The last statement about countability follows from Prop. 1.2.44. □

3.3.2 Orthogonal decomposition

Definition 3.3.8. Let U be a linear subspace of V. Let $\xi \in V$. An **orthogonal decomposition** of ξ with respect to U is an expression of the form

$$\xi = \eta + \psi \qquad \text{where } \eta \in U \text{ and } \psi \perp U$$

Orthogonal decompositions of ξ are unique if exist. We call η the **orthogonal projection** of ξ onto U.

Proof of uniqueness. Suppose that $\xi = \eta' + \psi'$ is another orthogonal decomposition. Then $\eta - \eta'$ equals $\psi' - \psi$. Let $\mu = \eta - \eta'$. Then $\mu \in U$ and $\mu \perp U$. So $\langle \mu|\mu\rangle = 0$, and hence $\mu = 0$. So $\eta = \eta'$ and $\psi = \psi'$. □

Definition 3.3.9. Let U be a linear subspace of V. We say that V **has a projection onto** U if every vector has an orthogonal decomposition with respect to U. In that case, we define the map

$$P : V \to V$$

determined by the fact that each $\xi \in V$ has orthogonal decomposition $\xi = P\xi + (\xi - P\xi)$ where $P\xi \in U$ and $\xi - P\xi \perp U$. Clearly P is linear. By the Pythagorean identity, we have $\|P\xi\| \leqslant \|\xi\|$, and hence

$$\|P\| \leqslant 1$$

Thus $P \in \mathfrak{L}(V)$. We say that P is the **projection (operator) associated to** U.

Example 3.3.10. Let $e_1, \dots, e_n$ be orthonormal vectors of V. Let $U = \mathrm{Span}\{e_1, \dots, e_n\}$. Choose any $\xi \in V$. Then by Gram-Schmidt,

$$\xi = \eta + w \qquad \text{where } \eta = \sum_{i=1}^{n} e_i \cdot \langle e_i|\xi\rangle \text{ and } \psi = \xi - \eta \tag{3.9}$$

is the orthogonal decomposition of ξ with respect to U. Therefore, the projection operator associated to U is

$$V \to V \qquad \xi \mapsto \sum_{i=1}^{n} e_i \cdot \langle e_i|\xi\rangle$$

Proposition 3.3.11. *Let U be a linear subspace of V. Suppose that $\xi \in V$ has orthogonal decomposition $\xi = \eta + \psi$ with respect to U. Then*

$$\|\xi - \eta\| = \inf_{\mu \in U} \|\xi - \mu\| \tag{3.10}$$

Proof. Clearly "$\geqslant$" holds. Choose any $\xi \in U$. Then $\xi - \mu = \xi - \eta + \eta - \mu = \psi + (\eta - \mu)$. Since $\eta - \mu \in U$, we have $\psi \perp \eta - \mu$. Thus, by Pythagorean, we have $\|\psi\| \leqslant \|\xi - \mu\|$. □

3.3.3 Direct sums and orthogonal decomposition

Next, we give a more explicit description of orthogonal decomposition in terms of direct sum.

Definition 3.3.12. Let $V_1, \dots, V_n$ be inner product spaces. Their **(orthogonal) direct sum $V_1 \oplus \cdots \oplus V_n$** is an inner product space defined as follows. As a set, $V_1 \oplus \cdots \oplus V_n$ equals $V_1 \times \cdots \times V_n$. So it consists of elements of the form $(\xi_1, \dots, \xi_n)$

where $\xi_i \in V_i$. We write $(\xi_1, \dots, \xi_n)$ as $\xi_1 \oplus \cdots \oplus \xi_n$. The linear structure is defined by

$$(\xi_1 \oplus \cdots \oplus \xi_n) + (\xi_1' \oplus \cdots \oplus \xi_n') = (\xi_1 + \xi_1') \oplus \cdots \oplus (\xi_n + \xi_n')$$
$$a(\xi_1 \oplus \cdots \oplus \xi_n) = a\xi_1 \oplus \cdots \oplus a\xi_n$$

where $\xi_i, \xi_i' \in V_i$ and $a \in \mathbb{C}$. The inner product is defined by

$$\langle \xi_1 \oplus \cdots \oplus \xi_n | \xi_1' \oplus \cdots \oplus \xi_n' \rangle = \langle \xi_1 | \xi_1' \rangle + \cdots + \langle \xi_n | \xi_n' \rangle$$

We view V_i as an inner product subspace of $V_1 \oplus \cdots \oplus V_n$ by identifying $\xi_i \in V_i$ with $0 \oplus \cdots \oplus \xi_i \oplus \cdots \oplus 0 \in V_1 \oplus \cdots \oplus V_n$. Then, it is clear that $V_i \perp V_j$ if $i \neq j$.

Remark 3.3.13. Suppose that $U_1, \dots, U_n$ are mutually orthogonal linear subspaces of V. Then we clearly have a linear isometry

$$U_1 \oplus \cdots \oplus U_n \longrightarrow V \qquad u_1 \oplus \cdots \oplus u_n \mapsto u_1 + \cdots + u_n \tag{3.11}$$

Therefore, if V is spanned by $U_1, \dots, U_n$, then (3.11) is surjective, and hence is an isomorphism of normed vector spaces. In that case, we say that (3.11) is the **canonical isomorphism** from $U_1 \oplus \cdots \oplus U_n$ to V. With abuse of notation, we also say that V "is" the direct sum $U_1 \oplus \cdots \oplus U_n$.

Example 3.3.14. Let U_1, U_2 be inner product spaces and $V = U_1 \oplus U_2$. Then V has a projection onto U_1. The projection operator associated to U_1 is defined by sending each $u_1 \oplus u_2$ to u_1.

We now show that any projection is unitarily equivalent to the one given in Exp. 3.3.14.

Definition 3.3.15. If $\mathfrak{S}$ is a subset of V, we define the **orthogonal complement** of $\mathfrak{S}$ (in V) to be

$$\boldsymbol{\mathfrak{S}^\perp} = \{\xi \in V : \langle \xi | u \rangle = 0 \text{ for all } u \in \mathfrak{S}\}$$

In the case that $\mathfrak{S}$ is a linear subspace U, we also write

$$\boldsymbol{V \ominus U} := U^\perp$$

Remark 3.3.16. Let U be a linear subspace of V. Then $U^\perp$ is closed in V, since it is the intersection of kernels of the bounded linear map $\xi \in V \mapsto \langle \xi | u \rangle$ over all $u \in U$. Moreover, by the continuity of $\langle \cdot | \cdot \rangle : V \times V \to \mathbb{C}$, a vector of V is orthogonal to U iff it is orthogonal to $\overline{U} = \mathrm{Cl}_V(U)$, that is,

$$U^\perp = \overline{U}^\perp$$

Example 3.3.17. If U_1, U_2 are inner product spaces, then U_1 and U_2 are the orthogonal complements of each other in $U_1 \oplus U_2$.

Proposition 3.3.18. *Let U be a linear subspace of V. Suppose that V has a projection onto U, and let P the projection operator onto U. Then V is canonically isomorphic to $U \oplus U^\perp$. Moreover, identifying $U \oplus U^\perp$ with V (by identifying $u \oplus v$ with $u + v$ if $u \in U, v \in U^\perp$), then*

$$P : U \oplus U^\perp \to U \oplus U^\perp \qquad u \oplus v \mapsto u = u \oplus 0$$

Consequently, $1 - P$ is the projection of V onto $U^\perp$, and we have

$$\mathrm{Rng}(P) = \mathrm{Ker}(1-P) = U \qquad \mathrm{Ker}(P) = \mathrm{Rng}(1-P) = U^\perp \tag{3.12}$$

It follows from $V = U \oplus U^\perp$ that U is the orthogonal complement of $U^\perp$, i.e., $U = U^{\perp\perp}$.

Proof. The surjectivity of the linear isometry

$$U \oplus U^\perp \to V \qquad u \oplus v \mapsto u + v$$

follows from the fact that V has a projection onto U. Clearly P sends $u + v$ to u. The rest of this proposition is obvious. □

Corollary 3.3.19. *Suppose that U is a finite-dimensional linear subspace of V. Then U is closed in V.*

Proof. By Exp. 3.3.10, there is a projection operator P of V onto U. By Prop. 3.3.18, U is the orthogonal complement of $U^\perp$, and hence is closed. □

Corollary 3.3.20. *Let U be a linear subspace of V, and suppose that V has a projection onto U. Let P be the projection operator associated to U. Then $P^2 = P$, and the sesquilinear form $\omega_P : V \times V \to \mathbb{C}$ defined by $\omega_P(\xi|\eta) = \langle \xi | P\eta \rangle$ is positive.*

Proof. By Prop. 3.3.18, we assume that $V = U \oplus U^\perp$, and P sends each $u \oplus v \in U \oplus U^\perp$ to $\xi = \xi \oplus 0$. Then it is easy to verify that $P^2 = P$. Moreover, $\omega_P(u \oplus v) = \|u\|^2 \geqslant 0$. □

3.3.4 Orthonormal basis

Definition 3.3.21. A set $\mathfrak{S}$ (or a family $(e_i)_{i \in I}$) of orthonormal vectors of V is called an **orthonormal basis** of V if it spans a dense subspace of V.

Example 3.3.22. If X is a set, by Prop. 2.8.1, $l^2(X)$ has an orthonormal basis $(\chi_{\{x\}})_{x \in X}$.

Example 3.3.23. If V is separable, then V has a countable orthonormal basis.

Proof. Let $\{v_1, v_2, \dots\}$ be a dense subset of V where $v_1 \neq 0$. Then by Gram-Schmidt (Rem. 3.3.6), we can find $e_1, e_2, \dots \in V$ such that the set $\{e_1, e_2, \dots\}$ is orthnormal (after removing the duplicated terms), and that $\mathrm{Span}\{v_1, \dots, v_n\} = \mathrm{Span}\{e_1, \dots, e_n\}$ for each n. Then $\{e_1, e_2, \dots\}$ clearly spans a dense subspace of V. □

We remark that there are non-separable and non-complete inner product spaces that do not have orthonormal bases. See [Gud74].

Theorem 3.3.24. *Suppose that $(e_i)_{i\in I}$ is an orthonormal basis of V. Then for each $\xi \in V$, the RHS of the following converges (in the norm of V) to the LHS:*

$$\xi = \sum_{i\in I} e_i \cdot \langle e_i|\xi\rangle \tag{3.13}$$

Proof. Note that for $J \in \mathrm{fin}(2^I)$, the expression

$$\Big\|\xi - \sum_{j\in J} e_j \cdot \langle e_j|\xi\rangle\Big\|^2 = \|\xi\|^2 - \sum_{j\in J} |\langle e_j|\xi\rangle|^2$$

decreases when J increases. Thus, it suffices to prove that the $\inf_{J\in\mathrm{fin}(2^I)}$ of this expression is 0.

By assumption, we can find $J \in \mathrm{fin}(2^I)$ and $(\lambda_j)_{j\in J}$ in $\mathbb{C}$ such that $\|\xi - \sum_{j\in J}\lambda_j e_j\|$ is small enough. On the other hand, applying Prop. 3.3.11 to the orthogonal projection $\xi = \eta + \psi$ where $\eta = \sum_{j\in J} e_j \cdot \langle e_j|\xi\rangle$ (cf. Exp. 3.3.10), we have

$$\Big\|\xi - \sum_{j\in J} e_j \cdot \langle e_j|\xi\rangle\Big\| \leqslant \Big\|\xi - \sum_{j\in J}\lambda_j e_j\Big\| \tag{3.14}$$

Thus, the infimum of the LHS over $J \in \mathrm{fin}(2^I)$ is zero. □

Corollary 3.3.25 (Parseval's identity). *Suppose that $(e_i)_{i\in I}$ is an orthonormal basis of V. Then for each $\xi, \eta \in V$ we have*

$$\langle \xi|\eta\rangle = \sum_{i\in I}\langle \xi|e_i\rangle \cdot \langle e_i|\eta\rangle \tag{3.15}$$

In particular,

$$\|\xi\|^2 = \sum_{i\in I} |\langle e_i|\xi\rangle|^2 \tag{3.16}$$

Proof. By Thm. 3.3.24, $\xi = \lim_{J\in\mathrm{fin}(2^I)} \xi_J$ where $\xi_J = \sum_{j\in J} e_j \cdot \langle \xi|e_j\rangle$. By the continuity of $\langle\cdot|\cdot\rangle : V \times V \to \mathbb{C}$ (Exp. 3.2.11), we have

$$\langle \eta|\xi\rangle = \lim_{J\in\mathrm{fin}(2^I)} \langle \eta|\xi_J\rangle = \lim_{J\in\mathrm{fin}(2^I)} \sum_{j\in J}\langle \eta|e_j\rangle \cdot \langle e_j|\xi\rangle = \sum_{i\in I}\langle \eta|e_i\rangle \cdot \langle e_i|\xi\rangle$$

□

Corollary 3.3.26. *Suppose that $(e_x)_{x\in X}$ is an orthonormal basis of V. Then there is a linear isometry*

$$\Phi : V \to l^2(X) \qquad \xi \mapsto (\langle e_x|\xi\rangle)_{x\in X} \tag{3.17}$$

whose range is dense in $l^2(X)$.

Proof. Parseval's identity shows that $(\langle e_x|\xi\rangle)_{x\in X}$ has finite l^2-norm $\|\xi\|$. So the map Φ defined by (3.17) is clearly a linear isometry. The density of the range of Φ follows from the fact that $l^2(X)$ contains all $\chi_{\{x\}} = \Phi(e_x)$, and that $\mathrm{Span}\{\chi_{\{x\}} : x \in X\}$ is dense in $l^2(X)$ (cf. Prop. 2.8.1). □

3.4 Hilbert spaces

Theorem 3.4.1. *Let $\mathcal{H}$ be an inner product space. Then the following three conditions are equivalent:*

(a) $\mathcal{H}$ is (Cauchy) complete.

(b) For each orthonormal family $(e_i)_{i\in I}$ in $\mathcal{H}$, and for each family $(a_i)_{i\in I}$ in $\mathbb{C}$ satisfying $\sum_{i\in I}|a_i|^2 < +\infty$, the unordered sum $\sum_{i\in I} a_ie_i$ converges (in the norm of $\mathcal{H}$).

(c) $\mathcal{H}$ is unitarily equivalent to $l^2(X)$ for some set X.

*If $\mathcal{H}$ satisfies any of these conditions, we say that $\mathcal{H}$ is a **Hilbert space**.*

Proof. (c)⇒(a): By Prop. 2.4.10, $l^2(X)$ is complete.

(a)⇒(b): Since $\sum_i |a_i|^2 < +\infty$, for each $\varepsilon > 0$ there exists $J \in \mathrm{fin}(2^I)$ such that for all finite $K \subset I\backslash J$ we have $\sum_{k\in K}|a_k|^2 < \varepsilon$, and hence, by the Pythagorean identity,

$$\Big\|\sum_{k\in K} a_ke_k\Big\|^2 = \sum_{k\in K}\|a_ke_k\|^2 < \varepsilon$$

Thus $(\sum_{j\in J} a_je_j)_{J\in\mathrm{fin}(2^I)}$ is a Cauchy net. Since any Cauchy net in a complete metric space converges (Thm. 1.2.15), by the completeness of $\mathcal{H}$, the unordered sum $\sum_{i\in I} a_ie_i$ must converge.

(b)⇒(c): Assume (b). We first show that $\mathcal{H}$ has an orthonormal basis. By Zorn's lemma, we can find a maximal (with respect to the partial order $\subset$) set of orthonormal vectors, written as a family $(e_i)_{i\in I}$. The maximality implies that every nonzero vector $\xi \in \mathcal{H}$ is not orthogonal to some e_i. (Otherwise, $\{e_i : i \in I\}$ can be extended to $\{e_i : i \in I\} \cup \{\xi/\|\xi\|\}$.)

Let us prove that $(e_i)_{i\in I}$ is an orthonormal basis. Suppose not. Then $U = \mathrm{Span}\{e_i : i \in I\}$ is not dense in $\mathcal{H}$. Let $\xi \in \mathcal{H}\backslash\overline{U}$. By Bessel's inequality, we have

$$\sum_{i\in I}|\langle e_i|\xi\rangle|^2 < +\infty$$

Therefore, by (b),

$$\sum_{i\in I} e_i\cdot\langle e_i|\xi\rangle \tag{3.18}$$

converges to some vector $\eta\in\mathcal{H}$. By the continuity of $\langle\cdot|\cdot\rangle$ (Exp. 3.2.11), we see that $\langle e_i|\eta\rangle=\langle e_i|\xi\rangle$ for all i, and hence

$$\langle e_i|\xi-\eta\rangle=0\qquad\text{for all } i\in I \tag{3.19}$$

Since $\eta\in\overline{U}$ and $\xi\notin\overline{U}$, we conclude that $\xi-\eta$ is a nonzero vector orthogonal to all e_i. This contradicts the maximality of $(e_i)_{i\in I}$.

Now we have an orthonormal basis $(e_i)_{i\in I}$. By Cor. 3.3.26, we have a linear isometry

$$\Phi:\mathcal{H}\to l^2(I)\qquad \xi\mapsto\big(\langle e_i|\xi\rangle\big)_{i\in I}$$

with dense range. If $(a_i)_{i\in I}$ belongs to $l^2(I)$, by (b), the unordered sum $\sum_{i\in I}a_ie_i$ converges to some $\xi\in\mathcal{H}$. Clearly $\Phi(\xi)=(a_i)_{i\in I}$. This proves that Φ is surjective, and hence is a unitary map. So $\mathcal{H}\simeq l^2(I)$. □

In the proof of Thm. 3.4.1, we use Zorn's lemma to show that every Hilbert space $\mathcal{H}$ admits an orthonormal basis. The same argument yields a stronger result:

Example 3.4.2. By Thm. 3.4.1, if X is a set, then $l^2(X)$ is a Hilbert space.

Example 3.4.3. Let (X,μ) be a measure space. By the Riesz-Fischer Thm. 1.6.14, the inner product space $L^2(X,\mu)$ is a Hilbert space.

Example 3.4.4. If V is a closed linear subspace of $\mathcal{H}$ whose inner product is inherited from that of $\mathcal{H}$, then V is a Hilbert space. This is either due to Thm. 3.4.1-(b), or due to the fact that a closed subset of a complete metric space is complete (Prop. 1.4.6). A closed linear subspace of the Hilbert space $\mathcal{H}$ is called a **Hilbert subspace** of $\mathcal{H}$.

Corollary 3.4.5. *Every Hilbert space $\mathcal{H}$ has an orthonormal basis. Moreover, $\mathcal{H}$ is separable iff the orthonormal basis can be chosen to be countable.*

Proof. That $\mathcal{H}$ has an orthonormal basis follows from the proof of Thm. 3.4.1 or from the fact that $l^2(X)$ has an orthonormal basis $(\chi_{\{x\}})_{x\in X}$. If X is countable, then $l^2(X)$ has dense subset $\mathrm{Span}_{\mathbb{Q}+\mathbf{i}\mathbb{Q}}\{\chi_{\{x\}}:x\in X\}$ and hence is separable. Conversely, we have proved in Exp. 3.3.23 that every separable inner product space has a countable orthonormal basis. □

Theorem 3.4.6. *Let $(e_x)_{x\in X}$ be an orthonormal basis of a Hilbert space $\mathcal{H}$. Then we have a unitary map*

$$\mathcal{H}\xrightarrow{\simeq} l^2(X)\qquad \xi\mapsto\big(\langle e_x|\xi\rangle\big)_{x\in X} \tag{3.20}$$

Proof. This is clear from the proof of Thm. 3.4.1. $\square$

Theorem 3.4.7. *Let V be a closed linear subspace of $\mathcal{H}$. Then $\mathcal{H}$ has a projection onto V. Consequently, by Prop. 3.3.18, $V \oplus V^\perp$ is canonically isomorphic to $\mathcal{H}$.*

Proof. By Exp. 3.4.4, V is a Hilbert space, and hence admits an orthonormal basis $(e_i)_{i\in I}$. For each $\xi \in \mathcal{H}$, since Bessel's inequality implies $\sum_i |\langle e_i|\xi\rangle|^2 < \|\xi\|^2 < +\infty$, by Thm. 3.4.1-(b), the following sum converges:

$$P\xi = \sum_{i\in I} e_i \cdot \langle e_i|\xi\rangle$$

and is clearly in V. Similar to the argument around (3.19), $\xi - P\xi$ is orthogonal to every e_i. Hence $V_0 := \mathrm{Span}\{e_i : i \in I\}$ is orthogonal to $\xi - P\xi$, i.e., $\xi - P\xi \in V_0^\perp$. Since V is the closure of V_0, by Rem. 3.3.16, we have $\xi - P\xi \in V^\perp$. Therefore, $\xi = P\xi + (\xi - P\xi)$ is the orthogonal decomposition of ξ with respect to V. $\square$

Corollary 3.4.8. *Let V be a linear subspace of $\mathcal{H}$. Then $(V^\perp)^\perp = \mathrm{Cl}_{\mathcal{H}}(V)$.*

Note that since $V^\perp$ is closed, Cor. 3.4.8 implies $V^{\perp\perp\perp} = V^\perp$.

Proof. By Rem. 3.3.16, we have $V^\perp = \overline{V}^\perp$. By Thm. 3.4.7, $\mathcal{H}$ has a projection onto $\overline{V}$. Therefore, by Prop. 3.3.18, we have $\mathcal{H} = \overline{V} \oplus \overline{V}^\perp = \overline{V} \oplus V^\perp$. Therefore, $\overline{V}$ is the orthogonal complement of $V^\perp$. $\square$

Corollary 3.4.9. *Let V be a linear subspace of $\mathcal{H}$. Then V is dense in $\mathcal{H}$ iff $V^\perp = \{0\}$.*

Proof. If V is dense, then $V^\perp = \overline{V}^\perp = \mathcal{H}^\perp = 0$. Conversely, if $V^\perp = \{0\}$, then $V^{\perp\perp} = 0^\perp = \mathcal{H}$. By Cor. 3.4.8, we have $\overline{V} = V^{\perp\perp} = \mathcal{H}$. Hence V is dense. $\square$

Exercise 3.4.10. Let V be a closed linear subspace of $\mathcal{H}$. Show that an orthonormal basis of V, together with an orthonormal basis of $V^\perp$, forms an orthonormal basis of $\mathcal{H}$. (Consequently, any set $\mathfrak{S}$ of orthonormal vectors of $\mathcal{H}$ can be extended to an orthonormal basis of $\mathcal{H}$ by choosing an orthonormal basis of $V^\perp$ where $V = \mathrm{Span}\mathfrak{S}$.)

3.5 Bounded linear maps and bounded sesquilinear forms

In this section, we let U, V, W be inner product spaces.

In Subsec. 2.5.2, we discussed the close relationship between bounded linear maps and bounded bilinear forms in the general setting of normed vector spaces. This connection allows us to combine the strengths of both perspectives. One key advantage of the perspective of linear operators is that the space $\mathfrak{L}(V)$ is particularly well-suited for symbolic calculus.

In this section, we explore this relationship in the context of inner product spaces and Hilbert spaces. We will see that the passage from $\mathfrak{L}(V)$ to bounded sesquilinear forms fundamentally relies on the Riesz-Fréchet theorem, a pivotal result that enables this correspondence.

3.5.1 The Riesz-Fréchet representation theorem

Definition 3.5.1. If $T \in \mathrm{Lin}(V, W)$, we let ω_T be the sesquilinear form

$$\omega_T : W \times V \to \mathbb{C} \qquad (w, v) \mapsto \langle w|Tv\rangle$$

Proposition 3.5.2. *For each $T \in \mathrm{Lin}(V, W)$, we have*

$$\|T\| = \|\omega_T\|$$

Consequently, T is bounded iff ω_T is so, and the map $T \in \mathrm{Lin}(V, W) \mapsto \omega_T$ is injective.

Proof. For each $v \in V, w \in W$, we have

$$|\omega_T(w|v)| = |\langle w|Tv\rangle| \leqslant \|Tv\| \cdot \|w\| \leqslant \|T\| \cdot \|v\| \cdot \|w\|$$

Applying $\sup$ over all v, w in the closed unit balls, we get $\|\omega_T\| \leqslant \|T\|$. Moreover,

$$\|Tv\|^2 = \omega_T(Tv|v) \leqslant \|\omega_T\| \cdot \|Tv\| \cdot \|v\|$$

and hence $\|Tv\| \leqslant \|\omega_T\| \cdot \|v\|$. Applying $\sup$ over all v in the closed unit ball, we get $\|T\| \leqslant \|\omega_T\|$. $\square$

Theorem 3.5.3 (Riesz-Fréchet representation theorem). *The following map is a linear isometry:*

$$\Phi : W \to (W^{\complement})^* \qquad \xi \mapsto \langle\overline{\xi}|-\rangle \tag{3.21a}$$

where $\langle\overline{\xi}|-\rangle$ denotes the bounded linear functional

$$\langle\overline{\xi}|-\rangle : W^{\complement} \to \mathbb{C} \qquad \overline{w} \mapsto \langle\overline{\xi}|\overline{w}\rangle_{W^{\complement}} = \langle w|\xi\rangle_W \tag{3.21b}$$

Moreover, W is a Hilbert space iff Φ is surjective (and hence an isomorphism of normed vector spaces).

In other words, Φ is determined by the fact that for each $w, \xi \in W$,

$$\langle\overline{w}, \Phi\xi\rangle = \langle w|\xi\rangle \tag{3.22}$$

Proof. First, note that for each $\xi \in W$,

$$\|\xi\| = \sup_{w \in \overline{B}_W(0,1)} |\langle w|\xi\rangle| \tag{3.23}$$

Indeed, the Cauchy-Schwarz inequality implies "$\geqslant$". The equality can be achieved by choosing $w = \xi/\|\xi\|$ if $\xi \neq 0$. Therefore,

$$\|\Phi(\xi)\| = \sup_{\overline{w} \in \overline{B}_{W^{\complement}}(0,1)} |\langle\overline{w}, \Phi(\xi)\rangle| = \sup_{w \in \overline{B}_W(0,1)} |\langle w|\xi\rangle| = \|\xi\|$$

This proves that Φ is a linear isometry.

If Φ is surjective, then the normed vector space W is isomorphic to the dual space $(W^{\complement})^*$ where the latter is complete by Cor. 2.4.11. Therefore, W is a Hilbert space.

Conversely, assume that W is a Hilbert space. By Thm. 3.4.1, we can assume that $W = l^2(X)$ for some set X. The surjectivity of Φ then follows from the surjectivity of the map

$$l^2(X) \to l^2(X)^* \qquad \xi \mapsto \langle -, \xi \rangle$$

due to Thm. 2.8.7. $\square$

Definition 3.5.4. The map Φ in Thm. 3.5.3 is called the **Riesz isometry** of W. If W is a Hilbert space, then Φ is called the **Riesz isomorphism** of W. An equivalent description of Φ is as follows: In view of the isomorphism

$$\mathfrak{Ses}(W) = \mathfrak{L}(W^{\complement} \times W, \mathbb{C}) \simeq \mathfrak{L}(W, (W^{\complement})^*)$$

due to Cor. 2.5.2, the Riesz isometry Φ is the element of $\mathfrak{L}(W, (W^{\complement})^*)$ corresponding the the inner product $\langle \cdot | \cdot \rangle_W$ as an element of $\mathfrak{Ses}(W)$.

3.5.2 Equivalence between bounded linear maps and bounded sesquilinear forms

With the help of the Riesz-Fréchet theorem, we can establish the equivalence between bounded linear maps and bounded sesquilinear forms.

Theorem 3.5.5. *Let $\mathcal{H}, \mathcal{K}$ be Hilbert spaces. Then we have an isomorphism of normed vector spaces*

$$\mathfrak{L}(\mathcal{H}, \mathcal{K}) \xrightarrow{\simeq} \mathfrak{Ses}(\mathcal{K}|\mathcal{H}) \qquad T \mapsto \omega_T \tag{3.24}$$

In particular, when $\mathcal{H} = \mathcal{K}$, the above isomorphism becomes

$$\mathfrak{L}(\mathcal{H}) \xrightarrow{\simeq} \mathfrak{Ses}(\mathcal{H}) \qquad T \mapsto \omega_T \tag{3.25}$$

Proof. By Cor. 2.5.2, we have

$$\mathfrak{L}(\mathcal{H}, (\mathcal{K}^{\complement})^*) \simeq \mathfrak{L}(\mathcal{K}^{\complement} \times \mathcal{H}, \mathbb{C}) = \mathfrak{Ses}(\mathcal{K}|\mathcal{H})$$

where each $S \in \mathfrak{L}(\mathcal{H}, (\mathcal{K}^{\complement})^*)$ corresponds to the bounded bilinear form

$$\mathcal{K}^{\complement} \times \mathcal{H} \to \mathbb{C} \qquad (\overline{\eta}, \xi) \mapsto \langle \overline{\eta}, S\xi \rangle$$

equivalently, the bounded sesquilinear form

$$\mathcal{K} \times \mathcal{H} \to \mathbb{C} \qquad (\eta, \xi) \mapsto \langle \overline{\eta}, S\xi \rangle$$

Now, suppose that $S = \Phi \circ T$ where $T \in \mathfrak{L}(\mathcal{H},\mathcal{K})$, and $\Phi : \mathcal{K} \xrightarrow{\simeq} (\mathcal{K}^{\complement})^*$ is the Riesz-isomorphism of $\mathcal{K}$ defined in Thm. 3.5.3. Then $\langle \overline{\eta}|\Phi\mu\rangle = \langle \eta|\mu\rangle$ for each $\mu,\eta \in \mathcal{K}$, and hence

$$\langle \overline{\eta}, S\xi\rangle = \langle \overline{\eta}, \Phi \circ T\xi\rangle = \langle \eta|T\xi\rangle = \omega_T(\eta|\xi)$$

Therefore, the isomorphism

$$\mathfrak{L}(\mathcal{H},\mathcal{K}) \xrightarrow[\simeq]{T\mapsto \Phi\circ T} \mathfrak{L}(\mathcal{H},(\mathcal{K}^{\complement})^*) \simeq \mathfrak{Ses}(\mathcal{K}|\mathcal{H})$$

sends T to ω_T. □

3.5.3 Adjoint operators, self-adjoint operators and positive operators

Let $\mathcal{H},\mathcal{K}$ be Hilbert spaces. With the help of Thm. 3.5.5, we can define adjoint operators:

Definition 3.5.6. Recall that for each $\omega \in \mathfrak{Ses}(\mathcal{K}|\mathcal{H})$, the **adjoint sesquilinear form** $\omega^* \in \mathfrak{Ses}(\mathcal{H}|\mathcal{K})$ is defined by $\omega^*(\xi|\eta) = \overline{\omega(\eta|\xi)}$ for each $\xi \in \mathcal{H},\eta \in \mathcal{K}$. It is clear that

$$\|\omega^*\| = \|\omega\|$$

Now, for each $T \in \mathfrak{L}(\mathcal{H},\mathcal{K})$, define the **adjoint operator** $T^* \in \mathfrak{L}(\mathcal{K},\mathcal{H})$ such that

$$\omega_{T^*} = (\omega_T)^*$$

More explicitly, T^* is determined by the fact that for each $\xi \in \mathcal{H},\eta \in \mathcal{K}$,

$$\langle \eta|T\xi\rangle = \langle T^*\eta|\xi\rangle$$

Then, we clearly also have $\|T\| = \|T^*\|$.

Exercise 3.5.7. Show that

$$* : \mathfrak{L}(\mathcal{H},\mathcal{K}) \to \mathfrak{L}(\mathcal{K},\mathcal{H}) \qquad T \mapsto T^*$$

is a bijective antilinear map, and that $(T^*)^* = T$. Prove that if $\mathcal{M}$ is a Hilbert space and $T \in \mathfrak{L}(\mathcal{H},\mathcal{K})$, $S \in \mathfrak{L}(\mathcal{K},\mathcal{M})$, then

$$(ST)^* = T^*S^*$$

Definition 3.5.8. A bounded linear operator $T \in \mathfrak{L}(\mathcal{H})$ is called **self-adjoint** if $T = T^*$, equivalently, if ω_T is Hermitian.

Definition 3.5.9. Let $A, B \in \mathfrak{L}(\mathcal{H})$ be self-adjoint. We write

$$\boldsymbol{A \leqslant B}$$

if $\omega_A \leqslant \omega_B$ in the sense of Def. 3.1.8, that is, $\langle \xi|A\xi\rangle \leqslant \langle \xi|B\xi\rangle$ for all $\xi \in \mathcal{H}$. We say that $A \in \mathfrak{L}(\mathcal{H})$ is **positive** if $A \geqslant 0$, equivalently, if ω_A is positive.

Example 3.5.10. Let $A \in \mathfrak{L}(\mathcal{H}, \mathcal{K})$. Then $A^*A \in \mathfrak{L}(\mathcal{H})$ is positive, because

$$\langle \xi | A^*A\xi \rangle = \|A\xi\|^2 \geqslant 0$$

Example 3.5.11. Let $A \in \mathfrak{L}(\mathcal{H})$, and let $a \geqslant 0$ such that $\|A\| \leqslant a$. Then $-a \leqslant A \leqslant a$.

Proof. Since $|\langle \eta | A\xi \rangle| \leqslant a\|\eta\| \cdot \|\xi\|$, we obtain $-a\|\xi\|^2 \leqslant \langle \xi | A\xi \rangle \leqslant a\|\xi\|^2$, and hence $-a \leqslant A \leqslant a$. □

3.5.4 Composition of bounded linear operators and bounded sesquilinear forms

Let $\mathcal{H}, \mathcal{K}, \mathcal{M}$ be Hilbert spaces.

One of the major advantages of working with bounded linear operators rather than bounded sesquilinear forms is the ease with which one can handle problems involving operator composition. This does not mean, however, that a notion of composition cannot be defined on the side of sesquilinear forms. In fact, the following lemma illustrates how such a composition can be defined.

Lemma 3.5.12. *Let $T \in \mathfrak{L}(\mathcal{K}, \mathcal{H})$ and $S \in \mathfrak{L}(\mathcal{M}, \mathcal{K})$. Let $(e_i)_{i \in I}$ be an orthonormal basis of $\mathcal{K}$. Then for each $\xi \in \mathcal{M}$, we have*

$$T \circ S\xi = \sum_{i \in I} Te_i \cdot \langle e_i | S\xi \rangle \tag{3.26}$$

where the unordered sum on the RHS converges in norm to the LHS.

Proof. By Thm. 3.3.24, we have $S\xi = \sum_i e_i \cdot \langle e_i | S\xi \rangle$. Therefore, by the linearity and the continuity of T, we get (3.26). □

Definition 3.5.13. Let $\omega \in \mathfrak{Ses}(\mathcal{H}|\mathcal{K})$ and $\sigma \in \mathfrak{Ses}(\mathcal{K}|\mathcal{M})$. Then the **composition** $\boldsymbol{\omega \circ \sigma}$ is the element of $\mathfrak{Ses}(\mathcal{M}|\mathcal{H})$ defined by [2]

$$(\omega \circ \sigma)(\psi|\xi) = \sum_{i \in I} \omega(\psi|e_i) \cdot \sigma(e_i|\xi) \qquad \text{for all } \psi \in \mathcal{H}, \xi \in \mathcal{M}$$

where $(e_i)_{i \in I}$ is a basis of $\mathcal{K}$. This definition is independent of the choice of basis (and applies even to bounded sesquilinear forms on inner product spaces). Moreover, by Lem. 3.5.12, for each $T \in \mathfrak{L}(\mathcal{K}, \mathcal{H})$ and $S \in \mathfrak{L}(\mathcal{M}, \mathcal{K})$, we have

$$\omega_{T \circ S} = \omega_T \circ \omega_S$$

However, many properties about composition that are straightforward from the perspective of bounded linear operators become far less transparent when viewed in terms of sesquilinear forms. For instance, consider the following basic result:

[2]This definition clearly also applies when $\mathcal{H}$ and $\mathcal{K}$ are merely inner product spaces admitting orthonormal bases, in particular, when $\mathcal{H}, \mathcal{K}$ are separable inner product spaces (cf. Exp. 3.3.23).

Proposition 3.5.14. *Let $\mathcal{U},\mathcal{V},\mathcal{W}$ be normed vector spaces. Then for any $T\in\mathfrak{L}(\mathcal{V},\mathcal{W})$ and $S\in\mathfrak{L}(\mathcal{U},\mathcal{V})$, we have*

$$\|TS\|\leqslant\|T\|\cdot\|S\|$$

In other words, the multiplication map

$$\mathfrak{L}(\mathcal{V},\mathcal{W})\times\mathfrak{L}(\mathcal{U},\mathcal{V})\rightarrow\mathfrak{L}(\mathcal{U},\mathcal{W})\qquad(T,S)\mapsto TS$$

is a bounded bilinear map with operator norm $\leqslant 1$.

Proof. Apply $\sup$ over all $\xi\in\overline{B}_{\mathcal{U}}(0,1)$ to

$$\|TS\xi\|\leqslant\|T\|\cdot\|S\xi\|\leqslant\|T\|\cdot\|S\|\cdot\|\xi\|$$

□

Corollary 3.5.15. *Let $\mathcal{V}$ be a normed $\mathbb{F}$-vector space. Let $\mathscr{A}$ be an $\mathbb{F}$-subalgebra of $\mathfrak{L}(\mathcal{V})$. Let $\overline{\mathscr{A}}$ be the closure of $\mathscr{A}$ in $\mathfrak{L}(\mathcal{V})$. Then $\mathscr{A}$ is also an $\mathbb{F}$-subalgebra of $\mathfrak{L}(\mathcal{V})$.*

Proof. Clearly $\overline{\mathscr{A}}$ is an $\mathbb{F}$-linear subspace of $\mathfrak{L}(\mathcal{V})$. If $T,S\in\mathscr{A}$, then there exist sequences (T_n) and (S_n) in $\mathscr{A}$ converging in the operator norm to T and S, respectively. By Prop. 3.5.14, the multiplication map $\mathfrak{L}(\mathcal{V})\times\mathfrak{L}(\mathcal{V})\rightarrow\mathfrak{L}(\mathcal{V})$ is continuous. Thus (T_nS_n) converges in the operator norm to TS. This proves that $TS\in\overline{\mathscr{A}}$. Therefore, $\overline{\mathscr{A}}$ is an $\mathbb{F}$-subalgebra. □

Corollary 3.5.16. *Let $T\in\mathfrak{L}(\mathcal{H})$. Let $\Omega=\{z\in\mathbb{C}:|z|>\|T\|\}$. Then for each $z\in\Omega$, the operator $z-T$ is invertible (cf. Def. 2.3.7). Moreover, for each $\xi,\eta\in\mathcal{H}$, the function*

$$z\in\Omega\mapsto\langle\eta|(z-T)^{-1}\xi\rangle=\omega_{(z-T)^{-1}}(\eta|\xi)$$

is holomorphic.

The expression $(z-T)^{-1}$ is called the **resolvent** of T.

Proof. By Prop. 3.5.14, we have $\|T^k\|\leqslant\|T\|^k$. Therefore, if $z\in\Omega$, then

$$\sum_{k=0}^{\infty}\|z^{-k-1}T^k\|\leqslant\sum_{k=0}^{\infty}|z|^{-k-1}\|T\|^k<+\infty$$

Therefore, if we define

$$S_n(z)=\sum_{k=0}^{n}z^{-k-1}T^k\tag{3.27}$$

Then $(S_n(z))_{n\in\mathbb{N}}$ is a Cauchy sequence in the normed vector space $\mathfrak{L}(\mathcal{H})$. By Cor. 2.4.9, $\mathfrak{L}(\mathcal{H}) \simeq \mathfrak{Ses}(\mathcal{H})$ is complete. Therefore, $(S_n(z))$ converges in the operator norm to some $S(z) \in \mathfrak{L}(\mathcal{H})$. Since

$$(z-T)S_n(z) = S_n(z)\cdot(z-T) = 1 - z^{-n-1}T^{n+1}$$

and since $\|z^{-n-1}T^{n+1}\| \leqslant |z|^{-n-1}\|T\|^{n+1} \to 0$, we have $(z-T)S(z) = S(z)(z-T) = 1$. This proves that $z - T$ is invertible.

For each $\xi, \eta \in \mathcal{H}$, and for each compact $K \subset \Omega$, we have

$$\sup_{z\in K}\sum_{k=0}^{\infty}|z^{-k-1}\langle\eta|T^k\xi\rangle| \leqslant \sup_{z\in K}\sum_{k=0}^{\infty}|z|^{-k-1}\|T\|^k\cdot\|\eta\|\cdot\|\xi\| < +\infty$$

Therefore, the series of functions

$$z\in\Omega \mapsto \sum_{k=0}^{\infty}\langle\eta|z^{-k-1}T^k\xi\rangle$$

converges absolutely and uniformly on compact subsets of Ω. Since the limit of this series of functions is $z \in \Omega \mapsto \omega_{S_n(z)}(\eta|\xi)$, the latter is holomorphic. □

Before we explore further examples, let us examine another foundational perspective that played a central role in the early development of functional analysis: the viewpoint of bounded matrices.

3.6 Bounded matrices

Let $\mathcal{H}, \mathcal{K}, \mathcal{M}$ be Hilbert spaces.

As mentioned in Subsec. 2.5.1, early developments in functional analysis focused primarily on bounded sesquilinear forms rather than bounded linear operators. Closely tied to this approach was the study of infinite matrices, which provided a concrete representation of these abstract objects. The notion of boundedness was first defined in this matrix context. Hilbert introduced this concept in [Hil06], where he also introduced the space $l^2(\mathbb{Z})$.

Definition 3.6.1. Let X, Y be sets. The elements of $\mathbb{C}^{X\times Y}$, which are of the form

$$A = (A(x,y))_{x\in X, y\in Y} \qquad \text{where } A(x,y) \in \mathbb{C}$$

are called ***X* × *Y* (complex) matrices**. For each $A \in \mathbb{C}^{X\times Y}$, the **norm $\|A\|$** is defined to be

$$\|A\| = \sup\Big\{\Big|\sum_{x\in X, y\in Y}\overline{\xi(x)}A(x,y)\eta(y)\Big| : \xi \in C_c(X), \eta \in C_c(Y), \|\xi\|_{l^2} \leqslant 1, \|\eta\|_{l^2} \leqslant 1\Big\} \tag{3.28}$$

where $C_c(X)$ and $C_c(Y)$ are the spaces of complex functions on X and Y with finite supports, respectively. We say that A is **bounded** if $\|A\| < +\infty$.

Definition 3.6.2. Suppose that $(e_x)_{x\in X}$ and $(e_y)_{y\in Y}$ are orthonormal basis of $\mathcal{H}$ and $\mathcal{K}$, respectively. The **matrix representation** of each $\omega \in \mathfrak{Ses}(\mathcal{H}|\mathcal{K})$ is the element $[\omega] \in \mathbb{C}^{X\times Y}$ defined by

$$[\omega](x,y) = \omega(e_x|e_y) \qquad \text{for each } x \in X, y \in Y$$

If $\omega = \omega_T$ where $T \in \mathfrak{L}(\mathcal{K},\mathcal{H})$, we also say that $[\omega]$ is the **matrix representation** of T and write it as $[T]$. In other words, we say that $[T] \in \mathbb{C}^{X\times Y}$ is the matrix representation of T if

$$[T](x,y) = \langle e_x|Te_y\rangle \qquad \text{for each } x \in X, y \in Y$$

Theorem 3.6.3. *Suppose that $(e_x)_{x\in X}$ and $(e_y)_{y\in Y}$ are orthonormal bases of $\mathcal{H}$ and $\mathcal{K}$, respectively. Then the linear map*

$$\mathfrak{Ses}(\mathcal{H}|\mathcal{K}) \to \mathbb{C}^{X\times Y} \qquad \omega \mapsto [\omega] \tag{3.29}$$

is injective, and its range is the set of all bounded matrices. Moreover, for each $\omega \in \mathfrak{Ses}(\mathcal{H}|\mathcal{K})$, we have

$$\|\omega\| = \|[\omega]\|$$

Proof. Step 1. We assume WLOG that $\mathcal{H} = l^2(X), \mathcal{K} = l^2(Y)$ and $e_x = \{\chi_{\{x\}}\}_{x\in X}$ and $e_y = \{\chi_{\{y\}}\}_{y\in Y}$. Recall that $C_c(X)$ is dense in $l^2(X)$ and $C_c(Y)$ is dense in $l^2(Y)$.

Let $\omega \in \mathfrak{Ses}(\mathcal{H}|\mathcal{K})$. From (3.28), it is clear that $\|[\omega]\|$ is the norm of the restriction of ω to $C_c(X) \times C_c(Y)$. Therefore, by the continuity of ω, we have $\|[\omega]\| = \|\omega\| < +\infty$. In particular, we have proved that the matrices in the range of (3.29) are bounded. Moreover, if $[\omega] = 0$, then $\|\omega\| = \|[\omega]\| = 0$, and hence $\omega = 0$. This proves that (3.29) is injective.

Step 2. Choose any bounded $A \in \mathbb{C}^{X\times Y}$. We want to find $\omega \in \mathfrak{Ses}(\mathcal{H}|\mathcal{K})$ such that $[\omega] = A$. Define a sesquilinear form $\omega : C_c(X) \times C_c(Y) \to \mathbb{C}$ by

$$\omega(\xi|\eta) = \sum_{x\in X, y\in Y} \overline{\xi(x)}A(x,y)\eta(y)$$

Then $\|\omega\| = \|A\|$, and hence ω is bounded. By Thm. 2.4.2, ω can be extended to a bounded sesquilinear form on $\mathcal{H} \times \mathcal{K}$ with $\|\omega\|$ unchanged. Clearly A is the matrix representation of ω. □

Definition 3.6.4. Let X, Y, Z be sets. Let $A \in \mathbb{C}^{X\times Y}$ and $B \in \mathbb{C}^{Y\times Z}$ be bounded matrices. Then the **matrix multiplication** $AB \in \mathbb{C}^{X\times Z}$ is defined to be

$$(AB)(x,z) = \sum_{y\in Y} A(x,y)B(y,z)$$

where the RHS is convergent for each $x \in X, z \in Z$. This definition is clearly compatible with Def. 3.5.13, that is, if $\mathcal{H}, \mathcal{K}, \mathcal{M}$ have orthonormal basis $(e_x)_{x\in X}, (e_y)_{y\in Y}, (e_z)_{z\in Z}$ respectively, and if $\omega \in \mathfrak{Ses}(\mathcal{H}|\mathcal{K}), \sigma \in \mathfrak{Ses}(\mathcal{K}|\mathcal{M})$, then the corresponding matrix representations satisfy

$$[\omega \circ \sigma] = [\omega] \cdot [\sigma]$$

We now return to the topic discussed at the end of Sec. 3.5.4: the subtlety of defining and understanding composition on the side of bounded sesquilinear forms—a subtlety that also arises in the context of bounded matrices. For simplicity, we restrict attention to a fixed Hilbert space $\mathcal{H}$ with orthonormal basis $(e_x)_{x\in X}$.

For $R, S, T \in \mathfrak{L}(\mathcal{H})$, associativity of composition,

$$(RS)T = R(ST)$$

is almost tautological. However, when working with bounded sesquilinear forms or bounded matrices, associativity is far less transparent. To see this, consider $\sigma, \omega, \tau \in \mathfrak{Ses}(\mathcal{H})$. Then the associativity $(\sigma\omega)\tau = \sigma(\omega\tau)$ amounts to the commutativity of the two unordered sums: for all $\xi, \eta \in \mathcal{H}$,

$$\sum_{y\in X}\sum_{x\in X} \sigma(\xi|e_x)\omega(e_x|e_y)\tau(e_y|\eta) = \sum_{x\in X}\sum_{y\in X} \sigma(\xi|e_x)\omega(e_x|e_y)\tau(e_y|\eta) \tag{3.30}$$

Similarly, if $A, B, C \in \mathbb{C}^{X\times X}$ are bounded matrices, associativity of matrix multiplications means that for each $i, j \in X$,

$$\sum_{y\in X}\sum_{x\in X} A(i,x)B(x,y)C(y,j) = \sum_{x\in X}\sum_{y\in X} A(i,x)B(x,y)C(y,j) \tag{3.31}$$

At first glance, the commutativity of $\sum_{x\in X}$ and $\sum_{y\in X}$ is not at all obvious.

The issue of commutativity of unordered sums—which appears in the frameworks of sesquilinear forms and matrices—disappears in the perspective of linear maps. Where is this Fubini-type property hidden in the linear map viewpoint? And how can one understand the commutativity of such unordered sums in a more general context? We will answer this question in the next section.

3.7 SOT and WOT

Let U, V, W be inner product spaces.

3.7.1 Convergence of vectors

Definition 3.7.1. The **weak topology** on V is defined to be the pullback of the weak-* topology on $(V^{\complement})^*$ by the Riesz isometry $V \to (V^{\complement})^*$. Therefore, a net (ξ_α)

in V converges weakly to $\xi \in V$ iff

$$\lim_\alpha \langle \eta|\xi_\alpha\rangle = \langle \eta|\xi\rangle \tag{3.32}$$

holds for each $\eta \in V$

It is clear that norm convergence implies weak convergence.

Remark 3.7.2. Let (ξ_α) be a uniformly bounded net in V, and let $\xi \in V$. Let U be a dense linear subspace of V. Applying Thm. 2.6.2 to the images of (ξ_α) and ξ in $(V^\complement)^*$, we see that (ξ_α) converges weakly to ξ iff (3.32) holds for each $\eta \in U$.

Proposition 3.7.3 (Fatou's lemma for weak convergence). *Let (ξ_α) be a net in V converging weakly to $\xi \in V$. Then*

$$\|\xi\| \leqslant \liminf_\alpha \|\xi_\alpha\| \tag{3.33}$$

Moreover, (ξ_α) converges in norm to ξ iff $\lim_\alpha \|\xi_\alpha\|$ converges to $\|\xi\|$.

Proof. The inequality (3.33) follows from Prop. 2.6.4 and the fact that the Riesz isometry $V \to (V^\complement)^*$ is an isometry. But it can also be proved directly:

$$\|\xi\|^2 = \lim_\alpha |\langle \xi|\xi_\alpha\rangle| = \liminf_\alpha |\langle \xi|\xi_\alpha\rangle| \leqslant \liminf_\alpha \|\xi\| \cdot \|\xi_\alpha\|$$

If (ξ_α) converges in norm to ξ, then by the continuity of $\|\cdot\| : V \to \mathbb{R}_{\geqslant 0}$, $\lim_\alpha \|\xi_\alpha\|$ converges to $\|\xi\|$. Conversely, suppose that $\lim_\alpha \|\xi_\alpha\| = \|\xi\|$. Then, since $\langle \xi|\xi_\alpha\rangle \to \langle \xi|\xi\rangle$, we have

$$\langle \xi - \xi_\alpha|\xi - \xi_\alpha\rangle = \|\xi\|^2 - 2\Re\langle \xi|\xi_\alpha\rangle + \|\xi_\alpha\|^2 \to \|\xi\|^2 - 2\Re\|\xi\|^2 + \|\xi\|^2 = 0$$

This shows that $\|\xi - \xi_\alpha\| \to 0$. □

Example 3.7.4. Let $\mathcal{H}$ be an infinite-dimensional Hilbert space. Let $(e_n)_{n\in\mathbb{Z}_+}$ be a sequence of orthonormal vectors. By Bessel's inequality (Cor. 3.3.7), for each $\xi \in \mathcal{H}$ we have $\sum_{n\in\mathbb{Z}_+} |\langle e_n|\xi\rangle|^2 < +\infty$, and hence $\lim_n \langle e_n|\xi\rangle = 0$. Therefore, the sequence $(e_n)_{n\in\mathbb{Z}_+}$ converges weakly to 0. However, it does not converge in norm to 0.

Theorem 3.7.5. *Let $\mathcal{H}$ be a Hilbert space. Then the closed unit ball $\overline{B}_{\mathcal{H}}(0,1)$ is weakly compact.*

Proof. This follows immediately from the Riesz isomorphism $\mathcal{H} \simeq (\mathcal{H}^\complement)^*$ (cf. Thm. 3.5.3) and the weak-* compactness of the closed unit ball of $(\mathcal{H}^\complement)^*$ (due to the Banach-Alaoglu Thm. 2.6.5).

Alternatively, one may assume that $\mathcal{H} = l^2(X)$ where X is a set. For each net (f_α) in the closed unit ball $\overline{B}_{l^2(X)}(0,1)$, by Tychonoff's theorem, (f_α) has a subnet (f_{α_ν}) converging pointwise on X. By Thm. 2.8.5, (f_{α_ν}) converges weak-* to some $f \in l^2(X)$. By Prop. 3.7.3, f has length $\leqslant 1$. This proves that any net (f_α) in $\overline{B}_{l^2(X)}(0,1)$ admits a weakly (i.e. weak-*) convergent subnet. Hence $\overline{B}_{l^2(X)}(0,1)$ is compact. □

3.7.2 Convergence of operators

Recall that the **norm topology** on $\mathfrak{L}(V, W)$ is the topology determined by the operator norm on $\mathfrak{L}(V, W)$.

Definition 3.7.6. The **strong operator topology (SOT)** on $\mathfrak{L}(V, W)$ is defined to be the pullback of the product topology on W^V by the inclusion map

$$\mathfrak{L}(V, W) \hookrightarrow W^V$$

The **weak-* topology** on $\mathfrak{Ses}(W|V)$ is defined to be the pullback of the product topology on $\mathbb{C}^{W \times V}$ by the inclusion map

$$\mathfrak{Ses}(W|V) \hookrightarrow \mathbb{C}^{W \times V}$$

The **weak operator topology (WOT)** on $\mathfrak{L}(V, W)$ is defined to be the pullback of the weak-* topology on $\mathfrak{Ses}(W|V)$ by the linear isometry

$$\mathfrak{L}(V, W) \rightarrow \mathfrak{Ses}(W|V) \qquad T \mapsto \omega_T$$

Remark 3.7.7. Let (T_α) be a net in $\mathfrak{L}(V, W)$ and $T \in \mathfrak{L}(V, W)$. Then (T_α) converges in SOT to T iff

$$\lim_\alpha T_\alpha \xi = T\xi \tag{3.34}$$

holds for each $\xi \in V$. (T_α) converges in WOT to T iff

$$\lim_\alpha \langle \eta | T_\alpha \xi \rangle = \langle \eta | T\xi \rangle \tag{3.35}$$

holds for each $\xi \in V, \eta \in W$. It is clear that

convergence in norm $\Rightarrow$ convergence in SOT $\Rightarrow$ convergence in WOT

Remark 3.7.8. Let (T_α) be a net in $\mathfrak{L}(V, W)$ and $T \in \mathfrak{L}(V, W)$. Then it is clear that (T_α) converges in WOT to T iff

$$\lim_\alpha T\xi_\alpha \quad \text{converges weakly to} \quad T\xi$$

for each $\xi \in V$. By Prop. 3.7.3, (T_α) converges in SOT to T iff the following two conditions hold:

(1) (T_α) converges in WOT to T.

(2) $\lim_\alpha \|T_\alpha \xi\| = \|T\xi\|$ for each $\xi \in V$.

Example 3.7.9. Suppose that V has a basis $(e_x)_{x\in X}$. For each $I \in \mathrm{fin}(2^X)$, let P_I be the projection of V onto $V_I = \mathrm{Span}\{e_x : x \in I\}$, that is,

$$P_I\xi = \sum_{x\in I} e_x \cdot \langle e_x|\xi\rangle \qquad \text{for all } \xi \in V$$

Then by Thm. 3.3.24, $\lim_{I\in\mathrm{fin}(2^X)} P_I$ converges in SOT to $\mathbf{1}$.

However, if X is infinite, then $\lim_{I\in\mathrm{fin}(2^X)} P_I$ does not converge in norm to $\mathbf{1}$. Indeed, for each I, choose $x \in X\backslash I$. Then $(\mathbf{1} - P_I)e_x = e_x$, and hence

$$\|\mathbf{1} - P_I\| \geqslant 1$$

Indeed, by Prop. 3.3.18, $\mathbf{1} - P_I$ is the projection operator associated to $U_I^\perp$, and hence has operator norm 1. □

3.7.3 SOT and composition of operators

Let us examine the commutativity of unordered sums from Sec. 3.6 through the lens of SOT.

Proposition 3.7.10. *Let $(T_\alpha)_{\alpha\in\mathscr{A}}$ be a net in $\mathfrak{L}(V, W)$ converging in SOT to $T \in \mathfrak{L}(V, W)$. Let $(\xi_\beta)_{\beta\in\mathscr{B}}$ be a net in V converging in norm to $\xi \in V$. Then*

$$\lim_\alpha \lim_\beta T_\alpha\xi_\beta = \lim_\beta \lim_\alpha T_\alpha\xi_\beta = T\xi \tag{3.36}$$

Proof. We compute that

$$\lim_\alpha \lim_\beta T_\alpha\xi_\beta = \lim_\alpha T_\alpha\xi = T\xi$$

and

$$\lim_\beta \lim_\alpha T_\alpha\xi_\beta = \lim_\beta T\xi_\beta = T\xi$$

where the last equality is due to the continuity of T. □

Corollary 3.7.11. *Let $(T_\alpha)_{\alpha\in\mathscr{A}}$ be a net in $\mathfrak{L}(V, W)$ converging in SOT to $T \in \mathfrak{L}(V, W)$. Let $(S_\beta)_{\beta\in\mathscr{B}}$ be a net in $\mathfrak{L}(U, V)$ converging in SOT to $S \in \mathfrak{L}(U, V)$. Then for each $u \in U$ we have*

$$\lim_\alpha \lim_\beta T_\alpha S_\beta u = \lim_\beta \lim_\alpha T_\alpha S_\beta u = TSu \tag{3.37}$$

Proof. Apply Prop. 3.7.10 to $\xi_\beta = S_\beta u$. □

Cor. 3.7.11 can be easily generalized to products of more than two nets of operators. We leave the details to the reader.

Example 3.7.12. Let $\mathcal{H}$ be a Hilbert space with basis $(e_x)_{x\in X}$. Let $(P_I)_{I\in \mathrm{fin}(2^X)}$ be the net of projections where P_I projects $\mathcal{H}$ onto $V_I = \mathrm{Span}\{e_x : x \in I\}$. By Exp. 3.7.9, $\lim_I P_I$ converges in SOT to $\mathbf{1}$. Choose any $R, S, T \in \mathfrak{L}(\mathcal{H})$. Then $\lim_I P_I S$ converges in SOT to S, and $\lim_I P_I T$ converges in SOT to T. Therefore, by Cor. 3.7.11, for each $\eta \in \mathcal{H}$ we have

$$\lim_{J\in\mathrm{fin}(2^X)} \lim_{I\in\mathrm{fin}(2^X)} RP_I SP_J T\eta = \lim_{I\in\mathrm{fin}(2^X)} \lim_{J\in\mathrm{fin}(2^X)} RP_I SP_J T\eta = RST\eta$$

Therefore, for each $\xi, \eta \in \mathcal{H}$ we have

$$\lim_{J\in\mathrm{fin}(2^X)} \lim_{I\in\mathrm{fin}(2^X)} \langle \xi | RP_I SP_J T\eta\rangle = \lim_{I\in\mathrm{fin}(2^X)} \lim_{J\in\mathrm{fin}(2^X)} \langle \xi | RP_I SP_J T\eta\rangle = \langle \xi | RST\eta\rangle \tag{3.38}$$

The commutativity of the two iterated limits in (3.38) is equivalent to the commutativity of the two limits in (3.30).

Next, we introduce an elementary yet fundamental property of SOT. This property was first highlighted by Riesz in [Rie13] as a key tool in his proof of the spectral theorem for bounded self-adjoint operators. In Ch. 5, we will also rely on it in our study of spectral theory, in particular when reproducing Riesz's argument from [Rie13] (see the proof of Thm. 5.5.17). This property roughly says that in Cor. 3.7.11, if the nets of operators are uniformly bounded, and if $\mathscr{A} = \mathscr{B}$, then $\lim_\alpha T_\alpha S_\alpha$ converges in SOT to TS.

Theorem 3.7.13. *Let $(T_\alpha)_{\alpha\in\mathscr{A}}$ be a net in $\mathfrak{L}(V, W)$ satisfying $\sup_\alpha \|T_\alpha\| < +\infty$ and converging in SOT to $T \in \mathfrak{L}(V, W)$. Let $(\xi_\beta)_{\beta\in\mathscr{B}}$ be a net in V converging to $\xi \in V$. Then the double net $(T_\alpha\xi_\beta)_{(\alpha,\beta)\in\mathscr{A}\times\mathscr{B}}$ converges to $T\xi$, that is,*

$$\lim_{(\alpha,\beta)\in\mathscr{A}\times\mathscr{B}} T_\alpha\xi_\beta = T\xi$$

In particular, if $\mathscr{A} = \mathscr{B}$, since $(T_\alpha\xi_\alpha)_{\alpha\in\mathscr{A}}$ is a subnet of $(T_\alpha\xi_\beta)_{(\alpha,\beta)\in\mathscr{A}\times\mathscr{A}}$ (defined by $\alpha \in \mathscr{A} \mapsto (\alpha, \alpha) \in \mathscr{A} \times \mathscr{A}$), we conclude that

$$\lim_\alpha T_\alpha\xi_\alpha = T\xi$$

Proof. Let $C = \sup_\alpha \|T_\alpha\|$, which is finite. We compute that

$$\|T\xi - T_\alpha\xi_\beta\| \leqslant \|T\xi - T_\alpha\xi\| + \|T_\alpha\xi - T_\alpha\xi_\beta\| \leqslant \|T\xi - T_\alpha\xi\| + C\|\xi - \xi_\beta\|$$

where the RHS converges to 0 under $\lim_{\alpha,\beta}$. □

Corollary 3.7.14. *Let $V_0, V_1, \dots, V_k$ be inner product spaces. For each $1 \leqslant i \leqslant k$, let $(T^i_{\alpha_i})_{\alpha_i\in\mathscr{A}_i}$ be a net in $\mathfrak{L}(V_{i-1}, V_i)$ converging in SOT to $T^i \in \mathfrak{L}(V_{i-1}, V_i)$. Assume that*

$$\sup_{\alpha_i\in\mathscr{A}_i} \|T^i_{\alpha_i}\| < +\infty \qquad \text{for all } 2 \leqslant i \leqslant k$$

Then for each $\xi \in V_0$, we have

$$\lim_{(\alpha_1,\dots,\alpha_k)\in\mathscr{A}_1\times\cdots\times\mathscr{A}_k} T^k_{\alpha_k}\cdots T^1_{\alpha_1}\xi = T^k\cdots T^1\xi$$

Similar to Thm. 3.7.13, if we also assume $\mathscr{A}_1 = \cdots = \mathscr{A}_k = \mathscr{A}$, then

$$\lim_{\alpha\in\mathscr{A}} T_\alpha^k \cdots T_\alpha^1 \xi = T^k \cdots T^1 \xi$$

Proof. This follows immediately from Thm. 3.7.13. □

Corollary 3.7.15. *Let $(T_\alpha)_{\alpha\in\mathscr{A}}$ be a net in $\mathfrak{L}(V)$ satisfying $\sup_\alpha \|T_\alpha\| < +\infty$ and converging in SOT to $T \in \mathfrak{L}(V)$. Then for each polynomial $f \in \mathbb{C}[z]$ and $\xi \in V$,*

$$\lim_\alpha f(T_\alpha)\xi = f(T)\xi$$

Proof. This is clear by Cor. 3.7.14. □

Example 3.7.16. In Exp. 3.7.12, since $\sup_{I\in\mathrm{fin}(2^X)} \|E_I\| = 1 < +\infty$, by Cor. 3.7.14 we have

$$\lim_{I\in\mathrm{fin}(2^X)} RP_I SP_I T\eta = RST\eta \qquad \text{for each } \eta \in \mathcal{H}$$

Example 3.7.17. Let $\mathcal{H}$ be a Hilbert space with basis $(e_x)_{x\in X}$. For each $I \in \mathrm{fin}(2^X)$, let P_I be the projection onto $U_I = \mathrm{Span}\{e_x : x \in I\}$. By Thm. 3.7.13, $\lim_I P_I T P_I$ converges in SOT to T, and

$$\sup_I \|P_I T P_I\| \leqslant \sup_I \|P_I\| \cdot \|T\| \cdot \|P_I\| \leqslant \|T\| < +\infty$$

Therefore, by Cor. 3.7.15, for each $T \in \mathfrak{L}(H)$ and $f \in \mathbb{C}[z]$,

$$\lim_{I\in\mathrm{fin}(2^X)} f(P_I T P_I) \qquad \text{converges in SOT to } f(T)$$

3.8 Problems

Let V be an inner product space.

Problem 3.1. Let ω be a Hermitian form on V. Prove the following sharpened polarization identity: For each $\xi, \eta \in V$ we have

$$\omega(\xi|\eta) = \frac{1}{4}\big(\omega(\xi+\eta|\xi+\eta) - \omega(\xi-\eta|\xi-\eta)\big) \tag{3.39}$$

Conclude that if $M \in \mathbb{R}_{\geqslant 0}$ and

$$|\omega(\xi|\xi)| \leqslant M\|\xi\|^2$$

for each $\xi \in V$, then $\|\omega\| \leqslant M$. This sharpens Prop. 3.2.12 in the Hermitian case.

Hint. First show $|\omega(\xi|\eta)| \leqslant (\|\xi\|^2 + \|\eta\|^2)/2$. □

Definition 3.8.1. Let V be an inner product space. A **Hilbert space completion** denotes a linear isometry $\Phi : V \to \mathcal{H}$ such that $\mathcal{H}$ is a Hilbert space and $\Phi(V)$ is dense in $\mathcal{H}$. Replacing V with $\Phi(V)$, one may view V as a dense inner product subspace of $\mathcal{H}$

† **Problem 3.2.** Prove that any inner product space V admits a Hilbert space completion.

Note that when V admits an orthonormal basis (e.g. when V is separable), the problem is directly solved by Cor. 3.3.26.

Hint. Let $\mathcal{H}$ be the metric space completion of V. Define vector additions, scalar multiplications, and the sesquilinear form $\langle\cdot|\cdot\rangle$ in an appropriate way, and prove that $\langle\cdot|\cdot\rangle$ is positive definite. $\square$

† **Problem 3.3.** Let $\mathcal{H}$ be a Hilbert space. Prove that any two orthonormal bases of $\mathcal{H}$ have the same cardinality. This cardinality is called the **dimension** of $\mathcal{H}$.

Hint. If E is an infinite set, then E and $\mathbb{N} \times E$ have the same cardinality. $\square$

4 The polynomial moment problem: a prehistory of spectral theory

4.1 Divergent series and the birth of the Stieltjes integral

4.1.1 Divergent series and the Padé approximation

In 1894, Stieltjes introduced the Stieltjes integral in [Sti94] (see [Sti-C, Vol. II] for an English translation) for the purpose of studying continued fractions. One key motivation for investigating continued fractions was to better understand the behavior of the series

$$\frac{c_0}{z} + \frac{c_1}{z^2} + \frac{c_2}{z^3} + \cdots \qquad \text{where each } c_n \geqslant 0 \tag{4.1}$$

in the divergent case—that is, when $\sum_n c_n r^n = +\infty$ for each $r > 0$. The core idea is that even when the power series in (4.1) diverges, it may still define a meaningful function outside a closed interval $I \subset \mathbb{R}$, provided we adopt a different notion of convergence.

Specifically, under suitable conditions on the sequence $(c_n)_{n\in\mathbb{N}}$, one can construct a sequence of rational functions $(f_n(z))_{n\in\mathbb{N}}$,

$$f_n(z) = \frac{q_n(z)}{p_{n+1}(z)} \qquad \text{where } p_{n+1}, q_n \in \mathbb{C}[z], \deg p_{n+1} = n+1, \deg q_n = n \tag{4.2}$$

Moreover, when $|z|$ is sufficiently large, $f_n(z)$ has Laurent expansion

$$f_n(z) = \sum_{m\in\mathbb{N}} c_{n,m} z^{-m-1} \qquad \text{with } c_{n,m} = c_m \text{ when } m \leqslant 2n+1 \tag{4.3a}$$

In other words,

$$f_n(z) = \frac{c_0}{z} + \frac{c_1}{z^2} + \cdots + \frac{c_{2n+1}}{z^{2n+2}} + \frac{?}{z^{2n+3}} + \cdots \tag{4.3b}$$

The sequence (f_n) (or a subsequence thereof) converges locally uniformly on $\mathbb{C}\backslash I$ to a holomorphic function f. This approximation is referred to as a **Padé approximation**. In such cases, we say that the holomorphic function $f(z)$ represents the series (4.1) and write

$$f(z) \sim \frac{c_0}{z} + \frac{c_1}{z^2} + \frac{c_2}{z^3} + \cdots$$

4.1.2 Stieltjes integral as the weak-* completion of finite sum

With the help of Padé approximation, we can understand how the Stieltjes integral naturally arises as the weak-* completion of finite sums. As we will see in the following sections, under suitable assumptions on the sequence (c_n), each rational function $f_n(z)$ has only simple poles. Consequently, $f_n(z)$ admits the representation

$$f_n(z) = \sum_i \frac{a_{n,i}}{z - \lambda_{n,i}} \tag{4.4}$$

where the sum is finite. Moreover, we have $a_{n,i} \geqslant 0$ and $\lambda_{n_i} \in I$.

The general Stieltjes integral appears when one tries to understand the behavior of the finite sum on the RHS of (4.4) under the limit $n \to +\infty$. To understand this behavior, we define an increasing function $\rho_n : I \to \mathbb{R}_{\geqslant 0}$ by setting

$$\rho_n(x) = \sum_{\lambda_{n,i} \leqslant x} a_{n,i}$$

That is, ρ_n is the right-continuous increasing function associated to the measure $\sum_i a_{n,i}\delta_{\lambda_{n,i}}$. Then (4.4) can be rewritten as

$$f_n(z) = \int_I \frac{d\rho_n(x)}{z - x}$$

The sequence (ρ_n) is indeed uniformly bounded. Therefore, by passing to a subsequence if necessary, we may assume that (ρ_n) almost converges to some increasing right-continuous function ρ. Therefore, by Thm. 2.9.6, (ρ_n) converges weak-* to ρ. It follows that (f_n) converges locally uniformly on $\mathbb{C}\backslash I$ to

$$f(z) = \int_I \frac{d\rho(x)}{z - x} \tag{4.5}$$

This gives a holomorphic function f on $\mathbb{C}\backslash I$ representing the series 2.9.3.

Definition 4.1.1. For each increasing function $\rho : I \to \mathbb{R}_{\geqslant 0}$, the function $f(z)$ defined by (4.5) is called the **Stieltjes transform** of ρ. If μ is a finite Borel measure on I, we also call the function

$$\mathbb{C}\backslash I \to \mathbb{C} \qquad z \mapsto \int_I \frac{d\mu(x)}{z - x}$$

the **Stieltjes transform** of μ.

We have thus seen that, via Padé approximation, the Stieltjes integral with respect to a general increasing function emerges as the (weak-*) limit of finite sums. This marks the first historical appearance of approximating a continuous spectrum by discrete spectra.

In the next section, we will see that the approximation of ρ by the sequence (ρ_n) can be interpreted as the finite-rank approximation of a (not necessarily bounded) Hermitian operator. Since this type of approximation plays a central role in the development of spectral theory by Hilbert and Riesz, it is crucial to understand its origin in the study of divergent series and its connection with Padé approximation.

4.1.3 The polynomial moment problem

We now investigate the following question: What assumption should we impose on $(c_n)_{n\in\mathbb{N}}$ so that the above strategy can be carried out? Note that when $|z| > |\lambda_{n,i}|$, we have

$$(z-\lambda_{n,i})^{-1} = \sum_{m\in\mathbb{N}} z^{-m-1}(\lambda_{n,i})^m$$

Therefore, (4.4) becomes $f_n(z) = \sum_{m\in\mathbb{N}}\sum_i z^{-m-1}\cdot a_{n,i}(\lambda_{n,i})^m$, and hence

$$f_n(z) = \sum_{m\in\mathbb{N}} z^{-m-1}\cdot \int_I x^m d\rho_n$$

Comparing this with (4.3), we obtain

$$c_{n,m} = \int_I x^m d\rho_n$$

On the one hand, (4.3) shows that $\lim_n c_{n,m} = c_n$. On the other hand, the weak-* convergence of $(d\rho_n)$ to $d\rho$ actually implies that $\lim_n \int_I x^m d\rho_n(\lambda)$ converges to $\int_I x^m d\rho(\lambda)$; see the proof of Thm. 4.2.19. Thus, we obtain

$$c_m = \int_I x^m d\rho \qquad \text{for all } m\in\mathbb{N} \tag{4.6}$$

Therefore, a necessary condition for the above strategy to work is the existence of an increasing function ρ satisfying (4.6). As we shall see in the next section, this condition is also sufficient. In this way, the problem of representing a divergent series and approximating it using rational functions is closely related to the polynomial moment problem.

4.2 Padé approximation via the finite-rank approximation of Hermitian operators

In this section, we let $I\subset\mathbb{R}$ be a proper closed interval. For simplicity, we assume that I is one of the following intervals

$$\mathbb{R} \qquad [0,+\infty) \qquad [0,1]$$

Let $\mathscr{R}r(I)$ be the set of functions $\rho: I\to\mathbb{R}_{\geqslant 0}$ satisfying condition (a) of the Riesz representation Thm. 1.10.1. That is,

$$\begin{aligned}\mathscr{R}r(I) = \{&\text{Increasing right continuous function } \rho: I\to\mathbb{R}_{\geqslant 0}\\ &\text{satisfying } \lim_{x\to-\infty}\rho(x) = 0 \text{ if } I=\mathbb{R}\}\end{aligned} \tag{4.7}$$

Fix a family $(c_n)_{n\in\mathbb{N}}$ in $\mathbb{R}$.

4.2.1 The goal

Problem 4.2.1 (Polynomial moment problem). Does there exists $\rho \in \mathcal{R}r(I)$ such that

$$c_n = \int_I x^n d\rho \qquad \text{for each } n \in \mathbb{N} \tag{4.8}$$

where the RHS is integrable, i.e., $\int_I |x|^n d\rho < +\infty$? Depending on whether I is $\mathbb{R}$, $\mathbb{R}_{\geqslant 0}$, or $[0,1]$, this problem is referred to as the **Hamburger moment problem**, the **Stieltjes moment problem**, or the **Hausdorff moment problem**, respectively.

The goal of this section is to give a complete solution of this problem; see Thm. 4.2.9. Moreover, when the moment problem is solvable for the given sequence (c_n), we will find $\rho \in \mathcal{R}r(I)$ whose Stieltjes transform

$$f(z) = \int_I \frac{d\rho(x)}{z-x}$$

represents the series $\sum_{n=0}^{\infty} c_n z^{-n-1}$ in the sense of Padé approximation. That is, there is a sequence of rational functions $(f_n(z))_{n\in\mathbb{N}}$ satisfying (4.3), and a subsequence $(f_{n_k}(z))$ converging locally uniformly on $\mathbb{C}\backslash I$ to $f(z)$. [1] Thus, the problem of representing the (possibly divergent) series $\sum_{n=0}^{\infty} c_n z^{-n-1}$ for such (c_n) is solved. See Thm. 4.2.20.

The classical construction of the sequence (f_n) relies crucially on the idea of orthogonal polynomials. The approach presented in this section reformulates that classical method using the language of inner product spaces—a modern reinterpretation shaped by the development of spectral theory in Hilbert spaces. Of course, this reformulation is a retrospective abstraction that emerged only after the development of spectral theory in Hilbert spaces. In this section, we adopt this modern perspective, while aiming to present it in a way that remains mindful of its historical origins. In the following sections, we will explain how this approach connects to continued fractions and the classical formulation using orthogonal polynomials.

4.2.2 Preliminary

Let us clarify the meaning of (4.8), since we have so far defined $\int_I f d\rho$ only for $f \in C_c(I)$.

Definition 4.2.2. Let f be a Borel function from I to $\mathbb{C}$ or $\overline{\mathbb{R}}_{\geqslant 0}$. Let $\rho \in \mathcal{R}r(I)$. Let μ_ρ be the finite Borel measure on I associated to ρ as in the Riesz representation

[1] When the solution to the moment problem is unique, the original sequence (f_n) indeed converges locally uniformly to f. See Pb. 7.14 for a criterion for the uniqueness of solutions to the moment problem.

Thm. 1.10.1. We define the **Stieltjes integral** $\int_I f d\rho$ to be

$$\int_I f d\rho := \int_I f d\mu_\rho$$

provided that the integral on the RHS exists.

Remark 4.2.3. When $f : I \to \mathbb{R}_{\geqslant 0}$ is continuous, the computation of the integral $\int_I f d\rho$ can be reduced to those of compactedly supported continuous functions. Indeed, for each $\lambda \geqslant 0$, let $\beta_\lambda : \mathbb{R} \to [0,1]$ be the (continuous) piecewise linear functions such that

$$\beta_\lambda|_{[-\lambda,\lambda]} = 1 \qquad \beta_\lambda|_{(-\infty,-\lambda-1]\cup[\lambda+1,+\infty)} = 0$$

Then $f\beta_\lambda \in C_c(I)$, and $\int_I f\beta_\lambda d\rho$ is increasing as λ increases. Therefore, by MCT (cf. Thm. 1.2.36), we have

$$\int_I f d\rho = \lim_{\lambda\to+\infty} \int_I f\beta_\lambda d\rho$$

4.2.3 The Hankel matrix

Definition 4.2.4. Let $H \in \mathbb{C}^{\mathbb{N}\times\mathbb{N}}$. Associate to H the unique sesquilinear form

$$\langle\cdot|\cdot\rangle : \mathbb{C}[x] \times \mathbb{C}[x] \to \mathbb{C} \qquad (x^m, x^n) \mapsto H(m,n)$$

We say that H is **Hermitian** (resp. **positive**, **positive definite**) if $\langle\cdot|\cdot\rangle$ is Hermitian (resp. positive (semi-definite), positive definite).

For each $n \in \mathbb{N}$, the ***n*-th truncation H_n** of H is the $(n+1) \times (n+1)$ matrix defined by the first $(n+1)$ rows and columns of H.

Remark 4.2.5. By linear algebra, H is positive definite iff $\det H_n > 0$ for each n; H is positive (semi-definite) iff the determinant of each principal submatrix is $\geqslant 0$.

Definition 4.2.6. The **Hankel matrix H** of $(c_n)_{n\in\mathbb{N}}$ is defined by

$$H(m,n) = c_{m+n}$$

That is,

$$H = \begin{pmatrix} c_0 & c_1 & c_2 & c_3 & \cdots \\ c_1 & c_2 & c_3 & c_4 & \cdots \\ c_2 & c_3 & c_4 & c_5 & \cdots \\ c_3 & c_4 & c_5 & c_6 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Equivalently, the associated sesquilinear form is determined by

$$\big\langle f\big|g\big\rangle = \big\langle 1\big|\overline{f}g\big\rangle \qquad \text{for each } f,g\in\mathbb{C}[x] \tag{4.9a}$$

$$\big\langle 1\big|x^n\big\rangle = c_n \qquad \text{for each } n\in\mathbb{N} \tag{4.9b}$$

Since each c_n is real, H is Hermitian.

We also let $H'\in\mathbb{R}^{\mathbb{N}\times\mathbb{N}}$ be the defined by $H'(m,n)=c_{m+n+1}$, that is,

$$H' = \begin{pmatrix} c_1 & c_2 & c_3 & c_4 & \cdots \\ c_2 & c_3 & c_4 & c_5 & \cdots \\ c_3 & c_4 & c_5 & c_6 & \cdots \\ c_4 & c_5 & c_6 & c_7 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

In other words, its associated sesquilinear form is

$$\mathbb{C}[x]\times\mathbb{C}[x]\to\mathbb{C} \qquad (f,g)\mapsto\langle f|xg\rangle \tag{4.10}$$

Since $c_n\in\mathbb{R}$, H' is also Hermitian. It follows that

$$\mathbb{C}[x]\times\mathbb{C}[x]\to\mathbb{C} \qquad (f,g)\mapsto\langle f|(1-x)g\rangle \tag{4.11}$$

is the (Hermitian) sesquilinear form associated to $H-H'$. □

Remark 4.2.7. Note that by (4.9a), for each $f,g,h\in\mathbb{C}[x]$ we have

$$\big\langle f\big|hg\big\rangle = \big\langle \overline{h}f\big|g\big\rangle \tag{4.12}$$

since $\langle f|hg\rangle = \langle 1|\overline{f}hg\rangle = \langle 1|\overline{\overline{h}f}g\rangle = \langle \overline{h}f|g\rangle$.

The construction of H and H' can be understood in a more general context; see Rem. 4.6.1.

In the rest of this section, we always let H denote the Hankel matrix of $(c_n)_{n\in\mathbb{N}}$, and equip $\mathbb{C}[x]$ with the sesquilinear form associated to H.

Proposition 4.2.8. *Suppose that Pb. 4.2.1 has a solution ρ. Then H must be positive, and the following are equivalent:*

(1) H is positive definite.

(2) The range of the solution ρ is not a finite set (equivalently, the associated measure μ_ρ is not supported in a finite set, cf. Lem. 1.6.5).

Moreover, if $I=[0,+\infty)$, then H' is positve; if $I=[0,1]$, then both H' and $H-H'$ are positive.

Proof. Let ρ be a solution of Pb. 4.2.1. Then for each $f,g\in\mathbb{C}[x]$, we have

$$\langle f|g\rangle=\int_I \overline{f}gd\rho \tag{4.13}$$

since this is clearly true when $f=x^m,g=x^n$. Thus $\langle f|f\rangle=\int_I |f|^2d\rho\geqslant 0$.

If H is not non-degenerated, then there exists $0\neq f\in\mathbb{C}[x]$ such that $\langle f|f\rangle=0$, and hence $\int_I |f|^2d\rho=0$. From this, one easily checks that $\mathrm{Supp}(\rho)$ is a subset of the finite set $f^{-1}(0)$. Conversely, if μ_ρ is supported in a finite set $E\subset I$, we choose a non-zero $f\in\mathbb{C}[x]$ such that $f|_E=0$. Then $\langle f|f\rangle=\int_I|f|^2d\mu_\rho=0$. This shows that H is not non-degenerate. This proves the equivalence of (1) and (2).

If $I=[0,+\infty)$, then for each $f\in\mathbb{C}[x]$, since $x|f|^2\geqslant 0$ on I, we have

$$\langle f|xf\rangle=\int_{[0,+\infty)}x|f|^2d\rho\geqslant 0$$

Therefore, H' is positive. Similarly, if $I=[0,1]$, then for each $f\in\mathbb{C}[x]$, both $x|f|^2$ and $(1-x)|f|^2$ are $\geqslant 0$ on I. Therefore

$$\langle f|xf\rangle=\int_{[0,1]}x|f|^2d\rho\geqslant 0\qquad \langle f|(1-x)f\rangle=\int_{[0,1]}(1-x)|f|^2d\rho\geqslant 0$$

This proves that both H' and $H-H'$ are positive. □

Therefore, to solve the polynomial moment problem, we should at least assume that H is positive. Indeed, Prop. 4.2.8 implies a half of the following theorem.

Theorem 4.2.9. *Let H be the Hankel matrix for $(c_n)_{n\in\mathbb{N}}$. The following are true*

1. *The **Hamburger moment problem** (i.e. $I=\mathbb{R}$) has a solution iff H is positive.*
2. *The **Stieltjes moment problem** (i.e. $I=[0,+\infty)$) has a solution iff H,H' are positive.*
3. *The **Hausdorff moment problem** (i.e. $I=[0,1]$) has a solution iff $H,H',H-H'$ are positive.*

Proof. The direction "⇒" follows from Prop. 4.2.8. The direction "⇒" will follow immediately from Thm. 4.2.19 once the latter has been established. □

4.2.4 The inner product space $V=\mathbb{C}[x]/\mathscr{N}$ associated to the positive Hankel matrix H

Assume that H is positive. In this subsection, we express the positivity of $H,H',H-H'$ in terms of the positivity of certain Hermitian operator T and the operator $1-T$ on an inner product space V.

Theorem 4.2.10. *Let $\langle\cdot|\cdot\rangle$ be the positive sesquilinear form on $\mathbb{C}[x]$ defined by the Hankel matrix H, which descends to an inner product on*

$$V = \mathbb{C}[x]/\mathscr{N}$$

where

$$\mathscr{N} := \{f \in \mathbb{C}[x] : \|f\|^2 = 0\}$$

is a linear subspace of $\mathbb{C}[x]$ due to Cor. 3.1.10. Then there is a (necessarily unique) linear map satisfying

$$T : V \to V \qquad f + \mathscr{N} \mapsto xf + \mathscr{N}$$

Moreover, ω_T is Hermitian, that is, $\langle\xi|T\eta\rangle = \langle T\xi|\eta\rangle$ for each $\xi, \eta \in V$.

Proof. If $f \in \mathscr{N}$, by Cauchy-Schwarz (cf. Thm. 3.1.9) and Rem. 4.2.7,

$$0 \leqslant \langle xf|xf\rangle = \langle x^2 f|f\rangle \leqslant \|x^2 f\| \cdot \|f\| = 0$$

and hence $xf \in \mathscr{N}$. Therefore, the linear map $f \in \mathbb{C}[x] \mapsto xf \in \mathbb{C}[x]$ descends to the linear map $T \in \mathrm{Lin}(V)$ sending each $f + \mathscr{N}$ to $xf + \mathscr{N}$. By Rem. 4.2.7, ω_T is Hermitian. □

Remark 4.2.11. Throughout this section, we set

$$\Omega := 1 + \mathscr{N}$$

which is a vector in V. Then for each $f, g \in \mathbb{C}[x]$, we clearly have

$$\langle f(T)\Omega|g(T)\Omega\rangle = \langle f|g\rangle \tag{4.14}$$

In particular, since $\langle 1|x^n\rangle = c_n$, we obtain

$$\langle\Omega|T^n\Omega\rangle = c_n \tag{4.15}$$

and hence

$$\|\Omega\|^2 = c_0 \tag{4.16}$$

By (4.10) and (4.11), we conclude that $\omega_T \geqslant 0$ iff $H' \geqslant 0$, and that $\omega_{\mathbf{1}-T} \geqslant 0$ iff $H - H' \geqslant 0$.

4.2.5 Solving the moment problem: the degenerate case

Assume $H \geqslant 0$.

In this subsection, we explain how the moment problem can be solved when $\dim V < +\infty$. This happens precisely when H is degenerate. Indeed, if H is degenerate, we can choose $f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0$ to be the polynomial with lowest degree satisfying $\langle f|f\rangle = 0$. Then $f(T)\Omega = 0$ by (4.14), and hence $T^k f(T)\Omega = 0$ for each $k \in \mathbb{N}$. It follows that V has basis $\Omega, T\Omega, \cdots, T^{n-1}\Omega$, and f is a minimal polynomial of T on V. Conversely, if V is finite-dimensional, then any $f \in \mathbb{C}[x]$ satisfying $f(T) = 0$ also satisfies $\langle f|f\rangle = 0$ due to (4.14).

Assume $\dim V < +\infty$. Then T is a Hermitian operator on the finite-dimensional inner product space V, and hence can be diagonalized. More precisely, there exists an orthonormal basis $e_1, \ldots, e_n$ of V such that $Te_j = \lambda_j e_j$ for all $1 \leqslant j \leqslant n$, where $\lambda_j \in \mathbb{R}$. Write

$$\Omega = a_1 e_1 + \cdots + a_n e_n$$

where $a_j \in \mathbb{C}$. Then, for each $f \in \mathbb{C}[x]$, we have

$$f(T)\Omega = \sum_{j=1}^n a_j f(\lambda_j) e_j$$

and hence

$$\langle \Omega | f(T)\Omega\rangle = \sum_{j=1}^n f(\lambda_j)|a_j|^2 = \int_{\mathbb{R}} f d\rho$$

where $\rho : \mathbb{R} \to \mathbb{R}_{\geqslant 0}$ is an increasing function corresponding to the measure $\mu = \sum_{j=1}^n |a_j|^2 \delta_{\lambda_j}$. In particular, by (4.15), we conclude

$$\int_{\mathbb{R}} x^n d\rho = \langle \Omega | T^n \Omega\rangle = c_n$$

Thus, the Hamburger moment problem is solved.

Next, assume $I = [0, +\infty)$ and $H' \geqslant 0$. By Rem. 4.2.11, T is positive, and hence each λ_j belongs to $[0, +\infty)$. Therefore, the measure μ constructed in the above paragraph is supported in $[0, +\infty)$, and hence $\int_{[0,+\infty)} x^n d\rho = c_n$. This solves the Stieltjes moment problem.

Finally, assume $I = [0, 1]$ and $H', H - H'$ are positive. By Rem. 4.2.11, T and $1 - T$ are positive. Therefore, each λ_j belongs to $[0, 1]$, and hence $\int_{[0,1]} x^n d\rho = c_n$. This solves the Hausdorff moment problem.

To summarize, we have proved Thm. 4.2.9 when $\dim V < +\infty$, equivalently, when H is positive but not positive definite.

4.2.6 Finite-rank approximation of T

In this subsection, we assume the following condition.

Condition 4.2.12. *Assume that $H \geqslant 0$. Moreover:*

- *If $I = [0, +\infty)$, assume that $H' \geqslant 0$ (equivalently, $\omega_T \geqslant 0$, cf. Rem. 4.2.11).*
- *If $I = [0, 1]$, assume that $H', H - H'$ are both positive (equivalently, ω_T and ω_{1-T} are positive, cf. Rem. 4.2.11).*

In this subsection, we construct a uniformly bounded sequence (ρ_n) in $\mathscr{R}r(I)$ whose Stieltjes transforms—or those of a subsequence—will provide the Padé approximation of a function representing the series $\sum_{n\in\mathbb{N}} c_n z^{-n-1}$.

Definition 4.2.13. Let T be as in Thm. 4.2.10. For each $n \in \mathbb{N}$, let

$$E_n \text{ be the projection of } V \text{ onto } V_n := \mathrm{Span}\{\Omega, T\Omega, \dots, T^n\Omega\}$$

Then we clearly have

$$\Omega \in V_0 \qquad TV_n \subset V_{n+1} \tag{4.17}$$

Remark 4.2.14. Since ω_T and ω_{E_n} are Hermitian (Cor. 3.3.20), for each n and $\xi \in V$ we have

$$\langle \xi | E_n T E_n \xi\rangle = \langle E_n\xi | T E_n \xi\rangle = \omega_T(E_n\xi|E_n\xi) \geqslant 0$$

It follows that $\omega_{E_nTE_n}$ is Hermitian. Similarly, if $H' \geqslant 0$ (equivalently, $\omega_T \geqslant 0$), then

$$\langle \xi | E_n T E_n \xi\rangle = \langle E_n\xi | T E_n \xi\rangle \geqslant 0$$

and hence $\omega_{E_nTE_n} \geqslant 0$. If both H' and $H - H'$ are positive, then

$$\langle \xi | E_n T E_n \xi\rangle = \langle E_n\xi | T E_n \xi\rangle \geqslant 0 \qquad \langle \xi | E_n(1-T)E_n\xi\rangle = \langle E_n\xi|(1-T)E_n\xi\rangle \geqslant 0$$

and hence $\omega_{E_nTE_n}, \omega_{E_n(1-T)E_n}$ are positive.

Definition 4.2.15. For each $n \in \mathbb{N}$, define a family $(c_{n,m})_{m\in\mathbb{N}}$ in $\mathbb{R}$ by

$$c_{n,m} = \langle \Omega | (E_nTE_n)^m\Omega\rangle \tag{4.18}$$

Remark 4.2.16. We have

$$c_{n,m} = c_m \qquad \text{if } m \leqslant 2n+1 \tag{4.19}$$

Consequently,

$$\lim_{n\to\infty} c_{n,m} = c_m \tag{4.20}$$

Proof. By (4.17), we have

$$(E_nTE_n)^k\Omega = T^k\Omega \in V_k \qquad \text{if } 0 \leqslant k \leqslant n \tag{4.21a}$$

$$(E_nTE_n)^{n+1}\Omega = E_nT^{n+1}\Omega \tag{4.21b}$$

where the first line is proved by induction on k. Thus, for each $m \leqslant 2n+1$, writing $m = a + b$ where $a, b \in \mathbb{N}$ and $a \leqslant n$ and $b \leqslant n+1$, we have

$$\begin{aligned}&\langle\Omega|(E_nTE_n)^m\Omega\rangle = \langle(E_nTE_n)^a\Omega|(E_nTE_n)^b\Omega\rangle = \langle T^a\Omega|E_nT^b\Omega\rangle = \langle E_nT^a\Omega|T^b\Omega\rangle\\ =&\langle T^a\Omega|T^b\Omega\rangle = \langle\Omega|T^{a+b}\Omega\rangle = c_m\end{aligned}$$

□

Proposition 4.2.17. *For each $n \in \mathbb{N}$, there exists $\rho_n \in \mathscr{R}r(I)$ such that $\mathrm{Rng}(\rho_n)$ is a finite subset of $[0, c_0]$, and that for all $m \in \mathbb{N}$ we have*

$$c_{n,m} = \int_I x^m d\rho_n \tag{4.22}$$

Proof. We view E_nTE_n as a linear operator T_n on V_n, i.e.,

$$T_n := E_nTE_n\big|_{V_n} \tag{4.23}$$

By Rem. 4.2.14, T_n is Hermitian. Therefore, by linear algebra, V_n has an orthonormal basis $e_{n,0}, e_{n,1}, \dots$ such that

$$T_ne_{n,i} = \lambda_{n,i}e_{n,i} \qquad \text{for all } i \tag{4.24}$$

where $\lambda_{n,i} \in \mathbb{R}$. Moreover, by Rem. 4.2.14, if $I = \mathbb{R}_{\geqslant 0}$ then $T_n \geqslant 0$, and hence $\lambda_{n,i} \geqslant 0$; if $I = [0,1]$ then $0 \leqslant T_n \leqslant \mathrm{id}_{V_n}$, and hence $0 \leqslant \lambda_{n,i} \leqslant 1$. It follows that

$$\lambda_{n,i} \in I$$

in all cases. Write

$$\Omega = \sum_i a_{n,i}e_{n,i} \tag{4.25}$$

where $a_n \in \mathbb{C}$. (So $a_{n,i} = \langle e_{n,i}|\Omega\rangle$.) Thus

$$c_{n,m} = \langle\Omega|(T_n)^m\Omega\rangle = \sum_i(\lambda_{n,i})^m \cdot |a_{n,i}|^2 = \int_I x^m d\rho_n$$

where ρ_n is the unique element in $\mathscr{R}r(I)$ whose associated finite Borel measure is $\sum_i |a_{n,i}|^2\delta_{\lambda_{n,i}}$, that is,

$$\rho_n(x) = \sum_{\substack{\text{all } i \text{ satisfying}\\ \lambda_{n,i}\leqslant x}} |a_{n,i}|^2 \tag{4.26}$$

In particular, by Parseval's identity, we have

$$0 \leqslant \rho_n(x) \leqslant \sum_i |a_{n,i}|^2 = \|\Omega\|^2 \xlongequal{(4.16)} c_0$$

□

4.2.7 From discrete spectra to continuous spectra

We continue to assume Condition 4.2.12. In this subsection, we solve the moment problem (i.e., complete the proof of Thm. 4.2.9) by proving Thm. 4.2.19.

Definition 4.2.18. Let $(\rho_n)_{n\in\mathbb{N}}$ be the uniformly bounded sequence in $\mathcal{Rr}(I)$ described by Prop. 4.2.17. By Helly's selection Thm. 2.9.3, (ρ_n) has a subsequence $(\rho_{n_k})_{k\in\mathbb{N}}$ converging pointwise to some increasing function $\widetilde{\rho}: I \to \mathbb{R}_{\geqslant 0}$. Define $\rho \in \mathcal{Rr}(I)$ as follows.

- If I is $\mathbb{R}_{\geqslant 0}$ or $[0,1]$, we let ρ be the right-continuous normalization of $\widetilde{\rho}$.
- If I is $\mathbb{R}$, we let ρ be the right-continuous normalization of $\widetilde{\rho} - \lim_{x\to-\infty}\widetilde{\rho}(x)$.

Then by Thm. 1.9.12 and Rem. 1.9.13, $d\rho$ and $d\widetilde{\rho}$ represent the same element of $C_c(I,\mathbb{F})^*$. By Thm. 2.9.6, the sequence $(d\rho_{n_k})$ converges weak-* to $d\rho$.

Theorem 4.2.19. *Assume Condition 4.2.12, and let $\rho \in \mathcal{Rr}(I)$ be as in Def. 4.2.18. Then for each $m \in \mathbb{N}$, we have*

$$c_m = \int_I x^m d\rho \tag{4.27}$$

where the RHS is integrable.

Proof. The easiest case is where $I = [0,1]$. In that case, $x^m \in C_c(I)$. Therefore, by the weak-* convergence of $(d\rho_{n_k})$ to $d\rho$,

$$\int_I x^m d\rho = \lim_k \int_I x^m d\rho_{n_k} \overset{(4.22)}{=\!=} \lim_k c_{n_k,m} \overset{(4.20)}{=\!=} c_m$$

Next, we consider the case where $I = \mathbb{R}$. Similar to the above argument, we have

$$\lim_k \int_{\mathbb{R}} x^m d\rho_{n_k} = c_m$$

Therefore, to prove (4.27), it suffices to prove that x^m is $d\rho$-integrable and

$$\lim_k \int_{\mathbb{R}} x^m d\rho_{n_k} = \int_{\mathbb{R}} x^m d\rho \tag{4.28}$$

Let β_λ be as in Rem. 4.2.3. By the weak-* convergence, we have

$$\lim_k \int_{\mathbb{R}} \beta_\lambda x^m d\rho_{n_k} = \int_{\mathbb{R}} \beta_\lambda x^m d\rho \tag{4.29}$$

for each $\lambda \geqslant 0$. Therefore, to prove (4.28) as well as $\int_{\mathbb{R}} |x^m| d\rho < +\infty$, due to

$$0 \leqslant 1 - \beta_\lambda \leqslant \chi_{J_\lambda} \qquad \text{where } J_\lambda = (-\infty, -\lambda] \cup [\lambda, +\infty)$$

it suffices to prove that for each $m \in \mathbb{N}$ and $\varepsilon > 0$,

$$\sup_n \int_{J_\lambda} |x^m| d\rho_n \leqslant \varepsilon \qquad \text{for sufficiently large } \lambda \tag{4.30a}$$

$$\int_{J_\lambda} |x^m| d\rho \leqslant \varepsilon \qquad \text{for sufficiently large } \lambda \tag{4.30b}$$

Since $|x^m| \leqslant 1 + |x^{m+1}|$, it suffices to prove (4.30) when m is an even number (and hence $|x^m| = x^m$).

When m is even, observe that $\lambda^2 \chi_{J_\lambda} \cdot x^m \leqslant x^{m+2}$, and hence

$$\int_{J_\lambda} x^m d\rho_n \leqslant \lambda^{-2} \int_{\mathbb{R}} x^{m+2} d\rho_n = \lambda^{-2} c_{n,m+2}$$

where the $\limsup_n$ of the RHS equals $\lambda^{-2} c_{m+2}$ due to (4.20). This proves (4.30a). By (4.29),

$$\int_{\mathbb{R}} \beta_\lambda x^m d\rho = \lim_k \int_{\mathbb{R}} \beta_\lambda x^m d\rho_{n_k} \leqslant \lim_k \int_{\mathbb{R}} x^m d\rho_{n_k} = \lim_k c_{n_k,m} \overset{(4.20)}{=\!=} c_m$$

Applying MCT to $\lim_{\lambda \to +\infty}$, we obtain $\int_{\mathbb{R}} x^m d\rho \leqslant c_m < +\infty$. DCT implies (4.30b).

The proof for the case that $I = \mathbb{R}_{\geqslant 0}$ is similar and is left to the reader. This case is even simpler, since $x^m \geqslant 0$ for all m. □

4.2.8 Padé approximation and the representation of divergence series

We continue to assume Condition 4.2.12. Let $(\rho_n)_{n \in \mathbb{N}}$ be described by Prop. 4.2.17. Let (ρ_{n_k}) be a subsequence as in Def. 4.2.18. That is, there exists $\rho \in \mathscr{R}r(I)$ such that $(d\rho_{n_k})$ converges weak-* to $d\rho$.

Theorem 4.2.20. *Let $f_n, f : \mathbb{C} \backslash I \to \mathbb{C}$ be the Stieltjes transforms of ρ_n and ρ, respectively. That is,*

$$f_n(z) = \int_I \frac{d\rho_n(x)}{z - x} \qquad f(x) = \int_I \frac{d\rho(x)}{z - x}$$

Then each $f_n(z)$ has the Laurent expansion

$$f_n(z) = \frac{c_0}{z} + \frac{c_1}{z^2} + \cdots + \frac{c_{2n+1}}{z^{2n+2}} + \frac{?}{z^{2n+3}} + \cdots \tag{4.31}$$

when $|z|$ is sufficiently large. Moreover, (f_{n_k}) converges locally uniformly on $\mathbb{C} \backslash I$ to f.

Therefore, the subsequence (f_{n_k}) provides a Padé approximation of f, and hence

$$f \sim \frac{c_0}{z} + \frac{c_1}{z^2} + \frac{c_2}{z^3} + \cdots$$

Proof. Recall from Prop. 4.2.17 that $c_{m,n} = \int_I x^m d\rho_n$. Therefore, by the argument in Subsec. 4.1.3, $f_n(z)$ has the Laurent expansion $\sum_n c_{m,n} z^{-n-1}$. Combined with Rem. 4.2.16, this establishes (4.31).

For each $z \in \mathbb{C}\backslash I$, let $g_z \in C_0(I)$ be defined by $g_z(x) = (z-x)^{-1}$. For each $z \in \mathbb{C}\backslash I$, one easily checks that $\lim_{\zeta\to z} g_\zeta$ converges uniformly on I to g_z. Thus, we have a continuous map

$$\Phi : \mathbb{C}\backslash I \to C_0(I) \qquad z \mapsto g_z$$

where $C_0(I)$ is equipped with the l^∞-norm. Therefore, for each compact $K \subset \mathbb{C}\backslash I$, the family $\Phi(K)$ is compact in $C_0(I)$. Since $(d\rho_{n_k})$ converges weak-* to $d\rho$ as linear functionals on $C_0(I)$ (cf. Rem. 2.9.5), it follows from the following Thm. 4.2.21 that $(d\rho_{n_k})$ converges uniformly to $d\rho$ when evaluated on functions in $\Phi(K)$. This proves that (f_{n_k}) converges locally uniformly to f. □

Theorem 4.2.21. *Let $\mathcal{V}, \mathcal{W}$ be normed vector spaces. Let (T_α) be a net in $\mathfrak{L}(\mathcal{V},\mathcal{W})$ satisfying $\sup_\alpha \|T_\alpha\| < +\infty$ and converging pointwise to some $T \in \mathfrak{L}(\mathcal{V},\mathcal{W})$. Let K be a precompact subset of V. Then (T_α) converges uniformly on K to T. That is,*

$$\lim_\alpha \sup_{\xi\in K} \|T\xi - T_\alpha\xi\| = 0$$

Proof. Replacing K with $\overline{K}$, we assume that K is compact. Since $C := \sup_\alpha \|T_\alpha\| < +\infty$ is a uniform Lipschitz of (T_α), the family (T_α) is equicontinuous. Therefore, by Thm. 1.4.36, (T_α) converges uniformly on K to T. □

4.3 Padé approximation via orthogonal polynomials

4.3.1 The setting

Fix $I \in \{\mathbb{R}, \mathbb{R}_{\geqslant 0}, [0,1]\}$, and choose a sequence $(c_n)_{n\in\mathbb{N}}$ in $\mathbb{R}$ satisfying Condition 4.2.12. Moreover, we assume that the Hankel matrix H of (c_n) is positive-definite. Therefore, the triple (V, Ω, T) in Thm. 4.2.10 can be described as follows: V is the vector space $\mathbb{C}[x]$ together with the inner product determined by H, the cyclic vector Ω is chosen to be the constant 1, and T is the multiplication by x. We assume for simplicity that

$$c_0 = 1$$

Therefore, $\|\Omega\| = 1$.

Recall that E_n is the projection operator of V onto $V_n = \mathrm{Span}\{1, x, \dots, x^n\} = \mathrm{Span}\{\Omega, T\Omega, \dots, T^n\Omega\}$. By Rem. 4.2.14,

$$T_n := E_n T E_n\big|_{V_n}$$

is a self-adjoint operator on V_n. Note that

$$\dim V_n = n+1$$

Recall (4.7) for the meaning of $\mathscr{R}r(I)$. Let $(\rho_n)_{n\in\mathbb{N}}$ be the sequence in $\mathscr{R}r(I)$ constructed in the proof of Prop. 4.2.17. That is,

$$\rho_n(x) = \sum_{\substack{\text{all } i \text{ satisfying} \\ \lambda_{n,i}\leqslant x}} |\langle e_{n,i}|\Omega\rangle|^2 \tag{4.32a}$$

where $e_{n,0},\dots,e_{n,n}$ form an orthonormal basis of V_n such that

$$T_n e_{n,i} = \lambda_{n,i} e_{n,i} \qquad \text{for all } i \tag{4.32b}$$

and $\lambda_{n,i}\in I$. As in Thm. 4.2.20, we let $f_n(z)$ be the Stieltjes transform of ρ_n, i.e.,

$$f_n(z) = \int_I \frac{d\rho_n(x)}{z-x}$$

Since ρ_n has finite range, $f_n(z)$ is a rational function.

4.3.2 Expressing $f_n(z)$ as a quotient of determinants

As shown in Thm. 4.2.20, a subsequence of (f_n) forms a Padé approximation to a holomorphic function f representing the series $\sum_{n\in\mathbb{N}} c_n z^{-n-1}$. The goal of this section is to provide an elementary description of f_n in terms of orthogonal polynomials. This description not only offers insight into the historical development of Padé approximation but is also essential for connecting Padé approximation to continued fractions, as we will see in the next section.

Proposition 4.3.1. *For sufficiently large $|z|$, we have*

$$f_n(z) = \big\langle \Omega \big| (z-T_n)^{-1}\Omega \big\rangle$$

Proof. By (4.32), we have $(z-T_n)^{-1}e_{n,i} = (z-\lambda_{n,i})^{-1}e_{n,i}$. Hence, the relation $\Omega = \sum_{i=0}^n e_{n,i}\cdot\langle e_{n,i}|\Omega\rangle$ implies

$$\big\langle \Omega \big| (z-T_n)^{-1}\Omega \big\rangle = \sum_i |\langle e_{n,i}|\Omega\rangle|^2\cdot (z-\lambda_{n,i})^{-1} = \int_I \frac{d\rho_n(x)}{z-x}$$

□

Therefore, if we extend the unit vector Ω to an orthonormal basis of V_n, then by Cramer's rule, $f_n(z)$ can be expressed as a quotient

$$f_n(z) = \frac{\widetilde{q}_n(z)}{\widetilde{p}_{n+1}(z)} \tag{4.33}$$

where $\widetilde{p}_{n+1}(z) = \det(z - T_n)$, and $\widetilde{q}_n(z)$ is a minor of $z - T_n$ of order n. In particular, we have

$$\deg \widetilde{p}_{n+1} = n + 1 \qquad \deg \widetilde{q}_n = n$$

Remarkably, the sequence $(\widetilde{p}_n)_{n\in\mathbb{N}}$ turns out to be orthogonal polynomials, as we will explain below.

4.3.3 Orthogonal polynomials

Definition 4.3.2. Let $(p_n)_{n\in\mathbb{N}}$ be a sequence in $\mathbb{C}[x]$. We say that (p_n) are **orthogonal polynomials** (resp. **orthonormal polynomials**) with respect to $(c_n)_{n\in\mathbb{N}}$ if the following conditions are satisfied:

(1) $\deg p_n = n$.

(2) (p_n) is orthogonal (resp. orthonormal) in the inner product space $\mathbb{C}[x]$ defined by the Hankel matrix H of c_n. Equivalently, $(p_n(T)\Omega)_{n\in\mathbb{N}}$ is an orthonormal (resp. orthogonal) basis of V.

Unless otherwise stated, we also assume that

$$\text{the leading coefficient of } p_n \text{ is } > 0 \tag{4.34}$$

We say that p_n is **monic** if the leading coefficient of p_n is 1.

Remark 4.3.3. The orthonormal polynomials (p_n) are uniquely determined by (c_n), and can be constructed by the Gram-Schmidt process. Since $c_0 = 1$, it is clear that

$$p_0 = 1$$

Moreover, since each c_n is real, the Gram-Schmidt process indicates that all the coefficients of p_n are real numbers.

Similarly, the monic orthogonal polynomials $(\widetilde{p}_n)$ are uniquely determined by (c_n). □

Remark 4.3.4. Moreover, if $\rho \in \mathscr{R}\mathrm{r}(I)$ solves the polynomial moment Problem 4.2.1 for (c_n), then by (4.13), condition (2) of Def. 4.3.2 is equivalent to

$$\int_I \overline{p_m} p_n d\rho = \delta_{m,n} \qquad \text{for each } m, n \in \mathbb{N} \tag{4.35}$$

(But note that $\overline{p_m} = p_m$.) In that case, we also say that (ρ_n) are **orthonormal polynomials** with respect to ρ.

Theorem 4.3.5. *Let $\widetilde{p}_{n+1}(z) = \det(z - T_n)$ and $\widetilde{p}_0(z) = 1$. Then $(\widetilde{p}_n(x))_{n\in\mathbb{N}}$ are the (unique) monic orthogonal polynomials with respect to (c_n).*

Proof. We want to show that $\widetilde{p}_{n+1}$ is orthogonal to V_n, equivalently, that $\widetilde{p}_{n+1}(T)\Omega \perp V_n$.

Applying the Cayley-Hamilton theorem to T_n, we have $\widetilde{p}_{n+1}(T_n) = 0$, equivalently, $\widetilde{p}_{n+1}(E_nTE_n) = 0$. Therefore, if $\widetilde{p}_{n+1}(x) = \sum_{k=0}^{n+1}\gamma_k x^k$ where $\gamma_i \in \mathbb{R}$, then

$$\sum_{k=0}^{n+1}\gamma_k(E_nTE_n)^k\Omega = 0$$

Together with (4.21), this implies

$$\sum_{k=0}^{n+1}\gamma_k E_nT^k\Omega = 0$$

and hence $E_n\widetilde{p}_{n+1}(T)\Omega = 0$. This proves $\widetilde{p}_{n+1}(T)\Omega \perp V_n$. $\square$

Corollary 4.3.6. *We have*

$$\det(z - T_n) = \begin{vmatrix} c_0 & c_1 & \cdots & c_{n+1} \\ c_1 & c_2 & \cdots & c_{n+2} \\ \vdots & \vdots & \ddots & \vdots \\ c_n & c_{n+1} & \cdots & c_{2n+1} \\ 1 & z & \cdots & z^{n+1} \end{vmatrix} \tag{4.36}$$

Proof. Denote the RHS by $\widetilde{p}_{n+1}(z)$. Then $\widetilde{p}_{n+1}(z)$ is a monic polynomial of degree $n+1$. Moreover, if we let $D_{i,j}$ be the (i,j)-th minor of the determinant on the RHS of (4.36) (where $0 \leqslant i, j \leqslant n+1$), then for each $0 \leqslant k \leqslant n$, we have

$$\langle x^k|\widetilde{p}_{n+1}(x)\rangle = \sum_{i=0}^{n+1} c_{k+i}D_{n+1,i} = \begin{vmatrix} c_0 & c_1 & \cdots & c_{n+1} \\ c_1 & c_2 & \cdots & c_{n+2} \\ \vdots & \vdots & \ddots & \vdots \\ c_n & c_{n+1} & \cdots & c_{2n+1} \\ c_k & c_{1+k} & \cdots & c_{n+k} \end{vmatrix} = 0$$

This proves that $\widetilde{p}_{n+1} \perp V_n$. $\square$

4.3.4 The Jacobi matrix

We now study the numerator $\widetilde{q}_n(z)$ in (4.33). As discussed in Subsec. 4.3.2, $\widetilde{q}_n(z)$ is a minor of $z - T_n$ of order n under the orthonormal basis $p_0, \ldots, p_n$ of V_n. To compute this minor, let us find the matrix representation of T under this basis.

Definition 4.3.7. Let $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ be sequences in $\mathbb{R}$ with $a_n > 0$ for each n. Define $J, J^+ \in \mathbb{R}^{\mathbb{N}\times\mathbb{N}}$ by

$$J = \begin{pmatrix} b_0 & a_0 & 0 & 0 & 0 & \cdots \\ a_0 & b_1 & a_1 & 0 & 0 & \cdots \\ 0 & a_1 & b_2 & a_2 & 0 & \cdots \\ 0 & 0 & a_2 & b_3 & a_3 & \cdots \\ 0 & 0 & 0 & a_3 & b_4 & \ddots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \ddots \end{pmatrix} \qquad J^+ = \begin{pmatrix} b_1 & a_1 & 0 & 0 & 0 & \cdots \\ a_1 & b_2 & a_3 & 0 & 0 & \cdots \\ 0 & a_2 & b_3 & a_3 & 0 & \cdots \\ 0 & 0 & a_3 & b_4 & a_4 & \cdots \\ 0 & 0 & 0 & a_4 & b_5 & \ddots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \ddots \end{pmatrix}$$

The matrix J is called the **Jacobi matrix** for (a_n) and (b_n). We also call (a_n) and (b_n), the **off-diagonal sequence** and the **diagonal sequence** of J, respectively.

By this definition, J^+ is the Jacobi matrix for

$$(a_n)^+ = (a_1, a_2, \dots) \qquad (b_n)^+ = (b_1, b_2, \dots) \tag{4.37}$$

Theorem 4.3.8. *There is a bijection between:*

(1) A sequence $(c_n)_{n\in\mathbb{N}}$ in $\mathbb{R}$ where $c_0 = 1$, and the associated Hankel matrix H is positive-definite.[2]

(2) A pair of sequences $(a_n)_{n\in\mathbb{N}}, (b_n)_{n\in\mathbb{N}}$ in $\mathbb{R}$ where $a_n > 0$ for each n.

The bijection is described as follows.

- *Given (c_n) satisfying (1), let $(p_n)_{n\in\mathbb{N}}$ be the orthonormal polynomials with respect to (c_n). Then the Jacobi matrix J for (a_n) and (b_n) is the matrix representation of $T : V \to V$ under (p_n).*

- *Given (a_n) and (b_n) satisfying (2), then c_n is the $(0,0)$-entry (i.e., the top-left entry) of the n-th power J^n.*

The Jacobi matrix J for $(a_n), (b_n)$ is called the **Jacobi matrix associated to the Hankel matrix of (c_n)**.

Remark 4.3.9. The fact that J is the matrix representation of T under (p_n) can be made explicit as follows: The sequence $(p_n)_{n\in\mathbb{N}}$ satisfies the **three-term recurrence relation**

$$xp_n(x) = a_n p_{n+1}(x) + b_n p_n(x) + a_{n-1}p_{n-1}(x) \qquad \text{for all } n \in \mathbb{N} \tag{4.38a}$$

where we set $p_{-1}(x) = 0$, and let a_{-1} be any number. Rewriting this relation as

$$a_n p_{n+1}(x) = (x - b_n)p_n(x) - a_{n-1}p_{n-1}(x)$$

[2]In other words, (c_n) satisfies the assumptions in Subsec. 4.3.1 for the case $I = \mathbb{R}$.

and noting that

$$p_{-1}(x) = 0 \qquad p_0(x) = 1 \tag{4.38b}$$

we see that (p_n) is uniquely determined by (a_n) and (b_n) through (4.38).

Proof of Thm. 4.3.8. Step 1. Given (c_n), let J be the matrix representation of T with respect to the orthonormal polynomials (p_n). Since $\deg p_k = k$, we have

$$xp_n(x) = a_n p_{n+1}(x) + b_n x^n + ?x^{n-1} + \cdots + ?x + ? \tag{4.39}$$

where $a_{n+1} \in \mathbb{R}_{>0}$ and $b_n \in \mathbb{R}$. Therefore, J is of the form

$$J = \begin{pmatrix} b_0 & ? & ? & ? & ? & \cdots \\ a_0 & b_1 & ? & ? & ? & \cdots \\ 0 & a_1 & b_2 & ? & ? & \cdots \\ 0 & 0 & a_2 & b_3 & ? & \cdots \\ 0 & 0 & 0 & a_3 & b_4 & \ddots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \ddots \end{pmatrix}$$

Since ω_T is Hermitian, J must be of the form given in Def. 4.3.7. This establishes the map

$$(c_n) \mapsto (a_n), (b_n) \tag{4.40}$$

equivalently, the map $H \mapsto J$. Moreover, we compute

$$c_n = \langle 1|x^n\rangle = \langle \Omega|T^n\Omega\rangle = \langle p_0|J^n p_0\rangle$$

This shows that c_n is the $(0,0)$-entry of J^n. Consequently, the map (4.40) is injective.

Step 2. It remains to prove that the map (4.40) is surjective. Let J be the Jacobi matrix of (a_n) and (b_n) satisfying (2). Let $C_c(\mathbb{N})$ be the set of functions $\mathbb{N} \to \mathbb{C}$ with finite supports. The inner product on $C_c(\mathbb{N})$ is chosen to be the one inherited from that of $l^2(\mathbb{N})$. Then J can be viewed as a linear operator on $C_c(\mathbb{N})$. Moreover, it is clear that

$$J^n \chi_{\{0\}} = a_0 \cdots a_{n-1}\chi_{\{n\}} + ?\chi_{\{n-1\}} + \cdots + ?\chi_{\{0\}}$$

where $a_0 \cdots a_{n-1}$ is understood as 1 if $n = 0$. Therefore, $(J^n\chi_{\{0\}})_{n\in\mathbb{N}}$ is a basis of $C_c(\mathbb{N})$, and there exists $p_n \in \mathbb{R}[x]$ satisfying

$$p_n(x) = (a_0 \cdots a_{n-1})^{-1}x^n + ?x^{n-1} + \cdots + ?x + ? \tag{4.41a}$$
$$\chi_{\{n\}} = p_n(J)\chi_{\{0\}} \tag{4.41b}$$

We define

$$c_n = \langle \chi_{\{0\}} | J^n \chi_{\{0\}} \rangle \tag{4.42}$$

Then the sesquilinear form $\langle \cdot | \cdot \rangle$ on $V = \mathbb{C}[x]$ determined by (c_n) satisfies

$$\langle g|h\rangle = \langle 1|\overline{g}h\rangle \xlongequal{(4.42)} \langle \chi_{\{0\}}|\overline{g}(J)h(J)\chi_{\{0\}}\rangle = \langle g(J)\chi_{\{0\}}|h(J)\chi_{\{0\}}\rangle \tag{4.43}$$

for each $g, h \in \mathbb{C}[x]$. (Note that in the last equality of (4.43), we have used the fact that $\langle \eta | J\xi \rangle = \langle J\eta | \xi \rangle$ for each $\xi, \eta \in C_c(\mathbb{N})$.) Therefore, $\langle g|g\rangle \geqslant 0$. If $g \neq 0$, by the fact that $(J^n \chi_{\{0\}})_{n\in\mathbb{N}}$ is a basis, we have $g(J)\chi_{\{0\}} \neq 0$, and hence

$$\langle g|g\rangle = \langle g(J)\chi_{\{0\}}|g(J)\chi_{\{0\}}\rangle > 0$$

Therefore, $\langle \cdot | \cdot \rangle$ is positive-definite. Thus, (c_n) satisfies (1).

Let us show that the map (4.40) sends (c_n) to $(a_n), (b_n)$. By (4.43), we have a unitary map

$$\Phi : V = \mathbb{C}[x] \to C_c(\mathbb{N}) \qquad g = g(T)\Omega \mapsto g(J)\chi_{\{0\}} \tag{4.44}$$

By (4.41b), Φ sends $\chi_{\{n\}}$ to p_n. Therefore, since $(\chi_{\{n\}})$ is an orthonormal basis of $C_c(\mathbb{N})$, the sequence $(p_n)_{n\in\mathbb{N}}$ is an orthonormal basis of V. This, together with (4.41a), shows that (p_n) is the orthogonal polynomials with respect to (c_n).

By (4.44), for each $g \in \mathbb{C}[x]$ we have

$$\Phi^{-1}J\Phi g = \Phi^{-1}Jg(J)\chi_{\{0\}} = \Phi^{-1} \cdot (xg)(J)\chi_{\{0\}} = xg = Tg$$

Therefore $\Phi^{-1}J\Phi = T$ as linear operators on V. In other words, J is the matrix representation of T under (p_n). This finishes the proof that (4.40) sends (c_n) to $(a_n), (b_n)$. □

Corollary 4.3.10. *Let $(p_n)_{n\in\mathbb{N}}$ be the orthonormal polynomials with respect to (c_n), and let (a_n) be the off-diagonal sequence of the Jacobi matrix associated to (c_n). Then the leading coefficient of p_n is $(a_0 \cdots a_{n-1})^{-1}$, understood to be 1 if $n = 0$.*

Proof. This is clear from the proof of Thm. 4.3.8, especially from (4.41). □

Corollary 4.3.11. *Let (a_n) and (b_n) be sequences in $\mathbb{R}$ with $a_n > 0$. Let J be the associated Jacobi matrix. Let $(p_n)_{n\in\mathbb{N}}$ be the sequence of polynomials satisfying (4.38). Then (p_n) can be described by*

$$a_0 \cdots a_n \cdot p_{n+1}(z) = \det(z - J_{n+1}) \qquad p_0(z) = 1 \tag{4.45}$$

where $J_{n+1} \in \mathbb{R}^{(n+1)\times(n+1)}$ is the matrix formed by taking the first $n+1$ rows and columns of J.

Proof. By Thm. 4.3.8, J is the Jacobi matrix associated to the positive-definite Hankel matrix of a sequence (c_n) satisfying $c_0 = 1$, and (p_n) is the orthonormal polynomials with respect to (c_n). Therefore, J_{n+1} is the matrix representation of T_n with respect to the orthonormal basis $p_0, \dots, p_n$ of V_n. It follows from Thm. 4.3.5 that the monic orthonormal polynomials $(\widetilde{p}_n)$ satisfy $\widetilde{p}_{n+1}(z) = \det(z - J_{n+1})$. By Cor. 4.3.10, we have $\widetilde{p}_n = a_0 \cdots a_n p_{n+1}(z)$. This finishes the proof. □

4.3.5 The main theorem

Recall the sequence of rational functions $(f_n(z))_{n\in\mathbb{N}}$ as in Subsec. 4.3.1.

Theorem 4.3.12. *Let J be the Jacobi matrix associated to the Hankel matrix of (c_n), and let (b_n) and (a_n) be the diagonal and off-diagonal sequences of J, respectively. Choose sequences of polynomials $(p_n(x))_{n\in\mathbb{N}}$ and $(q_n(x))_{n\in\mathbb{N}}$ determined by*

$$xp_n(x) = a_n p_{n+1}(x) + b_n p_n(x) + a_{n-1}p_{n-1}(x) \qquad p_{-1}(x) = 0 \quad p_0(x) = 1 \tag{4.46}$$

$$xq_n(x) = a_{n+1} q_{n+1}(x) + b_{n+1} q_n(x) + a_n q_{n-1}(x) \qquad q_{-1}(x) = 0 \quad q_0(x) = \frac{1}{a_0} \tag{4.47}$$

for each $n \in \mathbb{N}$. Then

$$f_n(z) = \frac{q_n(z)}{p_{n+1}(z)} \tag{4.48}$$

Proof. Let J_{n+1} (resp. J_n^+) be the $(n+1)\times(n+1)$ (resp. $n\times n$) matrix formed by taking the first $n+1$ (resp. first n) rows and columns of J (resp. J^+). By the description of J in Thm. 4.3.8, J_n is the matrix representation of T_n with respect to the orthonormal basis $p_0,\dots,p_n$. Therefore, since $p_0 = 1 = \Omega$, by Prop. 4.3.1 and Cramer's rule (or the inverse matrix formula), we have

$$f_n(z) = \frac{\det(z - J_n^+)}{\det(z - J_{n+1})} \tag{4.49}$$

Applying Cor. 4.3.11 to the Jacobi matrix J (resp. J^+) and the sequence (p_n) (resp. $(a_0 q_n)$), we see that $\det(z - J_{n+1}) = a_0 \cdots a_n p_{n+1}$ (resp. $\det(z - J_n^+) = a_1 \cdots a_n \cdot a_0 q_n(z)$). This establishes (4.48). □

4.4 Padé approximation via continued fractions

We continue to work in the setting of Subsec. 4.3.1 and freely use the notations recalled there. In particular, we consider the sequence of rational functions $(f_n(z))_{n\in\mathbb{N}}$ described by $f_n(z) = \langle\Omega|(z - T_n)^{-1}\Omega\rangle$ (cf. Prop. 4.3.1), which admits a subsequence that Padé-approximates a holomorphic function f representing the (possibly) divergent series $\sum_{n\in\mathbb{N}} c_n z^{-n-1}$.

In this section, we use Thm. 4.3.12 to express (f_n) as finite approximations of a continued fraction. In [Sti94], Stieltjes's reasoning proceeds in the opposite direction: he begins with a continued fraction, derives the three-term recurrence relation satisfied by its finite approximants, and then uses orthogonality to obtain the Padé approximation.

Theorem 4.4.1. *Let J be the Jacobi matrix associated to the Hankel matrix of (c_n), and let (b_n) and (a_n) be the diagonal and off-diagonal sequences of J, respectively. Then $f_n(z)$ is the n-th approximation of the continued fraction*

$$\cfrac{1}{z-b_0-\cfrac{a_0^2}{z-b_1-\cfrac{a_1^2}{z-b_2-\cfrac{a_2^2}{z-b_3-\ddots}}}}$$

That is, for each $n \in \mathbb{N}$, we have

$$f_n(z)=\cfrac{1}{z-b_0-\cfrac{a_0^2}{\ddots-\cfrac{a_{n-2}^2}{z-b_{n-1}-\cfrac{a_{n-1}^2}{z-b_n}}}} \tag{4.50}$$

Recall from Thm. 4.3.8 that any pair of real sequences $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ with $a_n>0$ arises from a unique sequence $(c_n)_{n\in\mathbb{N}}$ with positive-definite Hankel matrix and $c_0=1$ in the manner described in Thm. 4.4.1.

Proof. By Thm. 4.3.12, we have $f_n=q_n/p_{n+1}$ where (p_n) and (q_n) satisfy

$$a_np_{n+1}(z)=(z-b_n)p_n(z)-a_{n-1}p_{n-1}(z) \qquad p_{-1}(z)=0 \quad p_0(z)=1 \tag{4.51a}$$

$$a_nq_n(z)=(z-b_n)q_{n-1}(z)-a_{n-1}q_{n-2}(z) \qquad q_{-1}(z)=0 \quad q_0(z)=\frac{1}{a_0} \tag{4.51b}$$

In particular, we have $p_1(z)=(z-b_0)/a_0$, and hence $f_0(z)=1/(z-b_0)$. This proves (4.50) for $n=0$. Note that (4.51b) originally holds only when $n>0$. However, by setting $a_{-1}=1$ and $q_{-2}(z)=-1$, Eq. (4.51) also holds when $n=0$.

We denote the RHS of (4.50) by Υ_n. We view $(p_n),(q_n),(\Upsilon_n)$ as sequences of rational functions of $z,a_0,b_0,a_1,b_1,\dots$. Assume that (4.50) holds for $n-1$ where $n\in\mathbb{Z}_+$, i.e.,

$$\frac{q_{n-1}}{p_n}=\Upsilon_{n-1}$$

Note that Υ_n is obtained from Υ_{n-1} by replacing b_{n-1} with $b_{n-1}+\frac{a_{n-1}^2}{z-b_n}$. To prove that (4.50) holds for n, it suffices to show that $a_np_{n+1}/(z-b_n)$ (resp. $a_nq_n/(z-b_n)$) is also obtained from p_n (resp. q_{n-1}) by replacing b_{n-1} with $b_{n-1}+\frac{a_{n-1}^2}{z-b_n}$.

By (4.51a), we have

$$p_n=a_{n-1}^{-1}(z-b_{n-1})p_{n-1}-a_{n-1}^{-1}a_{n-2}p_{n-2} \tag{4.52}$$

and p_{n-1} does not involve b_{n-1}. Replacing the b_{n-1} on the RHS of (4.52) with $b_{n-1}+\frac{a_{n-1}^2}{z-b_n}$, what we want to prove is

$$a_np_{n+1}/(z-b_n)=a_{n-1}^{-1}\Big(z-b_{n-1}-\frac{a_{n-1}^2}{z-b_n}\Big)p_{n-1}-a_{n-1}^{-1}a_{n-2}p_{n-2}$$

equivalently,

$$a_{n-1}a_np_{n+1}=\Big((z-b_{n-1})(z-b_n)-a_{n-1}^2\Big)p_{n-1}-a_{n-2}(z-b_n)p_{n-2}$$

But this follows from the computation

$$\begin{aligned}
a_{n-1}a_np_{n+1}&\xlongequal{(4.51a)}(z-b_n)\cdot a_{n-1}p_n-a_{n-1}^2p_{n-1}\\
&\xlongequal{(4.52)}(z-b_n)\cdot\big((z-b_{n-1})p_{n-1}-a_{n-2}p_{n-2}\big)-a_{n-1}^2p_{n-1}\\
&=\Big((z-b_{n-1})(z-b_n)-a_{n-1}^2\Big)p_{n-1}-a_{n-2}(z-b_n)p_{n-2}
\end{aligned}$$

This proves the desired property for $a_np_{n+1}/(z-b_n)$. A similar argument proves the desired property for $a_nq_n/(z-b_n)$. □

4.5 Application: an alternative proof of the Riesz representation theorem

In Sec. 2.2.1, we noted that solving a moment problem and characterizing a dual space are often equivalent tasks. For the Hamburger and Stieltjes moment problems, the interval I is non-compact, so there is no suitable norm on $C(I)$ that would allow us to reformulate the moment problem as a dual space problem. In contrast, the Hausdorff moment problem can be translated into such a characterization because I is compact in this case.

Since we have already solved all types of polynomial moment problems in Thm. 4.2.9, it is natural to expect that this theorem also yields an alternative proof of the main part of the Riesz representation Thm. 1.10.1 for a compact interval I, namely, the classification of positive linear functionals on $C(I)$. This is exactly what we will do in this section.

For that purpose, Thm. 2.4.2 must be adapted so that the moment-problem interpretation of the classification of *bounded* linear functionals (discussed in Sec. 2.2.1 and relying crucially on Thm. 2.4.2) admits an analogue for *positive* linear functionals. This analogue is stated as Thm. 4.5.2. To prove it we first establish a preliminary result which abstracts the elementary fact that any positive integral operator is bounded whenever the constant function 1 is integrable.

Proposition 4.5.1. *Let X be a set. Let $\mathscr{A}$ be a unital *-$\mathbb{F}$-subalgebra of $l^\infty(X,\mathbb{F})$. Let $\Lambda:\mathscr{A}\to\mathbb{F}$ be a linear map which is **positive** in the sense that $\Lambda(f)\geqslant 0$ for each $f\geqslant 0$. Then Λ is bounded with operator norm $\leqslant\Lambda(1)$.*

Recall from Exp. 1.2.37 that the involution $*$ on $l^\infty(X,\mathbb{F})$ is defined by $f^* = \overline{f}$.

Proof. In the case that $\mathbb{F}=\mathbb{R}$, since $-\|f\|_{l^\infty}\leqslant f\leqslant \|f\|_{l^\infty}$, and since $\Lambda(a)=a\Lambda(1)$ for each scalar a, we obtain $-\|f\|_{l^\infty}\cdot\Lambda(1)\leqslant\Lambda(f)\leqslant\|f\|_{l^\infty}\cdot\Lambda(1)$, and hence $|\Lambda(f)|\leqslant\|f\|_{l^\infty}\cdot\Lambda(1)$.

In the case that $\mathbb{F}=\mathbb{C}$, by the positivity of Λ, the sesequilinear form

$$\mathscr{A}\times\mathscr{A}\rightarrow\mathbb{C}\qquad (f,g)\mapsto\Lambda(f^*g)$$

is positive. Therefore, by Cauchy Schwarz,

$$|\Lambda(f)|^2\leqslant\Lambda(f^*f)\Lambda(1)\leqslant\|f\|_{l^\infty}^2\cdot\Lambda(1)^2$$

where the last inequality is due to $f^*f\leqslant\|f\|_{l^\infty}^2$. □

Theorem 4.5.2. *Let X be a set. Let $\mathscr{A}$ be a unital *-$\mathbb{F}$-subalgebra of $l^\infty(X,\mathbb{F})$ with l^∞-closure $\overline{\mathscr{A}}$. Suppose that for each $f\in\overline{\mathscr{A}}$ satisfying $f\geqslant 0$, there exists $g\in\overline{\mathscr{A}}$ such that $f=\overline{g}g$ (i.e. $f=g^*g$). Then we have an $\mathbb{R}_{\geqslant 0}$-linear isomorphism*

$$\begin{gathered}\{\textit{positive linear functionals on }\overline{\mathscr{A}}\}\xrightarrow{\simeq}\{\textit{positive linear functionals on }\mathscr{A}\}\\ \Lambda\mapsto\Lambda|_{\mathscr{A}}\end{gathered}\tag{4.53}$$

Note that $\overline{\mathscr{A}}$ is also a unital *-$\mathbb{F}$-subalgebra of $l^\infty(X,\mathbb{F})$.

Proof. By Prop. 4.5.1, positive linear functionals on $\overline{\mathscr{A}}$ are bounded. Therefore, by Thm. 2.4.2, they are determined by their restrictions to $\mathscr{A}$. Hence the map (4.53) is injective.

To prove that (4.53) is surjective, we pick any positive linear functional $\Lambda:\mathscr{A}\rightarrow\mathbb{F}$, which is bounded (by Prop. 4.5.1) and hence can be extended to a bounded linear functional $\Lambda:\overline{\mathscr{A}}\rightarrow\mathbb{F}$ (by Thm. 2.4.2). Suppose that $f\in\overline{\mathscr{A}}$ satisfies $f\geqslant 0$. By assumption, $f=g^*g$ for some $g\in\overline{\mathscr{A}}$. Therefore, there is a sequence (g_n) in $\mathscr{A}$ converging uniformly to g. So $(g_n^*g_n)$ converges uniformly to f. Since Λ is continuous, and since $\Lambda(g_n^*g_n)\geqslant 0$, we conclude that $\Lambda(f)\geqslant 0$. This proves that $\Lambda:\overline{\mathscr{A}}\rightarrow\mathbb{F}$ is a positive linear functional. □

Theorem 4.5.3 (Riesz representation theorem). *Let I be a compact interval. Then positive linear functionals on $C(I)$ are precisely linear functionals of the form*

$$C(I)\rightarrow\mathbb{C}\qquad f\mapsto\int_I f\,d\rho$$

where $\rho:I\rightarrow\mathbb{R}_{\geqslant 0}$ is increasing.

Recall from Thm. 1.9.12 that replacing ρ with its right-continuous normalization does not change the values of the Stieltjes integrals. Therefore, in the above theorem, one may assume that ρ is right-continuous.

Proof. Clearly, any increasing function $\rho : I \to \mathbb{R}_{\geqslant 0}$ defines a positive linear functional. Conversely, let $\Lambda : C(I) \to \mathbb{C}$ be a positive linear functional. We assume WLOG that $I = [0,1]$. Define a sequence $(c_n)_{n\in\mathbb{N}}$ in $\mathbb{R}$ by

$$c_n = \Lambda(x^n)$$

Then, in view of Def. 4.2.6, the Hankel matrix H of (c_n) determines the sesquilinear form $\langle\cdot|\cdot\rangle$ on $\mathbb{C}[x]$ satisfying $\langle 1|f\rangle = \Lambda(f)$ for all $f \in \mathbb{C}[x]$, and hence

$$\langle f|g\rangle = \langle 1|f^*g\rangle = \Lambda(f^*g)$$

for all $f, g \in \mathbb{C}[x]$. Therefore, the positivity of Λ implies that $\langle\cdot|\cdot\rangle$ is positive. Moreover, for each $f \in \mathbb{C}[x]$ we have

$$\langle xf|f\rangle = \Lambda(xf^*f) \geqslant 0 \qquad \langle (1-x)f|f\rangle = \Lambda((1-x)f^*f) \geqslant 0$$

because xf^*f and $(1-x)f^*f$ belong to $C([0,1], \mathbb{R}_{\geqslant 0})$. Therefore, (c_n) satisfies the assumption of the Hausdorff moment problem. Hence, by Thm. 4.2.9, there exists an increasing $\rho : [0,1] \to \mathbb{R}_{\geqslant 0}$ such that $c_n = \int_I x^n d\rho$ for all n, and hence

$$\Lambda(f) = \int_I f d\rho$$

for all $f \in \mathbb{C}[x]$. In other words, the positive linear operators Λ and $\int_I d\rho$ agree on $\mathbb{C}[x]$. By Thm. 4.5.2, they agree on the closure of $\mathbb{C}[x]$, which is $C(I)$ by Stone-Weierstrass. $\square$

Remark 4.5.4. The above proof of the Riesz representation theorem aligns perfectly with Table 2.3. The reason is that the core of the proof of Thm. 4.5.3 is the solution of the polynomial moment problem, namely, the proof of Thm. 4.2.9. Our proof of Thm. 4.2.9 in Sec. 4.2 (especially Subsec. 4.2.7) is an excellent illustration of Table 2.3: Principle 2.2.10 is verified by approximating the linear functional

$$\Lambda : \mathbb{C}[x] \to \mathbb{C} \qquad x^m \mapsto c_m$$

with a sequence (ρ_n) of increasing functions (or a subsequence thereof), where each ρ_n (defined in (4.26)) has finite range. The equivalence between pointwise convergence of functions and convergence of moments—that is, the equivalence of the two shaded areas in Table 2.3—is captured by Thm. 2.9.6, which is invoked in Def. 4.2.18 in the process of solving the polynomial moment problem.

It is worth noting that our use of (ρ_n) to approximate Λ is an instance of approximating the infinite by the finite. This is not only because each ρ_n has finite range, but also because the definition of ρ_n arises from the diagonalization of the finite-rank operator T_n given in (4.23). In other words, approximating Λ by ρ_n is, at its core, an approximation of T (described in Thm. 4.2.10) by T_n.

Riesz, in contrast, proved the Riesz representation theorem using the method of linear extension rather than finite approximation. As we will discuss in Ch. 5, Riesz's treatment of the Riesz representation theorem and the spectral theorem of bounded self-adjoint operators marked a paradigm shift in functional analysis: the transition from finite approximation to linear extension. This paradigm shift will be one of the key themes of this course. □

4.6 Problems

Problem 4.1. Let $\mathscr{A}$ be a unital *-algebra. Let $\Lambda : \mathscr{A} \to \mathbb{C}$ be a linear map satisfying $\Lambda(x^*x) \geqslant 0$ for each $x \in \mathscr{A}$. In other words, the sesquilinear form

$$\langle\cdot|\cdot\rangle : \mathscr{A} \times \mathscr{A} \to \mathbb{C} \qquad (x,y) \mapsto \Lambda(x^*y) \tag{4.54}$$

is positive (semidefinite).

1. Let $V = \mathscr{A}/\mathscr{N}$ where $\mathscr{N} = \{g \in \mathscr{A} : \langle g|g\rangle = 0\}$ (cf. Cor. 3.1.10). Recall from Exe. 3.2.2 that $\langle\cdot|\cdot\rangle$ descends to an inner product of V. Prove that there is a linear map satisfying
$$\pi : \mathscr{A} \to \mathrm{Lin}(V) \qquad \pi(f)(g + \mathscr{N}) = fg + \mathscr{N}$$
2. Set $\Omega = 1_{\mathscr{A}} + \mathscr{N}$, which is an element of V. Prove that (π, V) is a pre-unitary representation of $\mathscr{A}$ (cf. Def. 5.5.3) satisfying
$$\begin{gathered} \pi(\mathscr{A})\Omega = V \qquad \|\Omega\|^2 = \Lambda(1_{\mathscr{A}}) \\ \langle\Omega|\pi(f)\Omega\rangle = \Lambda(f) \quad \text{for each } f \in \mathscr{A} \end{gathered} \tag{4.55}$$
3. Suppose that (π', V') is a pre-unitary representation of $\mathscr{A}$ (where V' is an inner product space), and $\Omega' \in V'$ satisfies
$$\begin{gathered} \pi'(\mathscr{A})\Omega' = V' \qquad \|\Omega'\|^2 = \Lambda(1_{\mathscr{A}}) \\ \langle\Omega'|\pi'(f)\Omega'\rangle = \Lambda(f) \quad \text{for each } f \in \mathscr{A} \end{gathered} \tag{4.56}$$
Prove that there exists a unique unitary map $\Phi : V \to V'$ such that
$$\Phi\Omega = \Omega' \qquad \Phi\pi(f) = \pi'(f)\Phi \quad \text{for each } f \in \mathscr{A}$$

Note. Compare this problem with the proof of Thm. 4.2.10. □

Remark 4.6.1. The pre-unitary representation (π, V) of $\mathscr{A}$ constructed in Pb. 4.1 is commonly known as the **GNS construction** (after Gelfand-Naimark-Segal). It may be regarded as an abstraction of the construction of the Hankel matrix H and the associated matrix H' from the moments (c_n) in Sec. 4.2.

To see how the GNS construction generalizes the classical setting, define $\Lambda : \mathbb{C}[x] \to \mathbb{C}$ to be the unique linear functional with $\Lambda(x^n) = c_n$ for each $n \in \mathbb{N}$. Then the sesquilinear form (4.54) is precisely the one whose Gram matrix, with respect to the basis $1, x, x^2, \ldots$, is the Hankel matrix H defined in Def. 4.2.6. Similarly, the matrix H', also from Def. 4.2.6, is the matrix representation of $\pi(x)$ under the same basis. $\square$

5 The spectral theorem for bounded self-adjoint operators

5.1 Hilbert's spectral theorem

5.1.1 Introduction

After Stieltjes' pioneering work on continued fractions [Sti94], the Stieltjes integral once again came into prominence with Hilbert's spectral theorem for bounded symmetric bilinear forms [Hil06]. It is fair to say that without Hilbert's discovery of the spectral theorem—and its subsequent refinement by later mathematicians, most notably F. Riesz—the Stieltjes integral might never have become the central and influential concept it is today. The reason modern readers are often unfamiliar with the Stieltjes integral is simply that its theory has been fully absorbed into modern measure theory. One should not forget that Stieltjes integrals are equivalent to integrals over intervals with respect to finite Borel measures.

The formulation of the spectral theorem for bounded self-adjoint operators on Hilbert spaces has undergone significant evolution throughout history. In this chapter, we will encounter four versions:

- Hilbert's original version.
- Riesz's version.
- The Borel functional calculus version.
- The multiplication operator version.

The relationships among these formulations are not immediate, nor are they obviously equivalent. In fact, to meaningfully compare them, we must take a practical perspective, that is, consider the problems that these versions can solve or help illuminate. Among them, the Borel functional calculus version and the multiplication operator version are the more modern and practically useful formulations. However, to fully appreciate their significance, we must not overlook their historical development, particularly Hilbert's and Riesz's versions, as well as the background of the polynomial moment problem discussed in the previous chapter.

In [Hil06], Hilbert introduced the Hilbert space $l^2(\mathbb{Z})$ and proved the spectral theorem for bounded Hermitian forms on it [Hil06, Satz 31]. We first present a formulation of this theorem, slightly adapted to modern terminology, and then provide some comments.

5.1.2 The spectral theorem of Hilbert

Theorem 5.1.1 (Hilbert's spectral theorem). *Let V be a separable inner product space, and let $\omega \in \mathfrak{Ses}(V)$ be Hermitian. Choose $r \geqslant 0$ such that $\|\omega\| \leqslant r$. Then for each $\xi \in V$, there is an increasing function $\rho_\xi : [-r,r] \to \mathbb{R}_{\geqslant 0}$ such that for $z \in \mathbb{C}$ satisfying $|z| > r$, the resolvent form $(z-\omega)^{-1} \in \mathfrak{Ses}(V)$ satisfies*

$$(z-\omega)^{-1}(\xi|\xi) = \int_{[-r,r]} \frac{d\rho_\xi(\lambda)}{z-\lambda} \tag{5.1}$$

The resolvent form $(z-\omega)^{-1}$ will be defined in the proof, and is related to the resolvent operator introduced in Cor. 3.5.16. The precise relationship will be discussed after the proof.

Proof. Let $u_1, u_2, \dots$ be an orthonormal basis of V. Similar to Def. 4.2.13, we let

$$V_n = \mathrm{Span}_{\mathbb{C}}\{u_1, \dots, u_n\}$$

We shall first establish (5.1) for ξ in a dense subset of V, and then extend it to all $\xi \in V$.

Step 1. Similar to (4.23), we let $\omega_n \in \mathfrak{Ses}(V_n)$ be the restriction of ω to $V_n \times V_n$. Then ω_n is Hermitian, and satisfies $\|\omega_n\| \leqslant r$. Hence $\omega_n = \omega_{T_n}$ for some self-adjoint $T_n \in \mathrm{Lin}(V_n)$, that is,

$$\omega_n(\xi|\eta) = \langle \xi | T_n \eta \rangle \qquad \text{for each } \xi, \eta \in V_n$$

and $\|T_n\| \leqslant r$.

By the spectral theorem for Hermitian operators on finite-dimensional inner product spaces, T_n is diagonalizable. Therefore, there exist $\lambda_{n,1}, \dots, \lambda_{n,n} \in \mathbb{R}$ and an orthonormal basis $e_{n,1}, \dots, e_{n,n}$ of V_n such that

$$T_n e_{n,i} = \lambda_{n,i} e_{n,i} \qquad \text{for each } 1 \leqslant i \leqslant n$$

Since $\|T\| \leqslant r$, we have $|\lambda_{n,i}| = \|T e_{n,i}\| \leqslant \|T\| \leqslant r$, and hence

$$\lambda_{n,1}, \dots, \lambda_{n,n} \in [-r, r]$$

Step 2. Let $\mathbb{K} = \mathbb{Q} + \mathbf{i}\mathbb{Q}$. Let

$$U = \mathrm{Span}_{\mathbb{K}}\{u_1, u_2, \dots\}$$

which is a dense $\mathbb{K}$-linear subspace of V with countable cardinality. In this step, we define $(z-\omega)^{-1}$ as a bounded $\mathbb{K}$-sesquilinear form on U admitting an integral representation as in (5.1). Roughly speaking, $(z-\omega)^{-1}$ on U will be defined to be the limit of a convergent subsequence of the sesquilinear forms associated to $(z-T_n)^{-1}$. The details are as follows.

For each $\xi \in U$, we have $V_n \ni \xi$ for sufficiently large n. Moreover, for such n, we clearly have $(z-T_n)^{-1}\xi = \sum_{i=1}^n (z-\lambda_{n,i})^{-1}e_{n,i}\cdot\langle e_{n,i}|\xi\rangle$, and hence

$$\langle \xi|(z-T_n)^{-1}\xi\rangle = \sum_{i=1}^n (z-\lambda_{n,i})^{-1}\langle \xi|e_{n,i}\rangle\cdot\langle e_{n,i}|\xi\rangle \tag{5.2}$$

Similar to (4.26), for each $\xi \in U$, we let $\rho_{\xi,n}:[-r,r]\to\mathbb{R}_{\geqslant 0}$ be

$$\rho_{\xi,n}(x) = \sum_{\substack{\text{all } i \text{ satisfying}\\ \lambda_{n,i}\leqslant x}} |\langle e_{n,i}|\xi\rangle|^2 \qquad \text{for all } V_n \ni \xi \tag{5.3}$$

We set $\rho_{\xi,n} = 0$ if $V_n \not\ni \xi$. Then $(\rho_{\xi,n})_{n\in\mathbb{Z}_+}$ is a uniformly bounded sequence, since Bessel's inequality implies $0 \leqslant \rho_{\xi,n}(x) \leqslant \|\xi\|^2$. It follows from (5.2) that

$$\langle \xi|(z-T_n)^{-1}\xi\rangle = \int_{[-r,r]} \frac{d\rho_{\xi,n}(\lambda)}{z-\lambda} \qquad \text{for all } n \text{ such that } \xi\in V_n$$

By Helly's selection Thm. 2.9.3 and the diagonal method (Rem. 1.4.18), there exist strictly positive integers $n_1 < n_2 < \dots$ such that [1]

$$\lim_k \rho_{\xi,n_k} \text{ converges pointwise to some function } \rho_\xi \qquad \text{for each } \xi\in U$$

with $0 \leqslant \rho_\xi \leqslant \|\xi\|^2$. By Thm. 2.9.6, for each $\xi \in U$, the sequence $(d\rho_{\xi,n_k})$ converges weak-* to $d\rho_\xi$ as positive linear functionals on $C([-r,r])$. Therefore

$$\lim_k \langle \xi|(z-T_{n_k})^{-1}\xi\rangle = \int_{[-r,r]} \frac{d\rho_\xi(\lambda)}{z-\lambda} \qquad \text{for all } \xi\in U \tag{5.4}$$

By the polarization identity (cf. Rem. 3.1.3), $\lim_k\langle \xi|(z-T_{n_k})^{-1}\eta\rangle$ converges for each $\xi,\eta\in U$. We can thus define a $\mathbb{K}$-sesquilinear form

$$(z-\omega)^{-1}: U\times U\to\mathbb{C} \qquad (z-\omega)^{-1}(\xi|\eta) = \lim_k \langle \xi|(z-T_{n_k})^{-1}\eta\rangle \tag{5.5}$$

By (5.4), the relation (5.1) is satisfied for all $\xi\in U$. Therefore, for each $\xi \in U$, since $0\leqslant\rho_\xi\leqslant\|\xi\|$, we have

$$\big|(z-\omega)^{-1}(\xi|\xi)\big| \leqslant \|\xi\|^2 / \inf_{\lambda\in[-r,r]} |z-\lambda| \tag{5.6}$$

---

[1] When Hilbert wrote [Hil06], the Helly selection theorem had not yet been discovered. His argument proceeded as follows. He first applied the Arzelà-Ascoli theorem to obtain a uniformly convergent subsequence of the antiderivatives of $(\rho_{\xi,n})$. The limit of this subsequence is a convex function, and its derivative (which exists outside a countable set) is then taken as the definition of ρ_ξ.

By the proof of Prop. 3.2.12, we conclude that $(z-\omega)^{-1}$ is a bounded $\mathbb{K}$-sesquilinear form.

Step 3. From the proof of Thm. 2.4.2, we know that $(z-\omega)^{-1}$ can be uniquely extended to a bounded ($\mathbb{C}$-)sesquilinear form on V. We know that (5.1) holds for all $\xi\in U$. Let us establish the integral representation (5.1) for any $\xi\in V$.

Suppose that $\xi\in V\backslash U$. Let (ξ_n) be a sequence in U converging to ξ. From the above proof, we know that $0\leqslant\rho_{\xi_n}\leqslant\|\xi_n\|$. In particular, the sequence (ρ_{ξ_n}) is bounded. By the Helly selection theorem, (ρ_{ξ_n}) has a subsequence $(\rho_{\xi_{n_k}})$ converging pointwise to some increasing function $\rho_\xi:[-r,r]\rightarrow\mathbb{R}_{\geqslant 0}$. By Thm. 2.9.6, $(d\rho_{\xi_{n_k}})$ converges to $d\rho_\xi$ when integrated against the function $\lambda\in[-r,r]\mapsto 1/(z-\lambda)$. Thus

$$
\begin{aligned}
&(z-\omega)^{-1}(\xi|\xi)=\lim_n(z-\omega)^{-1}(\xi_n|\xi_n)=\lim_k(z-\omega)^{-1}(\xi_{n_k}|\xi_{n_k})\\
=&\lim_k\int_{[-r,r]}\frac{d\rho_{\xi_{n_k}}(\lambda)}{z-\lambda}=\int_{[-r,r]}\frac{d\rho_\xi(\lambda)}{z-\lambda}
\end{aligned}
$$

This proves the existence of ρ_ξ satisfying (5.1). □

5.1.3 Q&A

We give some comments on Hilbert's spectral theorem (Thm. 5.1.1) in the form of Q&A.

Question 5.1.2. Assume that the inner product space V in Thm. 5.1.1 is a Hilbert space, so that there is a canonical isomorphism $\mathfrak{L}(V)\simeq\mathfrak{Ses}(V)$ (cf. Thm. 3.5.5). Write $\omega=\omega_T$ where $T\in\mathfrak{L}(V)$ is self-adjoint. Is $(z-\omega)^{-1}$ equal to the bounded sesquilinear form associated to the resolvent operator $(z-T)^{-1}$?

Answer. Yes, but this is not immediate. Let us first summarize how $(z-\omega)^{-1}$ is constructed in the proof of Thm. 5.1.1, now using the language of bounded operators.

Choose $r\in\mathbb{R}$ such that $r\geqslant\|\omega\|=\|T\|$. As in Def. 4.2.13, we let E_n be the projection of V onto $V_n=\{u_1,\dots,u_n\}$. Then E_nTE_n is a self-adjoint operator on V with operator norm $\leqslant r$. For $|z|>r$, the limit

$$
\lim_n(z-E_nTE_n)^{-1} \tag{5.7}
$$

indeed converges in SOT. However, even the WOT convergence of this limit was not known at the time of Hilbert's work [Hil06].

As seen in the proof of Thm. 5.1.1, Hilbert's idea was instead to show that for each $\xi\in V$, a subsequence of (5.7) converges when evaluated in $\langle\xi|-\xi\rangle$. One then selects a subsequence that converges simultaneously on a sufficiently large countable set of vectors, and finally uses the uniform boundedness of the operator norms of $(z - E_nTE_n)^{-1}$to conclude convergence on all pairs of vectors in V. (Compare this with Prop. 2.4.5.)

In summary, $(z-\omega)^{-1}$ is defined to be the bounded sesquilinear form obtained as the limit along a suitable subsequence:

$$(z-\omega)^{-1}(\xi|\eta) = \lim_k \langle \xi|(z - E_{n_k}TE_{n_k})^{-1}\eta\rangle \tag{5.8}$$

where the subsequence is independent of the choice of z, ξ, η.

Thus, the question reduces to whether $(z-\omega)^{-1}$ coincides with the bounded sesquilinear form associated to $(z-T)^{-1}$. The answer is yes, once one shows that

$$\lim_n (z - E_nTE_n)^{-1} = (z-T)^{-1} \tag{5.9}$$

in WOT. This will be proved in Rem. 5.11.2. □

Question 5.1.3. The proofs of Hilbert's spectral theorem (Thm. 5.1.1) and of the polynomial moment problems (Thm. 4.2.9) share many similarities. For instance, both make use of finite-rank approximations of Hermitian operators of the form $E_nTE_n \to T$; in both cases, the increasing function ρ_n for E_nTE_n is constructed by diagonalizing E_nTE_n; and in both, the increasing function ρ for T is obtained by taking a pointwise convergent subsequence of (ρ_n).

In view of these similarities, what exactly are the novelties in Hilbert's proof of his spectral theorem?

Answer. As noted in Sec. 4.4, Stieltjes' treatment of the polynomial moment problem did not rely on the diagonalization theory of linear algebra. He obtained Padé approximations not via finite-rank approximations of Hermitian operators, but through continued fractions and detailed analyses of determinants and polynomials. It was Hilbert's work that paved the way for the connection between the polynomial moment problem, inner product spaces, and spectral theory, thereby allowing us to approach polynomial moment problems from the perspective of spectral theory, as developed in Ch. 4.

I should also mention that although the resolvent form $(z-\omega)^{-1}$ considered by Hilbert bears a striking resemblance to the Stieltjes transforms arising in the polynomial moment problem with divergent series as background (see Thm. 4.2.20)—indeed, this similarity justifies viewing the resolvent of an operator T as its Stieltjes transform—the notion of resolvent in functional analysis actually first appeared in Fredholm's study of integral equations [Fre03].

In [Fre03], Fredholm sought to analyze the solutions of integral equations of the form

$$f(x) + \int_0^1 K(x,y)f(y) = g(y)$$

where K, g are given continuous functions and f is the unkown solution. Fredholm considered the resolvent of the integral operator $S : C([0, 1]) \to C([0, 1])$ defined by $(Sf)(x) = \int_0^1 K(x, y)f(y)$. As in Question 5.1.2, this resolvent was not defined directly as the inverse operator of $z-S$, but rather as the limit of $(z-S_n)^{-1}$, where (S_n) is a sequence of finite-rank matrices (with increasing ranks) obtained by partioning the interval $[0, 1]$. Moreover, the inverse $(z-S_n)^{-1}$ was expressed in terms of determinants, thanks to Cramer's rule/the inverse matrix formula. Fredholm studied the invertibility of $z - S$ by analyzing the zeros of the holomorphic function $\Delta(z) := \lim_n \det(z - S_n)$. See Subsec. 8.1.5 for further discussion.

Thus, another major novelty of Hilbert's proof of the spectral theorem was the way it connected Fredholm's notion of the resolvent, developed in the study of integral equations, with the Stieltjes transforms arising from the polynomial moment problem and divergent series. □

Question 5.1.4. I noticed that the more modern versions of the spectral theorem we will encounter later in this chapter (such as Riesz's version, the Borel functional calculus version, and the multiplication operator version) are more powerful and widely applicable than Hilbert's spectral theorem. In fact, Hilbert's version seems more like a special case of these modern results. So what is the significance of studying the proof of Hilbert's spectral theorem?

Answer. The proofs of modern spectral theorems share a common trait: they rely heavily on sophisticated algebraic machinery, often employing the Riesz representation theorem (for spaces of continuous functions) as a "black box" at critical junctures. Studying these proofs alone can leave learners puzzled:

- How did mathematicians first realize that the Riesz representation theorem could be used to prove spectral theorems?
- Why is it applied in this specific way?

While Hilbert's spectral theorem is less general than its modern counterparts, its proof avoids this complex abstraction. Instead, the connection with the Riesz representation theorem occupies most of the argument: As mentioned in Question 5.1.3, the proofs of Hilbert's spectral theorem (Thm. 5.1.1) and of the Hausdorff moment problem (cf. Thm. 4.2.9) run in close parallel. And as discussed in Subsec. 2.2.1 and Sec. 4.5, the latter problem is almost equivalent to classifying positive linear functionals on $C(I)$ for a compact interval I.

Thus, the real significance of Hilbert's spectral theorem is that it makes transparent why the Riesz representation theorem enters spectral theory in the first place.

Hilbert's proof of his spectral theorem, the earliest version of the spectral theorem, can be viewed as a linear-algebraic reinterpretation of all the key steps in the proof of the Hausdorff moment problem—and hence as an almost equivalent reformulation of the Riesz representation theorem for $C(I)$.

The additional layers found in modern spectral theorems—those not directly tied to the moment problem/Riesz representation paradigm—were introduced later, as part of the refinement and expansion of the theory. □

5.2 Towards Riesz's spectral theorem: projections

5.2.1 Hilbert's spectral theorem holds for inner product spaces

It is often said that one of the main differences between mathematicians and physicists in their approaches to the mathematics of quantum mechanics is that mathematicians stress the completeness of Hilbert spaces, emphasizing that they are more than just inner product spaces, whereas physicists find the notion of completeness largely irrelevant. Mathematicians commonly justify this emphasis by pointing out that the spectral theorem for self-adjoint operators requires completeness. This is certainly true for the spectral theorems developed after Hilbert. However, as we saw in Thm. 5.1.1, Hilbert's own spectral theorem already holds for general inner product spaces.—If this observation causes a degree of unease for the reader, then I have achieved my aim.

The fact that Hilbert's spectral theorem holds for all inner product spaces, while later versions hold only for Hilbert spaces, shows that Hilbert's version is less powerful and can therefore be established under weaker assumptions. Even so, completeness still plays a role in Hilbert's theorem—though not the Cauchy completeness of the inner product space. Rather, it is the weak-* completeness of Stieltjes integrals against increasing functions, already noted in Table 2.4. Since Hilbert's spectral theorem is the ancestor of all later versions, we may conclude that the truly central analytic condition underlying all spectral theorems is not the Cauchy completeness of Hilbert spaces (or, equivalently, the Riesz isomorphism $\mathcal{H} \simeq (\mathcal{H}^{\complement})^*$, cf. Thm. 3.5.3), but instead the weak-* completeness of increasing functions/finite Borel measures.

5.2.2 Why Riesz's spectral theorem requires Hilbert spaces

Riesz's spectral theorem, proved in [Rie13, Ch. V], is a significant improvement over Hilbert's. Beginning with this section, we prepare for the introduction of Riesz's version.

An important drawback of Hilbert's spectral theorem is that it is unclear how the increasing function ρ_ξ in Thm. 5.1.1 relies on ξ. The first highlight of Riesz's spectral theorem is that, under the assumption that $\mathcal{H}$ is a Hilbert space and $T \in \mathfrak{L}(\mathcal{H})$ is self-adjoint with $\|T\| \leqslant r$, he realizes ρ_ξ through an increasing net of projections $(E(\lambda))_{\lambda\in[-r,r]}$ associated to T (called the **spectral projections** of T), more precisely:

$$\langle\xi|E(\lambda)\xi\rangle = \rho_\xi(\lambda) \qquad \text{for all } \lambda \in [-r,r] \text{ and } \xi \in \mathcal{H}$$

As we will learn in this section, for each Hilbert space $\mathcal{H}$ there is an order-preserving bijection between projection operators and closed linear subspaces of $\mathcal{H}$, related by $P \mapsto \mathrm{Rng}(P)$. Therefore, Riesz's replacement of scalar-valued increasing functions ρ_ξ with projection-valued increasing functions E made it possible to incorporate the geometry of Hilbert spaces into the formulation of the spectral theorem. In fact, when $\dim \mathcal{H} < +\infty$, the subspace associated to $E(\lambda)$ is spanned by eigenvectors of T with eigenvalues $\leqslant \lambda$.

Interestingly, even for non-complete inner product spaces, the correspondence $P \mapsto \mathrm{Rng}(P)$ is still injective (cf. the proof of Thm. 5.2.3), which is sufficient for the spectral theorem. The essential reason Riesz's spectral theorem applies only to Hilbert spaces and not to general inner product spaces is that he constructs $E(\lambda)$ by first constructing its associated sesquilinear form $\omega_{E(\lambda)}$, as we will see in Sec. 5.4. However, if an inner product space V is not complete, there are even no natural injective maps from the set of **projection forms** (i.e., bounded Hermitian forms ω satisfying $\omega \circ \omega = \omega$, cf. Def. 3.5.13 for the definition of $\omega \circ \omega$) to the set of linear subspaces of V.

To turn a projection form into a projection operator, one needs the isomorphism $\mathfrak{Ses}(\mathcal{H}) \simeq \mathfrak{L}(\mathcal{H})$ in Thm. 3.5.5, which holds only for Hilbert spaces $\mathcal{H}$ due to the Riesz-Fréchet Thm. 3.5.3. This is one major reason why completeness is essential in Riesz's spectral theorem—though not in the form of Cauchy completeness, but rather through the duality $\mathcal{H} \simeq (\mathcal{H}^{\complement})^*$.

In Sec. 5.4, we will see another fundamental way in which the duality $\mathcal{H} \simeq (\mathcal{H}^{\complement})^*$ enters Riesz's proof of the spectral theorem, once again through the isomorphism $\mathfrak{Ses}(\mathcal{H}) \simeq \mathfrak{L}(\mathcal{H})$.

5.2.3 Projections

Definition 5.2.1. A **projection operator** (or simply a **projection**) on an inner product V is an element $P \in \mathfrak{L}(V)$ such that ω_P is Hermitian (i.e. $\langle\xi|P\eta\rangle = \langle P\xi|\eta\rangle$ for all $\xi,\eta \in V$) and $P^2 = \mathrm{id}_V$. It is easy to check that

$$P^\perp := 1 - P$$

is also a projection.

Recall that if V is a Hilbert space, ω_P being Hermitian is equivalent to $P^* = P$.

Example 5.2.2. Let U be a linear subspace of an inner product space V, and suppose that V has a projection onto U. Then by Cor. 3.3.20, the projection operator associated to U is a projection.

In what follows, we will mainly discuss projections on Hilbert spaces, although many of the results extend naturally to general inner product spaces.

Theorem 5.2.3. *Let $\mathcal{H}$ be a Hilbert space. We have a bijection*

$$\{\textit{projections on } \mathcal{H}\} \xrightarrow{\simeq} \{\textit{closed linear subspaces of } \mathcal{H}\} \qquad P \mapsto \mathrm{Rng}(P) \tag{5.10}$$

Moreover, P is the projection operator associated to $\mathrm{Rng}(P)$ in the sense of Def. 3.3.9. That is, for each $\xi \in \mathcal{H}$, we have $P\xi \in \mathrm{Rng}(P)$ and $\xi - P\xi \in \mathrm{Rng}(P)^\perp$.

Proof. If P is a projection on $\mathcal{H}$, then clearly $1-P$ is also a projection (i.e. $1-P$ is self-adjoint and $(1-P)^2 = 1-P$). Moreover, we have

$$\mathrm{Rng}(P) = \mathrm{Ker}(1-P) \tag{5.11}$$

Indeed, for each $\xi \in \mathcal{H}$ we have $(1-P)P\xi = P\xi - P^2\xi = 0$ and hence $\mathrm{Rng}(P) \subset \mathrm{Ker}(1-P)$; if $(1-P)\xi = 0$, then $\xi = P\xi$, and hence $\mathrm{Ker}(1-P) \subset \mathrm{Rng}(P)$. This proves $\mathrm{Rng}(P) = \mathrm{Ker}(1-P)$. We have thus proved that $\mathrm{Rng}(P)$ is a closed linear subspace of $\mathcal{H}$, since the kernel of any bounded linear operator is a close linear subspace (Cor. 2.3.11).

If $\xi \in \mathcal{H}$, then clearly $P\xi \in \mathrm{Rng}(P)$. For each $\eta \in \mathcal{H}$, we have $\langle P\eta|\xi - P\xi\rangle = \langle \eta|P\xi - P^2\xi\rangle = 0$. Thus $\xi - P\xi \in \mathrm{Rng}(P)^\perp$. This proves that P is the (unique) projection associated to $\mathrm{Rng}(P)$. In particular, P is determined by $\mathrm{Rng}(P)$, and hence the map (5.10) is injective. The surjectivity follows from Thm. 3.4.7. $\square$

Recall Def. 3.5.9 for the meaning of $A \leqslant B$ where A, B are bounded self-adjoint operators on a Hilbert space. The following property says that the bijection (5.10) is an isomorphism of partially ordered sets.

Theorem 5.2.4. *Let P, Q be projections on a Hilbert space $\mathcal{H}$. The following are equivalent.*

(1) $\mathrm{Rng}(P) \subset \mathrm{Rng}(Q)$.

(2) $QP = P$.

(2') $PQ = P$.

(3) $P \leqslant Q$, *namely,* $\langle \xi|P\xi\rangle \leqslant \langle \xi|Q\xi\rangle$ *for all* $\xi \in \mathcal{H}$.

(3') $\|P\xi\| \leqslant \|Q\xi\|$ *for all* $\xi \in \mathcal{H}$.

Proof. (1)⇒(2): Assume (1). Choose any $\xi \in \mathcal{H}$. Since $P\xi \in \mathrm{Rng}(Q)$, the unique orthogonal decomposition of $P\xi$ with respect to $\mathrm{Rng}(Q)$ is $P\xi = P\xi + 0$. Therefore $QP\xi = P\xi$. This proves (2).

(2)⇔(2′): This follows from $(QP)^* = PQ$.

(3)⇔(3′): This follows from $\|P\xi\|^2 = \langle P\xi|P\xi\rangle = \langle\xi|P^*P\xi\rangle = \langle\xi|P\xi\rangle$ and, similarly, $\|Q\xi\|^2 = \langle\xi|Q\xi\rangle$.

(2′)⇒(3′): Since $\xi = P\xi + (1-P)\xi$ where $P\xi \perp (1-P)\xi$, by the Pythagorean identity, we have $\|P\xi\| \leqslant \|\xi\|$. Replacing ξ with $Q\xi$, we get $\|PQ\xi\| \leqslant \|Q\xi\|$. Therefore, (3′) follows from (2′).

(3′)⇒(1′): Assume (3′). Then $\|Q\xi\| = 0$ implies $\|P\xi\| = 0$, i.e., $\mathrm{Ker}(Q) \subset \mathrm{Ker}(P)$. Therefore, $\mathrm{Ker}(P)^\perp \subset \mathrm{Ker}(Q)^\perp$. By (3.12), we have $\mathrm{Ker}(P)^\perp = \mathrm{Rng}(P)$ and $\mathrm{Ker}(Q)^\perp = \mathrm{Rng}(Q)$. This proves (1′). □

Corollary 5.2.5. *Let P, Q be projections on a Hilbert space $\mathcal{H}$. The following are equivalent.*

(1) $\mathrm{Rng}(P) \perp \mathrm{Rng}(Q)$.

(2) $QP = 0$.

(3) $PQ = 0$.

(4) $P + Q \leqslant 1$.

Moreover, if any of these conditions holds, then $P+Q$ is the projection onto the $\mathrm{Rng}(P) + \mathrm{Rng}(Q)$.

If P and Q satisfies one of these conditions, we say that P is **orthogonal** to Q.

Proof. Note that (1) is equivalent to $\mathrm{Rng}(P) \subset \mathrm{Rng}(Q)^\perp$. By (3.12), we have $\mathrm{Rng}(Q)^\perp = \mathrm{Rng}(1-Q)$. Thus (1) is equivalent to $\mathrm{Rng}(P) \subset \mathrm{Rng}(1-Q)$, and hence (by Thm. 5.2.4) is equivalent to each of the following three conditions: $P(1-Q) = P$, $(1-Q)P = P$, and $P \leqslant 1-Q$. This proves the equivalence of the four conditions.

Now, assume that these four conditions hold. Clearly $P+Q$ is a projection operator, and its range lies inside $\mathrm{Rng}(P) + \mathrm{Rng}(Q)$. Conversely, any vector in $\mathrm{Rng}(P) + \mathrm{Rng}(Q)$ can be written as $P\xi + Q\eta$ where $\xi, \eta \in \mathcal{H}$. Then

$$P\xi + Q\eta = (P+Q)(P\xi + Q\eta)$$

This proves that $\mathrm{Rng}(P) + \mathrm{Rng}(Q)$ is the range of $P+Q$. □

Corollary 5.2.6. *Let P, Q be projections on a Hilbert space such that $P \leqslant Q$. Then $Q-P$ is the projection operator associated to $\mathrm{Rng}(Q) \cap \mathrm{Rng}(P)^\perp$, that is,*

$$(Q-P)(\mathcal{H}) = Q(\mathcal{H}) \cap P(\mathcal{H})^\perp$$

Proof. Clearly $E := Q - P$ is a projection, and $PE = EP = 0$. By Cor. 5.2.5, we have $Q(\mathcal{H}) = P(\mathcal{H}) + E(\mathcal{H})$ where $P(\mathcal{H}) \perp E(\mathcal{H})$. From this one easily shows $E(\mathcal{H}) = Q(\mathcal{H}) \cap P(\mathcal{H})^\perp$. □

The convergence of an increasing net of projections also has a geometric meaning:

Theorem 5.2.7. *Let $(E_\alpha)_{\alpha\in I}$ be a net of projections on $\mathcal{H}$. Assume that (E_α) is increasing, i.e., $E_\alpha \leqslant E_\beta$ whenever $\alpha \leqslant \beta$. Let E be the projection operator such that*

$$\mathrm{Rng}(E) = \mathrm{Cl}_{\mathcal{H}}\Big(\bigcup_{\alpha\in I}\mathrm{Rng}(E_\alpha)\Big) \tag{5.12}$$

Then $E_\alpha \leqslant E$ for each α, and $\lim_\alpha E_\alpha$ converges in SOT to E.

Consequently, if an increasing net of projections (E_α) converges in WOT to some $F \in \mathfrak{L}(\mathcal{H})$, then clearly $F = E$. It follows that F is a projection, that (E_α) in SOT to E, and that $\mathrm{Rng}(F) = \mathrm{Cl}_{\mathcal{H}}\big(\bigcup_{\alpha\in I}\mathrm{Rng}(E_\alpha)\big)$.

Proof. Since $\mathrm{Rng}(E_\alpha) \subset \mathrm{Rng}(E)$, we have $E_\alpha \leqslant E$. Let $\xi \in \mathcal{H}$. By the definition of E, for each $\varepsilon > 0$ there exists $\alpha \in I$ such that

$$\|E\xi - E_\alpha\eta\| \leqslant \varepsilon$$

for some $\eta \in \mathcal{H}$. Since $E\xi - E_\alpha\eta = (E-E_\alpha)\xi + E_\alpha(\xi-\eta)$ with $(E-E_\alpha)\xi$ orthogonal to $E_\alpha(\xi - \eta)$ (cf. Cor. 5.2.6), the Pythagorean identity shows that $\|(E - E_\alpha)\xi\| \leqslant \|E\xi - E_\alpha\eta\|$ and hence

$$\|(E - E_\alpha)\xi\| \leqslant \varepsilon$$

For each $\beta \geqslant \alpha$, the inequality $E_\alpha \leqslant E_\beta$ implies $E - E_\beta \leqslant E - E_\alpha$, and hence (by Thm. 5.2.4) $\|(E - E_\beta)\xi\| \leqslant \|(E - E_\alpha)\xi\| \leqslant \varepsilon$. This proves $\lim_\alpha E_\alpha\xi = E\xi$. □

Corollary 5.2.8. *Let $(E_\alpha)_{\alpha\in I}$ be a decreasing net of projections on $\mathcal{H}$. Let E be the projection operator onto*

$$\mathrm{Rng}(E) = \bigcap_{\alpha\in I}\mathrm{Rng}(E_\alpha)$$

Then $E_\alpha \geqslant E$ for each α, and $\lim_\alpha E_\alpha$ converges in SOT to E.

Proof. Apply Thm. 5.2.7 to the increasing net $(E_\alpha^\perp)$ and use the following Exe. 5.2.9. □

Exercise 5.2.9. Let $(\mathcal{K}_\alpha)_{\alpha\in\mathscr{A}}$ be a family of closed linear subspaces of $\mathcal{H}$. Prove that

$$\mathrm{Cl}_{\mathcal{H}}\Big(\bigcup_\alpha \mathcal{K}_\alpha\Big)^\perp = \bigcap_\alpha \mathcal{K}_\alpha^\perp \qquad \Big(\bigcap_\alpha \mathcal{K}_\alpha\Big)^\perp = \mathrm{Cl}_{\mathcal{H}}\Big(\bigcup_\alpha \mathcal{K}_\alpha^\perp\Big) \tag{5.13}$$

(Note: Recall Cor. 3.4.8.)

5.3 Towards Riesz's spectral theorem: monotone convergence extension

5.3.1 Paradigm shift: from finite approximation to linear extension

Another of Riesz's innovations on Hilbert's spectral theorem is his entirely different approach to the polynomial moment problem/the Riesz representation theorem. One year after his proof of the spectral theorem in [Rie13], Riesz gave a new proof of the Riesz representation theorem in [Rie14]; this proof draws on key steps from his treatment of the spectral theorem in [Rie13] and simplifies the method he originally used in [Rie09].

We will not discuss [Rie09], since it offers little insight for this course. Instead we compare the method used in [Rie13, Rie14]—which I call the monotone convergence extension—with the Stieltjes-Hilbert method for treating the moment problem. (Recall again that the Riesz representation theorem for $C(I)$, with I a compact interval, is roughly equivalent to the Hausdorff moment problem, cf. Subsec. 2.2.1 and Sec. 4.5.) The transition from the Stieltjes-Hilbert method to Riesz's method marked a paradigm shift in the early development of functional analysis: the move from finite approximations to linear extensions. We explain this in more detail below.

As discussed in Sec. 2.2 and summarized in Table 2.3, the traditional approach to moment problems and to characterizing dual spaces proceeds in two steps: (1) establish the link between pointwise convergence of functions and convergence of moments; (2) show that any bounded (or positive) linear functional can be approximated in the weak-* topology by elementary functions. The treatment of polynomial moment problems in Ch. 4 exemplifies this strategy: the connection in step (1) is captured by Thm. 2.9.6 together with the Helly selection theorem (Thm. 2.9.3), while step (2) is achieved via Padé approximation, implemented as finite-rank approximations of Hermitian operators. This approach clearly belongs to the paradigm of finite approximation.

This finite-approximation paradigm gave way to the linear-extension paradigm, with F. Riesz as its prime mover. The guiding philosophy of this paradigm is:

> To characterize a (positive or bounded) linear functional Λ on a function space V, extend Λ to a suitable linear functional on a larger function space $\widetilde{V}$, whose linear functionals are easier to characterize.

In this course we will exhibit two main patterns of the linear-extension paradigm:

- The **monotone convergence extension**, the main subject of this section (and treated in greater detail in [Gui-A, Ch. 24-25]), which is closely tied to integration theory.

- The **bounded linear extension**, which is intimately connected with convexity in normed spaces and with various forms of the Hahn-Banach theorem.

The monotone convergence extension may be regarded as a gift from integration theory to functional analysis. Whereas Lebesgue developed integration by first defining measurable sets, the approach of monotone convergence extension—originally introduced by Young [You10, You13] as an alternative to Lebesgue's approach that appealed to more conservative contemporaries [2]—builds the integral by enlarging the class of integrable functions in such a way that the integral satisfies the monotone convergence theorem.

5.3.2 Monotone convergence extension as a theorem

Fix $\mathbb{K} \in \{\mathbb{R}_{\geqslant 0}, \mathbb{R}, \mathbb{C}\}$. Let X be a topological space. Recall from Sec. 1.1 that

$$\mathscr{Bor}_{\mathrm{b}}(X, \mathbb{K}) = \{\text{bounded Borel functions } X \to \mathbb{K}\} \tag{5.14}$$

As usual, $\mathscr{Bor}_{\mathrm{b}}(X)$ denotes $\mathscr{Bor}_{\mathrm{b}}(X, \mathbb{C})$.

Definition 5.3.1. A positive linear functional $\Lambda : \mathscr{Bor}_{\mathrm{b}}(X, \mathbb{K}) \to \mathbb{K}$ is called **normal** [3] if it satisfies the monotone convergence theorem, that is, if (f_n) is an increasing sequence in $\mathscr{Bor}_{\mathrm{b}}(X, \mathbb{R}_{\geqslant 0})$ converging pointwise to $f \in \mathscr{Bor}_{\mathrm{b}}(X, \mathbb{R}_{\geqslant 0})$, then

$$\lim_n \Lambda(f_n) = \Lambda(f)$$

Proposition 5.3.2. *We have a bijective map*

$$\begin{gathered}\{\textit{finite Borel measures on } X\} \xrightarrow{\simeq} \{\textit{normal positive linear functionals on } \mathscr{Bor}_{\mathrm{b}}(X, \mathbb{K})\} \\ \mu \mapsto \Lambda_\mu \end{gathered} \tag{5.15}$$

where $\Lambda_\mu(f) = \int_X f d\mu$ for each $f \in \mathscr{Bor}_{\mathrm{b}}(X, \mathbb{K})$. The measure μ is determined by Λ_μ by the relation $\mu(E) = \Lambda_\mu(\chi_E)$ for each Borel set $E \subset X$.

Proof. Given each finite Borel measure μ, Λ_μ satisfies the MCT. Therefore the map (5.15) is well-defined. Since Λ_μ is determined by its values on $\mathscr{Bor}_{\mathrm{b}}(X, \mathbb{R}_{\geqslant 0})$, and since each $f \in \mathscr{Bor}_{\mathrm{b}}(X, \mathbb{R}_{\geqslant 0})$ is the pointwise limit of an increasing sequence of simple functions, Λ_μ must be determined by the values $\Lambda_\mu(\chi_E) = \mu(E)$ for all any Borel set $E \subset X$. Therefore, the map (5.15) is injective.

[2] See [Pes, Sec. 6.6].

[3] This terminology is borrowed from the theory of von Neumann algebras. We avoid using the term "Borel", as it is reserved for describing maps between topological spaces: a map is called Borel if the preimage of every Borel set is itself Borel.

To prove that (5.15) is surjective, we pick an arbitrary normal positive linear functional $\Lambda : X \to \mathbb{K}$. Then Λ being normal implies that $\mu : E \in \mathfrak{B}_X \mapsto \Lambda(\chi_E) \in \mathbb{R}_{\geqslant 0}$ is a (Borel) measure on X. So Λ and Λ_μ agree on simple functions. Since both Λ and Λ_μ satisfy MCT, by the argument in the first paragraph, we conclude $\Lambda = \Lambda_\mu$. □

Example 5.3.3. Let Y be a topological space. Let $\Phi : X \to Y$ be a Borel map. Let μ be a finite Borel measure on X. Then the positive linear functional

$$f \in \mathscr{Bor}_{\mathrm{b}}(Y) \to \int_X (f \circ \Phi) d\mu$$

is clearly normal. Indeed, by Def. 1.6.6, this functional is represented by the pushforward measure $\Phi_*\mu$.

Prop. 5.3.2, which gives us a linear functional interpretation of measure theory, allows us to formulate the Riesz-Markov representation theorem (Thm. 1.7.6) for second-countable compact Hausdorff spaces in the form of monotone convergence extension.

Theorem 5.3.4 (Riesz-Markov representation theorem). *Let X be a second-countable compact Hausdorff space. Then we have an $\mathbb{R}_{\geqslant 0}$-linear isomorphism*

$$\begin{gathered}\{\text{normal positive linear functionals } \mathscr{Bor}_{\mathrm{b}}(X,\mathbb{F}) \to \mathbb{F}\} \\ \downarrow \simeq \\ \{\text{positive linear functionals } C(X,\mathbb{F}) \to \mathbb{F}\} \\ \Lambda \mapsto \Lambda|_{C(X,\mathbb{F})}\end{gathered} \tag{5.16}$$

Note that for second-countable compact Hausdorff spaces, finite Borel measures and finite Radon measures are synonymous (cf. Thm. 1.7.8).

Proof. This follows immediately from Thm. 1.7.6 and Prop. 5.3.2. □

Corollary 5.3.5 (Abstract Hausdorff moment theorem). *Let X be a second-countable compact Hausdorff space. Let $\mathscr{A}$ be a unital *-$\mathbb{F}$-subalgebra of $C(X,\mathbb{F})$ separating points of X. Then we have an $\mathbb{R}_{\geqslant 0}$-linear isomorphism*

$$\begin{gathered}\{\text{normal positive linear functionals } \mathscr{Bor}_{\mathrm{b}}(X,\mathbb{F}) \to \mathbb{F}\} \\ \downarrow \simeq \\ \{\text{positive linear functionals } \mathscr{A} \to \mathbb{F}\} \\ \Lambda \mapsto \Lambda|_{\mathscr{A}}\end{gathered} \tag{5.17}$$

Proof. By Stone-Weierstrass, the l^∞-closure of $\mathscr{A}$ is $C(X,\mathbb{F})$. Therefore, the corollary follows from Thm. 5.3.4 and 4.5.2. □

Remark 5.3.6. Suppose that X is an LCH space, not necessarily second-countable. A positive linear functional $\Lambda : \mathscr{Bor}_{\mathrm{b}}(X, \mathbb{F}) \rightarrow \mathbb{F}$ is called **Radon** if there exists a finite Radon measure μ on X such that

$$\Lambda(f) = \int_X f d\mu \qquad \text{for all } f \in \mathscr{Bor}_{\mathrm{b}}(X, \mathbb{F})$$

In view of Def. 1.7.3, Λ is Radon iff Λ is normal and satisfies the following two extra conditions:

(a) For each Borel set $E \subset X$ we have

$$\Lambda(\chi_E) = \inf \left\{ \Lambda(\chi_U) : U \in \mathcal{T}_X \text{ and } U \supset E \right\}$$

(b) For each open $U \subset X$ we have

$$\Lambda(\chi_U) = \sup \left\{ \Lambda(f) : f \in C(X, [0,1]) \right\}$$

It is clear that Thm. 5.3.4 and Cor. 5.3.5 can be generalized to this situation, with normal positive linear functionals replaced by Radon positive linear functionals.

5.3.3 Monotone convergence extension as a method

The connection between Thm. 5.3.4 and the monotone convergence extension is straightforward: the theorem asserts that every positive linear functional on $C(X, \mathbb{F})$ extends uniquely to a positive linear functional on $\mathscr{Bor}_{\mathrm{b}}(X, \mathbb{F})$ satisfying the monotone convergence theorem. However, the monotone convergence extension is not only the statement of a theorem, but also provides the mechanism for constructing the proof.

In what follows, we outline how this method of monotone convergence extension is applied to prove Thm. 5.3.4. For simplicity, we restrict attention to an $\mathbb{R}_{\geqslant 0}$-linear functional $\Lambda : C(X, \mathbb{R}_{\geqslant 0}) \rightarrow \mathbb{R}_{\geqslant 0}$, and explain how such a linear functional is extended.

The first main step is to extend Λ to an $\mathbb{R}_{\geqslant 0}$-linear functional on

$$\mathrm{LSC}_{\mathrm{b}}(X, \mathbb{R}_{\geqslant 0}) = \{\text{bounded lower semicontinuous functions } X \rightarrow \mathbb{R}_{\geqslant 0}\} \tag{5.18}$$

by the following procedure of **monotone convergence extension**:

(1) For each $f \in \mathrm{LSC}_{\mathrm{b}}(X, \mathbb{R}_{\geqslant 0})$, set

$$\Lambda(f) = \sup\{\Lambda(h) : h \leqslant f, h \in C(X, \mathbb{R}_{\geqslant 0})\}$$

(2) Prove the following version of the MCT: If (f_n) is a uniformly bounded increasing sequence in $\mathrm{LSC}_{\mathrm{b}}(X, \mathbb{R}_{\geqslant 0})$ converging pointwise to $f : X \rightarrow \mathbb{R}_{\geqslant 0}$. Then $\Lambda(f) = \lim_n \Lambda(f_n)$. (Note that $f \in \mathrm{LSC}_{\mathrm{b}}(X, \mathbb{R}_{\geqslant 0})$ by Prop. 1.4.10.)

(3) Show that any $f\in \mathrm{LSC}_{\mathrm{b}}(X,\mathbb{R}_{\geqslant 0})$ is the pointwise limit of an increasing sequence of functions in $C(X,\mathbb{R}_{\geqslant 0})$. Together with Step 2, this implies that the extended Λ is still $\mathbb{R}_{\geqslant 0}$-linear.

Readers familiar with measure theory will recognize that this is the same method used to define the integral on a measure space: one extends the integral from nonnegative simple functions to nonnegative measurable functions such that the MCT is satisfied. However, in [Rie13, Rie14], Riesz employed an equivalent but seemingly different procedure:

(a) Show that any $f\in \mathrm{LSC}_{\mathrm{b}}(X,\mathbb{R}_{\geqslant 0})$ is the pointwise limit of an increasing sequence of functions (f_n) in $C(X,\mathbb{R}_{\geqslant 0})$.

(b) Define $\Lambda(f)$ to be $\lim_n \Lambda(f_n)$ where (f_n) is any increasing sequence in $C(X,\mathbb{R}_{\geqslant 0})$ converging pointwise to $f\in \mathrm{LSC}_{\mathrm{b}}(X,\mathbb{R}_{\geqslant 0})$.

(c) Show that $\Lambda(f)$ is well-defined, i.e., independent of the choice of (f_n) approximating f. (The linearity of Λ is obvious.)

Step (c) plays the role of Step (2) mentioned above, since the arguments required in both cases are essentially the same.

The above approach (adapted to nets so that it applies to general locally compact Hausdorff spaces) is used in [Gui-A, Ch. 25] to prove the Riesz-Markov representation Thm. 1.7.6. To complete the proof, **the second main step** is of course to extend Λ from $\mathrm{LSC}_{\mathrm{b}}(X,\mathbb{R}_{\geqslant 0})$ to $\mathscr{Bor}_{\mathrm{b}}(X,\mathbb{F})$. In [Gui-A], this extension is carried out via a more measurable-set–based approach rather than the monotone convergence extension. Nevertheless, it is possible to proceed using monotone convergence extension as follows.

First, extend Λ from $\mathrm{LSC}_{\mathrm{b}}(X,\mathbb{R}_{\geqslant 0})$ to

$$\mathscr{C}_0=\{f+h: f\in \mathrm{LSC}_{\mathrm{b}}(X,\mathbb{R}_{\geqslant 0}), h\geqslant 0, h=0 \text{ almost everywhere}\}$$

by setting $\Lambda(f+h)=\Lambda(f)$ (with "almost everywhere" interpreted appropriately [4]). Then apply steps (a)-(c) above to extend Λ from $\mathscr{C}_0$ to the class $\mathscr{C}_1$ of functions that are pointwise limits of increasing sequences in $\mathscr{C}_0$. Finally, using Lem. 1.7.5 and Eq. (1.42), extend Λ to $\mathscr{C}_2:=\mathrm{Span}_{\mathbb{F}}(\mathscr{C}_1)$. One then verifies that $\mathscr{C}_2$ coincides with the space of bounded measurable functions.

The above approach was in fact used by Riesz and Sz.-Nagy to construct the Lebesgue integral on a compact interval $I\subset\mathbb{R}$, though with nonnegative step functions in place of $\mathrm{LSC}_{\mathrm{b}}(X,\mathbb{R}_{\geqslant 0})$. See [RN, Sec. 16-22] or [Apo, Ch. 10]. However, to prove the Riesz representation theorem for $C(I,\mathbb{F})$—that is, to represent a positive linear functional

$$\Lambda: C(I,\mathbb{F})\rightarrow\mathbb{F}$$

[4] One defines "almost everywhere" by defining a set $E\subset X$ to be null if it is contained in some open set $U\subset X$ with $\Lambda(\chi_U)=0$.

by a Stieltjes-integral against an increasing function $\rho : I \to \mathbb{R}_{\geqslant 0}$—it suffices to extend Λ to $\mathrm{LSC_b}(I, \mathbb{F})$ using the monotone convergence extension described in the first main step (namely, (a)-(c), or equivalently, (1)-(3)). By Lem. 1.7.5 and Eq. (1.42), Λ is further extended to $\mathrm{Span}_{\mathbb{F}}\mathrm{LSC_b}(I, \mathbb{F})$, which then allows us to define a desired increasing function $\rho : I = [a, b] \to \mathbb{R}_{\geqslant 0}$ by

$$\rho(x) = \Lambda(\chi_{[a,x]})$$

where $\chi_{I_{\leqslant x}}$ is an upper semicontinuous function (and hence lies to $\mathrm{Span}_{\mathbb{F}}\mathrm{LSC_b}(I, \mathbb{F})$). The MCT shows that this function ρ is right-continuous. This is precisely the approach taken by Riesz in [Rie13] to prove the Riesz representation theorem in the form needed for his spectral theorem.

For a more detailed discussion of the monotone convergence extension method, see [Gui-A, Ch. 25].

5.4 Riesz's spectral theorem: two paradigm shifts

5.4.1 Three paradigm shifts, and why Riesz's spectral theorem is related to the first two

The theme of this course is the three major paradigm shifts in functional analysis:

$$\text{From finite approximations to linear extensions} \tag{5.19a}$$
$$\text{From (muti)linear forms to linear operators} \tag{5.19b}$$
$$\text{From duality to Cauchy completeness} \tag{5.19c}$$

Riesz's proof of the spectral theorem in [Rie13] was a major milestone for the first two shifts. We have already discussed the first shift in detail in Sec. 5.3; we now turn to the second.

As discussed before (for instance, in Sec. 2.1 and 2.5), functional analysis moved its focus from scalar-valued functions (especially linear functionals and bilinear or sesquilinear forms) to vector-valued functions (linear operators acting on a normed or inner-product space V). This is the second paradigm shift mentioned above.

We have seen that Hilbert's spectral theorem (Thm. 5.1.1) is stated in the language of bilinear/sesquilinear forms. As noted in Sec. 5.2, one advantage of Riesz's spectral theorem being expressed in terms of linear operators is that projection operators correspond more naturally to linear subspaces of an inner-product space than do projection forms. Another reason—mentioned in Subsec. 2.5.2—is that symbolic calculus is easier to manipulate in the operator framework, i.e., one may replace the real or complex variable x in a function $f(x)$ by an operator or a sesquilinear form. We will explore this in more detail in the following sections. For now, we answer some questions that readers might naturally ask.

Question 5.4.1. What's the role played by symbolic calculus in Riesz's spectral theorem?

Answer. Let $T \in \mathfrak{L}(\mathcal{H})$ be a bounded self-adjoint operator on a Hilbert space $\mathcal{H}$. Starting with the polynomial functional calculus, i.e. the linear map

$$\pi_T : \mathbb{C}[x] \to \mathfrak{L}(\mathcal{H}) \qquad x^n \mapsto T^n$$

Riesz applied the Riesz representation theorem to extend π_T to a homomorphism

$$\pi_T : \mathrm{SpanLSC_b}([-r, r]) \to \mathfrak{L}(\mathcal{H})$$

where $r \in \mathbb{R}_{\geqslant 0}$ satisfies $r \geqslant \|T\|$. This extended linear map is not only linear but also multiplicative (i.e. $\pi_T(fg) = \pi_T(f)\pi_T(g)$) and intertwines the involutions (i.e. $\pi_T(\overline{f}) = \pi_T(f)^*$). Therefore, since for each $\lambda \in [-a, a]$ we have $\overline{\chi_{[a,\lambda]}} = \chi_{[a,\lambda]}$ and $\chi_{[a,\lambda]}^2 = \chi_{[a,\lambda]}$, the operator

$$E(\lambda) := \pi_T(\chi_{[a,\lambda]})$$

is a projection. This yields the construction of the spectral projections mentioned in Subsec. 5.2.2. See Rem. 5.6.1 for further discussion. □

Note that the above answer also explains how the Riesz representation theorem is used in the proof of Riesz's spectral theorem.

Question 5.4.2. Why does symbolic/functional calculus require the linear-operator perspective rather than the bilinear/sesquilinear form perspective?

Answer. Of course, to perform symbolic/functional calculus one must first define multiplication of operators, sesquilinear forms, or matrices. As noted in Subsec. 5.2.2, multiplication of bounded sesquilinear forms or bounded matrices can be defined as in Def. 3.5.13; this historical approach was indeed the one originally adopted by mathematicians.

The principal complexity with the sesquilinear-form and matrix perspectives—which does not arise in finite-dimensional linear algebra—is not their definition but the associativity of multiplication: one must address the Fubini-type issues for infinite sums appearing in (3.30), in (3.31), and in more complicated expressions. For bounded sesquilinear forms and bounded matrices, defining products in terms of orthonormal bases introduces many inconveniences that are absent in the finite-dimensional setting, whereas the operator framework avoids this subtlety. □

5.4.2 Riesz's spectral theorem, and why the third paradigm shift is missing

Fix a Hilbert space $\mathcal{H}$.

Definition 5.4.3. Let I be an interval. An **increasing net of projections** indexed by I is defined to be a function

$$E: I \to \{\text{projections on } \mathcal{H}\}$$

such that $E(\lambda) \leqslant E(\mu)$ (cf. Thm. 5.2.4) for all $\lambda, \mu \in I$ satisfying $\lambda \leqslant \mu$. We say that E is **right-continuous**, if for each $\lambda \in [a,b)$ we have

$$\lim_{\lambda \to \lambda_0} E(\lambda) = E(\lambda_0) \tag{5.20}$$

By Cor. 5.2.8, the SOT and WOT of the above limit are equivalent; moreover, this limit is equivalent to

$$\mathrm{Rng}(E(\lambda_0)) = \bigcap_{\lambda > \lambda_0} \mathrm{Rng}(E(\lambda))$$

The following theorem was proved in [Rie13, Ch. V, Sec. 94].

Theorem 5.4.4 (Riesz's spectral theorem). *Let $T \in \mathfrak{L}(\mathcal{H})$ be self-adjoint. Let $r \in \mathbb{R}_{\geqslant 0}$ such that $\|T\| \leqslant r$. Then there exists a right-continuous increasing net of projections $E: [-r,r] \to \mathfrak{L}(\mathcal{H})$ (called the **spectral projections**) such that for any $f \in C([-r,r])$, the continuous functional calculus $f(T)$ satisfies*

$$f(T) = \int_{[-r,r]} f(\lambda) dE(\lambda) \tag{5.21}$$

Right-continuity is not essential; it is imposed to ensure uniqueness of the net E satisfying (5.21). We will not need this uniqueness in the course.

Neither the continuous functional calculus nor the integral on the RHS of (5.21) has been defined yet. We will do this in Sec. 5.5.

Question 5.4.5. What is the essence of Riesz's spectral theorem?

Answer. As noted in the answer to Question 5.1.4, the proof of Hilbert's spectral theorem may be regarded as a linear-algebraic proof (see Sec. 4.2) of the Hausdorff moment problem/Riesz representation theorem, carried out in the paradigm of finite approximation. In contrast,

> Riesz's spectral theorem should be viewed as the operator-valued Riesz representation theorem for $C([-r,r])$, with its proof situated in the paradigm of linear extension (cf. the paradigm shift (5.19a)).

Indeed, just as the Riesz representation theorem (Thm. 1.10.1) expresses a positive linear functional Λ as a Stieltjes integral, Eq. (5.21) expresses the continuous functional calculus as an operator-valued Stieltjes integral. □

Question 5.4.6. Is the Cauchy completeness of $\mathcal{H}$ used in Riesz's proof of his spectral theorem?

Answer. No. What plays the essential role is the Riesz isomorphism $\mathcal{H} \simeq (\mathcal{H}^{\mathbb{C}})^*$ (cf. Thm. 3.5.3), or more precisely its consequence $\mathfrak{Ses}(\mathcal{H}) \simeq \mathfrak{L}(\mathcal{H})$ (cf. Thm. 3.5.5). This isomorphism makes it possible to pass seamlessly between the paradigms of bilinear/sesquilinear forms and linear operators, thereby taking advantage of the strengths of both: As noted in Subsec. 2.5.2, the bilinear/sesquilinear form paradigm is well suited for carrying out the monotone-convergence extension, while the linear operator paradigm provides the natural setting for symbolic calculus. □

Question 5.4.7. In Subsec. 2.5.1, you remarked that the transition from the bilinear/sesquilinear form paradigm to the linear-operator paradigm (cf. (5.19b)) necessarily increases the role of Cauchy-completeness in functional analysis, thereby leading to the paradigm shift (5.19c) from duality to Cauchy-completeness. Yet, according to the answers to the previous two questions, Riesz's proof of his spectral theorem exhibits the shift (5.19b) but not (5.19c). Why is that?

Answer. The fact that Riesz's proof of his spectral theorem reflects the shift (5.19b) but not (5.19c) shows that he did not fully abandon the bilinear/sesquilinear form paradigm in favor of the linear-operator paradigm. In other words, the paradigm shift (5.19b) in Riesz's proof of his spectral theorem is not complete. □

Question 5.4.8. From the perspective of modern functional analysis textbooks, does this incomplete realization of the paradigm shift (5.19b) imply that Riesz's approach to his spectral theorem is outdated?

Answer. In a sense, yes. However, this incompleteness, in my view, is precisely the merit of Riesz's approach. In his approach, neither of the two paradigms in (5.19b) is eliminated; instead, he allows them to meet and reinforce one another, rather than reducing one to a black-box result.

By contrast, many modern expositions of the spectral theorem, fully committed to the linear operator paradigm, typically invoke the Riesz representation Thm. 1.10.1 (or more generally, the Riesz-Markov representation Thm. 1.7.6) as an external result, without exposing its intimate connection to the paradigm of bilinear/sesquilinear form—the more natural paradigm for understanding the Riesz representation theorem. Riesz's own proof thus illustrates a broader methodological lesson: a good proof not only reaches the conclusion correctly, but also creates a setting where multiple conceptual frameworks can coexist and interact fruitfully. □

5.5 Proof of Riesz's spectral Thm. 5.4.4

In this section, we prove the Riesz spectral theorem (Thm. 5.4.4).

5.5.1 The Stieltjes integral against increasing projections

Let $-\infty < a < b < +\infty$.

Definition 5.5.1. Let $E : [a,b] \to \mathfrak{L}(\mathcal{H})$ be an increasing net of projections. Let $f \in C([a,b])$. For each tagged partition

$$(\sigma, \lambda_\bullet) = \big(\{a_0 = a < a_1 < \cdots < a_n = b\}, (\lambda_1, \dots, \lambda_n)\big)$$

(cf. Def. 1.9.3), define the Stieltjes sum

$$S_E(f, \sigma, \lambda_\bullet) = \sum_{j=1}^{n} f(\lambda_j)(E(a_j) - E(a_{j-1}))$$

Define the **operator-valued Stieltjes integrals**

$$\int_{(a,b]} f(\lambda)dE(\lambda) = \lim_{(\sigma,\lambda_\bullet)\in\mathcal{Q}([a,b])} S_E(f, \sigma, \lambda_\bullet) \tag{5.22}$$

$$\int_{[a,b]} f(\lambda)dE(\lambda) = f(a)E(a) + \int_{(a,b]} f(\lambda)dE(\lambda) \tag{5.23}$$

Proposition 5.5.2. *The limit on the RHS of (5.22) converges in the operator norm. Moreover, for each $\xi \in \mathcal{H}$, we have*

$$\Big\langle \xi \Big| \int_{[a,b]} f(\lambda)dE(\lambda)\xi \Big\rangle = \int_{[a,b]} f(\lambda)d\langle \xi|E(\lambda)\xi\rangle \tag{5.24}$$

where the RHS is the Stieltjes integral of f against the increasing function $\lambda \mapsto \langle \xi|E(\lambda)\xi\rangle$ on $[a,b]$.

Proof. By Cor. 2.4.9, $\mathfrak{Ses}(\mathcal{H})$ is Cauchy-complete (even if $\mathcal{H}$ is replaced by a general inner product space). Since we have a canonical isomorphism of normed vector spaces $\mathfrak{Ses}(\mathcal{H}) \simeq \mathfrak{L}(\mathcal{H})$ (cf. Thm. 3.5.5), we conclude that $\mathfrak{L}(\mathcal{H})$ is complete. [5] Therefore, by Thm. 1.2.15, it suffices to show that $(S_E(f, \sigma, \lambda_\bullet))_{(\sigma,\lambda_\bullet)\in\mathcal{Q}([a,b])}$ is a Cauchy net. Eq. (5.24) will follow from the easy fact that

$$\big\langle \xi \big| S_E(f, \sigma, \lambda_\bullet)\xi \big\rangle = S_{\rho_\xi}(f, \sigma, \lambda_\bullet) \tag{5.25}$$

where $\rho_\xi(\lambda) = \langle \xi|E(\lambda)\xi\rangle$.

Since f is uniformly continuous, for each $\varepsilon > 0$ there is $\delta > 0$ such that $|f(s) - f(t)| \leqslant \varepsilon$ whenever $|s - t| \leqslant \delta$. Choose any tagged partition $(\sigma, \lambda_\bullet)$ with mesh

[5] Of course, since $\mathcal{H}$ is complete, one can directly invoke Cor. 2.4.9 to conclude that $\mathfrak{L}(\mathcal{H})$ is complete. The argument provided here is to justify the answer to Question 5.4.6, namely, that the isomorphism $\mathcal{H} \simeq (\mathcal{H}^{\complement})^*$ rather than the Cauchy completeness of $\mathcal{H}$ is used to prove the Riesz spectral theorem.

$\leqslant \delta$. If a tagged partition $(\sigma', \lambda'_\bullet)$ is finer than $(\sigma, \lambda_\bullet)$, then for each $\xi \in \mathcal{H}$, applying (1.55) to the increasing function ρ_ξ, we get

$$\begin{aligned}
&\left|\langle\xi|(S_E(f,\sigma',\lambda'_\bullet) - S_E(f,\sigma,\lambda_\bullet))\xi\rangle\right| = |S_{\rho_\xi}(f,\sigma',\lambda'_\bullet) - S_{\rho_\xi}(f,\sigma,\lambda_\bullet)| \\
\leqslant&\varepsilon \cdot (\rho_\xi(b) - \rho_\xi(a)) \leqslant \varepsilon \cdot \rho_\xi(b) = \varepsilon\langle\xi|E(b)\xi\rangle \leqslant \varepsilon\|\xi\|^2
\end{aligned}$$

It follows from Prop. 3.2.12 that

$$\|S_E(f,\sigma',\lambda'_\bullet) - S_E(f,\sigma,\lambda_\bullet)\| \leqslant 4\varepsilon$$

finishing the proof of the Cauchyness. □

5.5.2 Radon/normal unitary representations

As noted in the answer to Question 5.4.5, proving Riesz's spectral theorem essentially amounts to establishing an operator-valued analogue of the Riesz representation theorem, and is therefore almost equivalent to proving an operator-valued analogue of the Hausdorff moment problem. We will approach this in a slightly more general framework: namely, we shall prove Thm. 5.5.13, an operator-valued version of the abstract Hausdorff moment theorem (Cor. 5.3.5).

In this subsection, we define an operator-valued version of Radon/normal positive linear functionals. Recall that *-algebras refer to *-$\mathbb{C}$-algebras.

Definition 5.5.3. Let $\mathscr{A}$ be a unital *-algebra. A **pre-unitary representation** of $\mathscr{A}$ denotes a pair (π, V) where V is an inner product space, and

$$\pi : \mathscr{A} \to \mathrm{Lin}(V)$$

is a linear map satisfying

$$\pi(xy) = \pi(x)\pi(y) \qquad \langle\eta|\pi(x)\xi\rangle = \langle\pi(x^*)\eta|\xi\rangle \qquad \pi(1_{\mathscr{A}}) = \mathrm{id}_V$$

for all $x, y \in \mathscr{A}$ and $\xi, \eta \in V$. If moreover V is a Hilbert space and $\pi(\mathscr{A}) \subset \mathfrak{L}(V)$, we say that (π, V) is a **unitary representation** of $\mathscr{A}$.

When $\mathscr{A}$ is a non-unital *-algebra, one can define pre-unitary and unitary representations in the same way, except that one does not assume $\pi(1_{\mathscr{A}}) = \mathrm{id}_V$. □

Convention 5.5.4. When the context is clear, a pre-unitary representation (π, V) is abbreviated to V, and $\pi(x)\xi$ is abbreviated to $x\xi$.

Definition 5.5.5. Let X be a set, and let $\mathscr{A}$ be a unital *-subalgebra of $l^\infty(X)$. A pre-unitary representation (π, V) of $\mathscr{A}$ is called **positive** if the linear functional $\langle\xi|\pi(-)\xi\rangle : \mathscr{A} \to \mathbb{C}$ is positive for each $\xi \in V$; in other words,

$$\langle\xi|\pi(f)\xi\rangle \geqslant 0 \qquad \text{for all } \xi \in V \text{ and } f \in \mathscr{A} \text{ satisfying } f \geqslant 0 \tag{5.26}$$

Proposition 5.5.6. *Let X be a set, and let $\mathscr{A}$ be a unital *-subalgebra of $l^\infty(X)$. Suppose that any $f \in \mathscr{A}$ with $f \geqslant 0$ can be written as $f = g^*g$ for some $g \in \mathscr{A}$. Then any pre-unitary representation (π, V) is positive.*

For example, if X is a compact Hausdorff (resp. LCH) sapce, any pre-unitary representation of $C(X)$ (resp. $\mathscr{Bor}_{\mathrm{b}}(X)$) is positive.

Proof. For each $f \in \mathscr{A}$ with $f \geqslant 0$, write $f = g^*g$. Then

$$\langle \xi | \pi(f)\xi \rangle = \langle \xi | \pi(g^*)\pi(g)\xi \rangle = \langle \pi(g)\xi | \pi(g)\xi \rangle \geqslant 0$$

□

Definition 5.5.7. Let X be an LCH space. A pre-unitary representation (π, V) of $\mathscr{Bor}_{\mathrm{b}}(X)$ (which is automatically positive) is called a **Radon pre-unitary representation** if for each $\xi \in V$, the linear functional

$$\mathscr{Bor}_{\mathrm{b}}(X) \to \mathbb{C} \qquad f \mapsto \langle \xi | \pi(f)\xi \rangle \tag{5.27}$$

is Radon in the sense of Rem. 5.3.6.

Definition 5.5.8. More generally, consider the case where X is a topological space. A pre-unitary representation (π, V) of $\mathscr{Bor}_{\mathrm{b}}(X)$ is called a **normal pre-unitary representation** if for each $\xi \in V$, the linear functional (5.27) is normal, that is,

$$\lim_n \langle \xi | \pi(f_n)\xi \rangle = \langle \xi | \pi(f)\xi \rangle$$

for any increasing sequence (f_n) in $\mathscr{Bor}_{\mathrm{b}}(X)$ converging pointwise to $f \in \mathscr{Bor}_{\mathrm{b}}(X)$.

By Prop. 5.3.2 and Thm. 1.7.8, when X is a second-countable compact Hausdorff space, a pre-unitary representation of $\mathscr{Bor}_{\mathrm{b}}(X)$ is Radon iff it is normal. □

Remark 5.5.9. Let X be a topological space, and let (π, V) be a normal pre-unitary representation of $\mathscr{Bor}_{\mathrm{b}}(X)$. By Prop. 5.3.2, the normality of π is equivalent to the fact that for each $\xi \in V$, there is a unique finite Borel measure $\boldsymbol{\mu_\xi}$ (called the **measure associated to ξ (and π)**) such that

$$\langle \xi | \pi(f)\xi \rangle = \int_X f d\mu_\xi \qquad \text{for each } f \in \mathscr{Bor}_{\mathrm{b}}(X)$$

For each $\xi, \eta \in V$ we define the complex Borel measure

$$\boldsymbol{\mu_{\eta,\xi}} := \frac{1}{4} \sum_{t=0,\frac{\pi}{2},\pi,\frac{3\pi}{2}} e^{\mathrm{i}t} \cdot \mu_{e^{\mathrm{i}t}\eta+\xi}$$

Then, by the polarization identity, $\mu_{\eta,\xi}$ satisfies

$$\langle \eta | \pi(f)\xi \rangle = \int_X f d\mu_{\eta,\xi} \qquad \text{for each } f \in \mathscr{Bor}_{\mathrm{b}}(X)$$

We call $\mu_{\eta,\xi}$ the **complex measure associated to η, ξ (and π)**.

5.5.3 Operator-valued abstract Hausdorff moment problem

In this subsection, unless otherwise stated, X denotes an LCH space. Our goal is to prove Thm. 5.5.13. We first need some preparation.

Definition 5.5.10. The **universal L^2-topology** on $\mathscr{Bor}_{\mathrm{b}}(X)$ is defined to be the pull-back topology of the map

$$\bigvee_{\mu}\Psi_\mu:\mathscr{Bor}_{\mathrm{b}}(X)\to\prod_{\mu}L^2(X,\mu)$$

where the product ranges over all finite Radon measures μ on X, (equivalently, all finite Borel measures when X is second countable, cf. Thm. 1.7.8), and where for each such μ, the map $\Psi_\mu:\mathscr{Bor}_{\mathrm{b}}(X)\to L^2(X,\mu)$ sends each f to (the equivalence class of) f.

Therefore, the universal L^2-topology is the unique topology such that a net (f_α) in $\mathscr{Bor}_{\mathrm{b}}(X)$ converges to $f\in\mathscr{Bor}_{\mathrm{b}}(X)$ iff

$$\lim_\alpha\int_X|f-f_\alpha|^2d\mu=0$$

for each finite Radon measure μ on X, iff

$$\lim_\alpha\Lambda(|f-f_\alpha|^2)=0$$

for each Radon positive linear functional Λ on $\mathscr{Bor}_{\mathrm{b}}(X)$. □

Remark 5.5.11. Suppose that (f_α) is a net in $\mathscr{Bor}_{\mathrm{b}}(X)$ converging universally-L^2 to $f\in\mathscr{Bor}_{\mathrm{b}}(X)$. Then for each finite Radon measure μ and each $g\in\mathscr{Bor}_{\mathrm{b}}(X)$, we have $\lim_\alpha\int_X f_\alpha gd\mu=\int_X fgd\mu$ (since (f_α) converges weakly to f in $L^2(X,\mu)$). In other words,

$$\lim_\alpha\Lambda(f_\alpha g)=\Lambda(fg)\tag{5.28}$$

for each finite Radon positive linear functional Λ on $\mathscr{Bor}_{\mathrm{b}}(X)$ and $g\in\mathscr{Bor}_{\mathrm{b}}(X)$.

Lemma 5.5.12. *Let $\mathscr{A}$ be a *-$\mathbb{F}$-subalgebra of $C_0(X,\mathbb{F})$ separating points of X and vanishing nowhere. Then $\overline{B}_{\mathscr{A}}(0,1)$ is dense in $\overline{B}_{\mathscr{Bor}_{\mathrm{b}}(X,\mathbb{F})}(0,1)$ in the universal L^2-topology, i.e.,*

$$\{f\in\mathscr{A}:\|f\|_{l^\infty}\leqslant1\}\qquad\textit{is dense in}\qquad\{f\in\mathscr{Bor}_{\mathrm{b}}(X,\mathbb{F}):\|f\|_{l^\infty}\leqslant1\}$$

Therefore, for any $f\in\mathscr{Bor}_{\mathrm{b}}(X,\mathbb{F})$ there exists a net (f_α) in $\mathscr{A}$ with $\sup_\alpha\|f_\alpha\|_{l^\infty}\leqslant\|f\|_{l^\infty}$ satisfying $\lim_\alpha\|f-f_\alpha\|_{L^2(X,\mu)}=0$ for each finite Radon measure μ on X.

Proof. By the Stone-Weierstrass Thm. 1.5.13, $\mathscr{A}$ is l^∞-dense in $C_0(X,\mathbb{F})$. Therefore, for each $f\in C_0(X,\mathbb{F})$ with $\|f\|_{l^\infty}\leqslant 1$ there is a net (f_α) in $\mathscr{A}$ converging uniformly to f. In particular, $\|f_\alpha\|_{l^\infty}$ converges to $\|f\|_{l^\infty}$. Therefore $f_\alpha\cdot\|f\|_{l^\infty}/\|f_\alpha\|_{l^\infty}$ converges uniformly to f. This proves that the closed unit ball of $\mathscr{A}$ is l^∞-dense (and hence universally L^2-dense) in the closed unit ball of $C_0(X,\mathbb{F})$.

It remains to prove that $\overline{B}_{C_0(X,\mathbb{F})}(0,1)$ is universally L^2-dense in $\overline{B}_{\mathscr{B}or_{\mathrm{b}}(X,\mathbb{F})}(0,1)$. Choose any $f\in\overline{B}_{\mathscr{B}or_{\mathrm{b}}(X,\mathbb{F})}(0,1)$. Let $\mathscr{I}$ be the set of finite Radon measures on X. For each finite subset $E\subset\mathscr{I}$ and $\varepsilon>0$, by applying Lusin's Thm. 1.7.9 (together with the Tietze extension Thm. 1.4.29) to the finite Radon measure $\mu_E:=\sum_{\mu\in E}\mu$, there exists $f_{E,\varepsilon}\in C_c(X,\mathbb{F})$ such that $\|f_{E,\varepsilon}\|_{l^\infty}\leqslant 1$ and

$$\mu\{x\in X:f(x)\neq f_{E,\varepsilon}\}\leqslant\varepsilon\qquad\text{for each }\mu\in E$$

Therefore

$$\int_X|f-f_{E,\varepsilon}|^2d\mu\leqslant 4\varepsilon\qquad\text{for each }\mu\in E$$

Hence, the net $(f_{E,\varepsilon})_{(E,\varepsilon)\in\mathrm{fin}(2^{\mathscr{I}})\times\mathbb{R}_{>0}}$ (where $\mathbb{R}_{>0}$ is given by the order $\geqslant$ instead of the usual one $\leqslant$) is a net in $\overline{B}_{C_0(X,\mathbb{F})}(0,1)$ and converges universally-L^2 to f. □

Theorem 5.5.13 (Operator-valued abstract Hausdorff moment theorem). *Let X be a compact Hausdorff space. Let $\mathscr{A}$ be a unital *-subalgebra of $C(X)$ separating points of X. Then there is a one-to-one correspondence between the following two classes of objects:* [6]

(1) Radon unitary representations of $\mathscr{B}or_{\mathrm{b}}(X)$;

(2) positive unitary representations of $\mathscr{A}$,

such that $(\pi,\mathcal{H})$ in the first class corresponds to $(\pi|_{\mathscr{A}},\mathcal{H})$ in the second class.

Let us repeat that when X is second countable, a unitary representation of $\mathscr{B}or_{\mathrm{b}}(X)$ is Radon iff it is normal.

Proof. Step 1. An object $(\pi,\mathcal{H})$ in the first class clearly restricts to an object $(\pi|_{\mathscr{A}},\mathcal{H})$ in the second class. Moreover, if $(\pi_1,\mathcal{H})$ and $(\pi_2,\mathcal{H})$ satisfy $\pi_1|_{\mathscr{A}}=\pi_2|_{\mathscr{A}}$, then for each $\xi\in\mathcal{H}$, the Radon positive linear functionals $\mathscr{A}\to\mathbb{C}$ defined by $f\mapsto\langle\xi|\pi_1(f)\xi\rangle$ and by $f\mapsto\langle\xi|\pi_2(f)\xi\rangle$ are equal on $\mathscr{A}$. Therefore, by Cor. 5.3.5 (and Rem. 5.3.6), these two linear functionals are equal. This proves $\pi_1=\pi_2$.

We have thus proved that the correspondence $(\pi,\mathcal{H})\mapsto(\pi|_{\mathscr{A}},\mathcal{H})$ from the first to the second class is injective. To prove that it is surjective, let us prove that any positive unitary representation $(\pi,\mathcal{H})$ of $\mathscr{A}$ can be extend to a Radon unitary representation $(\widehat{\pi},\mathcal{H})$ of $\mathscr{B}or_{\mathrm{b}}(X)$.

[6] Since these two classes are proper classes rather than sets, we avoid using the term "bijection".

Step 2. Fix $(\pi,\mathcal{H})$ in the second class. By Cor. 5.3.5, for each $\xi\in\mathcal{H}$, the positive linear functional

$$\Lambda_\xi:\mathscr{A}\rightarrow\mathbb{C}\qquad f\mapsto\langle\xi|\pi(f)\xi\rangle \tag{5.29}$$

can be extended uniquely to a normal positive linear functional

$$\Lambda_\xi:\mathscr{Bor}_{\mathrm{b}}(X)\rightarrow\mathbb{C}$$

The extension $(\widehat{\pi},\mathcal{H})$ will be constructed from the map

$$\begin{gathered}\Phi:\mathcal{H}^{\complement}\times\mathscr{Bor}_{\mathrm{b}}(X)\times\mathcal{H}\rightarrow\mathbb{C}\\ \Phi(\overline{\eta},f,\xi)=\frac{1}{4}\sum_{t=0,\frac{\pi}{2},\pi,\frac{3\pi}{2}}\Lambda_{e^{\mathbf{i}t}\eta+\xi}(f)\cdot e^{\mathbf{i}t}\end{gathered} \tag{5.30}$$

By the polarization identity, we have

$$\Phi(\overline{\eta},f,\xi)=\langle\eta|\pi(f)\xi\rangle\qquad\text{if } f\in\mathscr{A} \tag{5.31}$$

From the definition of Φ, and by Rem. 5.5.11, for each fixed $\xi,\eta\in\mathcal{H}$,

$$\begin{gathered}\text{if } (f_\alpha) \text{ is a net in } \mathscr{Bor}_{\mathrm{b}}(X) \text{ converging to } f\\ \text{in the universal } L^2\text{-topology, then } \lim_\alpha\Phi(\overline{\eta},f_\alpha,\xi)=\Phi(\overline{\eta},f,\xi)\end{gathered} \tag{5.32}$$

Therefore, for each $f\in\mathscr{Bor}_{\mathrm{b}}(X)$, if we choose a net (f_α) in $\mathscr{A}$ converging universally-L^2 to f (cf. Lem. 5.5.12), then the bilinearity of $(\overline{\eta},\xi)\mapsto\Phi(\overline{\eta},f_\alpha,\xi)$ (due to (5.31)) implies the bilinearity of $(\overline{\eta},\xi)\mapsto\Phi(\overline{\eta},f,\xi)$. Clearly Φ is also linear on the second variable. We have thus proved that Φ is a trilinear map.

Step 3. In this step, we construct the map $\widehat{\pi}$. Observe that

$$\Phi(\overline{\xi},f,\xi)\xlongequal{(5.32)}\lim_\alpha\Phi(\overline{\xi},f_\alpha,\xi)\xlongequal{(5.31)}\lim_\alpha\langle\xi|\pi(f_\alpha)\xi\rangle\xlongequal{(5.29)}\lim_\alpha\Lambda_\xi(f_\alpha)$$

By Rem. 5.5.11, the last term above equals $\Lambda_\xi(f)$. Therefore

$$\Phi(\overline{\xi},f,\xi)=\Lambda_\xi(f)\qquad\text{for all } f\in\mathscr{Bor}_{\mathrm{b}}(X) \tag{5.33}$$

Since Λ_ξ is positive, by Prop. 4.5.1, we have

$$|\Phi(\overline{\xi},f,\xi)|=|\Lambda_\xi(f)|\leqslant\|f\|_{l^\infty}\cdot\Lambda_\xi(1)=\|f\|_{l^\infty}\cdot\langle\xi|\pi(1)\xi\rangle=\|f\|_{l^\infty}\cdot\|\xi\|^2$$

It follows from Prop. 3.2.12 that the trilinear map $\Phi(-,f,-):\mathcal{H}^{\complement}\times\mathcal{H}\rightarrow\mathbb{C}$ is bounded. In other words, $\Phi(-,f,-)$ is a bounded sesquilinear form. Due to the isomorphism $\mathfrak{Ses}(\mathcal{H})\simeq\mathfrak{L}(\mathcal{H})$ (cf. Thm. 3.5.5), Φ gives rise to a linear map

$$\widehat{\pi}:\mathscr{Bor}_{\mathrm{b}}(X)\rightarrow\mathfrak{L}(\mathcal{H})\qquad\langle\eta|\widehat{\pi}(f)\xi\rangle=\Phi(\overline{\eta},f,\xi) \tag{5.34}$$

By (5.31), we have $\widehat{\pi}|_{\mathscr{A}} = \pi$.

Step 4. It remains to check that $(\widehat{\pi}, \mathcal{H})$ is a Radon unitary representation of $\mathscr{B}or_{\mathrm{b}}(X)$. By (5.33) and (5.34), we have

$$\langle\xi|\widehat{\pi}(f)\xi\rangle = \Lambda_\xi(f) \qquad \text{for all } f \in \mathscr{B}or_{\mathrm{b}}(X) \tag{5.35}$$

The Radon property of $\widehat{\pi}$ follows from (5.35) and the Radon property of Λ. To check $\widehat{\pi}(f^*) = \widehat{\pi}(f)^*$ for each $f \in \mathscr{B}or_{\mathrm{b}}(X)$, by linearity, it suffices to check that $\widehat{\pi}(f)^* = \widehat{\pi}(f)$ for all $f \in \mathscr{B}or_{\mathrm{b}}(X, \mathbb{R}_{\geqslant 0})$. In fact, we have $\widehat{\pi}(f) \geqslant 0$ due to (5.35) and the positivity of Λ_ξ. Since $\widehat{\pi}$ extends $\mathscr{A}$, we have $\widehat{\pi}(1) = \mathrm{id}_{\mathcal{H}}$.

It remains to check that $\widehat{\pi}(fg) = \widehat{\pi}(f)\widehat{\pi}(g)$ for all $f, g \in \mathscr{B}or_{\mathrm{b}}(X)$. To do this, note that by (5.34), the property (5.32) can be rephrased as follows:

$$\begin{gathered}\text{If } (f_\alpha) \text{ is a net in } \mathscr{B}or_{\mathrm{b}}(X) \text{ converging to } f\\ \text{in the universal } L^2\text{-topology, then } \lim_\alpha \widehat{\pi}(f_\alpha) = \widehat{\pi}(f) \text{ in WOT.}\end{gathered} \tag{5.36}$$

Now, choose nets $(f_\alpha)_{\alpha\in\mathscr{I}}$ and $(g_\beta)_{\beta\in\mathscr{J}}$ in $\mathscr{B}or_{\mathrm{b}}(X)$ converging uniformly-L^2 to f and g, respectively. Since π is multiplicative, we have $\widehat{\pi}(f_\alpha g_\beta) = \widehat{\pi}(f_\alpha)\widehat{\pi}(g_\beta)$. Clearly, for each $h \in \mathscr{B}or_{\mathrm{b}}(X)$, the nets $(f_\alpha h)$ and $(g_\beta h)$ converge uniformly-L^2 to fh and gh, respectively. It follows from (5.36) and the following Lem. 5.5.14 that

$$\widehat{\pi}(fg_\beta) = \lim_\alpha \widehat{\pi}(f_\alpha g_\beta) = \lim_\alpha \widehat{\pi}(f_\alpha)\widehat{\pi}(g_\beta) = \widehat{\pi}(f)\widehat{\pi}(g_\beta)$$

in WOT, and hence

$$\widehat{\pi}(fg) = \lim_\beta \widehat{\pi}(fg_\beta) = \lim_\beta \widehat{\pi}(f)\widehat{\pi}(g_\beta) = \widehat{\pi}(f)\widehat{\pi}(g)$$

This finishes the proof. □

Lemma 5.5.14. *Let $\mathcal{H}, \mathcal{K}$ be Hilbert spaces. Suppose that (T_α) is a net in $\mathfrak{L}(\mathcal{H}, \mathcal{K})$ converging in WOT to $T \in \mathfrak{L}(\mathcal{H}, \mathcal{K})$. Let $A \in \mathfrak{L}(\mathcal{H})$ and $B \in \mathfrak{L}(\mathcal{K})$. Then $(T_\alpha A)$ converges in WOT to TA, and (BT_α) converges in WOT to BT.*

Proof. For each $\xi \in \mathcal{H}$ and $\eta \in \mathcal{K}$, we have $\lim_\alpha\langle\eta|T_\alpha A\xi\rangle = \langle\eta|TA\xi\rangle$ and

$$\lim_\alpha\langle\eta|BT_\alpha\xi\rangle = \lim_\alpha\langle B^*\eta|T_\alpha\xi\rangle = \langle B^*\eta|T\xi\rangle = \langle\eta|BT\xi\rangle$$

□

We end this subsection with an easy observation that, thanks to the multiplicativity of $\widehat{\pi}$, the WOT convergence in (5.36) can in fact be strengthened to SOT convergence:

Proposition 5.5.15. *Let (π, V) be a Radon pre-unitary representation of $\mathscr{B}or_{\mathrm{b}}(X)$. Suppose that (f_α) is a net in $\mathscr{B}or_{\mathrm{b}}(X)$ converging to $f \in \mathscr{B}or_{\mathrm{b}}(X)$ in the universal L^2-topology. Then $\lim_\alpha \pi(f_\alpha)$ converges in SOT to $\pi(f)$.*

Proof. Choose any $\xi \in V$. Since $\Lambda_\xi : g \in \mathscr{Bor}_{\mathrm{b}}(X) \mapsto \langle \xi | \pi(g)\xi \rangle$ is a Radon positive linear funtional, we have

$$\begin{aligned}&\|\pi(f)\xi - \pi(f_\alpha)\xi\|^2 = \langle \pi(f - f_\alpha)\xi | \pi(f - f_\alpha)\xi \rangle = \langle \xi | \pi(f - f_\alpha)^* \pi(f - f_\alpha)\xi \rangle \\ =&\langle \xi | \pi(|f - f_\alpha|^2)\xi \rangle = \Lambda_\xi(|f - f_\alpha|^2) \to 0\end{aligned}$$

□

The proof of Thm. 5.5.13 also indicates that $\widehat{\pi} : \mathscr{Bor}_{\mathrm{b}}(X) \to \mathfrak{L}(\mathcal{H})$ is bounded. Therefore, if (f_α) converges uniformly to f, then $\lim_\alpha \pi(f_\alpha)$ converges in the operator norm to $\pi(f)$. We record this fact in a more general form, as the operator-valued analogue of Prop. 4.5.1.

Proposition 5.5.16. *Let X be a non-empty set. Let $\mathscr{A}$ be a unital *-subalgebra of $l^\infty(X)$, equipped with the l^∞-norm. Let (π, V) be a positive pre-unitary representation of $\mathscr{A}$. Then $\pi(\mathscr{A}) \subset \mathfrak{L}(V)$, and the linear map $\pi : \mathscr{A} \to \mathfrak{L}(V)$ is bounded with operator norm $\|\pi\| = 1$.*

Proof. By Prop. 4.5.1, for each $f \in \mathscr{A}$ and $\xi \in V$ we have

$$|\langle \xi | \pi(f)\xi \rangle| \leqslant \|f\|_{l^\infty} \cdot \langle \xi | \pi(1)\xi \rangle = \|f\|_{l^\infty} \cdot \|\xi\|^2$$

Thus

$$\langle \pi(f)\xi | \pi(f)\xi \rangle = \langle \xi | \pi(f^* f)\xi \rangle \leqslant \|f^* f\|_{l^\infty} \cdot \|\xi\|^2 = \|f\|_{l^\infty}^2 \cdot \|\xi\|^2$$

and hence $\|\pi(f)\| \leqslant \|f\|_{l^\infty}$. This proves that $\pi(\mathscr{A}) \subset \mathfrak{L}(V)$, and that the linear map $\pi : \mathscr{A} \to \mathfrak{L}(V)$ satisfies $\|\pi\| \leqslant 1$. Since $\|\pi(1)\| = 1$, we have $\|\pi\| = 1$. □

5.5.4 Borel functional calculus for bounded self-adjoint operators

Let $T \in \mathfrak{L}(\mathcal{H})$ be self-adjoint. Choose $r \in \mathbb{R}_{\geqslant 0}$ such that $\|T\| \leqslant r$. In this subsection we establish, in Thm. 5.5.17, the Borel functional calculus for T, which can be viewed the operator-valued (classical) Hausdorff moment theorem.

As we shall see, Thm. 5.5.17 follows directly from the operator-valued abstract Hausdorff moment Thm. 5.5.13, once a subtle difference has been addressed: namely, the distinction between the abstract Hausdorff moment theorem (Cor. 5.3.5)—applied in the special case $X = I = [a, b]$ with $-\infty < a < b < +\infty$ and $\mathscr{A}$ the polynomial algebra $\mathbb{C}[x]$ (viewed as a unital *-subalgebra of $C([-r, r])$ [7])—and the classical Hausdorff moment Thm. 4.2.9.

The solvability condition for the (classical) Hausdorff moment problem $\int_I x^n d\rho = c_n$ is that $H, H', H - H'$ are all positive. Define the linear functional

$$\Lambda : \mathscr{A} = \mathbb{C}[x] \to \mathbb{C} \qquad x^n \mapsto c_n$$

[7]To view $\mathbb{C}[x]$ as a subset of $C([-r, r])$ we must assume that $r > 0$. However, the case $r = 0$ is trivial since T must be zero.

Then the positivity of $H, H', H - H'$ means that $\Lambda(f) \geqslant 0$ whenever $f \in \mathbb{C}[x]$ has one of the following three forms:

$$f(x) = p(x)^* p(x) \qquad f(x) = xp(x)^* p(x) \qquad f(x) = (1-x)p(x)^* p(x)$$

with $p \in \mathbb{C}[x]$. Recall that this is the condition in the special case $I = [0,1]$. In the general case $I = [a,b]$, the corresponding solvability condition is that $\Lambda(f) \geqslant 0$ for all f of one of the following three forms (with $p \in \mathbb{C}[x]$):

$$f(x) = p(x)^* p(x) \qquad f(x) = (x-a)p(x)^* p(x) \qquad f(x) = (b-x)p(x)^* p(x) \tag{5.37}$$

By contrast, in the abstract Hausdorff moment theorem (Cor. 5.3.5), the condition required of Λ is that $\Lambda(f) \geqslant 0$ for all f belonging to the set

$$\{f \in \mathbb{C}[x], f|_I \geqslant 0\} \tag{5.38}$$

This condition is indeed equivalent to one above, since the set (5.38) is $\mathbb{R}_{\geqslant 0}$-spanned by polynomials of the form in (5.37); see Pb. 5.1.

We will not rely on this equivalence to establish the Borel functional calculus, because this method does not extend to the setting of several mutually commuting bounded self-adjoint operators. Instead, following Riesz's original idea in [Rie13], we will use finite-rank approximations of T (as in Subsec. 4.2.6 or the proof Hilbert's spectral Thm. 5.1.1) to prove the operator-valued version of this equivalence: the condition $-r \leqslant T \leqslant r$ is equivalent to $f(T) \geqslant 0$ for all $f \in \mathbb{C}[x]$ satisfying $f|_{[-r,r]} \geqslant 0$. [8]

Theorem 5.5.17 (Borel functional calculus). *There exists a unique normal unitary representation*

$$\pi_T : \mathscr{B}or_{\mathrm{b}}([-r,r]) \to \mathfrak{L}(\mathcal{H})$$

sending the function $x = \mathrm{id}_{[-r,r]}$ to T.

We call $(\pi_T, \mathcal{H})$ the **Borel functional calculus** of T. Its restriction to $C([-r,r])$ is called the **continuous functional calculus**. We write

$$\boldsymbol{f(T)} := \pi_T(f)$$

Proof. Let $\mathscr{A}$ be the set of polynomials on $[-r,r]$ (viewed as functions on $[-r,r]$). Then $\mathscr{A}$ is a unital *-subalgebra of $C([-r,r])$ separating points of $[-r,r]$. There is a unique unitary representation $(\pi_T, \mathcal{H})$ of $\mathscr{A}$ sending x to T. If we can prove that this representation is positive, then by Thm. 5.5.13, it can be extended uniquely to a normal unitary representation of $\mathscr{B}or_{\mathrm{b}}([-r,r])$, finishing the proof.

[8] The direction "$\Leftarrow$" follows directly by choosing $f(x) = x + r$. Therefore, it suffices to prove "$\Rightarrow$".

Let us prove that π_T is positive on $\mathscr{A}$, that is, for any polynomial f satisfying $f|_{[-r,r]} \geqslant 0$, the operator $f(T) := \pi_T(f)$ is positive. Let $(e_i)_{i\in\mathscr{I}}$ be an orthonormal basis of $\mathcal{H}$. For each $\alpha \in \mathrm{fin}(2^{\mathscr{I}})$, let E_α be the projection onto $\mathrm{Span}\{e_i : i \in \alpha\}$. By Exp. 3.7.9 (or more generally, by Thm. 5.2.7), $(E_\alpha)_{\alpha\in\mathrm{fin}(2^{\mathscr{I}})}$ is an increasing net of projections converging in SOT to $\mathrm{id}_{\mathcal{H}}$. By Exp. 3.7.17, the net $(f(E_\alpha T E_\alpha))_{\alpha\in\mathrm{fin}(2^{\mathscr{I}})}$ converges in SOT to $f(T)$. Therefore, it suffices to show that each $f(E_\alpha T E_\alpha)$ is positive.

The linear operator

$$T_\alpha = E_\alpha T E_\alpha\big|_{\mathrm{Rng}(E_\alpha)}$$

is self-adjoint since $\langle\xi|E_\alpha T E_\alpha\xi\rangle = \langle E_\alpha\xi|TE_\alpha\xi\rangle \in \mathbb{R}$. Therefore, by finite-dimensional linear algebra, T_α is diagonalizable. Thus, there exists an orthonormal basis $v_1, \dots, v_n$ of $\mathrm{Rng}(E_\alpha)$ such that $T_\alpha v_i = \lambda_i v_i$ with $\lambda_i \in \mathbb{R}$.

Extend $\{v_1, \dots, v_n\}$ to an orthonormal basis of $\mathcal{H}$ (cf. Exe. 3.4.10). Then, under this basis, the matrix representation of $E_\alpha T E_\alpha$ (cf. Def. 3.6.2) is diagonal, with the first n-diagonal terms being $\lambda_1, \dots, \lambda_n$ and the remaining terms being 0. Since $\|E_\alpha T E_\alpha\| \leqslant \|T\| \leqslant r$, we must have

$$\lambda_i \in [-r, r]$$

Therefore, the condition $f|_{[-r,r]} \geqslant 0$ implies that $f(\lambda_i) \geqslant 0$.

The matrix representation of $f(E_\alpha T E_\alpha)$ is also diagonal, with the first n-diagonal terms being $f(\lambda_1), \dots, f(\lambda_n)$ (all $\geqslant 0$) and the remaining terms being 0. Thus $f(E_\alpha T E_\alpha) \geqslant 0$. □

5.5.5 Proof of the Riesz spectral Thm. 5.4.4

Proof of Thm. 5.4.4. Define

$$E(\lambda) = \chi_{[-r,\lambda]}(T) = \pi_T(\chi_{[-r,\lambda]})$$

Then, since π_T is a unitary representation, and since $\chi_{[-r,\lambda]} = (\chi_{[-r,\lambda]})^* = (\chi_{[-r,\lambda]})^2$, we conclude that $E(\lambda)$ is a projection. Since π_T is positive (cf. Prop. 5.5.6), whenever $-r \leqslant \lambda_1 \leqslant \lambda_2 \leqslant r$ we have $\chi_{[-r,\lambda_2]} - \chi_{[-r,\lambda_1]} \geqslant 0$ and hence $E(\lambda_2) - E(\lambda_1) \geqslant 0$. This proves that the net $E := (E(\lambda))_{\lambda\in[-r,r]}$ is increasing.

Since π_T is normal, the linear functional $\Lambda_\xi : f \mapsto \langle\xi|f(T)\xi\rangle$ is normal (and positive). Therefore, by MCT or DCT (for first-countable nets, cf. Thm. 1.2.36 or 1.2.37), we have

$$\lim_{\lambda\searrow\lambda_0} \langle\xi|E(\lambda)\xi\rangle = \langle\xi|E(\lambda_0)\xi\rangle$$

That is, $\lim_{\lambda\searrow\lambda_0 E(\lambda)}$ converges in SOT to $E(\lambda_0)$. We have thus proved that E is right-continuous.

In view of Prop. 5.5.2 (and the polarization identity), to prove $f(T) = \int_{[-r,r]} f(\lambda)dE(\lambda)$, it suffices to show for each $\xi \in \mathcal{H}$ and $f \in C([-r,r])$ that

$$\Lambda_\xi(f) = \int_{[-r,r]} f(\lambda)d\langle\xi|E(\lambda)\xi\rangle$$

Since Λ_ξ is normal, by Prop. 5.3.2, we have

$$\Lambda_\xi(f) = \int_{[-r,r]} f d\mu_\xi \qquad \text{for each } f \in \mathscr{B}or_{\mathrm{b}}([-r,r])$$

where $\mu_\xi : \mathfrak{B}_{[-r,r]} \to \mathbb{R}_{\geqslant 0}$ is given by $\mu_\xi(E) = \Lambda_\xi(\chi_E)$. Recall from Thm. 1.10.1 that when $f \in C([-r,r])$, the integral $\int_{[-r,r]} f d\mu_\xi$ can be expresses by the Stieltjes integral of f against the increasing function sending each $\lambda \in [-r,r]$ to

$$\mu_\xi([-r,\lambda]) = \Lambda_\xi(\chi_{[-r,\lambda]}) = \langle\xi|\chi_{[-r,\lambda]}(T)\xi\rangle = \langle\xi|E(\lambda)\xi\rangle$$

This finishes the proof. □

5.6 Some concluding remarks on Riesz's spectral theorem

Let $T \in \mathfrak{L}(\mathcal{H})$ be self-adjoint where $\mathcal{H}$ is a Hilbert space. Let $r \geqslant \|T\|$.

Remark 5.6.1. We proved the Riesz spectral Thm. 5.4.4 by establishing the Borel functional calculus (Thm. 5.5.17). However, as noted in the answer to Question 5.4.1, it actually suffices—just as Riesz himself did in [Rie13]—to establish the **semicontinuous functional calculus**. That is, one proves the existence of a (unique) unitary representation

$$\pi_T : \mathrm{Span}_{\mathbb{C}}\mathrm{LSC}_{\mathrm{b}}([-r,r],\mathbb{R}_{\geqslant 0}) \to \mathfrak{L}(\mathcal{H})$$

(where $\mathrm{LSC}_{\mathrm{b}}(X,\mathbb{R}_{\geqslant 0})$ denotes the set of bounded lower-semicontinuous functions $X \to \mathbb{R}_{\geqslant 0}$) satisfying

$$\lim_n \langle\xi|\pi_T(f_n)\xi\rangle = \langle\xi|\pi_T(f)\xi\rangle$$

for each $\xi \in \mathcal{H}$ and each increasing sequence (f_n) in $\mathrm{LSC}_{\mathrm{b}}(X,\mathbb{R}_{\geqslant 0})$ converging pointwise to $f \in \mathrm{LSC}_{\mathrm{b}}(X,\mathbb{R}_{\geqslant 0})$.

One reason Riesz did not use the Borel functional calculus in [Rie13] is simply that measure theory had not yet matured at the time. In fact, the spectral theorem itself was a major driving force behind the development of modern measure theory.

For a more faithful presentation of Riesz's original proof of his spectral theorem in [Rie13], see Sec. 25.8 and 25.9 of [Gui-A]. We will discuss the origin of the Borel functional calculus in Subsec. 5.11.2. □

Remark 5.6.2. Although Riesz's treatment of the spectral theorem is framed primarily within the paradigm of linear extension and largely avoids the perspective of finite approximation, the latter is not completely absent: it is used to prove the positivity of the polynomial functional calculus, as seen in the proof of Thm. 5.5.17.

Modern approaches to the spectral theorem often abandon the finite approximation paradigm completely. One motivation is to develop methods that apply equally well to operators on Banach spaces, not just on Hilbert spaces. In these modern treatments, the primary non-trivial step—apart from applying the Riesz(-Markov) representation theorem—is still to prove the positivity of the polynomial functional calculus. The following is a list of mainstream approaches, along with the key theorems used to establish this crucial positivity.

- The approach using the notion of spectrum

  $$\sigma(T) = \{z \in \mathbb{C} : (z - T) \text{ is not invertible in } \mathfrak{L}(T)\}$$

  The key theorems are **Gelfand's spectral radius theorem**

  $$\sup\{|\lambda| : \lambda \in \sigma(T)\} = \lim_{n\to\infty} \|T^n\|^{1/n}$$

  (for any $T \in \mathfrak{L}(\mathcal{H})$, not necessarily self-adjoint) and the **spectral mapping theorem**: $\sigma(f(T)) \subset f(\sigma(T))$ for each $f \in \mathbb{C}[x]$. See [Lax, RS-1].

- The C^*-algebra approach. The key theorem is the (commutative) **Gelfand-Naimark theorem**, which states that any commutative normed-closed unital $*$-subalgebra of $\mathfrak{L}(\mathcal{H})$ is isometically isomorphic to $C(X)$ for a compact Hausdorff space X. This theorem is applied to the subalgebra generated by T, so that T can be viewed as a continuous function on X with range in $[-r, r]$. See [Rud-F].

- The approach using **holomorphic functional calculus**

  $$f(T) = \oint_C \frac{f(z)}{z - T} \cdot \frac{dz}{2\pi \mathbf{i}}$$

  where C is a suitable closed curve in $\mathbb{C}$ surrounding $[-r, r]$. The key theorem is $(fg)(T) = f(T)g(T)$ and $f^\dagger(T) = f(T)^*$ (where $f^\dagger(z) = \overline{f(\overline{z})}$), and that any holomorphic function f, defined and strictly positive on a neighborhood of $[-r, r]$, can be written as $h^\dagger h$ with h holomorphic. This allows the argument of Prop. 5.5.6 to apply. See [Gui-S].

- The approach reducing the spectral theorem for self-adjoint operators to that of unitary operators. The crucial point is that positivity of the polynomial functional calculus is easier to establish for unitary operators, due to the (nontrivial) fact that any strictly positive polynomial polynomial f on $\mathbb{S}^1$ can be written as g^*g for some polynomial g. This again allows the argument of Prop. 5.5.6 to go through. See [Xia].

□

5.7 Borel functional calculus for bounded normal operators

In this section, we fix a Hilbert space $\mathcal{H}$.

Definition 5.7.1. An operator $T \in \mathfrak{L}(\mathcal{H})$ is called **normal** if it satisfies $T^*T = TT^*$. This is equivalent to its real part $\mathrm{Re}(T)$ and its imaginary part $\mathrm{Im}(T)$ commuting with each other, where

$$\mathrm{Re}(T) = \frac{T + T^*}{2} \qquad \mathrm{Im}(T) = \frac{T - T^*}{2\mathbf{i}}$$

Example 5.7.2. Let $U \in \mathfrak{L}(\mathcal{H}, \mathcal{K})$ where $\mathcal{H}$ and $\mathcal{K}$ are Hilbert spaces. Then the following are true.

1. U is an isometry iff $U^*U = \mathbf{1}_{\mathcal{H}}$.

2. U is unitary iff $U^*U = \mathbf{1}_{\mathcal{H}}$ and $UU^* = \mathbf{1}_{\mathcal{K}}$.

It follows that unitary operators on $\mathcal{H}$ are normal.

It follows that $U^{-1} = U^*$ if U is a unitary operator between Hilbert spaces.

Proof. For each $\xi, \eta \in \mathcal{H}$, we have

$$\langle U\xi | U\eta \rangle = \langle \xi | U^*U\eta \rangle$$

Thus U is an isometry iff $\langle \xi | \eta \rangle = \langle \xi | U^*U\eta \rangle$ for each $\xi, \eta \in \mathcal{H}$. This proves the first equivalence.

By the first equivalence, U is unitary iff $U^*U = \mathbf{1}_{\mathcal{H}}$ and U is bijective (in which case the inverse U^{-1} must be an isometry, and hence bounded), iff $U^*U = \mathbf{1}_{\mathcal{H}}$ and U has a bounded linear inverse, iff U^* is the inverse of U, iff $U^*U = \mathbf{1}_{\mathcal{H}}$ and $UU^* = \mathbf{1}_{\mathcal{K}}$. □

5.7.1 Introduction

While Riesz's spectral Thm. 5.4.4 offers important historical context for modern versions of the spectral theorem, it's not the most practical approach for a reader who simply wants to learn how to apply the theorem quickly. The modern versions—specifically, the Borel functional calculus (Thm. 5.5.17) and the multiplication operator version—are more useful.

Proving Riesz's spectral theorem requires first establishing the Borel functional calculus, yet the functional calculus alone is sufficient for practical applications. The additional components in Riesz's theorem, such as the operator-valued Stieltjes integral (Subsec. 5.5.1) or the integral representation of functional calculus (Subsec. 5.5.5), are not essential for applying the theorem. Readers can safely skip these sections on first reading.

Furthermore, given that Riesz's spectral theory marks a transition from the paradigm of finite approximation to one of linear extension (cf. Subsec. 5.3.1), a complete paradigm shift entails abandoning the language of integration in favor of linear functionals—or, in the context of spectral theory, fully adopting the language of the Borel functional calculus. This is the path we will follow from this point forward, having fulfilled the historical mission of Riesz's spectral theorem.

The goal of this section is to generalize the Borel functional calculus for a single bounded self-adjoint operator (Thm. 5.5.17) to several mutually commuting bounded self-adjoint operators (Thm. 5.7.13). There are two main reasons for this generalization. First, in quantum mechanics, two commuting self-adjoint operators—say, A and B—represent two observables that can be measured simultaneously without uncertainty. (See Sec. 6.9.1 for a detailed discussion.) Second, this generalization is necessary to establish a spectral theorem for bounded normal operators. This case includes unitary operators, since every unitary operator is normal. In Ch. 6, we will establish the spectral theorem for unbounded self-adjoint operators by reducing it to the unitary case via the Cayley transform.

5.7.2 Borel functional calculus for mutually commuting bounded self-adjoint operators

Let $T_1,\dots,T_N\in\mathfrak{L}(\mathcal{H})$ be bounded self-adjoint operators. Assume that they commute with each other, i.e., $[T_i,T_j]=0$ for each $1\leqslant i,j\leqslant N$. Choose $R\in\mathbb{R}_{>0}$ such that

$$\|T_i\|\leqslant R$$

Let

$$\mathbb{C}[x_\bullet]=\mathbb{C}[x_1,\dots,x_N]=\{\text{complex polynomials of } x_1,\dots,x_N\}$$

Then we clearly have a unital *-homomophism

$$\pi_{T_\bullet}:\mathbb{C}[x_\bullet]\rightarrow\mathfrak{L}(\mathcal{H})\qquad x_i\mapsto T_i$$

called the **polynomial functional calculus**. This is the map sending each $f\in\mathbb{C}[x_\bullet]$ to $f(T_\bullet)$, e.g., sending $x_1^3x_3-4x_1x_2x_5^6$ to $T_1^3T_3-4T_1T_2T_5^6$.

We view $=\mathbb{C}[x_\bullet]$ as a unital *-subalgebra of $C(X)$ where

$$X=[-R,R]^N$$

As in the proof of Thm. 5.5.17, the main obstacle in applying the operator-valued abstract Hausdorff moment Thm. 5.5.13 is to prove that $\pi_{T_\bullet}$ is positive, i.e., that $f(T_\bullet) \geqslant 0$ for each $f \in \mathbb{C}[x_\bullet]$ satisfying $f|_X \geqslant 0$.

To overcome this difficulty, we adopt a different type of finite approximation, following von Neumann's treatment of the spectral theorem for bounded normal operators in [vN29a, Anhang 2]: The earlier method—approximating T_i by ET_iE where E is the projection onto a finite-dimensional subspace—fails here, since ET_iE and ET_jE need not commute. Instead, we use **finite-spectrum approximation** rather than **finite-rank approximation**. Specifically, we approximate each T_i by $f_i(T_i)$ where $f_i \in \mathscr{Bor}_{\mathrm{b}}([-R_i, R_i])$ has finite range. In this way, the operators $f_i(T_i)$ and $f_j(T_j)$ commute, as shown below.

Lemma 5.7.3. *Let $A, B \in \mathfrak{L}(A)$ be self-adjoint. Choose $a, b \in \mathbb{R}_{\geqslant 0}$ such that $\|A\| \leqslant a$ and $\|B\| \leqslant b$. Then $f(A)$ commutes with $g(B)$ for each $f \in \mathscr{Bor}_{\mathrm{b}}([-a, a])$ and $g \in \mathscr{Bor}_{\mathrm{b}}([-b, b])$.*

Proof. By Lem. 5.5.12, there is a net of polynomials (f_α) of $z_\bullet$ converging universally-L^2 to f. By Prop. 5.5.15, the net $(f_\alpha(A))$ converges in SOT to $f(A)$. Since each $f_\alpha(A)$ commutes with B, $f(A)$ also commutes with B. A similar argument shows that $f(A)$ commutes with $g(B)$. □

Theorem 5.7.4 (Borel functional calculus). *Let $T_1, \dots, T_N \in \mathfrak{L}(\mathcal{H})$ be mutually commuting self-adjoint operators. For each $1 \leqslant i \leqslant N$, choose $R \in \mathbb{R}_{>0}$ such that $\|T_i\| \leqslant R$ for each i. Let*

$$X = [-R, R]^N$$

Then there exists a unique normal unitary representation

$$\pi_{T_\bullet} : \mathscr{Bor}_{\mathrm{b}}(X) \to \mathfrak{L}(\mathcal{H})$$

*sending each x_i to T_i, where x_i the i-th **coordinate function** of $\mathbb{R}^N$, i.e., the projection $\mathbb{R}^N \to \mathbb{R}$ onto the i-th component.*

We call $(\pi_{T_\bullet}, \mathcal{H})$ the **Borel functional calculus** of $T_\bullet$, and write

$$\boldsymbol{f(T_\bullet)} := \pi_{T_\bullet}(f)$$

Proof. Step 1. Let $\mathscr{A}$ be the polynomial algebra $\mathbb{C}[x_\bullet]$, viewed as a unital *-subalgebra of $\mathscr{Bor}_{\mathrm{b}}(X)$. Then there exists a unique unitary representation (namely, the polynomial functional calculus)

$$\pi_{T_\bullet} : \mathscr{A} \to \mathfrak{L}(\mathcal{H})$$

sending each x_i to T_i. We abbreviate $\pi_{T_\bullet}$ to π for simplicity. With the help of Thm. 5.5.13, the proof of the current theorem is reduced to showing that π is positive.

In the following, we prove the positivity of π in the case $N = 2$; the general case follows from the same idea. We write $A = T_1$ and $B = T_2$.

Step 2. Let (g_n) be a sequence of simple functions in $\mathscr{Bor}_{\mathrm{b}}([-R,R],[-R,R])$ converging uniformly to the the identity function. Then, by Prop. 5.5.16, $g_n(A)$ and $g_n(B)$ converge in the operator norm to A and B respectively. Moreover, $g_n(A)$ commutes with $g_n(B)$ due to Lem. 5.7.3. By Prop. 3.5.14, for each $f \in \mathscr{A}$, the sequence $f(g_n(A), g_n(B))$ converges in norm to $f(A,B)$. Therefore, to prove the positivity of π, it remains to prove that $f(g(A), g(B)) \geqslant 0$ for each $f \in \mathscr{A}$ satisfying $f|_X \geqslant 0$, and for each simple function $g \in \mathscr{Bor}_{\mathrm{b}}([-R,R],[-R,R])$.

Step 3. Since g has finite range, we can write

$$g = c_1\chi_{I_1} + \cdots + c_n\chi_{I_n}$$

where $c_i \in [-R,R]$, and $I_1, \dots, I_n \in \mathfrak{B}([-R,R])$ are mutually disjoint with union $[-R,R]$. Then

$$g(A) = \sum_{i=1}^n c_iE_i \qquad g(B) = \sum_{j=1}^n c_jF_j \qquad \mathrm{id}_{\mathcal{H}} = \sum_{i=1}^n E_i = \sum_{j=1}^n F_j$$

where

$$E_i := \chi_{I_i}(A) \qquad F_j := \chi_{I_j}(B)$$

are projections, and

$$E_iE_j = \delta_{i,j}E_i \qquad F_iF_j = \delta_{i,j}F_j \qquad [E_i, F_j] = 0$$

where the first equality is due to $\chi_{I_i}(A)\chi_{I_j}(A) = (\chi_{I_i}\chi_{I_j})(A) = \delta_{i,j}\chi_{I_i}(A)$, the second equality is due to the same reasoning, and last equality is due to Lem. 5.7.3.

Define a linear map

$$\varpi : \mathscr{A} \to \mathfrak{L}(\mathcal{H}) \qquad f \mapsto \sum_{i,j=1}^n f(c_i, c_j)E_iF_j$$

Then $\varpi(1) = \sum_{i,j} E_iF_j = \mathbf{1}$, and

$$\varpi(x_1) = \sum_{i,j} x_1(c_i,c_j)E_iF_j = \sum_{i,j} c_iE_iF_j = \sum_{i,j} c_iE_i = g(A)$$

Similarly, we have $\varpi(x_2) = g(B)$. Clearly $\varpi(\overline{f}) = \varpi(f)^*$. For each $f, g \in \mathscr{A}$, we have

$$\varpi(f)\varpi(h) = \sum_{i,j,k,l} f(c_i,c_j)h(c_k,c_l)E_iF_jE_kF_l = \sum_{i,j,k,l} \delta_{i,k}\delta_{j,l}f(c_i,c_j)h(c_k,c_l)E_iF_j$$

$$= \sum_{i,j} f(c_i, c_j)h(c_i, c_j)E_iF_j = \varpi(fh)$$

Thus, ϖ is the unique unital *-homomorphism sending x_1 to $g(A)$ and x_2 to $g(B)$, that is, ϖ is the polynomial functional calculus of $g(A)$ and $g(B)$. Since $(c_i, c_j) \in X$, for each $f \in \mathscr{A}$ satisfying $f|_X \geqslant 0$ we must have $\varpi(f) \geqslant 0$. □

5.7.3 Borel functional calculus for adjointly-commuting bounded normal operators

Definition 5.7.5. Let $A, B \in \mathfrak{L}(\mathcal{H})$. We say that A **commutes adjointly** with B if

$$[A, B] = [A^*, B] = 0$$

Note that by taking adjoint, the above equalities are equivalent to

$$[A^*, B^*] = [A, B^*] = 0$$

Example 5.7.6. Let $A \in \mathfrak{L}(\mathcal{H})$. Then A commutes adjointly with A iff A is normal.

Example 5.7.7. Let $A, B \in \mathfrak{L}(\mathcal{H})$. The following are clearly equivalent:

(1) A and B are adjointly-commuting bounded normal operators.

(2) The self-adjoint operators $\mathrm{Re}(A), \mathrm{Im}(A), \mathrm{Re}(B), \mathrm{Im}(B)$ commute with each other.

Fuglede's theorem states that two bounded normal operators on $\mathcal{H}$ commute adjointly whenever they commute. We will not need the full result here and record only the following special case:

Example 5.7.8. Let $U \in \mathfrak{L}(\mathcal{H})$ be unitary, and let $T \in \mathfrak{L}(\mathcal{H})$. Then U and T commute iff they commute adjointly.

Proof. Adjoint commutativity clearly implies commutativity. Conversely, assume that $UT = TU$. Then $UTU^{-1} = T$, and hence $TU^{-1} = U^{-1}T$. Since U is unitary, we have $U^{-1} = U^*$, and hence $TU^* = U^*T$. □

The aim of this subsection—and indeed the final goal of this entire section—is to extend Thm. 5.7.4 to the Borel functional calculus for finitely many adjointly-commuting bounded normal operators. To formulate the result in a way that does not depend on the particular underlying space X, we introduce the following notion, which serves as the operator-theoretic analogue of the support of a Borel measure (cf. Def. 1.6.4).

Definition 5.7.9. Let X be a topological space. Let $(\pi,\mathcal{H})$ be a unitary representation of $\mathscr{Bor}_{\mathrm{b}}(X)$. The **support Supp($\pi$)** of π is defined to be

$$\mathrm{Supp}(\pi)=\{x\in X:\pi(\chi_U)\neq 0\text{ for each }U\in\mathrm{Nbh}_X(x)\}$$

Then $\mathrm{Supp}(\pi)$ is a closed subset of X, because we clearly have

$$X\backslash\mathrm{Supp}(\pi)=\bigcup_{U\in\mathcal{T}_X,\pi(\chi_U)=0}U$$

We also have

$$\mathrm{Supp}(\pi)=\mathrm{Cl}_X\Big(\bigcup_{\xi\in\mathcal{H}}\mathrm{Supp}(\mu_\xi)\Big) \tag{5.39}$$

We say that π is **compactly supported** (or that π has **compact support**) if the closed set $\mathrm{Supp}(\pi)$ is compact.

Proof of (5.39). Let $x\in X$. If $x\notin\mathrm{Supp}(\pi)$, then there exists $U\in\mathrm{Nbh}_X(x)$ such that $\pi(\chi_U)=0$, and hence

$$\mu_\xi(\chi_U)=\langle\xi|\pi(\chi_U)\xi\rangle=0$$

for all ξ, where μ_ξ is defined in Rem. 5.5.9. So $U\subset X\backslash\mathrm{Supp}(\mu_\xi)$ for all ξ, and hence x does not belong to the RHS of (5.39).

Conversely, suppose that x does not belong to the RHS of (5.39). There there exists $U\in\mathrm{Nbh}_X(x)$ such that $U\subset X\backslash\mathrm{Supp}(\mu_\xi)$ (and hence $\mu_\xi(\chi_U)=0$) for all ξ. By the above computation, we have $\langle\xi|\pi(\chi_U)\xi\rangle=0$ for all ξ, and hence $\pi(\chi_U)=0$. This proves $x\notin\mathrm{Supp}(\pi)$. $\square$

Remark 5.7.10. Assume that one of the following conditions is satisfied:

(1) X is second countable, and π is normal.

(2) X is LCH, and π is Radon.

Then $\pi(X\backslash\mathrm{Supp}(\pi))=0$. Therefore, $\mathrm{Supp}(\pi)$ is the smallest closed subset such that π vanishes on its complement.

Proof. For each $\xi\in V$, we have $\mathrm{Supp}(\mu_\xi)\subset\mathrm{Supp}(\pi)$ due to (5.39). By Def. 1.6.4, μ_ξ vanishes on $X\backslash\mathrm{Supp}(\mu_\xi)$, and hence on its subset $X\backslash\mathrm{Supp}(\pi)$. Since ξ is arbitrary, we conclude that π vanishes on $X\backslash\mathrm{Supp}(\pi)$. $\square$

Remark 5.7.11. Let $(\pi,\mathcal{H})$ be a unitary representation of $\mathscr{Bor}_{\mathrm{b}}(X)$ satisfying either (1) or (2) of Rem. 5.7.10. Choose any Borel subset $K\subset X$ containing $\mathrm{Supp}(\pi)$. Then for each $f,g\in\mathscr{Bor}_{\mathrm{b}}(X)$ we have

$$f|_K=g|_K\qquad\Longrightarrow\qquad\pi(f)=\pi(g)$$

Proof. By Rem. 5.7.10, we have $\pi(\mathrm{Supp}(\pi)^c) = 0$. Since $0 \leqslant \chi_{K^c} \leqslant \chi_{\mathrm{Supp}(\pi)^c}$, by the positivity of π (Prop. 5.5.6), we have $0 \leqslant \pi(\chi_{K^c}) \leqslant \pi(\chi_{\mathrm{Supp}(\pi)^c})$, and hence $\pi(\chi_{K^c}) = 0$. It follows that

$$\pi(f) = \pi(f\chi_K) + \pi(f\chi_{K^c}) = \pi(f\chi_K) + \pi(f)\pi(\chi_{K^c}) = \pi(f\chi_K)$$

Thus, if $f|_K = g|_K$, then $f\chi_K = g\chi_K$, and hence $\pi(f) = \pi(g)$. □

Proposition 5.7.12. *Assume that X is second-countable, and let $(\pi, \mathcal{H})$ be a normal unitary representation of $\mathscr{Bor}_{\mathrm{b}}(X)$. Let $K \subset X$ be a Borel set containing $\mathrm{Supp}(\pi)$. Then there is a unique normal unitary representation*

$$\boldsymbol{\pi|_K} : \mathscr{Bor}_{\mathrm{b}}(K) \to \mathfrak{L}(\mathcal{H}) \qquad \pi|_K(f|_K) = \pi(f) \tag{5.40}$$

(where $f \in \mathscr{Bor}_{\mathrm{b}}(X)$). We call $\pi|_K$ the ***restriction*** *of π to K.*

Proof. The uniqueness is obvious. Existence: For each $g \in \mathscr{Bor}_{\mathrm{b}}(K)$, extend g by zero to $\widetilde{g} : X \to \mathbb{C}$, and set $\pi|_K(g) = \pi(\widetilde{g})$. Then for each $f \in \mathscr{Bor}_{\mathrm{b}}(X)$, we have

$$\pi|_K(f|_K) = \pi(\widetilde{f|_K}) = \pi(f\chi_K) = \pi(f)$$

where the last step is by Rem. 5.7.11. In particular, $\pi|_K(1_K) = \pi(1|_X)$. It follows that $\pi|_K$ is a normal unitary representation satisfying (5.40). □

Theorem 5.7.13 (Borel functional calculus). *Let $T_1, \dots, T_N \in \mathfrak{L}(\mathcal{H})$ be adjointly-commuting bounded normal operators. Then there exists a unique normal unitary representation*

$$\pi_{T_\bullet} : \mathscr{Bor}_{\mathrm{b}}(\mathbb{C}^N) \to \mathfrak{L}(\mathcal{H})$$

with compact support such that $\pi_{T_\bullet}$ sends each z_i to T_i.

Here, z_i is the i-th **coordinate function** of $\mathbb{R}^N$, i.e., the projection $\mathbb{C}^N \to \mathbb{C}$ onto the i-th component. We call $(\pi_{T_\bullet}, \mathcal{H})$ the **Borel functional calculus** of $T_\bullet$, and write

$$\boldsymbol{f(T_\bullet)} := \pi_{T_\bullet}(f)$$

Proof. Uniqueness: Suppose that π_1, π_2 both satisfy the requirement for $\pi_{T_\bullet}$. Let K be a compact subset of $\mathbb{C}^N$ containing $\mathrm{Supp}(\pi_1) \cup \mathrm{Supp}(\pi_2)$. Let $\mathscr{A}$ be the polynoimal algebra $\mathbb{C}[z_\bullet, \overline{z_\bullet}]$ (cf. Exp. 1.5.6). By enlarging K, we assume that K has non-empty interior so that $\mathscr{A}$ can be viewed as a unital *-subalgebra of $C(K)$.

By Prop. 5.7.12,

$$\pi_1|_K(z_i|_K) = \pi_1(z_i) = T_i$$

and similarly $\pi_2|_K(z_i|_K) = T_i$. Thus $\pi_1|_K$ and $\pi_2|_K$ agree on $\mathscr{A}$, and hence $\pi_1|_K = \pi_2|_K$ by the uniqueness part of Thm. 5.5.13. By Prop. 5.7.12, we conclude

$\pi_1 = \pi_2$.

Existence: Let $A_i = \mathrm{Re}(T_i)$ and $B_i = \mathrm{Im}(T_i)$. Then $A_1, B_1, \dots, A_N, B_N$ are mutually commuting bounded self-adjoint operators on $\mathcal{H}$, and hence admit a Borel functional calculus by Thm. 5.7.4. We choose $R \in \mathbb{R}_{>0}$ such that $\|A_i\| \leqslant R$ and $\|B_i\| \leqslant R$ for each i. Let

$$K = ([-R, R]^2)^N$$

where, for each $1 \leqslant i \leqslant N$, the i-th component $[-R, R]^2$ is viewed as a subset of $\mathbb{C}$ via the correspondence $(x_i, y_i) \mapsto x_i + \mathbf{i}y_i$. Then K is a compact subset of $\mathbb{C}^N \simeq (\mathbb{R}^2)^N$. One checks easily that

$$\pi : \mathscr{Bor}_{\mathrm{b}}(\mathbb{C}^N) \to \mathfrak{L}(\mathcal{H}) \qquad f \mapsto (f|_K)(A_1, B_1, \dots, A_N, B_N)$$

is a normal unitary representation of $\mathscr{Bor}_{\mathrm{b}}(\mathbb{C}^N)$ with support $\mathrm{Supp}(\pi) \subset K$, and

$$\begin{aligned} &\pi(z_i) = (z_i|_K)(A_1, B_1, \dots, A_N, B_N) \\ =&(x_i|_K)(A_1, B_1, \dots, A_N, B_N) + \mathbf{i}(y_i|_K)(A_1, B_1, \dots, A_N, B_N) = A_i + \mathbf{i}B_i = T_i \end{aligned}$$

This finishes the proof. □

5.8 Joint spectrum $\mathrm{Sp}(T_\bullet)$ and the basic properties of Borel functional calculus

Fix a Hilbert space $\mathcal{H}$ and adjointly-commuting bounded normal operators $T_1, \dots, T_N \in \mathfrak{L}(\mathcal{H})$. Recall Exp. 1.5.6 for the meaning of $\mathbb{C}[z_\bullet, \overline{z_\bullet}]$. Recall that $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$.

Definition 5.8.1. The support of the Borel functional calculus $\pi_{T_\bullet} : \mathscr{Bor}_{\mathrm{b}}(\mathbb{C}^N) \to \mathfrak{L}(\mathcal{H})$ is denoted by **$\mathrm{Sp}(T_1, \dots, T_N)$** (abbreviated to **$\mathrm{Sp}(T_\bullet)$**) and called the **joint spectrum** of $T_1, \dots, T_N$. That is, $\lambda_\bullet \in \mathbb{C}^N$ belongs to $\mathrm{Sp}(T_\bullet)$ iff $\chi_U(T_\bullet) \neq 0$ for each $U \in \mathrm{Nbh}_{\mathbb{C}^N}(\lambda_\bullet)$.

Remark 5.8.2. Let $K \subset \mathbb{C}^N$ be a Borel set containing $\mathrm{Sp}(T_\bullet)$. By Rem. 5.7.11, for each $f, g \in \mathscr{Bor}_{\mathrm{b}}(\mathbb{C}^N)$, we have

$$f|_K = g|_K \qquad \Longrightarrow \qquad f(T_\bullet) = g(T_\bullet)$$

By Prop. 5.7.12, there is a unique normal unitary representation

$$\mathscr{Bor}_{\mathrm{b}}(K) \to \mathfrak{L}(\mathcal{H}) \qquad g \mapsto g(T_\bullet) \tag{5.41}$$

satisfying $(f|_K)(T_\bullet) = f(T_\bullet)$ for each $f \in \mathscr{Bor}_{\mathrm{b}}(K)$. The map (5.41) is also called a **Borel functional calculus** of $T_\bullet$.

Remark 5.8.3. If $\mathcal{H} \neq 0$, then $\mathrm{Sp}(T_\bullet)$ is non-empty; otherwise, by Rem. 5.8.2, we have $\mathbf{1} = 1(T_\bullet) = 0(T_\bullet) = 0$, which is impossible.

5.8.1 Basic properties of Borel functional calculus

Proposition 5.8.4. *For each $f \in \mathscr{Bor}_{\mathrm{b}}(\mathbb{C}^N)$, we have*

$$\|f(T_\bullet)\| \leqslant \|f\|_{l^\infty(\mathrm{Sp}(T_\bullet))}$$

Consequently, if (f_α) is a net in $\mathscr{Bor}_{\mathrm{b}}(\mathbb{C}^N)$ converging uniformly on $\mathrm{Sp}(T_\bullet)$ to $f \in \mathscr{Bor}_{\mathrm{b}}(\mathbb{C}^N)$, then $f_\alpha(T_\bullet)$ converges in norm to $f(T_\bullet)$.

Proof. This is an immediate consequence of Prop. 5.5.16. We can also prove it using the language of measure theory: Recall Rem. 5.5.9 for the meaning of μ_ξ where $\xi \in \mathcal{H}$. Then

$$\|f(T_\bullet)\xi\|^2 = \int_{\mathrm{Sp}(T_\bullet)} |f|^2 d\mu_\xi \tag{5.42a}$$

since $\|f(T_\bullet)\xi\|^2 = \langle\xi|f(T_\bullet)^*f(T_\bullet)\xi\rangle = \langle\xi|(f^*f)(T_\bullet)\xi\rangle = \int_{\mathrm{Sp}(T_\bullet)} |f|^2 d\mu_\xi$. In particular, since $1(T_\bullet) = \mathrm{id}_{\mathcal{H}}$, we obtain

$$\|\xi\|^2 = \mu_\xi(\mathrm{Sp}(T_\bullet)) \tag{5.42b}$$

Thus $\int_{\mathrm{Sp}(T_\bullet)} |f|^2 d\mu_\xi \leqslant \|f\|_{l^\infty(\mathrm{Sp}(T_\bullet))} \cdot \|\xi\|^2$. □

Proposition 5.8.5. *Let (f_α) be a net in $\mathscr{Bor}_{\mathrm{b}}(\mathbb{C}^N)$. Let $f \in \mathscr{Bor}_{\mathrm{b}}(\mathbb{C}^N)$. Suppose that (f_α) converges universally-L^2 to f when restricted to $\mathrm{Sp}(T_\bullet)$, that is,*

$$\lim_\alpha \int_{\mathrm{Sp}(T_\bullet)} |f - f_\alpha|^2 d\mu = 0$$

for each Radon measure (equivalently, finite Borel measure) μ on $\mathrm{Supp}(T_\bullet)$. Then $f_\alpha(T_\bullet)$ converges in SOT to $f(T_\bullet)$.

Proof. Apply Prop. 5.5.15 to $\pi_{T_\bullet}|_{\mathrm{Sp}(T_\bullet)}$. Alternatively, apply (5.42). □

Theorem 5.8.6. *Let $g_1, \dots, g_L \in \mathscr{Bor}_{\mathrm{b}}(\mathbb{C}^N)$ and $f \in \mathscr{Bor}_{\mathrm{b}}(\mathbb{C}^L)$. Then $g_1(T_\bullet), \dots, g_N(T_\bullet)$ are adjointly commuting normal operators, and*

$$f(g_1(T_\bullet), \dots, g_L(T_\bullet)) = (f \circ (g_1, \dots, g_L))(T_\bullet) \tag{5.43}$$

Proof. We have $[g_j(T_\bullet), g_k(T_\bullet)] = 0$ since

$$g_j(T_\bullet)g_k(T_\bullet) = (g_jg_k)(T_\bullet) = (g_kg_j)(T_\bullet) = g_k(T_\bullet)g_j(T_\bullet)$$

Since $g_k^*(T_\bullet) = g_k(T_\bullet)^*$, we have $[g_j(T_\bullet), g_k(T_\bullet)^*] = 0$. Thus $g_j(T_\bullet)$ commutes adjointly with $g_k(T_\bullet)$. Setting $k = j$, we see that $g_j(T_\bullet)$ is normal.

Since the map

$$\mathscr{Bor}_{\mathrm{b}}(\mathbb{C}^L) \longrightarrow \mathfrak{L}(\mathcal{H}) \qquad f \mapsto (f \circ (g_1, \dots, g_L))(T_\bullet)$$

is clearly a normal unitary representation of $\mathscr{Bor}_{\mathrm{b}}(\mathbb{C}^L)$, Eq. (5.43) holds. □

Corollary 5.8.7. *Let $1 \leqslant L \leqslant N$, and assume that $f \in \mathscr{Bor}_{\mathrm{b}}(\mathbb{C}^N)$ depends only on the first L variables so that f can also be viewed as a Borel function on $\mathbb{C}^L$. Then*

$$f(T_1, \ldots, T_L) = f(T_1, \ldots, T_N)$$

In particular, if f only depends on the first variable, then $f(T_1) = f(T_1, \ldots, T_N)$.

Proof. Apply Thm. 5.8.6 to the case that $f \in \mathscr{Bor}_{\mathrm{b}}(\mathbb{C}^L)$ and $g_1, \ldots, g_L$ are the first L coordinate functions of $\mathbb{C}^N$ (i.e., g_i sends $(p_1, \ldots, p_N) \in \mathbb{C}^N$ to p_i). □

We close this subsection with a functional-calculus description of eigenspaces.

Proposition 5.8.8. *Let $T \in \mathfrak{L}(\mathcal{H})$ be normal, and let $\lambda \in \mathbb{C}$. Then $\chi_{\{\lambda\}}(T)$ is the projection onto* $\mathrm{Ker}(T - \lambda)$.

Proof. Since $\chi_{\{\lambda\}} = \chi_{\{\lambda\}}^2 = \overline{\chi_{\{\lambda\}}}$, the operator $\chi_{\{\lambda\}}(T)$ is a projection. It remains to prove that $\mathrm{Ker}(T - \lambda)$ equals the range of $\chi_{\{\lambda\}}(T)$.

For each $\xi \in \mathcal{H}$, we compute that

$$\|(T - \lambda)\chi_{\{\lambda\}}(T)\xi\|^2 = \int_{\mathbb{C}} |z - \lambda|^2 \chi_{\{\lambda\}}^2 d\mu_\xi = 0$$

and hence $(T - \lambda)\chi_{\{\lambda\}}(T)\xi = 0$. This proves $\mathrm{Rng}(\chi_{\{\lambda\}}(T)) \subset \mathrm{Ker}(T - \lambda)$.

For each $\psi \in \mathrm{Ker}(T - \lambda)$, by (5.42),

$$0 = \|(T - \lambda)\psi\|^2 = \int_{\mathbb{C}} |z - \lambda|^2 d\mu_\psi$$

and hence $\mathbb{C}\backslash\lambda$ is μ_ψ-null by Prop. 1.6.1. It follows that

$$\|\chi_{\{\lambda\}}(T)\psi - \psi\|^2 = \int_{\mathbb{C}} |\chi_{\{\lambda\}} - 1|^2 d\mu_\psi = \int_{\{\lambda\}} |\chi_{\{\lambda\}} - 1|^2 d\mu_\psi = 0$$

and hence $\chi_{\{\lambda\}}(T)\psi = \psi$. This proves that $\mathrm{Ker}(T - \lambda) \subset \mathrm{Rng}(\chi_{\{\lambda\}}(T))$. □

5.8.2 Basic properties of joint spectra

Determining joint spectra is often highly useful. For instance, knowing that the spectrum of a positive operator is contained in $\mathbb{R}_{\geqslant 0}$ allows us to use the properties of integrals of positive functions. In this subsection, we present some basic methods for characterizing joint spectra and illustrate them with examples.

Proposition 5.8.9. *Let $1 \leqslant L < N$. Then*

$$\mathrm{Sp}(T_1, \ldots, T_N) \subset \mathrm{Sp}(T_1, \ldots, T_L) \times \mathrm{Sp}(T_{L+1}, \ldots, T_N) \tag{5.44}$$

By applying Prop. 5.8.9 repeatedly, we obtain

$$\mathrm{Sp}(T_1,\dots,T_N)\subset \mathrm{Sp}(T_1)\times\cdots\times\mathrm{Sp}(T_N)$$

Proof. Choose any $p\in\mathbb{C}^L, q\in\mathbb{C}^{N-L}$, and assume that (p,q) is outside $\mathrm{Sp}(T_1,\dots,T_L)\times\mathrm{Sp}(T_{L+1},\dots,T_N)$. Then either p is outside $\mathrm{Sp}(T_1,\dots,T_L)$ or q is outside $\mathrm{Sp}(T_{L+1},\dots,T_N)$. In the previous case, there exists $U\in\mathrm{Nbh}_{\mathbb{C}^L}(p)$ such that $\chi_U(T_1,\dots,T_L)=0$. By Cor. 5.8.7, we have

$$\chi_{U\times\mathbb{C}^{N-L}}(T_1,\dots,T_N)=\chi_U(T_1,\dots,T_L)=0$$

and hence $(p,q)\notin\mathrm{Sp}(T_1,\dots,T_N)$. The latter case can be addressed in the same way. $\square$

To some extend, Prop. 5.8.9 reduces the study of joint spectra to that of individual operators. The next proposition gives a particularly useful characterization of the spectrum of a single bounded normal operator.

Definition 5.8.10. Let $T\in\mathfrak{L}(\mathcal{V})$ where $\mathcal{V}$ is a normed vector space. The number $\lambda\in\mathbb{C}$ is called an **approximate eigenvalue** of T if one of the following (clearly) equivalent conditions holds:

(1) For each $\varepsilon>0$ there exists a non-zero vector $\xi\in\mathcal{V}$ such that

$$\|(T-\lambda)\xi\|\leqslant\varepsilon\|\xi\|$$

(2) There exists a sequence $(\xi_n)_{n\in\mathbb{Z}_+}$ of unit vectors in $\mathcal{V}$ such that

$$\lim_n(T-\lambda)\xi_n=0$$

Theorem 5.8.11. *Let $T\in\mathfrak{L}(\mathcal{H})$ be normal. Let $\lambda\in\mathbb{C}$. Then the following are equivalent:*

(1) $\lambda\in\mathrm{Sp}(T)$.

(2) λ is an approximate eigenvalue of T.

Proof. (1)⇒(2): Assume (1). Then for each $\varepsilon>0$, we have $\chi_{U_\varepsilon}(T)\neq 0$, where $U_\varepsilon=B_{\mathbb{C}}(\lambda,\varepsilon)$. Choose any non-zero $\xi\in\mathrm{Rng}(\chi_{U_\varepsilon}(T))$, noting that $\xi=\chi_{U_\varepsilon}(T)\xi$ because $\chi_{U_\varepsilon}^2=\chi_{U_\varepsilon}$. Let $f_\varepsilon\in\mathscr{B}or_{\mathrm{b}}(\mathbb{C})$ be defined by $f_\varepsilon(z)=(z-\lambda)\chi_{U_\varepsilon}(z)$. Then $\|f_\varepsilon\|_{l^\infty}\leqslant\varepsilon$, and hence

$$\|(T-\lambda)\xi\|=\|(T-\lambda)\chi_{U_\varepsilon}(T)\xi\|=\|f_\varepsilon(T)\xi\|\leqslant\|f_\varepsilon\|_{l^\infty}\cdot\|\xi\|=\varepsilon\|\xi\|$$

due to Prop. 5.8.4. This proves (2).

¬(1)⇒¬(2): Assume that $\lambda\notin\mathrm{Sp}(T)$. Since $\mathrm{Sp}(T)$ is compact, there exists $\varepsilon>0$ such that $\mathrm{Sp}(T)\subset\mathbb{C}\backslash U_\varepsilon$ where $U_\varepsilon=B_{\mathbb{C}}(\lambda,\varepsilon)$. For each $\xi\in\mathcal{H}$, by (5.42) we have

$$\|(T-\lambda)\xi\|^2=\int_{\mathrm{Sp}(T)}|z-\lambda|^2d\mu_\xi=\int_{\mathbb{C}\backslash U_\varepsilon}|z-\lambda|^2d\mu_\xi\geqslant\varepsilon^2\int_{\mathbb{C}\backslash U_\varepsilon}d\mu_\xi=\varepsilon^2\|\xi\|^2$$

Therefore, (2) is not true. $\square$

Exercise 5.8.12. Let $\lambda_\bullet = (\lambda_1, \ldots, \lambda_N) \in \mathbb{C}^N$. Prove that $\lambda_\bullet \in \mathrm{Sp}(T_\bullet)$ iff $\lambda_\bullet$ is a **joint approximate eigenvalue** of $T_\bullet$, that is, there exists a sequence of unit vectors (ξ_n) in $\mathcal{H}$ such that

$$\lim_n (T_i - \lambda_i)\xi_n = 0 \qquad \text{for all } 1 \leqslant i \leqslant N$$

Use this result to give an alternative proof of Prop. 5.8.9.

Example 5.8.13. Let $\lambda \in \mathbb{C}$. One easily sees that λ is the only approximate eigenvalue of $\lambda \cdot 1$. Therefore $\mathrm{Sp}(\lambda \cdot \mathbf{1}) = \lambda$.

Proposition 5.8.14. *Let $T \in \mathfrak{L}(\mathcal{H})$ be normal. Then $T \geqslant 0$ iff* $\mathrm{Sp}(T) \subset \mathbb{R}_{\geqslant 0}$.

Proof. Suppose that $T \geqslant 0$. Choose any $\lambda \in \mathbb{C}\backslash\mathbb{R}_{\geqslant 0}$. Then for any sequence (ξ_n) of unit vectors we have

$$\langle \xi_n | (T - \lambda)\xi_n \rangle = \langle \xi_n | T\xi_n \rangle - \lambda = a_n - \lambda$$

where $a_n \in \mathbb{R}_{\geqslant 0}$, and hence $a_n - \lambda \nrightarrow 0$. Thus λ is not an approximate eigenvalue of T. In view of Thm. 5.8.11, we conclude that $\mathrm{Sp}(T) \subset \mathbb{R}_{\geqslant 0}$.

Conversely, assume that $\mathrm{Sp}(T) \subset \mathbb{R}_{\geqslant 0}$. Then

$$\langle \xi | T\xi \rangle = \int_{\mathrm{Sp}(T)} zd \cdot \mu_\xi(z) \geqslant 0$$

for all $\xi \in \mathcal{H}$, where μ_ξ is defined in Rem. 5.5.9. Thus $T \geqslant 0$. □

Proposition 5.8.15. *Let $T \in \mathfrak{L}(\mathcal{H})$ be normal. Then T is self-adjoint iff* $\mathrm{Sp}(T) \subset \mathbb{R}$.

Proof. This can be proved in a same way as in Thm. 5.8.14. Here, we provide a different argument.

Note that $T = T^*$ iff $(z - \overline{z})(T) = 0$, iff $(z - \overline{z})(T)\xi = 0$ for all ξ, iff

$$\int_{\mathbb{C}} |z - \overline{z}|^2 d\mu_\xi \overset{(5.42)}{=\!=} \|(z - \overline{z})(T)\xi\|^2$$

equals zero for all ξ, iff (by Prop. 1.6.1) the set $\mathbb{C}\backslash\mathbb{R} = \{z - \overline{z} \neq 0\}$ is μ_ξ-null for each ξ. This is equivalent to $\mathrm{Supp}(\mu_\xi) \subset \mathbb{R}$ for all ξ, and hence equivalent to $\mathrm{Sp}(T) \subset \mathbb{R}$ due to (5.39). □

Recall from Exp. 5.7.2 that unitary operators on $\mathcal{H}$ are normal.

Proposition 5.8.16. *Let $U \in \mathfrak{L}(\mathcal{H})$ be normal. Then U is unitary iff* $\mathrm{Sp}(U) \subset \mathbb{S}^1$.

Proof. Suppose that U is unitary. Choose any $\lambda \in \mathbb{C}$ such that $|\lambda| \neq 1$. Then, for each sequence of unit vectors (ξ_n) in $\mathcal{H}$ we have

$$\|(U-\lambda)\xi_n\| \geqslant \|U\xi_n\| - \|\lambda\xi_n\| = 1 - |\lambda|$$

which does not converge to 0. Therefore, λ is not an approximate eigenvalue of U. By Thm. 5.8.11, we conclude $\mathrm{Sp}(U) \subset \mathbb{S}^1$.

Conversely, assume that $\mathrm{Sp}(U) \subset \mathbb{S}^1$. Then $z^*z|_{\mathrm{Sp}(\mathbb{S}^1)} = 1$, and hence $U^*U = z(U)^*z(U) = z^*z(U) = 1(U) = \mathbf{1}$ by Rem. 5.8.2. Since $[U, U^*] = 0$, we also have $UU^* = \mathbf{1}$. It follows from Exp. 5.7.2 that U is unitary. $\square$

Exercise 5.8.17. Let $T \in \mathfrak{L}(\mathcal{H})$ be normal. Show that $\mathrm{Sp}(T) \subset \overline{B}_{\mathbb{C}}(0, \|T\|)$. Show that if $T^* = T$ then $\mathrm{Sp}(T) \subset [-\|T\|, \|T\|]$.

Exercise 5.8.18. For each $T \in \mathfrak{L}(\mathcal{H})$, define

$$\boldsymbol{\sigma(T)} := \{\lambda \in \mathbb{C} : \lambda - T \text{ has no inverse in } \mathfrak{L}(\mathcal{H})\}$$

Prove that if $T \in \mathfrak{L}(\mathcal{H})$ is normal then $\mathrm{Sp}(T) = \sigma(T)$.

5.9 Application: von Neumann's mean ergodic theorem

In this section, we present an example that illustrates how the Borel functional calculus can be applied in practice. The results in this section will not be used elsewhere in this course.

Theorem 5.9.1 (Von Neumann's mean ergodic theorem). *Let $U \in \mathfrak{L}(\mathcal{H})$ be an isometry, equivalently, $U^*U = \mathbf{1}$. Let P be the projection onto $\mathrm{Ker}(\mathbf{1} - U)$. Then*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} U^k = P \qquad \text{in SOT} \tag{5.45}$$

A particularly important case occurs when $\mathrm{Ker}(\mathbf{1} - U)$ is one-dimensional, i.e., spanned by a unit vector $e \in \mathrm{Ker}(\mathbf{1} - U)$. In that case, (5.45) reads

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} U^k\xi = e \cdot \langle e|\xi\rangle \qquad \text{for each } \xi \in \mathcal{H} \tag{5.46}$$

Proof. Step 1. Let X be a set whose cardinality, added by the cardinality of an orthonormal basis of $\mathrm{Rng}(U)^\perp$, is equal to the cardinality of X. There there exists a unitary map $V : l^2(X) \to \mathrm{Rng}(U)^\perp \oplus l^2(X)$, which can thus be extended to a unitary map

$$W : \mathcal{H} \oplus l^2(X) \to \mathcal{H} \oplus l^2(X)$$

whose restriction to $\mathcal{H}$ equals U.

By Prop. 5.8.16, we have $\mathrm{Sp}(W) \subset \mathbb{S}^1$. On $\mathbb{S}^1$, the sequence of functions

$$f_n := \frac{1}{n}\sum_{k=0}^{n-1} z^k$$

which equals $(1-z^n)/n(1-z)$ outside $\{1\}$, is uniformly bounded and converges pointwise to $\chi_{\{1\}}$. By DCT, for any finite Borel measure μ on $\mathbb{S}^1$ we have

$$\lim_n \int_{\mathbb{S}^1} |f_n - \chi_{\{1\}}|^2 d\mu = 0$$

Therefore, by Thm. 5.8.5,

$$\lim_n \frac{1}{n}\sum_{k=0}^{n-1} W^k = Q \qquad \text{in SOT}$$

where $Q := \chi_{\{1\}}(W)$ is the projection of $\mathcal{H}\oplus l^2(X)$ onto $\mathrm{Ker}(\mathbf{1}-W)$ (cf. Prop. 5.8.8).

Step 2. According to Step 1, for each $\xi \in \mathcal{H}$, the limit $\lim_n n^{-1}(1+\cdots+U^{n-1})\xi$ converges in $\mathcal{H}$. In the case $\xi \in \mathrm{Ker}(\mathbf{1}-U)$, then $\xi \in \mathrm{Ker}(\mathbf{1}-W)$, and hence this limit converges to ξ. In the case $\xi \perp \mathrm{Ker}(\mathbf{1}-U)$, if we can prove that $\xi \perp \mathrm{Ker}(\mathbf{1}-W)$, then this limit converges to $Q\xi = 0$. Combining the two cases together, we obtain (5.45).

Let us fix $\xi \in \mathcal{H}$ orthogonal to $\mathrm{Ker}(\mathbf{1}-U)$, and prove that $\xi \perp \mathrm{Ker}(\mathbf{1}-W)$. Choose any $\psi + \eta \in \mathcal{H}\oplus l^2(X)$ (where $\psi \in \mathcal{H}$ and $\eta \in l^2(X)$), and assume that $W(\psi+\eta) = \psi+\eta$. Using the notation in Step 1, we have $\psi+\eta = U\psi + V\eta$. Let E be the projection of $\mathcal{H}\oplus l^2(X)$ onto $\mathcal{H}$. Then

$$\psi = E(\psi+\eta) = EU\psi + EV\eta = U\psi + EV\eta$$

Since V has range in $\mathrm{Rng}(U)^\perp \oplus l^2(X)$, we have $EV\eta \in \mathrm{Rng}(U)^\perp$, and hence $U\psi \perp EV\eta$. It follows from the Pythagorean identity that

$$\|\psi\|^2 = \|U\psi\|^2 + \|EV\eta\|^2$$

Since U is an isometry, we must have $\|EV\eta\|^2 = 0$. Thus $\psi = U\psi$, and hence $\psi \in \mathrm{Ker}(\mathbf{1}-U)$. We conclude that

$$\langle \xi|\psi+\eta\rangle = \langle \xi|\psi\rangle \subset \langle \xi|\mathrm{Ker}(\mathbf{1}-U)\rangle = \{0\}$$

This finishes the proof that $\xi \perp \mathrm{Ker}(\mathbf{1}-W)$. □

Example 5.9.2. Let $(X, \mathfrak{M}, \mu)$ be a **probability space**, i.e., a measure space satisfying $\mu(X) = 1$. Let $\phi : X \to X$ be a **measure preserving transform**. This means that ϕ is a measurable map satisfying $\phi_*\mu = \mu$, that is, $\mu(\phi^{-1}(A)) = \mu(A)$ for each measurable $A \subset X$ (cf. Def. 1.6.6). Then

$$U : L^2(X, \mu) \to L^2(X, \mu) \qquad f \mapsto f \circ \phi \tag{5.47}$$

is an isometry, since for all $f \in L^2(X, \mu)$ we have

$$\langle Uf|Uf\rangle = \int_X |f \circ \phi|^2 d\mu = \int_X |f|^2 d\phi_*\mu = \int_X |f|^2 d\mu = \langle f|f\rangle$$

Now, assume that the measure preserving $\phi : X \to X$ is **ergodic** , which means that the only measurable set $A \subset X$ satisfying [9]

$$\mu(\phi^{-1}(A)\triangle A) = 0$$

must satisfy either $\mu(A) = 0$ or $\mu(X\backslash A) = 0$. This implies that $\mathrm{Ker}(\mathbf{1} - U) = \mathrm{Span}\{1\}$, cf. Exe. 5.9.3. Thus, for each $f \in L^2(X, \mu)$, since $\langle 1|f\rangle = \int_X f d\mu$, it follows from the mean ergodic Thm. 5.9.1 that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f \circ \underbrace{\phi \circ \cdots \circ \phi}_{k \text{ times}} = \int_X f d\mu$$

where the LHS converges to the RHS in the L^2-norm. □

Exercise 5.9.3. Let $(X, \mathfrak{M}, \mu)$ be a probability space. Let $\phi : X \to X$ be a measure preserving transform. Prove that ϕ is ergodic iff the map U defined in (5.47) satisfies $\mathrm{Ker}(\mathbf{1} - U) = \mathrm{Span}\{1\}$.

Hint. "$\Leftarrow$": If $\mu(\phi^{-1}(A)\triangle A) = 0$ then $\chi_A \in \mathrm{Ker}(\mathbf{1} - U)$.

"$\Rightarrow$": Show that if $f, g : X \to \mathbb{C}$ are measurable functions satisfying $f = g$ a.e. in μ, then $h \circ f = h \circ g$ a.e. in μ for each function h on $\mathbb{C}$. Let $f \in \mathrm{Ker}(\mathbf{1} - U)$, that is, $f = f \circ \phi$ a.e.. Use the fact that $\chi_\Omega \circ f = \chi_\Omega \circ f \circ \phi$ a.e. for any open $\Omega \subset \mathbb{C}$ to conclude that the essential range of f (cf. Def. 1.6.6) is a single-point set. □

5.10 The multiplication-operator version of the spectral theorem

In this section, we introduce the final version of the Spectral Theorem, the multiplication operator version; see Thm. 5.10.22. For students already familiar with measure theory, this version is not only the most accessible but also the one most easily extended to the spectral theory of unbounded operators.

Historically, the multiplication-operator version appeared later for several reasons: First, measure theory was not yet fully developed when the earlier versions

[9] For any two sets A, B, the symmetric difference $A\triangle B$ is $(A\backslash B) \cup (B\backslash A)$.

of the theorem were established. Second, the spectral decomposition in this setting is not unique. Third, several of the ideas underlying this version grew out of representation theory and operator algebras. Finally, unlike the earlier formulations, this version adopts most completely the perspective of linear operators rather than that of bilinear or sesquilinear forms (cf. the paradigm shift (5.19b)). One drawback of this approach is that, because of this last reason, the proof by itself does not make clear how the Riesz representation theorem interacts with the rest of the argument, since the bilinear/sesquilinear form framework is more naturally suited to the Riesz representation theorem; cf. Subsec. 2.5.2.

Informally, the multiplication-operator version of the Spectral Theorem asserts that every bounded normal operator is unitarily equivalent to a multiplication operator on a direct sum of L^2-spaces. A detailed statement and explanation will be given in the following subsections.

5.10.1 Infinite direct sums of Hilbert spaces

Definition 5.10.1. Let $(V_\alpha)_{\alpha\in\mathscr{I}}$ be a family of $\mathbb{F}$-vector space. Then the $\mathbb{F}$-vector space structure on **direct product $\prod_{\alpha\in\mathscr{I}} V_\alpha$** is defined componentwisely. Namely, for each $(\xi_\alpha)_{\alpha\in\mathscr{I}}, (\eta_\alpha)_{\alpha\in\mathscr{I}}$ in $\prod_{\alpha\in\mathscr{I}}\mathcal{H}_\alpha$ and $\lambda\in\mathbb{C}$ we have

$$(\xi_\alpha)_{\alpha\in\mathscr{I}}+(\eta_\alpha)_{\alpha\in\mathscr{I}}=(\xi_\alpha+\eta_\alpha)_{\alpha\in\mathscr{I}}\qquad \lambda(\xi_\alpha)_{\alpha\in\mathscr{I}}=(\lambda\xi_\alpha)_{\alpha\in\mathscr{I}}$$

Definition 5.10.2. Let $(\mathcal{H}_\alpha)_{\alpha\in\mathscr{I}}$ be a family of Hilbert spaces. The **(orthogonal) Hilbert space direct sum** of $(\mathcal{H}_\alpha)_{\alpha\in\mathscr{I}}$ is defined by

$$\bigoplus_{\alpha\in\mathscr{I}}\mathcal{H}_\alpha=\Big\{(\xi_\alpha)_{\alpha\in\mathscr{I}}\in\prod_{\alpha\in\mathscr{I}}\mathcal{H}_\alpha:\sum_{\alpha\in\mathscr{I}}\Vert\xi_\alpha\Vert^2<+\infty\Big\}$$

We write any $\xi=(\xi_\alpha)_{\alpha\in\mathscr{I}}$ in $\bigoplus_{\alpha\in\mathscr{I}}\mathcal{H}_\alpha$ as $\oplus_{\alpha\in\mathscr{I}}\xi_\alpha$. Then $\bigoplus_{\alpha\in\mathscr{I}}\mathcal{H}_\alpha$ is equipped with the inner product

$$\big\langle\oplus_{\alpha\in\mathscr{I}}\eta_\alpha\big|\oplus_{\alpha\in\mathscr{I}}\xi_\alpha\big\rangle=\sum_{\alpha\in\mathscr{I}}\langle\eta_\alpha|\xi_\alpha\rangle \tag{5.48}$$

and is a Hilbert space. We view each $\mathcal{H}_\alpha$ as an inner product subspace of $\bigoplus_{\alpha\in\mathscr{I}}\mathcal{H}_\alpha$ by identifying each $\xi_\alpha\in\mathcal{H}_\alpha$ with the element of $\bigoplus_{\alpha\in\mathscr{I}}\mathcal{H}_\alpha$ whose α-component is ξ_α and whose other components are zero.

Explanation. For each finite set $I\subset\mathscr{I}$,

$$\begin{aligned}&\sum_{\alpha\in I}|\langle\eta_\alpha|\xi_\alpha\rangle|\leqslant\sum_{\alpha\in I}\Vert\eta_\alpha\Vert\cdot\Vert\xi_\alpha\Vert\leqslant\Big(\sum_{\alpha\in I}\Vert\eta_\alpha\Vert^2\Big)^{\frac12}\cdot\Big(\sum_{\alpha\in I}\Vert\xi_\alpha\Vert^2\Big)^{\frac12}\\&\leqslant\Big(\sum_{\alpha\in\mathscr{I}}\Vert\eta_\alpha\Vert^2\Big)^{\frac12}\cdot\Big(\sum_{\alpha\in\mathscr{I}}\Vert\xi_\alpha\Vert^2\Big)^{\frac12}=:C\end{aligned}$$

Applying $\sup_{I\in\mathrm{fin}(2^{\mathscr{I}})}$, we obtain $\sum_{\alpha\in\mathscr{I}}|\langle\eta_\alpha|\xi_\alpha\rangle|\leqslant C<+\infty$. Therefore, the RHS of (5.48) converges absolutely (and hence converges, cf. Prop. 1.2.45).

To prove that $\mathcal{H}:=\bigoplus_{\alpha\in\mathscr{I}}\mathcal{H}_\alpha$ is a Hilbert space, one can either show that $\mathcal{H}$ is Cauchy-complete, or show that any bounded linear functional $\Lambda:\mathcal{H}\to\mathbb{C}$ is realized by a pairing with some element of $\mathcal{H}$, cf. Thm. 3.5.3. We follow the second approach. The restriction of Λ to each $\mathcal{H}_\alpha$ is a bounded linear functional. Hence, by Thm. 3.5.3, there exists $\eta_\alpha\in\mathcal{H}_\alpha$ such that

$$\Lambda(\xi_\alpha)=\langle\eta_\alpha|\xi_\alpha\rangle\qquad\text{for any }\xi_\alpha\in\mathcal{H}_\alpha$$

Moreover, for each $I\in\mathrm{fin}(2^{\mathscr{I}})$, noting that $\sum_{\alpha\in I}\eta_\alpha$ equals $\oplus_{\alpha\in I}\eta_\alpha$, we have

$$\sum_{\alpha\in I}\|\eta_\alpha\|^2=\Lambda\big(\oplus_{\alpha\in I}\eta_\alpha\big)\leqslant\|\Lambda\|\cdot\big\|\oplus_{\alpha\in I}\eta_\alpha\big\|=\|\Lambda\|\cdot\Big(\sum_{\alpha\in I}\|\eta_\alpha\|^2\Big)^{\frac{1}{2}}$$

and hence $\sum_{\alpha\in I}\|\eta_\alpha\|^2\leqslant\|\Lambda\|^2$. Applying $\sup_{I\in\mathrm{fin}(2^{\mathscr{I}})}$, we obtain

$$\sum_{\alpha\in\mathscr{I}}\|\eta_\alpha\|^2\leqslant\|\Lambda\|^2<+\infty$$

This proves that $\eta:=\oplus_{\alpha\in\mathscr{I}}\eta_\alpha$ belongs to $\mathcal{H}$. Clearly Λ equals the pairing with η when restricted to the subspace spanned by all $\mathcal{H}_\alpha$ where $\alpha\in\mathscr{I}$. Since this subspace is dense, we conclude from Cor. 2.4.3 that Λ equals the pairing with η on the whole space $\mathcal{H}$. □

Remark 5.10.3. Let $(\xi_\alpha)_{\alpha\in\mathscr{I}}$ be an element of $\bigoplus_{\alpha\in\mathscr{I}}\mathcal{H}_\alpha$. Since $\sum_{\alpha\in\mathscr{I}}\|\xi_\alpha\|^2<+\infty$, by Prop. 1.2.44, we have $\xi_\alpha=0$ for all but countably many α.

Definition 5.10.4. Let $(\mathcal{H}_\alpha)_{\alpha\in\mathscr{I}}$ and $(\mathcal{K}_\alpha)_{\alpha\in\mathscr{I}}$ be families of Hilbert spaces. Let $(T_\alpha)_{\alpha\in\mathscr{I}}$ be a family where $T_\alpha\in\mathfrak{L}(\mathcal{H}_\alpha,\mathcal{K}_\alpha)$ for each $\alpha\in\mathscr{I}$. Assume that

$$\sup_{\alpha\in\mathscr{I}}\|T_\alpha\|<+\infty$$

Then

$$\boldsymbol{\oplus_{\alpha\in\mathscr{I}}T_\alpha}:\bigoplus_{\alpha\in\mathscr{I}}\mathcal{H}_\alpha\to\bigoplus_{\alpha\in\mathscr{I}}\mathcal{K}_\alpha\qquad\oplus_{\alpha\in\mathscr{I}}\xi_\alpha\mapsto\oplus_{\alpha\in\mathscr{I}}T_\alpha\xi_\alpha$$

is clearly a bounded linear map whose adjoint is $\oplus_{\alpha\in\mathscr{I}}T_\alpha^*$.

5.10.2 Orthogonal decompositions of Hilbert spaces and bounded linear operators

Let $\mathcal{H}$ be a Hilbert space. Recall from Exp. 3.4.4 that a Hilbert subspace of $\mathcal{H}$ refers to a closed (equivalently, complete) linear subspace of $\mathcal{H}$.

The following example shows that the direct sum of an orthogonal family of Hilbert subspaces of $\mathcal{H}$ can be viewed as a Hilbert subspace of $\mathcal{H}$.

Example 5.10.5. Let $(\mathcal{H}_\alpha)_{\alpha\in\mathscr{I}}$ be a family of *mutually-orthogonal* Hilbert subspaces of $\mathcal{H}$. Then we have a linear isometry

$$U:\bigoplus_{\alpha\in\mathscr{I}}\mathcal{H}_\alpha\to\mathcal{H}\qquad \oplus_{\alpha\in\mathcal{H}}\xi_\alpha\mapsto\sum_{\alpha\in\mathscr{I}}\xi_\alpha \tag{5.49}$$

where $\sum_{\alpha\in\mathscr{I}}\xi_\alpha$ is an unordered sum of mutually orthogonal vectors, and hence converges in $\mathcal{H}$ due to Thm. 3.4.1-(b). Since U is complete, $\mathrm{Rng}(U)$ is also complete. Since $\sum_{\alpha\in\mathscr{I}}\mathcal{H}_\alpha$ is dense in $\mathrm{Rng}(U)$, we conclude

$$\mathrm{Rng}(U)=\mathrm{Cl}_{\mathcal{H}}\Big(\sum_{\alpha\in\mathscr{I}}\mathcal{H}_\alpha\Big) \tag{5.50}$$

We can therefore view $\bigoplus_{\alpha\in\mathscr{I}}\mathcal{H}_\alpha$ as a complete (equivalently, closed) linear subspace of $\mathcal{H}$ by identifying the elements on the two sides via U. In that case, we have

$$\bigoplus_{\alpha\in\mathscr{I}}\mathcal{H}_\alpha=\mathrm{Cl}_{\mathcal{H}}\Big(\sum_{\alpha\in\mathscr{I}}\mathcal{H}_\alpha\Big) \tag{5.51}$$

due to (5.50). □

Remark 5.10.6. In Exp. 5.10.5, let $P_\alpha\in\mathfrak{L}(\mathcal{H})$ be the projection of $\mathcal{H}$ onto $\mathcal{H}_\alpha$. Note that by Cor. 5.2.5, for each $I\in\mathrm{fin}(2^{\mathscr{I}})$, the sum $\sum_{\alpha\in I}P_\alpha$ is the projection onto $\sum_{\alpha\in I}\mathcal{H}_\alpha$. Therefore,

$$\Big(\sum_{\alpha\in I}P_\alpha\Big)_{I\in\mathrm{fin}(2^{\mathscr{I}})}$$

is an increasing net of projections. It follows from Thm. 5.2.7 that the **sum of projections**

$$\sum_{\boldsymbol{\alpha\in\mathscr{I}}}\boldsymbol{P_\alpha}:=\lim_{I\in\mathrm{fin}(2^{\mathscr{I}})}\sum_{\alpha\in I}P_\alpha$$

converges in SOT to the projection onto the range of (5.49). In particular,

$$\sum_{\alpha\in\mathscr{I}}P_\alpha=\mathrm{id}_{\mathcal{H}}\qquad\Longleftrightarrow\qquad\bigoplus_{\alpha\in\mathscr{I}}\mathcal{H}_\alpha=\mathcal{H}$$

Definition 5.10.7. Let $\mathfrak{S}\subset\mathfrak{L}(\mathcal{H})$. Let V be a linear subspace of $\mathcal{H}$. We say that V is **$\mathfrak{S}$-invariant** if $TV\subset V$ for each $T\in\mathfrak{S}$. In that case, T restricts to

$$T|_V\in\mathfrak{L}(V)$$

Similarly, if $(\pi,\mathcal{H})$ is a unitary representation of a unital *-algebra $\mathscr{A}$, we say that a linear subspace V is **$\mathscr{A}$-invariant** if V is $\pi(\mathscr{A})$-invariant, i.e., $\pi(x)V\subset V$ for each $x\in\mathscr{A}$.

Example 5.10.8. Let $V \subset \mathcal{H}$ be a linear subspace. Suppose that $\mathfrak{S} \subset \mathfrak{L}(\mathcal{H})$ and V is $\mathfrak{S}$-invariant, then $V^\perp$ is invariant under $\mathfrak{S}^* = \{T^* : T \in \mathfrak{S}\}$. Suppose that $(\pi, \mathcal{H})$ is a unitary representation of a *-algebra $\mathscr{A}$ and V is $\mathscr{A}$-invariant, then $V^\perp$ is $\mathscr{A}$-invariant.

Proof. To prove that $V^\perp$ is $\mathfrak{S}^*$-invariant, we compute that

$$\langle V|T^*V^\perp\rangle = \langle TV|V^\perp\rangle \subset \langle V|V^\perp\rangle = \{0\}$$

for each $T \in \mathfrak{S}$, and hence $\mathfrak{S}^*V^\perp \subset V^\perp$. Thus, $V^\perp$ is invariant under $\pi(\mathscr{A})^* = \pi(\mathscr{A}^*) = \pi(\mathscr{A})$. □

Example 5.10.9. Let $(\mathcal{H}_\alpha)_{\alpha\in\mathscr{I}}$ be a family of mutually-orthogonal Hilbert subspaces of $\mathcal{H}$. Let $T \in \mathfrak{L}(\mathcal{H})$. Assume that each $\mathcal{H}_\alpha$ is T-invariant. If we view $\bigoplus_{\alpha\in\mathscr{I}} \mathcal{H}_\alpha$ as a Hilbert subspace of $\mathcal{H}$, then $\bigoplus_{\alpha\in\mathscr{I}} \mathcal{H}_\alpha$ is T-invariant, and

$$T\big|_{\bigoplus_{\alpha\in\mathscr{I}} \mathcal{H}_\alpha} = \oplus_{\alpha\in\mathscr{I}}(T|_{\mathcal{H}_\alpha}) \tag{5.52}$$

as bounded linear operators on $\bigoplus_{\alpha\in\mathscr{I}} \mathcal{H}_\alpha$.

Proof. By Cor. 1.2.19 and (5.51), we have

$$T\Big(\bigoplus_\alpha \mathcal{H}_\alpha\Big) = T\Big(\overline{\sum_\alpha \mathcal{H}_\alpha}\Big) \subset \overline{T\Big(\sum_\alpha \mathcal{H}_\alpha\Big)} \subset \overline{\sum_\alpha \mathcal{H}_\alpha} = \bigoplus_\alpha \mathcal{H}_\alpha$$

This proves that $\bigoplus_\alpha \mathcal{H}_\alpha$ is T-invariant. Since (5.52) holds when restricted to $\sum_\alpha \mathcal{H}_\alpha$, it holds on $\bigoplus_\alpha \mathcal{H}_\alpha$ due to Thm. 2.4.2. □

An especially important special case of Exp. 5.10.9 is when $\sum_\alpha \mathcal{H}_\alpha$ is dense in $\mathcal{H}$. In this situation, (5.52) reads

$$T = \oplus_{\alpha\in\mathscr{I}}(T|_{\mathcal{H}_\alpha})$$

which can be interpreted as writing T in block-diagonal form.

5.10.3 Orthogonal decompositions of unitary representations

Fix a unital $*$-algebra $\mathscr{A}$. In this subsection, we explain how a unitary representation of $\mathscr{A}$ can be decomposed into smaller components.

Definition 5.10.10. Let (π_1, V_1) and (π_2, V_2) be pre-unitary representations of $\mathscr{A}$. A bounded linear map $\Phi : V_1 \to V_2$ is called a **homomorphism** if

$$\Phi\pi_1(x) = \pi_2(x)\Phi \qquad \text{for each } x \in \mathscr{A}$$

A unitary homomorphism is called a **unitary equivalence**. If a unitary equivalence between (π_1, V_1) and (π_2, V_2) exists, we say that (π_1, V_1) and (π_2, V_2) are **unitarily equivalent**.

To simplify discussions, in the following, we focus on unitary representations.

Definition 5.10.11. Let $(\pi, \mathcal{H})$ be a unitary representation of $\mathscr{A}$. Suppose that $\mathcal{K}$ is an $\mathscr{A}$-invariant Hilbert subspace of $\mathcal{H}$. Then $(\pi|_{\mathcal{K}}, \mathcal{K})$ is also a unitary representation of $\mathscr{A}$, where

$$\pi|_{\mathcal{K}} : \mathscr{A} \to \mathfrak{L}(\mathcal{K}) \qquad x \mapsto \pi(x)|_{\mathcal{K}}$$

We say that $(\pi|_{\mathcal{K}}, \mathcal{K})$ is a **(unitary) subrepresentation** of $(\pi, \mathcal{H})$.

Definition 5.10.12. Let $(\pi_\alpha, \mathcal{H}_\alpha)_{\alpha\in\mathscr{I}}$ be a family of unitary representations of $\mathscr{A}$ satisfying

$$\sup_{\alpha\in\mathscr{I}} \|\pi_\alpha(x)\| < +\infty \qquad \text{for all } x \in \mathscr{A} \tag{5.53}$$

Then $(\oplus_{\alpha\in\mathscr{I}}\pi_\alpha, \bigoplus_{\alpha\in\mathscr{I}} \mathcal{H}_\alpha)$ is clearly a unitary representation of $\mathscr{A}$, where

$$\oplus_{\alpha\in\mathscr{I}}\pi_\alpha : \mathscr{A} \to \mathfrak{L}\Big(\bigoplus_{\alpha\in\mathscr{I}} \mathcal{H}_\alpha\Big) \qquad x \mapsto \oplus_{\alpha\in\mathscr{I}}\pi_\alpha(x)$$

We call $(\oplus_{\alpha\in\mathscr{I}}\pi_\alpha, \bigoplus_{\alpha\in\mathscr{I}} \mathcal{H}_\alpha)$ the **direct sum representation** of $(\pi_\alpha, \mathcal{H}_\alpha)_{\alpha\in\mathscr{I}}$.

Example 5.10.13. Suppose that $\mathscr{A}$ is a unital *-subalgebra of $l^\infty(X)$ where X is a set, and suppose that each $f \in \mathscr{A}$ with $f \geqslant 0$ can be written as $f = g^*g$ for some $g \in \mathscr{A}$. Let $(\pi_\alpha, \mathcal{H}_\alpha)_{\alpha\in\mathscr{I}}$ be a family of unitary representations of $\mathscr{A}$. By Prop. 5.5.6, each $(\pi_\alpha, \mathcal{H}_\alpha)$ is positive. Therefore, by Prop. 5.5.16, we have $\|\pi_\alpha(f)\| \leqslant \|f\|_{l^\infty}$ for each $f \in \mathscr{A}$. Thus (5.53) is satisfied.

Example 5.10.14. Let $(\pi, \mathcal{H})$ be a unitary representation of $\mathscr{A}$. Let $(\mathcal{H}_\alpha)_{\alpha\in\mathscr{I}}$ be a family of mutually-orthogonal $\mathscr{A}$-invariant Hilbert subspaces of $\mathcal{H}$. Viewing $\bigoplus_{\alpha\in\mathscr{I}} \mathcal{H}_\alpha$ as a $\mathscr{A}$-invariant Hilbert subspace of $\mathcal{H}$ (cf. Exp. 5.10.9), the subrepresentation $(\pi|_{\bigoplus_{\alpha\in\mathscr{I}} \mathcal{H}_\alpha}, \bigoplus_{\alpha\in\mathscr{I}} \mathcal{H}_\alpha)$ is equal to the direct sum representation of $(\pi|_{\mathcal{H}_\alpha}, \mathcal{H}_\alpha)_{\alpha\in\mathscr{I}}$.

In the important special case where $\sum_{\alpha\in\mathscr{I}} \mathcal{H}_\alpha$ is dense in $\mathcal{H}$, we therefore have

$$(\pi, \mathcal{H}) = \bigoplus_{\alpha\in\mathscr{I}} (\pi|_{\mathcal{H}_\alpha}, \mathcal{H}_\alpha) \tag{5.54}$$

which is abbreviated to $\mathcal{H} = \bigoplus_{\alpha\in\mathscr{I}} \mathcal{H}_\alpha$. In this case, we say that $(\pi, \mathcal{H})$ is the **(orthogonal) direct sum** of the family of subrepresentations $(\pi_\alpha, \mathcal{H}_\alpha)_{\alpha\in\mathscr{I}}$, or that $(\pi_\alpha, \mathcal{H}_\alpha)_{\alpha\in\mathscr{I}}$ is an **orthogonal decomposition** of $\mathcal{H}$. $\square$

Definition 5.10.15. Let $(\pi, \mathcal{H})$ be a unitary representation of $\mathscr{A}$. A vector $\xi \in \mathcal{H}$ is called a **cyclic vector** if the subspace $\pi(\mathscr{A})\xi$ is dense in $\mathcal{H}$. A unitary representation admitting a cyclic vector is called a **cyclic representation**. A subrepresentation admitting a cyclic vector is called a **cyclic subrepresentation** .

Proposition 5.10.16. *Let $(\pi, \mathcal{H})$ be a unitary representation of $\mathscr{A}$. Then $(\pi, \mathcal{H})$ is an orthogonal direct sum of cyclic subrepresentations.*

In other words, there exists a densely-spanning mutually-orthogonal family $(\mathcal{H}_\alpha)_{\alpha \in \mathscr{I}}$ of $\mathscr{A}$-invariant Hilbert subspaces of $\mathcal{H}$ such that each $\mathcal{H}_\alpha$ is cyclic.

Proof. Assume WLOG that $\mathcal{H} \neq 0$. Let $\mathscr{F}$ be the set of all $\mathscr{I} \in 2^{\mathcal{H}}$ where $\mathscr{I}$ is a set of mutually-orthogonal non-zero cyclic subrepresentations of $\mathcal{H}$. Then $(\mathscr{F}, \subset)$ is a partially ordered set. Clearly every totally ordered subset of $\mathscr{F}$ has an upper bound in $\mathscr{F}$ (defined by taking union). By Zorn's lemma, there exists a maximal element $\mathscr{I} \in \mathscr{F}$.

Let us prove that $(\mathcal{K})_{\mathcal{K} \in \mathscr{I}}$ is densely-spanning. Suppose not. Then by Cor. 3.4.8, $\mathcal{L} := \big(\sum_{\mathcal{K} \in \mathscr{I}} \mathcal{K}\big)^\perp$ is a non-zero Hilbert subspace of $\mathcal{H}$—the non-zeroness follows from the fact that $\mathcal{L}^\perp \neq \mathcal{H}$. Since $\sum_{\mathcal{K} \in \mathscr{I}} \mathcal{K}$ is $\mathscr{A}$-invariant, by Exp. 5.10.8, $\mathcal{L}$ is $\mathscr{A}$-invariant. Choose any non-zero $\xi \in \mathcal{L}$. Then $\mathscr{I} \cup \{\overline{\mathscr{A}\xi}\}$ belongs to $\mathscr{F}$ and is strictly larger than $\mathscr{I}$, contradicting the maximality of $\mathscr{I}$. □

Remark 5.10.17. If $\mathcal{H}$ is separable, Prop. 5.10.16 can be proved by induction, avoiding the use of Zorn's lemma. To see this, choose a countable sequence of vectors $\xi_1, \xi_2, \dots$ spanning a dense subspace of $\mathcal{H}$. Let $\psi_1 = \xi_1$. Suppose that ψ_n has been picked. Let P_n be the projection onto the closure of $\sum_{j=1}^n \mathscr{A}\psi_j$. Let $\psi_{n+1} = \xi_{n+1} - P_n\xi_{n+1}$. We leave it to the reader to check that $(\pi, \mathcal{H})$ has orthogonal decomposition

$$\mathcal{H} = \bigoplus_n \overline{\mathscr{A}\psi_n}$$

5.10.4 Classification of cyclic normal representations of $\mathscr{Bor}_{\mathrm{b}}(X)$

Example 5.10.18. Let (X, μ) be a measure space. Recall from the Riesz-Fischer Thm. 1.6.14, $L^2(X, \mu)$ is a Hilbert space. Then

$$\mathscr{Bor}_{\mathrm{b}}(X) \to \mathfrak{L}(L^2(X, \mu)) \qquad f \mapsto \mathbf{M}_f \tag{5.55a}$$

is a normal unitary representation of $\mathscr{Bor}_{\mathrm{b}}(X)$, where

$$\mathbf{M}_f : L^2(X, \mu) \to L^2(X, \mu) \qquad \xi \mapsto f\xi \tag{5.55b}$$

is called the **multiplication operator** of f. The representation $(\mathbf{M}, \mathscr{Bor}_{\mathrm{b}}(X))$ (abbreviated to $L^2(X, \mu)$) is called the **multiplication representation** of $\mathscr{Bor}_{\mathrm{b}}(X)$ with respect to μ.

Proof of normality. Let $(f_n)_{n \in \mathbb{Z}_+}$ be an increasing sequence in $\mathscr{Bor}_{\mathrm{b}}(X, \mathbb{R}_{\geqslant 0})$ converging pointwise to $f \in \mathscr{Bor}_{\mathrm{b}}(X, \mathbb{R}_{\geqslant 0})$. Then for each $\xi \in L^2(X, \mu)$, the sequence $(f_n|\xi|^2)_{n \in \mathbb{Z}_+}$ is increasing and converging pointwise fo $f|\xi|^2$, and hence

$$\lim_n \langle \xi | \mathbf{M}_{f_n} \xi \rangle = \lim_n \int_X f_n |\xi|^2 d\mu = \int_X f|\xi|^2 d\mu = \langle \xi | \mathbf{M}_f \xi \rangle$$

due to MCT. □

Proposition 5.10.19. *Let X be a topological space. Let $(\pi, \mathcal{H})$ be a normal unitary representation of $\mathscr{Bor}_{\mathrm{b}}(X)$. Let $\psi \in \mathcal{H}$. Then the following are equivalent.*

(1) ψ is a cyclic vector.

(2) There exists a finite Borel measure $\mu : \mathfrak{B}_X \to [0, +\infty]$ together with a unitary equivalence

$$\Phi : (\pi, \mathcal{H}) \xrightarrow{\simeq} (\mathbf{M}, L^2(X, \mu))$$

satisfying $\Phi\psi = 1$.

Proof. (2)⇒(1): Assume (2). Then we may well assume that $(\pi, \mathcal{H}) = (\mathbf{M}, L^2(X, \mu))$ (where μ is a finite Borel measure) and $\psi = 1$. Then $\pi(\mathscr{Bor}_{\mathrm{b}}(X))\psi$ equals $L^\infty(X, \mu)$ (viewed as a subspace of $L^2(X, \mu)$). Since simple functions are dense in $L^2(X, \mu)$, the space $L^\infty(X, \mu)$ is also dense in $L^2(X, \mu)$. Therefore ψ is a cyclic vector. This proves (1).

(1)⇒(2): Assume (1). Since $(\pi, \mathcal{H})$ is normal, for each $f \in \mathscr{Bor}_{\mathrm{b}}(X)$ we have

$$\langle \psi | \pi(f)\psi \rangle = \int_X f d\mu_\psi$$

where μ is the finite Borel measure associated to ξ, cf. Rem. 5.5.9. Let $\mu = \mu_\psi$. Then for each $f, g \in \mathscr{Bor}_{\mathrm{b}}(X)$ we have

$$\langle \pi(f)\psi | \pi(g)\psi \rangle = \langle \psi | \pi(\overline{f}g)\psi \rangle = \int_X \overline{f} g d\mu_\psi = \langle 1 | \mathbf{M}_{\overline{f}} \mathbf{M}_g 1 \rangle = \langle \mathbf{M}_f 1 | \mathbf{M}_g 1 \rangle \qquad (5.56)$$

It follows from the following Lem. 5.10.20 that there is a unique unitary map

$$\Phi : \mathcal{H} = \overline{\mathscr{Bor}_{\mathrm{b}}(X)\psi} \longrightarrow L^2(X, \mu)$$

sending $\pi(f)\psi$ to $f = \mathbf{M}_f 1$; in particular, $\Phi\psi = 1$.

For each $f, g \in \mathscr{Bor}_{\mathrm{b}}(X)$, we compute that

$$\Phi\pi(f)\pi(g)\psi = \Phi\pi(fg)\psi = fg = \mathbf{M}_f g = \mathbf{M}\Phi\pi(g)\psi$$

Therefore, the bounded linear operators $\Phi\pi(f)$ and $\mathbf{M}_f\Phi$ agree on $\mathscr{Bor}_{\mathrm{b}}(X)\psi$, and hence on $\mathcal{H}$ due to the continuity. Thus Φ is a homomorphism. □

Lemma 5.10.20. *Let $(\xi_\alpha)_{\alpha \in \mathscr{I}}$ be a densely spanning family in an inner product space V. Let $(\eta_\alpha)_{\alpha \in \mathscr{I}}$ be a family in a Hilbert space $\mathcal{K}$. Suppose that*

$$\langle \xi_\alpha | \xi_\beta \rangle = \langle \eta_\alpha | \eta_\beta \rangle \qquad \text{for each } \alpha, \beta \in \mathscr{I} \qquad (5.57)$$

Then there exists a unique linear isometry $\Phi : V \to \mathcal{K}$ sending each ξ_α to η_α. Moreover, if V is a Hilbert space and $(\eta_\alpha)_{\alpha \in \mathscr{I}}$ spans a dense subspace of $\mathcal{K}$, then Φ is unitary.

Proof. The uniqueness is obvious. Let us prove the existence. For each finite sum $\sum_\alpha c_\alpha\xi_\alpha$ (where $c_\alpha \in \mathbb{C}$), we have $\|\sum_\alpha c_\alpha\xi_\alpha\| = \|\sum_\alpha c_\alpha\eta_\alpha\|$ due to (5.57). Therefore, if we let $V_0 = \mathrm{Span}\{\xi_\alpha : \alpha \in \mathscr{I}\}$, then we have a well-defined linear map $\Phi : V_0 \to W$ sending each ξ_α to η_α. Moreover, by (5.57), the map Φ is a linear isometry; in particular, it is bounded. Since $\mathcal{K}$ is complete, by Thm. 2.4.2, Φ can be extended to a bounded linear map $\Phi : V \to \mathcal{K}$. Since $\|\Phi\xi\| = \|\xi\|$ holds for all $\xi \in V_0$, by the continuity of Φ, it follows that $\|\Phi\xi\| = \|\xi\|$ holds for all $\xi \in V$. Thus $\Phi : V \to \mathcal{K}$ is a linear isometry.

That Φ is surjective when V is a Hilbert space follows from the following Lem. 5.10.21. □

Lemma 5.10.21. *Let $T : V \to W$ be a bounded linear map between normed vector spaces, where V is complete. Assume that there exists $C \in \mathbb{R}_{>0}$ such that*

$$\|T\xi\| \geqslant C\|\xi\| \qquad \text{for all } \xi \in V \tag{5.58}$$

Then $T(V)$ is closed in W.

First proof. Let $Z = \overline{T(V)}$. By (5.58), we have $T\xi = 0 \Rightarrow \xi = 0$. Therefore T is injective. Let $S : T(V) \to V$ be its inverse. Condition (5.58) implies that S is bounded. Therefore, since V is complete, by Thm. 2.4.2, S can be extended to a bounded linear map $\widetilde{S} : Z \to V$. Since $TS = \mathrm{id}_{T(V)}$, we have $T\widetilde{S}|_{T(V)} = \mathrm{id}_{T(V)}$, and hence $T\widetilde{S} = \mathrm{id}_Z$ by the continuity of T and S. This proves that T has range $Z = \overline{T(V)}$, and hence $T(V)$ is closed. □

This proof makes explicit how the Cauchy completeness is used as a codomain condition (cf. Subsec. 2.5.1), although it initially appears to be a domain property. We now present a different argument.

Second proof. Since any Cauchy-complete subset of W is closed (cf. Prop. 1.4.6), it suffices to show that $T(V)$ is complete. Let (η_n) be a Cauchy sequence in W. Write $\eta_n = T\xi_n$ where $\xi_n \in V$. Then (5.58) indicates that (ξ_n) is a Cauchy sequence in V. Since V is complete, the sequence (ξ_n) converges to some $\xi \in V$. Since T is continuous, the limit $\lim_n \eta_n = \lim_n T\xi_n$ converges to $T\xi$. □

5.10.5 The multiplication-operator version of the spectral theorem

Let $\mathcal{H}$ be a Hilbert space, and let $T_1 \dots, T_N \in \mathfrak{L}(\mathcal{H})$ be adjointly-commuting normal operators.

Theorem 5.10.22 (Spectral theorem). *There exists a family $(\mu_\alpha)_{\alpha\in\mathscr{I}}$ of finite Borel measures on a compact set $X \subset \mathbb{C}^N$, together with a unitary map*

$$\Phi : \mathcal{H} \xrightarrow{\simeq} \bigoplus_{\alpha\in\mathscr{I}} L^2(X, \mu_\alpha)$$

such that for each $1 \leqslant i \leqslant N$, we have

$$\Phi T_i \Phi^{-1} = \oplus_{\alpha \in \mathscr{I}} \mathbf{M}_{z_i} \tag{5.59}$$

as bounded linear operators on $\bigoplus_{\alpha \in \mathscr{I}} L^2(X, \mu_\alpha)$.

As usual, z_i denotes the i-th coordinate function of $\mathbb{C}^N$. Then (5.59) means that for each $\oplus_\alpha \xi_\alpha \in \bigoplus_\alpha L^2(X, \mu_\alpha)$, we have

$$\Phi T_i \Phi^{-1}(\oplus_\alpha \xi_\alpha) = \oplus_\alpha z_i \xi_\alpha$$

Proof. Let $X = \mathrm{Sp}(T_\bullet)$, which is a compact subset of $\mathbb{C}^N$. Let $\pi : \mathscr{B}or_{\mathrm{b}}(X) \to \mathfrak{L}(\mathcal{H})$ be the Borel functional calculus (cf. Thm. 5.7.13 and Prop. 5.7.12). Then $(\pi, \mathcal{H})$ is a normal unitary representation of $\mathscr{B}or_{\mathrm{b}}(X)$.

By Prop. 5.10.16, $(\pi, \mathcal{H})$ admits an orthogonal decomposition into (automatically normal) cyclic subrepresentations $\mathcal{H} = \bigoplus_{\alpha \in \mathscr{I}} \mathcal{H}_\alpha$. By Prop. 5.10.19, each $(\pi|_{\mathcal{H}_\alpha}, \mathcal{H}_\alpha)$ is unitarily equivalent to the multiplication representation of $\mathscr{B}or_{\mathrm{b}}(X)$ with respect to some finite Borel measure μ_α on X. Therefore, $(\pi, \mathcal{H})$ is unitarily equivalent to $(\oplus_\alpha \mathbf{M}, \bigoplus_\alpha L^2(X, \mu_\alpha))$ via a unitary operator $\Phi : \mathcal{H} \to \bigoplus_\alpha L^2(X, \mu_\alpha)$. That is,

$$\Phi \pi(f) \Phi^{-1} = \oplus_\alpha \mathbf{M}_f \qquad \text{for each } f \in \mathscr{B}or_{\mathrm{b}}(X)$$

Since $\pi(z_i) = T_i$, we obtain (5.59). □

With the multiplication-operator version of the spectral theorem, one can give an explicit description of the Borel functional calculus:

Theorem 5.10.23. *Let $(X, \mathfrak{M})$ be a measurable space. Let $f_1, \dots, f_N : X \to \mathbb{C}$ be bounded measurable functions. Let $(\mu_\alpha)_{\alpha \in \mathscr{I}}$ be a family of measures on $\mathfrak{M}$. Let $\mathcal{H} = \bigoplus_{\alpha \in \mathscr{I}} L^2(X, \mu_\alpha)$. For each $1 \leqslant i \leqslant N$, define $T_i \in \mathfrak{L}(\mathcal{H})$ by*

$$T_i = \oplus_{\alpha \in \mathscr{I}} \mathbf{M}_{f_i}$$

Then $T_1, \dots, T_N$ are adjointly-commuting bounded normal operators. Moreover, for each $g \in \mathscr{B}or_{\mathrm{b}}(\mathbb{C}^N)$, the Borel functional calculus satisfies

$$g(T_1, \dots, T_N) = \oplus_{\alpha \in \mathscr{I}} \mathbf{M}_{g \circ f_i} \tag{5.60}$$

Note that since $\sup_\alpha \|\mathbf{M}_{f_i}\| \leqslant \|f_i\|_{l^\infty}$, the direct sum operator $\oplus_{\alpha \in \mathscr{I}} \mathbf{M}_{f_i}$ can be defined as a bounded linear operator.

Proof. Note that $(\oplus_{\alpha \in \mathscr{I}} \mathbf{M}, \mathcal{H})$ is a unitary representaiton of $\mathscr{B}or_{\mathrm{b}}(X)$, and that the range of any unitary representaiton of a commutative *-algebra consists of adjointly-commuting bounded normal operators. Therefore, $T_1, \dots, T_N$ are mutually-commuting normal operators. Moreover, the map

$$\pi : \mathscr{B}or_{\mathrm{b}}(\mathbb{C}^N) \to \mathfrak{L}(\mathcal{H}) \qquad g \mapsto \mathbf{M}_{g \circ f_i}$$

is clearly a unitary representation supported in K where $K \subset \mathbb{C}^N$ is any compact set containing $\mathrm{Rng}(f_\bullet) = \mathrm{Rng}(f_1, \dots, f_N)$. It remains to prove that $(\pi, \mathcal{H})$ is normal.

Choose an increasing sequence (g_n) in $\mathscr{Bor}_{\mathrm{b}}(\mathbb{C}^N, \mathbb{R}_{\geqslant 0})$ converging pointwise to $g \in \mathscr{Bor}_{\mathrm{b}}(\mathbb{C}^N, \mathbb{R}_{\geqslant 0})$. Let $\xi = \bigoplus_{\alpha \in \mathscr{I}} \xi_\alpha$ be an element of $\mathcal{H}$. MCT implies

$$\lim_n \langle \xi | \pi(g_n) \xi \rangle = \lim_n \sum_{\alpha \in \mathscr{I}} \int_X (g_n \circ f_\bullet) |\xi_\alpha|^2 d\mu_\alpha = \sum_{\alpha \in \mathscr{I}} \lim_n \int_X (g_n \circ f_\bullet) |\xi_\alpha|^2 d\mu_\alpha$$

$$= \sum_{\alpha \in \mathscr{I}} \int_X (g \circ f) |\xi_\alpha|^2 d\mu_\alpha = \langle \xi | \pi(g) \xi \rangle$$

recalling that $\sum_{\alpha \in \mathscr{I}}$ coincides with the integral on $\mathscr{I}$ with respect to the counting measure (so that MCT applies). □

Convention 5.10.24. Henceforth, when the context is clear, we abbreviate $\bigoplus_{\alpha \in \mathscr{I}} \mathbf{M}_f$ to $\mathbf{M}_f$, and also refer to it a **multiplication operator**.

Exercise 5.10.25. In the setting of Thm. 5.10.23, choose an element $\xi = \bigoplus_{\alpha \in \mathscr{I}} \xi_\alpha$ of $\mathcal{H}$. Recall the measure μ_ξ in Rem. 5.5.9. Show that

$$d\mu_\xi = \sum_{\alpha \in \mathscr{I}} (f_\bullet)_* \big(\|\xi_\alpha\|^2 d\mu_\alpha \big) \tag{5.61}$$

where all but countably many summands are zero due to Prop. 1.2.44. Note that in the special case where $X \subset \mathbb{C}^N$ and $f_i = z_i$, the above relation becomes

$$d\mu_\xi = \sum_{\alpha \in \mathscr{I}} \|\xi_\alpha\|^2 d\mu_\alpha \tag{5.62}$$

The joint spectra can also be described explicitly:

Exercise 5.10.26. In the setting of Thm. 5.10.23, assume that each μ_α is σ-finite. Prove that

$$\mathrm{Sp}(T_\bullet) = \mathrm{Cl}_{\mathbb{C}^N} \Big(\bigcup_{\alpha \in \mathscr{I}} \mathrm{Rng}^{\mathrm{ess}}(f_\bullet, \mu_\alpha) \Big) \tag{5.63}$$

where $\mathrm{Rng}^{\mathrm{ess}}(f_\bullet, \mu_\alpha)$ is the essential range (cf. Def. 1.6.6) of the map $f_\bullet = (f_1, \dots, f_N) : X \to \mathbb{C}^N$ with respect to μ_α. [10] Note that in the special case where $X \subset \mathbb{C}^N$ and $f_i = z_i$, the above relation becomes

$$\mathrm{Sp}(T_\bullet) = \mathrm{Cl}_{\mathbb{C}^N} \Big(\bigcup_{\alpha \in \mathscr{I}} \mathrm{Supp}(\mu_\alpha) \Big) \tag{5.64}$$

[10]Indeed, the relation "$\subset$" can be proved without assuming σ-finiteness.

We encourage the reader to use the results from this subsection to provide alternative proofs for the basic properties of the Borel functional calculus and joint spectra, namely those in Sec. 5.8. I must confess, when I was proving certain propositions in Sec. 5.8 (for example, Prop. 5.8.8), I often had in mind the concrete case of multiplication operators as in Thm. 5.10.22—sometimes even restricting to the situation where the index set $\mathscr{I}$ consists of a single point, thanks to Prop. 5.10.19—and then figured out the argument. This is my secret method for working with the Borel functional calculus.

5.11 Problems

In this section, we fix Hilbert spaces $\mathcal{H}, \mathcal{K}$.

Problem 5.1. Let $-\infty < a < b < +\infty$. Let $f \in \mathbb{C}[x]$. Prove that $f|_{[a,b]} \geqslant 0$ iff f can be written as

$$\overline{p_1(x)}p_1(x) + (x-a)\overline{p_2(x)}p_2(x) + (b-x)\overline{p_3(x)}p_3(x)$$

where $p_1, p_2, p_3 \in \mathbb{C}[x]$. (Assume for simplicity that $a = 0, b = 1$.)

Hint. Show that $f \in \mathbb{R}[x]$. Factor the polynomial f into linear factors. Show that any real root of f with odd multiplicity must be outside the open interval (a, b). □

The following problem complements Exp. 5.7.2.

† **Problem 5.2.** Let $\Phi \in \mathfrak{L}(\mathcal{H}, \mathcal{K})$. Prove that the following conditions are equivalent:

(1) There exists closed linear subspaces $V \subset \mathcal{H}$ and $W \subset \mathcal{K}$ such that Φ restricts to a unitary map $\Phi|_V : V \xrightarrow{\simeq} W$, and $\Phi|_{V^\perp}$ is zero.

(2) There exists closed linear subspaces $V' \subset \mathcal{H}$ and $W' \subset \mathcal{K}$ such that Φ^* restricts to a unitary map $\Phi^*|_{W'} : W' \xrightarrow{\simeq} V'$, and $\Phi^*|_{(W')^\perp}$ is zero.

(3) $\Phi^*\Phi$ is a projection operator on $\mathcal{H}$.

(4) $\Phi\Phi^*$ is a projection operator on $\mathcal{K}$.

Prove that if these conditions are satisfied, then

$$\mathrm{Rng}(\Phi^*\Phi) = V = V' \qquad \mathrm{Rng}(\Phi\Phi^*) = W = W' \tag{5.65}$$

Such Φ is called a **partial isometry**, and V and W are called the **source space** and the **target space** of Φ respectively. Clearly Φ^* is a partial isometry with source space W and target space V. □

Suggestion. Suppose that (1) holds. Prove that (2) holds with $V' = V$ and $W' = W$. Then prove that (3) holds and $\mathrm{Rng}(\Phi^*\Phi) = V$. A similar argument yields $\mathrm{Rng}(\Phi\Phi^*) = W$. In particular, V and W are uniquely determined by Φ.

Exchanging the roles of Φ and Φ^*, we conclude that (1)⇔(2), (1)⇒(3), (2)⇒(4), and that (5.65) holds. In particular, V, V', W, W' are uniquely determined by Φ.

Finally, prove (3)⇒(1) (and similarly, (4)⇒(2)) by computing $\langle\Phi\Phi^*\Phi\xi|\Phi\Phi^*\Phi\xi\rangle$ and $\langle\Phi(\mathbf{1}-\Phi^*\Phi)\xi|\Phi(\mathbf{1}-\Phi^*\Phi)\xi\rangle$. □

Definition 5.11.1. The **strong-* operator topology (SOT*)** on $\mathfrak{L}(\mathcal{H},\mathcal{K})$ is defined as the pullback topology along the map

$$\mathfrak{L}(\mathcal{H},\mathcal{K}) \to \mathfrak{L}(\mathcal{H},\mathcal{K}) \times \mathfrak{L}(\mathcal{K},\mathcal{H}) \qquad T \mapsto (T, T^*)$$

where the RHS carries the product topology of SOT on $\mathfrak{L}(\mathcal{H},\mathcal{K})$ and on $\mathfrak{L}(\mathcal{K},\mathcal{H})$, respectively. Equivalently, a net (T_α) in $\mathfrak{L}(\mathcal{H},\mathcal{K})$ converges to $T \in \mathfrak{L}(\mathcal{H},\mathcal{K})$ in SOT* iff $(T_\alpha) \to T$ in SOT and simultaneously $(T_\alpha^*) \to T^*$ in SOT.

Problem 5.3. Let (T_α) be a net in $\mathfrak{L}(\mathcal{H})$. Let $T \in \mathfrak{L}(\mathcal{H})$.

1. Assume that (T_α) converges in WOT to T. Prove that (T_α^*) converges in WOT to T^*.

2. Assume that each T_α is normal, and that (T_α) converges in SOT* to $T \in \mathfrak{L}(\mathcal{H})$. Prove that T is normal.

3. Assume that T and each T_α are normal, and that (T_α) converges in SOT to T. Prove that (T_α) converges in SOT* to T.

Note. Part 2. One cannot conclude that $(T_\alpha^* T_\alpha)$ converges in SOT to T^*T, since $\sup_{\alpha\in I}\|T_\alpha\|$ is not assumed to be finite. (Cf. Cor. 3.7.14.) Consider WOT instead.

Part 3. Recall Rem. 3.7.8. □

Problem 5.4. Let (T_α) be a net of bounded normal operators on $\mathcal{H}$ satisfying $\sup_\alpha\|T_\alpha\| \leqslant R$ for some $R \in \mathbb{R}_{\geqslant 0}$. Let $f \in C(\overline{B}_{\mathbb{C}}(0,R))$. Assume that (T_α) converges in SOT* to $T \in \mathfrak{L}(\mathcal{H})$. Note that T is normal by Pb. 5.3. Prove that

$$\lim_\alpha f(T_\alpha) = f(T) \qquad \text{in SOT*} \tag{5.66}$$

Hint. Approximate f uniformly by polynomials of z and $\overline{z}$. □

Remark 5.11.2. Let $T \in \mathfrak{L}(\mathcal{H})$ be self-adjoint and $R \geqslant \|T\|$. For each $z \in \mathbb{C}$ satisfying $|z| > R$, let $f_z(\lambda) = (z-\lambda)^{-1}$. Since the Borel functional calculus preserves the multiplication of functions, we have

$$f_z(T)(z-T) = (z-T)f_z(T) = \mathbf{1}$$

and hence $f_z(T) = (z - T)^{-1}$.

Now suppose that $\mathcal{H}$ is separable and infinite-dimensional, and let $e_1, e_2, \ldots$ be an orthonormal basis of $\mathcal{H}$. Let E_n be the projection of $\mathcal{H}$ onto $\mathrm{Span}\{e_1, \ldots, e_n\}$. Then $\|E_n T E_n\| \leqslant R$, and hence $f_z(E_n T E_n) = (z - E_n T E_n)^{-1}$. By Pb. 5.4, we have

$$\lim_n (z - E_n T E_n)^{-1} = (z - T)^{-1} \qquad \text{in SOT*}$$

This proves that Hilbert's resolvent is equal to the usual definition of resolvent for bounded linear operators, as discussed in the answer to Question 5.1.2. □

Problem 5.5. Let $P \in \mathfrak{L}(\mathcal{H})$ be normal. Prove that P is a projection iff $\mathrm{Sp}(P) \subset \{0, 1\}$.

5.11.1 Application of the spectral theorem to moment problems

In this chapter, we have used the polynomial moment problem in Sec. 4.2 as a central motivation for the spectral theorem. Conversely, the spectral theorem can in turn be applied to solve moment problems—specifically, to prove Thm. 4.2.9. In Exp. 5.11.3 below, we explain how the spectral theorem for bounded self-adjoint operators immediately establishes Part 3 of Thm. 4.2.9, i.e. the solution of the Hausdorff moment problem. In the future, we will see how the spectral theorem for unbounded self-adjoint operators provides alternative proofs for the first two parts of Thm. 4.2.9. (See Exp. 6.11.2 and Exe. 7.2.7.)

Example 5.11.3. Let $(c_n)_{n \in \mathbb{N}}$ be a sequence in $\mathbb{R}$. Define H and H' as in Def. 4.2.6. As in the proof of Prop. 4.2.8, if there is a finite Borel measure μ on $[0, 1]$ satisfying $\int_{[0,1]} x^n d\mu = c_n$ for all n, then necessarily $H, H', H - H' \geqslant 0$.

Conversely, assume that $H, H', H - H' \geqslant 0$. Let $\mathscr{A} = \mathbb{C}[x]$, and let $\Lambda : \mathscr{A} \to \mathbb{C}$ be the unique linear functional such that $\Lambda(x^n) = c_n$ for all n. Since $H \geqslant 0$, by Pb. 4.1 and Rem. 4.6.1, the space $\mathscr{A}$ admits a positive sesquilinear form defined by $\langle f | g \rangle = \Lambda(\overline{f} g)$, which descends to an inner product on $V = \mathscr{A}/\mathscr{N}$ where $\mathscr{N} = \{g \in \mathscr{A} : \langle g | g \rangle = 0\}$. Moreover, $\mathscr{A}$ has a pre-unitary representation (π, V) such that $\pi(f)(g + \mathscr{N}) = fg + \mathscr{N}$. Let $T = \pi(x)$. Then the condition $H', H - H' \geqslant 0$ implies that

$$0 \leqslant \langle \xi | T\xi \rangle \leqslant \langle \xi | \xi \rangle \qquad \text{for all } \xi \in V$$

By Prop. 3.2.12, we have $\|\omega_T\| \leqslant 4$, and hence (by Prop. 3.5.2) $\|T\| \leqslant 4$.

Let $\mathcal{H}$ be the Hilbert space completion of V (cf. Pb. 3.2). By Thm. 2.4.2, T can be extended uniquely to an element of $\mathfrak{L}(\mathcal{H})$, also denoted by T. We still have $0 \leqslant T \leqslant 1$, and hence $\mathrm{Sp}(T) \subset [0, 1]$ due to Prop. 5.8.14. Let $\Omega = 1 + \mathscr{N}$. By Rem. 5.5.9, there is a finite Borel measure $\mu = \mu_\Omega$ on $[0, 1]$ such that the Borel functional calculus of T satisfies

$$\langle \Omega | f(T)\Omega \rangle = \int_{[0,1]} f d\mu \qquad \text{for each } f \in \mathscr{B}\!or_{\mathrm{b}}([0, 1])$$

Taking $f(x) = x^n$, we obtain $c_n = \Lambda(x^n) = \langle\Omega|T^n\Omega\rangle = \int_{[0,1]} x^n d\mu$. Thus μ solves the Hausdorff moment problem. $\square$

Next, the reader is asked to solve a trigonometric moment problem by a similar method. Let $N \in \mathbb{Z}_+$. Recall that $\mathbb{T} = \mathbb{S}^1 = \{z \in \mathbb{C} : |z| = 1\}$. For each $z = (z_1, \ldots, z_N) \in \mathbb{T}^N$ and $n = (n_1, \ldots, n_N) \in \mathbb{Z}^N$, let

$$z^n = z_1^{n_1} \cdots z_N^{n_N} \tag{5.67}$$

Definition 5.11.4. For each finite Borel measure μ on $\mathbb{T}^N$, define the **Fourier series** $\widehat{\mu} : \mathbb{Z}^N \to \mathbb{C}$ by

$$\widehat{\mu}(n) = \int_{\mathbb{T}^N} z^{-n} d\mu(z)$$

Note that $|\widehat{\mu}(n)| \leqslant \mu(\mathbb{T}^N) < +\infty$, and hence $\widehat{\mu} \in l^\infty(\mathbb{Z}^N)$. Moreover, by Stone-Weierstrass, $\mathrm{Span}\{z^n : n \in \mathbb{Z}^N\}$ is dense in $C(\mathbb{T}^N)$. Thus μ is uniquely determined by its Fourier series $\widehat{\mu}$.

The following result is known as the **Carathéodory-Toeplitz theorem**.

Problem 5.6. (**Trigonometric moment problem**) Let $c_\bullet = (c_n)_{n\in\mathbb{Z}^N}$ be a family in $\mathbb{C}$. Prove that the following conditions are equivalent.

(1) There exists a finite Borel measure μ on $\mathbb{T}^N$ such that $\widehat{\mu} = c_\bullet$.

(2) $c_\bullet$ is **positive-definite** in the sense that the matrix $C \in \mathbb{C}^{\mathbb{Z}^N\times\mathbb{Z}^N}$ defined by $C(m,n) = c_{m-n}$ (where $m,n \in \mathbb{Z}^N$) is positive. That is, for each $f : \mathbb{Z}^N \to \mathbb{C}$ with finite support, we have

$$\sum_{m,n\in\mathbb{Z}^N} \overline{f(m)} c_{m-n} f(n) \geqslant 0$$

Hint for (2)⇒(1). Let $\mathscr{A}$ be the set of polynomials of $z_1, z_1^{-1}, \ldots, z_N, z_N^{-1}$, viewed as a unital *-subalgebra of $C(\mathbb{T}^N)$. Show that the linear functional

$$\Lambda : \mathscr{A} \to \mathbb{C} \qquad \Lambda(z^{-n}) = c_n$$

satisfies the condition in Pb. 4.1 so that the GNS construction gives a pre-unitary representation (π, V) of $\mathscr{A}$. Show that $U_j = \pi(z_j)$ (for $1 \leqslant j \leqslant N$) are unitary and (adjointly) commuting, and hence can be extended to adjointly-commuting unitary operators on the Hilbert space completion $\mathcal{H}$. Apply the spectral theorem to $U_1, \ldots, U_N$. $\square$

5.11.2 The origin of the Borel functional calculus and the abstract Riesz-Fischer theorem

As discussed in Rem. 5.6.1 and the answer to Question 5.4.1, Riesz used the semicontinuous functional calculus to establish his spectral theorem in [Rie13], rather than the more general Borel functional calculus. The Borel functional calculus was first developed by von Neumann in [vN31] to prove the following theorem.

Theorem 5.11.5. *Assume that $\mathcal{H}$ is separable, and let $T_1, T_2, \dots \in \mathfrak{L}(\mathcal{H})$ be a possibly finite sequence of mutually-commuting bounded self-adjoint operators. Then there exist a self-adjoint $S \in \mathfrak{L}(\mathcal{H})$ and $f_1, f_2, \dots \in \mathscr{Bor}_{\mathrm{b}}(\mathbb{R}, \mathbb{R})$ such that $T_i = f_i(S)$ for each i.*

Although von Neumann did not make this explicit, part of his motivation for studying Thm. 5.11.5 may have been to reduce the study of the functional calculus for several operators to that of a single operator: indeed, in Thm. 5.11.5 one has $g(T_1, T_2, \dots) = g \circ (f_1, f_2, \dots)(S)$ for $g \in \mathscr{Bor}_{\mathrm{b}}(\mathbb{R} \times \mathbb{R} \times \cdots)$. See [vN32a] (especially Sec. II.10, III.1, III.3) for von Neumann's use of this result in developing mathematical interpretations of quantum mechanics.

The Borel functional calculus is crucial to this theorem. In fact, as we will see in the proof, the operator S can also be written as $g(T_1, T_2, \dots)$ if the number of the operators $T_1, T_2, \dots$ is finite. This enhanced version of Thm. 5.11.5 does not hold if the functions f_i and g are only assumed to be semicontinuous—even in the case of just two operators T_1, T_2.

While proving Thm 5.11.5, von Neumann also established the abstract Riesz-Fischer theorem: the completeness of $L^2([a, b], \mu)$ where μ is a finite Borel measure on $[a, b]$. (See Pb. 5.8 for where this theorem is used.) His argument—similar to his proof of completeness in [vN27, Anhang 2] for $L^2(\Omega, m)$ where $\Omega \subset \mathbb{R}^n$ and m is the Lebesgue measure—is much closer to the modern proof of the Riesz-Fischer Thm. 1.6.14, and therefore extends naturally to arbitrary measure spaces. By contrast, the original proofs of Riesz and Fischer for the Lebesgue measure on $[a, b]$ relied heavily on the fundamental theorem of calculus (see [Gui-A, Sec. 27.3] for a discussion), and thus do not generalize.

The aim of this subsection is to give a proof of Thm. 5.11.5, following von Neumann's approach in [vN31]. The only exception is the proof of Thm. 5.11.11, where we follow the argument in [Dav, Lem. II.2.8], which differs from von Neumann's original proof in [vN29b].

Definition 5.11.6. Throughout this subsection, we let $\mathfrak{S}$ be a set of adjointly-commuting bounded normal operators on $\mathcal{H}$. Let $\mathbf{\mathbb{C}[\mathfrak{S} \cup \mathfrak{S}^*]}$ be the set of polynomials of elements of $\mathfrak{S} \cup \mathfrak{S}^*$, which is clearly a commutative unital *-subalgebra of $\mathfrak{L}(\mathcal{H})$. Let $\boldsymbol{W^*(\mathfrak{S})}$ be the closure of $\mathbb{C}[\mathfrak{S} \cup \mathfrak{S}^*]$ in SOT*.

In the following, we always let $\mathscr{A} = \mathbb{C}[\mathfrak{S} \cup \mathfrak{S}^*]$. Note that $W^*(\mathfrak{S}) = W^*(\mathscr{A})$.

Example 5.11.7. Let $T_1,\dots,T_N \in \mathfrak{S}$ and $f \in \mathscr{Bor}_{\mathrm{b}}(\mathbb{C}^N)$. Then $f(T_\bullet) \in W^*(\mathfrak{S})$.

Proof. By Lem. 5.5.12, there is a net (f_α) of polynomials of $z_1,\overline{z_1},\dots,z_N,\overline{z_N}$ converging to f in the universal L^2-topology on $\mathrm{Sp}(T_\bullet)$. By Prop. 5.8.5, $f_\alpha(T_\bullet)$ converges in SOT $f(T_\bullet)$. Since $f(T_\bullet)$ is normal (Thm. 5.8.6), the convergence is in SOT* due to Pb. 5.3. Since $f_\alpha(T_\bullet) \in \mathscr{A}$, we have $f(T_\bullet) \in W^*(\mathfrak{S})$. □

Remark 5.11.8. By Prop. 3.5.14, if $(A_\alpha)_{\alpha\in I}$ and $(B_\beta)_{\beta\in J}$ are nets in $\mathscr{A}$ converging in norm to $A,B \in \mathfrak{L}(\mathcal{H})$ respectively, then $(A_\alpha B_\beta)_{(\alpha,\beta)\in I\times J}$ converges in norm to AB. Thus, the norm closure of $\mathscr{A}$ is closed under multiplication, and hence is a unital *-subalgebra of $\mathfrak{L}(\mathcal{H})$. The proof that $W^*(\mathfrak{S})$ is closed under multiplication is much less straightforward, as we will establish in the following theorem.

Theorem 5.11.9. *$W^*(\mathfrak{S})$ is an (automatically SOT*-closed) unital *-subalgebra of $\mathfrak{L}(\mathcal{H})$.*

Problem 5.7. The goal of this problem is to prove Thm. 5.11.9.

1. Prove that any two elements of $W^*(\mathfrak{S})$ are normal and commute adjointly.

2. Let $S,T \in \mathfrak{L}(\mathcal{H})$ be normal and adjointly-commuting. Assume that $f \in \mathscr{Bor}(\mathbb{C})$ has Lipschitz constant C when restricted to $K := \mathrm{Sp}(S) \cup \mathrm{Sp}(T)$. Prove that
$$\|f(S)\xi - f(T)\xi\|^2 \leqslant C^2 \cdot \|S\xi - T\xi\|^2$$
(Note that $f|_K$ is bounded due to the Lipschitz-continuity. Thus $f(S), f(T)$ can be defined as bounded operators.)

3. Let $(T_\alpha)_{\alpha\in\mathscr{I}}$ be a net in $W^*(\mathfrak{S})$ converging in SOT* to $T \in \mathfrak{L}(\mathcal{H})$. (So $T \in W^*(\mathfrak{S})$.) Assume that $f \in \mathscr{Bor}(\mathbb{C})$ is Lipschitz continuous on a Borel set $K \subset \mathbb{C}$ containing $\mathrm{Sp}(T)$ and all $\mathrm{Sp}(T_\alpha)$. Prove that
$$\lim_\alpha f(T_\alpha) = f(T) \qquad \text{in SOT*}$$
(Note that Pb. 5.4 is not applicable here, since we do not assume $\sup_\alpha \|T_\alpha\| < +\infty$.)

4. Prove Thm. 5.11.9.

Hint. 2. Use the multiplication-operator version of the spectral theorem. Alternatively, apply the Borel functional calculus to S,T and the function $g(z,w) = f(z) - f(w)$ defined on $K \times K$.

4. Let $S,T \in W^*(\mathfrak{S})$. To show $ST \in W^*(\mathfrak{S})$, restrict to the case where S,T are self-adjoint. Choose nets $(S_\alpha)_{\alpha\in I}$ and $(T_\beta)_{\beta\in J}$ in $\mathscr{A}$ converging in SOT to S,T. Find a suitable bounded Lipschitz-continuous $f:\mathbb{R}\to\mathbb{R}$ such that $f(x)=x$ whenever $x \in \mathrm{Sp}(S)\cup\mathrm{Sp}(T)$. Use Cor. 3.7.14 to show that $\lim_{(\alpha,\beta)\in I\times J} f(S_\alpha)f(T_\beta)$ converges in SOT* to ST. □

Remark 5.11.10. It follows from Thm. 5.11.9 that $W^*(\mathfrak{S})$ is the smallest SOT*-closed unital *-subalgebra of $\mathfrak{L}(\mathcal{H})$ containing $\mathfrak{S}$. In particular, if $\mathfrak{T} \subset W^*(\mathfrak{S})$, then $W^*(\mathfrak{T}) \subset W^*(\mathfrak{S})$.

Theorem 5.11.11. *Assume that $\mathcal{H}$ is separable, and $\mathfrak{S} = \{T_1, \dots, T_N\}$ is a finite set. Then*

$$W^*(\mathfrak{S}) = \{f(T_\bullet) : f \in \mathscr{Bor}_{\mathrm{b}}(\mathbb{C}^N)\}$$

Problem 5.8. In this problem, we prove Thm. 5.11.11. By Exp. 5.11.7, it suffices to show that any $S \in W^*(\mathfrak{S})$ is a Borel functional calculus of $T_\bullet$. By considering $\mathrm{Re}(S)$ and $\mathrm{Im}(S)$ separately, it suffices to assume that $S = S^*$.

1. Show that there exists a net $(p_\alpha)_{\alpha \in \mathscr{I}}$ of polynomials of $z_1, \overline{z_1}, \dots, z_N, \overline{z_N}$ such that $(p_\alpha(T_\bullet))$ converges in SOT* to S, and that each p_α is real-valued on $\mathbb{C}^N$.
2. Show that there exists a net $(f_\alpha)_{\alpha \in \mathscr{I}}$ of Borel functions $\mathbb{C}^N \to \mathbb{R}$ with $\sup_\alpha \|f_\alpha\|_{l^\infty(\mathbb{C})} < +\infty$ such that $(f_\alpha(T_\bullet))$ converges in SOT* to S.
3. Let $(\psi_n)_{n \in \mathbb{Z}_+}$ be a densely-spanning sequence in $\mathcal{H}$ satisfying $\sum_n \|\psi_n\|^2 < +\infty$. Define a Borel measure on $\mathbb{C}^N$ by
$$\mu = \sum_n \mu_{\psi_n}$$
where μ_{ψ_n} is the finite Borel measure associated to ψ_n and $T_\bullet$. (That is, $\langle \psi_n | g(T_\bullet) \psi_n \rangle = \int_{\mathbb{C}^N} g d\mu_{\psi_n}$ for each $g \in \mathscr{Bor}_{\mathrm{b}}(\mathbb{C}^N)$, cf. Rem. 5.5.9.) Show that (f_α) is a Cauchy net in $L^2(\mathbb{C}^N, \mu)$, that is,
$$\lim_{\alpha, \beta \in \mathscr{I}} \int_{\mathbb{C}^N} |f_\alpha - f_\beta|^2 d\mu = 0$$
4. By the Riesz-Fischer Thm. 1.6.14 (and Thm. 1.2.15), there exists $f \in L^2(\mathbb{C}^N, \mu)$ such that $\lim_\alpha \|f - f_\alpha\|_{L^2(\mathbb{C}^N, \mu)} = 0$. Show that f can be chosen to be a bounded Borel function.
5. Show that $S = f(T_\bullet)$.

Hint. 2. Find a suitable bounded Lipschitz-continuous $\varphi : \mathbb{R} \to \mathbb{R}$ such that $\varphi(x) = x$ whenever $x \in [-\|S\|, \|S\|]$. Use Pb. 5.7 and Thm. 5.8.6 to show that $f_\alpha := \varphi \circ p_\alpha$ satisfies the desired property.

4. Choose a sequence $(\alpha_n)_{n \in \mathbb{Z}_+}$ in $\mathscr{I}$ such that $\lim_n \|f - f_{\alpha_n}\|_{L^2(\mu)} = 0$. Apply the second part of Thm. 1.6.14. □

Theorem 5.11.12. *Assume that $\mathfrak{S}$ is countable. Then there exists $S \in \mathfrak{L}(\mathcal{H})$ such that $S \geqslant 0$ and $W^*(\mathfrak{S}) = W^*(S)$.*

Problem 5.9. In this problem, we prove Thm. 5.11.12.

1. Prove that $W^*(\mathfrak{S})$ contains a countable subset $\mathfrak{E}$ of mutually-commuting projections such that $W^*(\mathfrak{S}) = W^*(\mathfrak{E})$.

2. Suppose that $P, H \in \mathfrak{L}(\mathcal{H})$ commute with each other, P is a projection, and $0 \leqslant H \leqslant 1/2$. Show that $P = \chi_{[1,+\infty)}(P + H)$.

3. Write $\mathfrak{E} = \{E_0, E_1, E_2, \dots\}$. For each $n \in \mathbb{N}$, let
$$S_n = \sum_{k=n}^{+\infty} 3^{-k} E_k$$
Show that the above limit converges in SOT*; in particular, $S_n \in W^*(\mathfrak{S})$.

4. Use Part 2 to prove that $E_n \in W^*(S_n)$. Conclude $E_0, E_1, E_2, \dots \in W^*(S_0)$.

5. Conclude that $W^*(\mathfrak{S}) = W^*(S_0)$.

Hint. 1. Consider $\chi_{(a,b]}(T)$ where $a, b \in \mathbb{Q}$ and $T \in \mathfrak{S}$.

2. View P, H as $\mathbf{M}_p$ and $\mathbf{M}_h$ where $p, h \in \mathscr{B}or_{\mathrm{b}}(\mathbb{C}^2)$. Show that p, h can be chosen so that p is a characteristic function, and $0 \leqslant h \leqslant 1/2$. $\square$

Proof of Thm. 5.11.5. Let $\mathfrak{S} = \{T_1, T_2, \dots\}$. By Thm. 5.11.12, there exists a self-adjoint $S \in \mathfrak{L}(\mathcal{H})$ such that $W^*(\mathfrak{S}) = W^*(S)$. Since each T_i belongs to $W^*(S)$, by Thm. 5.11.11, there exists $f_i \in \mathscr{B}or_{\mathrm{b}}(\mathbb{R})$ such that $T_i = f_i(S)$. Since $T_i = T_i^*$, we have $T_i = \overline{f_i}(S)$, and hence $T_i = (\mathrm{Re} f_i)(S)$. The proof is finished by replacing f_i with $\mathrm{Re} f_i$. $\square$

6 Unbounded operators and their spectral theorem

6.1 Von Neumann and the mathematical foundation of quantum mechanics

In this chapter we study unbounded operators and their spectral theory. Von Neumann was the principal developer of this theory; his chief motivation, beginning in 1927, was to place quantum mechanics on a firm mathematical footing. The aim of this section is to sketch that historical and conceptual background so the reader can appreciate von Neumann's concrete goals and methods.

6.1.1 The mathematical interpretation of quantum mechanics by physicists

The mathematical formulation of quantum mechanics, developed around 1925 by physicists such as Heisenberg, Born, Jordan, and Schrödinger, can be summarized as follows. A particle (for instance, an atom) may exist in different states, each represented by a vector. These vectors are often concrete objects, such as functions on $\mathbb{R}^3$ (i.e., wave functions). Since waves can superpose, quantum states also admit superposition, which mathematically corresponds to vector addition. Thus, the collection of all possible states of a particle forms a complex vector space $\mathcal{V}$. This space is in fact equipped with an inner product, which provides a notion of length. A genuine quantum state is then required to have length 1.

Another fundamental notion is that of observables, i.e., measurable quantities of a quantum system such as energy, momentum, position, and angular momentum. Mathematically, observables are represented by Hermitian operators; in finite dimensions these are simply Hermitian matrices. Actual quantum systems, however, are infinite-dimensional, and in that setting physicists did not initially adopt the same level of rigor as mathematicians in defining what precisely constitutes a "Hermitian operator". The need for such rigor, especially in giving a precise mathematical foundation to the notion of observables, became one of the main starting points for von Neumann's investigation of the spectral theory of unbounded operators, beginning with [vN27].

An observable cannot, in general, be measured with perfect accuracy in all states. If a state (i.e., a vector $\xi \in \mathcal{V}$) allows the observable T to be measured with perfect accuracy and value λ, then mathematically this means $T\xi = \lambda\xi$. Thus, the unit eigenvectors of a Hermitian operator correspond exactly to those states in which the observable can be measured precisely.

In general, at least when $\dim \mathcal{V} < +\infty$, a vector ξ can be written as a finite sum $\sum_i a_i e_i$ where $e_1, e_2, \dots$ are orthonormal eigenvectors of T corresponding to distinct eigenvalues $\lambda_1, \lambda_2, \dots \in \mathbb{R}$, and each $a_i \in \mathbb{C}$. This shows that the value of the observable T measured in state ξ cannot be predicted with certainty; rather, it follows a probability distribution: the probability of measuring the value λ_i

is $|a_i|^2$. By the Pythagorean identity, the probabilities sum to $\sum_i |a_i|^2 = 1$. The expectation value of measuring T in state ξ is $\langle\xi|T\xi\rangle$, since

$$\sum_i |a_i|^2 \lambda_i = \langle\xi|T\xi\rangle$$

For example, the multiplication operator $\mathbf{M}_{x_i}$ (abbreviated to x_i) for the i-th coordinate function of $\mathbb{R}^3$ is the observable measuring the i-th component of the position. Thus, for a function ξ on $\mathbb{R}^3$, the quantity $\langle\xi|x_i\xi\rangle$ gives the expectation value of the i-th component of the position of ξ. Similarly,

$$p_i := \frac{\partial}{\mathbf{i}\partial x_i}$$

is the momentum operator for the i-th component, and hence $\langle\xi|p_i\xi\rangle$ gives the expectation value of the i-th coordinate of the momentum of ξ. The **energy operator** (or **Hamiltonian**) $H = H_t$ typically takes the form

$$H_t = -\Delta + V = p_1^2 + p_2^2 + p_3^2 + V(t,x) \tag{6.1}$$

where V is the potential function defined on $\mathbb{R} \times \mathbb{R}^3$.

The time evolution of a state ξ, denoted $t \in \mathbb{R} \to \xi(t) \in \mathcal{V}$, is governed by the **Schrödinger equation** $\mathbf{i}\partial_t\xi(t) = H\xi(t)$ with initial condition $\xi(0) = \xi$, where H is the Hermitian operator for the energy (so that $\langle\xi|H\xi\rangle$ is the expected energy in state ξ.) In the important case where H is time-independent (e.g. the potential function $V(t,x)$ in (6.1) does not depend on t), the formal solution of the Schrödinger equation is

$$\xi(t) = e^{-\mathbf{i}tH}\xi$$

To compute the RHS explicitly, one may decompose $\xi = \sum_i u_i$ where $u_1, u_2, \dots$ are mutually orthogonal eigenvectors of H with corresponding eigenvalues $\lambda_1, \lambda_2, \dots \in \mathbb{R}$. Then

$$e^{-\mathbf{i}tH}\xi = \sum_i e^{-\mathbf{i}t\lambda_i} u_i$$

Of course, this argument relies on the diagonalization of Hermitian operators, which is valid when $\dim\mathcal{V} < +\infty$. In infinite dimensions, physicists still assume that Hermitian operators admit a kind of diagonalization, though quite different from mathematician's spectral theorems in Ch. 5. Instead, physicists like Dirac assumed that a Hermitian operator T can be "diagonalized" as a combination of a discrete sum and a continuous (Riemann) integral, as described below. (See [Dir30], especially Sec. 10, 16, and 17.)

6.1.2 Physicists' spectral decomposition

Let δ denote the **delta function** on $\mathbb{R}$, i.e., the idealized function satisfying $\delta(x) = 0$ when $x \neq 0$, $\delta(0) = +\infty$, and $\int \delta(x)dx = 1$. Then there exists a family of vectors (u_λ) indexed by real numbers λ, together with a countable family $v_{\lambda_1}, v_{\lambda_2}, \ldots$ for distinct real numbers $\lambda_1, \lambda_2, \ldots$, such that these two collections together form an "orthonormal basis" in the sense that

$$\langle u_\lambda | u_{\lambda'} \rangle = \delta(\lambda - \lambda') \qquad \langle u_\lambda | v_{\lambda_i} \rangle = 0 \qquad \langle v_{\lambda_i} | v_{\lambda_j} \rangle = \delta_{i,j} \tag{6.2}$$

and that for each $\xi, \eta \in \mathcal{V}$, the "Parseval identity"

$$\langle \xi | \eta \rangle = \int \langle \xi | u_\lambda \rangle \langle u_\lambda | \eta \rangle d\lambda + \sum_i \langle \xi | v_{\lambda_i} \rangle \langle v_{\lambda_i} | \eta \rangle$$

holds. These vectors are eigenvectors of T, in the sense that for each reasonable real-valued function f on $\mathbb{R}$, the following spectral decomposition holds:

$$\langle \xi | f(T) \eta \rangle = \int f(\lambda) \langle \xi | u_\lambda \rangle \langle u_\lambda | \eta \rangle d\lambda + \sum_i f(\lambda_i) \langle \xi | v_{\lambda_i} \rangle \langle v_{\lambda_i} | \eta \rangle$$

In particular, when T is the time-independent Hamiltonian H, setting $f(x) = e^{\mathbf{i}tx}$ gives the solution of the Schrödinger equation. Note that the integrals in the above formulas are understood in the classical sense—essentially what mathematicians would call Riemann integrals, rather than the more general Stieltjes integral (which was regarded as too abstract by physicists).

Physicists refer to the vectors u_λ as **scattering states**, since their squared lengths are $\delta(0) = +\infty$, in contrast to the unit vectors v_{λ_i}, which are called **bound states**. For example, when T is the Hamiltonion for the hydrogen atom

$$H = -\Delta - e^2/|x| \qquad e > 0$$

its spectral decomposition contains both continuous and discrete parts (scattering and bound eigenstates), cf. [Dir30, Sec. 39]. As another instance, assuming for simplicity that the space has one dimension (so that the wave functions are defined on $\mathbb{R}$), the eigenstates of the position and momentum operators x and $p = \frac{\partial}{\mathbf{i}\partial x}$ are scattering states: the λ-eigenstate of x is $\delta(x - \lambda)$, while the λ-eigenstate of p is $\frac{1}{\sqrt{2\pi}} e^{\mathbf{i}\lambda x}$.

6.1.3 Von Neumann's notion of abstract Hilbert spaces

Von Neumann rejected the use of vectors of infinite length (i.e., scattering states) such as delta functions in his effort to place quantum mechanics on a rigorous mathematical foundation. He insisted that wave functions must be square-integrable, and hence elements of $L^2(\mathbb{R}^n, m)$. Consequently, he opposed the physicists' practice of interpreting operators in quantum mechanics as continuous (or mixed continuous and discrete) versions of matrices, since that interpretation necessarily relied on "orthonormal bases" of the type in (6.2), which in turn presupposed scattering states.

For von Neumann, the concept that unifies ordinary matrices and operators on function spaces was the Riesz-Fischer theorem. To him, this theorem asserts that for any subset $\Omega \subset \mathbb{R}^n$, the space $L^2(\Omega, m)$ of Lebesgue square-integrable functions is an abstract, separable Hilbert space. This abstract definition of a Hilbert space was a major departure from the early 20th-century view, where Hilbert spaces were specifically understood as $l^2(\mathbb{Z})$. Introducing this abstract notion of Hilbert space was one of the central contributions of [vN27], von Neumann's first paper laying down the mathematical foundations of quantum mechanics.

This version of the Riesz-Fischer theorem did not exist before von Neumann's work. Riesz and Fischer had only proved (in 1907) the completeness of $L^2(I, m)$ where $I \subset \mathbb{R}$ is a compact interval. It was in [vN27] (see Anhang 2) that the completeness of $L^2(\Omega, m)$ for any higher-dimensional set $\Omega \subset \mathbb{R}^n$ was first established.[1] In addition, [vN27] proved that every separable (abstract) Hilbert space admits an orthonormal basis. Taken together, these results imply a unitary equivalence $L^2(\Omega, m) \simeq l^2(\mathbb{Z})$ (when $m(\Omega) > 0$), thus providing the mathematical foundation for unifying discrete matrices with operators on function spaces.

6.1.4 Von Neumann's spectral decomposition for unbounded operators

For von Neumann, another reason for adopting Hilbert spaces as the rigorous framework of quantum mechanics was that, as we have seen in Ch. 5, Hilbert and especially F. Riesz had already established spectral theorems for bounded self-adjoint operators. The only missing piece was that operators arising in quantum mechanics typically involve differential operators. Unlike bounded self-adjoint operators, these are unbounded and are defined only on dense linear subspaces of the Hilbert space.

Despite this, Riesz's spectral Thm. 5.4.4 provided a clear path forward. It states that any bounded self-adjoint operator T on a Hilbert space $\mathcal{H}$ admits a decomposition $T = \int_{[-a,a]} \lambda dE(\lambda)$. As von Neumann recognized in [vN27], this suggests that the spectral decomposition of an unbounded operator representing a quantum observable should have a similar form:

$$T = \int_{-\infty}^{+\infty} \lambda dE(\lambda) \tag{6.3}$$

[1] Unlike the original proofs, which depended heavily on the fundamental theorem of calculus, von Neumann's proof could be extended to general measure spaces and is closer in spirit to modern textbook proofs of the completeness of L^p spaces over abstract measure spaces. See also Subsec. 5.11.2.

where $E = (E(\lambda))_{\lambda \in \mathbb{R}}$ is a right-continuous increasing family of projections satisfying

$$\lim_{\lambda \to -\infty} E(\lambda) = 0 \qquad \lim_{\lambda \to +\infty} E(\lambda) = \mathbf{1}_{\mathcal{H}}$$

with both limits holding in SOT (cf. Thm. 5.2.7). More precisely, for each $\xi \in \mathcal{H}$, one defines

$$T\xi = \int_{-\infty}^{+\infty} \lambda dE(\lambda)\xi = \lim_{a \to +\infty} \int_{[-a,a]} \lambda dE(\lambda)\xi$$

whenever the limit exists; the domain of T is defined to be the set of all $\xi \in \mathcal{H}$ such that this limit exists.

For modern readers familiar with the multiplication-operator version of the spectral theorem (Thm. 5.10.22), the most suitable version for unbounded operators is that any such operator T is unitarily equivalent to a multiplication operators $\mathbf{M}_f$ on a direct sum of L^2-spaces, where f is a real-valued Borel function that is not necessarily bounded. (See Exp. 6.2.16 for the precise definition.)

6.1.5 Measurement with absolute precision vs. measurement with prescribed accuracy

Von Neumann's spectral decomposition for observables abandons the notion of scattering states and the idea that one can measure a quantity T with absolute precision, as is the case with bound states. The mathematically rigorous viewpoint is instead that, for $\lambda \in \mathbb{R}$ lying in the continuous spectrum, the quantity T can only be measured to within an arbitrarily prescribed accuracy.

To see how this rigorous viewpoint arises, one must understand von Neumann's interpretation of projections in [vN27], later expanded and refined in [vN32a].[2] A projection E on $\mathcal{H}$ represents a "proposition" (or "event") in a quantum system, i.e., an observable that can take only the values 0 and 1, interpreted as "false" and "true", respectively. Thus, $\mathrm{Rng}(E)$ is the subspace of states in which the proposition holds, and $\mathrm{Ker}(E)$ is the subspace of states in which it fails. It follows that for any unit vector $\xi \in \mathcal{H}$, the value $\langle \xi | E\xi \rangle$ gives the probability that the proposition is true in the state ξ.

Applying this interpretation to the spectral projections $E(\lambda) = \chi_{(-\infty,\lambda]}(T)$ in (6.3), one finds that for any unit vector ξ, the value $\langle \xi | E(\lambda)\xi \rangle$ is the probability that the observable T takes a value not exceeding λ in the state ξ. Therefore, if we define the projection

$$E(\lambda, \lambda') = E(\lambda') - E(\lambda)$$

[2] [vN32a] may be regarded as a greatly expanded version of [vN27]. There, von Neumann develops and clarifies his earlier treatment of the mathematical foundations of quantum mechanics.

for $\lambda \leqslant \lambda'$, then $\langle \xi | E(\lambda, \lambda')\xi \rangle$ is the probability that a measurement of T in the state ξ yields a value in the interval $(\lambda, \lambda']$. Accordingly, the unit vectors in the range of $E(\lambda, \lambda')$ correspond precisely to the states in which the measurement outcomes of T always lie within $(\lambda, \lambda']$.

As in Def. 5.8.1, a number $\lambda \in \mathbb{R}$ belongs to the spectrum of T precisely when $E(\lambda - \varepsilon, \lambda + \varepsilon) \neq 0$ for all $\varepsilon > 0$. For such λ, one can always choose a unit vector ξ in $\mathrm{Rng}(E(\lambda - \varepsilon, \lambda + \varepsilon))$ so that $\langle \xi | E(\lambda - \varepsilon, \lambda + \varepsilon)\xi \rangle = 1$. Therefore, in the state ξ, a measurement of T is not guaranteed to yield exactly λ, but must lie within the interval $(\lambda - \varepsilon, \lambda + \varepsilon]$. This is the precise meaning of **measurement with arbitrarily prescribed accuracy**.

6.1.6 Conclusion

We can now clarify the true starting point of von Neumann's study of unbounded operators. A common misconception is that his project began with a rigorous definition of unbounded self-adjoint operators, with the goal of proving a spectral theorem for them. In fact, such a rigorous definition was not available at the outset—the notion of the adjoint for unbounded operators is far subtler than in the bounded case. Rather, von Neumann's aim was to find a natural definition of observables under which the spectral decomposition of the form (6.3) would hold.

Following von Neumann's line of thought, we introduce in Sec. 6.3 the first such definition: the Hermitian operators T on $\mathcal{H}$ satisfying

$$\mathrm{Rng}(T + \mathbf{i}) = \mathrm{Rng}(T - \mathbf{i}) = \mathcal{H} \tag{6.4}$$

(See the beginning of Subsec. 6.3.3.) This is not yet the final definition of unbounded self-adjoint operators; indeed, in modern textbooks it often appears as a criterion for self-adjointness rather than a definition. Nevertheless, we will see that verifying condition (6.4) is frequently more convenient than checking self-adjointness directly,[3] and that many central ideas in the theory of unbounded operators—such as the Cayley transform and the concept of closed operators—are more naturally understood in terms of this condition.

6.2 Basic notions about unbounded operators

Let $\mathcal{H}, \mathcal{K}, \mathcal{L}, \mathcal{M}$ be Hilbert spaces.

6.2.1 Definitions and basic properties

Definition 6.2.1. An **unbounded operator** T from $\mathcal{H}$ to $\mathcal{K}$ (abbreviated to $T : \mathcal{H} \to \mathcal{K}$) is defined to be a linear map from a linear subspace $\mathscr{D}(T)$ of $\mathcal{H}$ to $\mathcal{K}$. We call

[3] See the proofs of Thm. 7.6.2 (which relies on Lem. 7.6.1), Stone's Thm. 7.3.3, and the Kato-Rellich Thm. 7.9.6.

$\mathscr{D}(T)$ the **domain** of T. If $\mathcal{H} = \mathcal{K}$, we say that T is an **unbounded operator on $\mathcal{H}$**. If $\mathscr{D}(T) = \mathcal{H}$, we say that T is everywhere-defined.

Convention 6.2.2. Unless otherwise stated, we assume that the domain $\mathscr{D}(T)$ of an unbounded operator $T : \mathcal{H} \to \mathcal{K}$ is dense. If this assumption is dropped, we will say that T is an **n.d.d. unbounded operator**, where "n.d.d." stands for "non-necessarily densely defined".

Note that every bounded linear operator is an unbounded operator in this sense. This is not contradictory: "unbounded" here simply means "not necessarily bounded".

Definition 6.2.3. Let $T : \mathcal{H} \to \mathcal{K}$ be an n.d.d. unbounded operator. If $\mathscr{D}_0$ is a linear subspace of $\mathscr{D}(T)$, we define the **restriction**

$$\boldsymbol{T|_{\mathscr{D}_0}} : \mathcal{H} \to \mathcal{K} \qquad \xi \mapsto T\xi$$

with $\mathscr{D}(T|_{\mathscr{D}_0}) = \mathscr{D}_0$.

Definition 6.2.4. Let $S, T : \mathcal{H} \to \mathcal{K}$ be n.d.d. unbounded operators. The addition $\boldsymbol{S + T}$ is defined to be

$$S + T : \mathscr{D}(S) \cap \mathscr{D}(T) \to \mathcal{K} \qquad \xi \mapsto S\xi + T\xi$$

which is an unbounded operator $\mathcal{H} \to \mathcal{K}$ whenever $\mathscr{D}(S + T) := \mathscr{D}(S) \cap \mathscr{D}(T)$ is dense in $\mathcal{H}$.

Definition 6.2.5. Let $S : \mathcal{K} \to \mathcal{L}$ and $T : \mathcal{H} \to \mathcal{K}$ be n.d.d. unbounded operators. The composition $\boldsymbol{ST}$ is defined to be

$$ST : \mathscr{D}(ST) \to \mathcal{L} \qquad \xi \mapsto ST\xi$$

where $\mathscr{D}(ST)$ is the space of all $\xi \in \mathscr{D}(T)$ satisfying $T\xi \in \mathscr{D}(S)$. Then ST is an unbounded operator $\mathcal{H} \to \mathcal{L}$ whenever $\mathscr{D}(ST)$ is dense in $\mathcal{H}$. If $\lambda \in \mathbb{C}$, we define the scalar product $\boldsymbol{\lambda T}$ (with domain $\mathscr{D}(\lambda T) := \mathscr{D}(T)$) by

$$\lambda T : \mathscr{D}(T) \to \mathcal{K} \qquad \xi \mapsto \lambda \cdot T\xi$$

Proposition 6.2.6. *Let A, B, C be n.d.d. unbounded operators $\mathcal{H} \to \mathcal{K}$. Then $(A + B) + C = A + (B + C)$.*

It is therefore legitimate to write $A + B + C$.

Proof. One checks that both sides have domain $\mathscr{D}(A) \cap \mathscr{D}(B) \cap \mathscr{D}(C)$, and that these two maps agree on this common domain. □

Proposition 6.2.7. *Let $A : \mathcal{L} \to \mathcal{M}$, $B : \mathcal{K} \to \mathcal{L}$, and $C : \mathcal{H} \to \mathcal{K}$ be n.d.d. unbounded operators. Then $(AB)C = A(BC)$.*

Due to this proposition, one can write ABC unambiguously.

Proof. The only non-trivial part is the compare their domains:

$$\mathscr{D}((AB)C) = \mathscr{D}(A(BC)) = \{\xi \in \mathscr{D}(C) : C\xi \in \mathscr{D}(B), BC\xi \in \mathscr{D}(A)\}$$

□

Proposition 6.2.8. *Let $A, B : \mathcal{K} \to \mathcal{L}$ and $C, D : \mathcal{H} \to \mathcal{K}$ be n.d.d. unbounded operators. Then*

$$(A+B)C = AC + BC \qquad A(C+D) \supset AC + AD \tag{6.5}$$

where the "$\supset$" in the second relation becomes "$=$" if A is everywhere-defined.

Proof. The only non-trivial part is to compare their domains:

$$\begin{gathered}\mathscr{D}((A+B)C) = \mathscr{D}(AC+BC) = \{\xi \in \mathscr{D}(C) : C\xi \in \mathscr{D}(A) \cap \mathscr{D}(B)\}\\ \mathscr{D}(A(C+D)) = \{\xi \in \mathscr{D}(C) \cap \mathscr{D}(D) : C\xi + D\xi \in \mathscr{D}(A)\}\\ \mathscr{D}(AC+AD) = \{\xi \in \mathscr{D}(C) \cap \mathscr{D}(D) : C\xi \in \mathscr{D}(A), D\xi \in \mathscr{D}(A)\}\end{gathered}$$

□

Example 6.2.9. Let A be an unbounded operator on $\mathcal{H}$ with $\mathscr{D}(A) \subsetneq \mathcal{H}$. Let $C = \mathbf{1}_{\mathcal{H}}$ and $D = -\mathbf{1}_{\mathcal{H}}$. Then $A(C+D) = A0 = 0$ has domain $\mathscr{D}(H)$, while $AC + AD = A + (-A)$ has domain $\mathscr{D}(A)$. Thus $A(C+D) \supsetneq AC + AD$.

Definition 6.2.10. Let S, T be n.d.d. unbounded operators $\mathcal{H} \to \mathcal{K}$. We say that T is an **extension** of S and write $\boldsymbol{S \subset T}$ if $\mathscr{D}(S) \subset \mathscr{D}(T)$, and if $S\xi = T\xi$ for each $\xi \in \mathscr{D}(S)$.

Definition 6.2.11. An unbounded operator T on $\mathcal{H}$ is called a **Hermitian operator** (or a **symmetric operator**) if for each $\xi, \eta \in \mathscr{D}(T)$ we have

$$\langle \eta | T\xi \rangle = \langle T\eta | \xi \rangle$$

In other words, T is Hermitian iff the sesquilinear form $\boldsymbol{\omega_T}$ associated to T, as defined by

$$\omega_T : \mathscr{D}(T) \times \mathscr{D}(T) \to \mathbb{C} \qquad (\eta, \xi) \mapsto \langle \eta | T\xi \rangle$$

is Hermitian.

6.2.2 Elementary examples of Hermitian operators

Example 6.2.12. Let $\Omega \subset \mathbb{R}^N$ be open. By Thm. 1.7.11, the space $C_c^\infty(\Omega)$ is dense in the Hilbert space $\mathcal{H} := L^2(\Omega, m)$. For each $1 \leqslant j \leqslant N$, let $\mathbf{p}_i$ be the unbounded operator on $\mathcal{H}$ defined by

$$\mathbf{p}_j = \frac{\partial}{\mathbf{i}\partial x_j} \qquad \mathscr{D}(\mathbf{p}_j) = C_c^\infty(\Omega)$$

Then $\mathbf{p}_j$ is Hermitian.

Proof. We consider for simplicity the case $j = 1$. For each $f, g \in C_c^\infty(\Omega)$, using integration by parts, we compute

$$\int_{\mathbb{R}} \overline{f}\partial_1 g dx_1 = \overline{f}g\Big|_{x_1=-\infty}^{+\infty} - \int_{\mathbb{R}} \overline{\partial_1 f} g dx_1 = -\int_{\mathbb{R}} \overline{\partial_1 f} g dx_1$$

as functions of $x_2, \dots, x_N$. Hence

$$\begin{aligned}\langle f|\mathbf{p}_1 g\rangle &= \int_{\mathbb{R}^{N-1}} \int_{\mathbb{R}} -\mathbf{i}\overline{f}\partial_1 g dx_1 \cdot dx_2 \cdots dx_N \\ &= \int_{\mathbb{R}^{N-1}} \int_{\mathbb{R}} \overline{-\mathbf{i}\partial_1 f} g dx_1 \cdot dx_2 \cdots dx_N = \langle \mathbf{p}_1 f|g\rangle\end{aligned}$$

□

Example 6.2.13. We work in the setting of Exp. 6.2.12. Let f be a real polynomial with N mutually-commuting variables, i.e., $f \in \mathbb{R}[x_1, \dots, x_N]$. Define the **polynomial functional calculus** $f(\mathbf{p}_\bullet) = f(\mathbf{p}_1, \dots, \mathbf{p}_N)$ in the obvious way, e.g., if $f = x_1 x_3^2 - 2x_4^3 x_5$ then $f(\mathbf{p}_\bullet) = \mathbf{p}_1\mathbf{p}_3^2 - 2\mathbf{p}_4^3\mathbf{p}_5$. According to Def. 6.2.4 and 6.2.5, $f(\mathbf{p}_\bullet)$ has domain $C_c^\infty(\Omega)$. Using the fact that each $\mathbf{p}_j$ is Hermitian, one easily checks that $f(\mathbf{p}_\bullet)$ is Hermitian. In particular, the **Laplacian**

$$\begin{gathered}\Delta = -\mathbf{p}_1^2 - \cdots - \mathbf{p}_N^2 = \partial_{x_1}^2 + \cdots + \partial_{x_N}^2 \\ \mathscr{D}(\Delta) = C_c^\infty(\Omega)\end{gathered}$$

is Hermitian. In this course, unless other stated, we always understand that Δ has domain $C_c^\infty(\Omega)$.

Definition 6.2.14. Let $(X, \mathfrak{M}, \mu)$ be a measure space. Let $f : X \to \mathbb{C}$ be measurable. The **multiplication operator $\mathbf{M}_f$** is an unbounded operator on $L^2(X, \mu)$ defined by

$$\begin{gathered}\mathbf{M}_f : \mathscr{D}(\mathbf{M}_f) \to L^2(X, \mu) \qquad \xi \mapsto f\xi \\ \mathscr{D}(\mathbf{M}_f) = \{\xi \in L^2(X, \mu) : f\xi \in L^2(X, \mu)\}\end{gathered}$$

If f is real-valued, then $\mathbf{M}_f$ is clearly Hermitian.

Proof. We need to prove that $\mathscr{D}(\mathbf{M}_f)$ is dense. Let

$$V_n = \{\xi \in L^2(X,\mu) : \xi \text{ vanishes outside } \{x \in X : |f(x)| \leqslant n\}\}$$

Then $\sum_{n\in\mathbb{Z}_+} V_n$ is dense in $L^2(X,\mu)$ (by DCT), and each V_n is contained in $\mathscr{D}(\mathbf{M}_f)$. Thus $\mathscr{D}(\mathbf{M}_f)$ is dense. $\square$

Definition 6.2.15. Let $(\mathcal{H}_\alpha)_{\alpha\in\mathscr{I}}$ and $(\mathcal{K}_\alpha)_{\alpha\in\mathscr{I}}$ be collections of Hilbert spaces. For each $\alpha \in \mathscr{I}$, let $T_\alpha : \mathcal{H}_\alpha \to \mathcal{K}_\alpha$ be an n.d.d. unbounded operators. Their **direct sum $\oplus_{\alpha\in\mathscr{I}} T_\alpha$** is an n.d.d. unbounded operator $\bigoplus_{\alpha\in\mathscr{I}} \mathcal{H}_\alpha \to \bigoplus_{\alpha\in\mathscr{I}} \mathcal{K}_\alpha$ defined by

$$\oplus_{\alpha\in\mathscr{I}} T_\alpha : \mathscr{D}(\oplus_{\alpha\in\mathscr{I}} T_\alpha) \to \bigoplus_{\alpha\in\mathscr{I}} \mathcal{K}_\alpha \qquad \oplus_{\alpha\in\mathscr{I}}\, \xi_\alpha \mapsto \oplus_{\alpha\in\mathscr{I}} T_\alpha \xi_\alpha$$

where

$$\mathscr{D}(\oplus_{\alpha\in\mathscr{I}} T_\alpha) = \Big\{ \oplus_\alpha\, \xi_\alpha \in \bigoplus_{\alpha\in\mathscr{I}} \mathcal{H}_\alpha : \ \xi_\alpha \in \mathscr{D}(T_\alpha) \text{ for all } \alpha \in \mathscr{I} \text{ and } \sum_{\alpha\in\mathscr{I}} \|T_\alpha \xi_\alpha\|^2 < +\infty \Big\}$$

Example 6.2.16. Let $(X, \mathfrak{M})$ be a measurable space. Let $(\mu_\alpha)_{\alpha\in\mathscr{I}}$ be a family of measures on $\mathfrak{M}$. Let $f : X \to \mathbb{C}$ be measurable. Then $\oplus_{\alpha\in\mathscr{I}} \mathbf{M}_f$ is an unbounded operator on $\bigoplus_{\alpha\in\mathscr{I}} L^2(X,\mu_\alpha)$ with domain

$$\mathscr{D}(\oplus_{\alpha\in\mathscr{I}} \mathbf{M}_f) = \Big\{ \oplus_\alpha\, \xi_\alpha \in \bigoplus_{\alpha\in\mathscr{I}} \mathcal{H}_\alpha : \oplus_\alpha f \xi_\alpha \in \bigoplus_{\alpha\in\mathscr{I}} \mathcal{H}_\alpha \Big\}$$

In other words, $\mathscr{D}(\oplus_{\alpha\in\mathscr{I}} \mathbf{M}_f)$ consists precisely of those elements $(\xi_\alpha)_{\alpha\in\mathscr{I}}$ in $\prod_{\alpha\in\mathscr{I}} \mathcal{H}_\alpha$ satisfying

$$\sum_{\alpha\in\mathscr{I}} \|\xi_\alpha\|^2_{L^2(\mu_\alpha)} < +\infty \qquad \sum_{\alpha\in\mathscr{I}} \|f\xi_\alpha\|^2_{L^2(\mu_\alpha)} < +\infty$$

When no confusion arises, we

$$\boxed{\text{abbreviate } \oplus_\alpha \mathbf{M}_f \text{ to } \mathbf{M}_f} \tag{6.6}$$

and also call $\mathbf{M}_f$ the **multiplication operator**. If f is real-valued, then $\mathbf{M}_f$ is clearly Hermitian.

Proof. Similar to Def. 6.2.14, we need to check that $\mathbf{M}_f$ has dense domain in $\mathcal{H} := \bigoplus_{\alpha\in\mathscr{I}} L^2(X,\mu_\alpha)$. Let V be the set of all $\oplus_\alpha \xi_\alpha$ in $\mathcal{H}$ satisfying the following conditions:

- There exists a finite set $I \subset \mathscr{I}$ such that $\xi_\alpha = 0$ for all $\alpha \notin I$.

- There exists $n \in \mathbb{Z}_+$ such that for each $\alpha \in I$, the function ξ_α vanishes outside $\{x \in X : |f(x)| \leqslant n\}$.

Then V is clearly a dense linear subspace of $\mathcal{H}$ contained in $\mathscr{D}(\mathbf{M}_f)$. $\square$

Remark 6.2.17. In Exp. 6.2.16, if $f, g : X \to \mathbb{C}$ are both measurable, then

$$\mathbf{M}_f \mathbf{M}_g \subset \mathbf{M}_{fg}$$

Moreover, if g is bounded (or more generally, if $\|g\|_{L^\infty(X,\mu_\alpha)} < +\infty$ for each α), then

$$\mathbf{M}_f \mathbf{M}_g = \mathbf{M}_{fg}$$

Proof. This follows immediately from the observation that $\mathscr{D}(\mathbf{M}_f \mathbf{M}_g)$ consists of vectors $(\xi_\alpha)_\alpha \in \prod_\alpha L^2(X, \mu_\alpha)$ satisfying

$$\sum_\alpha \|\xi_\alpha\|^2_{L^2(\mu_\alpha)} < +\infty \qquad \sum_\alpha \|g\xi_\alpha\|^2_{L^2(\mu_\alpha)} < +\infty \qquad \sum_\alpha \|fg\xi_\alpha\|^2_{L^2(\mu_\alpha)} < +\infty \tag{6.7a}$$

whereas $\mathscr{D}(\mathbf{M}_{fg})$ consists those satisfying

$$\sum_\alpha \|\xi_\alpha\|^2_{L^2(\mu_\alpha)} < +\infty \qquad \sum_\alpha \|fg\xi_\alpha\|^2_{L^2(\mu_\alpha)} < +\infty \tag{6.7b}$$

$\square$

Exercise 6.2.18. In the setting of Exp. 6.2.16, prove for each measurable function $f : X \to \mathbb{C}$ and each $n \in \mathbb{N}$ that

$$(\mathbf{M}_f)^n = \mathbf{M}_{f^n}$$

(See Pb. 6.3 for a related comment.)

Exercise 6.2.19. In the setting of Exp. 6.2.16, let $f_1, \dots, f_N : X \to \mathbb{C}$ be measurable functions. Let $\mathscr{D}_0$ be a linear subspace of $\mathcal{H}$ such that $\mathbf{M}_{f_j}\mathscr{D}_0 \subset \mathscr{D}_0$ for each j. Let $p \in \mathbb{C}[z_1, \dots, z_N]$. Prove that $\mathscr{D}_0 \subset \mathscr{D}(\mathbf{M}_{p\circ f_\bullet})$ and

$$p(\mathbf{M}_{f_1}, \dots, \mathbf{M}_{f_N})\big|_{\mathscr{D}_0} = \mathbf{M}_{p\circ f_\bullet}\big|_{\mathscr{D}_0}$$

6.2.3 Inverses of unbounded operators

Definition 6.2.20. Let $T : \mathcal{H} \to \mathcal{K}$ be an n.d.d. unbounded operator. Assume that T is injective, equivalently, that the kernel

$$\mathbf{Ker}(\boldsymbol{T}) := \{\xi \in \mathscr{D}(T) : T\xi = 0\}$$

is zero. The **inverse** of T is the n.d.d. unbounded operator $\mathcal{K} \to \mathcal{H}$ defined by

$$\boldsymbol{T^{-1}} : \mathrm{Rng}(T) \to \mathcal{H} \qquad T\xi \mapsto \xi$$

with $\mathscr{D}(T^{-1}) = \mathrm{Rng}(T)$. It is clear that T^{-1} is injective and $(T^{-1})^{-1} = T$. Note that T has dense range iff T^{-1} is a (densely-defined) unbounded operator.

It is useful to keep in mind that

$$\mathscr{D}(T^{-1}) = \mathrm{Rng}(T) \qquad \mathscr{D}(T) = \mathrm{Rng}(T^{-1}) \tag{6.8}$$

Proposition 6.2.21. *Let $A : \mathcal{K} \to \mathcal{L}$ and $B : \mathcal{H} \to \mathcal{K}$ be injective n.d.d. unbounded operators. Note that AB is clearly injective. We have*

$$(AB)^{-1} = B^{-1}A^{-1}$$

Proof. The only non-trivial part is to compare their domains. The domain of $(AB)^{-1}$ is $\mathrm{Rng}(AB)$. An element $\psi \in \mathcal{L}$ belongs to $\mathscr{D}(B^{-1}A^{-1})$ iff $\psi \in \mathscr{D}(A^{-1}) = \mathrm{Rng}(A)$ and $A^{-1}\psi \in \mathscr{D}(B^{-1}) = \mathrm{Rng}(B)$, iff $\psi = A\eta$ and $\eta = B\xi$ for some $\eta \in \mathscr{D}(A)$ and $\xi \in \mathscr{D}(B)$, iff $\psi \in \mathrm{Rng}(AB)$. □

Example 6.2.22. Let $(X, \mathfrak{M})$ be a measurable space. Let $(\mu_\alpha)_{\alpha\in\mathscr{I}}$ be a family of measures on $\mathfrak{M}$. Let $f : X \to \mathbb{C}$ be measurable. Let $\mathcal{H} = \bigoplus_{\alpha\in\mathscr{I}} L^2(X, \mu_\alpha)$. Then

$$\mathrm{Ker}(\mathbf{M}_f) = \Big\{ \oplus_\alpha \xi_\alpha \in \mathcal{H} : \text{each } \xi_\alpha \text{ is } \mu_\alpha\text{-a.e. zero outside } f^{-1}(0) \Big\} \tag{6.9}$$

Equivalently, $\mathrm{Ker}(\mathbf{M}_f)$ equals the range of $\mathbf{M}_{\chi_{f^{-1}(0)}}$. Consequently, $\mathbf{M}_f$ is injective iff $f^{-1}(0)$ is μ_α-null for each α.

It follows that $\mathbf{M}_f$ is injective iff there exists a measurable $g : X \to \mathbb{C}$ such that $fg = 1$ a.e. in μ_α for each $\alpha \in \mathscr{I}$.[4]

Proof of (6.9). An element $\oplus_\alpha\xi_\alpha \in \mathcal{H}$ belongs to $\mathrm{Ker}(\mathbf{M}_f)$ iff $f\xi_\alpha = 0$ in $L^2(X, \mu_\alpha)$ for each α, iff $f\xi_\alpha = 0$ a.e. in μ_α, iff $\oplus_\alpha\xi_\alpha$ belongs to the RHS of (6.9). □

Example 6.2.23. Let $(X, \mathfrak{M})$ be a measurable space. Let $(\mu_\alpha)_{\alpha\in\mathscr{I}}$ be a family of measures on $\mathfrak{M}$. Suppose that $f, g : X \to \mathbb{C}$ are measurable functions satisfying $fg = 1$ a.e. in μ_α for each $\alpha \in \mathscr{I}$. Then on $\mathcal{H} := \bigoplus_{\alpha\in\mathscr{I}} L^2(X, \mu_\alpha)$, both $\mathbf{M}_f$ and $\mathbf{M}_g$ are injective and have dense ranges. Moreover, we have $\mathbf{M}_g = \mathbf{M}_f^{-1}$, and hence $\mathbf{M}_f = \mathbf{M}_g^{-1}$.

Proof. We have

$$\begin{aligned}
\mathscr{D}(\mathbf{M}_f) &= \{\oplus_\alpha\xi_\alpha \in \mathcal{H} : \oplus_\alpha f\xi_\alpha \in \mathcal{H}\}\\
\mathrm{Rng}(\mathbf{M}_f) &= \{\oplus_\alpha\eta_\alpha \in \mathcal{H} : \oplus_\alpha\eta_\alpha = \oplus_\alpha f\xi_\alpha \text{ for some } \oplus_\alpha \xi_\alpha \in \mathcal{H}\}\\
\mathscr{D}(\mathbf{M}_g) &= \{\oplus_\alpha\eta_\alpha \in \mathcal{H} : \oplus_\alpha g\eta_\alpha \in \mathcal{H}\}\\
\mathrm{Rng}(\mathbf{M}_g) &= \{\oplus_\alpha\xi_\alpha \in \mathcal{H} : \oplus_\alpha\xi_\alpha = \oplus_\alpha g\eta_\alpha \text{ for some } \oplus_\alpha \eta_\alpha \in \mathcal{H}\}
\end{aligned}$$

It follows immediately that

$$\mathscr{D}(\mathbf{M}_g) = \mathrm{Rng}(\mathbf{M}_f) \qquad \mathrm{Rng}(\mathbf{M}_g) = \mathscr{D}(\mathbf{M}_f)$$

that $\mathbf{M}_f$ sends $\oplus_\alpha\xi_\alpha$ bijectively to $\oplus_\alpha\eta_\alpha$, and that $\mathbf{M}_g$ sends $\oplus_\alpha\eta_\alpha$ bijectively to $\oplus_\alpha\xi_\alpha$. Thus $\mathbf{M}_f : \mathscr{D}(\mathbf{M}_f) \to \mathrm{Rng}(\mathbf{M}_f)$ and $\mathbf{M}_g : \mathrm{Rng}(\mathbf{M}_f) \to \mathscr{D}(\mathbf{M}_f)$ are inverses of each other. In particular, $\mathbf{M}_g$ and $\mathbf{M}_f$ are injective with dense range. □

[4] The direction "⇐" follows from $f^{-1}(0) \subset \{x \in X : f(x)g(x) \neq 1\}$. The direction "⇒" follows by defining g to be $g(x) = 1/f(x)$ when $f(x) \neq 0$ and $g(x) = 0$ when $f(x) = 0$.

6.2.4 Inverses of bounded operators

Exp. 6.2.23 shows that $\oplus_{\alpha\in\mathscr{I}}\mathbf{M}_f$ has dense range if it is injective. In the case where f is bounded (and hence $\oplus_{\alpha\in\mathscr{I}}\mathbf{M}_f$ is a bounded normal operator), this phenomenon can be understood in a more general setting.

Definition 6.2.24. We say that an unbounded operator $T:\mathcal{H}\to\mathcal{K}$ has an **everywhere-defined bounded inverse** if $T=S^{-1}$ for some injective $S\in\mathfrak{L}(\mathcal{K},\mathcal{H})$ with dense range. (Note that such S must be unique, since $S=T^{-1}$.) Equivalently, T is injective, $\mathrm{Rng}(T)=\mathcal{K}$, and the inverse linear map $T^{-1}:\mathcal{K}\to\mathscr{D}(T)\subset\mathcal{H}$ is bounded. We adopt the notation:

$$\boldsymbol{T^{-1}\in\mathfrak{L}(\mathcal{K},\mathcal{H})}\text{ means that }T\text{ has an everywhere-defined bounded inverse}\qquad(6.10)$$

For bounded linear operators, injectivity and dense range are closely related:

Proposition 6.2.25. *Let $T\in\mathfrak{L}(\mathcal{H},\mathcal{K})$. Then*

$$\mathrm{Ker}(T)=\mathrm{Rng}(T^*)^\perp\qquad\overline{\mathrm{Rng}(T)}=\mathrm{Ker}(T^*)^\perp$$

It follows from Cor. 3.4.9 that T has dense range iff $\mathrm{Ker}(T^*)$ is injective.

Proof. For each $\xi\in\mathcal{H}$, we have $\xi\in\mathrm{Ker}(T)$ iff $T\xi=0$ iff $\langle\mathcal{K}|T\xi\rangle=0$ iff $\langle T^*\mathcal{K}|\xi\rangle=0$ iff $\xi\in\mathrm{Rng}(T^*)^\perp$. This proves the first identity. Replacing T with T^*, we obtain $\mathrm{Ker}(T^*)=\mathrm{Rng}(T)^\perp$, and hence

$$\overline{\mathrm{Rng}(T)}=\mathrm{Rng}(T)^{\perp\perp}=\mathrm{Ker}(T^*)^\perp$$

due to Cor. 3.4.8. □

Remark 6.2.26. Suppose that $T\in\mathfrak{L}(\mathcal{H})$ is normal. Then

$$\mathrm{Ker}(T)=\mathrm{Ker}(T^*)$$

since $\|T\xi\|^2=\|T^*\xi\|^2$ for each $\xi\in\mathcal{H}$. It follows from Prop. 6.2.25 that T is injective iff T has dense range.

6.3 Extensions of Hermitian operators I: reduction to extensions of unitary maps

Let $\mathcal{H}$ be a Hilbert space.

6.3.1 Introduction

In this section we follow von Neumann's idea in [vN29a] to investigate which Hermitian operators admit spectral decompositions. As noted in Subsec. 6.1.4, a spectral decomposition for a Hermitian operator T means that it can be written in the form $\int_{-\infty}^{+\infty} \lambda dE(\lambda)$. We will take a more modern, beginner-friendly viewpoint: we say that T has a spectral decomposition if it is unitarily equivalent to a multiplication operator on a direct sum of L^2-spaces (cf. Exp. 6.2.16).

Von Neumann's key idea was to use the Cayley transform to connect Hermitian operators and unitary operators. The Cayley transform itself goes back to Cayley, who gave a systematic way to construct (generic) real orthogonal matrices from skew-symmetric matrices via the formula

$$U = \frac{\mathbf{1} + A}{\mathbf{1} - A}$$

where A is skew-symmetric (so $\mathbf{1} - A$ is invertible). Here "generic" means that 1 is not an eigenvalue of U.

For von Neumann, the Cayley transform and its inverse

$$U = \frac{T - \mathbf{i}}{T + \mathbf{i}} \qquad T = \frac{U + \mathbf{1}}{\mathbf{i}(U - \mathbf{1})}$$

connect "generic" unitary operators U with Hermitian operators T that admit spectral decompositions. The meaning of "generic" becomes clear if we take T to be a multilication operator $\mathbf{M}_f$ where $f : X \to \mathbb{R}$ is measurable. Then U is expected to correspond to $\mathbf{M}_u$ where $u = (f + \mathbf{i})/(f - \mathbf{i})$. One can easily check that u takes values in $\mathbb{S}^1 \backslash \{1\}$, and conversely any measurable function $u : X \to \mathbb{S}^1 \backslash \{1\}$ arises in this way from some measurable $f : X \to \mathbb{R}$. In light of Exp. 6.2.22, it is natural to guess that "generic" here should mean that the unitary operator U satisfies $\mathrm{Ker}(U - \mathbf{1}) = \{0\}$.

Thus, a preliminary answer to the question "which Hermitian operators admit spectral decompositions?" is: those Hermitian operators that, via the Cayley transform (to be defined rigorously), correspond to unitary operators U with $\mathrm{Ker}(U - \mathbf{1}) = \{0\}$. As we will see in the following subsections, it is more convenient to work with the assumption that $\mathrm{Rng}(U - \mathbf{1})$ is dense in $\mathcal{H}$, though this condition is equivalent to $\mathrm{Ker}(U - \mathbf{1}) = \{0\}$ (for unitary operators $U \in \mathfrak{L}(\mathcal{H})$) due to Rem. 6.2.26.

In what follows we will make precise sense of these transforms. In doing so, we will see that the Cayley transform is well defined for any Hermitian operator, though the result need not be unitary on the whole space $\mathcal{H}$. This more general perspective will help us address the following deeper question: if a Hermitian operator does not admit a spectral decomposition, can it be extended to one that does? And if so, how can such extensions be characterized?

6.3.2 The Cayley transform

Lemma 6.3.1. *Let T be a Hermitian operator on $\mathcal{H}$. Then all eigenvalues of T are real numbers.*

Proof. Suppose that $a + b\mathbf{i}$ is an eigenvalue of T, where $a, b \in \mathbb{R}$ and $b \neq 0$. Then there exists a non-zero $\xi \in \mathrm{Ker}(T - a - b\mathbf{i})$, and hence hence $\langle \xi|(T-a)\xi\rangle = b\mathbf{i}\langle \xi|\xi\rangle$. This is impossible, since $\langle \xi|(T-a)\xi\rangle \in \mathbb{R}$ and $\langle \xi|\xi\rangle \in \mathbb{R}_{>0}$. $\square$

According to Lem. 6.3.1, we have $\mathrm{Ker}(T-\mathbf{i}) = \mathrm{Ker}(T+\mathbf{i}) = \{0\}$. Therefore, by Def. 6.2.20, $(T \pm \mathbf{i})^{-1}$ can be defined as an n.d.d. unbounded operators on $\mathcal{H}$. Moreover, since $\mathrm{Rng}(T+\mathbf{i})^{-1} = \mathscr{D}(T+\mathbf{i}) = \mathscr{D}(T) = \mathscr{D}(T-\mathbf{i})$, by Def. 6.2.5, the domain of $(T-\mathbf{i})(T+\mathbf{i})^{-1}$ is $\mathrm{Rng}(T+\mathbf{i})$.

Definition 6.3.2. Let T be a Hermitian operator on $\mathcal{H}$. The **Cayley transform** of T is the n.d.d. unbounded operator on $\mathcal{H}$ defined by

$$(T-\mathbf{i})(T+\mathbf{i})^{-1} : \mathrm{Rng}(T+\mathbf{i}) \to \mathcal{H}$$

and commonly denoted by $\frac{T-\mathbf{i}}{T+\mathbf{i}}$. In other words, the Cayley transform of T is the unique linear map satisfying

$$\frac{T-\mathbf{i}}{T+\mathbf{i}} : \mathrm{Rng}(T+\mathbf{i}) \to \mathcal{H} \qquad (T+\mathbf{i})\xi \mapsto (T-\mathbf{i})\xi$$

for each $\xi \in \mathscr{D}(T)$.

Proposition 6.3.3. *Let T be a Hermitian operator on $\mathcal{H}$. After restricting the codomain, the Cayley transform U_T of T is a unitary map*

$$\boxed{\frac{T-\mathbf{i}}{T+\mathbf{i}} : \mathrm{Rng}(T+\mathbf{i}) \xrightarrow{\simeq} \mathrm{Rng}(T-\mathbf{i}) \qquad (T+\mathbf{i})\xi \mapsto (T-\mathbf{i})\xi} \tag{6.11}$$

Moreover, the range of $U_T - \mathbf{1}$ equals $\mathscr{D}(T)$; in particular, it is a dense in $\mathcal{H}$.

Proof. For each $\xi \in \mathscr{D}(T)$, we compute that

$$\|(T \pm \mathbf{i})\xi\|^2 = \langle T\xi|T\xi\rangle + \langle \xi|\xi\rangle \tag{6.12}$$

since $\langle \xi|T\xi\rangle = \langle T\xi|\xi\rangle$ is real. Thus U_T is a linear isometry; its image is clearly $\mathrm{Rng}(T-\mathbf{i})$.

The map $U_T - \mathbf{1}$ sends each $(T+\mathbf{i})\xi$ to $(T-\mathbf{i})\xi - (T+\mathbf{i})\xi = -2\xi$ (where $\xi \in \mathscr{D}(T)$). Hence $\mathrm{Rng}(U_T - \mathbf{1}) = \mathscr{D}(T)$. $\square$

Our next goal is to determine which unitary maps arise from Hermitian operators via the Cayley transform.

Lemma 6.3.4. *Let $U : \mathscr{D}(U) \xrightarrow{\simeq} \mathrm{Rng}(U)$ be a unitary map between two linear subspaces of $\mathcal{H}$. Suppose that $\mathrm{Rng}(U-\mathbf{1})$ is dense in $\mathcal{H}$. Then $\mathrm{Ker}(U-\mathbf{1}) = \{0\}$.*

More precisely, if $(U-\mathbf{1})\mathscr{D}(U)$ is dense in $\mathcal{H}$, then the only vector $\xi \in \mathscr{D}(U)$ satisfying $U\xi = \xi$ is the zero vector.

Proof. Choose any $\xi \in \mathscr{D}(U)$ satisfying $U\xi = \xi$. Then for each $\eta \in \mathscr{D}(U)$, we have $\langle \eta|\xi\rangle = \langle U\eta|U\xi\rangle$, and hence

$$\langle (U-\mathbf{1})\eta|\xi\rangle = \langle U\eta|\xi\rangle - \langle \eta|\xi\rangle = \langle U\eta|\xi\rangle - \langle U\eta|U\xi\rangle = \langle U\eta|\xi - U\xi\rangle = 0$$

Since all such $(U-\mathbf{1})\eta$ form the dense linear subspace $\mathrm{Rng}(U-\mathbf{1})$, we conclude that $\xi = 0$. $\square$

Theorem 6.3.5. *There is a bijection between:*

(1) A Hermitian operator T on $\mathcal{H}$.

(2) A unitary map $U : \mathscr{D}(U) \xrightarrow{\simeq} \mathrm{Rng}(U)$ where $\mathscr{D}(U)$ and $\mathrm{Rng}(U)$ are linear subspaces of $\mathcal{H}$, and $\mathrm{Rng}(U-\mathbf{1})$ is dense in $\mathcal{H}$.

Viewing each U in (2) as an n.d.d. unbounded operator on $\mathcal{H}$, the correspondence is given by

$$U = (T-\mathbf{i})(T+\mathbf{i})^{-1} \qquad T = -\mathbf{i}(U+\mathbf{1})(U-\mathbf{1})^{-1}$$

Moreover, if $\widetilde{T}$ and $\widetilde{U}$ are also related by this bijection, then

$$T \subset \widetilde{T} \qquad \Longleftrightarrow \qquad U \subset \widetilde{U} \tag{6.13}$$

Note that $(U-\mathbf{1})^{-1}$ is well-defined by Lem. 6.3.4. As in Def. 6.3.2, it is customary to write $-\mathbf{i}(U+\mathbf{1})(U-\mathbf{1})^{-1}$ as $\frac{U+\mathbf{1}}{\mathbf{i}(U-\mathbf{1})}$ and call it the **inverse Cayley transform** of U. It is the unique map satisfying

$$\boxed{\frac{U+\mathbf{1}}{\mathbf{i}(U-\mathbf{1})} : \mathrm{Rng}(U-\mathbf{1}) \rightarrow \mathcal{H} \qquad \mathbf{i}(U\eta - \eta) \mapsto U\eta + \eta} \tag{6.14}$$

where $\eta \in \mathscr{D}(U)$.

Proof. Step 1. For each Hermitian T on $\mathcal{H}$, we denote its Cayley transform by U_T. Then U_T satisfies condition (2) by Prop. 6.3.3. For each U satisfying (2) we denote its inverse Cayley transform by T_U. Let us prove that T_U is a Hermitian operator on $\mathcal{H}$.

By (6.14), the domain of T_U is $\mathrm{Rng}(U-\mathbf{1})$, which is dense in $\mathcal{H}$ by assumption. For each $\xi = (U-\mathbf{1})\eta$ where $\eta \in \mathscr{D}(U)$, we compute that

$$\langle \xi|T_U\xi\rangle = \langle U\eta - \eta|T_U(U\eta-\eta)\rangle \xlongequal{(6.14)} -\mathbf{i}\langle U\eta - \eta|U\eta+\eta\rangle = -\mathbf{i}(\langle U\eta|\eta\rangle - \langle \eta|U\eta\rangle)$$

where $\langle U\eta|U\eta\rangle = \langle\eta|\eta\rangle$ is used. Since $\langle U\eta|\eta\rangle = \overline{\langle\eta|U\eta\rangle}$, the last term above is a real number. Thus $\langle\xi|T_U\xi\rangle$ is real, and hence T_U is Hermitian.

Step 2. Next we must prove that $T_{U_T} = T$ and $U_{T_U} = U$, and that the equivalence (6.13) holds. While these facts can be checked directly, the most convenient argument is to use operator graphs. We therefore defer the proof of Step 2 to Rem. 6.4.15. □

6.3.3 From extensions of unitary maps $U : \mathscr{D}(U) \to \mathrm{Rng}(U)$ to extensions of Hermitian operators

Thm. 6.3.5 suggests that the Hermitian operators T admitting spectral decompositions are precisely those whose Cayley transforms U_T are unitary operators on $\mathcal{H}$.[5] Equivalently, these are the operators for which

$$\mathscr{D}(U_T) \equiv \mathrm{Rng}(T+\mathbf{i}) = \mathcal{H} \qquad \mathrm{Rng}(U_T) \equiv \mathrm{Rng}(T-\mathbf{i}) = \mathcal{H} \tag{6.15}$$

which justifies the condition (6.4).

Even if the Hermitian operator T does not satisfy (6.15) (that is, if the Cayley transform U_T is not a unitary operator on the whole space $\mathcal{H}$), Thm. 6.3.5 shows that T can be extended to a Hermitian operator satisfying (6.15) iff U_T can be extended to a unitary operator on $\mathcal{H}$. Our next task, therefore, is to extend operators U satisfying condition (2) of Thm. 6.3.5, and to analyze how such extensions correspond, via the Cayley transform, to extensions of Hermitian operators.

The natural first step is to extend any such U to a bounded linear operator

$$\overline{U} : \overline{\mathscr{D}(U)} \longrightarrow \overline{\mathrm{Rng}(U)}$$

which exists uniquely by Thm. 2.4.2 (together with the completeness of $\overline{\mathrm{Rng}(U)}$). This extension remains a linear isometry, and hence is a unitary map from $\overline{\mathscr{D}(U)}$ to $\overline{\mathrm{Rng}(U)}$ by Lem. 5.10.21. The second step is then to extend $\overline{U}$ further; however, this extension is no longer unique.

Question 6.3.6. Let T be a Hermitian operator on $\mathcal{H}$. Let $\overline{T}$ denote the inverse Cayley transform of $\overline{U_T}$, the unique bounded linear extension of U_T to $\overline{\mathscr{D}(U_T)}$. Can we give a direct definition of $\overline{T}$ without appealing to its Cayley transform $\overline{U_T}$?

The idea for approaching this question is as follows. Note that $\mathscr{D}(\overline{U_T})$ consists of vectors $\eta \in \mathcal{H}$ that are limits of elements of $\mathscr{D}(U_T) = \mathrm{Rng}(T+\mathbf{i})$. Thus, $\eta \in \mathcal{H}$ belongs to $\mathscr{D}(\overline{U_T})$ iff there exists a sequence (ξ_n) in $\mathscr{D}(T)$ such that $\eta = \lim_n(T+\mathbf{i})\xi_n$. In view of (6.12), the fact that $(T\xi_n + \mathbf{i}\xi_n)$ is a Cauchy sequence is equivalent

[5] If $T = \mathbf{M}_f$ where $f : X \to \mathbb{R}$ is measurable, one can show that $U_T = \mathbf{M}_u$ where $u = (f-\mathbf{i})/(f+\mathbf{i})$, and hence is unitary on $\mathcal{H}$.

to that both (ξ_n) and $(T\xi_n)$ are Cauchy sequences. This suggests that $\xi := \lim_n \xi_n$ should be in the domain $\mathscr{D}(\overline{T})$, and that such ξ form the whole domain $\mathscr{D}(\overline{T})$. Moreover, it is easy to guess how $\overline{T}$ is defined on such ξ: the element $\overline{T}\xi$ should equal $\lim_n T\xi_n$. To summarize:

> Elements of $\mathscr{D}(\overline{T})$ are precisely those of the form $\xi = \lim_n \xi_n$ where (ξ_n) is a sequence/net in $\mathscr{D}(T)$ such that both (ξ_n) and $(T\xi_n)$ are Cauchy sequences. Moreover, $\overline{T}\xi = \lim_n T\xi_n$. (6.16)

Although the above reasoning can be made rigorous, in the next section we prefer to present the argument using the important concept of operator graphs, introduced implicitly by von Neumann in [vN29a] and later further explored in [vN32b]. In particular, as we will see in Def. 6.4.9, the approximation $(\xi_n, T\xi_n) \rightarrow (\xi, \overline{T}\xi)$ mentioned above can be understood geometrically in terms of operator graphs and their closures.

6.4 Extensions of Hermitian operators II: the unique part

Fix Hilbert spaces $\mathcal{H}, \mathcal{K}$.

In this section, we define closures of n.d.d. unbounded operators and explore their basic properties. Not every unbounded operator admits a closure; those that do are called closable. In Thm. 6.4.16, we will prove that every Hermitian operator T is closable, that its closure $\overline{T}$ is Hermitian, and that the Cayley transform of $\overline{T}$ is exactly the unique bounded linear extension of the Cayley transform U_T of T to $\overline{\mathscr{D}(U_T)}$. This provides a complete answer to Question 6.3.6.

6.4.1 Graphs and closures of unbounded operators

Definition 6.4.1. For each n.d.d. unbounded operator $T : \mathcal{H} \rightarrow \mathcal{K}$, the **graph** of T is defined to be

$$\mathscr{G}(T) = \{\xi \oplus T\xi : \xi \in \mathscr{D}(T)\}$$

which is clearly a linear subspace of $\mathcal{H} \oplus \mathcal{K}$. We equip $\mathscr{G}(T)$ with the inner product inherited from $\mathcal{H} \oplus \mathcal{K}$.

Clearly, for two n.d.d. unbounded operators $T_1, T_2 : \mathcal{H} \rightarrow \mathcal{K}$, we have $T_1 \subset T_2$ iff $\mathscr{G}(T_1) \subset \mathscr{G}(T_2)$.

Definition 6.4.2. Let $\mathfrak{G}$ be a linear subspace of $\mathcal{H} \oplus \mathcal{K}$. The **domain** of $\mathfrak{G}$ is defined by

$$\boldsymbol{\mathscr{D}(\mathfrak{G})} = \{\xi \in \mathcal{H} : \xi \oplus \eta \in \mathfrak{G} \text{ for some } \eta \in \mathcal{K}\}$$

which is clearly a linear subspace of $\mathcal{H}$. We say that $\mathfrak{G}$ is **densely-defined** if $\mathscr{D}(\mathfrak{G})$ is dense in $\mathcal{H}$.

It is helpful to view a linear subspace $\mathfrak{G} \in \mathcal{H} \oplus \mathcal{K}$ as a linear relation between elements of $\mathcal{H}$ and $\mathcal{K}$.

Example 6.4.3. Let $T : \mathcal{H} \to \mathcal{K}$ be an n.d.d. unbounded operator. Then

$$\mathscr{D}(T) = \mathscr{D}(\mathscr{G}(T))$$

Definition 6.4.4. Let $\mathfrak{G}$ be a linear subspace of $\mathcal{H} \oplus \mathcal{K}$. We say that $\mathfrak{G}$ is an **operator graph** if one of the following equivalent conditions hold:

(1) $\mathfrak{G} = \mathscr{G}(T)$ for some n.d.d. unbounded operator $T : \mathcal{H} \to \mathcal{K}$.

(2) For each $\xi \in \mathcal{H}$ there exists at most one $\eta \in \mathcal{H}$ such that $\xi \oplus \eta \in \mathfrak{G}$.

(3) Any element $\eta \in \mathcal{K}$ satisfying $0 \oplus \eta \in \mathfrak{G}$ must be zero.

Proof of equivalence. We clearly have (1)⇔(2) and (2)⇒(3). Assume (3). If both $\xi \oplus \eta$ and $\xi \oplus \eta'$ belong to $\mathfrak{G}$, then $0 \oplus (\eta - \eta') \in \mathfrak{G}$, and hence $\eta = \eta'$. This proves (2). □

Definition 6.4.5. Let $T : \mathcal{H} \to \mathcal{K}$ be an n.d.d. unbounded operator. We say that T is **closed** if $\mathscr{G}(T)$ is a closed subset of $\mathcal{H} \oplus \mathcal{K}$. We say that T is **closable** if one of the following equivalent conditions hold:

(1) $\overline{\mathscr{G}(T)} \equiv \mathrm{Cl}_{\mathcal{H}\oplus\mathcal{K}}(\mathscr{G}(T))$ is an operator graph.

(2) If (ξ_n) is a sequence in $\mathscr{D}(T)$ converging to 0 such that $(T\xi_n)$ converges, then $\lim_n T\xi_n = 0$.

(3) If (ξ_α) is a net in $\mathscr{D}(T)$ converging to 0 such that $(T\xi_\alpha)$ converges, then $\lim_\alpha T\xi_\alpha = 0$.

If T is closable, we let $\overline{\boldsymbol{T}}$ be the unique n.d.d. unbounded operator $\mathcal{H} \to \mathcal{K}$ extending T, and call $\overline{T}$ the **closure** of T. It is easy to check that $\overline{T}$ satisfies (6.16). □

Proof of equivalence. For each $\eta \in \mathcal{K}$, we have $0 \oplus \eta \in \overline{\mathscr{G}(T)}$ iff there exists a sequence $(\xi_n \oplus T\xi_n)$ in $\mathscr{G}(T)$ converging to $0 \oplus \eta$, iff there exists a sequence (ξ_n) in $\mathscr{D}(T)$ converging to 0 such that $\lim_n T\xi_n = \eta$. Thus, condition (3) of Def. 6.4.4 is equivalent to condition (2) of the current definition. This proves (1)⇔(2). A similar argument proves (1)⇔(3). □

Convention 6.4.6. As in Convention 6.2.2, a closable/closed operator T is automatically understood to be densely-defined. If this assumption is dropped, we will say that T is an **n.d.d. closable/closed operator**. However, if it has already been stated that an unbounded operator is n.d.d., then "closable" and "closed" are automatically taken in the n.d.d. sense.

Proposition 6.4.7. *Let $T : \mathcal{H} \to \mathcal{K}$ be an n.d.d. unbounded operator. The following are equivalent.*

(1) T is closable.

(2) There is an n.d.d. closed operator $\widetilde{T} : \mathcal{H} \to \mathcal{K}$ extending T.

Moreover, if T is closable, then $\overline{T}$ is the smallest n.d.d. closed operator extending T.

Proof. Assume (1). Then $\overline{T}$ is an n.d.d. closed operator extending T. This proves (2). If $\widetilde{T} : \mathcal{H} \to \mathcal{K}$ is also closed and extends T, then $\mathscr{G}(\widetilde{T})$ is a closed set containing the closure $\mathscr{G}(\overline{T})$ of $\mathscr{G}(T)$, and hence $\overline{T} \subset \widetilde{T}$. This proves that $\overline{T}$ is the smallest n.d.d. closed extension.

Assume (2), then each η in the second component of $\mathscr{G}(\widetilde{T})$ corresponds to at most one element of $\mathcal{H}$. The same property holds for any linear subspace of $\mathscr{G}(\widetilde{T})$. In particular, it holds for $\overline{\mathscr{G}(T)}$, since $\mathscr{G}(\widetilde{T})$ is closed and contains $\mathscr{G}(T)$. Therefore, $\overline{\mathscr{G}(T)}$ is an operator graph. This proves (1). □

Proposition 6.4.8. *Let $T : \mathcal{H} \to \mathcal{K}$ be an n.d.d. unbounded operator whose restriction $T|_{\mathscr{D}(T)} : \mathscr{D}(T) \to \mathcal{K}$ is bounded. Then T is closable, and the closure $\overline{T}$ is the unique bounded linear extension of $T|_{\mathscr{D}(T)}$ to $\overline{\mathscr{D}(T)} \to \mathcal{K}$. In particular, T is closed iff $\mathscr{D}(T)$ is a closed subspace of $\mathcal{H}$.*

Proof. Let $\widehat{T} : \overline{\mathscr{D}(T)} \to \mathcal{K}$ denote the unique bounded linear extension of T, which exists due to Thm. 2.4.2.

For each $\xi \oplus \eta \in \mathcal{H} \oplus \mathcal{K}$, we have $\xi \oplus \eta \in \overline{\mathscr{G}(T)}$ iff there exists a sequence (ξ_n) in $\mathscr{D}(T)$ such that $\xi_n \to \xi$ and $T\xi_n \to \eta$, iff there exists a sequence (ξ_n) in $\mathscr{D}(T)$ such that $\xi_n \to \xi$ and $\widehat{T}\xi_n \to \eta$, iff (by the boundedness of $\widehat{T}$) there exists a sequence (ξ_n) in $\mathscr{D}(T)$ such that $\xi_n \to \xi$ and $\widehat{T}\xi = \eta$, iff $\xi \oplus \eta \in \mathscr{G}(\widehat{T})$. □

Definition 6.4.9. Let $T : \mathcal{H} \to \mathcal{K}$ be an n.d.d. unbounded operator. A linear subspace $\mathscr{D}_0 \subset \mathscr{D}(T)$ is called a **core for *T*** if one of the following (clearly) equivalent conditions holds:

- $\mathscr{G}(T|_{\mathscr{D}_0})$ is dense in $\mathscr{G}(T)$.
- For each $\xi \in \mathscr{D}(T)$ there exists a sequence (ξ_n) in $\mathscr{D}_0$ such that $\xi = \lim_n \xi_n$ and $T\xi = \lim_n T\xi_n$.
- For each $\xi \in \mathscr{D}(T)$ there exists a net (ξ_α) in $\mathscr{D}_0$ such that $\xi = \lim_\alpha \xi_\alpha$ and $T\xi = \lim_\alpha T\xi_\alpha$.

For example, if T is closable, then $\mathscr{D}(T)$ is a core for $\overline{T}$.

Definition 6.4.10. If $\mathfrak{G}\subset\mathcal{H}\oplus\mathcal{K}$ is a linear subspace, we call

$$\mathfrak{G}^{-1}=\{\eta\oplus\xi:\xi\oplus\eta\in\mathfrak{G}\}$$

the **diagonal reflection** of $\mathfrak{G}$. In other words, $\mathfrak{G}^{-1}$ is the image of $\mathfrak{G}$ under the (clearly unitary) **diagonal reflection map**

$$\mathfrak{r}:\mathcal{H}\oplus\mathcal{K}\rightarrow\mathcal{K}\oplus\mathcal{H}\qquad\xi\oplus\eta\mapsto\eta\oplus\xi \tag{6.17}$$

6.4.2 Closures of Hermitian operators and their Cayley transforms

Definition 6.4.11. Define the **Cayley transform**

$$\mathscr{C}\!\mathit{ay}:\mathcal{H}\oplus\mathcal{H}\xrightarrow{\simeq}\mathcal{H}\oplus\mathcal{H}\qquad\xi\oplus\phi\mapsto\frac{\phi+\mathbf{i}\xi}{\sqrt{2}}\oplus\frac{\phi-\mathbf{i}\xi}{\sqrt{2}} \tag{6.18}$$

If $\mathfrak{G}\subset\mathcal{H}\oplus\mathcal{H}$, the set $\mathscr{C}\!\mathit{ay}(\mathfrak{G})$ is called the **Cayley transform** of $\mathfrak{G}$, and $\mathfrak{G}$ is called the **inverse Cayley transform** of $\mathscr{C}\!\mathit{ay}(\mathfrak{G})$. Clearly $\mathfrak{G}$ is the image of $\mathscr{C}\!\mathit{ay}(\mathfrak{G})$ under

$$\mathscr{C}\!\mathit{ay}^{-1}:\mathcal{H}\oplus\mathcal{H}\xrightarrow{\simeq}\mathcal{H}\oplus\mathcal{H}\qquad\eta\oplus\psi\mapsto\frac{\mathbf{i}(\psi-\eta)}{\sqrt{2}}\oplus\frac{\psi+\eta}{\sqrt{2}} \tag{6.19}$$

Proposition 6.4.12. *The Cayley transform $\mathscr{C}\!\mathit{ay}$ (defined in (6.18)) is a unitary map. Consequently, its inverse $\mathscr{C}\!\mathit{ay}^{-1}$ is also unitary.*

Proof. Clearly $\mathscr{C}\!\mathit{ay}$ is bijective. To show that it is a linear isometry, we compute that

$$\|\phi+\mathbf{i}\xi\|^2=\|\phi\|^2+\|\xi\|^2+\mathbf{i}\langle\phi|\xi\rangle-\mathbf{i}\langle\xi|\phi\rangle=\|\phi\|^2+\|\xi\|^2+2\mathbf{i}\mathrm{Re}\langle\phi|\xi\rangle$$

and similarly $\|\phi-\mathbf{i}\xi\|^2=\|\phi\|^2+\|\xi\|^2-2\mathbf{i}\mathrm{Re}\langle\phi|\xi\rangle$. $\square$

Example 6.4.13. Let T be a Hermitian operator on $\mathcal{H}$. Then $\mathscr{C}\!\mathit{ay}(\mathscr{G}(T))$ is the graph of the Cayley transform of T.

Proof. The Cayley transform U_T of T sends $T\xi+\mathbf{i}\xi$ to $T\xi-\mathbf{i}\xi$ where $\xi\in\mathscr{D}(T)$. Replacing ξ with $\xi/\sqrt{2}$, we see that $\mathscr{G}(U_T)$ consists of points of the form $(T\xi+\mathbf{i}\xi)/\sqrt{2}\oplus(T\xi-\mathbf{i}\xi)/\sqrt{2}$. This is compatible with (6.18) by setting $\phi=T\xi$. $\square$

Example 6.4.14. Let U satisfy Condition (2) of Thm. 6.3.5. Then $\mathscr{C}\!\mathit{ay}^{-1}(\mathscr{G}(U))$ is the graph of the inverse Cayley transform of T.

Proof. The inverse Cayley transform T_U of U sends $\mathbf{i}(U\eta-\eta)$ to $U\eta+\eta$ where $\eta\in\mathscr{D}(U)$. Replacing η with $\eta/\sqrt{2}$, we see that $\mathscr{G}(T_U)$ consists of points of the form $\mathbf{i}(U\eta-\eta)/\sqrt{2}\oplus(U\eta+\eta)/\sqrt{2}$. This is compatible with (6.19) by setting $\psi=U\eta$. $\square$

Remark 6.4.15. We are now ready to complete the second step in the proof of Thm. 6.3.5.

Proof. Let U_T be the Cayley transform of T, and let T_U be the inverse Cayley transform of U. By Exp. 6.4.13 and 6.4.14, the Cayley transform $\mathscr{Cay}$ maps the graph of T to the graph of U_T; the inverse Cayley transform $\mathscr{Cay}^{-1}$ maps the graph of U to the graph of T_U, and hence maps the graph of U_T to the graph of T_{U_T}. So the graphs of T and T_{U_T} are equal. This proves $T = T_{U_T}$. A similar argument shows $U = U_{T_U}$.

That $T \subset \widetilde{T}$ is equivalent to $U_T \subset U_{\widetilde{T}}$ follows directly from the fact that $\mathscr{Cay}$ maps a larger graph to a larger one. $\square$

Theorem 6.4.16. *Let T be a Hermitian operator on $\mathcal{H}$ with Cayley transform U_T. Then T is closable and $\overline{T}$ is Hermitian. Moreover, the Cayley transform of $\overline{T}$ is equal to the unique bounded linear extension of U_T to $\overline{\mathscr{D}(U_T)} \to \mathcal{H}$.*

It follows that T is closed iff $\mathscr{D}(U_T)$ is closed (equivalently, $\mathrm{Rng}(U_T)$ is closed).

Proof. By Prop. 6.4.8, the operator U_T is closable, and its closure $\overline{U_T}$ is precisely the unique bounded linear extension of U_T to $\overline{\mathscr{D}(U_T)} \to \mathcal{H}$. Since $U_T : \mathscr{D}(U_T) \to \mathcal{H}$ is a linear isometry, the bounded linear extension $\overline{U_T} : \overline{\mathscr{D}(U_T)} \to \mathcal{H}$ remains a linear isometry by continuity. Because $\mathrm{Rng}(U_T - \mathbf{1})$ is dense $\mathcal{H}$, the larger set $\mathrm{Rng}(\overline{U_T} - \mathbf{1})$ is also dense. Thus $\overline{U_T}$ satisfies Condition (2) of Thm. 6.3.5.

By Thm. 6.3.5, there exists a Hermitian operator $\widehat{T}$ on $\mathcal{H}$ whose Cayley transform $U_{\widehat{T}}$ equals $\overline{U_T}$. Since the map $\mathscr{Cay}$ is unitary (cf. Prop. 6.4.12), the fact that $\overline{U_T}$ equals the closure of U_T implies that (the graph of) $\widehat{T}$ equals the closure of (the graph of) T. This proves that T is closable, and that its closure $\widehat{T}$ is Hermitian with Cayley transform $\overline{U_T}$. $\square$

6.5 Extensions of Hermitian operators III: the non-unique part

Let $\mathcal{H}$ be a Hilbert space. In Thm. 6.3.5, we have related the extensions of a Hermitian operator T to the isometric extensions of its Cayley transform U_T. According to Thm. 6.4.16, the extension of T to $\overline{T}$ corresponds to the unique bounded linear extension of U_T to $\overline{\mathscr{D}(U_T)}$.

In this section, we study closed extensions of a closed Hermitian operator T, which corresponds to isometric extensions of U_T to larger closed linear subspaces of $\mathcal{H}$ containing $\mathscr{D}(U_T)$.

Theorem 6.5.1. *Let T be a closed Hermitian operator on $\mathcal{H}$. Then there exists a bijection between:*

(1) A closed Hermitian operator $\widehat{T}$ extending T.

(2) A unitary map $V : \mathscr{D}(V) \xrightarrow{\simeq} \mathrm{Rng}(V)$ where $\mathscr{D}(V)$ is a Hilbert subspace of $\mathrm{Rng}(T + \mathbf{i})^\perp$ and $\mathrm{Rng}(V)$ is a Hilbert subspace of $\mathrm{Rng}(T - \mathbf{i})^\perp$.

Viewing V as an n.d.d. unbounded operator on $\mathcal{H}$, the correspondence is given by

$$\begin{aligned}\mathscr{D}(\widehat{T}) &= \mathscr{D}(T) + \mathrm{Rng}(V - \mathbf{1})\\ \widehat{T}|_{\mathrm{Rng}(V-\mathbf{1})} &: \mathbf{i}(V\eta - \eta) \mapsto (V\eta + \eta)\end{aligned} \tag{6.20}$$

for each $\eta \in \mathscr{D}(V)$.

If one does not require that $\widehat{T}$ is closed, then $\mathscr{D}(V)$ and $\mathrm{Rng}(V)$ are not necessarily complete.

Proof. Let $U_T : \mathrm{Rng}(T + \mathbf{i}) \xrightarrow{\simeq} \mathrm{Rng}(T + \mathbf{i})$ be the Cayley transform of T. By Thm. 6.3.5 and 6.4.16, the Cayley transform establishes a bijection between $\widehat{T}$ and unitary $\widehat{U} : \mathscr{D}(\widehat{U}) \xrightarrow{\simeq} \mathrm{Rng}(\widehat{U})$ extending U_T, where $\mathscr{D}(\widehat{U})$ and $\mathrm{Rng}(\widehat{U})$ are closed linear subspaces of $\mathcal{H}$.

The bijection between $\widehat{U}$ and V is as follows. Given $\widehat{U}$, then $\mathscr{D}(V)$ is chosen to be the orthogonal complement of $\mathrm{Rng}(T + \mathbf{i})$ in $\mathscr{D}(\widehat{U})$, and V is defined to be $\widehat{U}|_{\mathscr{D}(V)}$. Conversely, given V, then $\mathscr{D}(\widehat{U})$ is set to be $\mathrm{Rng}(T + \mathbf{i}) + \mathscr{D}(V)$, and $\widehat{U}$ is defined to be the unique extension of the Cayley transform U_T whose restriction to $\mathscr{D}(V)$ equals V. Eq. (6.20) is easy to check. □

Definition 6.5.2. Let T be a Hermitian operator on $\mathcal{H}$. Let

$$n_+ = \dim \mathrm{Rng}(T + \mathbf{i})^\perp \qquad n_- = \dim \mathrm{Rng}(T - \mathbf{i})^\perp$$

More precisely, $n_\pm$ is the cardinality of any orthonormal basis of $\mathrm{Rng}(T \pm \mathbf{i})^\perp$ (cf. Pb. 3.3). We call (n_+, n_-) the (pair of) **deficiency indices** of T.

If $\mathcal{H}$ is separable, the only infinite cardinality is that of $\mathbb{N}$. In that case, it is unambiguous to denote this cardinality by ∞.

By Thm. 6.4.16, T and its closure $\overline{T}$ have the same deficiency indices. Therefore, it suffices to consider deficiency indices of closed Hermitian operators.

Remark 6.5.3. As discussed in Subsec. 6.3.3, the Hermitian operators T that admit spectral decompositions are those satisfying $\mathrm{Rng}(T + \mathbf{i}) = \mathrm{Rng}(T - \mathbf{i}) = \mathcal{H}$. Later we will show that these operators are precisely the self-adjoint ones. By Thm. 6.5.1, this implies that a Hermitian operator T has a self-adjoint extension iff its deficiency indices satisfy $n_+ = n_-$. See Cor. 6.7.11.

6.6 Adjoints of unbounded operators

Let $\mathcal{H}, \mathcal{K}$ be Hilbert spaces.

In [vN29a], von Neumann introduced the notion of the adjoint operator T^* for a Hermitian operator T.[6] The idea is straightforward: viewing T as a differential

[6] However, in [vN29a] von Neumann did not yet regard T^* as the adjoint of T. See Subsec. 6.7.1 for a detailed discussion.

operator, the extended operator T^* should have domain $\mathscr{D}(T^*)$ consisting of all $\eta \in \mathcal{H}$ for which the "weak derivative" $T^*\eta$ exists. Concretely, these are the vectors $\eta \in \mathcal{H}$ such that the linear functional

$$\mathscr{D}(T) \to \mathbb{C} \qquad \xi \mapsto \langle \eta | T\xi \rangle \tag{6.21}$$

is bounded. Since $\mathscr{D}(T)$ is dense, we have $\mathscr{D}(T)^* \simeq \mathcal{H}^*$ due to Cor. 2.4.3. Therefore, by the Riesz-Fréchet Thm. 3.5.3, there exists a unique element $T^*\eta \in \mathcal{H}$ such that

$$\langle \eta | T\xi \rangle = \langle T^*\eta | \xi \rangle \qquad \text{for all } \xi \in \mathscr{D}(T)$$

This definition of T^* clearly applies to any unbounded operator T on $\mathcal{H}$. However, without assuming that T is Hermitian, the adjoint operator T^* need not extend T. In fact, $\mathscr{D}(T^*)$ is not necessarily densely-defined, so T^* may be only an n.d.d. unbounded operator. The systematic study of adjoints of general (not necessarily Hermitian) unbounded operators was undertaken by von Neumann in [vN32b]. In this course, contrary to the historical order, we will first examine adjoints of general operators before specializing to Hermitian operators. The most surprising result we will see in this section is the equivalence of the closability of T and the density of $\mathscr{D}(T^*)$; see Thm. 6.6.9.

6.6.1 Adjoints of graphs

In Sec. 6.4, we have seen how the graphs $\mathscr{G}(T)$ of operators T illuminate both the algebraic and analytic features of the Cayley transform. Operator graphs are also a powerful tool for studying adjoints.

Definition 6.6.1. For each linear subspace $\mathfrak{G} \subset \mathcal{H} \oplus \mathcal{K}$, the **adjoint** of $\mathfrak{G}$ is defined to be [7]

$$\mathbf{Ad}(\mathfrak{G}) = \{\eta \oplus \psi \in \mathcal{K} \oplus \mathcal{H} : \langle \eta | \phi \rangle = \langle \psi | \xi \rangle \text{ for all } \xi \oplus \phi \in \mathfrak{G}\}$$

An equivalent description is as follows. Define a (clearly) unitary map

$$\mathbb{J}_{\mathcal{H},\mathcal{K}} : \mathcal{H} \oplus \mathcal{K} \xrightarrow{\simeq} \mathcal{K} \oplus \mathcal{H} \qquad \xi \oplus \phi \mapsto \mathbf{i}\phi \oplus (-\mathbf{i}\xi)$$

abbreviated to $\mathbb{J}$ when the context is clear. Then

$$\mathrm{Ad}(\mathfrak{G}) = (\mathbb{J}\mathfrak{G})^\perp$$

Proposition 6.6.2. *The unitary map $\mathbb{J}$ satisfies $\mathbb{J}_{\mathcal{K},\mathcal{H}}\mathbb{J}_{\mathcal{H},\mathcal{K}} = \mathbf{1}$ and hence*

$$\mathbb{J}_{\mathcal{H},\mathcal{K}}^* = \mathbb{J}_{\mathcal{H},\mathcal{K}}^{-1} = \mathbb{J}_{\mathcal{K},\mathcal{H}} \tag{6.22}$$

[7]We avoide the notation $\mathfrak{G}^*$ for the adjoint, since it is already reserved for dual spaces.

Proof. This is easy to check. □

Proposition 6.6.3. *For each linear subspace $\mathfrak{G} \subset \mathcal{H} \oplus \mathcal{K}$ we have*

$$(\mathbb{J}\mathfrak{G})^\perp = \mathbb{J}(\mathfrak{G}^\perp) \qquad (\mathbb{J}\mathfrak{G})^{-1} = \mathbb{J}(\mathfrak{G}^{-1}) \qquad (\mathfrak{G}^{-1})^\perp = (\mathfrak{G}^\perp)^{-1}$$

Consequently, we have

$$\mathrm{Ad}(\mathfrak{G}^{-1}) = (\mathrm{Ad}\mathfrak{G})^{-1}$$

Therefore, it is unambiguous to omit parentheses and write $\mathbb{J}\mathfrak{G}^\perp$ and $\mathbb{J}\mathfrak{G}^{-1}$.

Proof. The first and third relations follow from the fact that any unitary tranform commutes with taking orthogonal complements. The second one follows from the fact that the diagonal reflection map $\mathfrak{r}$ satisfies $\mathbb{J}\mathfrak{r} = -\mathfrak{r}\mathbb{J}$ and the fact that $-\mathfrak{V} = \mathfrak{V}$ for any linear subspace $\mathfrak{V}$. □

Theorem 6.6.4. *Let $\mathfrak{G} \subset \mathcal{H} \oplus \mathcal{K}$ be a linear subspace. The following are true.*

1. $\overline{\mathfrak{G}} = \mathrm{Ad}(\mathrm{Ad}(\mathfrak{G}))$. *In particular,* $\mathfrak{G} \subset \mathrm{Ad}(\mathrm{Ad}(\mathfrak{G}))$.

2. $\mathfrak{G}$ *is densely defined iff* $\mathrm{Ad}(\mathfrak{G})$ *is an operator graph.*

Part 2 is analogous to the fact that T has dense range iff T^* is injective, cf. Prop. 6.6.11.

Proof. Prop. 6.6.3 implies that

$$\mathbb{J}(\mathbb{J}\mathfrak{G}^\perp)^\perp = \mathbb{J}\mathbb{J}(\mathfrak{G}^{\perp\perp}) = \mathfrak{G}^{\perp\perp}$$

where the last term equals $\overline{\mathfrak{G}}$ by Cor. 3.4.8. This proves part 1.

Recall that $\mathscr{D}(\mathfrak{G})$ consists of all $\xi \in \mathcal{H}$ such that $\xi \oplus \phi \in \mathfrak{G}$. By Cor. 3.4.9, $\mathscr{D}(\mathfrak{G})$ is dense in $\mathcal{H}$ iff the only vector $\psi \in \mathcal{H}$ orthogonal to $\xi \in \mathcal{H}$ (for all $\xi \oplus \phi \in \mathfrak{G}$) is zero, iff the only vector $\psi \oplus 0 \in \mathcal{H} \oplus \mathcal{K}$ orthogonal to $\mathfrak{G}$ is zero, iff the only vector $0 \oplus (-\mathbf{i}\psi) = \mathbb{J}(\psi \oplus 0) \in \mathcal{K} \oplus \mathcal{H}$ orthogonal to $\mathbb{J}\mathfrak{G}$ is zero, iff the only vector $0 \oplus (-\mathbf{i}\psi) \in \mathcal{K} \oplus \mathcal{H}$ belonging to $\mathrm{Ad}(\mathfrak{G})$ is zero. This last statement is equivalent to that $\mathrm{Ad}(\mathfrak{G})$ is an operator graph, cf. Def. 6.4.4-(3). This proves part 2. □

Corollary 6.6.5. *Let $\mathfrak{G} \subset \mathcal{H} \oplus \mathcal{K}$ be a linear subspace. Then $\overline{\mathfrak{G}}$ is an operator graph iff* $\mathrm{Ad}(\mathfrak{G})$ *is densely defined.*

Proof. Part 2 of Thm. 6.6.4 implies that $\mathrm{Ad}(\mathfrak{G})$ is densely defined iff $\mathrm{Ad}(\mathrm{Ad}(\mathfrak{G}))$ is an operator graph, and part 1 says $\mathrm{Ad}(\mathrm{Ad}(\mathfrak{G})) = \overline{\mathfrak{G}}$. □

6.6.2 Adjoints of unbounded operators

Definition 6.6.6. Let $T : \mathcal{H} \to \mathcal{K}$ be an unbounded operator. By Thm. 6.6.4, there exists a (unique) n.d.d. unbounded operator $T^* : \mathcal{H} \to \mathcal{K}$ such that $\mathscr{G}(T^*) = \mathrm{Ad}(\mathscr{G}(T))$. We call T^* the **adjoint** of T.

Remark 6.6.7. Let $T : \mathcal{H} \to \mathcal{K}$ be an unbounded operator. An equivalent description of T^* is as follows: $\mathscr{D}(T)$ is the set of all $\eta \in \mathcal{K}$ such that there exists an element of $\mathcal{H}$, denoted by $T^*\eta$, satisfying

$$\langle \eta | T\xi \rangle = \langle T^*\eta | \xi \rangle \qquad \text{for all } \xi \in \mathscr{D}(T) \tag{6.23}$$

As discussed near (6.21), $\mathscr{D}(T^*)$ can also be described as the set of all $\eta \in \mathcal{K}$ such that the linear functional $\xi \in \mathscr{D}(T) \mapsto \langle \eta | T\xi \rangle \in \mathbb{C}$ is bounded.

Remark 6.6.8. Let $A, B : \mathcal{H} \to \mathcal{K}$ be unbounded operators satisfying $A \subset B$. Then $B^* \supset A^*$. This is due to the more general (obvious) fact that if $\mathfrak{G}, \mathfrak{K}$ are linear subspaces of $\mathcal{H} \oplus \mathcal{K}$, then

$$\mathfrak{G} \subset \mathfrak{K} \qquad \Longrightarrow \qquad \mathbb{J}\mathfrak{G}^\perp \supset \mathbb{J}\mathfrak{K}^\perp$$

Theorem 6.6.9. *Let $T : \mathcal{H} \to \mathcal{K}$ be an unbounded operator. Then the following are equivalent:*

(1) T is closable.

(2) T^* is densely-defined (i.e. T^* is an unbounded operator).

Moreover, if either (1) or (2) is true, then T^* is a closed operator, and

$$\overline{T} = T^{**} \qquad (\overline{T})^* = T^* = T^{***} \tag{6.24}$$

Proof. The equivalence of (1) and (2) is due to Cor. 6.6.5. If (1) or (2) is true, then $\overline{T} = T^{**}$ by Thm. 6.6.4-1, and the closedness of T^* is due to the closedness of $\mathscr{G}(T^*) = (\mathbb{J}\mathscr{G}(T))^\perp$. (Note that any orthogonal complement is closed; cf. Rem. 3.3.16.) Applying adjoints to both sides of $\overline{T} = T^{**}$ gives $(\overline{T})^* = T^{***}$. Replacing T by T^* in the relation $\overline{T} = T^{**}$ gives $\overline{T^*} = T^{***}$, and hence $T^* = T^{***}$ because T^* is closed.[8] $\square$

A typical situation in which Thm. 6.6.9 can be applied to prove closability is the following.

Example 6.6.10. Let $T : \mathcal{H} \to \mathcal{K}$ and $S : \mathcal{K} \to \mathcal{H}$ be unbounded operators. Then the following statements are clearly equivalent:

[8] Alternatively, (6.24) also follows from the fact that for each linear subspace V of a Hilbert space, the relations $\overline{V} = V^{\perp\perp}$ and $\overline{V}^\perp = V^\perp = V^{\perp\perp\perp}$ hold due to Rem. 3.3.16 and Cor. 3.4.8.

(1) For each $\xi \in \mathscr{D}(T)$ and $\eta \in \mathscr{D}(S)$ we have $\langle \eta | T\xi \rangle = \langle S\eta | \xi \rangle$.

(2) $S \subset T^*$.

(3) $T \subset S^*$.

Consequently, if (1) holds, then T^* is densely-defined (since $\mathscr{D}(T^*)$ contains the dense subspace $\mathscr{D}(S)$), and hence T is closable by Thm. 6.6.9.

As a special case, if T is an unbounded operator on $\mathcal{H}$, then T is Hermitian iff $T \subset T^*$. □

The following proposition generalizes Prop. 6.2.25.

Proposition 6.6.11. *Let $T : \mathcal{H} \to \mathcal{K}$ be an unbounded operator. Then*

$$\mathrm{Ker}(T^*) = \mathrm{Rng}(T)^\perp \tag{6.25}$$

Consequently, if T is closed, then

$$\mathrm{Ker}(T) = \mathrm{Rng}(T^*)^\perp \tag{6.26}$$

In particular, Eq. (6.25) implies that $\mathrm{Rng}(T)^\perp \subset \mathscr{D}(T^*)$; Eq. (6.26) implies that if T is closed then $\mathrm{Ker}(T)$ is closed.

Proof. Let $\eta \in \mathcal{K}$. Then $\eta \in \mathrm{Ker}(T^*)$ iff the linear functional $\eta \in \mathscr{D}(T) \mapsto \langle \eta | T\xi \rangle$ is bounded and constantly zero, iff $\eta \in \mathrm{Rng}(T)^\perp$. This proves the first relation. When T is closed, then by Thm. 6.6.9, the adjoint T^* is closed, and the first relation gives $\mathrm{Ker}(T^*) = \mathrm{Rng}(T^{***})^\perp = \mathrm{Rng}(T^*)^\perp$. □

Inverses of bounded injective operators with dense range provide a large class of examples of closed operators. The following proposition shows that the adjoints of such closed operators can be computed in terms of the adjoints of the original bounded operators.

Proposition 6.6.12. *Let $T : \mathcal{H} \to \mathcal{K}$ be an injective closed operator with dense range. Then $T^{-1}, T^* : \mathcal{K} \to \mathcal{H}$ are also injective closed operators with dense range, and*

$$(T^*)^{-1} = (T^{-1})^* \tag{6.27}$$

Proof. By Thm. 6.6.9, $T^* : \mathcal{K} \to \mathcal{H}$ is a closed operator. By Prop. 6.6.11, $\mathrm{Ker}(T^*) = \mathrm{Rng}(T)^\perp = \{0\}$ and $\overline{\mathrm{Rng}(T^*)} = \mathrm{Ker}(T)^\perp = \mathcal{H}$. Therefore T^* is injective with dense range. Clearly T^{-1} is injectve with dense range. Since the diagonal reflection map (6.17) is unitary, the operator T^{-1} is closed. Eq. (6.27) follows immediately from Prop. 6.6.3. □

6.6.3 Examples: Sobolev spaces

Let $\Omega \subset \mathbb{R}^N$ be open.

Definition 6.6.13. For each **multi-index**

$$\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$$

we define an unbounded operator

$$\boldsymbol{\partial^\alpha} = \partial_{x_1}^{\alpha_1} \cdots \partial_{x_N}^{\alpha_N} : L^2(\Omega, m) \to L^2(\Omega, m) \qquad \mathscr{D}(\partial^\alpha) = C_c^\infty(\Omega) \tag{6.28}$$

The **order** of α is defined to be

$$\boldsymbol{|\alpha|} = \alpha_1 + \cdots + \alpha_N$$

We also define the **gradient operator**

$$\begin{gathered} \nabla : L^2(\Omega, m) \to L^2(\Omega, m)^{\oplus N} \qquad f \mapsto \partial_{x_1} f \oplus \cdots \oplus \partial_{x_N} f \\ \mathscr{D}(\nabla) = C_c^\infty(\Omega) \end{gathered} \tag{6.29}$$

and the **divergence operator**

$$\begin{gathered} \mathrm{div} : L^2(\Omega, m)^{\oplus N} \to L^2(\Omega, m) \qquad f_1 \oplus \cdots \oplus f_N \mapsto \partial_{x_1} f_1 + \cdots + \partial_{x_N} f \\ \mathscr{D}(\mathrm{div}) = C_c^\infty(\Omega)^{\oplus N} \end{gathered} \tag{6.30}$$

both understood as unbounded operators.

Proposition 6.6.14. *The unbounded operators* $\partial^\alpha, \nabla, \mathrm{div}$ *are closable, and*

$$(\partial^\alpha)^* \supset (-1)^{|\alpha|} \partial^\alpha \qquad \nabla^* \supset -\mathrm{div} \qquad \mathrm{div}^* \supset -\nabla \tag{6.31}$$

Proof. As in the proof of Exp. 6.2.12, integration by parts gives $\langle f | \partial_{x_i} g \rangle = -\langle \partial_{x_i} f | g \rangle$ for each $f, g \in C_c^\infty(\Omega)$. Hence

$$\langle f | \partial^\alpha g \rangle = (-1)^{|\alpha|} \langle \partial^\alpha f | g \rangle \qquad \text{for each } f, g \in C_c^\infty(\Omega)$$

This is equivalent to the relation $(\partial^\alpha)^* \supset (-1)^{|\alpha|} \partial^\alpha$, cf. Exp. 6.6.10. The other two inclusion relations in (6.31) follow by a similar argument.

Eq. (6.31) implies that the adjoints of $\partial^\alpha, \nabla, \mathrm{div}$ are densely defined. Hence they are closable by Thm. 6.6.9. $\square$

Definition 6.6.15. Let $T : \mathcal{H} \to \mathcal{K}$ be an n.d.d. unbounded operator. The **graph inner product** (of T) on $\mathscr{D}(T)$ is defined by

$$\boldsymbol{\langle \cdot | \cdot \rangle_T} : \mathscr{D}(T) \times \mathscr{D}(T) \to \mathbb{C} \qquad \langle \xi | \eta \rangle_T = \langle \xi | \eta \rangle + \langle T\xi | T\eta \rangle$$

In other words, $\langle \cdot | \cdot \rangle_T$ is the pullback of the inner product of $\mathscr{G}(T)$ to $\mathscr{D}(T)$ via the linear bijection

$$\Psi : \mathscr{D}(T) \to \mathscr{G}(T) \qquad \xi \mapsto \xi \oplus T\xi \tag{6.32}$$

It follows that T is closed iff $\mathscr{D}(T)$ is complete with respect to the graph inner product.

Example 6.6.16. Define the **Sobolev spaces**

$$H_0^1(\Omega) = \mathscr{D}\big(\overline{\nabla}\big) \qquad H^1(\Omega) = \mathscr{D}\big(\mathrm{div}^*\big)$$

equipped with the graph inner products of $\overline{\nabla}$ and div^*, respectively. Then $H_0^1(\Omega)$ and $H^1(\Omega)$ are Hilbert spaces, because $\overline{\nabla}$ and div^* are closed.

In this course, we will focus primarily on first-order Sobolev spaces. For interested readers, the following describes how higher-order Sobolev spaces are defined.

Example 6.6.17. For each $k \in \mathbb{Z}_+$, define unbounded operators

$$\nabla^k : L^2(\Omega, m) \to \bigoplus_{1 \leqslant |\alpha| \leqslant k} L^2(\Omega, m) \qquad f \mapsto \oplus_\alpha \partial^\alpha f$$
$$\mathscr{D}(\nabla^k) = C_c^\infty(\Omega)$$

and

$$\mathrm{div}^k : \bigoplus_{1 \leqslant |\alpha| \leqslant k} L^2(\Omega, m) \to L^2(\Omega, m) \qquad \oplus_\alpha f_\alpha \mapsto \sum_\alpha \partial^\alpha f_\alpha$$
$$\mathscr{D}(\mathrm{div}^k) = \bigoplus_{1 \leqslant |\alpha| \leqslant k} C_c^\infty(\Omega)$$

Similar to Prop. 6.6.14, one easily checks that

$$(\nabla^k)^* \supset \mathrm{div}^k \cdot \mathbb{I}_k \qquad (\mathrm{div}^k)^* \supset \mathbb{I}_k \cdot \nabla^k \tag{6.33}$$

where $\mathbb{I}_k \in \mathfrak{L}\big(\bigoplus_{|\alpha| \leqslant k} L^2(\Omega, m)\big)$ is the unitary operator defined by $\mathbb{I}_k(\oplus_\alpha f_\alpha) = \oplus_\alpha (-1)^{|\alpha|} f_\alpha$. Hence ∇^k and div^k are closable by Thm. 6.6.9. Define the **Sobolev spaces**

$$H_0^k(\Omega) = \mathscr{D}\big(\overline{\nabla^k}\big) \qquad H^k(\Omega) = \mathscr{D}\big((\mathrm{div}^k)^*\big)$$

equipped with the graph inner products of $\overline{\nabla^k}$ and $(\mathrm{div}^k)^*$, respectively. Then these two spaces are Hilbert spaces, because $\overline{\nabla^k}$ and $(\mathrm{div}^k)^*$ are closed. Moreover, it is easy to check that

$$H^k(\Omega) = \bigcap_{|\alpha| \leqslant k} \mathscr{D}\big((\partial^\alpha)^*\big)$$

Remark 6.6.18. From the second relation in (6.33), we clearly have that $H_0^k(\Omega) \subset H^k(\Omega)$, and that the inner product of $H^k(\Omega)$ restricts to that of $H_0^k(\Omega)$. Moreover, by (6.33), the closure of $\mathscr{G}(\mathbb{I}_k \cdot \nabla^k)$ in $\mathscr{G}((\mathrm{div}^k)^*)$ is $\mathscr{G}(\mathbb{I}_k \cdot \overline{\nabla^k})$. It follows that $H_0^k(\Omega)$ is the closure of $C_c^\infty(\Omega)$ in $H^k(\Omega)$.

Remark 6.6.19. Heuristically, $H_0^k(\Omega)$ consists of those functions $f \in H^k(\Omega)$ such that $(\partial^\alpha)^* f$ "vanishes at the boundary of Ω" for each $|\alpha| \leqslant k - 1$.

6.7 Adjoints of closed Hermitian operators

Let $\mathcal{H}$ be a Hilbert space.

Definition 6.7.1. Let T be a Hermitian operator on $\mathcal{H}$. We say that T is **self-adjoint** if $T = T^*$. We say that T is **essentially self-adjoint** if $\overline{T}$ is self-adjoint.

Since adjoint operators are closed (Thm. 6.6.9), every self-adjoint operator is closed; in particular, it is essentially self-adjoint.

In this section, we show that a Hermitian operator T on $\mathcal{H}$ satisfies $T = T^*$ iff

$$\mathrm{Rng}(T+\mathbf{i}) = \mathrm{Rng}(T-\mathbf{i}) = \mathcal{H} \tag{6.34}$$

that is, iff the Cayley transform of T is a unitary operator on $\mathcal{H}$. The motivation for establishing this result requires some explanation, which we provide in the following subsection.

6.7.1 Introduction

The most fundamental result in the theory of unbounded operators is that Hermitian operators T satisfying $T = T^*$ admit spectral decompositions, as proved by von Neumann in [vN29a]. Today, we call unbounded operators with $T = T^*$ **self-adjoint** operators. This terminology, however, has often caused confusion about the actual development of the theory of unbounded operators.

The first main confusion is that one might think the notions of adjoint and self-adjointness for unbounded operators are straightforward, and that the real difficulty lies in proving that self-adjoint operators admit spectral decompositions. But this is not how the theory evolved. As emphasized earlier, having a spectral decomposition is not the ultimate goal, but rather the starting point. The true objective is to identify natural conditions on Hermitian operators that are equivalent to the existence of spectral decompositions. The first such condition encountered is (6.34), as discussed in Subsec. 6.3.3. The final condition is $T = T^*$. In other words, the condition $T = T^*$ is not the point of departure, but rather the culmination of the development of spectral theory for unbounded operators.

Why, then, did von Neumann in [vN29a] (and most later authors) not regard (6.34) as the definitive condition for Hermitian operators to admit spectral decompositions? This requires some explanation, especially since in Ch. 7 we will see that in many situations it is actually easier to verify (6.34) than to check $T = T^*$ directly. The reason lies in a second major source of confusion concerning the notion of self-adjoint operators, which I now discuss.

In [vN29a], von Neumann did not use the term "self-adjoint" for Hermitian operators satisfying $T = T^*$. Instead, he called them **hypermaximal operators**. Correspondingly, in that paper he did not regard T^* as the adjoint of T, but rather as an extension of T. (Recall that Hermitian operators T satisfy $T \subset T^*$; see

Exp. 6.6.10.) Of course, for non-Hermitian operators one does not have $T \subset T^*$—but this is precisely why it was only later, in [vN32b], where von Neumann considered T^* for general operators (not just Hermitian ones), that he began to interpret T^* as the adjoint of T.

There are many clues in [vN29a] that demonstrate von Neumann's original perspective of T^* as an extension of the Hermitian operator T, rather than its adjoint:

- In footnote 46, von Neumann used $\overline{\overline{T}}$ to denote what we now call T^*.
- A key concept in [vN29a] is **extension elements** (Erweiterungselemente), which are defined as those vectors belonging to the domain of (what we now call) T^*.
- In the second paragraph of Section I, von Neumann used both "self-adjoint" (selbstadjungiert) and "Hermitian" (Hermitsch) to describe Hermitian operators.

To explain why von Neumann regarded the condition $T = T^*$ (in his notation, $T = \overline{\overline{T}}$) rather than (6.34) as the definitive criterion for Hermitian operators admitting spectral decompositions, let's examine the following result; see [vN29a, Satz 33]:

Proposition 6.7.2. *Let T be a Hermitian operator on $\mathcal{H}$. Then T is **maximal** (i.e., any Hermitian operator extending T must be equal to T) if and only if*

$$\mathrm{Rng}(T + \mathbf{i}) = \mathcal{H} \qquad \textit{or} \qquad \mathrm{Rng}(T - \mathbf{i}) = \mathcal{H} \tag{6.35}$$

It follows from Thm. 6.4.16 that a Hermitian operator T is maximal precisely when it is closed and one of its deficiency indices is zero.

Proof. This follows directly from Thm. 6.3.5, since (6.35) means precisely that the Cayley transform of T has no proper extension satisfying condition (2) of Thm. 6.3.5. □

Of course, a maximal Hermitian operator need not satisfy (6.34). This naturally raises the following question:

Question 6.7.3. What is the analogue of Prop. 6.7.2 for condition (6.34)? In other words, what is the natural property corresponding to (6.34), in the same way that "maximality" corresponds to condition (6.35)?

Answer. The corresponding theorem is that a Hermitian operator T is hypermaximal iff (6.34) holds. The natural property sought is precisely hypermaximality $T = \overline{\overline{T}}$ (i.e. $T = T^*$). Heuristically, this means not only that T has no proper Hermitian extensions, but also that it has no proper "hyperextensions", if we interpret

T^* as the **hyperextension**[9] of T. The hyperextension is larger than any Hermitian extension $\widehat{T}$ of T, since $T \subset \widehat{T}$ implies $\widehat{T}^* \subset T^*$, and hence (due to $\widehat{T} \subset \widehat{T}^*$)

$$T \subset \widehat{T} \subset \widehat{T}^* \subset T^* \tag{6.36}$$

□

6.7.2 Characterization of adjoints of closed Hermitian operators

For a Hermitian operator T on $\mathcal{H}$, to relate the condition $T = T^*$ with (6.34), we need a clear understanding of T^* in terms of $\mathrm{Rng}(T \pm \mathbf{i})$. The goal of this subsection is to present such a characterization of T^*, originally due to [vN29a]. However, we will simplify the original technical computations by using the Cayley transform $\mathscr{Cay} : \mathcal{H} \oplus \mathcal{H} \to \mathcal{H} \oplus \mathcal{H}$ introduced in Def. 6.4.11.

Lemma 6.7.4. *Let T be a Hermitian operator on $\mathcal{H}$ with Cayley transform U_T. Then $\mathscr{Cay}(\mathscr{G}(T^*))$ consists precisely of elements $\alpha \oplus \beta \in \mathcal{H} \oplus \mathcal{H}$ satisfying*

$$\langle \alpha | \eta \rangle = \langle \beta | U_T \eta \rangle \qquad \text{for each } \eta \in \mathscr{D}(U_T) \tag{6.37}$$

Proof. On $\mathcal{H} \oplus \mathcal{H}$, one computes that

$$\mathscr{Cay} \circ \mathbb{J} \circ \mathscr{Cay}^{-1} = \Gamma : \mathcal{H} \oplus \mathcal{H} \to \mathcal{H} \oplus \mathcal{H} \qquad \eta \oplus \phi \mapsto (-\eta) \oplus \phi$$

Therefore, noting that any unitary transform commutes with taking orthogonal complements, we compute that

$$\mathscr{Cay}(\mathbb{J}\mathfrak{G}^\perp) = (\mathscr{Cay}\,\mathbb{J}\mathfrak{G})^\perp = (\Gamma\,\mathscr{Cay}\,\mathfrak{G})^\perp$$

and hence

$$\mathscr{Cay}(\mathrm{Ad}(\mathfrak{G})) = \{\alpha \oplus \beta \in \mathcal{H} \oplus \mathcal{H} : \langle \alpha | \eta \rangle = \langle \beta | \psi \rangle \text{ for each } \eta \oplus \psi \in \mathscr{Cay}(\mathfrak{G})\} \tag{6.38}$$

Eq. (6.37) follows by taking $\mathfrak{G} = \mathscr{G}(T)$. □

Proposition 6.7.5. *Let T be a closed Hermitian operator on $\mathcal{H}$ with Cayley transform U_T. Then*

$$\mathscr{Cay}(\mathscr{G}(T^*)) = \mathscr{G}(U_T) + (\mathrm{Rng}(T + \mathbf{i})^\perp \oplus 0) + (0 \oplus \mathrm{Rng}(T - \mathbf{i})^\perp) \tag{6.39}$$

Moreover, the three spaces $\mathscr{G}(U_T)$, $(\mathrm{Rng}(T + \mathbf{i})^\perp \oplus 0)$, $(0 \oplus \mathrm{Rng}(T - \mathbf{i})^\perp)$ are mutually orthogonal closed linear subspaces of $\mathcal{H} \oplus \mathcal{H}$. Therefore (6.39) is an (orthogonal) direct sum of Hilbert spaces.

[9] Von Neumann himself did not use the word "hyperextension" in his writings.

Proof. By (6.11), the projections of $\mathscr{C}ay(\mathscr{G}(T)) \equiv \mathscr{G}(U_T)$ onto the first and second components are $\mathscr{D}(U_T) = \mathrm{Rng}(T+\mathbf{i})$ and $\mathrm{Rng}(U_T) = \mathrm{Rng}(T-\mathbf{i})$, respectively. Thus, the three subspaces on the RHS of (6.39) are mutually orthogonal. They are clearly also closed. Moreover, one checks easily that any element in one of these three spaces satisfies the description of $\mathscr{C}ay(\mathscr{G}(T^*))$ in Lem. 6.7.4. This proves the relation "$\supset$" in (6.39).

To prove the relation "$\subset$", we choose any $\alpha \oplus \beta \in \mathscr{C}ay(\mathscr{G}(T^*))$ and prove that $\alpha \oplus \beta$ belongs to the RHS of (6.39). Since T is closed, by Thm. 6.4.16, $\mathrm{Rng}(T \pm \mathbf{i})$ is closed. Let $P_\pm \in \mathfrak{L}(\mathcal{H})$ be the projection of $\mathcal{H}$ onto $\mathrm{Rng}(T \pm \mathbf{i})$. Then

$$\alpha \oplus \beta = (P_+\alpha \oplus P_-\beta) + ((\mathbf{1}-P_+)\alpha \oplus 0) + (0 \oplus (\mathbf{1}-P_-)\beta)$$

where the second and third terms on the RHS belong to $(\mathrm{Rng}(T+\mathbf{i})^\perp \oplus 0)$ and $(0 \oplus \mathrm{Rng}(T-\mathbf{i})^\perp)$ respectively. Let $\alpha' = P_+\alpha$ and $\beta = P_-\beta$. Then $\alpha' \in \mathrm{Rng}(T+\mathbf{i})$ and $\beta' \in \mathrm{Rng}(T-\mathbf{i})$.

Since $\alpha' \oplus \beta' \in \mathscr{C}ay(\mathscr{G}(T^*))$ (because $\mathscr{C}ay(\mathscr{G}(T^*))$ is a linear subspace), the elements α', β' satisfy the description in Lem. 6.7.4, i.e., for each $\eta \in \mathscr{D}(U_T)$,

$$\langle \alpha' | \eta \rangle = \langle \beta' | U_T \eta \rangle$$

Since $\alpha' \oplus U_T\alpha'$ belongs to $\mathscr{G}(U_T) \subset \mathscr{C}ay(\mathscr{G}(T^*))$ and hence also satisfies the description in Lem. 6.7.4, for each $\eta \in \mathscr{D}(U_T)$ we have

$$\langle \alpha' | \eta \rangle = \langle U_T\alpha' | U_T \eta \rangle$$

and hence $\langle \beta' - U_T\alpha' | U_T\eta \rangle = 0$. Thus, the element $\beta' - U_T\alpha'$ (which belongs to $\mathrm{Rng}(U_T) = \mathrm{Rng}(T-\mathbf{i})$) is orthogonal to $\mathrm{Rng}(U_T)$. Hence $\beta' = U_T\alpha'$. This proves that $\alpha' \oplus \beta' \in \mathscr{G}(U_T)$, and hence $\alpha \oplus \beta$ belongs to the RHS of (6.39). □

Recall Def. 6.6.15 for the meaning of graph inner products.

Theorem 6.7.6. *Let T be a closed Hermitian operator on $\mathcal{H}$. Then, under the graph inner product of T^*, the Hilbert space $\mathscr{D}(T^*)$ admits an orthogonal decomposition into Hilbert subspaces*

$$\mathscr{D}(T^*) = \mathscr{D}(T) \oplus \mathrm{Rng}(T+\mathbf{i})^\perp \oplus \mathrm{Rng}(T-\mathbf{i})^\perp \tag{6.40}$$

Moreover, we have

$$T^*|_{\mathscr{D}(T)} = T \qquad T^*|_{\mathrm{Rng}(T+\mathbf{i})^\perp} = \mathbf{i} \qquad T^*|_{\mathrm{Rng}(T-\mathbf{i})^\perp} = -\mathbf{i} \tag{6.41}$$

Proof. Applying $\mathscr{C}ay^{-1} =$ (6.19) to Eq. (6.39) yields an orthogonal decomposition of $\mathscr{G}(T^*)$ into three Hilbert subspaces:

$$\begin{aligned}\mathscr{G}(T^*) =& \mathscr{G}(T) + \{-\mathbf{i}\eta \oplus \eta \in \mathcal{H}\oplus\mathcal{H} : \eta \in \mathrm{Rng}(T+\mathbf{i})^\perp\} \\ &+ \{\mathbf{i}\psi \oplus \psi \in \mathcal{H}\oplus\mathcal{H} : \psi \in \mathrm{Rng}(T-\mathbf{i})^\perp\}\end{aligned} \tag{6.42}$$

The unitary map $\Psi : \mathscr{D}(T^*) \to \mathscr{G}(T^*), \xi \mapsto \xi \oplus T^*\xi$ (where $\mathscr{D}(T^*)$ is equipped with the graph inner product) pulls the three spaces on the RHS of (6.42) back to those on the RHS of (6.40) respectively. Therefore, the Hilbert space decomposition (6.40) holds. Eq. (6.41) follows immediately from (6.42). □

Remark 6.7.7. By Prop. 6.6.11, we have

$$\mathrm{Ker}(T^* - \mathbf{i}) = \mathrm{Rng}(T + \mathbf{i})^\perp \qquad \mathrm{Ker}(T^* + \mathbf{i}) = \mathrm{Rng}(T - \mathbf{i})^\perp \tag{6.43}$$

This gives another proof of (6.41).

Corollary 6.7.8. *Let T be a Hermitian operator on $\mathcal{H}$. Then T is self-adjoint iff*

$$\mathrm{Rng}(T + \mathbf{i}) = \mathrm{Rng}(T - \mathbf{i}) = \mathcal{H}$$

Proof. Since self-adjoint operators are closed (by Thm. 6.6.9), and since Hermitian operators satisfying $\mathrm{Rng}(T+\mathbf{i}) = \mathrm{Rng}(T-\mathbf{i}) = \mathcal{H}$ are also closed (by Thm. 6.4.16), the two conditions can be compared under the assumption that T is a closed Hermitian operator. Then the equivalence follows immediately from Thm. 6.7.6. □

Corollary 6.7.9. *Let T be a Hermitian operator on $\mathcal{H}$. Then T is essentially self-adjoint iff both $\mathrm{Rng}(T + \mathbf{i})$ and $\mathrm{Rng}(T - \mathbf{i})$ are dense in $\mathcal{H}$.*

Proof. This follows immediately from Cor. 6.7.8, together with the fact that $\mathrm{Rng}(\overline{T} \pm \mathbf{i})$ is the closure of $\mathrm{Rng}(T \pm \mathbf{i})$ (due to Thm. 6.4.16 or Eq. (6.12)). □

Corollary 6.7.10. *Let T be an essentially self-adjoint operator on $\mathcal{H}$. Then $\overline{T}$ is the unique closed Hermitian extension of T, and $\overline{T}$ is self-adjoint.*

Consequently, if T is self-adjoint, then any Hermitian extension S must satisfy $\overline{S} = \overline{T}$.

First proof. Clearly $\overline{T}$ is a self-adjoint extension. By Cor. 6.7.9, the Cayley transform of T is a unitary map $U : \mathscr{D}(U) \to \mathrm{Rng}(U)$ where $\mathscr{D}(T) = \mathrm{Rng}(T + \mathbf{i})$ and $\mathrm{Rng}(U) = \mathrm{Rng}(T - \mathbf{i})$ are dense linear subspaces of $\mathcal{H}$. Thus U has at most one extension to a bounded linear map on a closed linear subspace of $\mathcal{H}$ containing $\mathscr{D}(U)$. Therefore, by Thm. 6.3.5 and 6.4.16, T has at most one closed Hermitian extension. □

Second proof. Clearly $\overline{T}$ is self-adjoint. If S is a closed Hermitian extension of T, then $T \subset S$ implies $\overline{T} \subset S$. Thus $S^* \subset \overline{T}^*$, and hence $\overline{T} \subset S \subset S^* \subset \overline{T}^*$. As $\overline{T} = \overline{T}^*$, we conclude $\overline{T} = S$. □

Corollary 6.7.11. *Let T be a Hermitian operator on $\mathcal{H}$ with deficiency indices (n_+, n_-). Then T has a self-adjoint extension iff $n_+ = n_-$.*

Proof. Any self-adjoint extension of T is closed, and hence extends $\overline{T}$. Therefore, replacing T with $\overline{T}$ (and noting that this procedure does not change the deficiency indices, cf. Thm. 6.4.16), it suffices to assume that T is closed.

By Thm. 6.5.1, the equality $n_+ = n_-$ holds iff T has a closed Hermitian extension $\widehat{T}$ satisfying $\mathrm{Rng}(\widehat{T}+\mathbf{i}) = \mathrm{Rng}(\widehat{T}-\mathbf{i}) = \mathcal{H}$. By Cor. 6.7.8, this later condition is equivalent to that T has a self-adjoint extension. □

6.7.3 Thm. 6.7.6 in the context of differential equations

In this subsection, we sketch an interpretation of Thm. 6.7.6 in the context of differential equations. First, we make the following observation:

Remark 6.7.12. In the setting of Thm. 6.7.6, we have

$$\mathrm{Ker}((T^*)^2+1) = \mathrm{Rng}(T+\mathbf{i})^\perp + \mathrm{Rng}(T-\mathbf{i})^\perp \tag{6.44}$$

Therefore, (6.40) can be written equivalently as

$$\mathscr{D}(T^*) = \mathscr{D}(T) \oplus \mathrm{Ker}((T^*)^2+1) \tag{6.45}$$

Proof. In view of Rem. 6.7.7, we clearly have "$\supset$" in (6.44). To prove "$\subset$", by (6.40), it suffices to show

$$\mathrm{Ker}((T^*)^2+1) \cap \mathscr{D}(T) = 0$$

Pick any ξ on the LHS. Then $T^*T\xi + \xi = 0$, and hence $\langle T\xi|T\xi\rangle + \langle \xi|\xi\rangle = 0$. This proves $\xi = 0$. □

Now, let $\Omega \subset \mathbb{R}^N$ be bounded and open with sufficiently regular (e.g. smooth) boundary $\partial\Omega$. Consider T_0 as a "Dirac operator" on $\mathcal{H} = L^2(\Omega, m)^{\oplus k}$ with domain $C_c^\infty(\Omega)^{\oplus k}$. This mean that T_0 is Hermitian and satisfies

$$\|T_0\xi\|^2 = -\langle \xi|\Delta\xi\rangle \qquad \text{for each } \xi \in C_c^\infty(\Omega)^{\oplus k}$$

For instance, when $N = k = 1$, the operator $-\mathbf{i}\frac{d}{dx}$ is a Dirac operator. For general N, let $\mathcal{H}$ be the space of Lebesgue-L^2 functions from Ω to $\bigoplus_{j\in\mathbb{N}} \bigwedge^j \mathbb{R}^N$, with $\mathbb{R}^N$ interpreted as the cotangent space of Ω. (In this case, $k = 2^N$.) The Dirac operator is defined by $T_0 = d + \delta$, where d is the exterior differential, and δ is the restriction of the adjoint d^* to the subspace of compactly supported smooth sections. The corresponding $\Delta = -(d+\delta)^2$ is called the (negative) Hodge-Laplacian.

In this setting, T_0 plays a role analogous to ∇ in Subsec. 6.6.3. Thus, as in Exp. 6.6.16, one can use the graph inner products of $T := \overline{T_0}$ and T^* to define Hilbert spaces

$$H_0^1(\Omega) = \mathscr{D}(T) \qquad H^1(\Omega) = \mathscr{D}(T^*)$$

Then Rem. 6.7.12 asserts that

$$H^1(\Omega) \ominus H_0^1(\Omega) = \mathrm{Ker}((T^*)^2 + 1) \tag{6.46}$$

On the other hand, one can define the Sobolev space $H^{\frac{1}{2}}(\partial\Omega)$ in an appropriate way, show that there exists a surjective bounded linear map (the trace map)

$$H^1(\Omega) \to H^{\frac{1}{2}}(\partial\Omega)$$

defined by "restriction to the boundary", and show that this map has kernel $H_0^1(\Omega)$; see e.g. [Eva, Sec. 5.5] for details. It follows that the trace map induces a bounded linear bijection[10]

$$\mathrm{Ker}((T^*)^2 + 1) \to H^{\frac{1}{2}}(\partial\Omega) \tag{6.47}$$

The LHS of (6.47) can be viewed as the space of $f \in H^1(\Omega)$ satisfying $(-\Delta + 1)f = 0$. Therefore, (6.47) establishes a bijection between $g \in H^{\frac{1}{2}}(\partial\Omega)$ and the solutions $f \in H^1(\Omega)$ of the Dirichlet problem

$$(-\Delta + 1)f = 0 \qquad f|_{\partial\Omega} = g$$

6.8 Example: self-adjoint extensions and the adjoint of $-\mathbf{i}\frac{d}{dx}$

Let $I = (a, b)$ be an open interval in $\mathbb{R}$ where $-\infty < a < b < +\infty$. Let T_0 be the unbounded (Hermitian) operator

$$T_0 = -\mathbf{i}\frac{d}{dx} : L^2(I, m) \to L^2(I, m) \qquad \mathscr{D}(T_0) = C_c^\infty(I)$$

Let $T = \overline{T_0}$. Recall from Subsec. 6.6.3 that

$$H_0^1(I) = \mathscr{D}(T) \qquad H^1(I) = \mathscr{D}(T^*)$$

Recall from Thm. 6.6.9 that $T^* = T_0^*$. In this section, we determine T^* and the self-adjoint extensions of T.

Due to Thm. 6.7.6 (and Rem. 6.7.7), to characterize T^* it suffices to determine $\mathrm{Ker}(T^* \mp \mathbf{i})$. For that purpose, let us first characterize $\mathrm{Ker}(T^*)$.

Lemma 6.8.1. *A function $f \in L^2(I, m)$ belongs to $\mathrm{Ker}(T^*) = \mathrm{Ker}(T_0^*)$ iff f is a.e. constant.*

[10] The inverse of the map (6.47) is also bounded by the inverse mapping theorem. Hence (6.47) implements an "equivalence" of Hilbert spaces, though not necessarily an isomorphism (i.e. unitary equivalence) of Hilbert spaces.

Proof. If f is a constant, then for each $g \in C_c^\infty(I)$, the obvious identity $\int g' = 0$ implies that $\langle f|T_0 g\rangle = 0$. Hence $f \in \mathscr{D}(T_0^*)$ and $T_0^* f = 0$. The converse follows from the following more general Lem. 6.8.2. □

Lemma 6.8.2. *Let $\Lambda : C_c^\infty(I) \to \mathbb{C}$ be a linear functional such that $\Lambda(g') = 0$ for each $g \in C_c^\infty(I)$. Then there exists $c \in \mathbb{C}$ such that $\Lambda(h) = c\int_I hdm$ for each $h \in C_c^\infty(I)$.*

Roughly speaking, this lemma says that any distribution with zero derivative must be a constant.

Proof. Define a linear functional $\Gamma : C_c^\infty(I) \to \mathbb{C}$ by $\Gamma(h) = \int_I h$. Then $\mathrm{Ker}(\Gamma) \subset \mathrm{Ker}(\Lambda)$. Indeed, if $h \in C_c^\infty(I)$ and $\int_I h = 0$, then h is the derivative of a function in $C_c^\infty(I)$, and hence $\Lambda(h) = 0$.

Write $V = C_c^\infty(I)$. Since $\mathrm{Rng}(\Gamma) = \mathbb{C}$, by linear algebra, Γ descends to a linear isomorphism $\Gamma' : V/\mathrm{Ker}(\Gamma) \xrightarrow{\simeq} \mathbb{C}$ (in particular $\dim V/\mathrm{Ker}(\Gamma) = 1$), and Λ descends to a linear map $\Lambda' : V/\mathrm{Ker}(\Gamma) \to \mathbb{C}$. Thus $\Lambda' = c\Gamma'$ for some $c \in \mathbb{C}$, and hence $\Lambda = c\Gamma$. □

Proposition 6.8.3. *We have*

$$\mathrm{Ker}(T^* - \mathbf{i}) = \mathrm{Span}\{e^{-x}\} \qquad \mathrm{Ker}(T^* + \mathbf{i}) = \mathrm{Span}\{e^{x}\} \tag{6.48}$$

In particular, the Hermitian operator T has deficiency indices $(1,1)$.

Proof. Clearly e^{-x} belongs to $\mathrm{Ker}(T^* - \mathbf{i})$. Conversely, asume that $f \in \mathrm{Ker}(T^* - \mathbf{i})$. In particular, $f \in \mathscr{D}(T^*)$. Let $g(x) = e^x f$. Then for each $h \in C_0^\infty(I)$, we compute

$$\begin{aligned}\langle g|T_0 h\rangle &= -\mathbf{i}\langle e^x f|h'\rangle = -\mathbf{i}\langle f|e^x h'\rangle = -\mathbf{i}\langle f|(e^x h)'\rangle + \mathbf{i}\langle f|e^x h\rangle \\ &= \langle f|T_0(e^x h)\rangle + \mathbf{i}\langle f|e^x h\rangle = \langle e^x T_0^* f|h\rangle + \mathbf{i}\langle e^x f|h\rangle = \langle e^x (T^* - \mathbf{i}) f|h\rangle\end{aligned}$$

This shows that $g \in \mathscr{D}(T^*) = \mathscr{D}(T_0^*)$ and $T^* g = e^x(T^* - \mathbf{i})f$. Since $f \in \mathrm{Ker}(T^* - \mathbf{i})$, we conclude $T^* g = 0$, and hence g is a constant by Lem. 6.8.1. This proves the first identity in (6.48). The second identity can be proved in the same way. □

Corollary 6.8.4. *The Hilbert space $H^1(I)$ has an orthogonal decomposition into Hilbert subspaces*

$$H^1(I) = H_0^1(I) \oplus \mathrm{Span}\{e^{-x}\} \oplus \mathrm{Span}\{e^{x}\} \tag{6.49}$$

Moreover, we have $T^* e^{-x} = \mathbf{i}e^{-x}$ and $T^* e^x = -\mathbf{i}e^x$.

Proof. This follows immediately from Prop. 6.8.3 and Thm. 6.7.6. □

Note that e^{-x} and e^x are orthogonal in the inner product of $H^1(I)$, rather than in the original one of $L^2(I, m)$.

Theorem 6.8.5. *Let $AC(\overline{I})$ be the set of absolutely continuous functions $\overline{I} \to \mathbb{C}$. Then*

$$H_0^1(I) = \{f \in AC(\overline{I}) : f' \in L^2(I, m), f(a) = f(b) = 0\} \tag{6.50a}$$

$$H^1(I) = \{f \in AC(\overline{I}) : f' \in L^2(I, m)\} \tag{6.50b}$$

Moreover, for each $f \in H^1(I)$ we have $T^* f = -\mathbf{i} f'$ (a.e.).

See Pb. 6.7 for a different proof strategy.

Proof. Step 1. Let $\mathcal{E}$ and $\mathcal{F}$ be the RHS of (6.50a) and (6.50b) respectively. In this step, we prove $H_0^1(I) \subset \mathcal{E}$.

Choose any $f \in H_0^1(I)$. Define $\widetilde{f} : \overline{I} \to \mathbb{C}$ by $\widetilde{f}(x) = \int_a^x \mathbf{i}Tf$. Then $\widetilde{f} \in AC(\overline{I})$, and $\widetilde{f}(a) = 0$. If we can prove $\widetilde{f}(b) = 0$ and $f = \widetilde{f}$, then $f \in \mathcal{E}$, and hence the proof of $H_0^1(I) \subset \mathcal{E}$ is finished.

Since T is the closure of T_0, there exists a sequence (f_n) in $C_c^\infty(I)$ such that $\|f - f_n\|_2 \to 0$ and $\|\mathbf{i}Tf - f_n'\|_2 \to 0$. Clearly $\int_a^b f_n' = 0$. Since $\|\mathbf{i}Tf - f_n'\|_1 \to 0$ (by Cauchy-Schwarz), we conclude that $\widetilde{f}(b) = \int_a^b \mathbf{i}Tf = \lim_n \int_a^b f_n' = 0$.

The fact that $\|\mathbf{i}Tf - f_n'\|_1 \to 0$ also implies that $\lim_n \int_a^x f_n' = \lim_n f_n(x)$ converges uniformly to $\int_a^x \mathbf{i}Tf = \widetilde{f}(x)$ over $x \in I$. Therefore $\lim_n \|\widetilde{f} - f_n\|_{L^2} = 0$, and hence $f = \widetilde{f}$.

Step 2. In the above step, we see that $\widetilde{f} = f$. By the fundamental theorem of algebra (for the Lebesgue integral), we have $\mathbf{i}Tf = \widetilde{f}'$ a.e.. Thus $\mathbf{i}Tf = f'$ a.e.. This prove $\mathbf{i}T^* f = f'$ (a.e) for all $f \in H_0^1(I)$. Since this relation also holds for $f = e^{\pm x}$, it follows from Cor. 6.8.4 that $\mathbf{i}T^* f = f'$ for all $f \in H^1(I)$.

Step 3. In this step, we prove $H^1(I) = \mathcal{F}$. The inclusion $H^1(I) \subset \mathcal{F}$ follows from $H_0^1(I) \subset \mathcal{E}$ and Cor. 6.8.4.

For each $f \in \mathcal{F}$, note that if $g \in C_c^\infty(I)$ then $\overline{f}g \in AC(\overline{I})$, and $(\overline{f}g)' = \overline{f}'g + \overline{f}g'$. The fundamental theorem of calculus implies $\langle f|g'\rangle = -\langle f'|g\rangle$. This shows that $f \in H^1(I)$ and $\mathbf{i}T^* f = f'$. We have thus proved $H^1(I) \supset \mathcal{F}$.

Step 4. Finally, we prove $H_0^1(I) \supset \mathcal{E}$. Choose any $f \in \mathcal{E}$. Since $f \in H^1(I)$ (because $\mathcal{E} \subset \mathcal{F} = H^1(I)$), by Cor. 6.8.4, we can write $f = f_1 + f_2$ where $f_1 \in H_0^1(I)$ and $f_2 = c_- e^{-x} + c_+ e^x$ for some $c_\pm \in \mathbb{C}$. By the definition of $\mathcal{E}$, we have $f(a) = f(b) = 0$. By Step 1, we have $f_1 \in \mathcal{E}$ and hence $f_1(a) = f_1(b) = 0$. Hence $f_2(a) = f_2(b) = 0$, which implies $c_- = c_+ = 0$. Therefore $f = f_1 \in H_0^1(I)$. This finishes the proof of $H_0^1(I) \supset \mathcal{E}$. □

Theorem 6.8.6. *There is a bijection between:*

(1) A self-adjoint extension $\widehat{T}$ of T.

(2) A number $\lambda \in \mathbb{S}^1$.

These two objects are related by $\widehat{T} = T^*|_{\mathscr{D}_\lambda}$ where

$$\mathscr{D}_\lambda = \{f \in H^1(I) : f(b) = \lambda f(a)\}$$

This theorem suggests that different self-adjoint extensions of a Hermitian operator correspond to different (Hermitian) boundary conditions.

Proof. We assume for simplicity that $I = (-1, 1)$ so that e^{-x} and e^x have the same L^2-norm. Thm. 6.5.1 (together wit Prop. 6.8.3) implies that the self-adjoint extensions $\widehat{T}$ of T correspond bijectively to $z \in \mathbb{S}^1$ (with the Cayley transform of $\widehat{T}$ sending e^{-x} to ze^x), and that

$$\mathscr{D}(\widehat{T}) = \mathscr{D}(T) + \mathrm{Span}(ze^x - e^{-x}) = H_0^1(I) + \mathrm{Span}(ze^x - e^{-x})$$

Comparing this with the description of $H^1(I)$ in (6.49) (and noting that functions in $H_0^1(I)$ are continuous on $\overline{I}$ and vanishing at a, b, cf. Thm. 6.8.5), it is clear that $\mathscr{D}(\widehat{T})$ equals

$$\mathscr{D}_\lambda = \{f \in H_0^1(I) + \mathbb{C}e^{-x} + \mathbb{C}e^x : f(-1) = \lambda f(1)\}$$

where $z \in \mathbb{S}^1$ and $\lambda \in \mathbb{S}^1$ are related by

$$\lambda = \frac{ze - e^{-1}}{ze^{-1} - e} \qquad z = \frac{\lambda e - e^{-1}}{\lambda e^{-1} - e}$$

Note that if $z \in \mathbb{S}^1$, then $\lambda \in \mathbb{S}^1$ since $\sqrt{z}e - \sqrt{z^{-1}}e^{-1}$ is the negative conjugate of $\sqrt{z}e^{-1} - \sqrt{z^{-1}}e$. Similarly, if $\lambda \in \mathbb{S}^1$ then $z \in \mathbb{S}^1$. □

6.9 Spectral theorem for strongly commuting self-adjoint operators

Let $\mathcal{H}$ be a Hilbert space.

6.9.1 Strong commutativity and simultaneous measurement

The goal of this section is to prove that an unbounded operator T on $\mathcal{H}$ is self-adjoint (equivalently, Hermitian and $\mathrm{Rng}(T + \mathbf{i}) = \mathrm{Rng}(T - \mathbf{i}) = \mathcal{H}$, cf. Cor. 6.7.8) iff T admits a spectral decomposition, i.e, T is unitarily equivalent to the multiplication operator of a real-valued measurable function. In fact, we will prove a more general result: the spectral theorem for a finite family of strongly commuting self-adjoint operators. The motivation for considering such operators requires some explanation.

The best-known phenomenon in quantum mechanics is the uncertainty principle: two observables (e.g. momentum $-\mathbf{i}\frac{d}{dx}$ and position x) cannot be simultaneously measured exactly. In finite-dimensional linear algebra, the possibility of measuring two observables simultaneously and with perfect precision is captured by the commutativity of Hermitian matrices, since commuting Hermitian matrices can be simultaneously diagonalized.

In infinite dimensions, greater care is needed. The first caveat is that if we restrict to bounded operators, then the spectral theorem shows that commutativity of two bounded self-adjoint operators does not mean that their observables can be **simultaneously measured with absolute precision**. Rather, as explained in Subsec. 6.1.5, they can only be **simultaneously measured with arbitrarily prescribed accuracy**.

More precisely, suppose $A, B \in \mathfrak{L}(\mathcal{H})$ are commuting self-adjoint operators. For each $\Omega, O \subset \mathbb{R}$, define spectral projections

$$E(\Omega) = \chi_\Omega(A) \qquad F(O) = \chi_O(B)$$

Then $E(\Omega)$ commutes with $F(O)$, so

$$P(\Omega, O) := E(\Omega)F(O)$$

equals the projection onto $E(\Omega)\mathcal{H} \cap F(O)\mathcal{H}$, the space of all states $\xi \in \mathcal{H}$ in which the measurements of A yield values in Ω, and the measurements of B yield values in O. Now fix $\varepsilon > 0$ (interpreted as a prescribed accuracy), we partition $\mathbb{R}$ into intervals

$$\mathbb{R} = \bigsqcup_{n \in \mathbb{Z}} \Omega_n \qquad \text{where } \Omega_n = (n\varepsilon, (n+1)\varepsilon]$$

Then for each unit vector $\xi \in \mathcal{H}$, the quantity $\langle \xi | P(\Omega_m, \Omega_n)\xi \rangle$ is the probability that the simultaneous measurements of A and B in the state ξ yield outcomes in Ω_m and Ω_n respectively—that is, a simultaneous measurement of A and B with prescribed accuracy ε. The identity

$$1 = \langle \xi | \xi \rangle = \sum_{m,n} \langle \xi | P(\Omega_m, \Omega_n)\xi \rangle \tag{6.51}$$

shows that the probabilities of all such ε-accurate simultaneous measurements sum to 1. By contrast, if A and B do not commute, then $E(\Omega_m)$ and $F(\Omega_n)$ may fail to commute. In that case, (6.51) may not hold, if we define $P(\Omega_m, \Omega_n)$ to be the projection onto $E(\Omega_m)\mathcal{H} \cap F(\Omega_n)\mathcal{H}$.

The second caveat arises when extending this reasoning to unbounded self-adjoint operators A, B. In this case, the appropriate mathematical condition for **simultaneous measurement with arbitrarily prescribed accuracy** is that the spectral projections of A commute with those of B. Equivalently, as we will see, the Cayley transforms of A and B must commute. We call this property **strong commutativity**.

However, strong commutativity is not equivalent to ordinary commutativity $AB = BA$. For example, any self-adjoint operator A commutes strongly with the zero operator, even though we only have $0A \subset A0$ (where $A0$ equals 0) but not $0A = A0$ if $\mathscr{D}(A) \neq \mathcal{H}$. Even if one adopts a weaker notion of commutativity—namely, that there exists a common core $\mathscr{D}_0$ for A and B such that $A\mathscr{D}_0 \subset \mathscr{D}_0$, $B\mathscr{D}_0 \subset \mathscr{D}_0$, and $AB|_{\mathscr{D}_0} = BA|_{\mathscr{D}_0}$—this still does not guarantee strong commutativity (cf. [RS-1, Sec. VIII.5]).

Due to this reason, strong commutativity, rather than the ordinary relation $AB = BA$, is the natural condition for the spectral theorem of unbounded self-adjoint operators. For von Neumann's detailed discussion of the connection between strong commutativity and simultaneous measurement with prescribed accuracy, see [vN32a] (especially Sec. III.3).

6.9.2 Spectral theorem for strongly commuting self-adjoint operators

Definition 6.9.1. Two self-adjoint operators A, B on $\mathcal{H}$ are said to **commute strongly** if their Cayley transforms U_A, U_B commute (and hence commute adjointly by Exp. 5.7.8).

Example 6.9.2. Let $(X, \mathfrak{M})$ be a measurable space. Let $(\mu_\alpha)_{\alpha\in\mathscr{I}}$ be a family of measures on $\mathfrak{M}$. Let $f_1, \dots, f_N : X \to \mathbb{R}$ be measurable. Let $\mathcal{H} = \bigoplus_{\alpha\in\mathscr{I}} L^2(X, \mu_\alpha)$. Then the multiplication operators $\mathbf{M}_{f_1}, \dots, \mathbf{M}_{f_N}$ (cf. Exp. 6.2.16) are strongly-commuting self-adjoint operators on $\mathcal{H}$. Moreover, the Cayley transform of $\mathbf{M}_{f_j}$ is $\mathbf{M}_{(f_j-\mathbf{i})/(f_j+\mathbf{i})}$.

Proof. If $f = f_j$, then $\mathbf{M}_f$ is Hermitian, because $\langle\xi|\mathbf{M}_f\xi\rangle = \sum_\alpha \int f|\xi_\alpha|^2 d\mu_\alpha \in \mathbb{R}$ for each $\xi = \oplus_\alpha \xi_\alpha \in \mathscr{D}(\mathbf{M}_f)$. It is also clear that $\mathbf{M}_f \pm \mathbf{i} = \mathbf{M}_{f\pm\mathbf{i}}$. By Exp. 6.2.23, $\mathbf{M}_{f\pm\mathbf{i}}$ is injective with dense range, and $\mathbf{M}_{f\pm\mathbf{i}}^{-1} = \mathbf{M}_{1/(f\pm\mathbf{i})}$. In particular, $\mathrm{Rng}(\mathbf{M}_{f\pm\mathbf{i}}) = \mathscr{D}(\mathbf{M}_{1/(f\pm\mathbf{i})}) = \mathcal{H}$. This proves that $\mathrm{Rng}(\mathbf{M}_f \pm \mathbf{i}) = \mathcal{H}$. Therefore, by Cor. 6.7.8, $\mathbf{M}_f$ is self-adjoint.

By Rem. 6.2.17, the Cayley transform of $\mathbf{M}_f$ equals

$$(\mathbf{M}_f - \mathbf{i})(\mathbf{M}_f + \mathbf{i})^{-1} = \mathbf{M}_{f-\mathbf{i}}\mathbf{M}_{1/(f+\mathbf{i})} = \mathbf{M}_u$$

where $u = (f - \mathbf{i})/(f + \mathbf{i}) : X \to \mathbb{S}^1$ is measurable. This proves that the Cayley transforms of $\mathbf{M}_{f_1}, \dots, \mathbf{M}_{f_N}$ are multiplication operators of unitary measurable functions, and hence commute. Thus $\mathbf{M}_{f_1}, \dots, \mathbf{M}_{f_N}$ commute strongly. $\square$

To prove that any finite family of strongly commuting self-adjoint operators are unitarily equivalent to multiplication operators, we need the following lemma. Recall Def. 1.6.8 for the notions of measurable isomorphisms and pullback measures.

Lemma 6.9.3. *Let $(X,\mathfrak{M})$ and $(Y,\mathfrak{N})$ be measurable spaces. Let $\phi : X \to Y$ be a measurable isomorphism. Let $(\nu_\alpha)_{\alpha\in\mathscr{I}}$ be a family of measures on $\mathfrak{N}$. Then the map*

$$\Gamma : \bigoplus_{\alpha\in\mathscr{I}} L^2(Y,\nu_\alpha) \xrightarrow{\simeq} \bigoplus_{\alpha\in\mathscr{I}} L^2(X,\phi^*\nu_\alpha) \qquad \oplus_\alpha \eta_\alpha \mapsto \oplus_\alpha(\eta_\alpha\circ\phi)$$

is unitary. Moreover, if $g : Y \to \mathbb{C}$ is measurable, then the identity

$$\Gamma \mathbf{M}_g \Gamma^{-1} = \mathbf{M}_{g\circ\phi}$$

holds as bounded linear operators on $\bigoplus_{\alpha\in\mathscr{I}} L^2(X,\phi^*\nu_\alpha)$. In particular, $\Gamma\mathscr{D}(\mathbf{M}_g) = \mathscr{D}(\mathbf{M}_{g\circ\phi})$.

Proof. The fact that Γ is an isometry follows from (1.31). It is easy to see that Γ is a bijection, and that $\Gamma\mathbf{M}_f\Gamma^{-1} = \mathbf{M}_{f\circ\phi}$ holds. $\square$

Theorem 6.9.4. *Let $T_1,\dots,T_N$ be strongly commuting self-adjoint operators on $\mathcal{H}$. Then there exists a family $(\mu_\alpha)_{\alpha\in\mathscr{I}}$ of finite Borel measures on $\mathbb{R}^N$, together with a unitary map*

$$\Phi : \mathcal{H} \xrightarrow{\simeq} \bigoplus_{\alpha\in\mathscr{I}} L^2(\mathbb{R}^N,\mu_\alpha)$$

such that for each $1\leqslant j\leqslant N$, we have

$$\Phi T_j \Phi^{-1} = \mathbf{M}_{x_j} \tag{6.52}$$

as unbounded operators on $\bigoplus_{\alpha\in\mathscr{I}} L^2(\mathbb{R}^N,\mu_\alpha)$.

Here, x_i denotes the i-th coordinate function of $\mathbb{R}^N$, i.e., the map sending $(a_1,\dots,a_N)\in\mathbb{R}^N$ to a_i.

Proof. Step 1. Let U_j be the Cayley transform of T_j, which is a unitary operator on $\mathcal{H}$ by Cor. 6.7.8. By Prop. 5.8.16, we have $\mathrm{Sp}(T_j)\subset\mathbb{T}\equiv\mathbb{S}^1$. Since $U_1,\dots,U_N$ commute adjointly, by the spectral Thm. 5.10.22, there exists a family $(\nu_\alpha)_{\alpha\in\mathscr{I}}$ of finite Borel measures on $\mathbb{T}^N$, together with a unitary map

$$\Psi : \mathcal{H} \to \bigoplus_{\alpha\in\mathscr{I}} L^2(\mathbb{T}^N,\nu_\alpha)$$

such that $\Psi U_j \Psi^{-1} = \mathbf{M}_{z_j}$ where $z_j : \mathbb{T}^N \to \mathbb{T}$ is the i-th coordinate function of $\mathbb{T}^N$.

Step 2. In this step, we reduce $\mathbb{T}^N$ to $(\mathbb{T}\backslash\{1\})^N$. Recall from Lem. 6.3.4 that 1 is not an eigenvalue of U_j. Hence 1 is not an eigenvalue of $\mathbf{M}_{z_j}$. Therefore, for each $\alpha\in\mathscr{I}$, we have $\nu_\alpha(z_j^{-1}(1)) = 0$. (Otherwise, there exists α such that $\chi_{z_j^{-1}(1)}$ is a non-zero vector of $L^2(\mathbb{T}^N,\nu_\alpha)$, but this vector is an eigenvector of $\mathbf{M}_{z_j}$ with eigenvalue 1.)

We have thus proved that ν_α vanishes on $\bigcup_j z_j^{-1}(1)$, the complement of $(\mathbb{T}\backslash\{1\})^N$. Therefore, we can restrict each ν_α to $(\mathbb{T}\backslash\{1\})^N$ so that the unitary map Ψ in Step 1 becomes

$$\Psi : \mathcal{H} \xrightarrow{\simeq} \bigoplus_{\alpha\in\mathscr{I}} L^2((\mathbb{T}\backslash\{1\})^N, \nu_\alpha)$$

and that

$$\Psi U_j \Psi^{-1} = \mathbf{M}_{z_j} \tag{6.53}$$

where

$$z_j : Y = (\mathbb{T}\backslash\{1\})^N \to \mathbb{T}$$

is the i-th coordinate function of $\mathbb{T}\backslash\{1\}$.

Step 3. Define a homeomorphism

$$u_j = \frac{x_j - \mathbf{i}}{x_j + \mathbf{i}} : \mathbb{R} \xrightarrow{\simeq} \mathbb{T}\backslash\{1\}$$

Applying Lem. 6.9.3 to the homeomorphism

$$(u_1, \dots, u_N) : \mathbb{R}^N \to (\mathbb{T}\backslash\{1\})^N$$

we obtain a unitary map

$$\Gamma : \bigoplus_{\alpha\in\mathscr{I}} L^2((\mathbb{T}\backslash\{1\})^N, \nu_\alpha) \xrightarrow{\simeq} \bigoplus_{\alpha\in\mathscr{I}} L^2(\mathbb{R}^N, \mu_\alpha)$$

where $\mu_\alpha = (u_1, \dots, u_N)^*\nu_\alpha$ is a finite Borel measure on $\mathbb{R}^N$, and

$$\Gamma \mathbf{M}_f \Gamma^{-1} = \mathbf{M}_{f\circ(u_1,\dots,u_N)} \qquad \text{for each Borel function } f : (\mathbb{T}\backslash\{1\})^N \to \mathbb{C}$$

Therefore $\Gamma \mathbf{M}_{z_j} \Gamma^{-1} = \mathbf{M}_{u_j} = \mathbf{M}_{(x_j - \mathbf{i})/(x_j + \mathbf{i})}$. This relation, together with (6.53), implies

$$\Phi U_j \Phi^{-1} = \mathbf{M}_{(x_j - \mathbf{i})/(x_j + \mathbf{i})} \tag{6.54}$$

where the unitary map Φ is defined to be

$$\Phi = \Gamma \circ \Psi : \mathcal{H} \xrightarrow{\simeq} \bigoplus_{\alpha\in\mathscr{I}} L^2(\mathbb{R}^N, \mu_\alpha)$$

Since U_j is the Cayley transform of T_j, the LHS of (6.54) is the Cayley transform of $\Phi T_j \Phi^{-1}$. By Exp. 6.9.2, the RHS of (6.54) is the Cayley transform of $\mathbf{M}_{x_j}$. This proves $\Phi T_j \Phi^{-1} = \mathbf{M}_{x_j}$. $\square$

6.10 Borel functional calculus and the joint spectrum $\mathrm{Sp}(T_\bullet)$

Let $\mathcal{H}$ be a Hilbert space.

The first goal of this section is to establish the Borel functional calculus for strongly-commuting self-adjoint operators. We remark that the notion of strong commutativity can be extended to arbitrary closed operators. Moreover, the spectral theorem in the form of multiplication operators, as well as the Borel functional calculus, can be established for strongly-commuting (unbounded) closed normal operators. This topic, however, lies beyond the scope of the present course. Interested readers are referred to [Gui-S] for further details.

6.10.1 Borel functional calculus

Recall that if X is a topological space, then $\mathscr{Bor}(X)$ (resp. $\mathscr{Bor}_{\mathrm{b}}(X)$) denotes the space of Borel functions (resp. bounded Borel functions) $X \rightarrow \mathbb{C}$. The main goal of this subsection is to prove the following theorem.

Theorem 6.10.1. *Let $T_1, \ldots, T_N$ be strongly commuting self-adjoint operators on $\mathcal{H}$. Then there exists a unique map*

$$\pi_{T_\bullet}: \mathscr{Bor}(\mathbb{R}^N) \rightarrow \{\textit{unbounded operators on } \mathcal{H}\}$$

satisfying the following three properties:

(1) The map $\pi_{T_\bullet}$ restricts to a normal unitary representation[11]

$$\pi_{T_\bullet}|_{\mathscr{Bor}_{\mathrm{b}}(\mathbb{R}^N)}: \mathscr{Bor}_{\mathrm{b}}(\mathbb{R}^N) \rightarrow \mathfrak{L}(\mathcal{H})$$

(2) For each $1 \leqslant j \leqslant N$, if $x_j: \mathbb{R}^N \rightarrow \mathbb{R}$ denotes the j-th coordinate function of $\mathbb{R}^N$, then

$$\pi_{T_\bullet}\Big(\frac{x_j - \mathbf{i}}{x_j + \mathbf{i}}\Big) = \frac{T_j - \mathbf{i}}{T_j + \mathbf{i}}$$

(3) For each $\xi \in \mathcal{H}$ and $f \in \mathscr{Bor}(\mathbb{R}^N)$, if μ_ξ denotes the finite Borel measure associated to ξ and $\pi_{T_\bullet}$ (cf. Rem. 5.5.9), then

$$\xi \in \mathscr{D}(\pi_{T_\bullet}(f)) \qquad \Longleftrightarrow \qquad \int_{\mathbb{R}^N} |f|^2 d\mu_\xi < +\infty \tag{6.55}$$

Moreover, if $\xi \in \mathscr{D}(\pi_{T_\bullet}(f))$, then [12]

$$\langle \xi | \pi_{T_\bullet}(f)\xi \rangle = \int_{\mathbb{R}^N} f d\mu_\xi \tag{6.56}$$

[11] Recall from Def. 5.5.8 that "normal" means that if (f_n) is an increasing sequence in $\mathscr{Bor}_{\mathrm{b}}(\mathbb{R}^N)$ converging to $f \in \mathscr{Bor}_{\mathrm{b}}(\mathbb{R}^N)$, then $\lim_n \langle \xi | \pi_{T_\bullet}(f_n)\xi \rangle = \langle \xi | \pi_{T_\bullet}(f)\xi \rangle$ for each $\xi \in \mathcal{H}$.

[12] Note that if $\int_{\mathbb{R}^N} |f|^2 d\mu_\xi < +\infty$, then $\int_{\mathbb{R}^N} |f| d\mu_\xi < +\infty$ by Cauchy-Schwarz and the fact that $\mu_\xi(\mathbb{R}^N) < +\infty$.

Moreover, for any $\pi_{T_\bullet}$ satisfying the above properties, we have

$$\|\pi_{T_\bullet}(f)\xi\|^2 = \int_{\mathbb{R}^N} |f|^2 d\mu_\xi \qquad \text{for each } \xi \in \mathscr{D}(\pi_{T_\bullet}(f)) \tag{6.57}$$

We call $(\pi_{T_\bullet}, \mathcal{H})$ the **Borel functional calculus** of $T_\bullet$, and write

$$\boldsymbol{f(T_\bullet)} := \pi_{T_\bullet}(f)$$

Note that (6.57) is particularly useful, as it allows us to reduce questions about the convergence of operators to corresponding questions about the convergence of functions.

Proof. Let us prove the uniqueness. Suppose that $\widetilde{\pi}_{T_\bullet}$ satisfies also conditions (1), (2), and (3). We extend $\pi_{T_\bullet}$ and $\widetilde{\pi}_{T_\bullet}$ by zero to

$$\pi_{T_\bullet}, \widetilde{\pi}_{T_\bullet} : \mathscr{Bor}_{\mathrm{b}}(\overline{\mathbb{R}}^N) \to \{\text{unbounded operators on } \mathcal{H}\}$$

In other words, if $f \in \mathscr{Bor}_{\mathrm{b}}(\overline{\mathbb{R}}^N)$, then $\pi_{T_\bullet}(f) = \pi_{T_\bullet}(f|_{\mathbb{R}^N})$, and $\widetilde{\pi}_{T_\bullet}$ is defined in a similar way.

By (2), the maps $\pi_{T_\bullet}$ and $\widetilde{\pi}_{T_\bullet}$ agree on the functions $u_j = (x_j - \mathbf{i})/(x_j + \mathbf{i})$. Here, u_j is viewed a continuous function on $\overline{\mathbb{R}}^N \simeq \mathbb{T}^N$, where the homeomorphism $\overline{\mathbb{R}} \simeq \mathbb{T}$ is defiend by the Cayley transform $x \mapsto \frac{x+\mathbf{i}}{x-\mathbf{i}}$. Since $u_1, \dots, u_N$ are the coordinate functions on $\mathbb{T}^N$, they separate the points of $\overline{\mathbb{R}}^N$. By the uniqueness part in Thm. 5.5.13, $\pi_{T_\bullet}$ and $\widetilde{\pi}_{T_\bullet}$ agree on $\mathscr{Bor}_{\mathrm{b}}(\overline{\mathbb{R}}^N)$, and hence on $\mathscr{Bor}_{\mathrm{b}}(\mathbb{R}^N)$. In particular, for each $\xi \in \mathcal{H}$, the measures $\mu_\xi, \widetilde{\mu}_\xi$ defined respectively by $\pi_{T_\bullet}, \widetilde{\pi}_{T_\bullet}$ are equal.

Now choose any $f \in \mathscr{Bor}(\mathbb{R}^N)$. By property (3), the unbounded operators $\pi_{T_\bullet}(f)$ and $\widetilde{\pi}_{T_\bullet}(f)$ have the same domain $\mathscr{D}_0$, and for any $\xi \in \mathscr{D}_0$ we have $\langle \xi | \pi_{T_\bullet}(f)\xi \rangle = \langle \xi | \widetilde{\pi}_{T_\bullet}(f)\xi \rangle$. By the polarization identiy, we have $\langle \eta | \pi_{T_\bullet}(f)\xi \rangle = \langle \eta | \widetilde{\pi}_{T_\bullet}(f)\xi \rangle$ for each $\eta, \xi \in \mathscr{D}_0$. The density of $\mathscr{D}_0$ implies $\pi_{T_\bullet}(f)\xi = \widetilde{\pi}_{T_\bullet}(f)\xi$. Thus $\pi_{T_\bullet}$ equals $\widetilde{\pi}_{T_\bullet}$ on the whole space $\mathscr{Bor}(\mathbb{R}^N)$.

To prove the existence, one may assume that $T_1, \dots, T_N$ are equal to the multiplication operators $\mathbf{M}_{x_1}, \dots, \mathbf{M}_{x_N}$ as described in the spectral Thm. 6.9.4. Then the existence follows immediately by applying the following Thm. 6.10.2 to $X = \mathbb{R}^N$ and $f_1 = x_1, \dots, f_N = x_N$. In this setting, (6.57) will also be proved in Step 3 of the proof of Thm. 6.10.2. $\square$

Theorem 6.10.2. *Let $(X, \mathfrak{M})$ be a measurable space. Let $f_1, \dots, f_N : X \to \mathbb{R}$ be measurable functions. Let $(\mu_\alpha)_{\alpha \in \mathscr{I}}$ be a family of measures on $\mathfrak{M}$. Let $\mathcal{H} = \bigoplus_{\alpha \in \mathscr{I}} L^2(X, \mu_\alpha)$. Then the map*

$$\pi_{T_\bullet} : \mathscr{Bor}(\mathbb{R}^N) \to \{\textit{unbounded operators on } \mathcal{H}\} \qquad g \mapsto \mathbf{M}_{g \circ f_\bullet} \tag{6.58}$$

(where $f_\bullet = (f_1, \dots, f_N) : \mathbb{R}^N \to \mathbb{R}^N$) is a Borel functional calculus of $\mathbf{M}_{f_1}, \dots, \mathbf{M}_{f_N}$.

Recall from Exp. 6.9.2 that $\mathbf{M}_{f_1}, \dots, \mathbf{M}_{f_N}$ are strongly commuting self-adjoint operators. In short, Thm. 6.10.2 says that for each $g \in \mathscr{Bor}(\mathbb{R}^N)$, we have

$$g(\mathbf{M}_{f_1}, \dots, \mathbf{M}_{f_N}) = \mathbf{M}_{g \circ f_\bullet}$$

Proof. Step 1. We need to prove that $\pi_{T_\bullet}$ satisfies the three properties in Thm. 6.10.1.

Property (1) is easy to check. In particular, for each $\xi = \oplus_\alpha \xi_\alpha$ in $\mathcal{H}$, noting that $\xi_\alpha = 0$ for all α outside a countable subset of $\mathscr{I}$ (Rem. 5.10.3), if we define[13]

$$\boxed{d\mu_\xi = \sum_{\alpha \in \mathscr{I}} (f_\bullet)_*(|\xi_\alpha|^2 d\mu_\alpha)} \tag{6.59}$$

then for each $g \in \mathscr{Bor}(\mathbb{R}^N, \overline{\mathbb{R}}_{\geqslant 0})$ we have (recalling Def. 1.6.6 for the basic properties of pushforward measures)

$$\sum_{\alpha \in \mathscr{I}} \int_X (g \circ f_\bullet)|\xi_\alpha|^2 d\mu_\alpha = \int_{\mathbb{R}^N} g d\mu_\xi \tag{6.60}$$

The LHS of (6.60) equals $\langle \xi | \pi_{T_\bullet}(g)\xi \rangle$ when $g \in \mathscr{Bor}_{\mathrm{b}}(\mathbb{R}^N, \mathbb{R}_{\geqslant 0})$. Thus $\langle \xi | \pi_{T_\bullet}(g)\xi \rangle = \int_{\mathbb{R}^N} g d\mu_\xi$ holds for all $g \in \mathscr{Bor}_{\mathrm{b}}(\mathbb{R}^N, \mathbb{R}_{\geqslant 0})$, and hence for all $g \in \mathscr{Bor}_{\mathrm{b}}(\mathbb{R}^N)$ by linearity. Thus μ_ξ is the finite Borel measure associated to ξ.

Step 2. Let us check property (2). According to the definition of $\pi_{T_\bullet}$,

$$\pi_{T_\bullet}\Big(\frac{x_j - \mathbf{i}}{x_j + \mathbf{i}}\Big) = \mathbf{M}_{\frac{x_j - \mathbf{i}}{x_j + \mathbf{i}} \circ f_\bullet} = \mathbf{M}_{(f_j - \mathbf{i})/(f_j + \mathbf{i})}$$

where the last term equals the Cayley transform of $\mathbf{M}_{f_j}$ by Exp. 6.9.2.

Step 3. We now check property (3). Choose any $g \in \mathscr{Bor}(\mathbb{R}^N)$ and $\xi = \oplus_\alpha \xi_\alpha \in \mathcal{H}$. From the definition of multiplication operators, we have $\xi \in \mathscr{D}(\mathbf{M}_{g \circ f_\bullet})$ iff

$$\sum_{\alpha \in \mathscr{I}} \int_X |g \circ f_\bullet|^2 \cdot |\xi_\alpha|^2 d\mu_\alpha \xlongequal{(6.60)} \int_{\mathbb{R}^N} |g|^2 d\mu_\xi \tag{6.61}$$

is finite; moreover, if $\xi \in \mathscr{D}(\mathbf{M}_{g \circ f_\bullet})$, then $\|\mathbf{M}_{g \circ f_\bullet}\xi\|^2$ equals the LHS (6.61). This proves the first half of (3) as well as Eq. (6.57).

By linearity, (6.60) holds when $g \in \mathscr{Bor}(\mathbb{R}^N, \mathbb{R})$, and when $g_+ = \max\{g, 0\}, g_- = \max\{-g, 0\}$ are μ_ξ-integrable (equivalently, $|g|$ is μ_ξ-integrable). Therefore, (6.60) holds when $g \in \mathscr{Bor}(\mathbb{R}^N)$, and when $\mathrm{Re}g, \mathrm{Im}g$ are μ_ξ-integrable (equivalently, $|g|$ is

[13] This formula is useful because it provides an explicit expression for μ_ξ in the context of multiplication operators.

μ_ξ-integrable). Therefore, it holds when $g \in \mathscr{Bor}(\mathbb{R}^N)$ and $|g|^2$ is μ_ξ-integral. This last statement is equivalent to

$$\langle \xi | \pi_{T_\bullet}(g)\xi \rangle = \int_{\mathbb{R}^N} g d\mu_\xi$$

whenever $\int_{\mathbb{R}^N} |g|^2 d\mu_\xi < +\infty$ (equivalently, $\xi \in \mathscr{D}(\mathbf{M}_{g\circ f_\bullet})$). This proves the second half of (3). □

Exercise 6.10.3. Let T be a self-adjoint operator on $\mathcal{H}$. Let $n \in \mathbb{N}$. Show that

$$T^n = f(T) \qquad \text{where } f(x) = x^n$$

by viewing T as a multiplication operator and applying Thm. 6.10.2 and Exe. 6.2.18.

Proposition 6.10.4. *Let $T_1, \dots, T_N \in \mathfrak{L}(\mathcal{H})$ be self-adjoint. Then the following two conditions are equivalent.*

(1) $T_1, \dots, T_N$ commute strongly.

(2) $T_1, \dots, T_N$ commute.

Moreover, if (1) and (2) hold and $f \in \mathscr{Bor}_b(\mathbb{R}^N)$, then the operator $f(T_\bullet)$ defined for strongly-commuting unbounded self-adjoint operators coincides with the one define for (adjointly) commuting bounded self-adjoint operators.

Proof. Assume (1). Then by Thm. 6.9.4, we can view $\mathcal{H}$ as $\bigoplus_\alpha L^2(\mathbb{R}^N, \mu_\alpha)$ where each μ_α is a finite Borel measure on $\mathbb{R}^N$, and $T_j = \mathbf{M}_{x_j}$. Let $R = \sqrt{\|T_1\|^2 + \cdots + \|T_N\|^2}$. For each $\delta > 0$ and each Borel set $A \subset \mathbb{R}^N \backslash \overline{B}_{\mathbb{R}^N}(0, R+\delta)$, by viewing χ_A as an element of $L^2(\mathbb{R}^N, \mu_\alpha)$, we have

$$\|T_1\chi_A\|^2 + \cdots + \|T_N\chi_A\|^2 = \int_A (x_1^2 + \cdots + x_N^2) d\mu_\alpha \geqslant (R+\delta)^2 \mu(A) = (R+\delta)^2 \|\chi_A\|^2$$

which forces χ_A to be zero in $L^2(\mathbb{R}^N, \mu_\alpha)$. Therefore, each μ_α is supported in $X := \overline{B}_{\mathbb{R}^N}(0, R)$, and hence $\mathcal{H} = \bigoplus_\alpha L^2(X, \mu_\alpha)$. It follows that $T_1 = \mathbf{M}_{x_1}, \dots, T_N = \mathbf{M}_{x_N}$ are mutually commuting bounded self-adjoint operators. Moreover, by Thm. 5.10.23 and 6.10.2, the two versions of $f(T_\bullet)$ are both equal to $\mathbf{M}_f$.

Conversely, assume (2). Let $X = \mathrm{Sp}(T_1) \times \cdots \times \mathrm{Sp}(T_N)$, which is a compact subset of $\mathbb{R}^N$ by Prop. 5.8.15. Then $\mathrm{Sp}(T_\bullet) \subset X$ by Prop. 5.8.9. By Thm. 5.10.22, one can view $\mathcal{H}$ as $\bigoplus_\alpha L^2(X, \mu_\alpha)$ and T_j as $\mathbf{M}_{x_j}$. Then the strong commutativity of $T_1, \dots, T_N$ follows from Exp. 6.9.2. □

6.10.2 Basic properties of Borel functional Calculus

Theorem 6.10.5. *Let $T_1, \ldots, T_N$ be strongly commuting self-adjoint operators on $\mathcal{H}$. Let $g_1, \ldots, g_L : \mathbb{R}^N \to \mathbb{R}$ be Borel functions. Then $g_1(T_\bullet), \ldots, g_L(T_\bullet)$ are strongly commuting self-adjoint operator on $\mathcal{H}$, and*

$$f(g_1(T_\bullet), \ldots, g_L(T_\bullet)) = (f \circ (g_1, \ldots, g_L))(T_\bullet) \tag{6.62}$$

holds for each $f \in \mathscr{Bor}(\mathbb{R}^L)$.

Proof. By Thm. 6.9.4, we may assume that $\mathcal{H} = \bigoplus_\alpha L^2(\mathbb{R}^N, \mu_\alpha)$ where (μ_α) is a family of finite Borel measures on $\mathbb{R}^N$, and $T_j = \mathbf{M}_{x_j}$. By Thm. 6.10.2, we have $g_k(T_\bullet) = \mathbf{M}_{g_k}$. Therefore, by Exp. 6.9.2, $g_1(T_\bullet), \ldots, g_L(T_\bullet)$ are strongly commuting self-adjoint operators.

Let $f \in \mathscr{Bor}(\mathbb{R}^L)$. By Thm. 6.10.2, we have

$$\begin{aligned}
&f(g_1(\mathbf{M}_{x_\bullet}), \ldots, g_L(\mathbf{M}_{x_\bullet})) = f(\mathbf{M}_{g_1}, \ldots, \mathbf{M}_{g_L})\\
=&\mathbf{M}_{f\circ(g_1,\ldots,g_L)} = (f \circ (g_1, \ldots, g_L))(\mathbf{M}_{x_\bullet})
\end{aligned}$$

This proves (6.62). □

Corollary 6.10.6. *Let $T_1, \ldots, T_N$ be strongly commuting self-adjoint operators on $\mathcal{H}$. Let $1 \leqslant L \leqslant N$. Assume that $f \in \mathscr{Bor}(\mathbb{R}^N)$ depends only on the first L variables so that f can also be viewed as a Borel function on $\mathbb{R}^L$. Then*

$$f(T_1, \ldots, T_L) = f(T_1, \ldots, T_N)$$

In particular, if f only depends on the first variable, then $f(T_1) = f(T_1, \ldots, T_N)$.

Proof. Apply Thm. 6.10.5 to the case that $f \in \mathscr{Bor}(\mathbb{R}^L)$ and $g_1, \ldots, g_L$ are the first L coordinate functions of $\mathbb{R}^N$. □

Next, we give some useful criteria for the strong commutativity of self-adjoint operators. More criteria will be given in Subsec. 7.3.4.

Proposition 6.10.7. *Let T be a self-adjoint operator on $\mathcal{H}$ with Cayley transform U_T. Let $V \in \mathfrak{L}(\mathcal{H})$ be unitary. Then $VT = TV$ iff V commutes with U_T.*

We emphasize that the commutativity of a unitary operator and a bounded linear operator implies the adjoint commutativity (cf. Exp. 5.7.8).

Proof. In general, if two self-adjoint operators are unitarily equivalent via a unitary operator V, then there Cayley transforms are also unitarily equivalent via V. Therefore, if $VTV^{-1} = T$, then $VU_TV^{-1} = U_T$. This proves "$\Rightarrow$".

Similarly, if two unitary operators U_1, U_2 are unitarily equivalent via a unitary operator V, and if $\mathrm{Rng}(U_1 - \mathbf{1})$ and $\mathrm{Rng}(U_2 - \mathbf{1})$ are dense in $\mathcal{H}$, then the inverse Cayley transforms of U_1, U_2 are also unitarily equivalent via V. This proves "$\Leftarrow$". □

Theorem 6.10.8. *Let T_1, T_2 be self-adjoint operators on $\mathcal{H}$ with Cayley transforms U_1, U_2 respectively. Then the following are equivalent.*

(1) T_1 and T_2 commute strongly, that is, $U_1U_2 = U_2U_1$.

(2) $U_1T_2 = T_2U_1$.

(3) $U_2T_1 = T_1U_2$.

(4) For each $f_1, f_2 \in \mathscr{B}or_{\mathrm{b}}(\mathbb{R}, \mathbb{R})$, the bounded self-adjoint operators $f_1(T_1), f_2(T_2)$ commute.

(5) For each $f_1, f_2 \in \mathscr{B}or_{\mathrm{b}}(\mathbb{R})$, the bounded normal operators $f_1(T_1), f_2(T_2)$ commute adjointly.

Proof. The equivalence of (1), (2), and (3) follow immediately from Prop. 6.10.7.

(1)⇒(4): Assume (1). By the spectral Thm. 6.9.4, we may assume that T_1, T_2 are equal to the multiplication operators $\mathbf{M}_{x_1}, \mathbf{M}_{x_2}$ on $\mathcal{H} = \bigoplus_\alpha L^2(\mathbb{R}^N, \mu_\alpha)$ where (μ_α) is a family of finite Borel measures. By Thm. 6.10.2, we have $f_1(T_1) = \mathbf{M}_{f_1}$ and $f_2(T_2) = \mathbf{M}_{f_2}$. These two bounded linear operators clearly commute.

(4)⇒(5): Since $f \in \mathscr{B}or_{\mathrm{b}}(\mathbb{R}) \to f(T_1) \in \mathfrak{L}(\mathcal{H})$ is a unitary representation (and in particular, linear), we have

$$f(T_1) = (\mathrm{Re} f)(T_1) + \mathbf{i}(\mathrm{Im} f)(T_1)$$

A similar relation holds for $f(T_2)$. By (4), the bounded self-adjoint operators $(\mathrm{Re} f_1)(T_1)$ and $(\mathrm{Im} f_1)(T_1)$ commute with $(\mathrm{Re} f_2)(T_2)$ and $(\mathrm{Im} f_2)(T_2)$. Therefore $f_1(T_1)$ commutes adjointly with $f_2(T_2)$. This proves (5).

(5)⇒(1): Assume (5). Let $f_1 = f_2 = f := \frac{x-\mathbf{i}}{x+\mathbf{i}}$. By property (2) in Thm. 6.10.1, $f(T_j)$ is the Cayley transform U_i of T_j. Therefore, U_1 commutes with U_2, and hence (1) holds. □

6.10.3 The joint spectrum $\mathrm{Sp}(T_\bullet)$

Definition 6.10.9. Let $T_1, \dots, T_N$ be strongly commuting self-adjoint operators on $\mathcal{H}$. Let $\pi_{T_\bullet}$ be the Borel functional calculus of $T_\bullet$ (cf. Thm. 6.10.1). The support (cf. Def. 5.7.9) of the restriction $\pi_{T_\bullet}|_{\mathscr{B}or_{\mathrm{b}}(\mathbb{R}^N)} : \mathscr{B}or_{\mathrm{b}}(\mathbb{R}^N) \to \mathfrak{L}(\mathcal{H})$ is denoted by $\mathbf{Sp}(\boldsymbol{T_\bullet})$ and called the **joint spectrum** of $T_1, \dots, T_N$. That is,

$$\mathrm{Sp}(T_\bullet) := \mathrm{Supp}(\pi_{T_\bullet}|_{\mathscr{B}or_{\mathrm{b}}(\mathbb{R}^N)}) \qquad \mathbb{R}^N$$

In other words, $\lambda_\bullet \in \mathbb{C}^N$ belongs to $\mathrm{Sp}(T_\bullet)$ iff $\chi_U(T_\bullet) \neq 0$ for each $U \in \mathrm{Nbh}_{\mathbb{C}^N}(\lambda_\bullet)$.

Note that $\mathrm{Sp}(T_\bullet)$ is a closed subset of $\mathbb{R}^N$, and hence is LCH.

Theorem 6.10.10. *Let $(X, \mathfrak{M})$ be a measurable space. Let $(\mu_\alpha)_{\alpha\in\mathscr{I}}$ be a family of σ-finite measures on $\mathfrak{M}$. Let $f_1, \dots, f_N : X \to \mathbb{R}$ be measurable. Let $\mathcal{H} = \bigoplus_{\alpha\in\mathscr{I}} L^2(X, \mu_\alpha)$. Let $T_1 = \mathbf{M}_{f_1}, \dots, T_N = \mathbf{M}_{f_N}$, which are strongly commuting self-adjoint operators by Exp. 6.9.2. Then* [14]

$$\mathrm{Sp}(T_\bullet) = \mathrm{Cl}_{\mathbb{C}^N}\Big(\bigcup_{\alpha\in\mathscr{I}} \mathrm{Rng}^{\mathrm{ess}}(f_\bullet, \mu_\alpha)\Big) \tag{6.63}$$

where $\mathrm{Rng}^{\mathrm{ess}}(f_\bullet, \mu_\alpha)$ is the essential range (cf. Def. 1.6.6) of the map $f_\bullet = (f_1, \dots, f_N) : X \to \mathbb{R}^N$ with respect to μ_α.

In the important special case that X is a Borel subset of $\mathbb{R}^N$ and $T_1 = \mathbf{M}_{x_1}, \dots, T_N = \mathbf{M}_{x_N}$, Eq. (6.63) becomes

$$\mathrm{Sp}(T_\bullet) = \mathrm{Cl}_{\mathbb{C}^N}\Big(\bigcup_{\alpha\in\mathscr{I}} \mathrm{Supp}(\mu_\alpha)\Big) \tag{6.64}$$

Proof. Denote the RHS of (6.63) by Y. For each open $U \subset \mathbb{R}^N$, we have

$$\chi_U(T_\bullet) = \mathbf{M}_{\chi_U\circ f_\bullet}$$

due to Thm. 6.10.2. Therefore, if U does not intersect Y, then U does not intersect $\mathrm{Supp}((f_\bullet)_*\mu_\alpha)$ for each α, and hence

$$\mu_\alpha\{x \in X : \chi_U \circ f_\bullet(x) \neq 0\} = \mu_\alpha((f_\bullet)^{-1}(U)) = ((f_\bullet)_*\mu_\alpha)(U) \tag{6.65}$$

equals zero. From this, it is clear that $\mathbf{M}_{\chi_U\circ f_\bullet} = 0$. We have thus proved that $\mathrm{Sp}(T_\bullet) \subset Y$.

Now choose any $y \in Y$. Then for each $U \in \mathrm{Nbh}_{\mathbb{R}^N}(y)$, there exists α such that U intersects $\mathrm{Supp}((f_\bullet)_*\mu_\alpha)$, and hence is not $(f_\bullet)_*\mu_\alpha$-null. It follows from (6.65) that the set $\{x \in X : \chi_U \circ f_\bullet(x) \neq 0\}$ is not μ_α-null. Since μ_α is σ-finite, there exists a measurable subset E of $\{x \in X : \chi_U \circ f_\bullet(x) \neq 0\}$ such that $0 < \mu_\alpha(E) < +\infty$. Let ξ_α be the function χ_E in $L^2(X, \mu_\alpha)$, viewed also as a vector of $\mathcal{H}$. Then

$$\langle \xi_\alpha | \mathbf{M}_{\chi_U\circ f_\bullet}\xi_\alpha\rangle = \int_E (\chi_U \circ f_\bullet) d\mu_\alpha = \mu_\alpha(E) > 0$$

Thus $\mathbf{M}_{\chi_U\circ f_\bullet} \neq 0$. This proves $Y \subset \mathrm{Sp}(T_\bullet)$. □

Remark 6.10.11. Let $T_1, \dots, T_N$ be strongly commuting self-adjoint operators on $\mathcal{H}$. Let $K \subset \mathbb{R}^N$ be a Borel set containing $\mathrm{Sp}(T_\bullet)$. Then there is a (clearly unique) map

$$\pi_{T_\bullet}|_K : \mathscr{B}or(K, \mathbb{R}) \to \{\text{unbounded operators on } \mathcal{H}\} \tag{6.66}$$

satisfying $\pi_{T_\bullet}|_K(f|_K) = \pi_{T_\bullet}(f)$ for each $f \in \mathscr{B}or(\mathbb{R}^N, \mathbb{R})$. The map (6.66) is also called a **Borel functional calculus** of $T_\bullet$. We write

$$\boldsymbol{f(T_\bullet)} := \pi_{T_\bullet}(f) \qquad \text{for each } f \in \mathscr{B}or(K, \mathbb{R})$$

[14]It makes no difference whether the closure is taken in $\mathbb{C}^N$ or in $\mathbb{R}^N$.

Proof. By the spectral Thm. 6.9.4, one may assume that $\mathcal{H} = \bigoplus_{\alpha\in\mathscr{I}} L^2(\mathbb{R}^N, \mu_\alpha)$ where each μ_α is a finite Borel measure on $\mathbb{R}^N$, and $T_1 = \mathbf{M}_{x_1}, \dots, T_N = \mathbf{M}_{x_N}$. It follows from Thm. 6.10.2 and the description (6.64) of $\mathrm{Sp}(T_\bullet)$ that the map (6.66) defined by $\pi_{T_\bullet}|_K(f) = \mathbf{M}_{\widetilde{f}}$ (where $\widetilde{f} \in \mathscr{Bor}(\mathbb{R}^N, \mathbb{R})$ is the zero extension of $f \in \mathscr{Bor}(K, \mathbb{R})$) satisfies the desired property. $\square$

Proposition 6.10.12. *Let $T_1, \dots, T_N$ be strongly commuting self-adjoint operators on $\mathcal{H}$. Let $1 \leqslant L < N$. Then*

$$\mathrm{Sp}(T_1, \dots, T_N) \subset \mathrm{Sp}(T_1, \dots, T_L) \times \mathrm{Sp}(T_{L+1}, \dots, T_N) \tag{6.67}$$

Proof. This can be proved in the same way as in Prop. 5.8.9. Alternatively, one may assume that $T_1, \dots, T_N$ are the multiplication operators $\mathbf{M}_{x_1}, \dots, \mathbf{M}_{x_N}$, and then apply Thm. 6.10.10. $\square$

We close this section by noting a close connection between $\mathrm{Sp}(T)$ and everywhere-defined bounded inverses.

Remark 6.10.13. Suppose that T is a self-ajoint operator on $\mathcal{H}$. Then for each $\lambda \in \mathbb{C}$ outside $\mathrm{Sp}(T)$, we have $(\lambda - T)^{-1} \in \mathfrak{L}(\mathcal{H})$ (cf. Def. 6.2.24).

Proof. By the spectral Thm. 6.9.4, we may assume $\mathcal{H} = \bigoplus_\alpha L^2(\mathrm{Sp}(T), \mu_\alpha)$ and $T = \mathbf{M}_x$. Since the function $(\lambda - x)^{-1}$ is bounded on $\mathrm{Sp}(T)$, the multiplication operator $\mathbf{M}_{1/(\lambda - x)}$ belongs to $\mathfrak{L}(\mathcal{H})$ and is the inverse of $\lambda - T = \mathbf{M}_{\lambda - x}$ due to Exp. 6.2.23. $\square$

The converse of Rem. 6.10.13 also holds: any $\lambda \in \mathbb{C}$ satisfying $(\lambda - T)^{-1} \in \mathfrak{L}(\mathcal{H})$ must lie outside $\mathrm{Sp}(T)$. This can be proved by first showing that $\mathrm{Sp}(T)$ equals the set of approximate eigenvalues of T; see Thm. 6.12.5.

6.11 Why the methods of Hilbert and Riesz fail

Let $\mathcal{H}$ be a Hilbert space.

6.11.1 The three paradigm shifts

In Sec. 5.4, we mentioned three major paradigm shifts as the central themes of this course. As this chapter illustrates, all three are deeply embodied in von Neumann's treatment of the spectral theory of unbounded operators. In this section, I would like to focus on (5.19a), the first of these shifts—the transition from finite approximation to linear extension. The pivotal role of linear extension in von Neumann's theory is most clearly seen in Thm. 6.3.5, where the Cayley transform is used to translate the problem of extending a Hermitian operator on $\mathcal{H}$ into that of extending a unitary map between two linear subspaces of $\mathcal{H}$.

The other two paradigm shifts, (5.19b) and (5.19c), follow naturally from this one. The move from the paradigm of bilinear/sesquilinear forms to that of linear operators is unavoidable because, even for bounded sesquilinear forms, defining multiplication necessarily involves finite approximation—as discussed in Subsec. 3.5.4—and so does the definition of resolvents, as seen in Hilbert's treatment of the resolvent in his proof of the spectral theorem in Sec. 5.1. Accordingly, the shift from the paradigm of duality to that of Cauchy completeness arises as a natural consequence of the focus on linear operators, as previously discussed in Subsec. 2.5.1.

6.11.2 Von Neumann's comments on the finite-approximation paradigm

In [vN29a], von Neumann's own comparison between the finite-approximation paradigm and the linear-extension paradigm appears in Section IX of the Introduction, immediately following his brief introduction of the Cayley transform method. The following is an English translation of his discussion:

> All other methods fail: The elegant procedures of Hellinger and of F. Riesz fail, because in them the operator R must be iterated, which is only permissible for everywhere-defined (hence bounded) Hermitian operators; for arbitrary Hermitian operators this is initially doubtful. All maximum-minimum methods, as well as Schmidt's method, are from the beginning restricted to cases without continuous spectrum. The original Hilbert method (approximation by finite-dimensional truncated operators) alone allows the derivation of certain results; but only with very great difficulties, and only a fraction of what we attain (footnote 23).

Von Neumann's objection to (a direct application of) Riesz's method is clear:[15] for a Hermitian operator T, one cannot simply perform the polynomial functional calculus and then extend to more general functional calculus, since arbitrary compositions of unbounded operators can be problematic. The correct approach—which von Neumann does not state explicitly in this paragraph—is to perform the functional calculus on the Cayley transform of T, once T has been extended to a self-adjoint operator.

The most revealing part of this quotation, however, is von Neumann's comment on Hilbert's method of finite-rank approximation. As we saw in Sec. 5.1, Hilbert derived his spectral theorem for a bounded bilinear (or sesquilinear) form ω by approximating ω with finite-rank forms. In the language of linear operators, this means approximating $T \in \mathfrak{L}(\mathcal{H})$ by E_nTE_n, where E_n is the projection of $\mathcal{H}$

[15] It is not clear whether Riesz himself has applied this method to unbounded operators.

(here assumed to be separable) onto $\mathrm{Span}\{e_1, \ldots, e_n\}$, with $e_1, e_2, \ldots$ an orthonormal basis of $\mathcal{H}$. The integral representation of the resolvent

$$\langle \xi | (z - T)^{-1} \xi \rangle = \int \frac{d\rho_\xi(\lambda)}{z - \lambda}$$

is then obtained as the limit of the corresponding representation for $\langle \xi | (z - E_n T E_n)^{-1} \xi \rangle$. In fact, as discussed in the answer to Question 5.1.2, in Hilbert's proof the resolvent $(z-T)^{-1}$ itself is defined as the limit of $(z-E_n T E_n)^{-1}$.

The last sentence of von Neumann's quotation can thus be understood as follows: When T is an unbounded Hermitian operator, defining the resolvent $(z - T)^{-1}$ and deriving its integral representation through the limits of $(z - E_n T E_n)^{-1}$ and their corresponding integral representations can yield only partial results compared with von Neumann's method via the Cayley transform—and even these partial results are obtained only with great difficulty.

What von Neumann meant by partial results (i.e., "a fraction of what we attain") can be inferred from the first half of Footnote 23 in [vN29a], where he comments on his earlier attempt to study the spectral theorem for Hermitian operators using Hilbert's method of finite approximation:

> By this method the author was able to treat the case of real Hermitian operators (cf. footnote 20), where the exceptional case indicated in VII had not yet appeared. However, the extension process for the Hermitian operators could not be as clearly understood as will be the case in Chapter VIII, for example, and the method was so non-constructive that, for instance, the well-ordering theorem had to be invoked (cf. the announcement in Jahresber. d. D. Math.-Ver. 37, 1-4 (1928), pp. 11-15); furthermore, the non-real Hermitian Operators could not be addressed in this way.

By "exceptional case" von Neumann meant the situation where exactly one of the two deficiency indices n_+, n_- of a Hermitian operator T is nonzero—or more generally, where $n_+ \neq n_-$. Recall from Cor. 6.7.11 that T admits a self-adjoint extension iff $n_+ = n_-$. This corollary appears in Satz 35 of [vN29a, Chapter VIII], which, according to the above quotation, provides a much clearer understanding of the extension process for Hermitian operators than Hilbert's finite-rank approximation method.

6.11.3 Why the linear-extension paradigm is superior to the finite-approximation paradigm

It might be too harsh to accuse Hilbert's method of failing to clarify the extension process for Hermitian operators—after all, according to the paradigm shift

(5.19a), the paradigm of linear extension directly opposes that of finite approximation. What can be said fairly, however, is that the finite-approximation paradigm does not provide a structural understanding of the non-uniqueness of spectral decompositions of Hermitian operators.

Indeed, as seen in Sec. 4.2, the solutions to the Hamburger and Stieltjes moment problems are not always unique, and this non-uniqueness corresponds to the fact that it is often a subsequence of $(z - E_n T E_n)^{-1}$, rather than the full sequence, that converges. Consequently, the resolvent $(z - T)^{-1}$, defined as the limit of a subsequence $(z - E_{n_k} T E_{n_k})^{-1}$, is not unique and depends on the chosen subsequence; the same holds for the spectral decomposition of T, expressed via the integral representation of T. In von Neumann's spectral theory, by contrast, this non-uniqueness appears as the non-uniqueness of the self-adjoint extensions $\widehat{T}$ of T.

In light of this comparison, the real advantage of von Neumann's theory over Hilbert's finite-approximation method becomes clear: Von Neumann's framework characterizes the non-uniqueness (and even the non-existence) of the resolvents and the spectral decompositions of a Hermitian operator T in a structural and conceptual way through the extensions of the Cayley transform of T and the deficiency indices n_+, n_-, rather than through the choice of subsequences of $(z - E_n T E_n)^{-1}$.

6.11.4 Real Hermitian operators have self-adjoint extensions

In the second quotation from von Neumann in Subsec. 6.11.2, it was mentioned that spectral theory can also be developed for real Hermitian operators, that is, Hermitian operators acting on real Hilbert spaces. For instance, the multiplication operator $\mathbf{M}_f$ associated with a measurable real-valued function f, defined on a dense subspace of a real L^2-space, is a real Hermitian operator. By contrast, $-\mathbf{i}\frac{d}{dx}$ is not a real Hermitian operator, since it maps real-valued functions to complex-valued ones.

The following theorem shows that every real Hermitian operator admits a self-adjoint extension.

Theorem 6.11.1. *Let T be a Hermitian operator on the (complex) Hilbert space $\mathcal{H}$. Assume that $C : \mathcal{H} \to \mathcal{H}$ is an anti-unitary map satisfying $C^2 = \mathrm{id}_{\mathcal{H}}$. Assume that*

$$\begin{aligned} C(\mathscr{D}(T)) &\subset \mathscr{D}(T) \\ TC\xi = CT\xi &\qquad \textit{for each } \xi \in \mathscr{D}(T) \end{aligned}$$

Then $C(\mathscr{D}(T)) = \mathscr{D}(T)$, and T can be extended to a self-adjoint operator on $\mathcal{H}$.

Here, the associated real Hilbert space is understood as $\mathcal{H}_{\mathbb{R}} = \{\xi \in \mathcal{H} : \xi = C\xi\}$. The corresponding real Hermitian operator is $T_{\mathbb{R}} : \mathcal{H}_{\mathbb{R}} \to \mathcal{H}_{\mathbb{R}}$ with domain $\mathscr{D}(T_{\mathbb{R}}) = \{\xi \in \mathscr{D}(T) : C\xi = \xi\}$ defined by $T_{\mathbb{R}}\xi = T\xi$.

Proof. In general, if $\Gamma : \mathcal{H} \to \mathcal{K}$ is an antiunitary map of Hilbert spaces, if S is an unbounded operator on $\mathcal{K}$ satisfying $\Gamma(\mathscr{D}(T)) = \mathscr{D}(S)$ and $S\Gamma\xi = \Gamma T\xi$ for each $\xi \in \mathscr{D}(T)$, then S is also self-adjoint, and $\Gamma(\mathrm{Rng}(T \pm \mathbf{i})^\perp) = \mathrm{Rng}(S \mp \mathbf{i})^\perp$.

In the setting of the current theorem, the condition $C(\mathscr{D}(T)) \subset \mathscr{D}(T)$ implies

$$\mathscr{D}(T) = C^{-1}C(\mathscr{D}(T)) \subset C^{-1}(\mathscr{D}(T))$$

and hence $\mathscr{D}(T) \subset C(\mathscr{D}(T))$ due to $C^2 = \mathbf{1}$. Therefore $C(\mathscr{D}(T)) = \mathscr{D}(T)$. Applying the above general result to $\mathcal{K} = \mathcal{H}$, $S = T$, and $\Gamma = C$, we obtain

$$C(\mathrm{Rng}(T + \mathbf{i})^\perp) = \mathrm{Rng}(T - \mathbf{i})^\perp$$

Therefore, the two deficiency indices n_+, n_- of T are equal. By Cor. 6.7.11, T has a self-adjoint extension. $\square$

The above theorem provides an alternative proof of the Hamburger moment theorem as stated in Thm. 4.2.9. (See Exp. 5.11.3 for the spectral-theoretic solution to the Hausdorff moment problem).

Example 6.11.2. Let $(c_n)_{n\in\mathbb{N}}$ be a sequence in $\mathbb{R}$. Let H be its Hankel matrix, cf. Def. 4.2.6. As in the proof of Prop. 4.2.8, if there is a finite Borel measure μ on $\mathbb{R}$ satisfying $\int_{\mathbb{R}} x^n d\mu = c_n$ for all n, then necessarily $H \geqslant 0$.

Conversely, assume $H \geqslant 0$. Let $\mathscr{A} = \mathbb{C}[x]$, and let $\Lambda : \mathscr{A} \to \mathbb{C}$ be the unique linear functional such that $\Lambda(x^n) = c_n$ for all n. By Pb. 4.1 and Rem. 4.6.1, the space $\mathscr{A}$ admits a positive sesquilinear form defined by $\langle f|g\rangle = \Lambda(\overline{f}g)$, which descends to an inner product on $V = \mathscr{A}/\mathscr{N}$ where $\mathscr{N} = \{g \in \mathscr{A} : \langle g|g\rangle = 0\}$. Moreover, $\mathscr{A}$ has a pre-unitary representation (π, V) such that $\pi(f)(g + \mathscr{N}) = fg + \mathscr{N}$.

Let $\mathcal{H}$ be the Hilbert space completion of V (cf. Pb. 3.2). Let T be the unbounded operator on $\mathcal{H}$ with domain $\mathscr{D}(T) = V$ defined by $T\xi = \pi(x)\xi$. (See also Thm. 4.2.10.) Then T is clearly a Hermitian operator. Moreover, we clearly have an antiunitary map $C : V \to V$ defined by sending $f + \mathscr{N}$ to $\overline{f} + \mathscr{N}$. It is extended uniquely to an antiunitary map $C : \mathcal{H} \to \mathcal{H}$. Clearly C satisfies the conditions in Thm. 6.11.1. Therefore, by that theorem, T has a self-adjoint extension $\widetilde{T}$.

The vector $\Omega = 1 + \mathscr{N}$ belongs to V, and hence belongs to $\mathscr{D}(\widetilde{T}^n)$ for any $n \in \mathbb{N}$. Let $\mu = \mu_\Omega$ be the finite Borel measure on $\mathbb{R}$ associated to Ω and the restriction of the Borel functional calculus $\pi_{\widetilde{T}}|_{\mathscr{B}or_{\mathrm{b}}(\mathbb{R})}$. Then property (3) of Thm. 6.10.1 yields

$$\langle \Omega | f(\widetilde{T})\Omega\rangle = \int_{\mathbb{R}} f d\mu$$

for each $f \in \mathscr{B}or(\mathbb{R})$ satisfying $\Omega \in \mathscr{D}(f(\widetilde{T}))$. Setting $f(x) = x^n$ and noting $f(\widetilde{T}) = \widetilde{T}^n$ (cf. Exe. 6.10.3), we obtain $c_n = \Lambda(x^n) = \langle \Omega | T^n\Omega\rangle = \int_{\mathbb{R}} x^n d\mu$. Thus μ solves the Hamburger moment problem. $\square$

6.12 Problems

Let $\mathcal{H}, \mathcal{K}, \mathcal{M}$ be Hilbert spaces.

† **Problem 6.1.** Let $A, B : \mathcal{H} \to \mathcal{K}$ and $C : \mathcal{M} \to \mathcal{H}$ be unbounded operators. Assume that $A + B : \mathcal{H} \to \mathcal{K}$ and $AC : \mathcal{M} \to \mathcal{K}$ are unbounded operators (i.e. they are densely-defined).

1. Prove that as n.d.d. unbounded operators, we have

$$(A+B)^* \supset A^* + B^* \qquad (AC)^* \supset C^* A^*$$

2. Prove that if $A \in \mathfrak{L}(\mathcal{H}, \mathcal{K})$, then $(A+B)^* = A^* + B^*$, and $(AC)^* = C^* A^*$.

The following result will be needed to solve Pb. 6.17.

Problem 6.2. Let A, B be n.d.d. injective unbounded operators on $\mathcal{H}$. Prove that

$$A^{-1} + B^{-1} \supset A^{-1}(A+B)B^{-1} \tag{6.68}$$

that the LHS has domain $\mathscr{D}(A^{-1}) \cap \mathscr{D}(B^{-1})$, and that the RHS has domain $\mathscr{D}(A^{-1}) \cap \mathscr{D}(AB^{-1})$. Conclude that the inclusion relation "$\supset$" in (6.68) becomes "$=$" if A is everywhere defined on $\mathcal{H}$.

† **Problem 6.3.** Solve Exe. 6.2.18.

Hint. This result is not as obvious as it appears. Let $E = \{x \in X : |f(x)| \leqslant 1\}$ and $F = X \backslash E$. For natural numbers $0 \leqslant k \leqslant n$ and $\xi \in \bigoplus_\alpha L^2(X, \mu_\alpha)$, show that $\sum_\alpha \|f^k \chi_F \xi\|^2_{L^2(\mu_\alpha)} \leqslant \sum_\alpha \|f^n \chi_F \xi\|^2_{L^2(\mu_\alpha)}$ and $\sum_\alpha \|f^n \chi_E \xi\|^2_{L^2(\mu_\alpha)} < +\infty$. □

Recall Pb. 3.3 for the definition of Hilbert space dimensions.

Definition 6.12.1. Assume that $\mathcal{H}$ is separable. Let $\omega : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$ be a bounded Hermitian form. Define $n_\pm$ to be the largest dimension of closed linear subspaces $\mathcal{K} \subset \mathcal{H}$ such that $\pm\omega|_\mathcal{K}$ is positive-definite (i.e. an inner product). We call n_+ and n_- the **positive index of inertia** and **negative index of inertia** of ω, respectively.

Remark 6.12.2. The indices of inertia depend only on the equivalence class of the Hilbert space inner product. In other words, if one replaces the inner product on $\mathcal{H}$ by any equivalent inner product (in the sense of Def. 7.2.13), then the indices of inertia of every bounded sesquilinear form ω remain unchanged.

Problem 6.4. Assume that $\mathcal{H}$ is separable. Let ω be a bounded Hermitian form on $\mathcal{H}$.

1. Prove that $\mathcal{H}$ admits an orthogonal decomposition into Hilbert subspaces $\mathcal{H} = \mathcal{H}_0 \oplus \mathcal{H}_+ \oplus \mathcal{H}_-$ such that $\omega|_{\mathcal{H}_+}$ and $-\omega|_{\mathcal{H}_-}$ are positive-definite, and that for each $\xi_0 \in \mathcal{H}_0, \xi_\pm \in \mathcal{H}_\pm$ we have

$$\omega(\xi_0 + \xi_+ + \xi_- | \xi_0 + \xi_+ + \xi_-) = \omega(\xi_+|\xi_+) + \omega(\xi_-|\xi_-)$$

2. For any decomposition as in Part 1, prove that $n_+ = \dim \mathcal{H}_+$ and $n_- = \dim \mathcal{H}_-$.

Hint. Part 1. Apply the spectral theorem to the bounded self-adjoint operator associated to ω.

Part 2. To prove $n_\pm \leqslant \dim \mathcal{H}_\pm$, show that if $\mathcal{K}$ is a closed linear subspace of $\mathcal{H}$ such that $\pm\omega|_\mathcal{K}$ is positive-definite, then the projection $\mathcal{H} \to \mathcal{H}_\pm$ restricts to an injective map $\mathcal{K} \to \mathcal{H}_\pm$. $\square$

Problem 6.5. Assume that $\mathcal{H}$ is separable. Let T be a closed Hermitian operator on $\mathcal{H}$. Define a sesquilinear form ω on $\mathscr{D}(T^*)$ by $\omega = (\omega_{T^*} - (\omega_{T^*})^*)/2\mathbf{i}$, that is,

$$\omega(\xi|\eta) = \frac{\langle\xi|T^*\eta\rangle - \overline{\langle\eta|T^*\xi\rangle}}{2\mathbf{i}} \qquad \text{for each } \xi, \eta \in \mathscr{D}(T^*)$$

(In other words, ω is the unique sesquilinear form on $\mathscr{D}(T^*)$ such that $\omega(\xi|\xi) = \mathrm{Im}\langle\xi|T^*\xi\rangle$ for each $\xi \in \mathscr{D}(T^*)$.)

1. Show that ω is a bounded Hermitian form with respect to the graph inner product of T^*. Moreover, prove that the positive and negative indices of inertia n_+ and n_- of ω coincide with the deficiency indices of T.

2. Prove that if $a \in \mathbb{R}_{>0}$ and $b \in \mathbb{R}$, then T and $aT + b$ have the same pair of deficiency indices.

Problem 6.6. Let $I = (a, b)$ where $-\infty \leqslant a < b \leqslant +\infty$. Let T_0 be the unbounded operator $-\mathbf{i}\frac{d}{dx}$ on $L^2(I, m)$ with domain $C_c^\infty(I)$. Let $T = \overline{T_0}$.

1. Prove that

$$\mathrm{Ker}(T^* - \mathbf{i}) = \begin{cases} \mathrm{Span}\{e^{-x}\} & \text{if } a > -\infty \\ 0 & \text{if } a = -\infty \end{cases}$$

$$\mathrm{Ker}(T^* + \mathbf{i}) = \begin{cases} \mathrm{Span}\{e^{x}\} & \text{if } b < +\infty \\ 0 & \text{if } b = +\infty \end{cases}$$

2. Conclude that if $I = (0, +\infty)$, the Hermitian operator T has deficiency indices $n_+ = 1, n_- = 0$, and hence has no self-adjoint extensions.

3. Conclude that if $I = \mathbb{R}$, then $H^1(I) = H_0^1(I)$, and T_0 is essentially self-adjoint.

Hint. Restrict $f \in \mathscr{D}(T^*)$ to bounded intervals, and apply the results in Sec. 6.8. $\square$

Problem 6.7. Let $-\infty < a < b < +\infty$ and $I = (a, b)$. Define the Hermitian operator $S = -\mathbf{i}\frac{d}{dx}$ on $\mathcal{H} = L^2(I, m)$ with $\mathscr{D}(S) = C_c^\infty(I)$. Let $\nu \in \mathbb{Z}_+$. Let

$$\boldsymbol{AC_0(\overline{I})} = \{f \in AC(\overline{I}) : f(a) = f(b) = 0\}$$

1. Prove that $\mathrm{Ker}((S^\nu)^*) = \mathrm{Span}\{1, x, \dots, x^{\nu-1}\}$.

2. Prove that

$$\mathscr{D}((S^\nu)^*) = \{f \in AC(\overline{I}) : f', \dots, f^{(\nu-1)} \in AC(\overline{I}), f^{(\nu)} \in L^2(I)\} \tag{6.69}$$

   and that $(S^\nu)^* f = (-\mathbf{i})^\nu f^{(\nu)}$ a.e. for each $f \in \mathscr{D}((S^\nu)^*)$.

3. Prove that

$$\mathscr{D}(\overline{S^\nu}) = \{f \in AC_0(\overline{I}) : f', \dots, f^{(\nu-1)} \in AC_0(\overline{I}), f^{(\nu)} \in L^2(I)\}$$

From the above results one easily deduces $\mathscr{D}((S^\nu)^*) = \mathscr{D}((S^*)^\nu) = H^\nu(I)$ and $\mathscr{D}(\overline{S^\nu}) = \mathscr{D}(\overline{S}^\nu) = H_0^\nu(I)$, cf. Exp. 6.6.17 for the definition of Sobolev spaces.

Hint. Part 1. Prove this by applying Lem. 6.8.2 successively. (The same method yields a more general statement: linear functionals on $C_c^\infty(I)$ whose ν-th distribution derivatives vanish are precisely linear combinations of $1, x, \dots, x^{\nu-1}$.)

Part 2. Denote the RHS of (6.69) by $\mathcal{F}$. Show that $\mathscr{D}((S^\nu)^*) \supset \mathcal{F}$ and $(S^\nu)^* f = (-\mathbf{i})^\nu f^{(\nu)}$ for each $f \in \mathcal{F}$. Define $\mathcal{J} : L^2(I, m) \to L^2(I, m)$ by $(\mathcal{J} f)(x) = \mathbf{i}\int_a^x f dm$. To prove "$\subset$", choose any $f \in \mathscr{D}((S^\nu)^*)$, let $\widetilde{f} = \mathcal{J}^\nu (S^\nu)^* f$, show that $\widetilde{f} \in \mathcal{F}$ and $(S^\nu)^*(f - \widetilde{f}) = 0$. Apply part 1.

Part 3. Use $\overline{S^\nu} = (S^\nu)^{**}$ to prove one direction. □

Problem 6.8. Let $-\infty < a < b < +\infty$ and $I = (a, b)$. Define the negative Laplacian $-\Delta = -\frac{d^2}{dx^2}$ with $\mathscr{D}(-\Delta) = C_c^\infty(I)$.

1. Prove that if $f \in \mathscr{D}(\Delta^*)$ and $-\Delta^* f = \lambda f$ for some $\lambda \in \mathbb{C}$, then $f \in C^\infty(I)$.

2. Find the deficiency indices of $-\Delta$.

Problem 6.9. Let $(X, \mathfrak{M})$ be a measurable space. Let $(\mu_\alpha)_{\alpha \in \mathscr{I}}$ be a collection of measures on $\mathfrak{M}$. Let $f : X \to \mathbb{C}$ be a measurable function. Let $\mathbf{M}_f$ be the multiplication operator on $\mathcal{H} := \bigoplus_{\alpha \in \mathscr{I}} L^2(X, \mu_\alpha)$ associated with f. Prove that

$$(\mathbf{M}_f)^* = \mathbf{M}_{\overline{f}}$$

Problem 6.10. Let $(\mathcal{H}_\alpha)_{\alpha \in \mathscr{A}}$ be a family of Hilbert spaces. For each $\alpha \in \mathscr{A}$, let (T_α) be a Hermitian operator on $\mathcal{H}_\alpha$. Prove that the direct sum operator $T := \oplus_\alpha T_\alpha$ (cf. Def. 6.2.15) is self-adjoint iff each T_α is self-adjoint.

The following problem shows that strong commutativity implies ordinary commutativity.

Problem 6.11. Let A, B be strongly-commuting self-adjoint operators on $\mathcal{H}$. Prove that for each $\xi \in \mathscr{D}(AB) \cap \mathscr{D}(A)$, we have $\xi \in \mathscr{D}(BA)$ and $AB\xi = BA\xi$.

Hint. View A, B as multiplication operators. □

Problem 6.12. Let A, B be self-adjoint operators on $\mathcal{H}$. Let $U \in \mathfrak{L}(\mathcal{H})$ be unitary. Define the **spectral projections** $E(\lambda) := \chi_{(-\infty,\lambda]}(A)$ and $F(\lambda) := \chi_{(-\infty,\lambda]}(B)$ where $\lambda \in \mathbb{R}$.

1. Prove that $UA = AU$ iff $UE(\lambda) = E(\lambda)U$ for each $\lambda \in \mathbb{R}$.
2. Prove that A commutes strongly with B iff $E(\lambda)F(\mu) = F(\mu)E(\lambda)$ for each $\lambda, \mu \in \mathbb{R}$.

Hint for "⇐". Part 1. Prove that for each $-\infty < a < b < +\infty$, $A\chi_{(a,b]}(A)$ can be recovered from the spectral projections of A. Show that $\mathscr{D}_1 := \sum_{-\infty<a<b<+\infty} \mathrm{Rng}(\chi_{(a,b]})(A)$ is a core for A, and that $A|_{\mathscr{D}_1}$ can be recovered from the spectral projections of A.

Part 2. Prove that the Cayley transform $V(\mu)$ of $F(\mu)$ commutes with $E(\lambda)$. Conclude from Part 1 that $V(\mu)$ commutes with A. □

Definition 6.12.3. Let $T_1, \ldots, T_N$ be strongly commuting self-adjoint operators on $\mathcal{H}$. We say that $\lambda = (\lambda_1, \ldots, \lambda_N) \in \mathbb{C}^N$ is a **joint approximate eigenvalue** of $T_1, \ldots, T_N$ if one of the following (clearly) equivalent conditions holds:

(1) For each $\varepsilon > 0$ there exists a nonzero vector $\xi \in \mathscr{D}(T_1) \cap \cdots \cap \mathscr{D}(T_N)$ such that
$$\sum_{j=1}^{N} \|T_j\xi - \lambda_j\xi\|^2 \leqslant \varepsilon\|\xi\|^2$$

(2) There exists a sequence $(\xi_n)_{n\in\mathbb{Z}_+}$ of unit vectors in $\mathscr{D}(T_1) \cap \cdots \cap \mathscr{D}(T_N)$ such that
$$\lim_n (T_j - \lambda_j)\xi_n = 0 \qquad \text{for each } 1 \leqslant j \leqslant N$$

† **Problem 6.13.** Let $T_1, \ldots, T_N$ be strongly commuting self-adjoint operators on $\mathcal{H}$. Let $\lambda_\bullet = (\lambda_1, \ldots, \lambda_N) \in \mathbb{C}^N$. Prove that $\lambda_\bullet \in \mathrm{Sp}(T_\bullet)$ (in particular, $\lambda_\bullet \in \mathbb{R}^N$) iff $\lambda_\bullet$ is a joint approximate eigenvalue of $T_1, \ldots, T_N$.

Definition 6.12.4. For each unbounded operator T on $\mathcal{H}$, define
$$\boldsymbol{\sigma(T)} = \{\lambda \in \mathbb{C} : (\lambda - T)^{-1} \notin \mathfrak{L}(\mathcal{H})\} \tag{6.70}$$

That is, $\lambda \notin \sigma(T)$ iff $\lambda - T$ has an everywhere-defined bounded inverse on $\mathcal{H}$ (cf. Def. 6.2.24).

The following result will be used to establish the lower-boundedness of certain Hamiltonian operators, cf. Cor. 7.9.15.

Theorem 6.12.5. *Let T be a self-adjoint operator on $\mathcal{H}$. Then*

$$\mathrm{Sp}(T) = \sigma(T)$$

Proof. Let $\lambda \in \mathbb{C}$. If $\lambda \notin \mathrm{Sp}(T)$, then Rem. 6.10.13 implies $\lambda \notin \sigma(T)$. Conversely, if $\lambda \in \mathrm{Sp}(T)$, then Pb. 6.13 implies that there is a sequence of unit vectors (ξ_n) in $\mathscr{D}(T)$ such that $\lim_n (T - \lambda)\xi_n = 0$. Therefore, for any $A \in \mathfrak{L}(\mathcal{H})$, we have $\lim_n A(T - \lambda)\xi_n = 0$, and hence $A(T - \lambda)$ cannot be $\mathbf{1}$ on $\mathscr{D}(T)$. Therefore, $T - \lambda$ has no everywhere-defined bounded inverse. □

Problem 6.14. Let $(U_\alpha)_{\alpha \in \mathscr{A}}$ be a net of unitary operators on $\mathcal{H}$ converging in SOT to a unitary operator $U \in \mathfrak{L}(\mathcal{H})$. Assume that $\mathrm{Ker}(U - \mathbf{1}) = 0$. Prove that for each bounded function $f : \mathbb{S}^1 \to \mathbb{C}$ continuous on $\mathbb{S}^1 \backslash \{1\}$, the net $(f(U_\alpha))_{\alpha \in \mathscr{A}}$ converges in SOT* to $f(U)$.

Hint. Use Pb. 5.4 to show that if $f \in C(\mathbb{S}^1)$, then $f(U_\alpha)$ converges in WOT to $f(U)$, and $f(U_\alpha)^* f(U_\alpha)$ converges in WOT to $f(U)^* f(U)$. Then apply Pb. 2.2 to remove the assumption that f is continuous at 1. □

6.12.1 Convergence in the strong resolvent sense

Definition 6.12.6. Let $(T_\alpha)_{\alpha \in \mathscr{A}}$ be a net of self-adjoint operators on $\mathcal{H}$. Let T be a self-adjoint operator on $\mathcal{H}$. We say that (T_α) converges to T in the **strong resolvent sense** if the Cayley transforms $(T_\alpha - \mathbf{i})(T_\alpha + \mathbf{i})^{-1}$ converge in SOT to $(T - \mathbf{i})(T + \mathbf{i})^{-1}$.

Problem 6.15. Let $(T_\alpha)_{\alpha \in \mathscr{A}}$ be a net of self-adjoint operators on $\mathcal{H}$ converging in the strong resolvent sense to a self-adjoint operator T.

1. Suppose that $f : \mathbb{R} \to \mathbb{C}$ is bounded and continuous. Use Pb. 6.14 to show that $f(T_\alpha)$ converges in SOT* to $f(T)$.

2. Suppose that $f : \mathbb{R} \to \mathbb{R}$ is continuous. Prove that $f(T_\alpha)$ converges in the strong resolvent sense to $f(T)$.

Problem 6.16. Let T be a self-adjoint operator on $\mathcal{H}$ with Cayley transform U_T. Let $\lambda \in \mathbb{C} \backslash \mathbb{R}$. Note from Rem. 6.10.13 that $(\lambda - T)^{-1} \in \mathfrak{L}(\mathcal{H})$.

1. Let $f : \mathbb{S}^1 \to \mathbb{C}$ be the unique map sending each $(t - \mathbf{i})/(t + \mathbf{i})$ (where $t \in \mathbb{R} \cup \{\infty\}$) to $1/(\lambda - t)$. Let $g : \mathrm{Rng}(f) \to \mathbb{S}^1$ be the inverse of f, extended by zero to $g : \mathbb{C} \to \mathbb{C}$. Prove that $f(U_T) = (\lambda - T)^{-1}$, and that $g((\lambda - T)^{-1}) = U_T$.

2. Let (T_α) be a net of self-adjoint operators on $\mathcal{H}$. Prove that T_α converges in the strong resolvent sense to T iff $(\lambda - T_\alpha)^{-1}$ converges in SOT to $(\lambda - T)^{-1}$.

Problem 6.17. Let $(T_\alpha)_{\alpha\in\mathscr{A}}$ be a net of self-adjoint operators on $\mathcal{H}$. Let T be a self-adjoint operator on $\mathcal{H}$. Suppose that $\mathscr{D}_0 \subset \mathscr{D}(T)$ is a core for T satisfying $\mathscr{D}_0 \subset \mathscr{D}(T_\alpha)$ for each $\alpha \in \mathscr{A}$. Suppose that

$$\lim_\alpha T_\alpha\xi = T\xi \qquad \text{for each } \xi \in \mathscr{D}_0$$

Prove that (T_α) converges in the strong resolvent sense to T.

Hint. Use Pb. 6.2 and part 2 of Pb. 6.16. □

The following corollary can be used to prove Thm. 5.11.9 without the assumption of adjoint commutativity; that is, it can be used to show that for any subset $\mathfrak{S} \subset \mathfrak{L}(\mathcal{H})$, the SOT* closure $W^*(\mathfrak{S})$ of $\mathbb{C}[\mathfrak{S} \cup \mathfrak{S}^*]$ is a unital *-subalgebra of $\mathfrak{L}(\mathcal{H})$.

Corollary 6.12.7. *Let (T_α) be a net of self-adjoint operators in $\mathfrak{L}(\mathcal{H})$ converging in SOT to a self-adjoint $T \in \mathfrak{L}(\mathcal{H})$. Then for each bounded continuous function $f : \mathbb{R} \to \mathbb{C}$, the net of bounded normal operators $(f(T_\alpha))$ converges in SOT* to $f(T)$.*

Proof. By Pb. 6.17, (T_α) converges in the strong resolvent sense to T. The conclusion then follows from Pb. 6.15. □

The following problem introduces the appropriate version of the finite-rank approximation for self-adjoint operators.

Problem 6.18. Assume that $\mathcal{H}$ is infinite-dimensional and separable. Let T be a self-adjoint operator on $\mathcal{H}$. Let $(f_n)_{n\in\mathbb{Z}_+}$ be an orthonormal basis of $\mathrm{Rng}(T+\mathbf{i})$. Let $e_n = (T+\mathbf{i})^{-1}f_n$. Let $E_n \in \mathfrak{L}(\mathcal{H})$ be the projection of $\mathcal{H}$ onto $\mathrm{Span}\{e_1, \dots, e_n\}$. Show that each E_nTE_n is a bounded self-adjoint operator on $\mathcal{H}$, and that $(E_nTE_n)_{n\in\mathbb{Z}_+}$ converges in the strong resolvent sense to T.

7 Examples of unbounded self-adjoint operators

7.1 The positive self-adjoint operator T^*T

Let $\mathcal{H},\mathcal{K}$ be Hilbert spaces.

Definition 7.1.1. An unbounded operator T on $\mathcal{H}$ is called a **positive operator** if $\langle\xi|T\xi\rangle \geqslant 0$ for each $\xi \in \mathcal{H}$. The operator T is called **strictly positive** if $T-\lambda$ is positive for some $\lambda \in \mathbb{R}_{>0}$. The operator T is called **lower bounded** if $T+\lambda$ is positive for some $\lambda \in \mathbb{R}$.

Positive operators are clearly Hermitian.

Many Hermitian operators arising in partial differential equations and quantum mechanics are positive. For example, if $\Omega \subset \mathbb{R}^n$ is open, then the negative Laplacian $-\Delta = -(\partial_{x_1}^2 + \cdots + \partial_{x_n}^2)$ on $L^2(\Omega, m)$ with domain $C_c^\infty(\Omega)$ is positive, since

$$-\langle\xi|\Delta\xi\rangle = \langle\nabla\xi|\nabla\xi\rangle \qquad \text{for each } \xi \in C_c^\infty(\Omega) \tag{7.1}$$

In quantum mechanics, one also considers the Hamitonian operator $H = -\Delta + V$ on $L^2(\mathbb{R}^N, m)$ with domain $C_c^\infty(\mathbb{R}^N)$, where V is the potential function. Of course, such H is in general not closed, and hence not self-adjoint. Therefore, in order to obtain a spectral theorem for H (so that, for example, one can solve the Schrödinger equation $\mathbf{i}\partial_t\xi(t) = H\xi(t)$ with initial condition $\xi(0) = \xi$ using the Borel functional calculus $\xi(t) = e^{-\mathbf{i}tH}\xi$), one must extend H to a positive self-adjoint operator.

The goal of this section is to show that T^*T is a positive self-adjoint operator if T is a closed operator, cf. Thm. 7.1.7. Using this result, we show that $-\Delta$ admits two canonical positive self-adjoint extensions, the Dirichlet and Neumann Laplacians. The study of positive self-adjoint extensions for the more general Hamiltonian operator $-\Delta + V$ will be addressed in Sec. 7.2.

7.1.1 Elementary properties of positive self-adjoint operators

Proposition 7.1.2. *Let T be an unbounded operator on $\mathcal{H}$. Let $\lambda \in \mathbb{R}_{>0}$. The following are equivalent.*

(1) T is self-adjoint and positive.

(2) T is self-adjoint and $\mathrm{Sp}(T) \subset \mathbb{R}_{\geqslant 0}$.

(3) T is positive, and $\mathrm{Rng}(\lambda + T) = \mathcal{H}$.

Proof. $\neg$(2)$\Rightarrow\neg$(1): Assume that T is self-adjoint and $\mathrm{Sp}(T) \not\subset \mathbb{R}_{\geqslant 0}$. By the spectral Thm. 6.9.4, one may assume that $\mathcal{H} = \bigoplus_{\alpha\in\mathscr{I}} L^2(\mathbb{R}, \mu_\alpha)$ where each μ_α is a finite

Borel measure on $\mathbb{R}$, and $T = \mathbf{M}_x$. By Thm. 6.10.10, there exists α such that $\mathrm{Supp}(\mu_\alpha) \nsubseteq \mathbb{R}_{\geqslant 0}$ and hence $\mu_\alpha(\mathbb{R}_{<0}) > 0$. Thus, viewing $\xi = \chi_{\mathbb{R}_{<0}}$ as an element of $L^2(\mathbb{R}, \mu_\alpha)$ (and hence an element of $\mathcal{H}$), we have $\langle\xi|T\xi\rangle = \int_{\mathbb{R}_{<0}} xd\mu_\alpha < 0$. So T is not positive.

(2)⇒(3): Since $-\lambda \notin \mathbb{R}_{\geqslant 0}$ and hence $-\lambda \notin \mathrm{Sp}(T)$, by Rem. 6.10.13 we have $(\lambda + T)^{-1} \in \mathfrak{L}(\mathcal{H})$, and hence $\mathrm{Rng}(\lambda + T) = \mathcal{H}$.

(3)⇒(1): Since T is positive, we have

$$\langle(\lambda + T)\xi|(\lambda + T)\xi\rangle \geqslant \lambda^2\langle\xi|\xi\rangle \qquad \text{for each } \xi \in \mathscr{D}(T)$$

In particular, if $(\lambda + T)\xi = 0$ then $\xi = 0$. Therefore, the unbounded operator $\lambda + T$ is injective and has range $\mathcal{H}$. Let $S : \mathcal{H} \to \mathcal{H}$ be its inverse, whose domain is $\mathscr{D}(S) = \mathrm{Rng}(\lambda + T) = \mathcal{H}$. Then the above inequality becomes

$$\langle\phi|\phi\rangle \geqslant \lambda^2\langle S\phi|S\phi\rangle \qquad \text{for each } \phi \in \mathcal{H}$$

Thus $S \in \mathfrak{L}(\mathcal{H})$. Moreover, since the diagonal reflection of a **positive graph** (i.e., a linear subspace $\mathfrak{G}$ of $\mathcal{H} \oplus \mathcal{H}$ satisfying $\langle\xi|\eta\rangle \geqslant 0$ for each $\xi \oplus \eta \in \mathfrak{G}$) is clearly also positive, we conclude that $S \geqslant 0$. In particular, we have $S^* = S$. By Prop. 6.6.12, we have $(\lambda + T)^* = (S^{-1})^* = (S^*)^{-1} = S^{-1} = \lambda + T$, and hence $\lambda + T^* = \lambda + T$. Therefore T is self-adjoint. □

Example 7.1.3. Let $T_1, \dots, T_N$ be strongly commuting self-adjoint operators on $\mathcal{H}$. Let $f : \mathbb{R}^N \to \mathbb{R}_{\geqslant 0}$ be Borel. By Thm. 6.9.4, $T_1, \dots, T_N$ can be viewed as $\mathbf{M}_{x_1}, \dots, \mathbf{M}_{x_N}$. Therefore, by Thm. 6.10.2, $f(T_\bullet)$ can be viewed as the multiplication operator $\mathbf{M}_f$. The positivity of $f(T_\bullet)$ is then easy to check.

Remark 7.1.4. Suppose that T is a positive self-adjoint operator on $\mathcal{H}$. The **positive square root** is defined by $\sqrt{T}$ using the Borel functional calculus for the function $f(x) = \sqrt{x}$ on $\mathbb{R}_{\geqslant 0}$. By Exp. 7.1.3, $\sqrt{T}$ is a positive self-adjoint operator.

By Exe. 6.10.3, we have $\sqrt{T} \cdot \sqrt{T} = (\sqrt{T})^2$ where $(\sqrt{T})^2$ is defined by the Borel functional calculus of $\sqrt{T}$ with respect to the function $g(x) = x^2$. It follows from the composition law (Thm. 6.10.5) that

$$\sqrt{T} \cdot \sqrt{T} = (\sqrt{T})^2 = T$$

A similar result holds for the n-th root of T where $n \in \mathbb{Z}_+$. □

Example 7.1.5. Let T be a positive self-adjoint operator on $\mathcal{H}$. Let $x \in \mathfrak{L}(H)$ be positive. Then $x + T$ is a positive self-adjoint operator on $\mathcal{H}$.

Proof. By Pb. 6.1, we have $(x + T)^* = T^* + x^* = x + T$. Thus $x + T$ is self-adjoint. Clearly $x + T$ is positive. □

7.1.2 The positive self-adjoint operator T^*T

Definition 7.1.6. For each n.d.d. unbounded operator $T : \mathcal{H} \to \mathcal{K}$, the **graph projection** denotes the linear map

$$\Pi_T : \mathscr{G}(T) \to \mathcal{H} \qquad \xi \oplus T\xi \mapsto \xi \tag{7.2}$$

We abbreviate Π_T to Π when no confusion arises.

The following theorem was proved by von Neumann in [vN32b].

Theorem 7.1.7. *Let $T : \mathcal{H} \to \mathcal{K}$ be a closed operator. Then T^*T is a positive self-adjoint operator on $\mathcal{H}$, $\mathscr{D}(T^*T)$ is a core for T, and $\Pi_T\Pi_T^* : \mathcal{H} \to \mathcal{H}$ is the everywhere-defined bounded inverse of $\mathbf{1} + T^*T$.*

Note that since T is closed, the graph $\mathscr{G}(T)$ is a Hilbert space. Thus Π_T is a bounded linear map between Hilbert spaces. Hence, its adjoint Π_T^* belongs to $\mathfrak{L}(\mathcal{H}, \mathscr{G}(T))$.

Proof. For each $\xi \in \mathscr{D}(T^*T) = \mathscr{D}(\mathbf{1} + T^*T)$, we clearly have

$$\langle \xi | (\mathbf{1} + T^*T)\xi \rangle = \langle \xi | \xi \rangle + \langle T\xi | T\xi \rangle \geqslant 0$$

Thus, $\mathbf{1} + T^*T$ is an n.d.d. positive operator on $\mathcal{H}$.

Step 1. Note that Π_T is injective with dense range $\mathrm{Rng}(\Pi_T) = \mathscr{D}(T)$. Its inverse is the unbounded operator

$$\Psi_T : \mathcal{H} \to \mathscr{G}(T) \qquad \xi \mapsto \xi \oplus T\xi$$
$$\mathscr{D}(\Psi_T) = \mathscr{D}(T)$$

We abbreviate Ψ_T to Ψ when no confusion arises. In this step, we prove

$$\Psi_T^* \Psi_T = \mathbf{1} + T^*T \tag{7.3}$$

Choose any $\xi \in \mathscr{D}(\Psi_T) = \mathscr{D}(T)$. Then for each $\eta \in \mathscr{D}(\Psi_T) = \mathscr{D}(T)$ we have

$$\langle \Psi\eta | \Psi\xi \rangle = \langle \eta | \xi \rangle + \langle T\eta | T\xi \rangle$$

If $\xi \in \mathscr{D}(T^*T)$, then $\langle T\eta | T\xi \rangle = \langle \eta | T^*T\xi \rangle$, and hence $\langle \Psi\eta | \Psi\xi \rangle = \langle \eta | (\mathbf{1} + T^*T)\xi \rangle$, showing that $\Psi\xi \in \mathscr{D}(\Psi^*)$ and $\Psi^*\Psi\xi = \xi + T^*T\xi$. Similarly, if $\xi \in \mathscr{D}(\Psi^*\Psi)$, then $\xi \in \mathscr{D}(T^*T)$ and $T^*T\xi = \xi - \Psi^*\Psi\xi$. This proves (7.3).

Step 2. Since $\Psi = \Pi^{-1}$, by Prop. 6.6.12 we have $\Psi^* = (\Pi^*)^{-1}$ where Ψ^* and Π^* are injective with dense range. Therefore, (7.3) implies

$$\mathbf{1} + T^*T = (\Pi^*)^{-1}\Pi^{-1} = (\Pi \cdot \Pi^*)^{-1} \tag{7.4}$$

where the last equality is due to Prop. 6.2.21. In particular,

$$\mathrm{Rng}(\mathbf{1} + T^*T) = \mathscr{D}(\Pi \cdot \Pi^*) = \mathcal{H}$$

This proves that $\mathbf{1} + T^*T$ has range $\mathcal{H}$. To prove that T^*T is positive self-adjoint, by Prop. 7.1.2, it remains to prove that $\mathscr{D}(T^*T)$ is dense. This will follow from the stronger fact that $\mathscr{D}(T^*T)$ is a core for T, to be proved in the next step.

Step 3. Observe that $\mathscr{D}(T^*T) = \mathscr{D}(\mathbf{1} + T^*T)$ equals $\mathrm{Rng}(\Pi \cdot \Pi^*)$ due to (7.4). On the other hand, since $\Psi = \Pi^{-1}$, by Prop. 6.2.7 we have

$$\Psi(\Pi \cdot \Pi^*) = (\Psi\Pi) \cdot \Pi^* = \mathrm{id}_{\mathscr{D}(\Pi)}\Pi^* = \mathrm{id}_{\mathcal{H}}\Pi^* = \Pi^*$$

Thus Ψ maps $\mathscr{D}(T^*T) = \mathrm{Rng}(\Pi \cdot \Pi^*)$ to $\mathrm{Rng}(\Pi^*)$. As observed in Step 2, Π^* has dense range in $\mathscr{G}(T)$. This shows that $\Psi\mathscr{D}(T^*T)$ is dense subspace of $\mathscr{G}(T)$. Since

$$\Psi\mathscr{D}(T^*T) = \{\xi \oplus T\xi : \xi \in \mathscr{D}(T^*T)\} = \mathscr{G}(T|_{\mathscr{D}(T^*T)})$$

the graph of $T|_{\mathscr{D}(T^*T)}$ is dense in that of T. Namely, $\mathscr{D}(T^*T)$ is a core for T. $\square$

Exercise 7.1.8. Let T be a self-adjoint operator on $\mathcal{H}$. Use the spectral theorem to give a direct proof that $\mathscr{D}(T^2)$ is a core for $\mathscr{D}(T)$. (More generally, prove that $\mathscr{D}(T^n)$ is a core for $\mathscr{D}(T)$ for each $n \in \mathbb{Z}_+$.)

Remark 7.1.9. We now interpret the fact that $\mathscr{D}(T^*T)$ is a core for T in the context of differential operators. By the proof of Thm. 7.1.7, this is equivalent to the statement that Π_T^* has dense range in $\mathscr{G}(T)$. Viewing $\mathscr{G}(T)^*$ as a Hilbert space unitarily equivalent to the conjugate space $\mathscr{G}(T)^{\complement}$ via the Riesz isomorphism (cf. Def. 3.5.4), the density of $\mathrm{Rng}(\Pi_T^*)$ in $\mathscr{G}(T)$ is equivalent to saying that

$$\{\psi \circ \Pi_T : \mathscr{G}(T) \to \mathbb{C} \text{ where } \psi \in \mathcal{H}^*\}$$

is a dense linear subspace of $\mathscr{G}(T)^*$.

In other words, if $\mathscr{D}(T)$ is equipped with the graph inner product of T, and if $\mathscr{D}(T)^*$ is equipped with the inner product corresponding to $\mathscr{D}(T)^{\complement}$ via the Riesz isomorphism, then the bounded linear functionals on $\mathcal{H}$, when restricted to functionals on $\mathscr{D}(T)$, form a dense linear subspace of the dual space $\mathscr{D}(T)^*$.

Now take $\mathcal{H} = L^2(\Omega, m)$ where $\Omega \subset \mathbb{R}^N$ is open, and set $T = \overline{\nabla}$. (See Subsec. 6.6.3 for the notation.) The above conclusion then says that pairings with elements of $L^2(\Omega, m)$ form a dense linear subspace of $\boldsymbol{H^{-1}(\Omega)}$, the dual space of $H_0^1(\Omega) = \mathscr{D}(\overline{\nabla})$. (Briefly speaking, $L^2(\Omega, m)$ is dense in $H^{-1}(\Omega)$.) $\square$

7.1.3 The polar decomposition

Definition 7.1.10. An element $\Phi \in \mathfrak{L}(\mathcal{H}, \mathcal{K})$ is called a **partial isometry** if there exists a closed linear subspace $V \subset \mathcal{H}$ such that $\Phi|_V : V \to \mathcal{K}$ is an isometry, and $\Phi|_{V^\perp} = 0$. Note that V is uniquely determined by Φ by the relation

$$V^\perp = \mathrm{Ker}(\Phi)$$

We call $V = \mathrm{Ker}(\Phi)^\perp$ the **source space** of Φ, and call $\mathrm{Rng}(\Phi)$ the **target space** of Φ.

The following Thm. 7.1.11, also due to [vN32b], is an easy consequence of Thm. 7.1.7. We will use Thm. 7.1.11 in the study of the Friedrichs extension (see Thm. 7.2.9). It will also play a crucial role in the analysis of unbounded operators with compact resolvents (see Thm. 8.8.10).

Note that the fact that $\mathscr{D}(T^*T)$ is a core for T will play a crucial role in the proof of Thm. 7.1.11.

Theorem 7.1.11. *Let $T : \mathcal{H} \to \mathcal{K}$ be a closed operator. Then $\sqrt{T^*T}$ is a positive self-adjoint operator on $\mathcal{H}$, and there exists a unique partial isometry $U \in \mathfrak{L}(\mathcal{H}, \mathcal{K})$ with source space $\overline{\mathrm{Rng}(\sqrt{T^*T})}$ satisfying*

$$T = U\sqrt{T^*T} \tag{7.5}$$

*We call (7.5) the **(left) polar decomposition** of T, and call $|T| := \sqrt{T^*T}$ the **absolute value** of T.*

According to this theorem, the closed operator T and the positive self-adjoint operator $\sqrt{T^*T}$ share the same domain and the same graph inner product. This implies, for example, that $H_0^1(\Omega) = \mathscr{D}(\sqrt{-\Delta_{\mathrm{D}}})$ (cf. Rem. 7.1.13).

Proof. Let $H = \sqrt{T^*T}$. Then H is clearly positive and self-adjoint, and hence $H^*H = H^2 = T^*T$. (Recall Rem. 7.1.4 for why we have $H^2 = T^*T$.) Therefore, by Thm. 7.1.7, $\mathscr{D}(T^*T)$ is a common core for H and T. Thus, $H\mathscr{D}(T^*T)$ is dense in $\mathrm{Rng}(H)$. Therefore, any bounded linear operator $\mathcal{H} \to \mathcal{K}$ is uniquely determined by its values on $H\mathscr{D}(T^*T)$ and on $\mathrm{Rng}(H^\perp)$. Thus, the elements $U \in \mathfrak{L}(\mathcal{H}, \mathcal{K})$ satisfying $T|_{\mathscr{D}(T^*T)} = UH|_{\mathscr{D}(T^*T)}$ and vanishing on $\mathrm{Rng}(H)^\perp$ are unique.

Let us prove the existence part. For each $\xi \in \mathscr{D}(T^*T) = \mathscr{D}(H^*H)$, we have

$$\langle H\xi|H\xi\rangle = \langle \xi|H^*H\xi\rangle = \langle \xi|T^*T\xi\rangle = \langle T\xi|T\xi\rangle$$

Therefore, by Lem. 5.10.20, there exits a linear isometry U from the closure of $H\mathscr{D}(T^*T)$ (which equals $\overline{\mathrm{Rng}(H)}$ due to the first paragraph) to $\mathcal{K}$ satisfying $T\xi = UH\xi$ for each $\xi \in \mathscr{D}(T^*T)$. By zero extension, the map U is extended to a partial isometry $U \in \mathfrak{L}(\mathcal{H}, \mathcal{K})$ with source space $\overline{\mathrm{Rng}(H)}$ satisfying

$$T|_{\mathscr{D}(T^*T)} = UH|_{\mathscr{D}(T^*T)}$$

Since $\mathscr{D}(T^*T)$ is a core for T and H, using (6.16) one easily checks that the closures of the two sides above are equal to T and UH respectively. This proves (7.5). $\square$

7.1.4 The Dirichlet and Neumann Laplacians

Let $\Omega \subset \mathbb{R}^N$ be an open subset. Recall Subsec. 6.6.3 for the Sobolev spaces $H_0^1(\Omega) = \mathscr{D}(\overline{\nabla})$ and $H^1(\Omega) = \mathscr{D}(\mathrm{div}^*)$, equipped with the (complete) graph inner products of $\overline{\nabla}$ and div^* respectively. Note that $(\overline{\nabla})^* = \nabla^*$ and $(\overline{\mathrm{div}})^* = \mathrm{div}^*$ due to Thm. 6.6.9.

Definition 7.1.12. The unbounded operators

$$\mathbf{\Delta_D} = -\overline{\nabla}^* \cdot \overline{\nabla} \qquad \mathbf{\Delta_N} = -\overline{\mathrm{div}} \cdot \overline{\mathrm{div}}^* \tag{7.6}$$

(which are negative and self-adjoint due to Thm. 7.1.7) are called the **Dirichlet Laplacian** and the **Neumann Laplacian** respectively.

Remark 7.1.13. Recall from Thm. 7.1.11 that $\mathscr{D}(T) = \mathscr{D}(\sqrt{T^*T})$ for any closed operator T. By letting T be $\overline{\nabla}$ and div^* respectively, we obtain

$$H_0^1(\Omega) = \mathscr{D}(\sqrt{-\Delta_{\mathrm{D}}}) \qquad H^1(\Omega) = \mathscr{D}(\sqrt{-\Delta_{\mathrm{N}}})$$

Remark 7.1.14. For each $\xi \in C_c^\infty(\Omega)$, we have $\overline{\nabla}\xi = \mathrm{div}^*\xi = \nabla\xi$. Therefore

$$\langle\eta|\nabla^*\overline{\nabla}\xi\rangle = \langle\nabla\eta|\nabla\xi\rangle = \langle\mathrm{div}^*\eta|\mathrm{div}^*\xi\rangle = \langle\eta|\overline{\mathrm{div}}\cdot\mathrm{div}^*\xi\rangle$$

for each $\xi, \eta \in C_c^\infty(\Omega)$. Moreover, integration by parts gives $\langle\nabla\eta|\nabla\xi\rangle = \langle\eta| - \Delta\xi\rangle$. This proves

$$\Delta_{\mathrm{D}}|_{C_c^\infty(\Omega)} = \Delta_{\mathrm{N}}|_{C_c^\infty(\Omega)} = \Delta \tag{7.7}$$

where Δ is the usual Laplacian operator on $L^2(\Omega, m)$ with domain $\mathscr{D}(\Delta) = C_c^\infty(\Omega)$. Thus, both $-\Delta_{\mathrm{D}}$ and $-\Delta_{\mathrm{N}}$ are positive self-adjoint extensions of $-\Delta$.

Remark 7.1.15. Recall from Rem. 6.6.19 that $\mathscr{D}(\overline{\nabla}) = H_0^1(\Omega)$ consists of functions $f \in H^1(\Omega)$ "vanishing on the boundary $\partial\Omega$". Therefore, if $f \in \mathscr{D}(\Delta_{\mathrm{D}})$, then $f \in H^1(\Omega)$ vanishes on $\partial\Omega$ (i.e., f satisfies the **Dirichlet boundary condition**).

If $f \in \mathscr{D}(\Delta_{\mathrm{N}})$, then $f \in \mathscr{D}(\Delta^*)$ (since $\Delta \subset \Delta_{\mathrm{N}}$ implies $\Delta_{\mathrm{N}} \subset \Delta^*$). Heuristically, Green's identity

$$\langle\mathrm{div}^* g|\mathrm{div}^* f\rangle = -\int_\Omega \overline{g}\Delta^* f + \int_{\partial\Omega} \overline{g}\partial_\nu f \tag{7.8}$$

holds for $g \in \mathscr{D}(\mathrm{div}^*) = H^1(\Omega)$, where $\partial_\nu f$ denotes the normal derivative of f along the boundary $\partial\Omega$. Since the LHS above equals $\langle g|\Delta_{\mathrm{N}} f\rangle$, the boundary term on the right must be bounded with respect to the $L^2(\Omega, m)$-norm of g. This implies the the **Neumann boundary condition**

$$\partial_\nu f|_{\partial\Omega} = 0$$

See [Sch-U, Sec. 10.6] for a rigorous treatment.

By Prop. 7.1.2, for each $\lambda \in \mathbb{R}_{>0}$, the operators $\lambda - \Delta_{\mathrm{D}}$ and $\lambda - \Delta_{\mathrm{N}}$ have range $L^2(\Omega, m)$. Intuitively, this means that for each $g \in L^2(\Omega, m)$, the differential equation

$$(\lambda - \Delta)f = g$$

subject to the Dirichlet boundary condition $f|_{\partial\Omega} = 0$ (resp. the Neumann boundary condition $\partial_\nu f|_{\partial\Omega} = 0$) admits a weak solution, i.e. a solution $f \in H^1(\Omega)$. $\square$

We will return to the discussion in Rem. 7.1.15 and develop more general results on the existence of weak solutions to differential equations in Subsec. 7.9.7.

The description of $\mathscr{D}(\Delta_{\mathrm{D}})$ and $\mathscr{D}(\Delta_{\mathrm{N}})$ can be made explicit when Ω is a bounded open interval, as illustrated by the following example.

Example 7.1.16. Assume that Ω is a bounded open interval $I = (a,b)$ in $\mathbb{R}$. Let $T_0 = -\mathbf{i}d/dx$ with domain $C_c^\infty(I)$. Recall Thm. 6.8.5 for the description of $T := \overline{T_0}$ and T^* in terms of absolutely continuous functions on $\overline{I} = [a,b]$. In particular, $\mathscr{D}(T) = H_0^1(I)$ and $\mathscr{D}(T^*) = H^1(I)$ are described by (6.50). Clearly

$$\Delta_{\mathrm{D}} = -T^*T \qquad \Delta_{\mathrm{N}} = -TT^*$$

From this one concludes:

- The domain $\mathscr{D}(\Delta_{\mathrm{D}})$ consists of all $f \in AC(\overline{I})$ satisfying

  $$f(a) = f(b) = 0 \qquad f' \in AC(\overline{I}) \qquad f'' \in L^2(I,m)$$

  For each such f we have $\Delta_{\mathrm{D}} f = -f''$.

- The domain $\mathscr{D}(\Delta_{\mathrm{N}})$ consists of all $f \in AC(\overline{I})$ satisfying

  $$f' \in AC(\overline{I}) \qquad f'(a) = f'(b) = 0 \qquad f'' \in L^2(I,m)$$

  For each such f we have $\Delta_{\mathrm{N}} f = -f''$.

We will see in Cor. 7.6.9 that when $\Omega = \mathbb{R}^N$, the Dirichlet and Neumann Laplacians are equal.

7.2 The Friedrichs extension

Let $\mathcal{H}$ be a Hilbert space.

The goal of this section is to show that every positive unbounded operator admits a canonical positive self-adjoint extension, known as the Friedrichs extension, first introduced by Friedrichs in [Fri34a].

As one may notice in Subsec. 7.1.4, the main reason that $-\Delta$ admits a positive self-adjoint extension is that $-\Delta$ can be written as $\nabla^*\nabla$. A positive self-adjoint extension can then be taken to be $\overline{\nabla}^*\overline{\nabla}$. This observation suggests that, in order to construct a positive self-adjoint extension for any positive unbounded operator A, one should first construct the analogous gradiant operator.

7.2.1 Abstract gradient operators

Definition 7.2.1. Let A be a positive unbounded operator on $\mathcal{H}$. An **abstract gradient operator** of A denotes an unbounded operator $T : \mathcal{H} \to \mathcal{K}$ with $\mathcal{K}$ a Hilbert space, such that

$$\mathscr{D}(T) = \mathscr{D}(A) \qquad A = T^*T \tag{7.9}$$

In practice, to check that T is an abstract gradient operator, it suffices to compute $\langle T\xi|T\eta\rangle$ rather than determining the adjoint T^*:

Lemma 7.2.2. *Let A be a positive unbounded operator on $\mathcal{H}$. Let $T:\mathcal{H}\to\mathcal{K}$ be an unbounded operator with domain $\mathscr{D}(T)=\mathscr{D}(A)$. Then T is an abstract gradient operator of A iff*

$$\langle\xi|A\eta\rangle=\langle T\xi|T\eta\rangle \qquad \text{for each } \xi,\eta\in\mathscr{D}(A) \tag{7.10}$$

Proof. If $A=T^*T$, then $\langle\xi|A\eta\rangle=\langle\xi|T^*T\eta\rangle=\langle T\xi|T\eta\rangle$ for each $\xi,\eta\in\mathscr{D}(A)=\mathscr{D}(T)$. Conversely, if (7.10) holds, then for each $\eta\in\mathscr{D}(A)$, since (7.10) holds for all $\xi\in\mathscr{D}(T)$, the definition of adjoint operators shows that $T\eta\in\mathscr{D}(T^*)$ and $T^*T\eta=A\eta$. Therefore $A=T^*T$. □

Proposition 7.2.3. *Let A be a positive unbounded operator on $\mathcal{H}$. Then any abstract gradient operator $T:\mathcal{H}\to\mathcal{K}$ of A is closable.*

Proof. Eq. (7.9) implies that $\mathrm{Rng}(T)\subset\mathscr{D}(T^*)$. By Prop. 6.6.11, we have $\mathrm{Rng}(T)^\perp\subset\mathscr{D}(T^*)$. Therefore, $\mathscr{D}(T^*)$ contains the dense subspace $\mathrm{Rng}(T)+\mathrm{Rng}(T)^\perp$ of $\mathcal{K}$. It follows from Thm. 6.6.9 that T is closable. □

Proposition 7.2.4. *Let A be a positive unbounded operator on $\mathcal{H}$. If $T:\mathcal{H}\to\mathcal{K}$ and $S:\mathcal{H}\to\mathcal{M}$ are abstract gradient operators of A, then there is a unique partial isometry $U:\mathcal{K}\to\mathcal{M}$ with source space $\overline{\mathrm{Rng}(T)}$ satisfying*

$$S=UT$$

Proof. Any such U must be determined by its values on $\mathrm{Rng}(T)$ and on $\mathrm{Rng}(T)^\perp$. Therefore such U are unique. As for the existence, note that for each $\xi\in\mathscr{D}(A)$ we have (cf. Lem. 7.2.2)

$$\langle S\xi|S\xi\rangle=\langle\xi|A\xi\rangle=\langle T\xi|T\xi\rangle$$

It follows from Lem. 5.10.20 that there is a linear isometry $U:\overline{\mathrm{Rng}(T)}\to\mathcal{M}$ sending each $T\xi$ to $S\xi$. By zero extension, the operator U is extended to the desired partial isometry $U\in\mathfrak{L}(\mathcal{K},\mathcal{M})$. □

Theorem 7.2.5. *Let A be a positive unbounded operator on $\mathcal{H}$. Then A admits an abstract gradient operator. Moreover, if $T:\mathcal{H}\to\mathcal{K}$ and $S:\mathcal{H}\to\mathcal{M}$ are abstract gradient operators of A (which are closable by Prop. 7.2.3), then $\overline{T}^*\overline{T}=\overline{S}^*\overline{S}$.*

Recall again from Thm. 6.6.9 that $T^*=(\overline{T})^*$ and $S^*=(\overline{S})^*$.

Proof. Since A is positive, the sesquilinear form

$$\omega_A:\mathscr{D}(A)\times\mathscr{D}(A)\to\mathbb{C}\qquad(\xi,\eta)\mapsto\langle\xi|A\eta\rangle$$

is positive. By Exe. 3.2.2, $\mathscr{N} = \{\xi \in \mathscr{D}(A) : \langle\xi|A\xi\rangle = 0\}$ is a linear subspace of $\mathscr{D}(A)$, and ω_A descends to an inner product on $\mathscr{D}(A)/\mathscr{N}$. Let $\mathcal{K}$ be the Hilbert space completion of this inner product space $\mathscr{D}(A)/\mathscr{N}$ (cf. Pb. 3.2). Define an unbounded operator

$$T : \mathcal{H} \to \mathcal{K} \qquad \xi \mapsto \xi + \mathscr{N}$$

with domain $\mathscr{D}(T) = \mathscr{D}(A)$. Then for each $\xi, \eta \in \mathscr{D}(A)$,

$$\langle T\xi|T\eta\rangle_{\mathcal{K}} = \langle \xi + \mathscr{N}|\eta + \mathscr{N}\rangle_{\mathcal{K}} = \omega_A(\xi|\eta) = \langle\xi|A\eta\rangle_{\mathcal{H}}$$

Therefore, by Lem. 7.2.2, T is an abstract gradient operator of A.

Next, assume that $S : \mathcal{H} \to \mathcal{M}$ is also an abstract gradient operator of A. Let $U \in \mathfrak{L}(\mathcal{K}, \mathcal{M})$ be as in Prop. 7.2.4. Using (6.16) one easily checks that $\overline{S} = U\overline{T}$. By Pb. 6.1, we have $(U\overline{T})^* = \overline{T}^*U^*$, and hence

$$\overline{S}^*\overline{S} = \overline{T}^*U^*U\overline{T} = \overline{T}^*\overline{T}$$

where the last identity is due to the easy fact that U^*U is the projection onto the source space $\overline{\mathrm{Rng}(T)}$ of U (cf. Pb. 5.2). □

7.2.2 The Friedrichs extension

Definition 7.2.6. Let A be a positive unbounded operator on $\mathcal{H}$. Let $T : \mathcal{H} \to \mathcal{K}$ be an abstract gradient operator of A. By Thm. 7.2.5, such T exists, and $T^*\overline{T}$ is independent of the choice of T. The unbounded operator $T^*\overline{T}$ is positive self-adjoint (due to Thm. 7.1.7) and extends A since

$$A = T^*T \subset T^*\overline{T}$$

We call $\boldsymbol{A_{\mathrm{F}}} := T^*\overline{T} \equiv \overline{T}^*\overline{T}$ the **Friedrichs extension** of A.

Exercise 7.2.7. Use the Friedrichs extension to give an alternative proof of the Stieltjes moment theorem (cf. Thm. 4.2.9-2), similar to the proof of the Hamburger moment theorem using Thm. 6.11.1 as presented in Exp. 6.11.2.

Example 7.2.8. Let $\Omega \subset \mathbb{R}^N$ be open. Since the negative smooth Laplacian $-\Delta = \nabla^*\nabla$ (with domain $\mathscr{D}(\Delta) = \mathscr{D}(\nabla) = C_c^\infty(\Omega)$) has abstract gradient operator ∇, the Friedrichs extension of $-\Delta$ is the negative Dirichlet Laplacian $-\Delta_{\mathrm{D}} = \nabla^*\overline{\nabla}$ (cf. Def. 7.1.12).

Since abstract gradient operators are not unique, but only unique up to partial isometries, it is desirable to give a characterization of the Friedrichs extension that is independent of the particular choice of abstract gradient operator:

Theorem 7.2.9. *Let A be a positive unbounded operator on $\mathcal{H}$. Let B be a positive self-adjoint operator on $\mathcal{H}$ extending A. Then the following are equivalent:*

(1) B is the Friedrichs extension of A.

(2) $\mathscr{D}(A)$ is a core for $\sqrt{B}$.

(3) $\mathscr{D}(A)$ is a core for $\sqrt{B}|_{\mathscr{D}(B)}$.

Note that a core for an unbounded operator is always assumed to be a linear subspace of its domain. Note also that $\sqrt{B}|_{\mathscr{D}(B)}$ can be defined since $B = \sqrt{B} \cdot \sqrt{B}$ implies that $\mathscr{D}(B) \subset \mathscr{D}(\sqrt{B})$.

Proof. (1)⇒(2): Assume (1). Then $B = \overline{T}^*\overline{T}$ where $T : \mathcal{H} \to \mathcal{K}$ is an abstract gradient operator of A. In particular, $\mathscr{D}(A) = \mathscr{D}(T)$. Clearly $\mathscr{D}(T)$ is a core for $\overline{T}$. By the polar decompsition of $\overline{T}$ (cf. Thm. 7.1.11), the operators $\overline{T}$ and $\sqrt{\overline{T}^*\overline{T}} = \sqrt{B}$ share the same domain and cores. Thus $\mathscr{D}(T)$ is a core for $\sqrt{B}$. This proves (2).

(2)⇔(3): This follows immediately from the fact that $\mathscr{D}(A) \subset \mathscr{D}(B)$ (because $A \subset B$), and that $\mathscr{D}(B)$ is a core for $\sqrt{B}$ (due to Thm. 7.1.7 or Exe. 7.1.8).

(2)⇒(1): Assume (2). Then $\mathscr{D}(A) \subset \mathscr{D}(\sqrt{B})$ and $\overline{\sqrt{B}|_{\mathscr{D}(A)}} = \sqrt{B}$. In particular, by Thm. 6.6.9, we have $(\sqrt{B}|_{\mathscr{D}(A)})^* = \sqrt{B}^* = \sqrt{B}$. Since $A \subset B$, we obtain

$$A = B|_{\mathscr{D}(A)} = \sqrt{B} \cdot \sqrt{B}|_{\mathscr{D}(A)} = (\sqrt{B}|_{\mathscr{D}(A)})^* \cdot \sqrt{B}|_{\mathscr{D}(A)}$$

This proves that $\sqrt{B}|_{\mathscr{D}(A)}$ is an abstract gradient operator of A. Thus, the Friedrichs extension of A is $\sqrt{B} \cdot \sqrt{B} = B$. □

Using Thm. 7.2.9, we can derive several basic properties of Friedrichs extensions. These properties can also be proved directly from Def. 7.2.6, though the proofs are somewhat longer.

Proposition 7.2.10. *Let A be a positive self-adjoint operator on $\mathcal{H}$. Then the Friedrichs extension of A is A itself.*

Proof. Since $\sqrt{A}\sqrt{A} = A$ implies $\mathscr{D}(A) \subset \mathscr{D}(\sqrt{A})$, the space $\mathscr{D}(A)$ is clearly a core for $\sqrt{A}|_{\mathscr{D}(A)}$. Thus, Thm. 7.2.9 implies that $A_{\mathrm{F}} = A$. □

Proposition 7.2.11. *Let A be a positive unbounded operator on $\mathcal{H}$. Then the Friedrichs extensions of A and $\overline{A}$ are equal.*

Note that the closure of a positive unbounded operator is clearly positive.

Proof. Let B be the Friedrichs extension of A. Since self-adjoint operators are closed, we have $\overline{A} \subset \overline{B} = B$. By Thm. 7.2.9, $\mathscr{D}(A)$ is a core for $\sqrt{B}|_{\mathscr{D}(B)}$. Since $\mathscr{D}(\overline{A}) \subset \mathscr{D}(B)$, the set $\mathscr{D}(\overline{A})$ is also a core for $\sqrt{B}|_{\mathscr{D}(B)}$. Thus Thm. 7.2.9 implies that B is the Friedrichs extension of $\overline{A}$. □

Proposition 7.2.12. *Let A be a positive unbounded operator on $\mathcal{H}$ with Friedrichs extension B. Let $x \in \mathfrak{L}(\mathcal{H})$ be positive. Then $x + B$ is the Friedrichs extension of $x + A$.*

In particular, for each $\lambda \geqslant 0$, $\lambda + B$ is the Friedrichs extension of $\lambda + A$.

Proof. Throughout the proof we use the obvious fact that

$$\mathscr{D}(A) = \mathscr{D}(x + A) \qquad \mathscr{D}(B) = \mathscr{D}(x + B)$$

By Exp. 7.1.5, $x + B$ is a positive self-adjoint operator. Since $A \subset B$, we have $x + A \subset x + B$. By Thm. 7.2.9, $\mathscr{D}(A)$ is a core for $\sqrt{B}|_{\mathscr{D}(B)}$, and it suffices to prove that $\mathscr{D}(A)$ is a core for $\sqrt{x + B}|_{\mathscr{D}(B)}$.

Note that a subspace being a core means that it is dense in the graph inner product. Let $\gamma = 1 + \|x\|$. For each $\xi \in \mathscr{D}(B) = \mathscr{D}(x + B)$, we have $\xi \in \mathscr{D}(\sqrt{B})$ (since $B = \sqrt{B}\sqrt{B}$) and similarly $\xi \in \mathscr{D}(\sqrt{x + B})$, and

$$\langle\xi|\xi\rangle + \langle\sqrt{B}\xi|\sqrt{B}\xi\rangle \leqslant \langle\xi|\xi\rangle + \langle\sqrt{x + B}\xi|\sqrt{x + B}\xi\rangle \leqslant \gamma(\langle\xi|\xi\rangle + \langle\sqrt{B}\xi|\sqrt{B}\xi\rangle)$$

since

$$\begin{aligned}&\langle\xi|\xi\rangle + \langle\xi|B\xi\rangle \leqslant \langle\xi|\xi\rangle + \langle\xi|(x + B)\xi\rangle \leqslant \langle\xi|\xi\rangle + \|x\|\langle\xi|\xi\rangle + \langle\xi|B\xi\rangle\\ \leqslant&(1 + \gamma)(\langle\xi|\xi\rangle + \langle\xi|B\xi\rangle)\end{aligned}$$

Hence, the graph inner products of $\sqrt{B}$ and $\sqrt{x + B}$ on $\mathscr{D}(B)$ are equivalent in the sense of Def. 7.2.13. Therefore, since $\mathscr{D}(A)$ is a core for $\sqrt{B}|_{\mathscr{D}(B)}$, it is also a core for $\sqrt{x + B}|_{\mathscr{D}(B)}$. □

Definition 7.2.13. Let V be a $\mathbb{C}$-vector space. Two inner products ω_1, ω_2 on V are said to be **equivalent** if there exist $\alpha, \beta \in \mathbb{R}_{>0}$ such that

$$\omega_1(\xi|\xi) \leqslant \alpha \cdot \omega_2(\xi|\xi) \leqslant \beta \cdot \omega_1(\xi|\xi) \qquad \text{for each } \xi \in V$$

More generally:

Definition 7.2.14. Two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ on an $\mathbb{F}$-vector space V are called **equivalent** if there exist $\alpha, \beta \in \mathbb{R}_{>0}$ such that

$$\|\xi\|_1 \leqslant \alpha\|\xi\|_2 \leqslant \beta\|\xi\|_1 \qquad \text{for each } \xi \in V$$

Equivalently, the identity map

$$(V, \|\cdot\|_1) \to (V, \|\cdot\|_2) \qquad \xi \mapsto \xi$$

is bounded, and so is its inverse.

Thanks to Prop. 7.2.12, the Friedrich extensions can be defined for lower bounded operators:

Definition 7.2.15. The **Friedrichs extension** of a lower bounded Hermitian operator A is defined to be

$$\boldsymbol{A_{\mathrm{F}}} := (A+\lambda)_{\mathrm{F}} - \lambda$$

where $\lambda \in \mathbb{R}$ is any number such that $A+\lambda$ is positive, and $(A+\lambda)_{\mathrm{F}}$ is the Friedrichs extension of $A+\lambda$. By Prop. 7.2.12, the definition of A_{F} is independent of the choice of λ. Moreover, if $\lambda \in \mathbb{R}$ is such that $A+\lambda$ is positive, then $A_{\mathrm{F}}+\lambda$ is also positive.

For example, if $H = -\Delta + V$ on $L^2(\Omega, m)$ where $\Omega \subset \mathbb{R}^N$ is open, $V : \Omega \to \mathbb{R}$ is a lower-bounded locally-L^2 function, and $\mathscr{D}(H) = C_c^\infty(\Omega)$, then H is a lower bounded Hermitian operator, and hence admits a Friedrichs extension.

7.3 Spectral theorem for unitary representations of the group $\mathbb{R}^N$

Fix a Hilbert space $\mathcal{H}$. Let $N \in \mathbb{Z}_+$. We adopt the notation

$$tx = t_1 x_N + \cdots + t_N x_N \qquad |x| = \sqrt{x_1^2 + \cdots + x_N^2}$$

if $x = (x_1, \dots, x_N) \in \mathbb{R}^N$ and $t = (t_1, \dots, t_N) \in \mathbb{R}^N$. For each $t \in \mathbb{R}^N$, define

$$\boldsymbol{e_t} : \mathbb{R}^N \to \mathbb{C} \qquad x \mapsto e^{\mathbf{i}tx} \tag{7.11}$$

A typical quantum system often possesses multiple continuous symmetries. For instance, its physical laws may remain invariant under time evolution, spatial translations along certain directions, or rotations about a fixed axis.

The purpose of this section is to examine the relationship between continuous symmetries and self-adjoint operators. Let T be a self-adjoint operator on $\mathcal{H}$. If T represents the time-independent Hamiltonian operator, then, as discussed at the beginning of Sec. 7.1, the family $t \in \mathbb{R} \to e^{-\mathbf{i}tT} \in \mathfrak{L}(\mathcal{H})$ provides the solution for the Schrödinger equation with Hamiltonian operator T, corresponding to the time-translation symmetry of the quantum system. If T is the momentum operator $-\mathbf{i}\partial_{x_i}$ in the direction of x_i, then the family $s \in \mathbb{R} \mapsto e^{\mathbf{i}sT}$ represents the spatial-translation symmetry in that direction. Likewise, if T is the angular-momentum operator about a given axis, then $s \mapsto e^{\mathbf{i}sT}$ describes the corresponding rotational symmetry.

In this section, we give a rigorous definition of an N parameter mutually-commuting symmetries in terms of strongly continuous unitary representations of $\mathbb{R}^N$, and we establish a one-to-one correspondence between such representations and tuples $(T_1, \dots, T_N)$ of strongly commuting self-adjoint operators on $\mathcal{H}$. As we will explain in Rem. 7.3.6, this correspondence can be viewed as establishing a spectral theorem for strongly continuous unitary representations of $\mathbb{R}^N$.

7.3.1 Stone's theorem

Definition 7.3.1. A **unitary representation of $\mathbb{R}^N$** on $\mathcal{H}$ denotes a map

$$U : \mathbb{R}^N \to \mathfrak{L}(\mathcal{H})$$

such that for each $t, s \in \mathbb{R}^N$, $U(t)$ is a unitary operator on $\mathcal{H}$, and

$$U(s+t) = U(s)U(t) \qquad U(0) = \mathbf{1}$$

A unitary representation U is called **strongly continuous** if the map $U : \mathbb{R}^N \to \mathfrak{L}(\mathcal{H})$ is continuous when $\mathfrak{L}(\mathcal{H})$ is equipped with the SOT.

In the literature, a unitary representation of $\mathbb{R}$ is usually called a **one parameter unitary group**.

Remark 7.3.2. Let U be a unitary representation of $\mathbb{R}^N$ on $\mathcal{H}$. Then we have

$$U(-t) = U(t)^{-1} = U(t)^*$$

since $U(t)U(-t) = U(0) = \mathbf{1}$ and similarly $U(-t)U(t) = \mathbf{1}$. Moreover, since $U(t + s) = U(t)U(s)$, to prove that U is strongly continuous, it suffices to check

$$\lim_{t\to 0} U(t) = \mathbf{1} \qquad \text{in SOT}$$

The following theorem by Marshall Stone is the main result of this section. (Stone only proved this theorem for the case $N = 1$.)

Theorem 7.3.3 (Stone's theorem). *There is a one-to-one correspondence between:*

(1) A tuple $(T_1, \dots, T_N)$ of strongly commuting self-adjoint operators on $\mathcal{H}$.

(2) A strongly continuous unitary representation U of $\mathbb{R}^N$ on $\mathcal{H}$.

The correspondence is given as follows: For each $t = (t_1, \dots, t_N) \in \mathbb{R}^N$, let

$$\boldsymbol{e}^{\mathbf{i}\boldsymbol{t}\boldsymbol{T}_\bullet} \equiv e^{\mathbf{i}t_1T_1} \cdots e^{\mathbf{i}t_NT_N}$$

Then $U(t) = \mathrm{e}^{\mathbf{i}tT_\bullet}$ for each $t \in \mathbb{R}^N$. Conversely, for each $1 \leqslant j \leqslant N$, the domain $\mathscr{D}(T_j)$ consists of all $\xi \in \mathcal{H}$ such that the limit

$$\partial_{t_j} U(t)\xi\big|_{t=0} = \lim_{t_j\to 0} \frac{U(0,\dots,t_j,\dots,0)\xi - \xi}{t_j}$$

exists, and for each $\xi \in \mathscr{D}(T_j)$ we have

$$\mathbf{i}T_j\xi = \partial_{t_j} U(t)\xi\big|_{t=0}$$

Definition 7.3.4. If U is a strongly continuous unitary representation of $\mathbb{R}^N$ on $\mathcal{H}$, the tuple $(T_1, \dots, T_N)$ of strongly commuting self-adjoint operators satisfying $U(t) = e^{\mathbf{i}tT}$ for all $t \in \mathbb{R}^N$ is called the (tuple of) **generators** of U.

Remark 7.3.5. By Thm. 6.10.8, the operators $e^{\mathbf{i}t_1T_1}, \dots, e^{\mathbf{i}t_NT_N}$ commute adjointly. Moreover, by Cor. 6.10.6, we have $e^{\mathbf{i}t_jT_j} = e_{t_j}(T_\bullet)$ where $e_{t_j} : \mathbb{R}^N \to \mathbb{C}$ is defined by $e_{t_j}(x_1, \dots, x_N) = e^{\mathbf{i}t_jx_j}$. Since the Borel functional calculus is multiplicative on $\mathscr{Bor}_{\mathrm{b}}(\mathbb{R}^N)$ (cf. part (1) of Thm. 6.10.1), we clearly have

$$e^{\mathbf{i}tT_\bullet} = e_t(T_\bullet)$$

Since e_t ranges in $\mathbb{S}^1$, we must have $e_t(T_\bullet)^*e_t(T_\bullet) = e_t(T_\bullet)e_t(T_\bullet)^* = |e_t|^2(T_\bullet) = \mathbf{1}$, and hence $e^{\mathbf{i}tT_\bullet}$ is unitary.

Remark 7.3.6. The nontrivial part of Stone's theorem lies in the surjectivity of (1)⇒(2), namely, that every strongly continuous unitary representation U can be expressed as $U(t) = e^{\mathbf{i}tT_\bullet}$ for some strong commuting self-adjoint operators $T_1, \dots, T_N$. Since spectral theorems have already been established for such operators, Stone's theorem may be regarded as a spectral theorem for strongly continuous unitary representations of $\mathbb{R}^N$.

More concretely, one formulation of the spectral theorem for $T_\bullet$ asserts a unitary equivalence

$$\mathcal{H} \simeq \bigoplus_{\alpha\in\mathscr{I}} L^2(\mathbb{R}^N, \mu_\alpha)$$

where each μ_α is a finite Borel measure on $\mathbb{R}^N$. Under this equivalence, each T_j acts as the multiplication operator $\mathbf{M}_{x_j}$ on $\mathcal{H}$, cf. Thm. 6.9.4. Therefore, the relation $U(t) = e^{\mathbf{i}tT_\bullet}$ implies that the unitary representation U is realized by multiplication operators; specifically

$$U(t) = \mathbf{M}_{e_t}$$

recall that $e_t(x) = e^{\mathbf{i}tx} = e^{\mathbf{i}t_1x_1} \cdots e^{\mathbf{i}t_Nx_N}$. □

Example 7.3.7. For each $t \in \mathbb{R}^N$, define the **translation operator**

$$U(t) : L^2(\mathbb{R}^N, m) \to L^2(\mathbb{R}^N, m) \qquad f \mapsto U(t)f$$
$$\text{where } (U(t)f)(x) = f(x - t)$$

Then $U(t)$ is clearly a unitary operator. Moreover, one checks easily that $U : t \in \mathbb{R} \to U(t) \in \mathfrak{L}(\mathcal{H})$ is a unitary representation of $\mathbb{R}^N$.

For each $f \in C_c(\mathbb{R}^N)$, one easily shows that the function $t \in \mathbb{R} \to U(t)f \in \mathcal{H}$ is continuous by using the uniform continuity of f. Since $C_c(\mathbb{R}^N)$ is dense in $L^2(\mathbb{R}^N, m)$ (cf. Thm. 1.7.10), it follows from Prop. 2.4.5 that U is strongly continuous. U is called the **(left) regular representation** of $\mathbb{R}^N$. [1] □

[1] The right regular representation denotes the map $t \mapsto U(-t)$.

7.3.2 From self-adjoint operators to unitary representations

The following lemma establishes the direction (1)⇒(2) in Thm. 7.3.3.

Lemma 7.3.8. *Let $T_1, \dots, T_N$ be strongly commuting self-adjoint operators on $\mathcal{H}$. Then $U : t \in \mathbb{R}^N \to e^{\mathbf{i}tT_\bullet} \in \mathfrak{L}(\mathcal{H})$ is a strongly continuous unitary representation of $\mathbb{R}^N$ on $\mathcal{H}$.*

Proof. As explained in Rem. 7.3.5, each $U(t) = e^{\mathbf{i}tT_\bullet}$ is unitary. Moreover, since the Borel functional calculus $\pi_{T_\bullet}$ restricts to a unitary representation of $\mathscr{Bor}_{\mathrm{b}}(\mathbb{R}^N)$, we clearly have $U(t+s) = U(t)U(s)$ and $U(0) = \mathbf{1}$. It remains to prove that U is strongly continuous (at 0). Choose any $\xi \in \mathcal{H}$. Let μ_ξ be the associated finite Borel measure on $\mathbb{R}^N$ associated to ξ, cf. Thm. 6.10.1. Then

$$\lim_{t\to 0} \|e^{\mathbf{i}tT_\bullet}\xi - \xi\|^2 \xlongequal{(6.57)} \lim_{t\to 0} \int_{\mathbb{R}^N} |e^{\mathbf{i}tx} - 1|^2 d\mu_\xi(x) = 0$$

by DCT. This proves that U is strongly continuous at 0. □

Next, we show that if (2) comes from (1), then (1) can be recovered from (2) by the process described at the end of Thm. 7.3.3. In particular, the direction (1)⇒(2) is injective.

Lemma 7.3.9. *Let $T_1, \dots, T_N$ be strongly commuting self-adjoint operators on $\mathcal{H}$. Then for each $1 \leqslant j \leqslant N$, $\mathscr{D}(T_j)$ is the set of all $\xi \in \mathcal{H}$ such that $\partial_{t_j} e^{\mathbf{i}t_jT_j}\xi\big|_{t=0}$ exists, and for any such ξ we have $\mathbf{i}T_j\xi = \partial_{t_j} e^{\mathbf{i}t_jT_j}\xi\big|_{t=0}$.*

Proof. This result has nothing to do with T_k if $k \neq j$. Therefore, we may assume that $N = 1$ and omit the subscript j in t_j and T_j. Let μ_ξ be the finite Borel measure on $\mathbb{R}^N$ associated to ξ and the Borel functional calculus of T, cf. Thm. 6.10.1. Then for each $\xi \in \mathscr{D}(T)$, we have $\int_{\mathbb{R}} x^2 d\mu_\xi(x) < +\infty$ by part (3) of Thm. 6.10.1, and

$$\Big\|\frac{e^{\mathbf{i}tT}\xi - \xi}{t} - \mathbf{i}T\xi\Big\|^2 = \int_{\mathbb{R}} \Big|\frac{e^{\mathbf{i}tx}-1}{t} - \mathbf{i}x\Big|^2 d\mu_\xi(x) = \int_{\mathbb{R}\backslash\{0\}} \Big|\frac{e^{\mathbf{i}tx}-1}{tx} - \mathbf{i}\Big|^2 \cdot x^2 d\mu_\xi(x)$$

by (6.57), where the integral domain can be restricted to $\mathbb{R}\backslash\{0\}$ because $t^{-1}(e^{\mathbf{i}tx} - 1) - \mathbf{i}x$ equals 0 when $x = 0$. To show that the RHS converges to 0 as $t \to 0$, by DCT, it suffices to show that the function

$$\varphi : \mathbb{R}\backslash\{0\} \to \mathbb{C} \qquad s \mapsto \frac{e^{\mathbf{i}s} - 1}{s} - \mathbf{i}$$

is bounded. This follows from the fact that φ can be extended to a continuous function $\varphi : \mathbb{R} \to \mathbb{C}$ by setting $\varphi(0) = 0$, and that $\lim_{|s|\to 0} \varphi(s) = -\mathbf{i}$.

It remains to prove that any $\xi \in \mathcal{H}$ whose limit $\partial_t e^{\mathbf{i}tT}\xi|_{t=0}$ exists belongs to $\mathscr{D}(T)$. Let $\psi = \partial_t e^{\mathbf{i}tT}\xi|_{t=0}$. Choose any $\eta \in \mathscr{D}(T)$. By what we have proved, we have $\partial_t e^{\mathbf{i}tT}\eta|_{t=0} = \mathbf{i}T\eta$. Therefore

$$\langle \mathbf{i}T\eta|\xi\rangle = \lim_{t\to 0} t^{-1}\langle (e^{\mathbf{i}tT} - \mathbf{1})\eta|\xi\rangle = \lim_{t\to 0} t^{-1}\langle \eta|(e^{-\mathbf{i}tT} - \mathbf{1})\xi\rangle = -\langle\eta|\psi\rangle$$

This proves that $\xi \in \mathscr{D}(T^*)$ and $(\mathbf{i}T)^*\xi = -\psi$. Since $T = T^*$, we conclude that $\xi \in \mathscr{D}(T)$. □

7.3.3 From unitary representations to self-adjoint operators: the case $N = 1$

In this subsection, we prove that the direction (1)⇒(2) in Thm. 7.3.3 is surjective in the special case where $N = 1$.

Lemma 7.3.10. *Let U be a strongly continuous unitary representation of $\mathbb{R}$ on $\mathcal{H}$. Then there exists a unique self-adjoint operator T on $\mathcal{H}$ such that $U(t) = e^{\mathbf{i}tT}$ for each $t \in \mathbb{R}$.*

Proof. Step 1. The uniqueness follows from Lem. 7.3.9. Let us prove the existence. Let $\mathscr{D}(T)$ be the subspace of all $\xi \in \mathcal{H}$ such that the limit $\partial_t U(t)\xi|_{t=0}$ exists. For any $\xi \in \mathscr{D}(T)$, let

$$T\xi = -\mathbf{i}\partial_t U(t)\xi\big|_{t=0}$$

Then T is an n.d.d. unbounded operator on $\mathcal{H}$ with domain $\mathscr{D}(T)$, called the **generator of U**. Moreover, for each $\xi, \eta \in \mathscr{D}(T)$, since $U(t)^* = U(-t)$, we have

$$\langle \eta | T\xi \rangle = -\mathbf{i}\partial_t \langle \eta | U(t)\xi \rangle\big|_{t=0} = -\mathbf{i}\partial_t \langle U(-t)\eta | \xi \rangle\big|_{t=0} = \langle T\eta | \xi \rangle$$

This proves that T is an n.d.d. Hermitian operator on $\mathcal{H}$. [2]

In the following steps, we show that $\mathscr{D}(T)$ is dense, that T is essentially self-adjoint (and hence $\overline{T}$ is self-adjoint), and $U(t) = e^{\mathbf{i}t\overline{T}}$ for each $t \in \mathbb{R}$.

Step 2. In this step, we show that $\mathscr{D}(T)$ is dense.[3] Let $\xi \in \mathcal{H}$. For each $h \in C_c^\infty(\mathbb{R})$, let

$$\xi(h) = \int_{\mathbb{R}} h(x)U(x)\xi dx$$

where the integral is understood as in Sec. 7.A. Since for each $t \in \mathbb{R}$ we have

$$U(t)\xi(h) = \int_{\mathbb{R}} h(x)U(x+t)\xi dx = \int_{\mathbb{R}} h(y-t)U(y)\xi dy \tag{7.12}$$

it follows from Cor. 7.A.12 that $\xi(h) \in \mathscr{D}(T)$ and

$$T\xi(h) = -\mathbf{i}\partial_t U(t)\xi(h)\big|_{t=0} = \mathbf{i}\xi(h')$$

Now, we assume moreover that $h \geqslant 0$ and $\int h = 1$. For each $\varepsilon > 0$, let $h_\varepsilon(x) = \varepsilon^{-1}h(\varepsilon^{-1}x)$. From the fact that $x \mapsto U(x)\xi$ is continuous and $U(0)\xi = \xi$, that $\int_{\mathbb{R}} h_\varepsilon = 1$, and that

$$\lim_{\varepsilon\to 0} \int_{\mathbb{R}\backslash[-\delta,\delta]} h_\varepsilon = \lim_{\varepsilon\to 0} \int_{\mathbb{R}\backslash[-\varepsilon^{-1}\delta,\varepsilon^{-1}\delta]} h = 0$$

[2] Alternatively, one may use $\frac{d}{dt}\langle U(t)\eta | U(t)\xi \rangle\big|_{t=0} = 0$ to show $\langle \eta | T\xi \rangle - \langle T\eta | \xi \rangle = 0$.

[3] The idea is as follows. Consider the special case where U is the regular representation of $\mathbb{R}$ on $L^2(\mathbb{R}, m)$. Then any $\xi \in L^2(\mathbb{R}, m)$ can be approximated by the convolution $h * \xi$ where $h \in C_c^\infty(\mathbb{R})$. Note that $h * \xi = \int_{\mathbb{R}} h(x)U(x)\xi dx$. Thus, this approximation can be generalized to the abstract setting.

for each $\delta > 0$, one easily deduces that $\lim_{\varepsilon\to 0} U(h_\varepsilon)\xi = \xi$. We have thus proved that any $\xi \in \mathcal{H}$ can be approximated by a net $(U(h_\varepsilon)\xi)_{\varepsilon\in\mathbb{R}_{>0}}$ in $\mathscr{D}(T)$. Therefore $\mathscr{D}(T)$ is dense in $\mathcal{H}$. (In fact, by (7.12), $\mathscr{D}(T)$ contains a dense U-invariant subspace $\mathrm{Span}\{\xi(h) : \xi \in \mathcal{H}, h \in C_c^\infty(\mathbb{R})\}$ of $\mathcal{H}$.)

Step 3. To show that T is essentially self-adjoint, by Cor. 6.7.9, it suffices to show that $\mathrm{Rng}(T + \mathbf{i})$ and $\mathrm{Rng}(T - \mathbf{i})$ are dense in $\mathcal{H}$. We shall show the equivalent fact that $\mathrm{Rng}(T \pm \mathbf{i})^\perp = 0$, that is, $\mathrm{Ker}(T^* \mp \mathbf{i}) = 0$ (cf. Prop. 6.6.11).[4]

Choose any $\xi \in \mathrm{Ker}(T^* - \mathbf{i})$. By Step 2, $\mathcal{H}$ has a dense U-invariant subspace $\mathscr{D}_0$ contained in $\mathscr{D}(T)$. Then for each $\eta \in \mathscr{D}_0$ we have $\mathbf{i}TU(t)\eta = \partial_t U(t)\eta$ (since $U(s+t) = U(s)U(t)$). Thus

$$\partial_t\langle\xi|U(t)\eta\rangle = \mathbf{i}\langle\xi|TU(t)\eta\rangle = \mathbf{i}\langle T^*\xi|U(t)\eta\rangle = \langle\xi|U(t)\eta\rangle$$

Therefore, $f : t \in \mathbb{R} \mapsto \langle\xi|U(t)\eta\rangle$ is a differentiable function satisfying the differential equation $f' = f$, and hence is of the form $f(t) = ce^t$. Since f is bounded (because $\lVert U(t)\eta\rVert = \lVert\eta\rVert$), we must have $c = 0$. Therefore $\langle\xi|\eta\rangle = 0$ for all $\eta \in \mathscr{D}_0$, and hence $\xi = 0$. This proves $\mathrm{Ker}(T^* - \mathbf{i}) = 0$. A similar argument yields $\mathrm{Ker}(T^* - \mathbf{i}) = 0$.

Step 4. It remains to prove $U(t) = e^{\mathbf{i}t\overline{T}}$ for each $t \in \mathbb{R}$. Recall that T is the generator of U as defined in Step 1. Therefore, if $\Phi : \mathcal{H} \to \mathcal{K}$ is a unitary map, then $t \mapsto \Phi U(t)\Phi^{-1}$ is a strongly continuous unitary representation of $\mathbb{R}$ on $\mathcal{K}$, and its generator is $\Phi T\Phi^{-1}$ with domain $\mathscr{D}(\Phi T\Phi^{-1}) = \Phi\mathscr{D}(T)$.

Now, we take $\mathcal{K} = \mathcal{H}$ and $\Phi = U(s)$ for a fixed $s \in \mathbb{R}$. The above paragraph thus implies that the generator of $t \in \mathbb{R} \mapsto U(s)U(t)U(s)^{-1}$ has domain $U(s)\mathscr{D}(T)$. But this generator is also T, since $U(s)U(t)U(s)^{-1} = U(s+t-s) = U(t)$. Therefore

$$U(s)\mathscr{D}(T) = \mathscr{D}(T) \qquad \text{for each } s \in \mathbb{R}$$

Similarly, since (by Lem. 7.3.8 and 7.3.9) $t \in \mathbb{R} \mapsto e^{\mathbf{i}t\overline{T}}$ is a strongly continuous unitary representation of $\mathbb{R}$ with generator $\overline{T}$, we have

$$e^{\mathbf{i}s\overline{T}}\mathscr{D}(\overline{T}) = \mathscr{D}(\overline{T}) \qquad \text{for each } s \in \mathbb{R}$$

For each $\xi \in \mathscr{D}(\overline{T})$ and $\eta \in \mathscr{D}(T)$, by the following Exe. 7.3.11 and the fact that $\overline{T}^* = \overline{T}$, we have

$$\partial_t\langle\xi|e^{-\mathbf{i}t\overline{T}}U(t)\eta\rangle\big|_{t=0} = \partial_t\langle e^{\mathbf{i}t\overline{T}}\xi|U(t)\eta\rangle\big|_{t=0} = -\mathbf{i}\langle\overline{T}\xi|\eta\rangle + \mathbf{i}\langle\xi|T\eta\rangle$$

[4]The idea for showing $\mathrm{Ker}(T^* - \mathbf{i}) = 0$ is as follows. Consider the case that U is the regular representation of $\mathbb{R}$ on $L^2(\mathbb{R}, m)$. Then T extends the Hermitian operator $\mathbf{i}\frac{d}{dx}$ with domain $\mathscr{D}(\mathbf{i}\frac{d}{dx}) = C_c^\infty(\mathbb{R})$. By Prop. 6.8.3, any $\xi \in \mathcal{H}$ killed by $T^* - \mathbf{i}$ must be locally proportional to the function e^x, and hence globally proportional to e^x. Since $e^x \notin L^2$, we must have $\xi = 0$. This motivates the following proof for the general case.

$$= -\mathbf{i}\langle\xi|\overline{T}\eta\rangle + \mathbf{i}\langle\xi|T\eta\rangle = 0$$

By what we have proved in this step, for each $s \in \mathbb{R}$ we have $e^{\mathbf{i}s\overline{T}}\xi \in \mathscr{D}(\overline{T})$ and $U(s)\eta \in \mathscr{D}(T)$. Therefore, the above equality still holds if ξ, η are replaced by $e^{\mathbf{i}s\overline{T}}\xi$ and $U(s)\eta$. That is,

$$0 = \partial_t \langle e^{\mathbf{i}s\overline{T}}\xi|e^{-\mathbf{i}t\overline{T}}U(t)U(s)\eta\rangle\big|_{t=0} = \partial_t \langle \xi|e^{-\mathbf{i}t\overline{T}}U(t)\eta\rangle\big|_{t=s}$$

We have thus proved that the function $f : t \in \mathbb{R} \mapsto \langle\xi|e^{-\mathbf{i}t\overline{T}}U(t)\eta\rangle$ has derivative 0 everywhere, and hence is equal to $f(0) = \langle\eta|\xi\rangle$. Therefore $e^{-\mathbf{i}t\overline{T}}U(t)$ equals $\mathbf{1}$ when evaluated between $\xi \in \mathscr{D}(\overline{T})$ and $\eta \in \mathscr{D}(T)$. Since $\mathscr{D}(T)$ is dense (by Step 2), $e^{-\mathbf{i}t\overline{T}}U(t)$ must be equal to $\mathbf{1}$ on the whole domain $\mathcal{H}$. □

Exercise 7.3.11. Let $I \subset \mathbb{R}$ be a proper interval. Let $\xi, \eta : I \to \mathcal{H}$ be functions differentiable at $t_0 \in I$. Prove that

$$\partial_t \langle\xi(t)|\eta(t)\rangle\big|_{t=t_0} = \langle\partial_t\xi(t_0)|\eta(t_0)\rangle + \langle\xi(t_0)|\partial_t\eta(t_0)\rangle$$

7.3.4 From unitary representations to self-adjoint operators: the general case

We have already proved Stone's Thm. 7.3.3 for the case $N = 1$. In this subsection, we prove that the direction (1)⇒(2) in Thm. 7.3.3 is surjective for arbitrary N. To prove that any strongly continuous unitary representation of $\mathbb{R}^N$ arises from a tuple of strongly commuting self-adjoint operators $(T_1, \dots, T_N)$, we need some criteria for the strong commutativity.

The following property is analogous to Prop. 6.10.7.

Proposition 7.3.12. *Let T be a self-adjoint operator on $\mathcal{H}$. Let $V \in \mathfrak{L}(\mathcal{H})$ be unitary. Then $VT = TV$ iff V commutes with $e^{\mathbf{i}tT}$ for each $t \in \mathbb{R}$.*

Proof. In general, if $V : \mathcal{H} \to \mathcal{K}$ is a unitary map and $S = VTV^{-1}$ (which is self-adjoint), then $e^{\mathbf{i}tS} = Ve^{\mathbf{i}tT}V^{-1}$ holds for each $t \in \mathbb{R}$. (Indeed, we have $f(S) = Vf(T)V^{-1}$ for each $f \in \mathscr{B}or(\mathbb{R})$ by the uniqueness of the Borel functional calculus.)

We now consider the special case where $\mathcal{K} = \mathcal{H}$. Then $VT = TV$ iff $S = T$. If $S = T$, then $e^{\mathbf{i}tT} = Ve^{\mathbf{i}tT}V^{-1}$ by the first paragraph, and hence V commutes with every $e^{\mathbf{i}tT}$. Conversely, assume that V commutes with every $e^{\mathbf{i}tT}$. Then $e^{\mathbf{i}tT} = Ve^{\mathbf{i}tT}V^{-1}$, and hence $e^{\mathbf{i}tS} = e^{\mathbf{i}tT}$ for all $t \in \mathbb{R}$. Since we have proved Thm. 7.3.3 for the case $N = 1$, the uniqueness part of that theorem shows that $S = T$. □

The following theorem complements Thm. 6.10.8.

Theorem 7.3.13. *Let T_1, T_2 be self-adjoint operators on $\mathcal{H}$. Then the following are equivalent.*

(1) T_1 and T_2 commute strongly.

(2) $e^{\mathbf{i}tT_1}T_2 = T_2e^{\mathbf{i}tT_1}$ *for each* $t \in \mathbb{R}$.

(3) $e^{\mathbf{i}sT_2}T_1 = T_1e^{\mathbf{i}sT_2}$ *for each* $s \in \mathbb{R}$.

(4) $e^{\mathbf{i}tT_1}$ *commutes with* $e^{\mathbf{i}sT_2}$ *for each* $t, s \in \mathbb{R}$.

Proof. (2)⇔(3)⇔(4): This follows immediately from Prop. 7.3.12.

(1)⇒(4): This follows from the direction (1)⇒(5) in Thm. 6.10.8.

(2)⇒(1): Assume (2). Let U_2 be the Cayley transform of T_2. By Prop. 6.10.7, $e^{\mathbf{i}tT_1}$ commutes with U_2 for each $t \in \mathbb{R}$. By Prop. 7.3.12, we conclude that $U_2T_1 = T_1U_2$. Thm. 6.10.8 then implies the strong commutativity of T_1, T_2. □

Proof of Thm. 7.3.3. By Lem. 7.3.8 and 7.3.9, it remains to prove that the direction (1)⇒(2) is surjective. Let U be a strongly continuous unitary representation of $\mathbb{R}^N$ on $\mathcal{H}$. Let $\varepsilon_1, \dots, \varepsilon_N$ be the standard basis of $\mathbb{R}^N$. Then

$$U_j : \mathbb{R} \to \mathfrak{L}(\mathcal{H}) \qquad t \mapsto U(t\varepsilon_j)$$

is a strongly continuous unitary representation of $\mathbb{R}$ on $\mathbb{M}$, and hence can be written as $U_j(t) = e^{\mathbf{i}tT_j}$ for some self-adjoint operator T_j due to Thm. 7.3.10. Since

$$e^{\mathbf{i}t_jT_j}e^{\mathbf{i}t_kT_k} = U_j(t_j)U_k(t_k) = U(t_j\varepsilon_j)U(t_k\varepsilon_k) = U(t_j\varepsilon_j + t_k\varepsilon_k)$$

and similarly $e^{\mathbf{i}t_kT_k}e^{\mathbf{i}t_jT_j} = U(t_j\varepsilon_j + t_k\varepsilon_k)$, it follows from Thm. 7.3.13 that T_j and T_k commute strongly. For each $t = (t_1, \dots, t_N) \in \mathbb{R}^N$, we have

$$U(t) = U(t_1\varepsilon_1 + \cdots + t_N\varepsilon_N) = U_1(t_1)\cdots U_N(t_N) = e^{\mathbf{i}t_1T_1}\cdots e^{\mathbf{i}t_NT_N}$$

This proves that $U(t) = e^{\mathbf{i}tT_\bullet}$. □

The proof of Stone's Thm. 7.3.3 is now complete.

7.4 Harmonic analysis on the group $\mathbb{R}^N$

Let $N \in \mathbb{Z}_+$. We continue to use the notation at the beginning of Sec. 7.3. Let $\mathfrak{m}_N$ be strictly-positively proportional to the Lebesgue measure m_N, that is,

$$d\mathfrak{m}_N = ?dm_N \qquad \text{where } ? \in \mathbb{R}_{>0}$$

We call $\mathfrak{m}_N$ a **Haar measure** of $\mathbb{R}^N$ and abbreviate $\mathfrak{m}_N$ to $\mathfrak{m}$ when no confusion arises.

Convention 7.4.1. In this section, integrals of the form $\int_{\mathbb{R}^N} f d\mathfrak{m} = \int_{\mathbb{R}^N} f(x)d\mathfrak{m}(x)$ will also be written as $\int_{\mathbb{R}^N} f(x)dx$.

7.4.1 Introduction

Representation theory and harmonic analysis are closely connected fields. In this subsection, we briefly illustrate how harmonic analysis on $\mathbb{R}^N$ aids the study of strongly continuous unitary representations of $\mathbb{R}^N$. Although our main focus in this course is on the group $\mathbb{R}^N$, we present this discussion in the more general setting of locally compact abelian (LCA) groups. For detailed treatments of locally compact groups and abstract harmonic analysis, see [DE-P], [Fol-A], and [Sim-O, Ch. 6].

A locally compact abelian toplogical group (abbreviated **LCA group**) is an abelian group G endowed with an LCH topology compatible with its group operation. We omit the precise definition here and list several standard examples:

- $\mathbb{R}$ with the Euclidean topology and ordinary addition;
- $\mathbb{T} = \mathbb{S}^1$ with the Euclidean topology and multiplication as the group operation;
- $\mathbb{Z}$ with the discrete topology and ordinary addition;
- $\mathbb{Z}/p\mathbb{Z}$ with $p = 2, 3, \ldots$, with the discrete topology and the usual modular addition;
- Finite products of the above examples.

For an LCA group G, we denote the group operation by addition, so that the product of $x, y \in G$ is written as $x + y$, the inverse of x is written as $-x$, and the identity element is denoted by 0.

There exists a unique (up to a positive scalar multiple) Borel measure μ on G satisfying $\mu(E + x) = \mu(E)$ for all Borel $E \subset G$ and $x \in G$, called the **Haar measure** of G. The Haar measure is unimodular, meaning $\mu(-E) = \mu(E)$ for each Borel $E \subset G$. To be consistent with the convention for Lebesgue measure, we take the completion of μ. Moreover, if G is compact, we normalize μ so that $\mu(G) = 1$.

A strongly continuous unitary representation $U : G \to \mathfrak{L}(\mathcal{H})$ induces a unitary representation of the **convolution algebra** $L^1(G, \mu)$, whose multiplication is defined by the **convolution**

$$(f * g)(x) = \int_G f(x - y)g(y)d\mu(y) \equiv \int_G f(y)g(x - y)\mu(y)$$

and whose *-operation is given by

$$f^\dagger(x) = \overline{f(-x)}$$

Then U gives rise to a unitary representation

$$U : L^1(G, \mu) \to \mathfrak{L}(\mathcal{H}) \qquad U(f) = \int_G f(x)U(x)d\mu(x)$$

To understand intuitively why $U(f * g) = U(f)U(g)$, it is helpful to consider the simple case where G is a finite abelian group. In that case,

$$U(f) = \sum_{x \in G} f(x)U(x)$$

and hence

$$U(f)U(g) = \sum_{x \in G} f(x)U(x) \sum_{y \in G} g(y)U(y) = \sum_{x,y \in G} f(x)g(y)U(x+y)$$
$$= \sum_{x,z \in G} f(x)g(z-x)U(z) = U(f * g)$$

This computation, in fact, motivates the definition of convolution in the general setting.

A central idea in harmonic analysis is that the Fourier transform converts the unitary representation U of the convolution algebra $L^1(G, \mu)$ into a unitary representation of a dense *-subalgebra of $C_0(\widehat{G})$, where $\widehat{G}$ is the **Pontryagin dual** of G. As a set, $\widehat{G}$ consists of all continuous group homomorphisms $\chi : G \to \mathbb{T}$, called the **characters** of G. The group operation on $\widehat{G}$ is defined by pointwise multiplication:

$$(\chi_1\chi_2)(x) = \chi_1(x)\chi_2(x) \qquad \text{for each } x \in G \text{ and } \chi_1, \chi_2 \in \widehat{G}$$

Endowed with the compact-open topology inherited from $C(G, \mathbb{T})$ (that is, the topology in which a net converges iff it converges uniformly on compact subsets of G), the dual group $\widehat{G}$ becomes an LCA group. Typical examples include:

- The dual of $\mathbb{R}$ is $\mathbb{R}$, where each $t \in \mathbb{R}$ is viewed as the character $x \in \mathbb{R} \mapsto e^{\mathbf{i}tx}$.
- The dual of $\mathbb{T}$ is $\mathbb{Z}$, where each $n \in \mathbb{Z}$ is viewed as the character $z \in \mathbb{T} \mapsto z^n$.
- The dual of $\mathbb{Z}$ is $\mathbb{T}$, where each $z \in \mathbb{T}$ is viewed as the character $n \in \mathbb{Z} \mapsto z^n$.
- The dual of $\mathbb{Z}/p\mathbb{Z}$ is $\mathbb{Z}/p\mathbb{Z}$, where each $n \in \mathbb{Z}/p\mathbb{Z}$ is viewed as the character $m \in \mathbb{Z}/p\mathbb{Z} \mapsto e^{2\pi\mathbf{i}\cdot mn/p}$.
- $(G_1 \times \cdots \times G_N)^\wedge \simeq \widehat{G}_1 \times \cdots \times \widehat{G}_N$.

For $f \in L^1(G, \mu)$, the **Fourier transform** $\widehat{f} : \widehat{G} \to \mathbb{C}$ is defined by

$$\widehat{f}(\chi) = \int_G f(x)\overline{\chi(x)}d\mu(x) \qquad \text{for each } \chi \in \widehat{G}$$

A fundamental result, proved in this section for $G = \mathbb{R}^N$ (cf. Thm. 7.4.7), states that the Fourier transform defines a *-algebra homomorphism

$$\mathcal{F} : L^1(G) \to C_0(\widehat{G}) \qquad f \mapsto \widehat{f}$$

which is injective and whose range $\mathscr{A} := \mathcal{F}(L^1(G))$ is dense in $C_0(\widehat{G})$ under the l^∞-norm. In particular, $\mathcal{F}$ intertwines the algebra mutiplications:

$$\mathcal{F}(f * g) = \mathcal{F}(f) \cdot \mathcal{F}(g) \qquad \text{for each } f, g \in L^1(G)$$

Consequently, one obtains a unitary representation

$$\pi : \mathscr{A} \to \mathfrak{L}(\mathcal{H}) \qquad \pi(\widehat{f}) = U(f) \tag{7.13}$$

Since $\mathscr{A}$ is dense in $C_0(\widehat{G})$, it is natural to expect—by analogy with the operator-valued abstract Hausdorff moment theorem (Thm. 5.5.13)—that the representation (7.13) can be extended to a Radon (or, if G is second countable, a normal) unitary representation $\pi : \mathscr{Bor}_{\mathrm{b}}(\widehat{G}) \to \mathfrak{L}(\mathcal{H})$. Such an extension may be regarded as the spectral decomposition of the group representation $U : G \to \mathfrak{L}(\mathcal{H})$.[5] For example, combining Prop. 5.10.16 with Prop. 5.10.19 yields a unitary equivalence

$$\mathcal{H} \simeq \bigoplus_{\alpha \in \mathscr{I}} L^2(\widehat{G}, \mu_\alpha) \tag{7.14a}$$

where each μ_α is a finite Radon measure on $\widehat{G}$. Under this equivalence, we have

$$\pi(g) = \mathbf{M}_g \qquad \text{for each } g \in \mathscr{Bor}_{\mathrm{b}}(\widehat{G})$$

From this, it is not hard to show that the representation $U : G \to \mathfrak{L}(\mathcal{H})$ can be represented by the multiplication operators

$$\begin{gathered} U(x) = \mathbf{M}_{x^\sharp} \\ \text{where } x^\sharp : \widehat{G} \to \mathbb{C} \qquad \chi \mapsto \chi(x) \end{gathered} \tag{7.14b}$$

Such an extension $\pi : \mathscr{Bor}_{\mathrm{b}}(\widehat{G}) \to \mathfrak{L}(\mathcal{H})$ indeed exists for every LCA group G, though the proof is nontrivial. Fortunately, as we will see in Thm. 7.5.2 (Sec. 7.5), in the case $G = \mathbb{R}^N$, one can use Stone's theorem (Thm. 7.3.3) to extend the representation (7.13) to a normal unitary representation

$$\pi : \mathscr{Bor}_{\mathrm{b}}(\mathbb{R}^N) \to \mathfrak{L}(\mathcal{H}) \qquad g \mapsto g(T_\bullet)$$

defined by the Borel functional calculus of the generators $T_1, \dots, T_N$ of the representation $U : \mathbb{R}^N \to \mathfrak{L}(\mathcal{H})$. This is unsurprising because, as noted in Rem. 7.3.6, Stone's theorem provides a spectral theorem for strongly continuous unitary representations of $\mathbb{R}^N$ similar to the form (7.14).

[5] In particular, the support of $\pi : \mathscr{Bor}_{\mathrm{b}}(\widehat{G}) \to \mathfrak{L}(\mathcal{H})$ can be viewed as the spectrum of the representation U of G on $\mathcal{H}$, called the **Arveson spectrum** of U.

7.4.2 Reading guide

In the remainder of this section, we provide rigorous proofs of all statements mentioned in Subsec. 7.4.1 for the case $G = \mathbb{R}^N$, except that the extension of the representation (7.13) to $\mathscr{Bor}_{\mathrm{b}}(\mathbb{R}^N)$ will be postponed to Thm. 7.5.2 in Sec. 7.5.

Many of the results presented here are not logically required later in the course, though they are conceptually illuminating and part of the standard background knowledge. In fact, besides the easy inequality $\|\widehat{f}\|_{L^\infty} \leqslant \|f\|_{L^1}$, the only theorem that will be used subsequently is that the Fourier transforms of functions in $C_c(\mathbb{R}^N)$ form a dense subspace of $C_0(\mathbb{R}^N)$, cf. Thm. 7.4.8. Hence, any result not needed later may be safely skipped on a first reading.

Accordingly, in Sec. 7.5 we will restrict attention to integrals $U(f) = \int f(x)U(x)dx$ only for $f \in C_c(\mathbb{R}^N)$. The theory for such integrals is simpler than for general $f \in L^1(\mathbb{R}^N)$, since one may, for instance, define $U(f)\xi = \int f(x)U(x)\xi dx$ for $\xi \in \mathcal{H}$ using Riemann integrals over bounded boxes in $\mathbb{R}^N$. See Subsec. 7.A.1 for a detailed discussion.

In the following subsections, we present proofs of the results using arguments that also apply to general LCA groups. For example, our reasoning does not depend on explicit functions on $\mathbb{R}^N$, except for the exponential $e_t(x) = e^{\mathbf{i}tx}$ (which generalizes to group characters) and compactly supported continuous functions obtained from Urysohn's lemma. In particular, we make no use of the smooth structure of $\mathbb{R}^N$. Differential aspects of the Fourier transform will be treated separately in Sec. 7.8.

7.4.3 The convolution algebra $L^1(\mathbb{R}^N, \mathfrak{m})$

Proposition 7.4.2. *Let $f, g \in L^1(\mathbb{R}^N, \mathfrak{m})$. The **convolution***

$$(f * g)(x) = \int_{\mathbb{R}^N} f(x-y)g(y)dy \equiv \int_{\mathbb{R}^N} f(y)g(x-y)dy$$

exists for almost every x. Moreover, we have $f * g \in L^1(\mathbb{R}^N, \mathfrak{m})$ and

$$\|f * g\|_{L^1(\mathbb{R}^N,\mathfrak{m})} \leqslant \|f\|_{L^1(\mathbb{R}^N,\mathfrak{m})}\|g\|_{L^1(\mathbb{R}^N,\mathfrak{m})} \tag{7.15}$$

Therefore, the convolution map $L^1(\mathbb{R}^N, \mathfrak{m}) \times L^1(\mathbb{R}^N, \mathfrak{m}) \to L^1(\mathbb{R}^N, \mathfrak{m})$ is bounded with operator norm $\leqslant 1$.

Proof. If $f, g \in L^1(\mathbb{R}^N, \mathfrak{m})$, then since $\int\int |f(x-y)g(y)|dxdy = \|f\|_{L^1} \cdot \|g\|_{L^1} < +\infty$, the theorems of Tonelli and Fubini imply that $f * g$ can be defined a.e., and that $f * g$ has L^1-norm $\leqslant \|f\|_{L^1}\|g\|_{L^1}$. $\square$

Theorem 7.4.3. *The space $L^1(\mathbb{R}^N, \mathfrak{m})$, equipped with the algebra multiplication defined by the convolution $f * g$ and the *-structure defined by $f \mapsto f^\dagger$ with*

$$f^\dagger(x) = \overline{f(-x)}$$

*is a commutative *-algebra, called the* **convolution (L^1-)algebra** *of* $\mathbb{R}^N$. *Moreover,* $C_c(\mathbb{R}^N)$ *is a *-subalgebra of* $L^1(\mathbb{R}^N,\mathfrak{m})$.

Proof. For each $f,g,h\in C_c(\mathbb{R}^N,\mathfrak{m})$ we have

$$
\begin{aligned}
&((f*g)*h)(z)=\int(f*g)(x)h(z-x)dx\\
=&\int\int f(y)g(x-y)h(z-x)dydx=\int\int f(y)g(x-y)h(z-x)dxdy\\
=&\int f(y)(g*h)(z-y)dy=(f*(g*h))(z)
\end{aligned}
$$

This proves $((f*g)*h)=(f*(g*h))$ whenever $f,g,h\in C_c(\mathbb{R}^N,\mathfrak{m})$. In the general case that $f,g,h\in L^1(\mathbb{R}^N,\mathfrak{m})$, by Thm. 1.7.10, there exist sequences $(f_n),(g_n),(h_n)$ in $C_c(\mathbb{R}^N,\mathfrak{m})$ converging in L^1 to f,g,h respectively. Since the bilinear convolution map is bounded in the L^1-norm, we have

$$
((f*g)*h)=\lim_n((f_n*g_n)*h_n)=\lim_n(f_n*(g_n*h_n))=f*(g*h)
$$

This proves that the convolution is associative.

The commutativity $f*g=g*f$ is obvious. Moreover,

$$
\begin{aligned}
&(f*g)^\dagger(x)=\overline{f*g(-x)}=\overline{\int f(-x-y)g(y)dy}=\int\overline{f(-x-y)}\cdot\overline{g(y)}dy\\
=&\int f^\dagger(x+y)g^\dagger(-y)dy=(g^\dagger*f^\dagger)(x)
\end{aligned}
$$

This proves that $L^1(\mathbb{R}^N,\mathfrak{m})$ is a commutative *-algebra.

Clearly $C_c(\mathbb{R}^N)$ is closed under convolution and taking †. Thus it is a *-subalgebra of $L^1(\mathbb{R}^N,\mathfrak{m})$. □

The main reason for introducing the convolution algebra is that it turns a strongly continuous unitary representation of the group $\mathbb{R}^N$ into a unitary representation of $L^1(\mathbb{R}^N,\mathfrak{m})$, as shown by the following theorem. For simplicity, when dealing with vector-valued integrals, readers may replace $L^1(\mathbb{R}^N,\mathfrak{m})$ by the more elementary space $C_c(\mathbb{R}^N)$.

Theorem 7.4.4. *Let* $U:\mathbb{R}^N\to\mathcal{H}$ *be a strongly continuous unitary representation of* $\mathbb{R}^N$. *Then we have a unitary representation of* $L^1(\mathbb{R}^N,\mathfrak{m})$, *also denoted by* U, *defined by*

$$
U:L^1(\mathbb{R}^N,\mathfrak{m})\to\mathfrak{L}(\mathcal{H})\qquad f\mapsto\int_{\mathbb{R}^N}f(x)U(-x)dx \tag{7.16}
$$

Here, $U(f):=\int_{\mathbb{R}^N}f(x)U(-x)dx$ *is the bounded linear operator on* $\mathcal{H}$ *defined by*

$$
U(f):\mathcal{H}\to\mathcal{H}\qquad \xi\mapsto\int_{\mathbb{R}^N}f(x)U(-x)\xi dx
$$

where $\int_{\mathbb{R}^N} f(x)U(-x)\xi dx$ is defined as in Sec. 7.A. Moreover, the linear map (7.16) is bounded with operator norm $\leqslant 1$.

Note that our definition of $U(f)$ differs from that in Subsec. 7.4.1 by the presence of the negative sign in $U(-x)$; this convention is adopted to maintain consistency with the Fourier transform convention $\int f(x)e^{-\mathbf{i}tx}dx$.

Proof. Since $U : \mathbb{R}^N \to \mathcal{H}$ is strongly continuous, for each $\xi \in \mathcal{H}$ the function $x \mapsto f(x)U(-x)\xi$ is integrable, and hence the integral $\int f(x)U(-x)\xi dx$ exists by Thm. 7.A.10. Moreover, since $\|U(x)\xi\| = \|\xi\|$, Prop. 7.A.6 shows that $\|U(f)\xi\| \leqslant \|f\|_{L^1} \cdot \|\xi\|$. Therefore $U(f) \in \mathfrak{L}(\mathcal{H})$, and $\|U\| \leqslant 1$.

For each $f, g \in L^1(\mathbb{R}^N, \mathfrak{m})$ and $\xi, \eta \in \mathcal{H}$, we compute that

$$\langle U(f)\eta|\xi\rangle = \int \overline{f(x)}\langle U(-x)\eta|\xi\rangle dx = \int \overline{f(x)}\langle \eta|U(x)\xi\rangle dx = \langle \eta|U(f^\dagger)\eta\rangle$$

This proves $U(f)^* = U(f^\dagger)$. We also compute that

$$\langle \eta|U(f * g)\xi\rangle = \int (f * g)(x)\langle \eta|U(-x)\xi\rangle dx = \int\int f(y)g(x-y)\langle \eta|U(-x)\xi\rangle dydx$$

By Exp. 7.A.2, $U(\mathbb{R}^N)$ is separable. Therefore, we can apply Prop. 7.A.13:

$$\langle \eta|U(f)U(g)\xi\rangle = \langle U(f^\dagger)\eta|U(g)\xi\rangle = \int\int f(-s)g(t)\langle U(-s)\xi|U(-t)\eta\rangle dsdt$$
$$= \int\int f(-s)g(t)\langle \xi|U(s-t)\eta\rangle dsdt$$

The change of variables $y = -s, x = t - s$ implies $U(f * g) = U(f)U(g)$. □

7.4.4 The Fourier transform $L^1(\mathbb{R}^N, \mathfrak{m}) \to C_0(\mathbb{R}^N)$ as a *-homomorphism

Definition 7.4.5. Given $f \in L^1(\mathbb{R}^N, \mathfrak{m})$, the **Fourier transform $\widehat{f}$** and the **inverse Fourier transform $\check{f}$** are functions $\mathbb{R}^N \to \mathbb{C}$ defined by

$$\widehat{f}(t) = \int_{\mathbb{R}^N} f(x)e^{-\mathbf{i}tx}dx \qquad \check{f}(t) = \widehat{f}(-t) = \int_{\mathbb{R}^N} f(x)e^{\mathbf{i}tx}dx$$

The following proposition prepares us for the proof of Thm. 7.4.7, the main result of this subsection.

Proposition 7.4.6. *For each function $f \in L^1(\mathbb{R}^N, \mathfrak{m})$ and $y \in \mathbb{R}^N$, the **translation** $\boldsymbol{\tau_y f} \in L^1(\mathbb{R}^N, \mathfrak{m})$ is defined by $(\tau_y f)(x) = f(x-y)$. Then for each $y, s \in \mathbb{R}^N$ we have*

$$\widehat{\tau_y f} = e_{-y} \cdot \widehat{f} \qquad \widehat{e_s \cdot f} = \tau_s \widehat{f}$$

Proof. This follows easily from a change of coordinates in the computation of the integrals. □

Theorem 7.4.7. *The Fourier transform defines a map*

$$\mathcal{F} : L^1(\mathbb{R}^N, \mathfrak{m}) \to C_0(\mathbb{R}^N) \qquad f \mapsto \widehat{f} \tag{7.17}$$

The map $\mathcal{F}$ is a ***-homomorphism***, *that is, $\mathcal{F}$ is linear and satisfies*

$$\widehat{f * g} = \widehat{f} \cdot \widehat{g} \qquad \widehat{f^\dagger} = \overline{\widehat{f}} \qquad \text{for each } f, g \in L^1(\mathbb{R}^N, \mathfrak{m}) \tag{7.18}$$

Moreover, the linear map $\mathcal{F}$ is bounded with operator norm $\leqslant 1$, that is,

$$\|\widehat{f}\|_{l^\infty} \leqslant \|f\|_{L^1(\mathbb{R}^N,\mathfrak{m})} \qquad \text{for each } f \in L^1(\mathbb{R}^N, \mathfrak{m}) \tag{7.19}$$

In addition, the map $\mathcal{F}$ is injective with dense range.

The same conclusions hold for the inverse Fourier transform $f \mapsto \check{f}$.

Proof. Step 1. The map $\mathcal{F} : f \in L^1(\mathbb{R}^N) \mapsto \widehat{f}$ is clearly linear, and the relation (7.19) is obvious. By DCT, the function $\widehat{f}$ is continuous. Therefore, to show that $\mathcal{F}$ has range in $C_0(\mathbb{R}^N)$, it suffices to show that for each $f \in L^1(\mathbb{R}^N, \mathfrak{m})$, the function $\widehat{f}$ satisfies $\lim_{|t|\to+\infty} \widehat{f}(t) = 0$. (This property is known as the **Riemann-Lebesgue lemma**.)

We first consider the special case that $f \in C_c(\mathbb{R}^N)$. Since f is uniformly continuous, for any $\varepsilon > 0$ there exists δ such that for all $y \in \mathbb{R}^N$ with $|y| \leqslant \delta$ we have $\|f - \tau_y f\|_{L^1} \leqslant \varepsilon$, and hence $\|(1 - e_y)\widehat{f}\|_{l^\infty} \leqslant \varepsilon$ by Prop. 7.4.6 and (7.19). Thus, for all $0 < |y| \leqslant \delta$ and $t \in \mathbb{R}^N$ we have

$$|\widehat{f}(t)| = \frac{|(1 - e^{\mathbf{i}ty})\widehat{f}(t)|}{|1 - e^{\mathbf{i}ty}|} \leqslant \frac{\varepsilon}{|1 - e^{\mathbf{i}ty}|}$$

Therefore, for each $|t| \geqslant \pi/\delta$, by choosing $y = \pi t/|t|^2$ (so that $0 < |y| = \pi/|t| \leqslant \delta$ and $ty = \pi$), we obtain $|\widehat{f}(t)| \leqslant \varepsilon/2$. This proves $\widehat{f} \in C_0(\mathbb{R}^N)$.[6]

In the general case, by Thm. 1.7.10, there exists a sequence (f_n) in $C_c(\mathbb{R}^N)$ such that $\lim_n \|f - f_n\|_{L^1} = 0$. By (7.19), we have $\lim_n \|\widehat{f} - \widehat{f}_n\|_{l^\infty} = 0$. Since $\widehat{f}_n \in C_0(\mathbb{R}^N)$, and since $C_0(\mathbb{R}^N)$ is closed in $l^\infty(\mathbb{R}^N)$, the function $\widehat{f}$ belongs to $C_0(\mathbb{R}^N)$.

[6] An alternative proof, which does not generalize to LCA groups, is as follows. Let B be a bounded closed box containing $\mathrm{Supp}(f)$. By Stone-Weierstrass, there exists a sequence (f_n) in $\mathrm{Span}\{\chi_B e_t : t \in \mathbb{R}^N\}$ converging uniformly to f, and hence also in L^1. Therefore, by (7.19), $\widehat{f}_n$ converges uniformly to $\widehat{f}$. Since each $\widehat{f}_n$ vanishes at infinity (as can be verified by computing the Fourier transform of $\chi_B e_t$), so does $\widehat{f}$.

Step 2. In the previous step, we have proved that $\mathcal{F}$ is a linear map with range in $C_0(\mathbb{R}^N)$ satisfying the inequality (7.19). In this step, we prove (7.18).

For each $f, g \in L^1(\mathbb{R}^N, \mathfrak{m})$ and $t \in \mathbb{R}^N$, we have

$$\begin{aligned}\widehat{f * g}(t) &= \int (f * g)(x) e^{-\mathbf{i}tx} dx = \int\int f(x-y)g(y)e^{-\mathbf{i}tx} dy dx \\ &= \int\int f(x-y)g(y)e^{-\mathbf{i}tx} dx dy = \int\int f(u)g(y)e^{-\mathbf{i}tu}e^{-\mathbf{i}ty} du dy = \widehat{f}(t)\widehat{g}(t)\end{aligned}$$

where the theorems of Tonelli and Fubini are applicable since $\int\int |f(x-y)g(y)e^{\mathbf{i}tx}| dx dy = \|f\|_{L^1}\|g\|_{L^1} < +\infty$. The relation $\widehat{f^\dagger} = \overline{\widehat{f}}$ is easy to check.

Step 3. It remains to check that $\mathcal{F}$ has dense range, and that $\mathcal{F}$ is injective. This follows from Thm. 7.4.8 and Lem. 7.4.10 respectively. $\square$

Theorem 7.4.8. *The Fourier transform $\mathcal{F} : L^1(\mathbb{R}^N, \mathfrak{m}) \to C_0(\mathbb{R}^N)$ sends $C_c(\mathbb{R}^N)$ to an l^∞-dense $*$-subalgebra of $C_0(\mathbb{R}^N)$.*

Proof. Since $\mathscr{A} := C_c(\mathbb{R}^N)$ is a *-subalgebra of $L^1(\mathbb{R}^N, \mathfrak{m})$, and since we have proved in Thm. 7.4.7 that $\mathcal{F}$ is a *-homomorphism with range in $C_0(\mathbb{R}^N)$, it follows that $\widehat{\mathscr{A}} := \mathcal{F}(C_c(\mathbb{R}^N))$ is a $*$-subalgebra of $C_0(\mathbb{R}^N)$.

For each $t \in \mathbb{R}^N$, one can find $f \in C_c(\mathbb{R}^N, \mathbb{R}_{\geqslant 0})$ supported in a neighborhood of 0 with $\int f d\mathfrak{m} = 1$ such that $\widehat{f}(t) = \int f(x)e^{-\mathbf{i}tx}dx$ is close to $e^{-\mathbf{i}t\cdot 0} = 1$. Therefore, $\widehat{\mathscr{A}}$ vanishes nowhere. For each $s, t \in \mathbb{R}^N$ such that $s \neq t$, choose any $y \in \mathbb{R}^N$ such that $e^{-\mathbf{i}sy} \neq e^{-\mathbf{i}ty}$. Again, one can find $f \in C_c(\mathbb{R}^N, \mathbb{R}_{\geqslant 0})$ supported in a neighborhood of y with $\int f d\mathfrak{m} = 1$ such that $\widehat{f}(t)$ is close to $e^{-\mathbf{i}ty}$, and that $\widehat{f}(s)$ is close to $e^{-\mathbf{i}sy}$. In particular, one can find f such that $\widehat{f}(t) \neq \widehat{f}(s)$. This proves that $\widehat{\mathscr{A}}$ separates points of $\mathbb{R}^N$. Therefore, by the Stone-Weierstrass Thm. 1.5.13, $\widehat{\mathscr{A}}$ is dense in $C_0(\mathbb{R}^N)$. $\square$

To show that the Fourier transform is injective we need the following result.

Proposition 7.4.9. *Let $f, g \in L^1(\mathbb{R}^N, \mathfrak{m})$. Then*

$$\int_{\mathbb{R}^N} \widehat{f} g d\mathfrak{m} = \int_{\mathbb{R}^N} f \widehat{g} d\mathfrak{m}$$

Proof. We compute that

$$\int \widehat{f}(t)g(t)dt = \int\int f(x)e^{-\mathbf{i}tx}g(t)dxdt = \int\int f(x)e^{-\mathbf{i}tx}g(t)dtdx = \int f(x)\widehat{g}(x)dx$$

where the integrals can be exchanged due to $\int\int |f(x)e^{-\mathbf{i}tx}g(t)| dt dx = \|f\|_{L^1}\|g\|_{L^1} < +\infty$. $\square$

Lemma 7.4.10. *If $f \in L^1(\mathbb{R}^N, \mathfrak{m})$ and $\widehat{f} = 0$, then $f = 0$ (a.e.).*

Proof. For each $g \in C_c(\mathbb{R}^N)$, by Prop. 7.4.9 we have $\int f\widehat{g} = \int \widehat{f}g = 0$. By Thm. 7.4.8, all such $\widehat{g}$ form a dense linear subspace of $C_0(\mathbb{R}^N)$. It follows that $\int fh = 0$ for each $h \in C_0(\mathbb{R}^N)$, and hence for each $h \in C_c(\mathbb{R}^N)$. By Thm. 2.7.4, $C_c(\mathbb{R}^N)$ is weak-* dense in $L^\infty(\mathbb{R}^N, \mathfrak{m})$. Therefore $\int fh = 0$ for each $h \in L^\infty(\mathbb{R}^N, \mathfrak{m})$. It follows easily (e.g. from Lem. 1.6.15, from the isomorphism $L^\infty \simeq (L^1)^*$, or from an elementary measure-theoretic argument) that $\|f\|_{L^1} = 0$. □

7.5 Unitary subrepresentations of the group $\mathbb{R}^N$

Let $\mathcal{H}$ be a Hilbert space. We continue to use the notation at the beginning of Sec. 7.3.

Definition 7.5.1. Let $U : \mathbb{R}^N \to \mathcal{H}$ be a strongly continuous unitary representation of $\mathbb{R}^N$. Suppose that $\mathcal{K}$ is an **invariant Hilbert subspace** of $\mathcal{H}$, that is, $U(x)\mathcal{K} \subset \mathcal{K}$ for each $x \in \mathbb{R}^N$. Then

$$U|_\mathcal{K} : \mathbb{R}^N \to \mathfrak{L}(\mathcal{K}) \qquad x \mapsto U(x)|_\mathcal{K}$$

is also a strongly continuous unitary representation of $\mathbb{R}^N$, called a **(unitary) subrepresentation** of U.

In this section, we apply the Fourier transform to study unitary subrepresentations of $\mathbb{R}^N$. Specifically, we first establish Thm. 7.5.2, which connects the Borel functional calculus with the Fourier transform, and then derive Thm. 7.5.4, providing equivalent characterizations of invariant subspaces. Thm. 7.5.4 will be used in Sec. 7.6 to prove a powerful criterion for determining when a dense subspace is a core.

7.5.1 Borel functional calculus for unitary representations of the group $\mathbb{R}^N$

Let $\mathfrak{m}$ be a Haar measure on $\mathbb{R}^N$, and use it to define the Fourier transform $\widehat{f}$ for each $f \in L^1(\mathbb{R}^N, \mathfrak{m})$

Theorem 7.5.2. *Let $T_1, \dots, T_N$ be strongly commuting self-adjoint operators on $\mathcal{H}$. Let $f \in L^1(\mathbb{R}^N, \mathfrak{m})$. Then*

$$\int_{\mathbb{R}^N} f(t)e^{-\mathbf{i}tT_\bullet} d\mathfrak{m}(t) = \widehat{f}(T_\bullet)$$

where the LHS is defined to be the bounded linear map on $\mathcal{H}$ sending each $\xi \in \mathcal{H}$ to $\int_{\mathbb{R}^N} f(t)e^{-\mathbf{i}tT_\bullet}\xi d\mathfrak{m}(t)$.

Proof. Write $d\mathfrak{m}(t)$ as dt. We first note that the linear operator $\int f(t)e^{-\mathbf{i}tT_\bullet}dt$ is bounded, because for each $\xi \in \mathcal{H}$ we have

$$\Big\| \int f(t)e^{-\mathbf{i}tT_\bullet}\xi dt \Big\| \leqslant \|f\|_{L^1} \cdot \|\xi\|$$

due to Prop. 7.A.6. Moreover, if we let μ_ξ be the finite Borel measure on $\mathbb{R}^N$ associated to ξ and the Borel functional calculus of $T_\bullet$ (cf. Thm. 6.10.1), then

$$\Big\langle \xi \Big| \int f(t)e^{-\mathrm{i}tT_\bullet}dt \cdot \xi \Big\rangle = \int \langle \xi | f(t)e^{-\mathrm{i}tT_\bullet}\xi\rangle dt = \int\int f(t)e^{-\mathrm{i}tx}d\mu_\xi(x)dt$$
$$= \int\int f(t)e^{-\mathrm{i}tx}dtd\mu_\xi(x) = \int \widehat{f}(x)d\mu_\xi(x) = \langle \xi | \widehat{f}(T_\bullet)\xi\rangle$$

where the theorems of Tonelli and Fubini are applicable because

$$\int\int |f(t)e^{-\mathrm{i}tx}|dtd\mu_\xi(x) = \|f\|_{L^1}\cdot \mu_\xi(\mathbb{R}^N) < +\infty$$

This finishes the proof, thanks to the polarization identity. □

The following interpretation of Thm. 7.5.2 echoes the discussion on harmonic analysis in Subsec. 7.4.1, and establishes the property asserted at the end of Subsection 7.4.1 for the case $G = \mathbb{R}^N$, though it will not be used later in this course.

Remark 7.5.3. Define $U : \mathbb{R}^N \to \mathfrak{L}(\mathcal{H})$ by $U(t) = e^{-\mathrm{i}tT_\bullet}$. Then, using the notation from Subsec. 7.4.1, Thm. 7.5.2 states that $U(f) = \widehat{f}(T_\bullet)$. Consequently, the unitary representation $\pi : \mathscr{A} = \mathcal{F}(L^1(\mathbb{R}^N, \mathfrak{m})) \to \mathfrak{L}(\mathcal{H})$ defined by $\pi(\widehat{f}) = U(f)$ (cf. (7.13)) satisfies $\pi(\widehat{f}) = \widehat{f}(T_\bullet)$. It follows that $\pi : \mathscr{A} \to \mathfrak{L}(\mathcal{H})$ extends to a normal unitary representation $\mathscr{Bor}_{\mathrm{b}}(\mathbb{R}^N) \to \mathfrak{L}(\mathcal{H})$, given precisely by the (restricted) Borel functional calculus $\pi_{T_\bullet}|_{\mathscr{Bor}_{\mathrm{b}}(\mathbb{R}^N)}$ of $T_1, \dots, T_N$. Therefore, Thm. 7.5.2 may be interpreted as stating that the Fourier transform converts the unitary representation of the convolution algebra $L^1(\mathbb{R}^N, \mathfrak{m})$ into the Borel functional calculus of $T_\bullet$.

7.5.2 Unitary subrepresentations

Theorem 7.5.4. *Let $\mathcal{K}$ be a closed linear subspace of $\mathcal{H}$. Let $U : \mathbb{R}^N \to \mathfrak{L}(\mathcal{H})$ be a strongly continuous unitary representation of $\mathbb{R}^N$ with generators $T_1, \dots, T_N$. Then the following are equivalent.*

(1) $U(t)\mathcal{K} \subset \mathcal{K}$ for each $t \in \mathbb{R}^N$.

(2) $f(T_\bullet)\mathcal{K} \subset \mathcal{K}$ for each $f \in \mathscr{Bor}_{\mathrm{b}}(\mathbb{R}^N)$.

Proof. (1)⇒(2): Assume (1). By Prop. 7.A.8, for each $g \in C_c(\mathbb{R}^N)$ we have $\int g(t)U(-t)\mathcal{K} \subset \mathcal{K}$. Let

$$\mathscr{A} = \{\widehat{g} : g \in C_c(\mathbb{R}^N)\}$$

Then, by Thm. 7.5.2, we have $h(T_\bullet)\mathcal{K} \subset \mathcal{K}$ for each $h \in \mathscr{A}$. By Thm. 7.4.8, $\mathscr{A}$ is a dense *-subalgebra of $C_0(\mathbb{R}^N)$, and hence separates points of $\mathbb{R}^N$ and vanishes nowhere. Therefore, by Lem. 5.5.12, for each $f \in \mathscr{Bor}_{\mathrm{b}}(\mathbb{R}^N)$ there exists a net

$(f_\alpha)_{\alpha\in\mathscr{I}}$ in $\mathscr{A}$ converging to f in the universal L^2-topology. Prop. 5.5.15 (or Eq. (6.57)) then implies that $f_\alpha(T_\bullet)$ converges in SOT to $f(T_\bullet)$. Since each $f_\alpha(T_\bullet)$ leaves $\mathcal{K}$ invariant, we conclude that $f(T_\bullet)\mathcal{K}\subset\mathcal{K}$.

(2)⇒(1): Choose $f(x)=e^{\mathbf{i}tx}$. □

Remark 7.5.5. Let $E\in\mathfrak{L}(\mathcal{H})$ be the projection onto $\mathcal{K}$. Then conditions (1) and (2) in Thm. 7.5.4 are indeed equivalent to any one of the following conditions:

(3) $ET_j\subset T_jE$ for each $1\leqslant j\leqslant N$.

(4) $E\cdot\frac{T_j-\mathbf{i}}{T_j+\mathbf{i}}=\frac{T_j-\mathbf{i}}{T_j+\mathbf{i}}\cdot E$ for each $1\leqslant j\leqslant N$.

(5) $\frac{T_j-\mathbf{i}}{T_j+\mathbf{i}}\mathcal{K}=\mathcal{K}$ for each $1\leqslant j\leqslant N$.

See Pb. 7.7 for an application of condition (3).

Establishing this equivalence—and, in particular, developing the appropriate framework to understand naturally why conditions (3), (4), and (5) are all equivalent to (1) and (2)—requires extending the notion of strong commutativity to unbounded closed operators that are not necessarily self-adjoint. This, in turn, calls for the language of von Neumann algebras. Since this topic lies beyond the scope of the present course, we refer the interested reader to [Gui-S] for a comprehensive discussion. □

7.6 Application: a criterion for cores

Let $\mathcal{H}$ be a Hilbert space. The primary goal of this section is to prove Thm. 7.6.2. Its proof provides an excellent application of Thm. 7.5.4 (which, in turn, relies on the Fourier transform).

7.6.1 The main result

Lemma 7.6.1. *Let T be a self-adjoint operator on $\mathcal{H}$. Let $\mathscr{D}_0$ be a linear subspace of $\mathscr{D}(T)$. Then the following are equivalent.*

(1) $\mathscr{D}_0$ is a core for T.

(2) $\mathrm{Rng}(\mathbf{i}+T|_{\mathscr{D}_0})$ is dense in $\mathcal{H}$.

Proof. By Cor. 6.7.8, we have $\mathrm{Rng}(\mathbf{i}+T)=\mathcal{H}$. Therefore, by (6.12), we have a uniary map

$$\Phi:\mathscr{G}(T)\to\mathcal{H}\qquad\xi\oplus T\xi\mapsto(\mathbf{i}+T)\xi$$

Thus, $\mathscr{D}_0$ is a core for T iff $\mathscr{G}(T|_{\mathscr{D}_0})$ is dense in $\mathscr{G}(T)$, iff $\Phi(\mathscr{G}(T|_{\mathscr{D}_0}))$ (which equals $\mathrm{Rng}(\mathbf{i}+T|_{\mathscr{D}_0})$) is dense in $\mathcal{H}$. □

Theorem 7.6.2. *Let $U:\mathbb{R}^N\rightarrow\mathfrak{L}(\mathcal{H})$ be a strongly continuous unitary representation with generators $T_1,\dots,T_N$, and let $f\in\mathscr{B}or(\mathbb{R}^N,\mathbb{R})$. Suppose $\mathscr{D}_0$ is a dense linear subspace of $\mathscr{D}(f(T_\bullet))$ (with respect to the inner product of $\mathcal{H}$) such that*

$$U(t)\mathscr{D}_0\subset\mathscr{D}_0\qquad\text{for all } t\in\mathbb{R}^N \tag{7.20}$$

Then $\mathscr{D}_0$ is a core for $f(T_\bullet)$.

Note that by Thm. 6.10.5, $f(T_\bullet)$ is a self-adjoint operator on $\mathcal{H}$. Therefore, as a consequence of this theorem, the Hermitian operator $f(T_\bullet)|_{\mathscr{D}_0}$ is essentially self-adjoint.

Proof. Step 1. To prove that $f(T_\bullet)$ is the closure of $f(T_\bullet)|_{\mathscr{D}_0}$, by Lem. 7.6.1, it suffices to show that $\mathrm{Rng}(\mathbf{i}+f(T_\bullet)|_{\mathscr{D}_0})$ is dense in $\mathcal{H}$. Therefore, by Cor. 3.4.9, it suffices to show that the following closed linear subspace is zero:

$$\mathcal{K}:=\mathrm{Rng}(\mathbf{i}+f(T_\bullet)|_{\mathscr{D}_0})^\perp=\{\xi\in\mathcal{H}:\langle\xi|(\mathbf{i}+f(T_\bullet))\mathscr{D}_0\rangle=0\}$$

We first note that

$$\mathcal{K}\cap\mathscr{D}(f(T_\bullet))=0$$

Indeed, if $\xi\in\mathcal{K}\cap\mathscr{D}(f(T_\bullet))$, then

$$\langle(\mathbf{i}+f(T_\bullet))\xi|\mathscr{D}_0\rangle=\langle\xi|(\mathbf{i}+f(T_\bullet))\mathscr{D}_0\rangle=0$$

and hence $\xi\in\mathrm{Ker}(\mathbf{i}+f(T_\bullet))$ by the density of $\mathscr{D}_0$ in $\mathcal{H}$. Thus

$$\|\xi\|^2+\|f(T_\bullet)\xi\|^2=\|(\mathbf{i}+f(T_\bullet))\xi\|^2=0$$

and hence $\xi=0$.

Step 2. In this step, we prove that $\mathcal{K}$ is invariant under $U(t)=e^{\mathbf{i}tT_\bullet}$ for all $t\in\mathbb{R}^N$, which will prepare us for the more difficult task of showing that any $\xi\in\mathcal{K}$ (not necessarily in $\mathscr{D}(f(T_\bullet))$) must be zero.

By Thm. 6.10.5, each T_j commute strongly with $f(T_\bullet)$. Hence, by Thm. 7.3.13, we have $e^{-\mathbf{i}t_jT_j}f(T_\bullet)=f(T_\bullet)e^{-\mathbf{i}t_jT_j}$ for each j, and hence

$$e^{-\mathbf{i}tT_\bullet}f(T_\bullet)=f(T_\bullet)e^{-\mathbf{i}tT_\bullet}$$

Thus, for each $\xi\in\mathcal{K}$ we have

$$\begin{aligned}&\langle e^{\mathbf{i}tT_\bullet}\xi|(\mathbf{i}+f(T_\bullet))\mathscr{D}_0\rangle=\langle\xi|e^{-\mathbf{i}tT_\bullet}(\mathbf{i}+f(T_\bullet))\mathscr{D}_0\rangle=\langle\xi|(\mathbf{i}+f(T_\bullet))e^{-\mathbf{i}tT_\bullet}\mathscr{D}_0\rangle\\ \subset&\langle\xi|(\mathbf{i}+f(T_\bullet))\mathscr{D}_0\rangle=0\end{aligned}$$

where condition (7.20) is used. Thus $e^{\mathbf{i}tT_\bullet}\xi\in\mathcal{K}$.

Step 3. Having established that $\mathcal{K}$ is U-invariant, we invoke Theorem 7.5.4, which shows that $g(T_\bullet)\mathcal{K} \subset \mathcal{K}$ for each $g \in \mathscr{Bor}_{\mathrm{b}}(\mathbb{R}^N)$. For each $n \in \mathbb{N}$, we set

$$g_n = \chi_{f^{-1}[-n,n]}$$

Then $g_n(T_\bullet)$ is a projection because $g_n = \overline{g_n} = g_n^2$.

We now show that any $\xi \in \mathcal{K}$ must be zero by showing:

(a) $\lim_n g_n(T_\bullet)\xi$ converges to ξ.

(b) For each n, we have $g_n(T_\bullet)\xi \in \mathcal{K} \cap \mathscr{D}(f(T_\bullet))$, and thus $g_n(T_\bullet)\xi = 0$ by Step 1.

Proof of (a): By DCT, we have

$$\lim_n \|\xi - g_n(T_\bullet)\xi\|^2 \xlongequal{(6.57)} \lim_n \int_{\mathbb{R}^N} |1 - g_n|^2 d\mu_\xi = 0$$

Proof of (b): Since $g_n(T_\bullet)\mathcal{K} \subset \mathcal{K}$, we have $g_n(T_\bullet)\xi \in \mathcal{K}$. Since fg_n is bounded, by Lemma 7.6.3 below we have $g_n(T_\bullet)\xi \in \mathscr{D}(f(T_\bullet))$.

The proof is now complete. □

Lemma 7.6.3. *Let $T_1, \dots, T_N$ be strongly commuting self-adjoint operators on $\mathcal{H}$. Let $f, g \in \mathscr{Bor}(\mathbb{R}^N)$ such that $\|g\|_{l^\infty} < +\infty$. Then*

$$f(T_\bullet)g(T_\bullet) = (fg)(T_\bullet) \tag{7.21}$$

In particular, if fg is bounded, then $f(T_\bullet)g(T_\bullet) \in \mathfrak{L}(\mathcal{H})$, and hence $g(T_\bullet)\mathcal{H} \subset \mathscr{D}(f(T_\bullet))$.

Proof. By Thm. 6.10.2 and the spectral Thm. 6.9.4, we can view $f(T_\bullet)$ and $g(T_\bullet)$ as multiplication operators $\mathbf{M}_f$ and $\mathbf{M}_g$ on a direct sum of L^2-spaces. By Rem. 6.2.17, we have $\mathbf{M}_f\mathbf{M}_g = \mathbf{M}_{fg}$, which is equivalent to (7.21). □

Remark 7.6.4. We briefly explain the main idea behind the proof of Thm. 7.6.2. Let $A = f(T_\bullet)|_{\mathscr{D}_0}$. Our goal is to show that $\mathrm{Ker}(A^* - \mathbf{i}) \equiv \mathrm{Rng}(A + \mathbf{i})^\perp$ is zero, that is, every "distribution solution" (or weak solution) of the equation

$$(\mathbf{i} + f(T_\bullet))\xi = 0$$

is zero. Step 1 shows that any true solution is zero. Therefore, it remains to show that any distribution solution can be approximated by true solutions. By Step 2, the space of distribution solutions is invariant under $e^{\mathbf{i}tT_\bullet}$. Therefore, any convolution of a distribution solution ξ is itself a distribution solution.

In Step 3, we choose the convolution to be $g_n(T_\bullet)\xi$ where $g_n = \chi_{f^{-1}[-n,n]}$. Indeed, by Thm. 7.5.2, $g_n(T_\bullet)\xi$ can be viewed as the convolution $\int \check{g}_n e^{-\mathbf{i}tT_\bullet}\xi$ where $\check{g}_n$ is the (possibly only formal) inverse Fourier transform of g_n. One the one hand, since $\lim g_n = 1$, the convolution converges to ξ as $n \to +\infty$. On the other hand, the boundedness of fg_n ensures that the convolution $g_n(T_\bullet)\xi$ belongs to the domain of $f(T_\bullet)$, and hence is a true solution. This finishes the proof.

We remark that when f is a polynomial, one may instead take g_n to be the Fourier transform of $h_n(x) = n^N h(nx)$ where $h \in C_c^\infty(\mathbb{R}^N)$ satisfies $\int h = 1$. Then $g_n(T_\bullet)\xi = \int h_n(t)e^{-\mathbf{i}tT_\bullet}\xi dt$ converges to ξ as in the proof of Lem. 7.3.10. Moreover, $g_n(T_\bullet)\xi$ is a true solution, because $g_n \in \mathcal{S}(\mathbb{R}^N)$ (cf. Prop. 7.8.2) and hence fg_n is bounded. □

As the following exercise shows, Thm. 7.6.2 also holds for complex-valued functions f, though we will not need this fact later in this course.

Exercise 7.6.5. In Thm. 7.6.2, suppose instead that $f \in \mathscr{Bor}(\mathbb{R}^N) \equiv \mathscr{Bor}(\mathbb{R}^N, \mathbb{C})$, without assuming that f is real-valued. Choose any $u \in \mathscr{Bor}_{\mathrm{b}}(\mathbb{R}^N, \mathbb{S}^1)$ such that $u(x) = f(x)/|f|(x)$ whenever $x \in \mathbb{R}^N$ satisfies $f(x) \neq 0$. Use Thm. 6.10.2 to show that $u(T_\bullet)$ is unitary, and that

$$f(T_\bullet) = u(T_\bullet) \cdot |f|(T_\bullet) \tag{7.22}$$

Conclude that $f(T_\bullet)$ is closed, that $f(T_\bullet)$ and $|f|(T_\bullet)$ share the same cores, and hence that $\mathscr{D}_0$ is a core for $f(T_\bullet)$.

7.6.2 Applications

Example 7.6.6. Let $\mathcal{H} = \bigoplus_{\alpha \in \mathscr{I}} L^2(\mathbb{R}^N, \mu_\alpha)$, where each μ_α is a Borel measure on $\mathbb{R}^N$. Let $(E_n)_{n \in \mathbb{Z}_+}$ be an increasing sequence of bounded Borel subsets of $\mathbb{R}^N$ such that $\mathbb{R}^N = \bigcup_n E_n$. By DCT, the subspace

$$\mathscr{D}_0 = \bigcup_{n \in \mathbb{Z}_+} \mathbf{M}_{\chi_{E_n}} \mathcal{H}$$

is dense in $\mathcal{H}$. Let $f : \mathbb{R}^N \to \mathbb{R}$ be a Borel function that is bounded on compact subsets of $\mathbb{R}^N$. One readily verifies that $\mathscr{D}_0 \subset \mathscr{D}(\mathbf{M}_f)$. Since $\mathbf{M}_{e_t}\mathscr{D}_0 \subset \mathscr{D}_0$ for each $t \in \mathbb{R}^N$, Thm. 7.6.2 implies that $\mathscr{D}_0$ is a core for $\mathbf{M}_f$.

Corollary 7.6.7. *For each $1 \leqslant j \leqslant N$, the Hermitian operator $\mathbf{i}\frac{\partial}{\partial x_j}$ on $L^2(\mathbb{R}^N, m)$ with domain $C_c^\infty(\mathbb{R}^N)$ is essentially self-adjoint. Moreover, if we let T_j be the closure of $\mathbf{i}\frac{\partial}{\partial x_j}$, then $T_1, \dots, T_N$ are strongly commuting, and*

$$U : \mathbb{R}^N \to \mathfrak{L}(L^2(\mathbb{R}^N, m)) \qquad t \mapsto e^{\mathbf{i}tT_\bullet}$$

is equal to the regular representation of $\mathbb{R}^N$ (cf. Exp. 7.3.7).

It follows from Cor. 6.7.10 that any Hermitian extension of $\mathbf{i}\frac{\partial}{\partial x_j}$ is essentially self-adjoint with closure T_j.

Proof. Let W be the regular representation of $\mathbb{R}^N$. That is, $(W(t)\xi)(x) = \xi(x-t)$ for each $\xi \in L^2(\mathbb{R}^N, m)$. By Stone's Thm. 7.3.3, the generators $S_1, \dots, S_N$ of W commute strongly and are given by vector-valued derivatives. For each $\xi \in C_c^\infty(\mathbb{R}^N)$, by using DCT one easily checks that

$$\partial_{t_j} W(t)\xi\big|_{t=0} = -\partial_{x_j} f$$

Hence $f \in \mathscr{D}(S_j)$ and $S_j f = \mathbf{i}\partial_{x_j} f$. We conclude that $\mathbf{i}\frac{\partial}{\partial x_j} \subset S_j$.

Clearly $C_c^\infty(\mathbb{R}^N)$ is W-invariant. Therefore, by Thm. 7.6.2, $C_c^\infty(\mathbb{R}^N)$ is a core for S_j. This proves that $T_j = S_j$. Since S_j is self-adjoint, $\mathbf{i}\frac{\partial}{\partial x_j}$ is essentially self-adjoint. Since $S_1, \dots, S_N$ commuting strongly, so do $T_1, \dots, T_N$. $\square$

Corollary 7.6.8. *Choose a polynomial $f \in \mathbb{R}[x_1, \dots, x_N]$. Then the Hermitian operator*

$$A = f\Big(\mathbf{i}\frac{\partial}{\partial x_1}, \dots, \mathbf{i}\frac{\partial}{\partial x_N}\Big) \qquad \mathscr{D}(A) = C_c^\infty(\mathbb{R}^N)$$

on $L^2(\mathbb{R}^N, m)$, defined by the usual polynomial functional calculus of linear operators (cf. Exp. 6.2.13), is essentially self-adjoint. The closure $\overline{A}$ is equal to the Borel functional calculus $f(T_\bullet)$ where each T_j is the closure of $\mathbf{i}\frac{\partial}{\partial x_i}$ with $\mathscr{D}(\mathbf{i}\frac{\partial}{\partial x_i}) = C_c^\infty(\mathbb{R}^N)$.

It follows from Cor. 6.7.10 that any Hermitian extension of A is essentially self-adjoint with closure $\overline{A}$. For example, if we define B to be $f(\mathbf{i}\frac{\partial}{\partial x_1}, \dots, \mathbf{i}\frac{\partial}{\partial x_N})$ with domain $\mathcal{S}(\mathbb{R}^N)$ (cf. Def. 7.8.1), then $\overline{A} = \overline{B}$.

Proof. We use the notation in the proof of Cor. 7.6.7. Then $T_j = S_j$. Since $C_c^\infty(\mathbb{R}^N)$ is W-invariant, by Thm. 7.6.2, $C_c^\infty(\mathbb{R}^N)$ is a core for the self-adjoint operator $f(S_\bullet)$ defined by the Borel functional calculus. Therefore $f(S_\bullet)|_{C_c^\infty(\mathbb{R}^N)}$ is essentially self-adjoint.

By viewing $S_1, \dots, S_N$ as multiplication operators (so that $f(S_\bullet)$ is also a multiplication operator, cf. Thm. 6.10.2), one concludes from Exe. 6.2.19 that $f(S_\bullet)|_{C_c^\infty(\mathbb{R}^N)}$ equals A. Therefore A is essentially self-adjoint with closure $f(S_\bullet) = f(T_\bullet)$. $\square$

Recall Def. 7.1.12 for the Dirichlet and Neumann Laplacians.

Corollary 7.6.9. *The smooth Laplacian Δ on $L^2(\mathbb{R}^N, m)$ with domain $C_c^\infty(\mathbb{R}^N)$ is essentially self-adjoint, and its closure satisfies*

$$\overline{\Delta} = \Delta_{\mathrm{D}} = \Delta_{\mathrm{N}}$$

Consequently, by Rem. 7.1.13 we have

$$H_0^1(\mathbb{R}^N) = H^1(\mathbb{R}^N) = \mathscr{D}(\sqrt{-\overline{\Delta}})$$

Proof. By Cor. 7.6.8, Δ is essentially self-adjoint, and hence any Hermitian extension of Δ has closure $\overline{\Delta}$. Since both Δ_{D} and Δ_{N} are self-adjoint (and hence closed) extensions of Δ, they must be equal to $\overline{\Delta}$. $\square$

The results obtained so far in this subsection can also be proved directly via the Fourier transform, without invoking Thm. 7.6.2; see Pb. 7.11 and 7.12. However, using Thm. 7.6.2, one can establish the following generalization of Cor. 7.6.9, which is difficult to prove by a direct Fourier transform approach: the Laplacian on a complete Riemannian manifold M with domain $C_c^\infty(M)$ is essentially self-adjoint. See [Tay-2, Sec. 8.2] or [Che73].

In the following, we present a simple example whose solution via the Fourier transform is considerably more complicated than that obtained using Thm. 7.6.2.

Example 7.6.10. Let $x_0, x_1, \dots, x_N$ be the standard coordinates of $\mathbb{R} \times \mathbb{R}^N$. Let Ω be an open subset of $\mathbb{R}^N$. Let μ be a Radon measure on Ω. Let $\mathcal{H} = L^2(\mathbb{R} \times \Omega, m \times \mu)$. Then the Hermitian operator $\mathbf{i}\frac{\partial}{\partial x_0}$ on $\mathcal{H}$ with domain $C_c^\infty(\mathbb{R} \times \Omega)$ is essentially self-adjoint.

See Subsec. 1.7.5 for details on product measures.

Proof. The map $U : \mathbb{R} \to \mathfrak{L}(\mathcal{H})$ defined by

$$(U(t)f)(x_0, x_1, \dots, x_N) = f(x_0 - t, x_1, \dots, x_N)$$

is a unitary representation of $\mathbb{R}$ on $\mathcal{H}$. Since U is strongly continuous on the dense linear subspace $C_c(\mathbb{R} \times \Omega)$, it is strongly continuous on $\mathcal{H}$. Let T be the generator of U. Then by Stone's Thm. 7.3.3, for each $f \in C_c^\infty(\mathbb{R} \times \Omega)$ we have $f \in \mathscr{D}(T)$ and $\mathbf{i}Tf = \frac{d}{dt}U(t) = -\partial_{x_0} f$. This proves $T|_{C_c^\infty(\mathbb{R}\times\Omega)} = \mathbf{i}\frac{\partial}{\partial x_0}$. Since $C_c^\infty(\mathbb{R} \times \Omega)$ is dense in $\mathcal{H}$ and U-variant, it follows from Thm. 7.6.2 that $C_c^\infty(\mathbb{R} \times \Omega)$ is a core for T. Therefore $\mathbf{i}\frac{\partial}{\partial x_0}$ is essentially self-adjoint. $\square$

Remark 7.6.11. Readers might think that Exp. 7.6.10 can be solved using a "partial Fourier transform", although they may not readily find such a theory explicitly formulated in the literature. In fact, the theory of strongly continuous unitary representations $U : \mathbb{R}^N \to \mathcal{H}$ developed in this course is precisely a general framework for partial Fourier transforms, since U can be viewed as parametrizing the Hilbert space $\mathcal{H}$ (or its underlying geometric space) by N real parameters!

This general framework also enables us to address more intricate dynamical problems. For instance, suppose a smooth manifold M endowed with a Radon measure μ admits a smooth flow $t \in \mathbb{R} \mapsto \alpha_t$ that preserves μ.[7] Then Thm. 7.6.2 implies that the Hermitian operator S on $L^2(M, \mu)$ with domain $\mathscr{D}(S) = C_c^\infty(M)$ defined by $Sf = -\mathbf{i}\frac{d}{dt}(f \circ \alpha_t)\big|_{t=0}$ is essentially self-adjoint, since $C_c^\infty(M)$ is a core for the generator of the unitary representation $U : \mathbb{R} \to \mathfrak{L}(L^2(M, \mu))$ given by $U(t)f = f \circ \alpha_t$. By the same reasoning, any real polynomial of S is essentially self-adjoint. $\square$

[7] Here, each α_t is a diffeomorphism of M.

7.7 Application: dense subspaces of $L^2(\mathbb{R}^N,\mu)$

In this section, we present another application of Thm. 7.5.4 by establishing two useful criteria for dense linear subspaces of $L^2(\mathbb{R}^N,\mu)$, where μ is a finite Borel measure on $\mathbb{R}^N$. Recall from (7.11) that $e_t(x)=e^{\mathbf{i}tx}$.

7.7.1 Density of trigonometric functions

Proposition 7.7.1. *Let μ be a finite Borel measure on $\mathbb{R}^N$. Then*

$$\mathrm{Span}\{e_t:t\in\mathbb{R}^N\}$$

is dense in $L^2(\mathbb{R}^N,\mu)$.

Proof. By viewing 1 as an element of $\mathcal{H}:=L^2(\mathbb{R}^N,\mu)$, we have

$$\mathrm{Span}\{e_t:t\in\mathbb{R}^N\}=\mathrm{Span}\{e^{\mathbf{i}t\mathbf{M}_x}1:t\in\mathbb{R}^N\}$$

Let $\mathcal{K}$ denote the closure of this space. We need to show that $\mathcal{K}=\mathcal{H}$.

Since $\mathcal{K}$ is $e^{\mathbf{i}t\mathbf{M}_x}$-invariant, by Thm. 7.5.4 we have $\beta(\mathbf{M}_x)\mathcal{K}\subset\mathcal{K}$ for each $\beta\in\mathscr{B}or_{\mathrm{b}}(\mathbb{R}^N)$, and hence $\beta\equiv\beta(\mathbf{M}_x)1$ belongs to $\mathcal{K}$. Therefore $\mathcal{K}$ contains $L^\infty(\mathbb{R}^N,\mu)$. Since $L^\infty(\mathbb{R}^N,\mu)$ is dense in $L^2(\mathbb{R}^N,\mu)$, we conclude that $\mathcal{K}=\mathcal{H}$. □

The following result is a continuous analogue of the fact that $\mathbb{C}[z_1^{\pm1},\dots,z_N^{\pm1}]$ is l^∞-dense in $C(\mathbb{T}^N)$ by the Stone-Weierstrass theorem. Can you guess how to extend this result to an arbitrary LCA group?

Exercise 7.7.2. Show that $\mathrm{Span}\{e_t:t\in\mathbb{R}^N\}$ is dense in $\mathscr{B}or_{\mathrm{b}}(\mathbb{R}^N)$ in the universal L^2-topology (cf. Def. 5.5.10) by adapting the proof of Lem. 5.5.12, replacing the use of Lusin's theorem with an application of Prop. 7.7.1.

7.7.2 Density of polynomials

Definition 7.7.3. Let $\Omega\subset\mathbb{R}^N$ be open. A function $f:\Omega\to\mathbb{C}$ is called **holomorphic** if f is continuous, and if f is holomorphic in each variable.

With the help of Prop. 7.7.1 we prove:

Theorem 7.7.4. *Let μ be a finite Borel measure on $\mathbb{R}^N$. Assume that there exists $r\in\mathbb{R}_{>0}$ such that*

$$\int_{\mathbb{R}^N}e^{r(|x_1|+\cdots+|x_N|)}d\mu(x)<+\infty\tag{7.23a}$$

equivalently (by MCT), that

$$\sum_{n_1,\dots,n_N\in\mathbb{N}}\frac{r^{n_1+\cdots+n_N}}{n_1!\cdots n_N!}\int_{\mathbb{R}^N}|x_1|^{n_1}\cdots|x_N|^{n_N}d\mu(x)<+\infty\tag{7.23b}$$

Then $\mathbb{C}[x_1,\dots,x_N]$ is dense in $L^2(\mathbb{R}^N,\mu)$.

Proof. By Cor. 3.4.9, it suffices to show that any vector orthogonal to $\mathbb{C}[x_\bullet] \equiv \mathbb{C}[x_1, \dots, x_N]$ must be zero. Let us fix $f \in L^2(\mathbb{R}^N, \mu)$ orthogonal to $\mathbb{C}[x_\bullet]$. To prove $f = 0$, by Prop. 7.7.1, it suffices to show that $f \perp e_t$ for each $t \in \mathbb{R}^N$. We establish this in the following steps.

Step 1. In this step, we construct a holomorphic function

$$g : \Omega^N \to \mathbb{C} \qquad g(z) = \int f(x) e^{zx} d\mu(x)$$
$$\text{where } \Omega = \{\zeta \in \mathbb{C} : |\mathrm{Re}(\zeta)| < r/2\}$$

First, we need to check that the family of functions $x \mapsto f(x)e^{zx}$ (indexed by $z \in \Omega^N$) is dominated by a positive L^1 function so that DCT can be applied. Let $\varphi : \mathbb{R}^N \to \mathbb{R}_{\geqslant 0}$ be

$$\varphi(x) = e^{r(|x_1| + \cdots + |x_N|)/2}$$

Then $\varphi \in L^2(\mathbb{R}^N, \mu)$ by (7.23a). Since $f \in L^2$, we have $f\varphi \in L^1(\mathbb{R}^N, \mu)$. Since

$$|e^{zx}| \leqslant |\varphi(x)|$$

the family of functions $f(x)e^{zx}$ is dominated $|f\varphi|$. Therefore, by DCT, g is well-defined and continuous.

Moreover, we can also apply Fubini's theorem, which shows that whenever $z_1, \dots, z_{j-1}, z_{j+1}, \dots, z_N$ are fixed, the contour integral of g over the variable z_j on any triangle in Ω is zero. Thus, Morera's theorem (applied to each variable of g) shows that g is holomorphic.

Step 2. We now prove that $g = 0$ on Ω^N; the special case $g(-\mathbf{i}t) = 0$ for all $t \in \mathbb{R}$ then implies $f \perp e_t$, completing the proof.

Since g is holomorphic on each variable, it suffices to show that $g = 0$ on the polydisk $B_{\mathbb{C}}(0, r/2)^N$. For each $z \in B_{\mathbb{C}}(0, r/2)^N$, since

$$\sum_{n_1, \dots, n_N \in \mathbb{N}} |f(x)| \cdot \frac{|z_1 x_1|^{n_1} \cdots |z_N x_N|^{n_N}}{n_1! \cdots n_N!} = |f(x)| e^{|z_1 x_1| + \cdots + |z_N x_N|} \leqslant |f(x)\varphi(x)|$$

and since $f\varphi \in L^1(\mathbb{R}^N, \mu)$, by DCT we have

$$g(z) = \sum_{n_1, \dots, n_N \in \mathbb{N}} \int f(x) \cdot \frac{(z_1 x_1)^{n_1} \cdots (z_N x_N)^{n_N}}{n_1! \cdots n_N!} d\mu(x)$$

where the RHS equals zero by the assumption that $f \perp \mathbb{C}[x_\bullet]$. This proves $g = 0$ on $B_{\mathbb{C}}(0, r/2)^N$. $\square$

Exercise 7.7.5. Let $(c_{n_\bullet})_{n_1, \dots, n_N \in \mathbb{N}}$ be a family in $\mathbb{C}$. Show that the following are equivalent:

(1) There exists $r \in \mathbb{R}_{>0}$ such that

$$\sum_{n_1,\dots,n_N\in\mathbb{N}} r^{n_1+\cdots+n_N}\cdot|c_{n_\bullet}| < +\infty$$

(2) There exists $\rho \in \mathbb{R}_{>0}$ such that

$$\sup_{n_1,\dots,n_N\in\mathbb{N}} \rho^{n_1+\cdots+n_N}\cdot|c_{n_\bullet}| < +\infty$$

Conclude that in Thm. 7.7.4, the existence of $r \in \mathbb{R}_{>0}$ satisfying (7.23) is equivalent to the existence of $\rho \in \mathbb{R}_{>0}$ satisfying

$$\sup_{n_1,\dots,n_N\in\mathbb{N}} \frac{\rho^{n_1+\cdots+n_N}}{n_1!\cdots n_N!}\int_{\mathbb{R}^N}|x_1|^{n_1}\cdots|x_N|^{n_N}d\mu(x) < +\infty \tag{7.24}$$

7.7.3 Application: density of $\mathcal{S}_0(\mathbb{R}^N)$ in $L^2(\mathbb{R}^N,m)$

As an application of Thm. 7.7.4, we prove the density of $\mathcal{S}_0(\mathbb{R}^N)$ in $L^2(\mathbb{R}^N,m)$. See Pb. 7.14 for another application of Thm. 7.7.4 to the uniqueness of solutions of the polynomial moment problem.

Definition 7.7.6. $\mathcal{S}_0(\mathbb{R}^N)$ is the space spanned by functions of the form

$$f(x) = x_1^{n_1}\cdots x_N^{n_N}e^{-\frac{|x|^2}{2}} \qquad \text{where } n_1,\dots,n_N \in \mathbb{N} \tag{7.25}$$

Clearly $\mathcal{S}_0(\mathbb{R}^N)$ is a subspace of $L^2(\mathbb{R}^N,m)$, which will play an important role in the study of Fourier transforms in Sec. 7.8. In particular, we will need the density of $\mathcal{S}_0(\mathbb{R}^N)$ in $L^2(\mathbb{R}^N,m)$ in the proof of Thm. 7.8.4.

Corollary 7.7.7. *$\mathcal{S}_0(\mathbb{R}^N)$ is dense in $L^2(\mathbb{R}^N,m)$.*

Proof. Let $h(x) = e^{-|x|^2/2}$ and $d\mu = hdm$. Then $L^2(\mathbb{R}^N,m) \subset L^2(\mathbb{R}^N,\mu)$. Choose any $f \in L^2(\mathbb{R}^N,m)$ orthogonal to $\mathcal{S}_0(\mathbb{R}^N)$ in $L^2(\mathbb{R}^N,m)$. Then f is orthogonal to $\mathbb{C}[x_\bullet]$ in $L^2(\mathbb{R}^N,\mu)$. By Thm. 7.7.4, we conclude that $f=0$ in $L^2(\mathbb{R}^N,\mu)$, and hence $f=0$ in $L^2(\mathbb{R}^N,m)$. □

Example 7.7.8. One important reason for studying $\mathcal{S}_0(\mathbb{R}^N)$ is that the Hamiltonian operator for the **quantum harmonic oscillator**

$$H = -\Delta + \mathbf{M}_{|x|^2} \qquad \mathscr{D}(H) = \mathcal{S}_0(\mathbb{R}^N)$$

is diagonalizable. Since $\mathcal{S}_0(\mathbb{R}^N)$ is dense, it follows from the following Prop. 7.7.9 that H is essentially self-adjoint.

Proof. We show that H is diagonalizable. Define the **raising operator**

$$A_j = \mathbf{M}_{x_j} - \frac{\partial}{\partial x_j} \qquad \mathscr{D}(A_j) = \mathcal{S}_0(\mathbb{R}^N).$$

A direct computation shows that

$$[H, A_j] = 2A_j.$$

Thus, if $H\xi = \lambda\xi$ then $HA_j\xi = (\lambda+2)A_j\xi$. Therefore, since $He^{-|x|^2/2} = e^{-|x|^2/2}$, for each $n_1, \dots, n_N \in \mathbb{N}$ we have

$$HA_\bullet^{n_\bullet} e^{-\frac{|x|^2}{2}} = (1 + 2n_1 + \cdots + 2n_N) A_\bullet^{n_\bullet} e^{-\frac{|x|^2}{2}}$$

where $A_\bullet^{n_\bullet}$ abbreviates $A_1^{n_1} \cdots A_N^{n_N}$. Since $A_\bullet^{n_\bullet} e^{-\frac{|x|^2}{2}}$ is equal to $e^{-\frac{|x|^2}{2}}$ multiplied by a polynomial of $x_\bullet$ with multi-degree $(n_1, \dots, n_N)$, it follows that

$$\mathcal{S}_0(\mathbb{R}^N) = \mathrm{Span}\{A_\bullet^{n_\bullet} e^{-\frac{|x|^2}{2}} : n_1, \dots, n_N \in \mathbb{N}\}.$$

Thus $\mathcal{S}_0(\mathbb{R}^N)$ is spanned by eigenvectors of H. □

Proposition 7.7.9. *Let T be a Hermitian operator on $\mathcal{H}$. Suppose that $(e_i)_{i\in I}$ is an orthonormal basis of $\mathcal{H}$ such that each e_i belongs to $\mathscr{D}(T)$, and $Te_i = \lambda_i e_i$ with $\lambda_i \in \mathbb{R}$. Then T is essentially self-adjoint. Moreover, for each $f \in \mathscr{Bor}(\mathbb{R})$, the Borel functional calculus $f(\overline{T})$ satisfies $e_i \in \mathscr{D}(f(\overline{T}))$ and*

$$f(\overline{T})e_i = f(\lambda_i)e_i$$

for each $i \in I$.

Proof. We may identify $\mathcal{H}$ with $\bigoplus_{i\in I} L^2(\mathbb{R}, \delta_{\lambda_i})$ and identify each e_i ith $1 \in L^2(\mathbb{R}, \delta_{\lambda_i})$, where δ_{λ_i} is the Dirac measure at λ_i. Then $\mathscr{D}_0 := \mathrm{Span}\{e_i : i \in I\}$ equals $\sum_{i\in I} L^2(\mathbb{R}, \delta_{\lambda_i})$. By Exp. 7.6.6 or Thm. 7.6.2 (or by a direct argument), $\mathscr{D}_0$ is a core for $\mathbf{M}_x$. Thus $T|_{\mathscr{D}_0} = \mathbf{M}_x|_{\mathscr{D}_0}$ is essentially self-adjoint with closure $\mathbf{M}_x$. By Cor. 6.7.10, any Hermitian extension of $T|_{\mathscr{D}_0}$ must have closure $\overline{T|_{\mathscr{D}_0}} = \mathbf{M}_x$. Therefore $\overline{T} = \mathbf{M}_x$; in particular, T is essentially self-adjoint.

By Thm. 6.10.2, we have $f(\mathbf{M}_x) = \mathbf{M}_f$. Therefore each e_i belongs to the domain of $\mathbf{M}_f$, and since $\mathbf{M}_f e_i = f(\lambda_i)e_i$, the same hold for $f(\mathbf{M}_x) = f(\overline{T})$. □

The topic of this section will be continued in Subsec. 7.10.1. In particular, Exp. 7.10.8 of that subsection will show that $\mathcal{S}_0(\mathbb{R}^N)$ is a core for real polynomials of $\mathbf{M}_{x_1}, \dots, \mathbf{M}_{x_N}$ and for (the closures of) those of $-\mathbf{i}\frac{\partial}{\partial x_1}, \dots, -\mathbf{i}\frac{\partial}{\partial x_N}$.

7.8 Fourier transform as a spectral theorem for the regular representation of the group $\mathbb{R}^N$

In this section, we adopt the Haar measure on $\mathbb{R}^N$ given by

$$d\mathfrak{m} = (2\pi)^{-\frac{N}{2}} dm$$

and continue to follow Convention 7.4.1 by writing $d\mathfrak{m}(x)$ as dx. The main reason for choosing this Haar measure is because

$$\int_{\mathbb{R}^N} e^{-\frac{|x|^2}{2}} d\mathfrak{m}(x) = 1 \tag{7.26}$$

By Cor. 7.6.7, the Hermitian operators $-\mathbf{i}\frac{\partial}{\partial x_1}, \dots, -\mathbf{i}\frac{\partial}{\partial x_N}$ on $L^2(\mathbb{R}^N, \mathfrak{m})$ with domain $C_c^\infty(\mathbb{R}^N)$ have self-adjoint closures $T_1, \dots, T_N$. The spectral theorem then implies that $T_1, \dots, T_N$ are unitarily equivalent to the multiplication operators $\mathbf{M}_{x_1}, \dots, \mathbf{M}_{x_N}$ on a Hilbert space of the form $\bigoplus_\alpha L^2(\mathbb{R}^N, \mu_\alpha)$.

In this section, we show that the Fourier transform defines a unitary operator $\mathcal{F} : L^2(\mathbb{R}^N, \mathfrak{m}) \to L^2(\mathbb{R}^N, \mathfrak{m})$ satisfying

$$\mathcal{F}T_j\mathcal{F}^{-1} = \mathbf{M}_{x_j} \qquad \text{on } L^2(\mathbb{R}^N, \mathfrak{m})$$

thereby providing an explicit spectral decomposition for the differential operators $T_1, \dots, T_N$. Since Cor. 7.6.7 identifies $-T_1, \dots, -T_N$ as the generators of the regular representation U of $\mathbb{R}^N$, the Fourier transform on $L^2(\mathbb{R}^N, \mathfrak{m})$ also yields a spectral decomposition for U in the form of (7.14): [8]

$$\mathcal{F}U(t)\mathcal{F}^{-1} = \mathbf{M}_{e_{-t}} \qquad \text{on } L^2(\mathbb{R}^N, \mathfrak{m})$$

7.8.1 The Schwartz space $\mathcal{S}(\mathbb{R}^N)$

Definition 7.8.1. A smooth function $f : \mathbb{R}^N \to \mathbb{C}$ is called a **Schwartz function** (or **rapidly decreasing function**) if for each $m_1, \dots, m_N, n_1, \dots, n_N \in \mathbb{N}$ we have

$$\left\| x_1^{m_1} \cdots x_N^{m_N} \partial_{x_1}^{n_1} \cdots \partial_{x_N}^{n_N} f \right\|_{l^\infty(\mathbb{R}^N)} < +\infty$$

The space of Schwartz functions is denoted by $\boldsymbol{\mathcal{S}(\mathbb{R}^N)}$.

Intuitively, if $f \in \mathcal{S}(\mathbb{R}^N)$, then all partial derivatives of f grow lower than arbitrary polynomials. In particular, we have $\mathcal{S}(\mathbb{R}^N) \subset L^p(\mathbb{R}^N, \mathfrak{m})$ for each $1 \leqslant p \leqslant +\infty$.

[8] Note, however, one distinction: in (7.14) each measure μ_α is finite, whereas here the Haar measure $\mathfrak{m}$ is σ-finite but not finite.

Proposition 7.8.2. *For each $1 \leqslant i \leqslant N$ we have*

$$\partial_{x_i}\mathcal{S}(\mathbb{R}^N) \subset \mathcal{S}(\mathbb{R}^N) \qquad x_i\mathcal{S}(\mathbb{R}^N) \subset \mathcal{S}(\mathbb{R}^N) \qquad \widehat{\mathcal{S}(\mathbb{R}^N)} \subset \mathcal{S}(\mathbb{R}^N)$$

Moreover, for each $f \in \mathcal{S}(\mathbb{R}^N)$ and $t = (t_1, \ldots, t_N)$ we have

$$\widehat{\partial_{x_i} f}(t) = \mathbf{i}t_i\widehat{f}(t) \qquad \widehat{x_i f}(t) = \mathbf{i}\partial_{t_i}\widehat{f}(t) \tag{7.27}$$

Similar to (7.27), we also have

$$\widecheck{\partial_{x_i} f}(t) = -\mathbf{i}t_i\widehat{f}(t) \qquad \widecheck{x_i f}(t) = -\mathbf{i}\partial_{t_i}\widehat{f}(t)$$

Proof. The space $\mathcal{S}(\mathbb{R}^N)$ is clearly invariant under partial derivatives and multiplication by polynomials. If $f \in \mathcal{S}(\mathbb{R}^N)$, integration by parts gives

$$\widehat{\partial_{x_i} f}(t) = \int \partial_{x_i} f(x) e^{-\mathbf{i}tx}dx = -\int f(x)\partial_{x_i} e^{-\mathbf{i}tx}dx = \mathbf{i}t_i\widehat{f}(t)$$

DCT shows that $\partial_{t_i}\widehat{f}$ exists everywhere and

$$\partial_{t_i}\widehat{f}(t) = \int f(x)\partial_{t_i} e^{-\mathbf{i}tx}dx = -\mathbf{i}\int x_i f(x) e^{-\mathbf{i}tx}dx = -\mathbf{i}\widehat{x_i f}(t)$$

This proves (7.27). By applying (7.27) repeatedly, we see that $\widehat{f}$ is smooth, and that any expression of the form $g = t_1^{m_1}\cdots t_N^{m_N}\partial_{t_1}^{n_1}\cdots\partial_{t_N}^{n_N}\widehat{f}$ can be written as the Fourier transform of some $h \in \mathcal{S}(\mathbb{R}^N)$. Therefore, by (7.19), the function g is bounded since $\|g\|_{l^\infty} \leqslant \|h\|_{L^1} < +\infty$. Hence $\widehat{f} \in \mathcal{S}(\mathbb{R}^N)$. We have thus proves that $\mathcal{S}(\mathbb{R}^N)$ is invariant under the Fourier transform. $\square$

7.8.2 The Fourier transform on $\mathcal{S}_0(\mathbb{R}^N)$

The space $\mathcal{S}_0(\mathbb{R}^N)$ defined in Def. 7.7.6 is clearly a linear subspace of $\mathcal{S}(\mathbb{R}^N)$ invariant under partial derivatives and multiplication by polynomials. An advantage of working with $\mathcal{S}_0(\mathbb{R}^N)$ is that the Fourier transform can be explicitly computed in this space.

Proposition 7.8.3. *Let f be defined by (7.25), that is, $f(x) = x_1^{n_1}\cdots x_N^{n_N}e^{-\frac{|x|^2}{2}}$. Then*

$$\widehat{f} = \mathbf{i}^{n_1+\cdots+n_N}f \qquad \widecheck{f} = (-\mathbf{i})^{n_1+\cdots+n_N}f$$

Proof. By Prop. 7.8.2, it suffices to prove that $\widehat{f} = f$ when $f(x) = e^{-|x|^2/2}$. Moreover, it suffices to prove the case $N = 1$; the general case follows from Fubini's theorem, since the function $e^{-|x|^2/2-\mathbf{i}tx}$ is a product of functions depending only on $x_1, \ldots, x_N$ respectively.

Assume $N = 1$, and consider the function $g : \mathbb{C} \to \mathbb{C}$ defined by

$$g(z) := \int_{\mathbb{R}} e^{-\frac{x^2}{2}} e^{-zx} d\mathfrak{m}(x)$$

Since $|e^{-x^2/2}e^{-zx}| = |e^{-x^2/2} \cdot e^{-x\mathrm{Re}z}|$ is integrable, $g(z)$ can be defined for all $z \in \mathbb{C}$. Moreover, by DCT, g is continuous. By Fubini's theorem, the contour integral of g on any triangle in $\mathbb{C}$ is zero. Therefore, Morera's theorem implies that g is a holomorphic function.

When $z \in \mathbb{R}$, by (7.26) we have

$$g(z) = \int e^{-\frac{(x+z)^2}{2}} \cdot e^{\frac{z^2}{2}} d\mathfrak{m}(x) = e^{\frac{z^2}{2}} \int e^{-\frac{x^2}{2}} d\mathfrak{m}(x) = e^{\frac{z^2}{2}}$$

By the holomorphicity, we conclude that $g(z) = e^{z^2/2}$ for all $z \in \mathbb{C}$, and hence $g(-\mathbf{i}t) = e^{-t^2/2}$ for all $t \in \mathbb{R}$. This proves $\widehat{f} = f$. □

7.8.3 The Fourier transform on $\mathcal{S}(\mathbb{R}^N)$ and $L^2(\mathbb{R}^N)$

Define a clearly unitary map

$$\Theta : L^2(\mathbb{R}^N, \mathfrak{m}) \to L^2(\mathbb{R}^N, \mathfrak{m}) \qquad (\Theta f)(x) = f(-x) \tag{7.28}$$

Theorem 7.8.4. *There exists a unique bounded linear map*

$$\mathcal{F} : L^2(\mathbb{R}^N, \mathfrak{m}) \to L^2(\mathbb{R}^N, \mathfrak{m}) \tag{7.29a}$$

called the ***Fourier transform on*** $L^2(\mathbb{R}^N, \mathfrak{m})$*, satisfying*

$$\mathcal{F}f = \widehat{f} \qquad \text{for each } f \in \mathcal{S}(\mathbb{R}^N) \tag{7.29b}$$

Moreover, the map (7.29) *is unitary, satisfies*

$$\mathcal{F}^2 = \Theta \qquad \mathcal{F}^4 = \mathbf{1}$$
$$\mathcal{F}^3 f = \check{f} \quad \text{if } f \in \mathcal{S}(\mathbb{R}^N)$$

and restricts to bijections $\mathcal{S}(\mathbb{R}^N) \xrightarrow{\simeq} \mathcal{S}(\mathbb{R}^N)$ *and* $\mathcal{S}_0(\mathbb{R}^N) \xrightarrow{\simeq} \mathcal{S}_0(\mathbb{R}^N)$*.*

Proof. Step 1. Uniqueness follows from the density of $C_c^\infty(\mathbb{R}^N) \subset \mathcal{S}(\mathbb{R}^N)$ in $L^2(\mathbb{R}^N, \mathfrak{m})$ (cf. Thm. 1.7.11). We now establish existence.

By Prop. 7.8.2, we have a linear map $\mathcal{F} : \mathcal{S}(\mathbb{R}^N) \to \mathcal{S}(\mathbb{R}^N)$ defined by $\mathcal{F}f = \widehat{f}$. By Prop. 7.8.3, this map restricts to a linear bijection $\mathcal{F}_0 : \mathcal{S}_0(\mathbb{R}^N) \to \mathcal{S}_0(\mathbb{R}^N)$. Moreover, $\mathcal{F}_0$ is unitary, since Prop. 7.8.3 implies $\overline{\widehat{f}} = \check{\overline{f}}$ and $\widehat{\check{f}} = f$ for each $f \in \mathcal{S}_0(\mathbb{R}^N)$, and hence

$$\int \overline{\widehat{f}}\widehat{f} d\mathfrak{m} = \int \check{\overline{f}}\widehat{f} d\mathfrak{m} \xlongequal{\text{Prop.7.4.9}} \int \widehat{\check{\overline{f}}} f d\mathfrak{m} = \int \overline{f} f d\mathfrak{m}$$

To extend $\mathcal{F}$ to a bounded linear operator on $L^2(\mathbb{R}^N, \mathfrak{m})$, by Thm. 2.4.2, it suffices to prove that $\mathcal{F} : \mathcal{S}(\mathbb{R}^N) \to \mathcal{S}(\mathbb{R}^N)$ is bounded.

Recall from Cor. 7.7.7 that $\mathcal{S}_0(\mathbb{R}^N)$ is dense in $L^2(\mathbb{R}^N, \mathfrak{m})$. For any $f \in \mathcal{S}(\mathbb{R}^N)$ and $g \in \mathcal{S}_0(\mathbb{R}^N)$, since $\|\widehat{g}\|_{L^2} = \|g\|_{L^2}$ (because $\mathcal{F}_0$ is unitary), we have

$$\Big|\int \widehat{f} g d\mathfrak{m}\Big| \xlongequal{\text{Prop.7.4.9}} \Big|\int f\widehat{g} d\mathfrak{m}\Big| \leqslant \|f\|_{L^2} \cdot \|\widehat{g}\|_{L^2} = \|f\|_{L^2} \cdot \|g\|_{L^2}$$

Thus, the linear functional on $L^2(\mathbb{R}^N, \mathfrak{m})$ defined by integration against $\widehat{f}$ has operator norm at most $\|f\|_{L^2}$ when restricted to the dense subspace $\mathcal{S}_0(\mathbb{R}^N)$, and hence on all of $L^2(\mathbb{R}^N, \mathfrak{m})$. This implies $\|\widehat{f}\|_{L^2} \leqslant \|f\|_{L^2}$, so $\mathcal{F} : \mathcal{S}(\mathbb{R}^N) \to \mathcal{S}(\mathbb{R}^N)$ is bounded.

Step 2. We have proved the existence of the map (7.29) and the fact that (7.29) restricts to a unitary operator on $\mathcal{S}_0(\mathbb{R}^N)$. In this step, we prove the remaining properties.

Since $\mathcal{S}_0(\mathbb{R}^N)$ is dense, the map (7.29) must be isometric and surjective (due to Lem. 5.10.21), and hence is unitary. Prop. 7.8.3 also shows that the relation $\mathcal{F}^2 = \Theta$ and $\mathcal{F}^4 = \mathbf{1}$ hold when restricted to the dense subspace $\mathcal{S}_0(\mathbb{R}^N)$. So they hold on $L^2(\mathbb{R}^N, \mathfrak{m})$.

The map (7.29) clearly restricts to a linear isometry

$$\mathcal{F} : \mathcal{S}(\mathbb{R}^N) \to \mathcal{S}(\mathbb{R}^N) \tag{7.30}$$

It remains to show that (7.30) is surjective (and hence unitary), and that its inverse is given by $f \mapsto \check{f}$. To prove this, note that similar to what we have proved, there is a linear isometry

$$\mathcal{G} : \mathcal{S}(\mathbb{R}^N) \to \mathcal{S}(\mathbb{R}^N) \qquad f \mapsto \check{f}$$

By Prop. 7.8.3, the relation $\mathcal{F}\mathcal{G} = \mathbf{1}$ holds when restricted to the dense subspace $\mathcal{S}_0(\mathbb{R}^N)$. So it holds on $\mathcal{S}(\mathbb{R}^N)$. This proves that (7.30) is surjective, and that $\mathcal{G}$ is its inverse. $\square$

We are ready to establish the canonical spectral decomposition for the regular representation U of $\mathbb{R}^N$. Recall from Exp. 7.3.7 that $(U(t)f)(x) = f(x - t)$.

Theorem 7.8.5. *Let $\mathcal{F} : L^2(\mathbb{R}^N, \mathfrak{m}) \xrightarrow{\simeq} L^2(\mathbb{R}^N, \mathfrak{m})$ be the Fourier transform. Let $U : \mathbb{R}^N \to \mathfrak{L}(L^2(\mathbb{R}^N, \mathfrak{m}))$ be the regular representation of $\mathbb{R}^N$. Then for each $t \in \mathbb{R}^N$ we have*

$$\mathcal{F}U(t)\mathcal{F}^{-1} = \mathbf{M}_{e_{-t}}$$

Moreover, if T_j denotes the closure of the operator $-\mathbf{i}\frac{\partial}{\partial x_j}$ with domain $\mathscr{D}(-\mathbf{i}\frac{\partial}{\partial x_j}) = C_c^\infty(\mathbb{R}^N)$, then

$$\mathcal{F}T_j\mathcal{F}^{-1} = \mathbf{M}_{x_j}$$

as unbounded operators on $L^2(\mathbb{R}^N, \mathfrak{m})$.

Proof. By Prop. 7.4.6, the relation $\mathcal{F}U(t)\mathcal{F}^{-1} = \mathbf{M}_{e_{-t}}$ holds on $\mathcal{S}(\mathbb{R}^N)$, and hence on $L^2(\mathbb{R}^N, \mathfrak{m})$. By Cor. 7.6.7, $-T_1, \dots, -T_N$ are the generators of U. By Thm. 6.10.2, $-\mathbf{M}_{x_1}, \dots, -\mathbf{M}_{x_N}$ are the generators of $t \mapsto \mathbf{M}_{e_{-t}}$. Therefore the relation $\mathcal{F}T_j\mathcal{F}^{-1} = \mathbf{M}_{x_j}$ holds. □

Exercise 7.8.6. Let $T_1, \dots, T_N$ be as in Thm. 7.8.5. Prove that $\mathcal{S}(\mathbb{R}^N) \subset \mathscr{D}(T_j)$, and that $T_j f = -\mathbf{i}\partial_{x_j} f$ for each $f \in \mathcal{S}(\mathbb{R}^N)$.

7.9 Perturbation of self-adjoint operators

Let $\mathcal{H}, \mathcal{K}, \mathcal{M}$ be Hilbert spaces.

7.9.1 Introduction

Consider the setting of quantum mechanics, where the Hamiltonian is given by $H = -\Delta + V$ on $L^2(\mathbb{R}^N, m)$ with domain $\mathscr{D}(H) = C_c^\infty(\mathbb{R}^N)$, and $V : \mathbb{R}^N \to \mathbb{R}$ is a potential function. By Thm. 6.11.1, H admits at least one self-adjoint extension $\widehat{H}$. Any such extension $\widehat{H}$ satisfies the spectral theorem and therefore yields the unitary evolution $e^{-\mathbf{i}t\widehat{H}}$ as the solution to the Schrödinger equation. However, there is no canonical self-adjoint extension, and thus no canonical solution to the Schrödinger equation. One must therefore identify the appropriate extension according to the physical context.

The good news is that a physically meaningful Hamiltonian must be lower bounded. In this case, by the Friedrichs extension (cf. Def. 7.2.15), H possesses a canonical lower-bounded self-adjoint extension H_{F}. Nevertheless, determining the exact domain and cores of H_{F} is often difficult, unless if one can show that H, with its original domain $C_c^\infty(\mathbb{R}^N)$, is essentially self-adjoint. If this holds, then by Cor. 6.7.10, the closure $\overline{H}$ is the unique self-adjoint extension of H; in particular $H_{\mathrm{F}} = \overline{H}$.

In this section, we develop general methods for determining whether H is essentially self-adjoint. Moreover, these methods will show that in favorable cases we indeed have $\mathscr{D}(\overline{H}) = \mathscr{D}(-\overline{\Delta})$, noting that $\mathscr{D}(-\overline{\Delta})$ can be described via the Fourier transform (cf. Rem. 7.9.12).

The basic idea is that in well-behaved cases (e.g., the hydrogen atom Hamiltonian; see Exp. 7.9.13), the potential V (or more precisely, its multiplication operator $\mathbf{M}_V$) is relatively small with respect to $-\overline{\Delta}$. The rigorous meaning of "relatively small" will be made precise later in this section. In such cases, $-\overline{\Delta} + V$ can be viewed as a perturbation of $-\overline{\Delta}$. Using the fact that $\pm y\mathbf{i} - \overline{\Delta}$ are surjective for all $y \in \mathbb{R}_{>0}$, one shows that $\pm\mathbf{i}y - \overline{\Delta} + V$ are surjective when y is sufficiently large. It then follows from Cor. 6.7.8 that $-\overline{\Delta} + V$ is self-adjoint.

Thus, the key to proving the self-adjointness of the Hamiltonian is to establish that, for each $g \in L^2(\mathbb{R}^N, m)$, the partial differential equation

$$(\pm\mathbf{i}y - \overline{\Delta} + V)f = g$$

has a solution. A similar situation arises in the theory of differential equations (though unrelated to quantum mechanics): given an open set $\Omega \subset \mathbb{R}^N$ and functions $a_1, \dots, a_N, b \in L^\infty(\Omega, m)$, one seeks, for each $g \in L^2(\Omega, m)$, a solution

$$\Big(\lambda - \Delta + \sum_{j=1}^N a_j \partial_{x_j} + b\Big) f = g$$

where $\lambda \in \mathbb{R}_{\geqslant 0}$, subject to either the Dirichlet boundary condition $f|_{\partial\Omega} = 0$ or the Neumann boundary condition $\partial_\nu f|_{\partial\Omega} = 0$. The same techniques used to establish the self-adjointness of Hamiltonians show that "weak solutions" exist for sufficiently large λ. In Subsec. 8.9.2, we will see that under slightly stronger assumptions, such solutions exist even for $\lambda = 0$.

In the following subsections, we first present a general result on the existence of (weak) solutions under perturbations of unbounded operators (Thm. 7.9.4). We then introduce the Kato-Rellich theorem (Thm. 7.9.6), which can be interpreted both as a self-adjointness criterion and as a statement on the existence of weak solutions to differential equations on $\mathbb{R}^N$. We illustrate how this theorem applies to proving the self-adjointness of the hydrogen atom Hamiltonian (with core $C_c^\infty(\mathbb{R}^3)$). Finally, we state Thm. 7.9.14, which, in the same spirit as the Kato-Rellich theorem, applies to differential equations on open subsets $\Omega \subset \mathbb{R}^N$, not necessarily the entire Euclidean space $\mathbb{R}^N$. The lower-boundedness of the hydrogen atom Hamiltonian will also be addressed with the help of Thm. 7.9.14.

7.9.2 A-boundedness

Definition 7.9.1. Let $A : \mathcal{H} \to \mathcal{K}$ and $B : \mathcal{H} \to \mathcal{M}$ be unbounded operators. We say that B is ***A*-bounded** if the following conditions hold:

(a) $\mathscr{D}(A) \subset \mathscr{D}(B)$.

(b) There exist $\alpha, \beta \in \mathbb{R}_{\geqslant 0}$ such that for each $\xi \in \mathscr{D}(A)$ we have

$$\|B\xi\| \leqslant \alpha\|A\xi\| + \beta\|\xi\| \tag{7.31}$$

The number α is called an ***A*-bound** (or simply a **bound**). If (a) and (b) hold with $\beta = 0$, we say that B is **strictly *A*-bounded**.

Note that condition (a) is redundant if we set $\|B\xi\| = +\infty$ (resp. $\|A\xi\| = +\infty$) when $\xi \in \mathcal{H}\backslash\mathscr{D}(B)$ (resp. $\xi \in \mathcal{H}\backslash\mathscr{D}(A)$).

Remark 7.9.2. Let $A : \mathcal{H} \to \mathcal{K}$ and $B : \mathcal{H} \to \mathcal{M}$ be unbounded operators satisfying $\mathscr{D}(A) \subset \mathscr{D}(B)$. Let $a, b \in \mathbb{R}_{\geqslant 0}$. Suppose that each $\xi \in \mathscr{D}(A)$ satisfies

$$\|B\xi\|^2 \leqslant a^2\|A\xi\|^2 + b^2\|\xi\|^2 \tag{7.32}$$

Then ξ clearly satisfies (7.31) with $\alpha = a$ and $\beta = b$.

Conversely, suppose that each $\xi \in \mathscr{D}(A)$ satisfies (7.31). Then for each $a > \alpha$ there exists $b \in \mathbb{R}_{\geqslant 0}$ such that each $\xi \in \mathscr{D}(A)$ satisfies (7.32). $\square$

Proof. The first paragraph is clear. Let us prove the second paragraph. Assume that (7.31) holds for all $\xi \in \mathscr{D}(A)$. Then for each $\varepsilon > 0$, we have $2\|A\xi\| \cdot \|\xi\| = 2\varepsilon\|A\xi\| \cdot \varepsilon^{-1}\|\xi\| \leqslant \varepsilon^2\|A\xi\|^2 + \varepsilon^{-2}\|\xi\|^2$, and hence

$$\|B\xi\|^2 \leqslant \alpha^2\|A\xi\|^2 + 2\alpha\beta\|A\xi\| \cdot \|\xi\| + \beta^2\|\xi\|^2 \leqslant (\alpha^2 + \varepsilon^2)\|A\xi\|^2 + (\beta^2 + \varepsilon^{-2})\|\xi\|^2$$

The second paragraph follows easily. $\square$

Very often, to verify (7.31) for all $\xi \in \mathscr{D}(A)$, it suffices to verify the inequality (7.31) for all ξ in a core for A:

Proposition 7.9.3. *Let $A : \mathcal{H} \to \mathcal{K}$ and $B : \mathcal{H} \to \mathcal{M}$ be unbounded operators where B is closed. Let $\alpha, \beta \in \mathbb{R}_{\geqslant 0}$. Suppose that $\mathscr{D}_0 \subset \mathscr{D}(A) \cap \mathscr{D}(B)$ is a core for A, and that inequality (7.31) holds for all $\xi \in \mathscr{D}_0$. Then conditions (a) and (b) of Def. 7.9.1 are satisfied.*

Proof. Choose any $\xi \in \mathscr{D}(A)$. Since $\mathscr{D}_0$ is a core for A, there is a sequence (ξ_n) in $\mathscr{D}_0$ converging to ξ such that $(A\xi_n)$ converges to $A\xi$. Since (7.31) holds for $\xi_m - \xi_n$, the sequence $(B\xi_n)$ is Cauchy, and hence converges to some ψ. Since B is closed and $\xi_n \to \xi$, we must have $\xi \in \mathscr{D}(B)$ (and hence $\mathscr{D}(A) \subset \mathscr{D}(B)$) and $\psi = B\xi$. Therefore $B\xi = \lim_n B\xi_n$, and hence

$$\|B\xi\| = \lim_n \|B\xi_n\| \leqslant \lim_n(\alpha\|A\xi_n\| + \beta\|\xi_n\|) = \alpha\|A\xi\| + \beta\|\xi\|$$

$\square$

7.9.3 Existence of solutions to linear equations under perturbation

Recall from Def. 6.2.24 that $A^{-1} \in \mathfrak{L}(\mathcal{H})$ means that A has an everywhere-defined bounded inverse.

Theorem 7.9.4. *Let A, B be unbounded operators on $\mathcal{H}$. Assume that $A^{-1} \in \mathfrak{L}(\mathcal{H})$. Assume also that B is strictly A-bounded with bound $0 \leqslant \alpha < 1$. Then $(A+B)^{-1} \in \mathfrak{L}(\mathcal{H})$.*

Proof. Recall that $\mathrm{Rng}(A^{-1}) = \mathscr{D}(A)$. Therefore, condition (7.31) (with $\beta = 0$) can be equivalently described by

$$\|BA^{-1}\psi\| \leqslant \alpha\|\psi\|$$

for each $\psi \in \mathcal{H}$. Thus BA^{-1} has operator norm no more than α. In particular $BA^{-1} \in \mathfrak{L}(\mathcal{H})$. Since $0 \leqslant \alpha < 1$, by Cor. 3.5.16, $1 + BA^{-1}$ invertible in $\mathfrak{L}(\mathcal{H})$.

By the distributive law (Prop. 6.2.8), we have

$$(1 + BA^{-1})A = A + B\mathbf{1}_{\mathscr{D}(A)} = A + B$$

where the last identity is due to $\mathscr{D}(A) \subset \mathscr{D}(B)$. Therefore, Cor. 6.2.21 implies that $A + B$ is injective, and

$$(A + B)^{-1} = A^{-1}(1 + BA^{-1})^{-1}$$

as n.d.d. unbounded operators. Since both A^{-1} and $(1 + BA^{-1})^{-1}$ belong to $\mathfrak{L}(\mathcal{H})$, we conclude that $(A + B)^{-1} \in \mathfrak{L}(\mathcal{H})$. □

7.9.4 Self-adjointness under perturbation

Proposition 7.9.5. *Let A, B be unbounded operators on $\mathcal{H}$. Assume that A is self-adjoint, and that B is A-bounded with bound $0 \leqslant \alpha < 1$. Then there exists $y_0 \in \mathbb{R}_{\geqslant 0}$ such that for each $y \geqslant y_0$, we have $(y\mathbf{i} + A + B)^{-1} \in \mathfrak{L}(\mathcal{H})$ and $(-y\mathbf{i} + A + B)^{-1} \in \mathfrak{L}(\mathcal{H})$.*

Proof. Let $y > 0$. By (6.12), we have

$$\|(A \pm y\mathbf{i})\xi\|^2 = \|A\xi\|^2 + y^2\|\xi\|^2 \tag{7.33}$$

By Rem. 7.9.2, there exist $0 \leqslant a < 1$ and $b \geqslant 0$ such that $\|B\xi\|^2 \leqslant a^2\|A\xi\|^2 + b^2\|\xi\|^2$ for each $\xi \in \mathscr{D}(A)$. Combining these two facts together, we see that for sufficiently large y, B is strictly $(A \pm y\mathbf{i})$-bounded with bound < 1. Since $(A \pm y\mathbf{i})^{-1} \in \mathfrak{L}(\mathcal{H})$ (due to $\mathrm{Sp}(A) \subset \mathbb{R}$ and Rem. 6.10.13), it follows from Thm. 7.9.4 that for sufficiently large y we have $(A \pm y\mathbf{i} + B)^{-1} \in \mathfrak{L}(\mathcal{H})$. □

Theorem 7.9.6 (Kato-Rellich theorem). *Let A be a self-adjoint operator on $\mathcal{H}$, and let B be a Hermitian operator on $\mathcal{H}$. Suppose that B is A-bounded with bound $0 \leqslant \alpha < 1$. Then $A + B$ is a self-adjoint operator on $\mathcal{H}$ having the same cores as A.*

Note that since $\mathscr{D}(A) \subset \mathscr{D}(B)$, by Def. 6.2.4 we have $\mathscr{D}(A) = \mathscr{D}(A + B)$.

Proof. For each $y \in \mathbb{R}_{>0}$, observe that $A + B$ is self-adjoint iff $y^{-1}(A + B)$ is self-adjoint. Therefore, by Cor. 6.7.8, it suffices to show that there exists $y \in \mathbb{R}_{>0}$ such that $\mathrm{Rng}(A \pm y\mathbf{i} + B) = \mathcal{H}$. But this follows from Prop. 7.9.5. That A and $A + B$ share the same cores follows from the following Prop. 7.9.7. □

Proposition 7.9.7. *Let $A, B : \mathcal{H} \to \mathcal{K}$ be unbounded operators. Assume that B is A-bounded with bound $0 \leqslant \alpha < 1$. Then the graph inner products on $\mathscr{D}(A) = \mathscr{D}(A + B)$ defined by A and $A + B$ are equivalent. Consequently, A and $A + B$ have the same cores, and (by Def. 6.4.5-(2)) A is closable iff $A + B$ is closable.*

Proof. Since (7.31) holds with $0 \leqslant \alpha < 1$, for each $\xi \in \mathscr{D}(A)$ we have that

$$\|(A+B)\xi\| \leqslant (1+\alpha)\|A\xi\| + \beta\|\xi\|$$

and that

$$\|(A+B)\xi\| \geqslant \|A\xi\| - \|B\xi\| \geqslant (1-\alpha)\|A\xi\| - \beta\|\xi\|$$

from which one deduces $\|A\xi\| \leqslant (1-\alpha)^{-1}(\|(A+B)\xi\| + \beta\|\xi\|)$. □

Corollary 7.9.8 (Kato-Rellich). *Let A be an essentially self-adjoint operator on $\mathcal{H}$, and let B be a Hermitian operator on $\mathcal{H}$. Suppose that B is A-bounded with bound $0 \leqslant \alpha < 1$. Then $A+B$ is essentially self-adjoint.*

Proof. Note that Hermitian operators are closable (Thm. 6.4.16). Since $\mathscr{D}(A) \subset \mathscr{D}(B)$, we have $\mathscr{D}(A) \subset \mathscr{D}(\overline{A}) \cap \mathscr{D}(\overline{B})$, and $\mathscr{D}(A)$ is a core for A. Therefore, by Prop. 7.9.3, $\overline{B}$ is $\overline{A}$-bounded with bound α. Thus, by Thm. 7.9.6, $\overline{A}+\overline{B}$ is a self-adjoint operator with core $\mathscr{D}(A)$. It follows that $A+B$, being the restriction of $\overline{A}+\overline{B}$ to $\mathscr{D}(A)$, is essentially self-adjoint. □

7.9.5 Application to quantum mechanics

To show that the hydrogen atom Hamiltonian satisfies the assumption in the Kato-Rellich theorem, we need to following estimate.

Lemma 7.9.9 (Sobolev inequality). *Assume that $N \in \{1,2,3\}$. Then for each $\alpha \in \mathbb{R}_{>0}$, there exists $\beta \in \mathbb{R}_{>0}$ such that for each $f \in \mathcal{S}(\mathbb{R}^N)$ we have*

$$\|f\|_{l^\infty} \leqslant \alpha\|\Delta f\|_{L^2(\mathbb{R}^N,m)} + \beta\|f\|_{L^2(\mathbb{R}^N,m)} \tag{7.34}$$

Proof. It suffices to prove this with f replaced by its Fourier transform $\widehat{f}$. Therefore, by Inequality (7.19), Prop. 7.8.2, and Thm. 7.8.4, it suffices to prove that for each $\alpha > 0$ there exists $\beta > 0$ such that

$$\|f\|_{L^1} \leqslant \alpha\||x|^2 f\|_{L^2} + \beta\|f\|_{L^2}$$

Since $N \leqslant 3$, we have $\int_1^{+\infty} t^{N-5}dt < +\infty$, and hence the number

$$C_r = \int_{\mathbb{R}^N} (r+|x|^4)^{-1}dx$$

is finite for each $r > 0$. Therefore

$$\int |f| = \int (r+|x|^4)^{-\frac{1}{2}}(r+|x|^4)^{\frac{1}{2}}|f| \leqslant \sqrt{C_r}\sqrt{\int (r+|x|^4)|f|^2}$$

Since $\sqrt{a+b} \leqslant \sqrt{a}+\sqrt{b}$ for each $a,b \geqslant 0$, we have

$$\int |f| \leqslant \sqrt{C_r}\cdot\||x|^2 f\|_{L^2} + \sqrt{rC_r}\cdot\|f\|_{L^2}$$

The proof is finished by noting that $\lim_{r\to+\infty} C_r = 0$ due to DCT. □

Exercise 7.9.10. Let $N \in \mathbb{Z}_+ \backslash \{4\}$. Suppose that there exist $\alpha_0, \beta_0 \in \mathbb{R}_{>0}$ such that for every $f \in \mathcal{S}(\mathbb{R}^N)$, inequality (7.34) holds with α, β replaced by α_0, β_0. Prove directly that for any $\alpha \in \mathbb{R}_{>0}$, there exists $\beta \in \mathbb{R}_{>0}$ such that for every $f \in \mathcal{S}(\mathbb{R}^N)$, inequality (7.34) holds.

Hint. Consider $f_\lambda(x) = f(\lambda x)$ where $\lambda > 0$. □

Theorem 7.9.11. *Let $N \in \{1, 2, 3\}$. Let $V \in L^\infty(\mathbb{R}^N, m) + L^2(\mathbb{R}^N, m)$ be real-valued, that is, $V = V_1 + V_2$ where $V_1 \in L^\infty(\mathbb{R}^N, m)$ and $V_2 \in L^2(\mathbb{R}^N, m)$ are both real-valued. Let Δ be the Laplacian operator on $L^2(\mathbb{R}^N, m)$ with $\mathscr{D}(\Delta) = C_c^\infty(\mathbb{R}^N)$. Then $\mathscr{D}(\overline{\Delta}) \subset \mathscr{D}(\mathbf{M}_V)$, and*

$$T := -\overline{\Delta} + \mathbf{M}_V$$

is self-adjoint with core $C_c^\infty(\mathbb{R}^N)$.

Proof. Recall from Cor. 7.6.9 that $\overline{\Delta}$ is self-adjoint. Write $V = V_1 + V_2$ where $V_1 \in L^\infty$ and $V_2 \in L^2$ are real-valued. Then $\mathbf{M}_{V_1}$ is Δ-bounded with bound 0. By Lem. 7.9.9, for each $\alpha > 0$ there exists $\beta > 0$ such that for any $f \in C_c^\infty(\mathbb{R}^N)$ we have

$$\|V_2 f\|_{L^2} = \|V_2\|_{L^2}\|f\|_{L^\infty} \leqslant \|V_2\|_{L^2}(\alpha\|\Delta f\|_{L^2} + \beta\|f\|_{L^2})$$

Thus, $\mathbf{M}_{V_2}$ is Δ-bounded with arbitrarily small bound. The same is true for $\mathbf{M}_V$. Hence, by Prop. 7.9.3, $\mathbf{M}_V$ is $\overline{\Delta}$-bounded with arbitrarily small bound. In particular, $\mathscr{D}(\overline{\Delta}) \subset \mathscr{D}(\mathbf{M}_V)$. By the Kato-Rellich Thm 7.9.6, T is self-adjoint and shares the same cores with $\overline{\Delta}$. In particular, $C_c^\infty(\mathbb{R}^N)$ is a core for T. □

The domain $\mathscr{D}(\overline{\Delta})$ can be described by the Fourier transform:

Remark 7.9.12. By Cor. 7.6.8, the operator $\overline{\Delta}$ is equal to the Borel functional calculus $f(T_\bullet)$ where $f(x) = -|x|^2$ and T_j is the closure of $-\mathbf{i}\frac{\partial}{\partial x_j}$ with $\mathscr{D}(-\mathbf{i}\frac{\partial}{\partial x_j}) = C_c^\infty(\mathbb{R}^N)$. Therefore, Thm. 7.8.5 implies

$$\mathcal{F}\overline{\Delta}\mathcal{F}^{-1} = \mathbf{M}_{-|x|^2}$$

where $\mathcal{F} : L^2(\mathbb{R}^N, \mathfrak{m}) \to L^2(\mathbb{R}^N, \mathfrak{m})$ is the Fourier transform and $d\mathfrak{m} = (2\pi)^{-N/2}dm$. In particular,

$$\mathscr{D}(\overline{\Delta}) = \mathcal{F}^{-1}\mathscr{D}(\mathbf{M}_{-|x|^2}) = \{f \in L^2(\mathbb{R}^N, m) : |x|^2\mathcal{F}f \in L^2(\mathbb{R}^N, m)\}$$

Example 7.9.13. On $\mathbb{R}^3$, the function $V(x) = -e^2/|x|$ (where $e \in \mathbb{R}$) belongs to $L^\infty + L^2$, since $V\chi_{\{|x|\geqslant 1\}}$ belongs to L^∞ and $V\chi_{|x|<1}$ belongs to L^2. Therefore, by Thm. 7.9.11, we have $\mathscr{D}(\overline{\Delta}) \subset \mathscr{D}(\mathbf{M}_{e^2/|x|})$,

$$T := -\overline{\Delta} - \mathbf{M}_{e^2/|x|}$$

is a self-adjoint operator with domain $\mathscr{D}(T) = \mathscr{D}(\overline{\Delta})$, and $C_c^\infty(\mathbb{R}^N)$ is a core for T. The operator T is known as the **hydrogen atom Hamiltonian** with electron charge e.

7.9.6 Lower-boundedness under perturbation

The idea underlying the proof of Prop. 7.9.5 leads directly to Thm. 7.9.14, a powerful result demonstrating that many self-adjoint Hamiltonian operators in quantum mechanics are lower-bounded. In Subsec. 7.9.7, we will use Thm. 7.9.14 to establish Thm. 7.9.17 and 7.9.18 on the existence of weak solutions to certain second-order linear differential equations.

Our treatment of partial differential equations in this course is operator-theoretic in nature: we work with linear operators and their algebraic relations rather than with sesquilinear forms, which are more common in the PDE literature. For comparison, see Subsec. 6.2.1 and 6.2.2 of [Eva], where Thm. 7.9.17 is proved via the method of sesquilinear forms, using the Lax-Milgram theorem in place of Thm. 7.9.4.

Theorem 7.9.14. *Let A, B be unbounded operators on $\mathcal{H}$. Assume that A is positive and self-adjoint, and that B is A-bounded with bound $0 \leqslant \alpha < 1$. Then there exists $\lambda_0 \in \mathbb{R}_{\geqslant 0}$ such that for each $\lambda \geqslant \lambda_0$, we have $(\lambda + A + B)^{-1} \in \mathfrak{L}(\mathcal{H})$.*

Proof. Let $\lambda > 0$. For each $\xi \in \mathscr{D}(A)$, we have

$$\|(\lambda + A)\xi\|^2 = \|A\xi\|^2 + 2\langle\xi|A\xi\rangle + \lambda^2\|\xi\|^2 \geqslant \|A\xi\|^2 + \lambda^2\|\xi\|^2$$

By Rem. 7.9.2, there exist $0 \leqslant a < 1$ and $b \geqslant 0$ such that $\|B\xi\|^2 \leqslant a^2\|A\xi\|^2 + b^2\|\xi\|^2$ for each $\xi \in \mathscr{D}(A)$. Therefore, B is strictly $(\lambda + A)$-bounded with bound < 1 for sufficiently large λ. Since $(\lambda + A)^{-1} \in \mathfrak{L}(\mathcal{H})$ (due to $\mathrm{Sp}(A) \subset \mathbb{R}_{\geqslant 0}$ and Rem. 6.10.13), it follows from Thm. 7.9.4 that we have $(\lambda + A + B)^{-1} \in \mathfrak{L}(\mathcal{H})$ for sufficiently large λ. □

Corollary 7.9.15. *In Thm. 7.9.14, assume moreover that B is Hermitian. Then $A + B$ is a lower-bounded self-adjoint operator.*

Consequently, the self-adjoint operator T in Thm. 7.9.11 is lower-bounded. In particular, the hydrogen atom Hamiltonian is lower-bounded.

Proof. By the Kato-Rellich Thm. 7.9.6, $T := A + B$ is self-adjoint. By Thm. 7.9.14 and 6.12.5, there exists $\lambda_0 \geqslant 0$ such that $-\lambda \notin \mathrm{Sp}(T)$ for each $\lambda > \lambda_0$. Thus $\mathrm{Sp}(\lambda_0 + T)$ is contained in $\mathbb{R}_{\geqslant 0}$. Hence $\lambda_0 + T$ is positive by Prop. 7.1.2. □

In Subsec. 7.9.7, we shall apply Thm. 7.9.14 to the case that A is the negative Laplacian and B is a first-order differential operator. To establish the A-bounds in this situation, we need the following proposition.

Proposition 7.9.16. *Let A, B be unbounded operators on $\mathcal{H}$. Assume that A is positive and self-adjoint, and that B is $\sqrt{A}$-bounded. Then for each $\alpha \in \mathbb{R}_{>0}$, B is A-bounded with bound α.*

Proof. It suffices to show that $\sqrt{A}$ is A-bounded with arbitrarily small bound. For each $\xi \in \mathscr{D}(A)$, we compute that

$$\|\sqrt{A}\xi\|^2 = \langle \xi|A\xi\rangle \leqslant \|A\xi\| \cdot \|\xi\| = \varepsilon\|A\xi\| \cdot \varepsilon^{-1}\|\xi\| \leqslant \frac{\varepsilon^2}{2}\|A\xi\|^2 + \frac{1}{2\varepsilon^2}\|\xi\|^2$$

for each $\varepsilon > 0$. Hence $\|\sqrt{A}\xi\| \leqslant \frac{\varepsilon}{\sqrt{2}}\|A\xi\| + \frac{1}{\sqrt{2}\varepsilon}\|\xi\|$, finishing the proof. □

7.9.7 Application to partial differential equations

Recall Def. 7.1.12 for the meanings of the Dirichlet Laplacian Δ_{D} and the Neumann Laplacian Δ_{N}.

Theorem 7.9.17. *Let $\Omega \subset \mathbb{R}^N$ be open. Let $a_1, \ldots, a_N, b \in L^\infty(\Omega, m)$. Let X_j be the closure of ∂_{x_j} with $\mathscr{D}(\partial_{x_j}) = C_c^\infty(\Omega)$.*[9] *Then $\mathscr{D}(\sqrt{-\Delta_{\mathrm{D}}}) \subset \mathscr{D}(X_j)$ (and hence $\mathscr{D}(\Delta_{\mathrm{D}}) \subset \mathscr{D}(X_j)$). Moreover, if we define*

$$T := -\Delta_{\mathrm{D}} + \sum_{j=1}^N \mathbf{M}_{a_j} X_j + \mathbf{M}_b$$

then $\mathscr{D}(T) = \mathscr{D}(\Delta_{\mathrm{D}})$, and there exists $\lambda_0 \in \mathbb{R}_{\geqslant 0}$ such that $(\lambda + T)^{-1} \in \mathfrak{L}(L^2(\Omega, m))$ for each $\lambda \geqslant \lambda_0$.

Proof. Since ∂_{x_j} is strictly ∇-bounded, by Prop. 7.9.3, X_j is strictly $\overline{\nabla}$-bounded. Since $-\Delta_{\mathrm{D}} = \overline{\nabla}^*\overline{\nabla}$, by the polar decomposition for $\overline{\nabla}$ (cf. Thm. 7.1.11), X_j is strictly $\sqrt{-\Delta_{\mathrm{D}}}$-bounded. In particular, we have $\mathscr{D}(\sqrt{-\Delta_{\mathrm{D}}}) \subset \mathscr{D}(X_j)$. Since $\mathscr{D}(\Delta_{\mathrm{D}}) \subset \mathscr{D}(\sqrt{-\Delta_{\mathrm{D}}})$, we conclude $\mathscr{D}(\Delta_{\mathrm{D}}) \subset \mathscr{D}(X_j)$, and hence $\mathscr{D}(T) = \mathscr{D}(\Delta_{\mathrm{D}})$.

By Prop. 7.9.16, X_j is $-\Delta_{\mathrm{D}}$-bounded with arbitrarily small bound. Thus $\sum_j \mathbf{M}_{a_j} X_j + \mathbf{M}_b$ is $-\Delta_{\mathrm{D}}$-bounded with arbitrarily small bound. Hence, Thm. 7.9.14 implies that $(\lambda + T)^{-1} \in \mathfrak{L}(L^2(\Omega, m))$ for large enough $\lambda \in \mathbb{R}_{\geqslant 0}$. □

Theorem 7.9.18. *In the setting of Thm. 7.9.17, we have $\mathscr{D}(\sqrt{-\Delta_{\mathrm{N}}}) \subset \mathscr{D}(X_j^*)$ (and hence $\mathscr{D}(\Delta_{\mathrm{N}}) \subset \mathscr{D}(X_j^*)$). Moreover, if we define*

$$S := -\Delta_{\mathrm{N}} - \sum_{j=1}^N \mathbf{M}_{a_j} X_j^* + \mathbf{M}_b$$

then $\mathscr{D}(S) = \mathscr{D}(\Delta_{\mathrm{N}})$, and there exists $\lambda_0 \in \mathbb{R}_{\geqslant 0}$ such that $(\lambda + S)^{-1} \in \mathfrak{L}(L^2(\Omega, m))$ for each $\lambda \geqslant \lambda_0$.

Proof. It is clear that every $f \in \mathscr{D}(\mathrm{div}^*)$ lies in $\mathscr{D}(X_j^*)$, and that $\mathrm{div}^* f = (X_1^* f, \ldots, X_N^* f)$. Therefore X_j^* is div^*-bounded. Since $-\Delta_{\mathrm{N}} = \overline{\mathrm{div}} \cdot \mathrm{div}^*$, we can apply the same reasoning as in Thm. 7.9.17 to establish the present theorem. □

[9]Note that $\mathrm{i}\partial_{x_j}$ is Hermitian, and hence is closable by Thm. 6.4.16.

Remark 7.9.19. The surjectivity of $\lambda+T$ in Thm. 7.9.17 (for sufficiently large $\lambda\geqslant 0$) can be interpreted as follows: for every $g\in L^2(\Omega,m)$, the differential equation

$$\Big(\lambda-\Delta+\sum_j a_j\partial_{x_j}+b\Big)f=g\qquad f|_{\partial\Omega}=0 \tag{7.35}$$

admits a "weak solution" $f\in\mathscr{D}(\Delta_{\mathrm{D}})$. In particular, each solution f belongs to $\mathscr{D}(\sqrt{-\Delta_{\mathrm{D}}})=H_0^1(\Omega)$ (cf. Rem. 7.1.13). Similarly, the surjectivity of $\lambda+S$ in Thm. 7.9.18 means that

$$\Big(\lambda-\Delta+\sum_j a_j\partial_{x_j}+b\Big)f=g\qquad \partial_\nu f|_{\partial\Omega}=0 \tag{7.36}$$

has a "weak solution" f in $H^1(\Omega)$. See Rem. 7.1.15 for a discussion of how the various Laplacians correspond to their respective boundary conditions.

7.10 Problems

Let $\mathcal{H},\mathcal{K}$ be Hilbert spaces.

Problem 7.1. Let T be a positive unbounded operator on $\mathcal{H}$. Let $\mathscr{D}_0$ be a linear subspace of $\mathscr{D}(T)$. Let $\lambda\in\mathbb{R}_{>0}$.

1. Prove that $\mathscr{D}_0$ is a core for T iff $(\lambda+T)\mathscr{D}_0$ is dense in $(\lambda+T)\mathscr{D}(T)$.

2. Prove that T is closed iff $(\lambda+T)\mathscr{D}(T)$ is closed in $\mathcal{H}$.

Hint. Show that the graph inner product of T on $\mathscr{D}(T)$ is equivalent (cf. Def. 7.2.13) to $(\xi,\eta)\mapsto\langle(\lambda+T)\xi|(\lambda+T)\eta\rangle$. □

Problem 7.2. Let A be a positive self-adjoint operator on $\mathcal{H}$. Let $n\in\mathbb{Z}_+$. Prove that there exists a unique positive self-adjoint operator B on $\mathcal{H}$ satisfying $B^n=A$.

Hint for the uniqueness. Apply the composition law (cf. Thm. 6.10.5) to the Borel functional calculus of B. □

Problem 7.3. Solve the following problems.

1. Let A be a positive self-adjoint operator on $\mathcal{H}$. Prove that if $0\leqslant\alpha\leqslant\beta<+\infty$, then $\mathscr{D}(A^\beta)\subset\mathscr{D}(A^\alpha)$, and any core for A^β is a core for A^α.

2. Let $T:\mathcal{H}\to\mathcal{K}$ be a closed operator. Prove that any core for T^*T is a core for T.

3. Let Δ_{D} be the Dirichlet Laplacian for the bounded open interval $I=(a,b)$. Show that $C_c^\infty(I)$ is a core for $\sqrt{-\Delta_{\mathrm{D}}}$, that $C_c^\infty(I)\subset\mathscr{D}(-\Delta_{\mathrm{D}})$, but that $C_c^\infty(I)$ is not a core for $-\Delta_{\mathrm{D}}$.

Hint. 1. See the hint of Pb. 6.3 for a related estimate.

3. Recall Exp. 7.1.16 for an explicit description of Δ_{D} for I. $\square$

Problem 7.4. Let $T : \mathcal{H} \to \mathcal{K}$ be an unbounded operator. Prove that the following are equivalent.

(1) T is closable.

(2) Equipping $\mathscr{D}(T)$ with the inner product of $\mathcal{H}$, the function

$$\mathscr{D}(T) \to [0, +\infty] \qquad \xi \mapsto \|T\xi\|$$

is lower semicontinuous. In other words (cf. Prop. 1.4.11), if (ξ_n) is a sequence/net in $\mathscr{D}(T)$ converging to $\xi \in \mathscr{D}(T)$, then $\|T\xi\| \leqslant \liminf_n \|T\xi_n\|$.

Hint. (1)⇒(2): Method 1: By the polar decomposition Thm. 7.1.11 and the spectral theorem, one may assume that T is a multiplication operator. Apply Thm. 1.6.14 and Fatou's lemma in measure theorey. Method 2: Assume WLOG that $\liminf_n \|T\xi_n\| < +\infty$. Use Banach-Alaoglu to choose a weak convergent subnet of $(T\xi_n)$. Then apply Fatou's lemma for weak convergence (Prop. 3.7.3). Where is the closability used in this method?

(2)⇒(1): Use Def. 6.4.5-(2) to show that T is closable: Let (ξ_n) be a sequence in $\mathscr{D}(T)$ converging to 0 such that $(T\xi_n)$ converges. Consider $\|T(\xi_m - \xi_n)\|$ for large enough m, n. Apply $\liminf_n$. $\square$

Definition 7.10.1. If A, B are positive self-adjoint operators on $\mathcal{H}$, we write $\boldsymbol{A \leqslant B}$ if $\mathscr{D}(\sqrt{B}) \subset \mathscr{D}(\sqrt{A})$ and $\langle\sqrt{A}\xi|\sqrt{A}\xi\rangle \leqslant \langle\sqrt{B}\xi|\sqrt{B}\xi\rangle$ for each $\xi \in \mathscr{D}(B)$. This definition can be simplified to

$$\langle\sqrt{A}\xi|\sqrt{A}\xi\rangle \leqslant \langle\sqrt{B}\xi|\sqrt{B}\xi\rangle \qquad \text{for each } \xi \in \mathscr{D}(\mathcal{H})$$

if for each closed operator $T : \mathcal{H} \to \mathcal{K}$ and $\xi \in \mathcal{H}$ we let

$$\|T\xi\| = +\infty \qquad \text{when } \xi \in \mathcal{H}\backslash\mathscr{D}(T)$$

Problem 7.5. Let A, B be unbounded positive operators on $\mathcal{H}$ with Friedrichs extensions $A_{\mathrm{F}}, B_{\mathrm{F}}$.

1. Assume that $\mathscr{D}(B) \subset \mathscr{D}(A)$ and $\langle\xi|A\xi\rangle \leqslant \langle\xi|B\xi\rangle$ for each $\xi \in \mathscr{D}(B)$. Prove that $A_{\mathrm{F}} \leqslant B_{\mathrm{F}}$.

2. Conclude that A_{F} is the largest positive self-adjoint extension of A with respect to the partial order $\leqslant$. (That is, if $A \subset \widehat{A}$ where $\widehat{A}$ is a positive self-adjoint operator, then $\widehat{A} \leqslant A_{\mathrm{F}}$.)

Problem 7.6. Let A be a strictly positive closed operator on $\mathcal{H}$. Solve the problem following the given hint.

1. Show that the inner product ω on $\mathscr{D}(A)$ defined by $\omega(\xi|\eta) = \langle\xi|A\eta\rangle + \langle A\xi|A\eta\rangle$ is equivalent to the graph inner product on $\mathscr{D}(A)$, and hence is complete.

2. Show that for each $\psi \in \mathcal{H}$ there exists $\xi \in \mathscr{D}(A)$ such that $\psi - \xi - A\xi \in \mathrm{Rng}(A)^\perp$.

3. Show that

$$A_{\mathrm{K}} := A^*|_{\mathscr{D}(A)+\mathrm{Ker}(A^*)} \tag{7.37}$$

is a positive self-adjoint operator on $\mathcal{H}$ extending A. We call A_{K} the **Krein-von Neumann extension** of A.

Hint. 2. Show that the linear functional $\eta \in \mathscr{D}(A) \to \langle\psi|A\eta\rangle$ is bounded in the inner product ω. Then apply the Riesz-Fréchet Thm. 3.5.3.
3. Show that A_{K} is positive and $\mathbf{1} + A_{\mathrm{K}}$ has range $\mathcal{H}$. □

Remark 7.10.2. Note that by Pb. 7.1, for any closed strictly-positive operator A the range $\mathrm{Rng}(A)$ is closed. The following problem provides a more operator-theoretic proof that the Krein-von Neumann extension A_{K} is self-adjoint, following von Neumann's original method in [vN29a]. The key idea is as follows:

Given a projection $P \in \mathfrak{L}(\mathcal{H})$ and an unbounded operator T on $\mathcal{H}$, the relation

$$PT \subset TP$$

captures precisely the situation in which T decomposes as the direct sum (cf. Def. 6.2.15) of $T|_{\mathrm{Rng}(P)}$ (with domain $P\mathscr{D}(T)$) and $T|_{\mathrm{Rng}(P)^\perp}$ (with domain $(\mathbf{1}-P)\mathscr{D}(T)$). See Rem. 7.5.5 for further discussion. □

Problem 7.7. Let A be a Hermitian operator on $\mathcal{H}$ with closed range $\mathrm{Rng}(A)$. The goal of this problem is to show that $A_{\mathrm{K}} := A^*|_{\mathscr{D}(A)+\mathrm{Ker}(A^*)}$ is a self-adjoint extension of A. Let $P \in \mathfrak{L}(\mathcal{H})$ be the projection onto $\mathrm{Rng}(A)$.

1. Show that A_{K} is Hermitian, and that

$$P^\perp A_{\mathrm{K}} \subset A_{\mathrm{K}} P^\perp$$

(That is, $P^\perp \mathscr{D}(A_{\mathrm{K}}) \subset \mathscr{D}(A_{\mathrm{K}})$, and $P^\perp A_{\mathrm{K}}\xi = A_{\mathrm{K}} P^\perp \xi$ for each $\xi \in \mathscr{D}(A_{\mathrm{K}})$.) Conclude from Prop. 6.2.8 that $PA_{\mathrm{K}} \subset A_{\mathrm{K}} P$.

2. Use $PA_{\mathrm{K}} \subset A_{\mathrm{K}} P$ to show that we have a (densely-defined) unbounded operator B on the Hilbert space $\mathrm{Rng}(A) = \mathrm{Rng}(P)$ by

$$B : \mathrm{Rng}(P) \to \mathrm{Rng}(P) \qquad \psi \mapsto A_{\mathrm{K}}\psi$$
$$\mathscr{D}(B) = P\mathscr{D}(A_{\mathrm{K}})$$

3. Show that A_{K} equals the direct sum operator (cf. Def. 6.2.15) of B on $\mathrm{Rng}(A)$ and 0 on $\mathrm{Rng}(A)^\perp$. (Note that for any linear subspace $V \subset \mathcal{H}$, we have $V = PV \oplus P^\perp V$ iff $PV \subset V$ iff $P^\perp V \subset V$.) Conclude that B is Hermitian and $\mathrm{Rng}(B) = \mathrm{Rng}(A)$.

Since B is a surjective Hermitian operator, by Lem. 7.10.3, B is self-adjoint. It follows from Pb. 6.10 that $A_{\mathrm{K}} = B \oplus 0$ is self-adjoint.

Lemma 7.10.3. *Let A be a Hermitian operator on $\mathcal{H}$. Suppose that $\mathrm{Rng}(A) = \mathcal{H}$. Then A is self-adjoint.*

Proof. Since $\mathrm{Rng}(A) = \mathcal{H}$, for each $\xi \in \mathscr{D}(A^*)$ there exists $\eta \in \mathscr{D}(A)$ such that $A^*\xi = A\eta$, and hence $\xi - \eta \in \mathrm{Ker}(A^*)$. Since $\mathrm{Ker}(A^*) = \mathrm{Rng}(A)^\perp = 0$ (Cor. 6.6.11), we conclude that $\xi = \eta$. Therefore $\mathscr{D}(A^*) \subset \mathscr{D}(A)$, and hence $A = A^*$. $\square$

Problem 7.8. Let T be an unbounded positive operator on $\mathcal{H}$ with deficiency indices n_+ and n_-. Assume that $\mathcal{H}$ is separable. Prove that

$$\dim(\mathrm{Rng}(T+1)^\perp) = n_+ = n_-$$

Hint. Assume WLOG that T is closed. Use Thm. 6.5.1 to show that any self-adjoint extension $\widehat{T}$ of T satisfies $\dim(\mathscr{D}(\widehat{T})/\mathscr{D}(T)) = n_+ = n_-$. Construct $\widehat{T}$ using the Krein-von Neumann extension. $\square$

The following result is also known as **von Neumann's mean ergodic theorem**.

Problem 7.9. Let $U : \mathbb{R} \to \mathfrak{L}(\mathcal{H})$ be a strongly-continuous unitary representation of $\mathbb{R}$ with generator T.

1. Let $\mathcal{K} = \{\xi \in \mathcal{H} : U(t)\xi = \xi \text{ for each } t \in \mathbb{R}\}$. Prove that $\mathcal{K} = \mathrm{Ker}T$.

2. For each $\lambda \in \mathbb{R}_{>0}$, define $f_\lambda : \mathbb{R} \to \mathbb{C}$ by

$$f_\lambda(x) = \begin{cases} 1 & \text{if } x = 0 \\ \dfrac{1}{\mathbf{i}\lambda x}(e^{\mathbf{i}\lambda x} - 1) & \text{if } x \neq 0 \end{cases}$$

   Prove for each $\xi \in \mathcal{H}$ that

$$f_\lambda(T)\xi = \frac{1}{\lambda}\int_0^\lambda U(t)\xi dt$$

3. Let $E \in \mathfrak{L}(\mathcal{H})$ be the projection onto $\mathcal{K}$. Prove the SOT convergence

$$\lim_{\lambda \to +\infty} \frac{1}{\lambda}\int_0^\lambda U(t)dt = E$$

   where $\int_0^\lambda U(t)dt$ denotes the linear map sending each $\xi \in \mathcal{H}$ to $\int_0^\lambda U(t)\xi dt$.

Recall that for each $t \in \mathbb{R}^N$, the function $e_t : \mathbb{R}^N \to \mathbb{C}$ is defined by $e_t(x) = e^{\mathbf{i}tx} = e^{\mathbf{i}(t_1x_1+\cdots+t_Nx_N)}$

Definition 7.10.4. For each finite Borel measure μ on $\mathbb{R}^N$, define the **Fourier transform** $\widehat{\mu} : \mathbb{R}^N \to \mathbb{C}$ by

$$\widehat{\mu}(t) = \int_{\mathbb{R}^N} e_{-t} d\mu$$

Clearly $\widehat{\mu}(t) \leqslant \mu(\mathbb{R}^N) < +\infty$, and $\widehat{\mu}$ is continuous by DCT. Thus $\widehat{\mu}$ is a bounded continuous function.

Definition 7.10.5. A function $\gamma : \mathbb{R}^N \to \mathbb{C}$ is called **positive-definite** if the matrix $\Gamma \in \mathbb{C}^{\mathbb{R}^N \times \mathbb{R}^N}$ defined by $\Gamma(s,t) = \gamma(s-t)$ (where $s,t \in \mathbb{R}^N$) is positive. In other words, for each $f : \mathbb{R}^N \to \mathbb{C}$ vanishing outside a finite set, we have

$$\sum_{s,t\in\mathbb{R}^N} \overline{f(s)}\gamma(s-t)f(t) \geqslant 0$$

The following result, known as the **Bochner theorem**, can be viewed as the continuous analogue of the trigonometric moment theorem presented in Pb. 5.6.

Problem 7.10. Let $\gamma : \mathbb{R}^N \to \mathbb{C}$. Prove that the following conditions are equivalent.

(1) There exists a finite Borel measure μ on $\mathbb{R}^N$ such that $\widehat{\mu} = \gamma$.

(2) The function γ is continuous at 0 and positive-definite.

Show that if γ satisfies either (1) or (2), then γ is continuous on $\mathbb{R}^N$.

Hint. Adapt the method used to solve Pb. 5.6 to the continuous setting. More precisely, define a linear functional Λ on $\mathscr{A} = \mathrm{Span}\{e_t : t \in \mathbb{R}^N\}$ by $\Lambda(e_t) = \gamma(-t)$. (You must justify that for distinct $t_1, \dots, t_n \in \mathbb{R}^N$, the functions $e_{t_1}, \dots, e_{t_n}$ are linearly independent. To see this, take $f \in \mathcal{S}(\mathbb{R}^N)$ whose Fourier transform vanishes at all but one of $-t_1, \dots, -t_n$.)

Apply the GNS construction to obtain a corresponding inner product space, take its Hilbert space completion, and construct a unitary representation U of $\mathbb{R}^N$ on this Hilbert space. To check the strong continuity of U, first verify continuity when U acts on each e_t. □

Problem 7.11. Let X be a measurable space. Let $\mathcal{H} = \bigoplus_{\alpha\in\mathscr{I}} L^2(X,\mu_\alpha)$ where each μ_α is a measure on $\mathfrak{M}$. Let $f : X \to \mathbb{C}$ be measurable. Let $(E_n)_{n\in\mathbb{Z}_+}$ be an increasing sequence of measurable subsets of X such that $X = \bigcup_n E_n$, and that $f|_{E_n}$ is bounded for each E_n. Prove that $\mathscr{D}_0$ is a core for M_f, where

$$\mathscr{D}_0 := \bigcup_n \mathrm{M}_{\chi_{E_n}}\mathcal{H} = \bigcup_n \bigoplus_{\alpha\in\mathscr{I}} L^2(E_n, \mu_\alpha)$$

Problem 7.12. Let $\mathcal{H} = L^2(\mathbb{R}^N, m)$. Choose a polynomial $f \in \mathbb{R}[x_1, \dots, x_N]$. Give direct solutions of the following problems without using Thm. 7.6.2.

1. Let $E_n = \overline{B}_{\mathbb{R}^N}(0, n)$ and $\mathscr{D}_0 = \bigcup_n \mathbf{M}_{\chi_{E_n}}\mathcal{H}$. Prove that $C_c^\infty(\mathbb{R}^N)$ is a core for $\mathbf{M}_f|_{\mathscr{D}_0}$. Conclude from Pb. 7.11 that $C_c^\infty(\mathbb{R}^N)$ is a core for $\mathbf{M}_f$.

2. For each $1 \leqslant j \leqslant N$, let $T_j = -\mathbf{i}\frac{\partial}{\partial x_j}$ with $\mathscr{D}(T_j) = \mathcal{S}(\mathbb{R}^N)$. Let $A = f(T_\bullet) = f(T_1, \dots, T_N)$ with $\mathscr{D}(A) = \mathcal{S}(\mathbb{R}^N)$, where $f(T_\bullet)$ is interpreted via the usual polynomial functional calculus of linear operators (rather than the Borel functional calculus). Use part 1 and the Fourier transform on $\mathcal{S}(\mathbb{R}^N)$ (i.e. Thm. 7.8.4 and Prop. 7.8.2) to show that A is essentially self-adjoint.

3. Show that $C_c^\infty(\mathbb{R}^N)$ is a core for A. (Equivalently, show that $A|_{C_c^\infty(\mathbb{R}^N)}$ is essentially self-adjoint.)

Hint for part 3. Choose any $\varphi \in C_c^\infty(\mathbb{R}^N)$ such that $0 \leqslant \varphi \leqslant 1$, and that $\varphi = 1$ on a neighborhood of 0. For each $\varepsilon > 0$, let $\varphi_\varepsilon(x) = \varphi(\varepsilon x)$. For each $f \in \mathcal{S}(\mathbb{R}^N)$, show that $\varphi_\varepsilon f \to f$ and $A(\varphi_\varepsilon f) \to Af$ as $\varepsilon \to 0$. □

Recall Def. 7.9.1 for the meaning of A-boundedness.

Problem 7.13. Let $-\infty < a < b < +\infty$ and $I = (a, b)$. Let $c \in I$. Let $\mathcal{H} := L^2(I, m)$. Let A, B be unbounded operators on $\mathcal{H}$ with $\mathscr{D}(A) = \mathscr{D}(B) = C_c^\infty(I)$ defined by

$$Af = f(c) \qquad Bf = f' + f(c)$$

1. Prove that A is not closable.

2. Prove that B is closable and $\mathscr{D}(\overline{B}) = H_0^1(I)$.

Hint for Part 2. Prove a Sobolev inequality for functions in $\mathcal{S}(\mathbb{R})$ analogous to Lem. 7.9.9, with Δ replaced by d/dx. Then apply Prop. 7.9.7. □

The following problem gives a sufficient condition for the uniqueness of solutions to the polynomial moment problem.

Problem 7.14. Let $(c_n)_{n\in\mathbb{N}}$ be a sequence in $\mathbb{R}$ satisfying

$$\sup_{n\in\mathbb{N}} \frac{\rho^{2n}}{(2n)!}|c_{2n}| < +\infty$$

for some $\rho > 0$. Prove that there exists at most one finite Borel measure μ on $\mathbb{R}$ satisfying for each $n \in \mathbb{N}$ the condition

$$\int_{\mathbb{R}} |x|^n d\mu < +\infty \qquad \int_{\mathbb{R}} x^n d\mu = c_n \tag{7.38}$$

Hint. Suppose μ_1, μ_2 both satisfy (7.38). Let $\mu = \mu_1 + \mu_2$. Use the Radon-Nikodym Thm. 1.6.12 to write $d\mu_j = f_j d\mu$ where $f_j : \mathbb{R} \to [0,1]$ is Borel. Show that $f := f_1 - f_2$ belongs to $L^2(\mathbb{R},\mu)$ and is orthogonal to polynomials. Prove that μ satisfies the assumption in Thm. 7.7.4, noting Exe. 7.7.5 and

$$\frac{(2n)!}{(n!)^2} = \binom{2n}{n} \leqslant \sum_{k=0}^{2n} \binom{2n}{k} = (1+1)^{2n} = 4^n \tag{7.39}$$

□

7.10.1 Analytic vectors

Definition 7.10.6. Let $T_1, \dots, T_N$ be unbounded operators on $\mathcal{H}$. Let

$$\boldsymbol{C^\infty(T_\bullet)} = \bigcap_{k \in \mathbb{Z}_+} \bigcap_{1 \leqslant i_1, \dots, i_k \leqslant N} \mathscr{D}(T_{i_1} \cdots T_{i_k})$$

Note that $T_j C^\infty(T_\bullet) \subset C^\infty(T_\bullet)$. Elements of $C^\infty(T_\bullet)$ are called **joint smooth vectors** of $T_\bullet$.

Definition 7.10.7. Let $T_1, \dots, T_N$ be strongly commuting self-adjoint operators on $\mathcal{H}$. Note that

$$T_i T_j = T_j T_i \qquad \text{on } C^\infty(T_\bullet)$$

by Pb. 6.11. An element $\xi \in C^\infty(T_\bullet)$ is called a **joint analytic vector** for $T_\bullet$ if there exists $r \in \mathbb{R}_{>0}$ such that

$$\sum_{n_1,\dots,n_N \in \mathbb{N}} \frac{r^{n_1+\cdots+n_N}}{n_1! \cdots n_N!} \|T_1^{n_1} \cdots T_N^{n_N} \xi\| < +\infty \tag{7.40a}$$

equivalently (cf. Exe. 7.7.5), if there exists $\rho \in \mathbb{R}_{>0}$ such that

$$\sup_{n_1,\dots,n_N \in \mathbb{N}} \frac{\rho^{n_1+\cdots+n_N}}{n_1! \cdots n_N!} \|T_1^{n_1} \cdots T_N^{n_N} \xi\| < +\infty \tag{7.40b}$$

The set of joint analytic vectors for $T_\bullet$ is denoted by $\boldsymbol{\mathfrak{A}(T_\bullet)}$.

Problem 7.15. Let $T_1, \dots, T_N$ be strongly commuting self-adjoint operators on $\mathcal{H}$.

1. Prove that $T_j \mathfrak{A}(T_\bullet) \subset \mathfrak{A}(T_\bullet)$ for each $1 \leqslant j \leqslant N$.

2. Let $\xi \in \mathcal{H}$. Let μ_ξ be the finite Borel measure on $\mathbb{R}^N$ associated to ξ and $T_\bullet$ (cf. Thm. 6.10.1). Prove that $\xi \in \mathfrak{A}(T_\bullet)$ iff μ_ξ satisfies the condition in Thm. 7.7.4, that is,

$$\int_{\mathbb{R}^N} e^{r(|x_1|+\cdots+|x_N|)} d\mu_\xi(x) < +\infty$$

Hint for part 2. Use Cauchy-Schwarz for one direction. Use (7.39) for the other direction. □

The following property is a variant of Prop. 7.5.4.

Problem 7.16. Let $T_1, \dots, T_N$ be strongly commuting self-adjoint operators on $\mathcal{H}$. Let $\mathscr{D}_0$ be a linear subspace of $\mathfrak{A}(T_\bullet)$ such that $T_j\mathscr{D}_0 \subset \mathscr{D}_0$ for all j. Prove that the closure $\overline{\mathscr{D}_0}$ satisfies $f(T_\bullet)\overline{\mathscr{D}_0} \subset \overline{\mathscr{D}_0}$ for each $f \in \mathscr{B}or_{\mathrm{b}}(\mathbb{R}^N)$.

Hint. For each $\xi \in \mathscr{D}_0$, let $\mathcal{K}$ be the closure of $\mathscr{B}or_{\mathrm{b}}(\mathbb{R}^N)\xi$, where $\mathscr{B}or_{\mathrm{b}}(\mathbb{R}^N)$ acts on $\mathcal{H}$ via the Borel functional calculus of $T_\bullet$. By Prop. 5.10.19, $\mathcal{K}$ can be identified with $L^2(\mathbb{R}^N, \mu)$. Use Thm. 7.7.4 to show that vectors of the form $T_1^{n_1} \cdots T_N^{n_N}\xi$ span a dense subspace of $\mathcal{K}$.

(This result can be viewed as an operator-theoretic interpretation of Thm. 7.7.4.) □

The following property can be viewed as a variant of Thm. 7.6.2.

Problem 7.17. Let $T_1, \dots, T_N$ be strongly commuting self-adjoint operators on $\mathcal{H}$. Let $f \in \mathscr{B}or(\mathbb{R}^N, \mathbb{R})$. Suppose that $\mathscr{D}_0$ satisfies the following conditions:

(a) $\mathscr{D}_0 \subset \mathfrak{A}(T_\bullet)$, and $\mathscr{D}_0$ is a dense linear subspace of $\mathcal{H}$.

(b) $T_j\mathscr{D}_0 \subset \mathscr{D}_0$ for each j.

(c) $\mathscr{D}_0 \subset \mathscr{D}(f(T_\bullet))$, and $f(T_\bullet)\mathscr{D}_0 \subset \mathfrak{A}(T_\bullet)$.

Prove that $\mathscr{D}_0$ is a core for $f(T_\bullet)$.

Note that condition (c) is redundant if $f \in \mathbb{R}[x_1, \dots, x_N]$.

Hint. Recall Pb. 6.11. Apply Pb. 7.16 to $(\mathbf{i} + f(T_\bullet))\mathscr{D}_0$ instead of $\mathscr{D}_0$, and mimic the proof of Thm. 7.6.2. □

Example 7.10.8. Let $p \in \mathbb{R}[x_1, \dots, x_N]$. The space $\mathcal{S}_0(\mathbb{R}^N)$ defined in 7.7.6 is a core for the self-adjoint operator $p(\mathrm{M}_{x_1}, \dots, \mathrm{M}_{x_N}) = \mathrm{M}_p$ on $\mathcal{H} := L^2(\mathbb{R}^N, m)$. If T_j denotes the closure of the operator $-\mathbf{i}\frac{\partial}{\partial x_j}$ with domain $\mathscr{D}(-\mathbf{i}\frac{\partial}{\partial x_j}) = C_c^\infty(\mathbb{R}^N)$, then $\mathcal{S}_0(\mathbb{R}^N)$ is a core for $p(T_\bullet)$.

Recall that by Cor. 7.6.7, $T_1, \dots, T_N$ are strongly commuting self-adjoint operators.

Proof. By Pb. 7.15, $\mathcal{S}_0(\mathbb{R}^N)$ is a subspace of $\mathfrak{A}(\mathrm{M}_{x_1}, \dots, \mathrm{M}_{x_N})$. By Cor. 7.7.7 (or by Pb. 7.16), $\mathcal{S}_0(\mathbb{R}^N)$ is dense in $\mathcal{H}$. Therefore, by Pb. 7.17, $\mathcal{S}_0(\mathbb{R}^N)$ is a core for $p(\mathrm{M}_{x_\bullet}) = \mathrm{M}_p$. Since the L^2-Fourier transform intertwines M_{x_j} and T_j (cf. Thm. 7.8.5) and restricts to a unitary operator on $\mathcal{S}_0(\mathbb{R}^N)$ (cf. Thm. 7.8.4), the space $\mathcal{S}_0(\mathbb{R}^N)$ is also a core for $p(T_\bullet)$. □

Example 7.10.9. Let $f : \mathbb{R}^N \to \mathbb{R}$ be defined by $f(x) = x_1^{n_1} \cdots x_N^{n_N} e^{a|x|^2 + tx}$ where $n_j \in \mathbb{N}$, $a < 1/2$, and $t \in \mathbb{R}^N$. Then by Pb. 7.17, $\mathcal{S}_0(\mathbb{R}^N)$ is a core for $f(\mathbf{M}_{x_\bullet}) = \mathbf{M}_f$.

Problem 7.18. Let T be a Hermitian operator on $\mathcal{H}$. Suppose that $(\mathscr{D}_\alpha)_{\alpha \in \mathscr{A}}$ is a family of linear subspaces of $\mathscr{D}(T)$ satisfying the following conditions:

(a) $\sum_\alpha \mathscr{D}_\alpha$ is dense in $\mathcal{H}$.

(b) $T\mathscr{D}_\alpha \subset \mathscr{D}_\alpha$ for each α.

(c) Let T_α be the Hermitian operator on $\mathcal{H}_\alpha := \overline{\mathscr{D}_\alpha}$ defined by $\mathscr{D}(T_\alpha) = \mathscr{D}_\alpha$ and $T_\alpha = T|_{\mathscr{D}_\alpha}$. Then T_α is essentially self-adjoint.

Prove that T is essentially self-adjoint.

Hint. Show that $T + \mathbf{i}$ and $T - \mathbf{i}$ have dense ranges. □

The following result is due to Nelson.

Problem 7.19. Let T be a Hermitian operator on $\mathcal{H}$ satisfying $T\mathscr{D}(T) \subset \mathscr{D}(T)$. Assume that any $\xi \in \mathscr{D}(T)$ is analytic, i.e.,

$$\sum_{n \in \mathbb{N}} \frac{r^n}{n!} \|T^n \xi\| < +\infty$$

for some $r \in \mathbb{R}_{>0}$ (possibly depending on ξ). Prove that T is essentially self-adjoint.

Hint. First consider the special case where $\mathscr{D}(T) = \mathrm{Span}\{T^n\xi : n \in \mathbb{N}\}$ for some $\xi \in \mathscr{D}(T)$. Show that T admits a self-adjoint extension $\widetilde{T}$ by relating this situation to the spectral-theoretic interpretation of the Hamburger moment problem in Exp. 6.11.2. Then apply Pb. 7.17 to $\widetilde{T}$ to conclude that T is essentially self-adjoint.

For the general case, reduce to this special case using Pb. 7.18. □

7.A Integrals of measurable functions valued in Hilbert spaces

In this appendix section, we let $\mathcal{H}, \mathcal{K}$ be Hilbert spaces, and let $(X, \mathfrak{M}, \mu)$ be a measure space.

7.A.1 Introduction

In Sec. 7.3, 7.4, and 7.5 we encountered integrals of the form $\int_{\mathbb{R}^N} f dm$ where $f : \mathbb{R}^N \to \mathcal{H}$ is a measurable function satisfying $\int |f| dm < +\infty$. As mentioned in Subsec. 7.4.2, for our purposes in this course it suffices to consider the case $f \in C_c(\mathbb{R}^N, \mathcal{H})$, in which the integral $\int f dm$ can be defined by the Riemann integral (as in Def. 1.9.4)—specifically, as the limit of Riemann sums obtained by partitioning a bounded box in $\mathbb{R}^N$ containing $\mathrm{Supp}(f)$. In this setting, many useful properties follow directly from approximation by Riemann sums; for example, the inequality

$$\Big\|\int_{\mathbb{R}^N} f dm\Big\| \leqslant \int_{\mathbb{R}^N} |f| dm \tag{7.41}$$

follows immediately from the corresponding inequality for Riemann sums.

From (7.41), one readily derives a version of the dominated convergence theorem: Let (f_n) be a sequence in $C_c(\mathbb{R}^N,\mathcal{H})$ converging pointwise to $f \in C_c(\mathbb{R}^N,\mathcal{H})$. If there exists an integrable $g:\mathbb{R}^N \to \mathbb{R}_{\geqslant 0}$ such that $|f_n| \leqslant g$ for all n, then

$$\lim_n \int_{\mathbb{R}^N} f_n dm = \int_{\mathbb{R}^N} f dm$$

This follows by applying the standard DCT to the sequence $|f-f_n|$ and noting the inequality $\|\int f - \int f_n\| \leqslant \int |f-f_n|$ due to (7.41).

Although in practice it is sufficient to work with integrals of functions in $C_c(\mathbb{R}^N,\mathcal{H})$, in this section we develop the general integration theory for integrable functions, i.e. measurable functions $f: X \to \mathcal{H}$ satisfying $\int_X |f| d\mu < +\infty$. As discussed in Subsec. 7.A.2, a subtlety arises because the sum of two measurable functions need not be measurable unless their ranges are separable. Hence, it is natural to restrict attention to integrable functions with separable (essential) ranges—these are known as Bochner integrable functions.

7.A.2 Criteria for measurability

Proposition 7.A.1. *Let $f: X \to \mathcal{H}$ with separable range $f(X)$. Then the following are equivalent.*

(1) f is measurable.

(2) For each $\eta \in \mathcal{H}$, the function $\langle \eta | f\rangle : X \to \mathbb{C}$ defined by $x \mapsto \langle \eta | f(x)\rangle$ is measurable.

Proof. (1)⇒(2) is obvious, since $\langle \eta | f(x)\rangle$ is the composition of a continuous linear functional and the measurable map f. Let us prove (2)⇒(1).

Let E be a countable dense subset of $f(X)$. Then $V := \overline{\mathrm{Span}_{\mathbb{C}} E}$ is separable, since $\mathrm{Span}_{\mathbb{Q}+\mathbf{i}\mathbb{Q}} E$ is a countable dense subspace of V. Moreover, V contains $f(X)$. Since the inclusion map $V \hookrightarrow \mathcal{H}$ is continuous, to show that $f: X \to \mathcal{H}$ is measurable, it suffices to show that the restriction $f: X \to V$ is measurable.

From now on, we view V as the codomain of f. We need to show that $f^{-1}(\Omega)$ is measurable for each open set $\Omega \subset V$. Since Ω is a separable metric space, it is second countable and hence Lindelöf. Therefore, Ω is a countable union of open balls. Thus, it suffices to prove that $f^{-1}(B_V(v,r))$ is measurable for each $v \in V$ and $r \in \mathbb{R}_{>0}$.

Let $\xi_1, \xi_2, \ldots$ be a dense sequence in the closed unit ball $\overline{B}_V(0,1)$. Then $\|\eta\| = \sup_n |\langle \xi_n|\eta\rangle|$ for each $\eta \in V$. Therefore, the function

$$|f - v| : X \to \mathbb{R}_{\geqslant 0} \qquad x \mapsto \|f(x) - v\| = \sup_n |\langle \xi_n|f(x) - v\rangle|$$

is measurable, because each $\langle \xi_n|f - v\rangle$ is measurable. The inverse image of $[0, r)$ under the map $|f - v|$ is precisely $f^{-1}(B_V(v, r))$. Therefore $f^{-1}(B_V(v, r))$ is measurable. □

Example 7.A.2. Assume that X is a σ-compact topological space. Then any continuous map $f : X \to \mathcal{H}$ has separable range.

Proof. Write $X = \bigcup_{n\in\mathbb{Z}_+} K_n$ where each K_n is compact. Then $f(K_n)$ is a compact subset of $\mathcal{H}$, and hence is metrizable. By Thm. 1.5.14, $f(K_n)$ is second countable (and hence separable). Therefore $f(X) = \bigcup_n f(K_n)$ is separable. □

Given measurable functions $f, g : X \to \mathcal{H}$, the sum $f + g$ is not necessarily measurable.[10] However, if $f(X), g(X)$ are separable, then $f + g$ is indeed measurable.

Corollary 7.A.3. *Let $f, g : X \to \mathcal{H}$ be measurable functions with separable range. Let $\gamma : X \to \mathbb{C}$ be measurable. Then the functions $f + g : X \to \mathcal{H}$ and $\gamma f : X \to \mathcal{H}$ and $\langle f|g\rangle : X \to \mathbb{C}$ are measurable.*

Proof. By Prop. 7.A.1, for each $\eta \in \mathcal{H}$, the $\mathbb{C}$-valued functions $\langle \eta|f\rangle$ and $\langle \eta|g\rangle$ are measurable. So $\langle \eta|f + g\rangle$ and $\langle \eta|\gamma f\rangle$ are measurable. By Prop. 7.A.1 again, $f + g$ and γ are measurable.

The function $|f|^2$ is the composition of f and a continuous function, and hence is measurable. Similarly, for each $\lambda \in \mathbb{C}$, the function $|\lambda f + g|^2$ is measurable since $\lambda f + g$ is measurable. Therefore $\langle f|g\rangle$ is measurable by the polarization identity. □

Corollary 7.A.4. *Let (f_n) be a sequence of measurable functions $X \to \mathcal{H}$ converging pointwise to $f : X \to \mathcal{H}$. Assume that $f_n(X)$ is separable for each n. Then f is measurable with separable range $f(X)$.*

Proof. The closure of $\bigcup_n f_n(X)$ is a separable set containing $f(X)$. Therefore, we can apply Prop. 7.A.1: For each $\eta \in \mathcal{H}$, the function $\langle \eta|f_n\rangle$ is measurable for each n. Hence $\langle \eta|f\rangle$ is measurable. Thus f is measurable. □

[10] More generally, $f \vee g : x \in X \mapsto (f(x), g(x)) \in \mathcal{H}^2$ is not necessarily measurable. This is related to the fact that not all open subsets $\Omega \subset \mathcal{H}^2$ can be written as a countable union $\bigcup_n A_n \times B_n$ where $A_n, B_n \subset \mathcal{H}$ are open.

7.A.3 Weakly integrable functions

Definition 7.A.5. A measurable function $f : X \to \mathcal{H}$ is called **weakly integrable** if there exists a vector $\int_X f d\mu \in \mathcal{H}$ such that for each $\eta \in \mathcal{H}$, we have

$$\left\langle \eta \middle| \int_X f d\mu \right\rangle = \int_X \langle \eta | f(x) \rangle d\mu(x)$$

where the RHS is μ-integrable. If such $\int_X f d\mu$ exists, then it must be unique, and we call it the **integral** of f on X with respect to μ.

Clearly, the set of weakly integrable functions form a linear subspace of the space of measurable functions. Moreover, the integral operator $\int_X$ on this subspace is linear.

Proposition 7.A.6. *Let $f : X \to \mathcal{H}$ be weakly integrable. Then*

$$\Big\| \int_X f d\mu \Big\| \leqslant \|f\|_{L^1(X,\mu)}$$

Proof. The inequality is obvious when $\xi := \int_X f d\mu$ is zero. Thus, we assume WLOG that $\xi \neq 0$. Then

$$\|\xi\|^2 = \int_X \langle \xi | f \rangle d\mu \leqslant \int_X \|\xi\| \cdot |f| d\mu = \|\xi\| \int_X |f| d\mu$$

Dividing both sides by $\|\xi\|$ yields the desired inequality. □

Proposition 7.A.7. *Let $f : X \to \mathcal{H}$ be weakly integrable. Let $T \in \mathfrak{L}(\mathcal{H}, \mathcal{K})$. Then $T \circ f$ is weakly integrable, and*

$$T \int_X f d\mu = \int_X T \circ f d\mu$$

Proof. Clearly $T \circ f$ is measurable. Let $\xi = \int_X f d\mu$. For each $\eta \in \mathcal{H}$, we have

$$\int_X \langle \eta | T \circ f \rangle d\mu = \int_X \langle T^* \eta | f \rangle d\mu = \langle T^* \eta | \xi \rangle = \langle \eta | T\xi \rangle$$

which shows that $T \circ f$ is weakly integrable with integral $T\xi$. □

Proposition 7.A.8. *Let $f : X \to \mathcal{H}$ be weakly integrable. Then $\int_X f d\mu$ belongs to the closure of* $\mathrm{Span}_{\mathbb{C}}(f(X))$.

Proof. Let $\xi = \int_X f d\mu$. Let V be the closure of $\mathrm{Span}_{\mathbb{C}}(f(X))$. For each $\eta \in V^\perp$, we have $\eta \perp f(X)$, and hence

$$\langle \eta | \xi \rangle = \int_X \langle \eta | f \rangle d\mu = 0$$

This proves that $\xi \in (V^\perp)^\perp$. Since $(V^\perp)^\perp = V$ due to Cor. 3.4.8, we have $\xi \in V$. □

7.A.4 Integrable and Bochner integrable functions

Definition 7.A.9. Let $f : X \to \mathcal{H}$ be a map. We say that f is **integrable** if f is measurable and $\|f\|_{L^1(X,\mu)} < +\infty$. We say that f is **Bochner integrable** if f is integrable and $f(X)$ is separable.

In the literature, the definition of Bochner integrability is slightly weaker: it only requires the existence of a null set $\Delta \subset X$ such that $f(X\backslash\Delta)$ is separable. We adopt the stronger version above, which is sufficient for our purposes and simplifies the exposition.

Theorem 7.A.10. *Any integrable function $f : X \to \mathcal{H}$ is weakly integrable.*

Proof. The linear functional

$$\mathcal{H}^{\complement} \to \mathbb{C} \qquad \overline{\eta} \mapsto \int_X \langle \eta | f(x) \rangle d\mu(x)$$

is bounded since $|\int_X \langle \eta | f \rangle d\mu| \leqslant \|\eta\| \int_X |f| d\mu = \|\eta\| \cdot \|f\|_{L^1}$. Therefore, by the Riesz-Fréchet Thm. 3.5.3, this linear functional is represented by some $\xi \in \mathcal{H}$. Thus f is weakly integrable with integral ξ. □

Recall Def. 1.2.32 for the definition of first-countable nets.

Theorem 7.A.11 (Dominated convergence theorem). *Let $(f_\alpha)_{\alpha \in \mathscr{I}}$ be a first-countable net of Bochner integrable maps $X \to \mathcal{H}$ converging pointwise to a map $f : X \to \mathcal{H}$. Assume that there exists a measurable $g : X \to \overline{\mathbb{R}}_{\geqslant 0}$ with $\int_X g d\mu < +\infty$ such that $|f_\alpha| \leqslant g$ holds for each $\alpha \in \mathscr{I}$. Then f is Bochner integrable, and*

$$\lim_\alpha \int_X f_\alpha d\mu = \int_X f d\mu \tag{7.42}$$

Note that the convergence of the LHS of (7.42) is also part of the conclusion of Thm. 7.A.11.

Proof. By Def. 1.2.32, there exists an increasing sequence $(\alpha_n)_{n \in \mathbb{Z}_+}$ in $\mathscr{I}$ eventually larger than any element of $\mathscr{I}$. Therefore $(f_{\alpha_n})_{n \in \mathbb{Z}_+}$ converges pointwise to f. Since each f_{α_n} is measurable with separable range, by Cor. 7.A.4, f is measurable with separable range. Clearly $|f| \leqslant g$, and hence $\int_X |f| d\mu \leqslant \int_X g d\mu < +\infty$. Thus f is Bochner integrable.

By Cor. 7.A.3, each $f - f_\alpha$ is measurable; in particular, $|f - f_\alpha|$ is measurable. Since $|f - f_\alpha| \leqslant 2g$, $f - f_\alpha$ is integrable, and hence is weakly integrable due to Thm. 7.A.10. By Prop. 7.A.6,

$$\Big\| \int_X f d\mu - \int_X f_\alpha d\mu \Big\| = \Big\| \int_X (f - f_\alpha) d\mu \Big\| \leqslant \int_X |f - f_\alpha| d\mu$$

and the RHS converges under $\lim_\alpha$ to 0 due to DCT for first countable nets (cf. Thm. 1.2.37). This proves the limit (7.42). □

Corollary 7.A.12. *Let $I \subset \mathbb{R}$ be a proper interval. Assume that $f : X \times I \to \mathcal{H}$ satisfies the following properties:*

(a) For each $t \in I$, the following function is Bochner integrable.

$$f(-,t) : X \to \mathcal{H} \qquad x \mapsto f(x,t)$$

(b) The partial derivative function $\partial_I f(x,t)$ over the variable on I exists for all $x \in X, t \in I$. Moreover, for each $x \in X$, the following function is continuous

$$\partial_I f(x,-) : I \to \mathcal{H} \qquad t \mapsto \partial_I f(x,t)$$

(c) There exists a measurable $g : X \to \mathbb{R}_{\geqslant 0}$ with $\int_X g d\mu < +\infty$ such that for all $t \in I$, the function

$$\partial_I f(-,t) : X \to \mathcal{H} \qquad x \mapsto \partial_I f(x,t)$$

satisfies

$$|\partial_I f(-,t)| \leqslant g$$

Then $\int_X f(x,t)d\mu(x)$ is differentiable over t; for each $t \in I$, the function $\partial_I f(-,t)$ is Bochner integrable, and

$$\frac{d}{dt}\int_X f(x,t)d\mu(x) = \int_X \partial_I f(x,t)d\mu(x) \tag{7.43}$$

Proof. Fix $t \in I$. Consider the first-countable net of measurable functions $(\varphi_s)_{s \in I\backslash\{t\}}$ from X to $\mathcal{H}$ defined by

$$\varphi_s(x) = \frac{f(x,s) - f(x,t)}{s-t}$$

The preorder on $I\backslash\{t\}$ is defined such that larger elements are closer to t. Then each $\varphi_s(x)$ is measurable (by Cor. 7.A.3) and has separable range. By the fundamental theorem of calculus, for each $\eta \in \mathcal{H}$ we have

$$\begin{aligned}|\langle \eta|\varphi_s(x)\rangle| &= \left|\frac{\langle \eta|f(x,s)\rangle - \langle \eta|f(x,t)\rangle}{s-t}\right| = |s-t|^{-1}\left|\int_t^s \partial_I\langle \eta|f(x,u)\rangle du\right| \\ &= |s-t|^{-1}\left|\int_t^s \langle \eta|\partial_I f(x,u)\rangle du\right| \leqslant |s-t|^{-1}\int_{\min\{s,t\}}^{\max\{s,t\}} \|\eta\| \cdot g(x)du = \|\eta\| \cdot g(x)\end{aligned}$$

and hence $|\varphi_s| \leqslant g$. Therefore Thm. 7.A.11 is applicable, which shows that $\partial_I f(-,t)$ is Bochner integrable and $\lim_{s\to t}\int_X \varphi_s d\mu = \int_X \partial_I f(-,t)d\mu$. This proves (7.43). $\square$

The following proposition is used solely in the proof of Thm. 7.4.4.

Proposition 7.A.13. *Let $f,g:X\rightarrow\mathcal{H}$ be Bochner integrable. Then*

$$\Big\langle\int_X fd\mu\Big|\int_X gd\mu\Big\rangle=\int_X\int_X\langle f(x)|g(y)\rangle d\mu(x)d\mu(y)=\int_X\int_X\langle f(x)|g(y)\rangle d\mu(y)d\mu(x)$$

Proof. Let V be the closure of $\mathrm{Span}(\mathrm{Rng}(f))+\mathrm{Span}(\mathrm{Rng}(g))$, which is separable. By Prop. 7.A.8, the vectors $\xi=\int_X fd\mu$ and $\eta\in\int_X gd\mu$ belong to V. Let $(e_n)_{n\in\mathbb{Z}_+}$ be a countable orthonormal basis of V. Then

$$\langle\xi|\eta\rangle=\sum_n\langle\xi|e_n\rangle\langle e_n|\eta\rangle=\sum_n\int\langle\xi(x)|e_n\rangle d\mu(x)\cdot\int\langle e_i|\eta(y)\rangle d\mu(y)$$
$$=\sum_n\int\int\langle\xi(x)|e_n\rangle\langle e_n|\eta(y)\rangle d\mu(x)d\mu(y)$$

To justify interchanging the sum with the two integrals, we verify that

$$\int\sum_n|\langle\xi(x)|e_n\rangle\langle e_n|\eta(y)\rangle|d\mu(x)\leqslant\int\sqrt{\sum_n|\langle\xi(x)|e_n\rangle|^2}\sqrt{\sum_n|\langle e_n|\eta(y)\rangle|^2}d\mu(x)$$
$$=\int\|\xi(x)\|\cdot\|\eta(y)\|d\mu(x)=\|\xi\|_{L^1}\cdot\|\eta(y)\|$$

for each $y\in X$, and hence

$$\int\int\sum_n|\langle\xi(x)|e_n\rangle\langle e_n|\eta(y)\rangle|d\mu(x)d\mu(y)\leqslant\|\xi\|_{L^1}\|\eta\|_{L^1}<+\infty$$

Therefore, by DCT,

$$\langle\xi|\eta\rangle=\int\int\sum_n\langle\xi(x)|e_n\rangle\langle e_n|\eta(y)\rangle d\mu(x)d\mu(y)=\int\int\langle f(x)|g(y)\rangle d\mu(x)d\mu(y)$$

which proves the first desired equality. The second follows by the same argument. □

8 Compact operators and unbounded operators with compact resolvents

8.1 The historical context of compact operators

8.1.1 From completely continuous forms to compact operators

In this chapter, we study an important class of bounded linear operators—*compact operators*. Compactness is a natural analytic condition that guarantees the discreteness of the spectrum of an operator. The theory of compact operators is among the earliest developed in linear functional analysis. A landmark result in this theory is the diagonalization theorem for compact self-adjoint operators (the so-called *Hilbert-Schmidt theorem*), first established by Hilbert in [Hil06]. In fact, Hilbert introduced the Hilbert space $l^2(\mathbb{Z})$ and the fundamental notion of weak convergence precisely to lay the groundwork for the diagonalization of compact self-adjoint operators, thereby advancing the theory of integral equations—a central topic in mathematics at the time.

Despite their importance, compact operators are one of the most conceptually confusing notions in functional analysis. As mentioned earlier (cf. the paradigm shift (5.19b)), early developments in the field were framed in terms of bilinear/sesquilinear forms rather than linear operators. Consequently, in [Hil06] (and in many subsequent works by members of the Hilbert school), Hilbert did not use the modern operator-theoretic viewpoint. Instead, he considered the corresponding concept of *completely continuous forms*. The notion of a compact operator, as we now understand it, emerged only later.

In fact, the more direct translation of completely continuous forms into operator language is not the compact operator, but the completely continuous operator, first introduced by Riesz in [Rie10]. It was only later, in [Rie18], that Riesz introduced compact operators, recognizing that the earlier notions of completely continuous forms and operators were not suitable for normed vector spaces lacking natural dual structures (such as $C[a, b]$). The historical progression of these ideas can be summarized as follows:

$$\begin{gathered}\text{completely continuous forms}\\ \downarrow\\ \text{completely continuous operators}\\ \downarrow\\ \text{compact operators}\end{gathered}$$

Although the three concepts above are equivalent in the setting of Hilbert spaces (cf. Thm. 8.4.7), their definitions appear quite different. Consequently, the ways they are applied in proofs can differ significantly, even when proving the same theorem. It was through this process of conceptual evolution that the originally intuitive idea of a completely continuous form gradually transformed into the more abstract and technically intricate notion of a compact operator—a concept whose definition and interpretation have since become considerably less transparent.

8.1.2 The modern approach to the Dirichlet problem

Another reason the theory of compact operators may appear obscure is that its modern usage, particularly in partial differential equations, differs significantly from its historical application. To illustrate this, let us compare the two approaches using the Dirichlet problem. Let Ω be a bounded open subset of $\mathbb{R}^N$ with smooth boundary $\partial\Omega$. Given $g \in C(\partial\Omega)$, the **Dirichlet problem** seeks a function $u \in C(\overline{\Omega})$ that is smooth on Ω and satisfies

$$-\Delta u|_\Omega = 0 \qquad u|_{\partial\Omega} = g \tag{8.1}$$

To explain the modern approach, assume for simplicity that $g \in C^\infty(\partial\Omega)$. Then g can be extended to a function $\widetilde{g} \in C_c^\infty(\mathbb{R}^N)$. Set $\varphi = \Delta\widetilde{g}|_\Omega$. Then $\xi := u - \widetilde{g}$ satisfies

$$-\Delta\xi|_\Omega = \varphi \qquad \xi|_{\partial\Omega} = 0 \tag{8.2}$$

Hence, solving the original problem (8.1) reduces to showing that for each $\varphi \in C^\infty(\overline{\Omega})$, there exists $\xi \in C^\infty(\overline{\Omega})$ satisfying (8.2).

Recall from Def. 7.1.12 that the negative Dirichlet Laplacian $-\Delta_{\mathrm{D}}$ is a positive self-adjoint operator on $\mathcal{H} := L^2(\Omega, m)$. As we will see later in Thm. 8.9.5, the boundedness of Ω implies that $-\Delta_{\mathrm{D}}$ is strictly positive. By Prop. 7.1.2, it follows that $-\Delta_{\mathrm{D}}^{-1} \in \mathfrak{L}(\mathcal{H})$, and hence for every $\varphi \in C^\infty(\overline{\Omega})$ there exists $\xi \in H_0^1(\Omega)$ such that $-\Delta_{\mathrm{D}}\xi = \varphi$. Regularity results in Sobolev spaces then imply that such ξ actually lies in $C^\infty(\overline{\Omega})$ and satisfies $\xi|_{\partial\Omega} = 0$. Details of this approach can be found in most modern PDE textbooks, e.g. [Eva, Ch. 5,6], [Tay-1, Ch. 4,5], or [Yu].

The reader immediately notices that the modern approach to the Dirichlet problem described above does not rely on the theory of compact operators. Indeed, compact operators are used not in solving the harmonic equation (8.1) itself, but rather in showing that $-\Delta_{\mathrm{D}}$ is diagonalizable, thereby providing solutions to the **Helmholtz equation**

$$-\Delta\xi|_\Omega = \lambda\xi|_\Omega \qquad \xi|_{\partial\Omega} = 0$$

for $\lambda \in \mathbb{R}_{\geqslant 0}$. This is achieved by showing that the bounded positive operator $(\mathbf{1} - \Delta_{\mathrm{D}})^{-1}$ (or simply $(-\Delta_{\mathrm{D}})^{-1}$) is a compact operator on $\mathcal{H}$, which can then be diagonalized via the Hilbert-Schmidt theorem. Another application of Sobolev regularity theory shows that all eigenfunctions of Δ_{D} are smooth and vanish on $\partial\Omega$.

As we shall see in the following subsections, in the historical approach to the Dirichlet problem (8.1), compact operators play a crucial and central role. In fact, the study of the Dirichlet problem was one of the main motivations for introducing completely continuous forms/operators and compact operators.

8.1.3 The integral equation approach to the Dirichlet problem

The historical approach to the Dirichlet problem (8.1) reduces it to the solution of an associated integral equation using the method of the **double layer potential**. We briefly sketch this method in this and the following subsection. Detailed proofs of the statements below can be found in [Sim-O, Sec. 3.3], [Fol-P, Ch. 3], or [Tay-2, Sec. 7.11].

Recall that given $g \in C(\partial\Omega)$, we seek a function $u \in C(\overline{\Omega})$ that is smooth on Ω and satisfies (8.1). Define the **fundamental solution**

$$\Phi(x) = \begin{cases} -\dfrac{1}{2\pi}\log|x| & \text{if } N = 2 \\ \dfrac{1}{(N-2)\sigma_{N-1}\cdot|x|^{N-2}} & \text{if } N \geqslant 3 \end{cases} \tag{8.3}$$

where σ_{N-1} denotes the area of the $(N-1)$-dimensional sphere $\mathbb{S}^{N-1}$. For each $f \in C(\partial\Omega)$, define a function $\mathcal{T}f$ on $\mathbb{R}^N$ by

$$(\mathcal{T}f)(x) = \int_{\partial\Omega} \langle \nabla\Phi(x-y), \mathbf{n}_y\rangle \cdot f(y)dy \tag{8.4}$$

where $\mathbf{n}_y$ is the unit normal vector at $y \in \partial\Omega$ pointing outward. One computes

$$\langle \nabla\Phi(x-y), \mathbf{n}_y\rangle = -\frac{\langle x-y, \mathbf{n}_y\rangle}{\sigma_{N-1}|x-y|^N}$$

From this relation, we conclude

$$\sup_{x\in\mathbb{R}^N\setminus\{y\}} |x-y|^{N-2}|\langle \nabla\Phi(x-y), \mathbf{n}_y\rangle| < +\infty \tag{8.5}$$

and hence the convergence of the integral in (8.4).

One can show that $\mathcal{T}f$ is (smooth and) harmonic on $\mathbb{R}^N\setminus\partial\Omega$, and that the restriction $\mathcal{T}f|_{\overline{\Omega}}$ has a jump discontinuity of size $-\frac{1}{2}f$ at $\partial\Omega$. The latter means that the function $u : \overline{\Omega} \to \mathbb{C}$ defined by

$$u(x) = \begin{cases} (\mathcal{T}f)(x) & \text{if } x \in \Omega \\ \frac{1}{2}f(x) + (\mathcal{T}f)(x) & \text{if } x \in \partial\Omega \end{cases} \tag{8.6}$$

is continuous on the entire domain $\overline{\Omega}$. Therefore, this function is a continuous function on $\overline{\Omega}$, smooth on Ω, and satisfies the differential equation (8.1) with boundary condition $g = \frac{1}{2}f + \mathcal{T}f$.

Consequently, the Dirichlet problem (8.1) is reduced to finding $f \in C(\partial\Omega)$ that satisfies the **integral equation**

$$\frac{1}{2}f(x) + \int_{\partial\Omega} K(x,y)f(y)dy = g(x) \qquad \text{with } K(x,y) = \langle \nabla\Phi(x-y), \mathbf{n}_y \rangle \tag{8.7}$$

The physical meaning of the above construction is as follows.

Remark 8.1.1. If a point charge Q is located at $y \in \mathbb{R}^N$, then the electric potential at $x \in \mathbb{R}^N$ generated by Q is $V(x) = Q \cdot \Phi(x-y)$. Thus, if $\varphi : \mathbb{R}^N \to \mathbb{R}$ is a charge distribution function, the potential function generated by φ is $V(x) = \int_{\mathbb{R}^N} \varphi(y)\Phi(x-y)dy$, and the associated electric field is given by $\mathbf{E} = -\nabla V$.

According to Gauss's law, the charge distribution function φ is equal to the divergence $\mathrm{div}\mathbf{E}$ of $\mathbf{E}$. Hence $\varphi = -\Delta V$. In particular, the potential function V satisfies the harmonic equation $-\Delta V = 0$ in regions where no electric charges are present.

Now, place a slightly smaller $(N-1)$-dimensional surface $\partial'\Omega$ inside $\partial\Omega$ such that the distance between $\partial'\Omega$ and $\partial\Omega$ is a very small positive number ε. (Thus, $\partial'\Omega$ and $\partial\Omega$ form a thin "**double layer**".) On $\partial\Omega$ and on $\partial'\Omega$, we put $(N-1)$-dimensional charge distribution functions $Q = Q(x)$ and $Q' = Q'(x)$ respectively satisfying $Q'(x) = -Q(x) = f(x)/\varepsilon$. (Note that since the two surfaces are very close, the points x on $\partial'\Omega$ can be identified with those on $\partial\Omega$, allowing Q' and Q to be viewed as functions on $\partial\Omega$.) Then $\mathcal{T}f$ represents the potential generated by this pair of opposite surface charges.

Since there are no charges inside $\partial'\Omega$, it follows from the above discussion that $-\Delta\mathcal{T}f|_\Omega = 0$. Eq. (8.6) expresses the fact that $\mathcal{T}f$ exhibits a jump discontinuity of magnitude $-\frac{1}{2}f$ when moving from inside $\partial'\Omega$ to midway between $\partial'\Omega$ and $\partial\Omega$. This gives a heuristic explanation of why the function u defined in (8.6) is continuous on $\overline{\Omega}$ and harmonic on Ω. $\square$

8.1.4 Uniqueness of solutions to the integral equation

In finite-dimensional linear algebra, an injective linear map is necessarily surjective. In the setting of the double layer potential introduced in the previous subsection, the injectivity of the operator $\frac{1}{2} + T_K$ is much easier to establish than its surjectivity, where

$$T_K : C(\partial\Omega) \to C(\partial\Omega) \qquad (T_K f)(x) = (\mathcal{T}f)(x) = \int_{\partial\Omega} K(x,y)f(y)dy \tag{8.8}$$

is the integral operator associated with the kernel K given in (8.7). (The fact that $f \in C(\partial\Omega)$ implies $T_K f \in C(\partial\Omega)$ follows from (8.5); see Pb. 8.6.)

The injectivity of $\frac{1}{2}+T_K$ rests on the key observation that for any $f \in C(\partial\Omega)$, the normal derivative $\langle \nabla\mathcal{T}f, \mathbf{n}\rangle$ does not exhibit a jump discontinuity when passing from the interior to the exterior of $\partial\Omega$. Physically (cf. Rem. 8.1.1), this means that the electric field $-\nabla\mathcal{T}f$ generated by the double layer potential $\mathcal{T}f$ remains continuous in the normal direction across $\partial\Omega$. A detailed proof of this property can be found in the references cited in Subsec. 8.1.3.

We now explain why $\frac{1}{2}+T_K$ is injective. Suppose $f\in C(\partial\Omega)$ satisfies $\frac{1}{2}f+T_Kf=0$. Then the continuous function $u:\overline{\Omega}\to\mathbb{C}$ defined in (8.6) vanishes on the boundary $\partial\Omega$. Since u is harmonic on Ω, the maximum principle implies $u=0$ on Ω, and hence $\mathcal{T}f=0$ on Ω. In particular, $\langle\nabla\mathcal{T}f,\mathbf{n}\rangle$ vanishes inside and near $\partial\Omega$. From the previous paragraph, it then follows that $\langle\nabla\mathcal{T}f,\mathbf{n}\rangle$ vanishes outside and near $\partial\Omega$.

Choose $R>0$ such that $\overline{\Omega}\subset B_{\mathbb{R}^N}(0,R)$. By Green's identity (7.8), and using the fact that $\mathcal{T}f$ is harmonic outside $\partial\Omega$, we obtain

$$\begin{aligned}&\int_{B_{\mathbb{R}^N}(0,R)\backslash\overline{\Omega}}|\nabla\mathcal{T}f|^2\\ =&\int_{|y|=R}\overline{\mathcal{T}f(y)}\cdot\langle\nabla\mathcal{T}f(y),y/|y|\rangle dy-\int_{\partial\Omega}\overline{\mathcal{T}f(y)}\cdot\langle\nabla\mathcal{T}f(y),\mathbf{n}_y\rangle dy\end{aligned}$$

The last integral on the right-hand side vanishes by the previous argument. One can easily show that the second last integral tends to zero as $R\to+\infty$. It follows that outside $\partial\Omega$, the gradient of $\mathcal{T}f$ is zero, and hence $\mathcal{T}f$ is constant. Since $\mathcal{T}f$ vanishes at infinity (as can be easily verified), we conclude that $\mathcal{T}f|_{\mathbb{R}^N\backslash\overline{\Omega}}=0$.

Similar to (8.6), one can show that the function $v:\mathbb{R}^N\backslash\Omega\to\mathbb{C}$ defined by

$$v(x)=\begin{cases}(\mathcal{T}f)(x) & \text{if } x\in\mathbb{R}^N\backslash\overline{\Omega}\\ -\frac{1}{2}f(x)+(\mathcal{T}f)(x) & \text{if } x\in\partial\Omega\end{cases}$$

ia continuous. Therefore, $-\frac{1}{2}f+\mathcal{T}f=0$ on $\partial\Omega$, i.e., $-\frac{1}{2}f+T_Kf=0$. Combining this with $\frac{1}{2}f+T_Kf=0$, we obtain $f=0$. This proves the injectivity of $\frac{1}{2}+T_K$.

8.1.5 Existence of solutions to the integral equation: the Fredholm alternative

In general, an injective bounded linear operator is not necessarily surjective. However, as later discovered, the integral operator T_K is compact (see Pb. 8.6), and thus enjoys the remarkable property known as the **Fredholm alternative**. This principle states that for any *nonzero* $\lambda\in\mathbb{C}$, exactly one of the following two cases occurs:

- $\mathrm{Ker}(\lambda-T_K)\neq 0$; that is, the homogeneous equation $\lambda f-T_Kf=0$ has a non-zero solution f.
- $\lambda-T_K$ is surjective; that is, for each g, the inhomogeneous equation $\lambda f-T_Kf=g$ admits a solution f.

In short, $\lambda - T_K$ is injective iff it is surjective.

The Fredholm alternative for compact operators on Banach spaces was established by Riesz in [Rie18], and hence applies to the integral equation (8.7) in arbitrary dimension N. More precisely, since the integral operator T_K arising from the double layer potential satisfies $\mathrm{Ker}(\frac{1}{2}+T_K)=0$ (as explained in Subsec. 8.1.4), it follows that $\frac{1}{2}+T_K$ has full range $C(\partial\Omega)$. Consequently, by the discussion in Subsec. 8.1.3, the Dirichlet problem (8.1) is solvable for every boundary condition $g\in C(\partial\Omega)$.

In the early development of functional analysis, however, attention was focused on the simplest nontrivial case $N=2$, i.e. when $\dim\partial\Omega=1$. In this setting, by parametrizing the curve $\partial\Omega$, the integral operator can be expressed in the form

$$T_K: C(I)\to C(I) \qquad (T_K f)(x)=\int_I K(x,y)f(y)dy \tag{8.9}$$

where $I\subset\mathbb{R}$ is a compact integral and $K\in C(I\times I)$. The Fredholm alternative was first established by Fredholm in [Fre03] for this one-dimensional case.

It is important to note that, consistent with the early tradition of functional analysis, Fredholm did not formulate integral equations in terms of linear operators—and unlike Hilbert, who came later, he also did not use the language of bilinear forms. Rather, as mentioned in the answer to Question 5.1.3, Fredholm introduced the resolvent $(\lambda-T_K)^{-1}$ as the limit of inverses of finite-rank matrices, expressed via determinants and their minors using Cramer's rule. Accordingly, Fredholm viewed the resolvent $(\lambda-T_K)^{-1}$ as a kernel function defined on $I\times I$ of the form $\Delta(\lambda,x,y)/\Delta(\lambda)$ where $\Delta(\lambda)$ is the determinant of $\lambda-K$, and $\Delta(\lambda,x,y)$ is the minor function of $\lambda-K$.

Fredholm established the Fredholm alternative by proving the following:

- If $\Delta(\lambda)=0$, then the homogeneous equation $\lambda f-T_Kf=0$ has solution space of dimension $0<m<+\infty$, and the solutions can be expressed in terms of determinants. Moreover, m equals the order of the zero of the holomorphic function $\Delta(z)$ at $z=\lambda$.

- If $\Delta(\lambda)\neq 0$, then the inhomogeneous equation $\lambda f-T_Kf=g$ has a solution given by $f(x)=\Delta(\lambda)^{-1}\int_I\Delta(\lambda,x,y)g(y)dy$.

We will not pursue Fredholm's approach in this course. Interested readers may refer to [Lax, Ch. 24] for details. In the following sections, we will first study Hilbert's approach to completely continuous Hermitian forms, particularly the Hilbert-Schmidt theorem on the diagonalizability of such forms. Next, we study Riesz's proof of the Fredholm alternative for compact operators on Banach spaces. Finally, we apply these results to unbounded operators.

8.2 Completely continuous sesquilinear forms

Let $\mathcal{H}, \mathcal{K}$ be Hilbert spaces. Recall Def. 3.7.1 for the notion of weak topology in inner product spaces.

8.2.1 The definition of completely continuous sesquilinear forms

Definition 8.2.1. A bounded sesquilinear form $\omega : \mathcal{K}\times\mathcal{H} \to \mathbb{C}$ is called **completely continuous** if its restriction to $\overline{B}_{\mathcal{K}}(0,1) \times \overline{B}_{\mathcal{H}}(0,1) \to \mathbb{C}$ is weakly continuous, i.e., continuous with respect to the product topology of the weak topologies of $\overline{B}_{\mathcal{K}}(0,1)$ and $\overline{B}_{\mathcal{H}}(0,1)$.

Remark 8.2.2. A completely continuous form $\omega : \mathcal{K} \times \mathcal{H} \to \mathbb{C}$ is clearly (norm-)continuous when restricted to $B_{\mathcal{K}}(0,1) \times B_{\mathcal{H}}(0,1)$, and hence is (norm-)continuous at $(0,0)$. Therefore, $\omega \in \mathfrak{Ses}(\mathcal{K}|\mathcal{H})$ by Prop. 2.3.10.

Proposition 8.2.3. *Let $\omega : \mathcal{H} \times \mathcal{K} \to \mathbb{C}$ be a completely continuous sesquilinear form. Then the adjoint form $\omega^* : \mathcal{K} \times \mathcal{H} \to \mathbb{C}$ is completely continuous.*[1]

Proof. This is obvious. □

To understand why complete continuity is a natural property, we must address two questions:

(1) In what situations can this property be easily verified?

(2) What theorems can be proved under this assumption?

We answer the first question in this section, and defer the second to Sec. 8.3.

8.2.2 The infinite matrix perspective

In Exp. 3.7.17, we observed that bounded matrices can be approximated in SOT by their finite-rank truncations. In this subsection, we shall see that completely continuous forms are precisely those that can be approximated *in the operator norm* by their finite-rank truncations; see Thm. 8.2.9. Some preparation is needed to prove this result.

Definition 8.2.4. A sesquilinear form $\omega : \mathcal{K} \times \mathcal{H} \to \mathbb{C}$ is called **finite rank** if ω is a finite sum (equivalently, a finite linear combination) of sesquilinear forms of the form

$$\boldsymbol{\omega_{\psi,\phi}} : \mathcal{K} \times \mathcal{H} \to \mathbb{C} \qquad (\eta, \xi) \mapsto \langle \eta|\psi\rangle\langle\phi|\xi\rangle \tag{8.10}$$

where $\psi \in \mathcal{K}$ and $\phi \in \mathcal{H}$.

[1] Recall from Def. 3.1.4 that $\omega^*(\eta|\xi) = \overline{\omega(\xi|\eta)}$

Theorem 8.2.5. *Any finite rank sesquilinear form $\omega : \mathcal{K} \times \mathcal{H} \to \mathbb{C}$ is completely continuous.*

Proof. Assume that ω is of the form (8.10). Since the linear functionals $\xi \in \mathcal{H} \mapsto \langle\phi|\xi\rangle$ and $\eta \in \mathcal{K} \mapsto \langle\psi|\eta\rangle$ are weakly continuous, ω is weakly continuous on the whole domain $\mathcal{K} \times \mathcal{H}$, and hence is completely continuous. □

Theorem 8.2.6. *Let $(\omega_\alpha)_{\alpha\in\mathscr{I}}$ be a net of completely continuous sesquilinear forms $\mathcal{K} \times \mathcal{H} \to \mathbb{C}$ converging in the operator norm to a sesquilinear form $\omega : \mathcal{K} \times \mathcal{H} \to \mathbb{C}$. Then ω is completely continuous.*

Proof. By assumption, when restricted to $\overline{B}_{\mathcal{K}}(0,1) \times \overline{B}_{\mathcal{H}}(0,1)$, the net (ω_α) converges uniformly to ω. Since each ω_α is weakly continuous on $\overline{B}_{\mathcal{K}}(0,1) \times \overline{B}_{\mathcal{H}}(0,1)$, and since the uniform limit of a net of continuous functions on a compact space X is continuous (that is, $C(X)$ is a l^∞-closed subset of $l^\infty(X)$), ω is weakly continuous on $\overline{B}_{\mathcal{K}}(0,1) \times \overline{B}_{\mathcal{H}}(0,1)$. Hence ω is completely continuous. □

Definition 8.2.7. A sesquilinear form $\omega : \mathcal{K} \times \mathcal{H} \to \mathbb{C}$ is called **approximable** if there is a net (equivalently, a sequence)[2] of completely continuous sesquilinear forms $\mathcal{K} \times \mathcal{H} \to \mathbb{C}$ converging in norm to ω.

Corollary 8.2.8. *Every approximable sesquilinear form on $\mathcal{K} \times \mathcal{H}$ is completely continuous.*

Proof. This follows immediately from Thm. 8.2.5 and 8.2.6. □

We are now ready to prove that a sesquilinear form is completely continuous iff it is the norm-limit of its finite-rank truncations.

Theorem 8.2.9. *Let $\omega : \mathcal{K} \times \mathcal{H} \to \mathbb{C}$ be a sesquilinear form. Let $(e_x)_{x\in X}$ and $(f_y)_{y\in Y}$ be orthonormal bases of $\mathcal{K}$ and $\mathcal{H}$ respectively. For each $I \subset \mathrm{fin}(2^X)$ and $J \in \mathrm{fin}(2^Y)$, define a finite-rank sesquilinear form on $\mathcal{K} \times \mathcal{H}$ by* [3]

$$\omega_{I,J} = \sum_{i\in I, j\in J} \omega(e_i|f_j) \cdot \omega_{e_i,f_j}$$

Then ω is completely continuous iff

$$\lim_{I\in\mathrm{fin}(2^X), J\in\mathrm{fin}(2^Y)} \|\omega - \omega_{I,J}\| = 0 \tag{8.11}$$

That condition (8.11) implies the complete continuity of ω follows immediately from Cor. 8.2.8. The converse direction is slightly more difficult to prove. Indeed, we will not use this harder implication logically in the course, although knowing the equivalence is helpful for motivation. Readers who prefer to skip the proof may safely omit the following subsection, except that Lem. 8.2.11 will be needed in Sec. 8.4 to establish the equivalence between completely continuous forms and operators.

[2]The equivalence is due to the fact that the norm topology of $\mathfrak{Ses}(\mathcal{K}|\mathcal{H})$ is metrizable.

[3]Here, the sesquilinear form ω_{e_i,f_j} is defined as in (8.10)

8.2.3 Proof of Thm. 8.2.9

***Proof of Thm.** 8.2.9.* The direction "$\Leftarrow$" follows directly from Cor. 8.2.8. Assume that ω is completely continuous. Let $P_I \in \mathfrak{L}(\mathcal{K})$ be the projection onto $\mathrm{Span}\{e_i : i \in I\}$. Let $Q_J \in \mathfrak{L}(\mathcal{H})$ be the projection onto $\mathrm{Span}\{f_j : j \in J\}$. Then $\omega_{I,J}(\eta|\xi) = \omega(P_I\eta|Q_J\xi)$. By Exp. 3.7.9, we have $\lim_I P_I = \mathbf{1}_{\mathcal{K}}$ and $\lim_J Q_J = \mathbf{1}_{\mathcal{H}}$ in SOT. The convergence $\lim_{I,J} \|\omega - \omega_{I,J}\| = 0$ follows from the following Lem. 8.2.10. □

Lemma 8.2.10. *Let $\omega : \mathcal{K} \times \mathcal{H} \to \mathbb{C}$ be a completely continuous sesquilinear form. Let $A \in \mathfrak{L}(\mathcal{K})$ and $B \in \mathfrak{L}(\mathcal{H})$. Let $(A_\alpha)_{\alpha\in\mathscr{I}}$ be a net in $\mathfrak{L}(\mathcal{K})$ whose adjoint converges in SOT to A^*. Let $(B_\alpha)_{\alpha\in\mathscr{I}}$ be a net in $\mathfrak{L}(\mathcal{H})$ whose adjoint converges in SOT to B^*. Assume that $\sup_\alpha \|A_\alpha\| < +\infty$ and $\sup_\alpha \|B_\alpha\| < +\infty$. Let*

$$\omega_\alpha : \mathcal{K} \times \mathcal{H} \to \mathbb{C} \qquad \omega_\alpha(\eta|\xi) = \omega(A_\alpha\eta|B_\alpha\xi)$$

Then $(\omega_\alpha)_{\alpha\in\mathscr{I}}$ converges in norm to the sesquilinear form ω_0 defined by

$$\omega_0 : \mathcal{K} \times \mathcal{H} \to \mathbb{C} \qquad \omega_0(\eta|\xi) = \omega(A\eta|B\xi)$$

Proof. Suppose that (ω_α) does not converge in norm to ω_0. Then there exists $\varepsilon > 0$ such that

$$I := \{\alpha \in \mathscr{I} : \|\omega_\alpha - \omega_0\| \geqslant 2\varepsilon\}$$

is a cofinal subset of $\mathscr{I}$. For each $\alpha \in I$, choose $\xi_\alpha \in \overline{B}_{\mathcal{H}}(0,1)$ and $\eta_\alpha \in \overline{B}_{\mathcal{K}}(0,1)$ such that

$$|\omega(A_\alpha\eta_\alpha|B_\alpha\xi_\alpha) - \omega(A\eta_\alpha|B\xi_\alpha)| \geqslant \varepsilon$$

By Thm. 3.7.5, there exists a subnet (η_ν, ξ_ν) of $(\eta_\alpha, \xi_\alpha)_{\alpha\in I}$ converging weakly to some $(\eta, \xi) \in \overline{B}_{\mathcal{K}}(0,1) \times \overline{B}_{\mathcal{H}}(0,1)$. Clearly $A\eta_\nu$ converges weakly to $A\eta$, and $B\xi_\nu$ converges weakly to $B\xi$. Therefore, by the complete continuity of ω we have

$$\lim_\nu \omega(A\eta_\nu|B\xi_\nu) = \omega(A\eta|B\xi)$$

Suppose we can prove that net $(B_\nu\xi_\nu)$ (which is bounded by the assumption $\sup_\alpha \|B_\alpha\| < +\infty$) converges weakly to $B\xi$, and similarly, the bounded net $(A_\nu\eta_\nu)$ converges weakly to $A\eta$. Then

$$\lim_\nu \omega(A_\nu\eta_\nu|B_\nu\xi_\nu) = \omega(A\eta|B\xi)$$

which contradicts the above inequality, and hence finishes the proof.

The weak convergence $\lim_\nu B_\nu\xi_\nu = B\xi$ follows from the observation that for any $\phi \in \mathcal{H}$ we have $\lim_\nu B_\nu^*\phi = B^*\phi$ (by the SOT convergence of B_ν^* to B^*), and hence

$$\lim_\nu \langle\phi|B_\nu\xi_\nu\rangle = \lim_\nu \langle B_\nu^*\phi|\xi_\nu\rangle \xlongequal{\text{Lem.8.2.11}} \langle B^*\phi|\xi\rangle = \langle\phi|B\xi\rangle$$

The weak convergence $\lim_\nu A_\nu\eta_\nu = A\eta$ is proved in the same way. □

The following lemma is used at the end of the proof of Lem. 8.2.10 and will also be needed in the proof of Prop. 8.4.3.

Lemma 8.2.11. *Let $(\phi_\alpha)_{\alpha\in I}$ be a net in $\overline{B}_{\mathcal{H}}(0,1)$ converging weakly to $\phi \in \overline{B}_{\mathcal{H}}(0,1)$. Then the following are equivalent.*

(1) $(\phi_\alpha)_{\alpha\in I}$ converges to ϕ.

(2) For each net $(\xi_\alpha)_{\alpha\in I}$ in $\overline{B}_{\mathcal{H}}(0,1)$ converging weakly to some $\xi \in \overline{B}_{\mathcal{H}}(0,1)$ we have $\lim_\alpha \langle\phi_\alpha|\xi_\alpha\rangle = \langle\phi|\xi\rangle$.

Proof. Assume (1). Then

$$|\langle\phi|\xi\rangle - \langle\phi_\alpha|\xi_\alpha\rangle| \leqslant |\langle\phi|\xi-\xi_\alpha\rangle| + |\langle\phi-\phi_\alpha|\xi_\alpha\rangle| \leqslant |\langle\phi|\xi-\xi_\alpha\rangle| + \|\phi-\phi_\alpha\|$$

where the RHS converges to 0. This proves (2). Conversely, assume (2). Then $\lim_\alpha\langle\xi_\alpha|\xi_\alpha\rangle = \langle\xi|\xi\rangle$, and hence (1) holds by Prop. 3.7.3. □

8.2.4 The L^2-kernel perspective

It is purely a matter of good fortune that the integral kernel $K : I \times I \to \mathbb{C}$ in (8.9), which arises from the Dirichlet problem in $\mathbb{R}^2$, happens to be a continuous function, and hence L^2-integrable. The L^2-integrability fails for kernels arising from Dirichlet problems in dimensions $N \geqslant 3$, where the function $|x-y|^{2-N}$ is not L^2-integrable on $M \times M$ for any (nonempty) hypersurfaces $M \subset \mathbb{R}^N$. In these higher-dimensional settings, the associated integral operators T_K remain compact, as we will see in Pb. 8.6. However, for such non L^2-integrable kernels, it was far more difficult for early researchers such as Hilbert to recognize and formulate the notion of complete continuity.

The advantage of the L^2-integrability of K is that the matrix representation $[T_K]$ of T_K is l^2-summable; in fact, by Prop. 8.2.12, the L^2-norm of K coincides with the l^2-norm of $[T_K]$. It is not hard to show that the operator norm of any l^2-summable matrix is bounded by its l^2-norm. Consequently, the finite-rank truncations of $[T_K]$ converge in operator norm to $[T_K]$, proving that T_K is completely continuous. In the following, we translate this argument to the setting of sesquilinear forms. See [Gui-A, Sec. 22.2-22.4] for a direct matrix-theoretic approach.

In the remainder of this subsection, we assume that X and Y are LCH spaces, and that μ and ν are σ-finite Radon measures on the Borel σ-algebras of X and Y, respectively. [4] The σ-finiteness is needed for the purpose of using the theorems of Tonelli and Fubini, cf. Subsec. 1.7.5. Recall Def. 1.7.19 for the definition of $\mu \times \nu$. Here, we assume that $\mu \times \nu$ is defined on the Borel σ-algebra of $X \times Y$.

[4] For simplification, readers may assume that X and Y are closed or open subsets of Euclidean spaces, and μ, ν are Lebesgue measures.

Proposition 8.2.12. *Let $(e_i)_{i\in\mathscr{A}}$ be an orthonormal basis of $L^2(X,\mu)$, and let $(f_j)_{j\in\mathscr{B}}$ be an orthonormal basis of $L^2(Y,\nu)$. Define*

$$e_if_j : X\times Y\to\mathbb{C}\qquad (x,y)\mapsto e_i(x)f_j(y)$$

Then $(e_if_j)_{i\in\mathscr{A},j\in\mathscr{B}}$ is an orthonormal basis of $L^2(X\times Y,\mu\times\nu)$.

Since e_i and f_j are Borel functions on X and Y respectively, e_if_j is a Borel function on $X\times Y$.

Proof. It is easy to check that (e_if_j) is an orthonormal family. Thus, it remains to show that this family spans a dense linear subspace of $L^2(X\times Y,\mu\times\nu)$. By Thm. 1.7.10, $C_c(X\times Y)$ is dense in $L^2(X\times Y,\mu\times\nu)$. Therefore, it suffices to show that each $\varphi\in C_c(X\times Y)$ can be L^2-approximated by linear combinations of (e_if_j).

Since $\mathrm{Supp}(\varphi)$ is compact, its projections onto X and Y are both compact, and hence have finite measures. By the outer regularity of Radon measures (cf. Def. 1.7.3), there exist open subsets $U\subset X$ and $V\subset Y$ such that $\mu(U)<+\infty$, $\nu(V)<+\infty$, and $\mathrm{Supp}(\varphi)\subset U\times V$. By the Stone-Weierstrass Thm. 1.5.13 (applied to $U\times V$), φ can be approximated uniformly (and hence in $L^2(X\times Y,\mu\times\nu)$) by linear combinations of functions of the form $\alpha\beta$ where $\alpha\in C_c(U)$ and $\beta\in C_c(V)$.

Since (e_i) is an orthonormal basis of $L^2(X,\mu)$, α can be L^2-approximated by linear combinations of (e_i). Similarly, β can be L^2-approximated by linear combinations of (f_j). This completes the proof. □

Theorem 8.2.13. *Let $K\in L^2(X\times Y,\mu\times\nu)$. Then the sesquilinear form*

$$\begin{gathered}\omega : L^2(X,\mu)\times L^2(Y,\nu)\to\mathbb{C}\\ \omega(f|g)=\int_{X\times Y}\overline{f(x)}K(x,y)g(y)d(\mu\times\nu)(x,y)\end{gathered}\tag{8.12}$$

is completely continuous, and $\lVert\omega\rVert\leqslant\lVert K\rVert_{L^2}$.

Note that by the theorems of Tonelli and Fubini, we have

$$\omega(f|g)=\int_X\int_Y\overline{f}Kgd\nu d\mu=\int_Y\int_X\overline{f}Kgd\mu d\nu$$

Proof. For each $f\in L^2(X,\mu)$ and $g\in L^2(Y,\nu)$, by Fubini-Tonelli we have

$$\int_{X\times Y}|f(x)g(y)|^2d(\mu\times\nu)=\int_X\int_Y|f(x)g(y)|^2d\nu d\mu=\lVert f\rVert_{L^2}^2\cdot\lVert g\rVert_{L^2}^2$$

In particular, we have $fg\in L^2(X\times Y,\mu\times\nu)$, and hence the integral defining $\omega(f|g)$ makes sense. Moreover, by Cauchy Schwarz, we have

$$|\omega(f|g)|\leqslant\lVert f\rVert_{L^2}\cdot\lVert g\rVert_{L^2}\cdot\lVert K\rVert_{L^2}$$

This proves $\|\omega\| \leqslant \|K\|_{L^2}$.

Let $(e_i)_{i\in\mathscr{A}}$ be an orthonormal basis of $L^2(X,\mu)$, and let $(f_j)_{j\in\mathscr{B}}$ be an orthonormal basis of $L^2(Y,\nu)$. By Prop. 8.2.12, $(e_if_j)_{i\in\mathscr{A},j\in\mathscr{B}}$ is an orthonormal basis of $L^2(X\times Y,\mu\times\nu)$. Therefore, Thm. 3.3.24 implies

$$\lim_{I\in\mathrm{fin}(2^{\mathscr{A}}),J\in\mathrm{fin}(2^{\mathscr{B}})}\|K-K_{I,J}\|_{L^2}^2=0$$

where $K_{I,J}$ is the projection of K onto the subspace of $L^2(X\times Y,\mu\times\nu)$ spanned by $(e_if_j)_{i\in I,f\in J}$. Let $\omega_{I,J}$ be the sesquilinear form defined by $K_{I,J}$ instead of K. Then, by the above paragraph, we have

$$\lim_{I,J}\|\omega-\omega_{I,J}\|\leqslant\lim_{I,J}\|K-K_{I,J}\|_{L^2}=0$$

Since each $\omega_{I,J}$ has finite rank, by Cor. 8.2.8, ω is completely continuous. □

8.3 Hilbert-Schmidt theorem: the form perspective

In this section, we address the second question posed in Sec. 8.2.1 by proving that every completely continuous Hermitian form admits a diagonalization. Incidentally, this theorem also yields the spectral theorem for self-adjoint operators on finite-dimensional Hilbert spaces. This finite-dimensional result is in fact used in our proof of the spectral theorem for general bounded self-adjoint operators on Hilbert spaces; see the proof of Thm. 5.5.17.

Fix a Hilbert space $\mathcal{H}$. Recall Def. 8.2.4 for the meaning of $\omega_{\psi,\phi}$.

8.3.1 The Hilbert-Schmidt theorem

Theorem 8.3.1 (Hilbert-Schmidt theorem). *Let $T\in\mathfrak{L}(\mathcal{H})$. Then the following statements are equivalent.*

(1) The associated sesquilinear form $\omega_T:\mathcal{H}\times\mathcal{H}\to\mathbb{C}$ (defined by $\omega_T(\xi|\eta)=\langle\xi|T\eta\rangle$) is Hermitian and completely continuous.

(2) $\mathcal{H}$ admits an orthonormal basis $(e_1,e_2,\dots)\cup(f_j)_{j\in J}$, where the countable family $(e_1,e_2,\dots)$ is possibly finite, such that:

(2a) For each n, we have $Te_n=\lambda_ne_n$ for some $\lambda_n\in\mathbb{R}$, and $Tf_j=0$ for each j.

(2b) If the sequence $(e_1,e_2,\dots)$ is infinite, then $\lim_{n\to\infty}\lambda_n=0$.

Proof. Step 1. In this step, we prove (2)⇒(1). Assume (2). We treat the case where $(e_1,e_2,\dots)$ is infinite, as the finite case is similar and simpler. Define $T_n\in\mathfrak{L}(\mathcal{H})$ by

$$T_n=\sum_{k=1}^n\lambda_kP_{e_k}\tag{8.13}$$

where P_{e_k} is the orthogonal projection onto $\mathrm{Span}\{e_k\}$. Then T_n is self-adjoint (since each P_{e_k} is self-adjoint), and $\lim_n T_n\xi = T\xi$ for each ξ in the union of $(e_1, e_2, \dots)$ and $(f_j)_{j\in J}$. Since $\sup_n \|T_n\| \leqslant \sup_n |\lambda_n| < +\infty$, Prop. 2.4.5 implies that (T_n) converges to T in the SOT.

The norm of

$$T - T_n = \sum_{k=n+1}^{\infty} \lambda_k P_{e_k}$$

is bounded by $\sup_{k\geqslant n+1} |\lambda_k|$, and hence converges to zero as $n \to \infty$. We have thus proved that (T_n) converges in norm to T. Therefore, since each T_n is self-adjoint, T is self-adjoint. Since ω_{T_n} has finite rank, ω_T is approximable, and hence completely continuous by Cor. 8.2.8. This proves (1).

Step 2. In the following steps, we prove (1)⇒(2). We treat only the case in which $\mathcal{H}$ is infinite-dimensional; the finite-dimensional case follows by a similar (and easier) method.

Since ω_T is completely continuous, the function

$$f : \overline{B}_{\mathcal{H}}(0,1) \to \mathbb{R} \qquad f(\xi) = \omega_T(\xi|\xi) = \langle \xi|T\xi\rangle$$

is weakly continuous. Since $\overline{B}_{\mathcal{H}}(0,1)$ is weakly compact (Thm. 3.7.5), $|f|$ attains its maximum at some $e_1 \in \overline{B}_{\mathcal{H}}(0,1)$. Moreover, we may assume that $\|e_1\| = 1$; otherwise, replacing e_1 with $e_1/\|e_1\|$ will enlarge the value of $|f|$.

We now show that e_1 is an eigenvector of T with eigenvalue $\lambda_1 := f(e_1)$. In the case where $\lambda_1 \geqslant 0$, f achieves its maximum at λ_1, and hence for each unit vector $\xi \in \mathcal{H}$ we have

$$\langle \xi|T\xi\rangle \leqslant \lambda_1 = \lambda_1\|\xi\|^2$$

By sesquilinearity, the inequality $\langle \xi|T\xi\rangle \leqslant \lambda_1\|\xi\|^2$ holds for all $\xi \in \mathcal{H}$. Therefore, the sesquilinear form $\omega_{\lambda_1 - T}$ associated to $\lambda_1 - T$ is positive. Moreover, since $f(e_1) = \lambda_1$, we have

$$\omega_{\lambda_1 - T}(e_1|e_1) = 0$$

It follows from Cor. 3.1.10 that for each $\xi \in \mathcal{H}$ we have $\omega_{\lambda_1-T}(\xi|e_1) = 0$, namely $\langle \xi|(\lambda_1 - T)e_1\rangle = 0$. Therefore $Te_1 = \lambda_1 e_1$. In the case where $\lambda_1 \leqslant 0$, replacing T with $-T$ yields again the fact that $Te_1 = \lambda_1 e_1$.

To summarize, for each self-adjoint $T \in \mathfrak{L}(\mathcal{H})$ whose associated Hermitian form is completely continuous, there exists a unit vector $e_1 \in \mathcal{H}$ satisfying $Te_1 = \lambda_1 e_1$ where $\lambda_1 \in \mathbb{R}$, and

$$\langle \xi|T\xi\rangle \leqslant |\lambda_1| \cdot \|\xi\|^2 \qquad \text{for each } \xi \in \mathcal{H}$$

Step 3. Set

$$V_n = \mathrm{Span}\{e_1, \ldots, e_n\}$$

By repeating the process in Step 2, we obtain an orthonormal sesquence $(e_n)_{n\in\mathbb{Z}_+}$ in $\mathcal{H}$ such that $Te_n = \lambda_n e_n$ for each n, and that

$$\langle \xi | T\xi \rangle \leqslant |\lambda_n| \cdot \|\xi\|^2 \qquad \text{for each } \xi \in V_{n-1}^\perp \tag{8.14}$$

Indeed, suppose $e_1, \ldots, e_{n-1}$ have been constructed. Note that V_{n-1} is T-invariant, and hence $V_{n-1}^\perp$ is T-invariant due to Exp. 5.10.8. The vector e_n is obtained by applying Step 2 to $T|_{V_{n-1}^\perp}$, noting that the sesquilinear form associated to $T|_{V_{n-1}^\perp}$ is clearly Hermitian and completely continuous.

Step 4. Since the infinite sequence (e_n) is orthonormal, it converges weakly to 0 by Exp. 3.7.4. Therefore, since $\lambda_n = \langle e_n | Te_n \rangle = \omega_T(e_n|e_n)$, and since ω_T is completely continuous, we obtain $\lim_n \lambda_n = 0$.

By (8.14) and Prop. 3.2.12, we have $\|T|_{V_{n-1}^\perp}\| \leqslant 4|\lambda_n|$. By Exe. 3.4.10, $(e_n)_{n\in\mathbb{Z}_+}$ can be extended to an orthonormal basis $(e_n)_{n\in\mathbb{Z}_+} \cup (f_j)_{j\in J}$ of $\mathcal{H}$. Since each f_j belongs to $V_{n-1}^\perp$ for all n, we have $\|Tf_j\| \leqslant 4|\lambda_n|$ for all n, and hence $Tf_j = 0$. This proves (2). □

8.3.2 Comments on the proof

In the proof of the Hilbert-Schmidt theorem we made use of linear operators—primarily in Steps 3 and 4—to streamline the inductive construction of the sequence of eigenvectors arising from the argument in Step 2. Nevertheless, the essential heart of the proof lies in Step 2, where we apply the **maximum-minimum method** (a term used in the first quotation from von Neumann in Subsec. 6.11.2) to obtain an eigenvector corresponding to the largest eigenvalue. This method belongs to the paradigm of sesquilinear forms rather than linear operators, and indeed it is Hilbert's original method in [Hil06].

This maximum-minimum method is particularly transparent, as it makes explicit how complete continuity enters the argument: combined with the weak compactness of $\overline{B}_{\mathcal{H}}(0,1)$, it ensures that the quadratic form

$$\overline{B}_{\mathcal{H}}(0,1) \to \mathbb{R}_{\geqslant 0} \qquad \xi \mapsto \omega_T(\xi|\xi)$$

attains both its maximum and its minimum. It was precisely for the purpose of applying this maximum-minimum method that Hilbert introduced the notion of weak convergence in [Hil06], the first explicit appearance of weak/weak-* convergence in the history of mathematics. In contrast, the proof we will see in Sec. 8.5, which relies solely on linear operator methods, is considerably less straightforward.

8.4 Completely continuous operators and compact operators

Let $\mathcal{H}, \mathcal{K}$ be Hilbert spaces. Let $\mathcal{V}, \mathcal{W}$ be normed vector spaces over $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$.

In [Rie10], Riesz introduced the notion of completely continuous operators on Hilbert spaces, a concept that is equivalent to the completely continuous sesquilinear forms considered earlier by Hilbert. Riesz was an early and influential advocate of the operator viewpoint, and it was natural for him to reinterpret Hilbert's results in operator language, in accordance with his own mathematical inclinations. In Ch. 5, we already saw how Riesz's operator-theoretic interpretation of Hilbert's spectral theory for bounded bilinear forms had a profound and lasting influence on the development of functional analysis. In the following sections, we will see that Riesz's reinterpretation of Hilbert's theory of completely continuous forms was equally revolutionary.

The operator perspective frees us from two significant limitations in Hilbert's original framework: the requirement that the form be Hermitian (which was essential for applying the maximum-minimum method in the proof of the Hilbert-Schmidt theorem), and the restriction to Hilbert spaces rather than more general Banach spaces. Riesz's proof of the Fredholm alternative for compact operators on Banach spaces in [Rie18] stands as a landmark achievement that emerged from this paradigm shift.

In this section, we aim to translate the notion of a completely continuous sesquilinear form into the corresponding operator-theoretic condition on normed vector spaces—namely, that of a compact operator. We will also compare the strategies used to establish key properties from the perspectives of forms and operators. We begin with the case of operators on Hilbert spaces, for which the natural notion is that of a completely continuous operator.

8.4.1 Completely continuous operators

Definition 8.4.1. A linear map $T : \mathcal{H} \to \mathcal{K}$ is called **completely continuous** if the restriction

$$T|_{\overline{B}_{\mathcal{H}}(0,1)} : \overline{B}_{\mathcal{H}}(0,1) \to \mathcal{K}$$

is continuous when $\overline{B}_{\mathcal{H}}(0,1)$ is equipped with the weak topology and $\mathcal{H}$ is equipped with the norm topology. Similar to Rem. 8.2.2, any completely continuous map is bounded.

It is helpful to compare this definition with the following property.

Proposition 8.4.2. *Let $T \in \mathfrak{L}(\mathcal{H}, \mathcal{K})$. Then $T : \mathcal{H} \to \mathcal{K}$ is continuous when both $\mathcal{H}$ and $\mathcal{K}$ are equipped with their weak topologies.*

Proof. Let (ξ_α) be a net in $\mathcal{H}$ converging weakly to ξ. Then for each $\eta \in \mathcal{K}$ we have

$$\lim_\alpha \langle \eta | T\xi_\alpha \rangle = \lim_\alpha \langle T^*\eta | \xi_\alpha \rangle = \langle T^*\eta | \xi \rangle = \langle \eta | T\xi \rangle$$

Therefore $T\xi_\alpha$ converges weakly to $T\xi$. $\square$

We now prove the equivalence between completely continuous forms and completely continuous operators.

Proposition 8.4.3. *Let $T : \mathcal{H} \to \mathcal{K}$ be a linear map. Then T is completely continuous iff its associated sesquilinear form $\omega_T : \mathcal{K} \times \mathcal{H} \to \mathbb{C}$ (cf. Def. 3.5.1) is completely continuous.*

Proof. T is completely continuous iff for each weak convergence $\xi_\alpha \to \xi$ in $\overline{B}_{\mathcal{H}}(0,1)$, the weak convergence $T\xi_\alpha \to T\xi$ (due to Prop. 8.4.2) must be a norm convergence. By Lem. 8.2.11, this is equivalent to $\langle\eta_\alpha|T\xi_\alpha\rangle \to \langle\eta|T\xi\rangle$ for each net $(\xi_\alpha)_{\alpha\in I}$ converging to ξ in $\overline{B}_{\mathcal{H}}(0,1)$ and each net $(\eta_\alpha)_{\alpha\in I}$ converging to η in $\overline{B}_{\mathcal{K}}(0,1)$, and hence equivalent to the complete continuity of ω_T. $\square$

Of course, completely continuous forms and completely continuous operators can be unified under the broader notion of a completely continuous multilinear map: a multilinear map $T : \mathcal{H}_1 \times \cdots \times \mathcal{H}_N \to \mathcal{K}$ (where each $\mathcal{H}_j$ is a Hilbert space) is called **completely continuous** if its restriction to $\overline{B}_{\mathcal{H}_1}(0,1) \times \cdots \times \overline{B}_{\mathcal{H}_N}(0,1) \to \mathcal{K}$ is continuous when the domain carries the weak topology and $\mathcal{K}$ carries the norm topology. We refrain from adopting this general formulation, since such a general notion will not be used later.

8.4.2 Compact operators

The notion of complete continuity can be extended to linear maps $T : \mathcal{V}^* \to \mathcal{W}$, meaning that the restriction

$$T|_{\overline{B}_{\mathcal{V}^*}(0,1)} : \overline{B}_{\mathcal{V}^*}(0,1) \to \mathcal{W}$$

is continuous when $\overline{B}_{\mathcal{V}^*}(0,1)$ is equipped with the weak-* topology and $\mathcal{W}$ with the norm topology. However, even this generalization does not encompass linear operators acting on spaces of continuous functions on compact sets—operators that play a crucial role in solving the Dirichlet problem via double layer potentials (cf. Sec. 8.1). This motivates the need for the following definition.

Definition 8.4.4. A linear map $T : \mathcal{V} \to \mathcal{W}$ is called a **compact operator** (or a **compact map**) if one of the following equivalent conditions hold:

(1) $T(\overline{B}_{\mathcal{V}}(0,1))$ is precompact in $\mathcal{W}$ (with respect to the norm topology).

(2) For each net (ξ_α) in $\mathcal{H}$ satisfying $\sup_\alpha \|\xi_\alpha\| < +\infty$, the net $(T\xi_\alpha)$ has a subnet converging in $\mathcal{W}$.

(3) For each sequence (ξ_n) in $\mathcal{H}$ satisfying $\sup_n \|\xi_n\| < +\infty$, the sequence $(T\xi_n)$ has a subsequence converging in $\mathcal{W}$.

The equivalence is due to Prop. 1.4.22.

Remark 8.4.5. Since any compact subset of a metric space is bounded (since it can be covered by finitely many open balls), every compact operator must be bounded.

Exercise 8.4.6. Suppose that $S, T : \mathcal{V} \to \mathcal{W}$ are compact operators, and $a, b \in \mathbb{F}$. Show that $aS + bT$ is a compact operator.

Theorem 8.4.7. *Let $T : \mathcal{H} \to \mathcal{K}$ be a linear map. The following are equivalent.*

(1) The associated sesquilinear form ω_T (cf. Def. 3.5.1) is completely continuous.

(2) T is completely continuous.

(3) T is compact.

Proof. (1)⇔(2): By Prop. 8.4.3.

(2)⇒(3): Let (ξ_α) be a bounded net in $\mathcal{H}$. By Thm. 3.7.5, it has a subnet (ξ_ν) converging weakly to some $\xi \in \mathcal{H}$. Since T is completely continuous, the subnet $(T\xi_\nu)$ of $(T\xi_\alpha)$ converges (in norm) to $T\xi$. Thus T is compact. (Or simply: the image of any compact set under a continuous map is compact, and hence precompact.)

(3)⇒(2). Let $X = \overline{B}_{\mathcal{H}}(0, 1)$. Let (ξ_α) be a net in X converging weakly to $\xi \in X$. Since T is bounded (Rem. 8.4.5), by Prop. 8.4.2, $(T\xi_\alpha)$ converges weakly to $T\xi$. Therefore, any subnet of $(T\xi_\alpha)$ converges weakly to $T\xi$. In particular, any norm-convergent subnet of $(T\xi_\alpha)$ converges weakly to $T\xi$, and hence converges in norm to $T\xi$.

Since $Y := \overline{T(X)}$ is compact (because T is compact) and contains the net $(T\xi_\alpha)$ and the element $T\xi$, it follows from Thm. 1.3.9 that $(T\xi_\alpha)$ converges in norm to $T\xi$. This proves (2). □

In the above proof, the direction (2)⇒(3) is more straightforward than (3)⇒(2). Consequently, translating a proof about compact operators into one about completely continuous operators is typically easier than working in the opposite direction.

It is not immediately clear why the adjoint of a compact operator must also be compact. However, with Thm. 8.4.7 and the elementary fact that the adjoint of a completely continuous sesquilinear form remains completely continuous (Prop. 8.2.3), we obtain the following result at once:

Corollary 8.4.8. *Let $T : \mathcal{H} \to \mathcal{K}$ be a compact linear map. Then $T^* : \mathcal{K} \to \mathcal{H}$ is compact.*

8.4.3 Comments on the definition of compact operators

Remark 8.4.9. The reader may notice a key difference between the notions of completely continuous forms/operators and compact operators. The former is purely a continuity condition and, by itself, has nothing to do with compactness. Compactness enters only through applications—for instance, in the proof of the Hilbert-Schmidt theorem, where the maximum-minimum method relies on compactness.

In contrast, the notion of a compact operator is not a continuity condition but a compactness condition. This is natural to understand, since compact operators were introduced to extend the notion of complete continuity to normed vector spaces that lack a duality structure (and hence lack weak-* compactness). When the underlying spaces no longer provide compactness, compactness must instead be built into the definition of the operators themselves. □

Remark 8.4.10. The proof of (2)⇒(3) in Thm. 8.4.7 actually shows that if $T:\mathcal{H}\rightarrow\mathcal{K}$ is completely continuous, then $T(\overline{B}_{\mathcal{H}}(0,1))$ is compact (rather than merely precompact). However, defining compact operators on Banach spaces by requiring the image of the closed unit ball to be compact would be too restrictive: important results such as Thm. 8.4.15 could not be established under such a strong definition. In particular, one could not prove with this definition the crucial example (cf. Pb. 8.4) that every continuous kernel $K: I\times I\rightarrow\mathbb{C}$ (where I is a compact interval) defines a compact operator on $C(I)$.

8.4.4 Criteria for compact operators

We now extend several results from Subsec. 8.2.2, originally proved for completely continuous sesquilinear forms, to the setting of compact operators. The first such result, generalizing Thm. 8.2.5, concerns finite-rank forms and operators.

Definition 8.4.11. A linear map $T:\mathcal{V}\rightarrow\mathcal{W}$ is called of **finite rank** if T is bounded and $\dim\mathrm{Rng}(T)<+\infty$.

Proposition 8.4.12. *Let $T:\mathcal{H}\rightarrow\mathcal{K}$ be a linear map. The following are equivalent.*

(1) The associated sesquilinear form ω_T has finite rank.

(2) There exist finitely many vectors $\phi_1,\dots,\phi_n\in\mathcal{H}$ and $\psi_1,\dots,\psi_n\in\mathcal{K}$ such that

$$T\xi=\sum_{j=1}^n\psi_j\cdot\langle\phi_j|\xi\rangle\qquad\text{for each }\xi\in\mathcal{H}\tag{8.15}$$

(3) T has finite rank.

Proof. (1)⇔(2) and (2)⇒(3) are obvious.

(3)⇒(2): Assume (3). Let $\psi_1, \dots, \psi_n$ be an orthonormal basis of the finite-dimensional Hilbert space $\mathrm{Rng}(T)$. By Thm. 3.3.24,

$$T\xi = \sum_{j=1}^n \psi_j \cdot \langle \psi_j | T\xi \rangle = \sum_{j=1}^n \psi_j \cdot \langle T^*\psi_j | \xi \rangle$$

Thus (8.15) holds with $\phi_j = T^*\psi_j$. □

Theorem 8.4.13. *Let $T : \mathcal{V} \to \mathcal{W}$ be a linear operator of finite-rank. Then T is compact.*

First proof for Hilbert spaces. Assume that $\mathcal{V}, \mathcal{W}$ are Hilbert spaces. By Prop. 8.4.12, ω_T is of finite rank, and hence is completely continuous due to Thm. 8.2.5. (Alternatively, by (8.15), T is completely continuous.) Therefore, by Thm. 8.4.7, T is compact. □

Second proof. Since T is bounded, the set $T(\overline{B}_{\mathcal{V}}(0,1))$ is bounded. By the following Thm. 8.4.14, every bounded closed subset of $\mathrm{Rng}(T)$ is compact. Thus $T(\overline{B}_{\mathcal{V}}(0,1))$ is precompact, and hence T is compact. □

Theorem 8.4.14. *Let $\mathcal{X}$ be a finite-dimensional $\mathbb{F}$-vector space. Then any two norms on $\mathcal{X}$ are equivalent (cf. Def. 7.2.14). Consequently, any bounded closed subset of $\mathcal{X}$ is compact.*

In particular, every finite-dimensional normed $\mathbb{F}$-vector space V is complete, since its norm is equivalent to the pushforward of the Euclidean norm under any linear isomorphism $\Phi : \mathbb{F}^N \to V$.

Proof. Fix a linear bijection $\Phi : \mathbb{F}^N \to \mathcal{X}$ where $\mathbb{F}^N$ is equipped with the Euclidean norm. If we show that both Φ and Φ^{-1} are bounded with respect to any norm of $\mathcal{X}$, then every norm on $\mathcal{X}$ is equivalent to the pushforward of the Euclidean norm under Φ; hence any two norms on $\mathcal{X}$ are equivalent. The compactness statement then follows from the Heine-Borel theorem.

We first note that Φ is bounded, since for each $a = (a_1, \dots, a_N) \in \mathbb{F}^N$ we have

$$\|\Phi(a_1, \dots, a_N)\| = \Big\| \sum_{j=1}^N a_j \Phi(e_j) \Big\| \leqslant \|a\|_{l^\infty} \cdot \sum_j \|\Phi(e_j)\| \leqslant \|a\|_{l^2} \cdot \sum_j \|\Phi(e_j)\|$$

where $e_1, \dots, e_N$ are the standard basis of $\mathbb{F}^N$.

Suppose that Φ^{-1} is not bounded. Then for each $n \in \mathbb{Z}_+$ there exists a nonzero $v_n \in \mathbb{F}^N$ such that $\|v_n\| = \|\Phi^{-1}(\Phi(v_n))\| \geqslant n\|\Phi(v_n)\|$. By scaling each v_n we assume $\|v_n\| = 1$. Then $\|\Phi(v_n)\| \leqslant 1/n$, and hence any subsequence of $(\Phi(v_n))$ converges to 0. Since the unit sphere of $\mathbb{R}^N$ is compact, (v_n) has a subsequence (v_{n_k}) converging to some $v \in \mathbb{R}^N$ with $\|v\| = 1$. By the continuity of Φ we have $\lim_k \Phi(v_{n_k}) = \Phi(v)$, where $\Phi(v)$ is nonzero because Φ is bijective and v is nonzero. This contradiction shows that Φ^{-1} must be bounded. □

The following generalization of Thm. 8.2.6 is crucial for proving that the integral operator arising from the double layer potentials (cf. Subsec. 8.1.3) is compact; see Pb. 8.6.

Theorem 8.4.15. *Let $(T_\alpha)_{\alpha\in\mathscr{I}}$ be a net of compact operators $\mathcal{V}\to\mathcal{W}$ converging in the operator norm to a linear map $T:\mathcal{V}\to\mathcal{W}$. Assume that $\mathcal{W}$ is complete. Then T is compact.*

First proof for Hilbert spaces. Assume that $\mathcal{V},\mathcal{W}$ are Hilbert spaces. Then ω_T is the norm limit of the net (ω_{T_α}) of completely continuous forms, and hence is completely continuous by Thm. 8.2.6. Therefore, T is compact. □

Second proof for Hilbert spaces. Assume that $\mathcal{V},\mathcal{W}$ are Hilbert spaces. Let $X=\overline{B}_{\mathcal{H}}(0,1)$, equipped with the weak topology. By assumption, the net $(T_\alpha|_X)$ converges uniformly to $T|_X$. Since each $T_\alpha|_X$ is continuous, $T|_X$ must be continuous. Therefore T is completely continuous, i.e. compact. □

Third proof. Since our task is equivalently to show that the set of compact operators is closed in the metric space $\mathfrak{L}(\mathcal{V},\mathcal{W})$, Rem. 1.2.23 allows us to restrict attention to the case where (T_α) is a sequence $(T_n)_{n\in\mathbb{Z}_+}$.

Let $X=\overline{B}_{\mathcal{V}}(0,1)$ and $Y_n=\overline{T_n(X)}$. Since T_n is compact, each Y_n is a compact metric space. Therefore, by the diagonal method (cf. Rem. 1.4.18), the product space $\prod_m Y_m$ is sequentially compact.

Let (ξ_n) be a sequence in X. For each n, let

$$\mathfrak{y}_n=(T_1\xi_n,T_2\xi_n,\dots)\qquad\in\prod_m Y_m$$

Then the sequence $(\mathfrak{y}_n)$ has a convergent subsequence $(\mathfrak{y}_{n_k})$. Therefore, for each $m\in\mathbb{Z}_+$, the sequence $(T_m\xi_{n_k})_{k\in\mathbb{Z}_+}$ converges, and hence is a Cauchy sequence. From the convergence $\lim_m\|T_m-T\|=0$ one easily deduces that $(T\xi_{n_k})_{k\in\mathbb{Z}_+}$ is a Cauchy sequence, and hence converges by the completeness of $\mathcal{W}$. We have thus proved that $(T\xi_n)$ admits a convergent subsequence. Therefore T is compact. □

Definition 8.4.16. A linear map $T:\mathcal{V}\to\mathcal{W}$ is called **approximable** if it lies in the norm-closure of the set of finite-rank operators in $\mathfrak{L}(\mathcal{V},\mathcal{W})$. Equivalently, T is the limit of a norm convergent net/sequence of finite-rank operators.

Corollary 8.4.17. *Assume that $T:\mathcal{V}\to\mathcal{W}$ is an approximable linear map and $\mathcal{W}$ is complete. Then T is compact.*

Proof. This follows immediately from Thm. 8.4.13 and 8.4.15. □

8.5 Hilbert-Schmidt theorem: the operator perspective

Let $\mathcal{H}$ be a Hilbert space.

In this section, following Riesz's approach in [Rie10, §14], we translate Hilbert's proof of the Hilbert-Schmidt theorem—originally formulated in the language of sesquilinear forms and the maximum-minimum method—into the language of operators. From a historical perspective, Riesz's contribution here was transitional, serving as a reinterpretation of Hilbert's existing work. Moreover, its impact was less revolutionary than Riesz's operator-theoretic interpretation of Hilbert's spectral theorem for bounded symmetric bilinear forms, as presented and discussed in Ch. 5.

Nevertheless, this contribution should not be underestimated. Many highly original advances in mathematics begin as reinterpretations of earlier ideas. Without understanding the techniques developed in such transitional works, one cannot fully appreciate the deeper innovations that followed—such as Riesz's proof of the Fredholm alternative for compact operators on arbitrary Banach spaces.

8.5.1 Riesz's method: achieving approximate eigenvalues

Let $T \in \mathfrak{L}(\mathcal{H})$ be completely continuous and self-adjoint. As we saw in Sec. 8.3, to find the unit eigenvector e_1 whose eigenvalue λ_1 has the largest absolute value $|\lambda_1|$, one chooses $e_1 \in \overline{B}_{\mathcal{H}}(0,1)$ maximizing the quantity $|\langle e_1|Te_1\rangle|$; the associated eigenvalue is then $\lambda_1 = \langle e_1|Te_1\rangle$. In the operator setting, the natural analogue would be to choose $e_1 \in \overline{B}_{\mathcal{H}}(0,1)$ maximizing $\|Te_1\|$, so that $\|Te_1\| = \|T\|$. Unfortunately, such a vector need not be an eigenvector of T:

Example 8.5.1. Let T be the diagonal matrix on $\mathbb{C}^2$ with diagonal entries 1 and -1. Then $e_1 = \frac{1}{\sqrt{2}}(1,1)$ is a vector in $\overline{B}_{\mathcal{H}}(0,1)$, and $\|Te_1\| = 1 = \|T\|$. However, e_1 is not an eigenvector of T.

Instead of employing the maximum-minimum argument, Riesz worked with approximate eigenvalues. To motivate this approach, consider the following elementary fact:

Exercise 8.5.2. Let $A \in \mathfrak{L}(\mathcal{H})$ be normal. Use the spectral theorem to prove that

$$\|A\| = \sup_{\lambda \in \mathrm{Sp}(A)} |\lambda| \tag{8.16}$$

(For example, view A as a multiplication operator.) Conclude that if $A^* = A$, then one of $\|A\|$ and $-\|A\|$ belongs to $\mathrm{Sp}(A)$.

It then follows from Thm. 5.8.11 and the above exercise that one of $\|A\|$ and $-\|A\|$ is an approximate eigenvalue A. Of course, Riesz's approach in [Rie10] does not rely on the general spectral theorem for bounded self-adjoint operators or bounded Hermitian forms, since Hilbert's original proof of the Hilbert-Schmidt theorem did not. Riesz's proof of this fact is more elementary, though quite technical. We include his argument below, not with the expectation that the reader should master the detail; indeed, we discourage this. Rather, the purpose is to give a sense of how such a result can be obtained through elementary arguments.

Lemma 8.5.3. *Let $A \in \mathfrak{L}(\mathcal{H})$ be self-adjoint. Then one of $\|A\|$ and $-\|A\|$ is an approximate eigenvalue A.*

Proof. Assume WLOG that $\|A\| = 1$. Suppose that neither 1 nor -1 is an approximate eigenvalue of A. Then there exists $\varepsilon \in \mathbb{R}_{>0}$ such that

$$\|\xi - A\xi\|^2 \geqslant \varepsilon\|\xi\|^2 \qquad \|\eta + A\eta\|^2 \geqslant \varepsilon\|\eta\|^2$$

for all $\xi, \eta \in \mathcal{H}$. Set $\eta = \xi - A\xi$. Then

$$\|\xi - A^2\xi\|^2 = \|(\xi - A\xi) + A(\xi - A\xi)\|^2 \geqslant \varepsilon\|\xi - A\xi\|^2 \geqslant \varepsilon^2\|\xi\|^2$$

Since $\|A\| = 1$, we have $\|A^2\xi\|^2 \leqslant \|\xi\|^2$, and there exists $\xi \in \mathcal{H}$ such that

$$\|A\xi\|^2 > \left(1 - \frac{\varepsilon^2}{2}\right)\|\xi\|^2$$

For such a vector ξ, we compute

$$\begin{aligned}\|\xi - A^2\xi\|^2 &= \|\xi\|^2 - 2\mathrm{Re}\langle\xi|A^2\xi\rangle + \|A^2\xi\|^2 \xlongequal{A^*=A} \|\xi\|^2 - 2\|A\xi\|^2 + \|A^2\xi\|^2 \\ &< \|\xi\|^2 - (2 - \varepsilon^2)\|\xi\|^2 + \|\xi\|^2 = \varepsilon^2\|\xi\|^2\end{aligned}$$

This contradicts the previously obtained inequality $\|\xi - A^2\xi\|^2 \geqslant \varepsilon^2\|\xi\|^2$. □

8.5.2 Riesz's method: from approximate eigenvalues to genuine eigenvalues

Up to this point we have not used the assumption that T is completely continuous. In [Rie10], Riesz exploits this assumption in a crucial way, and the same idea later becomes fundamental in his proof of the Fredholm alternative in [Rie18].

Theorem 8.5.4. *Let $\mathcal{V}$ be a normed $\mathbb{F}$-vector space. Let $A : \mathcal{V} \to \mathcal{V}$ be a compact operator. Let $\lambda \in \mathbb{F}\backslash\{0\}$. Then the following are equivalent.*

(1) λ is an eigenvalue of A.

(2) λ is an approximate eigenvalue of A.

We give two proofs below. The first applies only to Hilbert spaces, but is more straightforward.

First proof for Hilbert spaces. Assume that $\mathcal{V}$ is a Hilbert space. Then A is completely continuous by Thm. 8.4.7. Let us prove (2)⇒(1), as the other direction is obvious.

Assume (2). There there exists a net (indeed, a sequence) of unit vectors (ξ_n) in $\mathcal{V}$ such that

$$\lim_n \|A\xi_n - \lambda\xi_n\| = 0 \tag{8.17}$$

Since $\overline{B}_{\mathcal{V}}(0,1)$ is weakly compact (cf. Thm. 3.7.5), by passing to a subnet we may assume that (ξ_n) converges weakly to some $\xi \in \mathcal{H}$. Then $(A-\lambda)\xi_n$ converges weakly to $(A-\lambda)\xi$ by Prop. 8.4.2. Together with (8.17), this implies $(A-\lambda)\xi = 0$.

To conclude (1), it remains to show that $\xi \neq 0$. Here we use the fact that A is completely continuous, which implies

$$\lim_n \|A\xi_n - A\xi\| = 0$$

This, together with (8.17), implies $\|A\xi\| = \lim_n \|\lambda\xi_n\| = |\lambda| \neq 0$ and hence $\xi \neq 0$. □

In the above proof, we observe that after passing to a subnet, we have $A\xi_n \to A\xi$ and $A\xi = \lambda\xi$, therefore $A\xi_n \to \lambda\xi$. This suggests that one can obtain the eigenvector ξ by choosing a convergent subsequence/subnet of $(A\xi_n)$ without first selecting a weakly convergent subnet of (ξ_n). This motivates the following proof in the general case.

Second proof of Thm. 8.5.4. Let us prove the nontrivial direction (2)⇒(1). Assume (2). Then there exists a sequence of unit vectors (ξ_n) in $\mathcal{V}$ satisfying (8.17). Unlike in the first proof, we cannot obtain a nonzero eigenvector by selecting a weakly convergent subnet of (ξ_n). However, the compactness of A implies that, after passing to a subsequence, we have

$$\lim_n \|A\xi_n - \lambda\xi\| = 0$$

for some $\xi \in \mathcal{V}$ (here we use the fact that $\lambda \neq 0$).

Combining this limit with (8.17), we obtain $\lim_n \xi_n = \xi$; in particular, $\|\xi\| = \lim_n \|\xi_n\| = 1 \neq 0$. Using $\lim_n \xi_n = \xi$ together with (8.17), we conclude that $A\xi = \lambda\xi$. This proves (1). □

8.5.3 Proof of the Hilbert-Schmidt theorem

We now return to the setting where $T \in \mathfrak{L}(\mathcal{H})$ is completely continuous and self-adjoint, and explain how Riesz proceeded to prove the Hilbert-Schmidt theorem.

Assume for simplicity that $\mathcal{H}$ is infinite-dimensional. By Lem. 8.5.3 and Thm. 8.5.4, T has a unit eigenvector e_1 with eigenvalue $\lambda_1 \in \{\|T\|, -\|T\|\}$. Applying the same method to $T|_{\mathrm{Span}(e_1)}$, we obtain another unit eigenvector e_2. Repeating this argument yields an orthonormal sequence (e_n) in $\mathcal{H}$ such that $Te_n = \lambda_n e_n$ where $\lambda_n \in \mathbb{R}$, and

$$\|T\xi\| \leqslant |\lambda_n| \cdot \|\xi\|^2 \qquad \text{for each } \xi \in \mathrm{Span}\{e_1, \dots, e_{n-1}\}^\perp \tag{8.18}$$

Arguing as in the proof of Thm. 8.3.1, one shows $\lim_n \lambda_n = 0$ and $Tf_j = 0$ where $(f_j)_{j\in J}$ is an orthonormal basis of $\{e_n : n \in \mathbb{Z}_+\}^\perp$. This finishes the proof.

8.6 Compact triangular matrices

Let $\mathcal{H}$ be a Hilbert space. Let $\mathcal{V}$ be a normed $\mathbb{F}$-vector space.

The goal of this section is to discuss several basic features of compact operators in preparation for the proof of the Fredholm alternative in Sec. 8.7. We begin with the following observation.

8.6.1 An elementary example

Example 8.6.1. Let $(e_n)_{n\in\mathbb{Z}_+}$ be an orthonormal sequence in $\mathcal{H}$. Let $T : \mathcal{H} \to \mathcal{H}$ be a compact operator. Then for sequences $(m_k)_{k\in\mathbb{Z}_+}$ and $(n_k)_{k\in\mathbb{Z}_+}$ in $\mathbb{Z}_+$ converging to ∞, we have

$$\lim_{k\to\infty} \langle e_{m_k} | Te_{n_k} \rangle = 0 \tag{8.19}$$

This is because $(e_{m_k})_{k\in\mathbb{Z}_+}$ and $(e_{n_k})_{k\in\mathbb{Z}_+}$ converge weakly to 0 (cf. Exp. 3.7.4), and the sesquilinear form ω_T is completely continuous.[5]

The matrix interpretation of this example is as follows. Let $A \in \mathbb{C}^{\mathbb{Z}_+\times\mathbb{Z}_+}$ be the matrix representation of a compact operator on $l^2(\mathbb{Z}_+)$. Then

$$\lim_{k\to\infty} A(m_k, n_k) = 0$$

In particular, the diagonal entries of A converge to 0. □

We will be primarily interested in Exp. 8.6.1 in the case where A is an upper or lower triangular matrix, for two reasons. First, this special case implies that a compact operator $T \in \mathfrak{L}(\mathcal{H})$ cannot have a Jordan block of infinite size corresponding to a nonzero eigenvalue, nor can it have infinite-dimensional eigenspaces for such eigenvalues. Second, the triangular-matrix setting admits a natural generalization to compact operators on normed vector spaces.

[5]In fact, it suffices to assume that one of $(m_k)_{k\in\mathbb{Z}_+}$ and $(n_k)_{k\in\mathbb{Z}_+}$, say $(n_k)_{k\in\mathbb{Z}_+}$, converges to ∞. Then $\lim_k \|Te_{n_k}\| = 0$ by the complete continuity of T. Eq. (8.19) then follows from the inequality $|\langle e_{m_k}|Te_{n_k}\rangle| \leqslant \|Te_{n_k}\|$.

To motivate this generalization, we begin the following subsection by presenting an alternative proof of Exp. 8.6.1 for upper triangular matrices using operator language rather than completely continuous sesquilinear forms.

8.6.2 Compact upper triangular matrices

Example 8.6.2. Let $A \in \mathbb{C}^{\mathbb{Z}_+ \times \mathbb{Z}_+}$ be the matrix representation of a compact operator on $l^2(\mathbb{Z}_+)$. Assume that A is **upper triangular** with diagonal sequence $(\lambda_n)_{n\in\mathbb{Z}_+}$ in $\mathbb{C}$, i.e., A is of the form

$$A = \begin{pmatrix} \lambda_1 & ? & ? & ? & \cdots \\ 0 & \lambda_2 & ? & ? & \cdots \\ 0 & 0 & \lambda_3 & ? & \cdots \\ 0 & 0 & 0 & \lambda_4 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

where each $?$ denotes a scalar. Then $\lim_n \lambda_n = 0$.

In the following proofs, we let A be the matrix representation of $T \in \mathfrak{L}(l^2(\mathbb{Z}_+))$ under the basis $(e_n)_{n\in\mathbb{Z}_+}$ where $e_n = \chi_{\{n\}}$. Then

$$Te_n = \lambda_n e_n + ?e_{n-1} + \cdots + ?e_1 \tag{8.20}$$

First proof. Since (e_n) converges weakly to zero and T is completely continuous, we have $\|Te_n\| \to 0$. Since

$$\|Te_n\|^2 = |\lambda_n|^2 + |?|^2 + \cdots + |?|^2 \geqslant |\lambda_n|^2$$

we must have $\lambda_n \to 0$. □

Second proof. By (8.20), for $m < n$ in $\mathbb{Z}_+$ we have

$$\|Te_n - Te_m\|^2 = \|\lambda_n e_n + ?e_{n-1} + \cdots + ?e_1\|^2 \geqslant \|\lambda_n\|^2 \tag{8.21}$$

If $\lambda_n \nrightarrow 0$, then there exists $\delta \in \mathbb{R}_{>0}$ and a subsequence (λ_{n_k}) satisfying $|\lambda_{n_k}| \geqslant \delta$ for each k. Therefore

$$\|Te_{n_k} - Te_{n_l}\|^2 \geqslant \delta^2$$

for each $l < k$ in $\mathbb{Z}_+$. Thus $(Te_{n_k})_{k\in\mathbb{Z}_+}$ has no convergent subsequence, contradicting the compactness of T. □

We now formulate and generalize Exp. 8.6.2 in a basis-independent way.

Theorem 8.6.3. *Let $T : \mathcal{V} \to \mathcal{V}$ be a compact operator. Suppose there exists a strictly increasing sequence*

$$V_0 \subsetneq V_1 \subsetneq V_2 \subsetneq V_3 \subsetneq \cdots$$

of closed linear subspaces of $\mathcal{V}$, together with a sequence (λ_n) in $\mathbb{F}$, such that

$$(T - \lambda_n)V_n \subset V_{n-1} \qquad \text{for each } n \in \mathbb{Z}_+$$

Then $\lim\limits_{n\to\infty} \lambda_n = 0$.

In the case where $\mathcal{V}$ is a Hilbert space, the operator T, when restricted to $\overline{\bigcup_n V_n}$, may be viewed as a block upper-triangular matrix with respect to the decomposition

$$V_0 \oplus (V_1 \ominus V_0) \oplus (V_2 \ominus V_1) \oplus \cdots$$

Thus, the following proofs are motivated by the those of Exp. 8.6.2.

First proof for Hilbert spaces. Suppose $\mathcal{V}$ is a Hilbert space. For each $n \in \mathbb{Z}_+$, choose a unit vector $e_n \in V_n$ orthogonal to V_{n-1}; such e_n exists by the closedness of V_{n-1} and V_n. Then

$$Te_n = \lambda_n e_n + \text{ a vector in } V_{n-1} \tag{8.22}$$

and hence $\|Te_n\|^2 \geqslant |\lambda_n|^2$ by the Pythagorean identity. Since the sequence (e_n) is orthonormal, it converges weakly to 0 by Exp. 3.7.4. It follows from the complete continuity of T that $\|Te_n\|^2 \to 0$, and hence $\lambda_n \to 0$. $\square$

Second proof for Hilbert space. Suppose $\mathcal{V}$ is a Hilbert space. Let (e_n) be as in the first proof. Then for each $m < n$ in $\mathbb{Z}_+$, applying (8.22) to Te_m gives

$$Te_m \in V_m \subset V_{n-1}$$

Combining this relation with (8.22), we obtain

$$Te_n - Te_m = \lambda_n e_n + \text{ a vector in } V_{n-1} \tag{8.23}$$

Therefore, by the Pythagorean identity, we have

$$\|Te_n - Te_m\|^2 \geqslant \|\lambda_n\|^2 \tag{8.24}$$

which is analogous to (8.21). Arguing as in the second proof of Exp. 8.6.2, we conclude that $\lambda_n \to 0$. $\square$

To generalize the above second proof to compact operators on normed vector spaces, we must select vectors $e_n \in V_n$ for which an inequality analogous to (8.24) holds. To achieve this, we first need an appropriate analogue of the notion of a vector orthogonal to a linear subspace:

Definition 8.6.4. Let V be a normed $\mathbb{F}$-vector space with linear subspace U. Let $\varepsilon \in \mathbb{R}_{\geqslant 0}$. A vector $\xi \in V$ is called ***ε*-orthogonal** to U if

$$\mathrm{dist}(\xi, \mathcal{U}) \geqslant (1-\varepsilon)\|\xi\|$$

that is, if

$$\|\xi - \eta\| \geqslant (1-\varepsilon)\|\xi\| \qquad \text{for each } \eta \in \mathcal{U}$$

In the case where V is a Hilbert space with closed linear subspace U, 0-orthogonality coincides with ordinary orthogonality.

Theorem 8.6.5 (Riesz lemma). *Let $\mathcal{U} \subsetneq \mathcal{V}$ be a proper closed linear subspace of $\mathcal{V}$. Then for each $\psi \in \mathcal{V}\backslash\mathcal{U}$ and $0 < \varepsilon < 1$, there exists a non-zero $\xi \in \psi + \mathcal{U}$ that is ε-orthogonal to $\mathcal{U}$.*

Proof. Since $\mathcal{U}$ is closed, we have $\mathrm{dist}(\psi, \mathcal{U}) > 0$. Choose $u \in \mathcal{U}$ such that $0 < \|\psi - u\| \leqslant \mathrm{dist}(\psi, \mathcal{U})/(1-\varepsilon)$. Then $\xi := \psi - u$ satisfies

$$\mathrm{dist}(\xi, \mathcal{U}) = \mathrm{dist}(\psi, \mathcal{U}) \geqslant (1-\varepsilon)\|\xi\|$$

□

Third proof of Thm. 8.6.3. Choose any $0 < \varepsilon < 1$, for example $\varepsilon = \frac{1}{2}$. Since each V_{n-1} is closed in V_n, by the Riesz lemma (Thm. 8.6.5), there exists a unit vector $e_n \in V_n \backslash V_{n-1}$ satisfying

$$\|e_n - \text{ any vector in } V_{n-1}\| \geqslant 1-\varepsilon$$

For each $m < n$, since (8.23) holds, the above inequality implies

$$\|Te_n - Te_m\| \geqslant |\lambda_n|(1-\varepsilon)$$

Arguing as in the second proof of Exp. 8.6.2, we conclude that $\lambda_n \to 0$. □

8.6.3 Compact lower triangular matrices

We now discuss the analogous result for lower triangular matrices.

Example 8.6.6. Suppose that $T \in \mathfrak{L}(l^2(\mathbb{Z}_+))$ is a compact operator whose matrix representation B is **lower triangular** with diagonal sequence (λ_n), that is,

$$B = \begin{pmatrix} \lambda_1 & 0 & 0 & 0 & \cdots \\ ? & \lambda_2 & 0 & 0 & \cdots \\ ? & ? & \lambda_3 & 0 & \cdots \\ ? & ? & ? & \lambda_4 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Then for each $m > n$ (rather than $m < n$) in $\mathbb{Z}_+$, we have

$$\|Te_n - Te_m\|^2 = \|\lambda_n e_n + ?e_{n+1} + ?e_{n+2} + \cdots\|^2 \geqslant \|\lambda_n\|^2 \tag{8.25}$$

Thus one can repeat the argument from the second proof of Exp. 8.6.2 to conclude that $\lambda_n \to 0$. This motivates the proof of the following theorem.

Theorem 8.6.7. *Let $T : \mathcal{V} \to \mathcal{V}$ be a compact operator. Suppose there exists a strictly decreasing sequence*

$$V_1 \supsetneq V_2 \supsetneq V_3 \supsetneq V_4 \supsetneq \cdots$$

of closed linear subspaces of $\mathcal{V}$, together with a sequence (λ_n) in $\mathbb{F}$, such that

$$(T - \lambda_n)V_n \subset V_{n+1} \qquad \text{for each } n \in \mathbb{Z}_+$$

Then $\lim\limits_{n\to\infty} \lambda_n = 0$.

In the case where $\mathcal{V}$ is a Hilbert space, $T|_{V_1}$ may be viewed as a block lower-triangular matrix with respect to the decomposition

$$(V_1 \ominus V_2) \oplus (V_2 \ominus V_3) \oplus \cdots \oplus V_\infty \quad \text{where } V_\infty = \bigcap_n V_n$$

Proof of Thm. 8.6.7. Choose any $0 < \varepsilon < 1$, for example $\varepsilon = \frac{1}{2}$. Since each V_{n+1} is closed in V_n, by the Riesz lemma (Thm. 8.6.5), there exists a unit vector $e_n \in V_n \backslash V_{n+1}$ satisfying

$$\|e_n - \text{ any vector in } V_{n+1}\| \geqslant 1 - \varepsilon$$

By assumption, we have

$$Te_n = \lambda_n e_n + \text{ a vector in } V_{n+1} \tag{8.26}$$

For each $n < m$, applying (8.26) to Te_m gives

$$Te_m \in V_m \subset V_{n+1}$$

Combining this relation with (8.26), we get

$$Te_n - Te_m = \lambda_n e_n + \text{ a vector in } V_{n+1}$$

This relation, together with the above inequality, implies

$$\|Te_n - Te_m\| \geqslant |\lambda_n|(1 - \varepsilon)$$

Arguing as in the second proof of Exp. 8.6.2, we conclude that $\lambda_n \to 0$. □

Remark 8.6.8. Assume that $\mathcal{V}$ is a Hilbert space. Then Thm. 8.6.7 can be proved using completely continuous operators, in a manner similar to the first proof of Thm. 8.6.3. We leave the details to the reader.

Alternatively, one may prove Thm. 8.6.7 by applying Thm. 8.6.3 to the adjoint operator T^* (which is compact by Cor. 8.4.8) and the increasing chain of subspaces $V_1^\perp \subsetneq V_2^\perp \subsetneq \cdots$. □

8.6.4 Conclusion

We conclude this section by restating Thm. 8.6.3 and 8.6.7 in a form that is more convenient for certain applications. In particular, we weaken the assumption that the chain of subspaces must be strictly increasing (or strictly decreasing).

Corollary 8.6.9. *Let $T : \mathcal{V} \to \mathcal{V}$ be a compact operator. Suppose that*

$$V_0 \subset V_1 \subset V_2 \subset V_3 \subset \cdots$$

is an increasing sequence of closed linear subspaces of $\mathcal{V}$ satisfying

$$(T - \lambda_n)V_n \subset V_{n-1} \qquad \textit{for each } n \in \mathbb{Z}_+$$

where $\lambda_n \in \mathbb{F}$. Suppose that (V_n) is not eventually stable, that is, for each n there exists $m > n$ such that $V_n \subsetneq V_m$. Then

$$\liminf_{n\to\infty} |\lambda_n| = 0$$

Proof. Since (V_n) is not eventually stable, there exists a strictly increasing sequence $(n_k)_{k\in\mathbb{N}}$ in $\mathbb{N}$ such that [6]

$$V_{n_0} \subsetneq V_{n_1} \subsetneq V_{n_2} \subsetneq V_{n_3} \subsetneq \cdots$$

and that

$$(T - \lambda_{n_k})V_{n_k} \subset V_{n_{k-1}} \qquad \text{for each } k \in \mathbb{Z}_+$$

Then $\lim_k \lambda_{n_k} = 0$ by Thm. 8.6.3, and hence $\liminf_{n\to\infty} |\lambda_n| = 0$. □

Corollary 8.6.10. *Let $T : \mathcal{V} \to \mathcal{V}$ be a compact operator. Suppose that*

$$V_1 \supset V_2 \supset V_3 \supset V_4 \supset \cdots$$

is a decreasing sequence of closed linear subspaces of $\mathcal{V}$ satisfying

$$(T - \lambda_n)V_n \subset V_{n+1} \qquad \textit{for each } n \in \mathbb{Z}_+$$

where $\lambda_n \in \mathbb{F}$. Suppose that (V_n) is not eventually stable, that is, for each n there exists $m > n$ such that $V_n \supsetneq V_m$. Then

$$\liminf_{n\to\infty} |\lambda_n| = 0$$

Proof. This follows from Thm. 8.6.7 in the same way that Cor. 8.6.9 follows from Thm. 8.6.3. □

[6] Choose $n_0 = 0$. Having chosen n_k, let n_{k+1} be the smallest integer $> n_k$ with $V_{n_k} \subsetneq V_{n_{k+1}}$.

8.7 Application of compact triangular matrices: Fredholm alternative

Fix a normed vector space $\mathcal{V}$ over $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$.

In this section we prove the Fredholm alternative for compact operators on Banach spaces, originally proved by Riesz in [Rie18]. When applied to the integral operator on $C(\partial\Omega)$ arising from the method of double layer potentials, this result yields the solvability of the Dirichlet problem on $\Omega \subset \mathbb{R}^N$, as discussed in Subsec. 8.1.3-8.1.5. We refer the reader to those subsections for a detailed account of the history of the Fredholm alternative. Here we add a few remarks concerning the choice of function spaces.

8.7.1 Why non-self-adjoint operators? Why general Banach spaces?

Recall the integral kernel $K : \partial\Omega \times \partial\Omega \to \mathbb{C}$ defined in (8.7) (i.e. $K(x, y) = \langle \nabla\Phi(x - y), \mathbf{n}_y \rangle$). Using this kernel, one defines the integral operator T_K by $(T_K f)(x) = \int_{\partial\Omega} K(x, y)f(y)dy$. Because of the estimate (8.5) describing the behavior of K near the diagonal, one can show that T_K is compact both on $C(\partial\Omega)$ and on $L^2(\partial\Omega)$; see Pb. 8.6. When regarded as a compact operator on $C(\partial\Omega)$, the solution of the integral equation

$$\frac{1}{2}f + T_K f = g \qquad \text{where } g \in C(\partial\Omega) \tag{8.27}$$

gives the solution of the Dirichlet problem (8.1), as explained in the subsections mentioned above.

Historically, the development of compact-operator theory concentrated on Hilbert spaces rather than on the space $C(\partial\Omega)$, with particular emphasis on compact self-adjoint operators—for instance, in the Hilbert-Schmidt theorem. In modern treatments of differential equations (cf. Subsec. 8.1.2), the compact operators $(-\Delta_{\mathrm{D}})^{-1}$ and $(1 - \Delta_{\mathrm{D}})^{-1}$ are self-adjoint, and the Hilbert-Schmidt theorem applies directly. However, in the classical integral-equation approach, the operator T_K on $L^2(\partial\Omega)$ is typically not self-adjoint unless Ω has a highly symmetric shape. Thus one must study compact operators that are not self-adjoint.

A further limitation appears in the Hilbert-space setting when $N > 2$: the kernel K is not continuous, and hence a function $f \in L^2(\partial\Omega)$ satisfying (8.27) need not be continuous, and therefore may fail to provide a solution to the original Dirichlet problem. In fact, K is not even in L^2. The situation improves dramatically when $N = 2$, since in that case the kernel K is continuous . Hence, $T_K f$ is automatically continuous for any $f \in L^2(\partial\Omega)$, and any L^2-solution of (8.27) must therefore belong to $C(\partial\Omega)$, thereby solving the original Dirichlet problem.

It is fortunate that the early study of the Dirichlet problem was restricted to the case $N = 2$, the dimension in which the notion of Hilbert space first appeared in

[Hil06]. Had the initial investigations treated arbitrary N, the discovery of Hilbert spaces would likely have been significantly delayed—if not missed entirely.

However, for the study of integral equations in arbitrary dimension, one must eventually abandon the assumption of self-adjointness and instead work in the Banach space $C(X)$ of continuous functions on a compact Hausdorff space X.[7] The Banach space $C(X)$ is not a dual space; thus the only analytic property one can rely on is its completeness. This illustrates the paradigm shift described in (5.19c).

I hope these remarks help convey that the study of non-self-adjoint compact operators on general Banach spaces is not pursued merely for the sake of generality; rather, it is motivated by concrete problems. The same is true for many results that arise in the theory of compact operators but may at first seem unrelated—for instance, Riesz's lemma (Thm. 8.6.5), the equivalence of any two norms on a finite-dimensional vector space (Thm. 8.4.14), and the equivalence between finite dimensionality and compactness of the closed unit ball (Cor. 8.7.6).

8.7.2 Fredholm alternative for compact operators on Banach spaces

Definition 8.7.1. For each $T \in \mathfrak{L}(\mathcal{V})$, define the **spectrum**

$$\boldsymbol{\sigma(T)} = \{\lambda \in \mathbb{F} : T - \lambda \text{ is not invertible in } \mathfrak{L}(\mathcal{V})\} \tag{8.28}$$

That is, $\mathbb{F}\backslash\sigma(T)$ is the set of all $\lambda \in \mathbb{F}$ such that there exists $A \in \mathfrak{L}(\mathcal{V})$ satisfying $A(T-\lambda) = (T-\lambda)A = \mathrm{id}_{\mathcal{V}}$.

In other words, $\mathbb{F}\backslash\sigma(T)$ is the set of all $\lambda \in \mathbb{F}$ such that $T-\lambda : \mathcal{V} \to \mathcal{V}$ is bijective, and that its inverse $(T-\lambda)^{-1} : \mathcal{V} \to \mathcal{V}$ is bounded.

Theorem 8.7.2 (Fredholm alternative). *Assume that $\mathcal{V}$ is complete and T is a compact operator on $\mathcal{V}$. Let $\lambda \in \mathbb{F}\backslash\{0\}$. Then the following are equivalent.*

(1) $\mathrm{Ker}(T-\lambda) = 0$.

(2) $\lambda \notin \sigma(T)$.

(3) $\mathrm{Rng}(T-\lambda) = \mathcal{V}$.

In its original meaning, the Fredholm alternative denotes the equivalence of (1) and (3). See Subsec. 8.1.5 for more explanation of the terminology.

Note that the direction (1)⇔(2) asserts that the set of non-zero eigenvalues of T coincides with $\sigma(T)\backslash\{0\}$.

[7] In [Rie18], Riesz worked with the case where X is a compact interval. But he explicitly noted that his method applies directly to more general spaces, namely, Banach spaces (a notion introduced many years after [Rie18]).

Proof. The direction (2)⇒(3) is obvious. In the first two steps we prove (1)⇒(2), the most important direction of the theorem. In Step 3 we prove (3)⇒(1).

Step 1. Assume (1). Then λ is not an eigenvalue of T, and hence not an approximate eigenvalue by Thm. 8.5.4. Therefore, there exists $\varepsilon \in \mathbb{R}_{>0}$ such that

$$\|(T-\lambda)\xi\| \geqslant \varepsilon\|\xi\| \qquad \text{for each } \xi \in \mathcal{H}$$

Therefore, if we can show that $\mathrm{Rng}(T-\lambda) = \mathcal{V}$, then $(T-\lambda)^{-1}$ is clearly a bounded linear map on $\mathcal{V}$, and hence $\lambda \in \sigma(T)$.

Step 2. Let us prove that $T-\lambda$ is surjective. Note that the above inequality implies for each $n \in \mathbb{Z}_+$ that

$$\|(T-\lambda)^n\xi\| \geqslant \varepsilon^n\|\xi\| \qquad \text{for each } \xi \in \mathcal{H}$$

and hence that $V_n := (T-\lambda)^n\mathcal{V}$ is a closed linear subspace of $\mathcal{V}$ by Lem. 5.10.21 and the completeness of $\mathcal{V}$. We thus obtain a decreasing sequence of closed linear subspaces

$$V_1 \supset V_2 \supset V_3 \supset \cdots$$

satisfying $(T-\lambda)V_n \subset V_{n+1}$ (indeed $(T-\lambda)V_n = V_{n+1}$) for each n.

Suppose that $T-\lambda$ is not surjective. Then $V_1 \subsetneq \mathcal{V}$, and for each injective linear map $A : \mathcal{V} \to \mathcal{V}$ we have $AV_1 \subsetneq A\mathcal{V}$. Taking $A = (T-\lambda)^n$, we obtain $V_{n+1} \subsetneq V_n$. Therefore, the chain (V_n) is strictly decreasing. By Thm. 8.6.7, this forces $\lambda = 0$, a contradiction. This completes the proof of (1)⇒(2).

Step 3. Let us prove (3)⇒(1). Assume (3). If $\mathrm{Ker}(T-\lambda) \neq 0$, there there exists a nonzero $\xi_1 \in \mathrm{Ker}(T-\lambda)$. For each $n \in \mathbb{Z}_+$, since $(T-\lambda)^n$ is surjective, there exists $\xi_{n+1} \in \mathcal{H}$ such that $(T-\lambda)^n\xi_{n+1} = \xi_1$. Since

$$(T-\lambda)^n\xi_{n+1} = \xi_0 \neq 0 \qquad (T-\lambda)^{n+1}\xi_{n+1} = (T-\lambda)\xi_0 = 0$$

the vector ξ_n lies in $\mathrm{Ker}((T-\lambda)^n)$ but not in $\mathrm{Ker}((T-\lambda)^{n-1})$. This proves

$$\mathrm{Ker}((T-\lambda)^{n-1}) \subsetneq \mathrm{Ker}((T-\lambda)^n)$$

for each n, contradicting the following Prop. 8.7.3. □

Proposition 8.7.3. *Let T be a compact operator on $\mathcal{V}$. Let $\lambda \in \mathbb{F}\backslash\{0\}$. Then there exists $n \in \mathbb{Z}_+$ such that*

$$\mathrm{Ker}((T-\lambda)^n) = \mathrm{Ker}((T-\lambda)^{n+1}) = \mathrm{Ker}((T-\lambda)^{n+2}) = \cdots$$

Proof. Apply Cor. 8.6.9 to the increasing chain $(V_n)_{n\in\mathbb{N}}$ where each $V_n = \mathrm{Ker}((T-\lambda)^n)$ is closed by Cor. 2.3.11. □

Exercise 8.7.4. In the special case where $\mathcal{V}$ is a Hilbert space and T is a compact self-adjoint operator on $\mathcal{V}$, explain why Thm. 8.7.2 is an easy consequence of the Hilbert-Schmidt Thm. 8.3.1.

8.7.3 Supplementary results on compact operators

We close this section with a few properties of compact operators that follow easily from the theorems established in Sec. 8.6. Note that we do not assume completeness of $\mathcal{V}$, just as in Prop. 8.7.3. See Pb. 8.8 and 8.9 for additional complementary results.

Proposition 8.7.5. *Let T be a compact operator on $\mathcal{V}$. Let $\lambda \in \mathbb{F}\backslash\{0\}$. Then $\mathrm{Ker}((T-\lambda)^n)$ is finite-dimensional for each $n \in \mathbb{Z}_+$.*

Proof. We begin with the case $n = 1$. Suppose $\mathrm{Ker}(T - \lambda)$ is infinite dimensional. Then there exists a strictly increasing sequence $(V_k)_{k\in\mathbb{Z}_+}$ of finite-dimensional linear subspaces of $\mathrm{Ker}(T - \lambda)$. Since $(T - \lambda)V_{k+1} = 0$, we have $(T - \lambda)V_{k+1} \subset V_k$. By Thm. 8.4.14, V_k is complete and therefore closed in $\mathrm{Ker}(T - \lambda)$. We may then apply Thm. 8.4 to reach a contradiction.

We now treat the general case by induction. Suppose $\mathrm{Ker}((T - \lambda)^n)$ is finite-dimensional but $\mathrm{Ker}((T - \lambda)^{n+1})$ is not. Then we can choose a strictly increasing sequence (V_k) of finite-dimensional linear subspaces of $\mathrm{Ker}((T-\lambda)^{n+1})$ containing $\mathrm{Ker}((T-\lambda)^n)$. The operator $(T-\lambda)$ sends each V_{k+1} into $\mathrm{Ker}((T-\lambda)^n)$, and hence into V_k. The same argument as in the case $n = 1$ then yields a contradiction.[8]

Alternatively, one can reduce the general case to the special case by observing that $(T - \lambda)^n = TA + (-\lambda)^n$ for some $A \in \mathfrak{L}(\mathcal{V})$, and TA is clearly compact. $\square$

Corollary 8.7.6. *$\mathcal{V}$ is finite-dimensional iff $\overline{B}_{\mathcal{V}}(0, 1)$ is compact (in the norm topology).*

Proof. "$\Rightarrow$": By Thm. 8.4.14.

"$\Leftarrow$": The compactness of $\overline{B}_{\mathcal{V}}(0, 1)$ is equivalent to the compactness of the identity operator $\mathbf{1}$ of $\mathcal{V}$. By Prop. 8.7.5, if $\mathbf{1}$ is compact, then $\mathcal{V} = \mathrm{Ker}(\mathrm{id} - \mathrm{id})$ is finite-dimensional. $\square$

Exercise 8.7.7. Use the Riesz lemma (Thm. 8.6.5) to give a direct proof of the direction "$\Leftarrow$" of Cor. 8.7.6, without appealing to compact operators.

Proposition 8.7.8. *Let T be a compact operator on $\mathcal{V}$. Then the set of non-zero eigenvalues has no accumulation points in $\mathbb{F}\backslash\{0\}$.*

Proof. Suppose, on the contrary, that $\lambda \in \mathbb{F}\backslash\{0\}$ is the limit of a sequence $(\lambda_n)_{n\in\mathbb{Z}_+}$ of non-zero eigenvalues of T where $\lambda_m \neq \lambda_n$ for each $m \neq n$ in $\mathbb{Z}_+$. For each n, choose $0 \neq \xi_n \in \mathrm{Ker}(T - \lambda_n)$, and let $V_n = \mathrm{Span}\{\xi_1, \dots, \xi_n\}$. Then (V_n) is an increasing sequence of finite-dimensional linear subspaces of $\mathcal{V}$. By the following Lem. 8.7.9, this sequence is strictly increasing.

By Thm. 8.4.14, each V_n is complete, and hence is closed in $\mathcal{V}$. Since

$$(T - \lambda_n)V_n \subset V_{n-1}$$

Thm. 8.6.3 implies that $\lim_n \lambda_n = 0$, contradicting $\lambda_n \to \lambda \neq 0$. $\square$

[8]In other words, we are applying the special case to $T - \lambda$ acting on the quotient space $\mathcal{V}/\mathrm{Ker}((T - \lambda)^n)$.

Lemma 8.7.9. *Let V be an $\mathbb{F}$-vector space. Let $T \in \mathrm{Lin}(V)$. Suppose that $\xi_1, \dots, \xi_n$ are non-zero vectors of V satisfying $T\xi_j = \lambda_j \xi_j$ for distinct scalars $\lambda_1, \dots, \lambda_n \in \mathbb{F}$. Then $\xi_1, \dots, \xi_n$ are linearly independent.*

Proof. This can be proved by induction on n. The case $n = 1$ is obvious. Suppose case $n - 1$ has been proved where $n \geqslant 2$. Then, in the setting of this lemma, $\xi_1, \dots, \xi_{n-1}$ are linearly independent. To prove that $\xi_1, \dots, \xi_n$ are linearly independent, it suffices to show $\xi_n \notin \mathrm{Span}\{\xi_1, \dots, \xi_{n-1}\}$.

Suppose not. Then $\xi_n = a_1\xi_1 + \cdots + a_{n-1}\xi_{n-1}$ for some $a_1, \dots, a_{n-1} \in \mathbb{F}$. Applying $T - \lambda_n$ to both sides gives

$$0 = a_1(\lambda_1 - \lambda_n)\xi_1 + \cdots + a_{n-1}(\lambda_{n-1} - \lambda_n)\xi_{n-1}$$

and hence $a_1 = \cdots = a_{n-1} = 0$ by the linear independence of $\xi_1, \dots, \xi_{n-1}$. This contradicts $\xi_n \neq 0$. □

8.8 Unbounded operators with compact graph projections and compact resolvents

Let $\mathcal{H}, \mathcal{K}$ be Hilbert spaces. In this section we apply the theory of compact operators to the study of unbounded operators. The results established in this section will be applied to partial differential equations in Sec. 8.9.

8.8.1 Unbounded operators with compact graph projections

Definition 8.8.1. Let $T : \mathcal{H} \to \mathcal{K}$ be an n.d.d. unbounded operator. We say that T has **compact graph projection** if the graph projection

$$\Pi_T : \mathscr{G}(T) \to \mathcal{H} \qquad \xi \oplus T\xi \mapsto \xi$$

is compact. Equivalently, T is said to have compact graph projection if the map

$$\mathscr{D}(T) \to \mathcal{H} \qquad \xi \mapsto \xi$$

is compact when $\mathscr{D}(T)$ is equipped with the graph inner product.

Recall from Def. 7.1.6 that Π_T is abbreviated to Π when no confusion arises.

In the following, we give some simple criteria for compact graph projection.

Proposition 8.8.2. *Let $T : \mathcal{H} \to \mathcal{K}$ be an unbounded operator with core $\mathscr{D}_0 \subset \mathscr{D}(T)$. Then T has compact graph projection iff $T|_{\mathscr{D}_0}$ does.*

Recall Def. 6.2.3 for the meaning of the restriction of unbounded operators.

Proof. Since $\mathscr{G}(T|_{\mathscr{D}_0})$ is dense in $\mathscr{G}(T)$, the proposition follows immediately from the following lemma. □

Lemma 8.8.3. *Let $\mathcal{V}, \mathcal{W}$ be normed vector spaces where $\mathcal{W}$ is complete. Let $A \in \mathfrak{L}(\mathcal{V}, \mathcal{W})$, and let $\mathcal{U}$ be a dense linear subspace of $\mathcal{V}$. Then A is compact iff the restriction $A|_{\mathcal{U}} : \mathcal{U} \to \mathcal{W}$ is compact.*

Proof. The direction "$\Rightarrow$" is obvious. Conversely, assume that $A|_{\mathcal{U}}$ is compact. Let (ξ_n) be a bounded sequence in $\mathcal{V}$. Choose $\eta_n \in \mathcal{U}$ such that $\|\xi_n - \eta_n\| \leqslant 1/n$. The compactness of $A|_{\mathcal{U}}$ implies that $(A\eta_n)$ has a convergence subsequence $(A\eta_{n_k})_{k\in\mathbb{Z}_+}$, which must be a Cauchy sequence. It follows that $(A\xi_{n_k})$ is also Cauchy, and hence converges by the completeness of $\mathcal{W}$. Thus A is compact. □

Recall Def. 7.9.1 for the meaning of A-bounds.

Proposition 8.8.4. *Let $T, S : \mathcal{H} \to \mathcal{K}$ be unbounded operator. Assume that S is T-bounded with bound $0 \leqslant \alpha < 1$. Then T has compact graph projection iff $T + S$ does.*

Proof. This is an immediate consequence of the fact that $\mathscr{D}(T) = \mathscr{D}(T+S)$ (because $\mathscr{D}(T) \subset \mathscr{D}(S)$ by assumption), and that the graph inner products of T and $T + S$ are equivalent (cf. Prop. 7.9.7). □

8.8.2 Unbounded operators with compact resolvents

Although the notion of compact graph projections is widely applicable, it yields compact operators whose domains and codomains differ. However, many of the most important results concerning compact operators (such as the Hilbert-Schmidt theorem and the Fredholm alternative) apply to compact operators acting on a single space.

To make these key results available for unbounded operators, we introduce a more convenient notion: that of compact resolvents. This notion applies to a smaller class of unbounded operators; for example, it does not apply to the closed gradient operator $\overline{\nabla}$ for open sets in $\mathbb{R}^N$ when $N > 1$, even though the notion of compact graph projection does apply.

Recall Def. 6.12.4 for the definition of $\sigma(T)$ for unbounded operators T on $\mathcal{H}$.

Theorem 8.8.5. *Let T be an unbounded operator on $\mathcal{H}$ such that $\sigma(T) \subsetneq \mathbb{C}$. Then T is closed. Moreover, the following conditions are equivalent.*

(1) The operator $(T - \lambda)^{-1} \in \mathfrak{L}(\mathcal{H})$ is compact for each $\lambda \in \mathbb{C}\backslash\sigma(T)$.

(2) The operator $(T - \lambda)^{-1} \in \mathfrak{L}(\mathcal{H})$ is compact for some $\lambda \in \mathbb{C}\backslash\sigma(T)$.

(3) T has compact graph projection.

If T satisfies one of these conditions, we say that T has ***compact resolvent****.*

In other words, for unbounded operators T on $\mathcal{H}$,

$$\text{compact solvent} \;=\; \text{compact graph projection} \;+\; \text{“}\sigma(T) \neq \mathbb{C}\text{”}$$

Proof. For each $\lambda \in \mathbb{C}\backslash\sigma(T)$, since $(T-\lambda)^{-1}$ belongs to $\mathfrak{L}(\mathcal{H})$ and hence is closed (cf. Prop. 6.4.8), it follows from the unitarity of the diagonal reflection map (cf. Def. 6.4.10) that $T-\lambda$ is closed, and hence T is closed.

To prove the equivalence of the three conditions, it suffices to prove that for each $\lambda \in \mathbb{C}\backslash\sigma(T)$, $(T-\lambda)^{-1}$ is compact iff T has compact graph projection. Moreover, since $\lambda \cdot \mathbf{1}$ is T-bounded with bound 0, Prop. 8.8.4 implies that T has compact graph projection iff $T-\lambda$ does. Therefore, replacing T with $T-\lambda$, it suffices to prove that if $0 \in \mathbb{C}\backslash\sigma(T)$ (that is, $T^{-1} \in \mathfrak{L}(\mathcal{H})$), then T^{-1} is compact iff the graph projection Π_T is compact.

Assume $T^{-1} \in \mathfrak{L}(\mathcal{H})$. Since $T^{-1}: \mathcal{H} \to \mathcal{H}$ is bounded, there exists $\delta \in \mathbb{R}_{>0}$ such that for each $\xi \in \mathscr{D}(T)$ we have

$$\|T\xi\| \geqslant \delta\|\xi\|$$

and hence

$$\|T\xi\|^2 \leqslant \|T\xi\|^2 + \|\xi\|^2 \leqslant (1+\delta^{-2})\|T\xi\|^2$$

It follows that there is a bounded linear bijection

$$\Phi : \mathscr{G}(T) \to \mathcal{H} \qquad \xi \oplus T\xi \mapsto T\xi$$

whose inverse is also bounded. Since $\Pi_T = T^{-1} \circ \Phi$, it follows that T^{-1} is compact iff Π_T is so. $\square$

Corollary 8.8.6. *Let T and S be unbounded operators on $\mathcal{H}$ where T is self-adjoint. Assume that S is T-bounded with bound $0 \leqslant \alpha < 1$. Then T has compact resolvent iff $T+S$ does.*

Proof. By Prop. 7.9.5, the spectrum $\sigma(T+S)$ is not equal to $\mathbb{C}$. We also have $\sigma(T) \subset \mathbb{R}$ by Rem. 6.10.13. Since, by Prop. 8.8.4, T has compact graph projection iff S does, it follows from Thm. 8.8.5 that T has compact resolvent iff $T+S$ does. $\square$

8.8.3 Hilbert-Schmidt and Fredholm alternative

We now establish the unbounded-operator versions of the Hilbert-Schmidt theorem and the Fredholm alternative. The following result is the Hilbert-Schmidt theorem for unbounded operators.

Theorem 8.8.7 (Hilbert-Schmidt). *Let T be a positive self-adjoint operator on $\mathcal{H}$. Then the following conditions are equivalent.*

(1) T has compact resolvent.

(2) *There exists a (possibly finite) sequence* (e_n) *of orthonormal basis of* $\mathcal{H}$ *such that each* e_n *is an eigenvector of* T *with eigenvalue* $\lambda_n \in \mathbb{R}_{\geqslant 0}$, [9] *and*

$$\lim_n \lambda_n = +\infty$$

if the sequence (e_n) *is infinite.*

Note that $\sigma(T) \subset \mathbb{R}_{\geqslant 0}$ by Rem. 6.10.13. Therefore, by Thm. 8.8.5, T has compact graph projection iff $(\lambda + T)^{-1}$ is compact for some/all $\lambda \in \mathbb{C}\backslash\mathbb{R}_{\geqslant 0}$.

Proof. Since the case $\dim \mathcal{H} < +\infty$ is trivial, we assume $\mathcal{H}$ is infinite-dimensional. Since $(\mathbf{1} + T)^{-1} \in \mathfrak{L}(\mathcal{H})$ is an injective self-adjoint operator, the Hilbert-Schmidt Thm. 8.3.1 shows that condition (1) is equivalent to the existence of an orthonormal basis (e_n) of $\mathcal{H}$ such that

$$(\mathbf{1} + T)^{-1}e_n = \kappa_n e_n \qquad \kappa_n \in \mathbb{R}\backslash\{0\} \qquad \lim_{n\to\infty} \kappa_n = 0$$

This is clearly equivalent to (2) upon relating λ_n and κ_n via $\kappa_n = 1/(1 + \lambda_n)$ and $\lambda_n = \kappa_n^{-1} - 1$, with the sole difference that condition (2) initially allows $\lambda_n \in \mathbb{R}\backslash\{-1\}$. However, since T is positive, all eigenvalues of T lie in $\mathbb{R}_{\geqslant 0}$. Thus the condition $\lambda_n \in \mathbb{R}\backslash\{-1\}$ may be strengthened to $\lambda_n \in \mathbb{R}_{\geqslant 0}$, completing the proof. $\square$

Remark 8.8.8. Thm. 8.8.7 can be extended easily to arbitrary self-adjoint operators, not necessarily positive, by applying the Hilbert-Schmidt theorem for compact normal operators (cf. Thm. 8.A.2) to $(\mathbf{i} + T)^{-1}$.

Corollary 8.8.9. *Let* T *be a positive operator on* $\mathcal{H}$. *Let* $r \in \mathbb{R}_{>0}$. *Then* T *has compact resolvent iff* T^r *does.*

Proof. Suppose that T has compact resolvent. Then T can be described by condition (2) of Thm. 8.8.7. Hence, by Prop. 7.7.9, we have $T^r e_n = \lambda_n^r e_n$ for each n. Thus T^r has compact resolvent. This proves one direction. The other direction follows from the composition law $T = (T^r)^{1/r}$ (cf. Thm. 6.10.5). $\square$

The following consequence of Cor. 8.8.9, together with the polar decomposition, allows us to relate the compactness properties of $\overline{\nabla}$ and $-\Delta_{\mathrm{D}}$ (a topic we will discuss in detail in Subsection 8.9.1).

Theorem 8.8.10. *Let* $T : \mathcal{H} \to \mathcal{K}$ *be a closed operator. Then the following are equivalent.*

(1) *T has compact graph projection.*

(2) *T^*T has compact resolvent.*

[9] In particular, $e_n \in \mathscr{D}(T)$.

(3) The absolute value $|T| := \sqrt{T^*T}$ has compact resolvent.

Proof. (2)⇔(3): This is due to Cor. 8.8.9.

(1)⇔(3): By polar decomposition (Thm. 7.1.11), the two domains $\mathscr{D}(T)$ and $\mathscr{D}(|T|)$ coincide, and that their graph inner products also coincide. Therefore, T has compact graph projection iff $|T|$ does. □

The following result is the Fredholm alternative for unbounded operators.

Theorem 8.8.11 (Fredholm alternative). *Let T be an unbounded operator on $\mathcal{H}$ with compact resolvent. Then the following are equivalent.*

(1) $\mathrm{Ker}(T) = 0$.

(2) $T^{-1} \in \mathfrak{L}(\mathcal{H})$.

(3) $\mathrm{Rng}(T) = \mathcal{H}$.

Proof. By assumption, there exists $\lambda \in \mathbb{C}$ such that $(T-\lambda)^{-1} \in \mathfrak{L}(\mathcal{H})$, and that $A := (T-\lambda)^{-1}$ is compact. If $\lambda = 0$, then (2) holds, and hence both (1) and (3) follow. Thus, it remains to treat the case $\lambda \neq 0$.

Since $T - \lambda = A^{-1}$, we have

$$T = A^{-1} + \lambda = A^{-1} + \lambda \cdot \mathrm{id}_{\mathscr{D}(A^{-1})} = A^{-1} + \lambda A A^{-1} = (\mathbf{1} + \lambda A)A^{-1} \tag{8.29}$$

where Prop. 6.2.8 is used in the last identity. Since A^{-1} is a bijection from $\mathscr{D}(A^{-1})$ to $\mathrm{Rng}(A^{-1}) = \mathscr{D}(A) = \mathcal{H}$, we conclude that T is injective (resp. surjective) iff $\mathbf{1} + \lambda A$ is injective (resp. surjective).

(1)⇒(2): Assume (1). Then $\mathbf{1} + \lambda A$ is injective. By the Fredholm alternative (Thm. 8.7.2), $(\mathbf{1} + \lambda A)$ is invertible in $\mathfrak{L}(\mathcal{H})$. By (8.29) and Prop. 6.2.21, we have

$$T^{-1} = A(\mathbf{1} + \lambda A)^{-1}$$

which belongs to $\mathfrak{L}(\mathcal{H})$. This proves (2).

(2)⇒(3): This is obvious.

(3)⇒(1): Assume (3). Then $\mathbf{1} + \lambda A$ is surjective. By the Fredholm alternative (Thm. 8.7.2), $\mathbf{1} + \lambda A$ is injective. Therefore T is injective. □

8.9 Application to partial differential equations

Let $\Omega \subset \mathbb{R}^N$ be a bounded open subset of $\mathbb{R}^N$. Recall from Subsec. 6.6.3 that the (smooth) gradient operator

$$\nabla : L^2(\Omega, m) \to L^2(\Omega, m)^{\oplus N} \qquad \mathscr{D}(\nabla) = C_c^\infty(\Omega)$$

is closable. Recall also from Subsec. 7.1.4 the negative Dirichlet Laplacian is defined by $-\Delta_{\mathrm{D}} = \overline{\nabla}^*\overline{\nabla}$, and is a positive self-adjoint operator on $L^2(\Omega, m)$ by Thm. 7.1.7.

8.9.1 Diagonalization of the Laplacian

Theorem 8.9.1 (Rellich compactness theorem). *The positive self-adjoint operator $-\Delta_{\mathrm{D}}$ has compact resolvent.*

Consequently, $-\Delta_{\mathrm{D}}$ admits a diagonalization of the form described in Thm. 8.8.7.

Proof. By Thm. 8.8.10, the fact that $-\Delta_{\mathrm{D}}$ has compact resolvent is equivalent to $\overline{\nabla}$ having compact graph projection. By Prop. 8.8.2, this is further equivalent to ∇ having compact graph projection.

Thus it suffices to show that the graph projection Π_∇ is compact, equivalently, the map

$$C_c^\infty(\Omega) \to L^2(\Omega, m) \qquad f \mapsto f$$

is compact where $C_c^\infty(\Omega)$ is equipped with the graph inner product of ∇:

$$\langle f|g\rangle_\nabla = \langle f|g\rangle + \langle \nabla f|\nabla g\rangle \qquad \text{for each } f, g \in C_c^\infty(\Omega)$$

To prove this, we may assume WLOG that $\overline{\Omega} \subset (-\pi, \pi)^N$, so that Ω can be regarded as an open subset of the torus $\mathbb{T}^N \simeq (\mathbb{R}/2\pi\mathbb{Z})^N$. (Note that functions on $\mathbb{T}^N$ are equivalently 2π-periodic functions on $\mathbb{R}^N$.) Then $C_c^\infty(\Omega) \subset C_c^\infty(\mathbb{T}^N)$, and hence it suffices to prove the compactness of

$$C_c^\infty(\mathbb{T}^N) \to L^2(\mathbb{T}, m) \qquad f \mapsto f$$

where $C_c^\infty(\mathbb{T}^N)$ is equipped with the graph inner product of the smooth gradient operator on $\mathbb{T}^N$:

$$\widetilde{\nabla} : L^2(\mathbb{T}^N, m) \to L^2(\mathbb{T}^N, m)^{\oplus N} \qquad \mathscr{D}(\widetilde{\nabla}) = C_c^\infty(\mathbb{T}^N)$$

By Prop. 8.8.2 and Thm. 8.8.10, it suffices to show that the unbounded operator

$$-\widetilde{\Delta}_{\mathrm{D}} = \overline{\widetilde{\nabla}}^* \overline{\widetilde{\nabla}} : L^2(\mathbb{T}^N, m) \to L^2(\mathbb{T}^N, m)$$

(which is positive and self-adjoint by Thm. 7.1.7) has compact resolvent.

Clearly, $-\widetilde{\Delta}_{\mathrm{D}}$ extends the smooth negative Laplacian on $L^2(\mathbb{T}^N, m)$ with domain $C_c^\infty(\mathbb{T}^N, m)$. Therefore, for each $n \in \mathbb{Z}^N$, the function

$$e_n : \mathbb{T}^N \to \mathbb{C} \qquad e_n(x) = e^{\mathbf{i}nx}$$

is an eigenvector of $-\widetilde{\Delta}_{\mathrm{D}}$ with eigenvalue $|n|^2$. Since $(e_n)_{n\in\mathbb{Z}^N}$ is orthogonal, and since it spans an l^∞-dense (and hence L^2-dense) subspace of $C(\mathbb{T}^N)$ by the Stone-Weierstrass Thm. 1.5.12, it follows that (e_n) is (after scaling each element) an orthonormal basis of $L^2(\mathbb{T}^N, m)$. Thus, by Thm. 8.8.7, $-\widetilde{\Delta}_{\mathrm{D}}$ has compact resolvent. This completes the proof. □

Remark 8.9.2. Assume that Ω has smooth boundary. Then one can show that the negative Neumann Laplacian $-\Delta_{\mathrm{N}} = \overline{\mathrm{div}} \cdot \overline{\mathrm{div}}^*$ also has compact resolvent. As in the proof of Thm. 8.9.1, this is equivalent to showing that div^* has compact graph projection, equivalently, that the linear map

$$H^1(\Omega) \to L^2(\Omega, m) \qquad f \mapsto f$$

is compact where $H^1(\Omega) = \mathscr{D}(\mathrm{div}^*)$ is equipped with the graph inner product of div^*.

To prove this, we view Ω as an open subset of $\mathbb{T}^N$ with smooth boundary. The key point is that there exists a bounded linear map $H^1(\Omega) \to H^1(\mathbb{T}^N)$ whose composition with the restriction map $H^1(\mathbb{T}^N) \to H^1(\Omega)$ is the identity on $H^1(\Omega)$. Consequently, it suffices to show that the adjoint of the divergence operator on $\mathbb{T}^N$ has compact graph projection; by Thm. 8.8.10, this reduces to proving that the Neumann Laplacian on $\mathbb{T}^N$ has compact resolvent.

This can be established in exactly the same way as for the Dirichlet Laplacian on $\mathbb{T}^N$. Indeed, the two Laplacians coincide: both extend the smooth Laplacian on $\mathbb{T}^N$, which is diagonalizable and therefore essentially self-adjoint by Prop. 7.7.9. For details, we refer the reader to [Tay-1, Sec. 4.4]. $\square$

8.9.2 Weak solutions

Corollary 8.9.3. *Let $a_1, \dots, a_N, b \in L^\infty(\Omega, m)$. Consider the unbounded operator*

$$T := -\Delta_{\mathrm{D}} + \sum_{j=1}^{N} \mathbf{M}_{a_j} X_j + \mathbf{M}_b$$

on $L^2(\Omega, m)$, as described in Thm. 7.9.17, whose domain satisfies $\mathscr{D}(T) = \mathscr{D}(\Delta_{\mathrm{D}})$. Then T has compact resolvent. Consequently, by Thm. 8.8.11, T is injective iff T is surjective.

Proof. By Thm. 8.9.1, $-\Delta_{\mathrm{D}}$ has compact resolvent. As noted in the proof of Thm. 7.9.17, the operator $\sum_{j=1}^{N} \mathbf{M}_{a_j} X_j + \mathbf{M}_b$ is $-\Delta_{\mathrm{D}}$-bounded with arbitrarily small bound. Therefore T has compact resolvent by Cor. 8.8.6. $\square$

Remark 8.9.4. In Cor. 8.9.3, assume that Ω has smooth boundary, that $b = 0$, and that $a_1, \dots, a_N$ are real-valued smooth functions on an open set containing $\overline{\Omega}$. In this situation one can show that T is injective, and hence is surjective by Cor. 8.9.3. Therefore, for each $g \in L^2(\Omega, m)$, the differential equation (7.35) admits a weak solution f.

The injectivity of T follows from the following facts. By the regularity theory of Sobolev spaces, any $f \in \mathrm{Ker}(T)$ must in fact lie in $C^k(\overline{\Omega})$ for each $k \in \mathbb{N}$, and $f|_{\partial\Omega} = 0$; see [Tay-1, Sec. 5.1]. Moreover, by the maximum principle, the maximum of $|f|$ must occur on the boundary; see [Tay-1, Sec. 5.2]. Hence $f = 0$. $\square$

In the special case $a_1 = \cdots = a_N = b = 0$, the surjectivity of T can be established more directly, as shown below.

Theorem 8.9.5 (Poincaré inequality). *The self-adjoint operator $-\Delta_{\mathrm{D}}$ is strictly positive. Hence $\mathrm{Rng}(-\Delta_{\mathrm{D}}) = \mathcal{H}$ by Prop. 7.1.2.*

Proof. It suffices to find $\lambda > 0$ and verify $\langle \overline{\nabla} f | \overline{\nabla} f \rangle \geqslant \lambda \langle f | f \rangle$ for each $f \in \mathscr{D}(\overline{\nabla})$. Moreover, since $\overline{\nabla}$ is the closure of ∇, it suffices to prove

$$\langle \nabla f | \nabla f \rangle \geqslant \lambda \langle f | f \rangle \tag{8.30}$$

for each $f \in C_c^\infty(\Omega)$. Indeed, we will prove this inequality for all $f \in C_c^\infty(B)$ where $B = (-R, R)^N$ is a bounded open cube containing Ω.

In the case $N = 1$, for each $x \in I$ we have $f(x) = \int_{-R}^{x} f'$, and hence

$$|f(x)| \leqslant \int_{-R}^{R} |f'| \leqslant \sqrt{2R} \cdot \sqrt{\int_{-R}^{R} |f'|^2}$$

where the second inequality is due to Cauchy-Schwarz. Thus

$$\int_{-R}^{R} |f|^2 \leqslant 2R \sup_{-R<x<R} |f(x)|^2 \leqslant 4R^2 \int_{-R}^{R} |f'|^2$$

For general N, the same reasoning yields

$$\int_{-R}^{R} |f|^2 dx_1 \leqslant 4R^2 \int_{-R}^{R} |\partial_{x_1} f|^2 dx_1 \leqslant 4R^2 \int_{-R}^{R} |\nabla f|^2 dx_1$$

Integrating over the remaining variables gives

$$\int_B |f|^2 \leqslant 4R^2 \int_B |\nabla f|^2$$

This proves the desired inequality with $\lambda = 1/(4R^2)$. □

8.10 Problems

Let $\mathcal{U}, \mathcal{V}, \mathcal{W}, \mathcal{X}$ be normed vector spaces over $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. Let $\mathcal{H}, \mathcal{K}, \mathcal{M}$ be Hilbert spaces.

Problem 8.1. Let $T : \mathcal{V} \to \mathcal{W}$ be a compact operator. Let $A : \mathcal{W} \to \mathcal{X}$ and $B : \mathcal{U} \to \mathcal{V}$ be bounded linear maps.

1. Prove that $ATB : \mathcal{U} \to \mathcal{X}$ is a compact operator.

2. Assume that $\mathcal{U}, \mathcal{V}, \mathcal{W}, \mathcal{X}$ are Hilbert spaces. Use the complete continuity of T to prove that ATB is completely continuous.

Problem 8.2. Prove that any bounded sequence (ξ_n) in $\mathcal{H}$ admits a weakly convergent subsequence.

Hint. When $\mathcal{H}$ is separable, apply Banach-Alaoglu and Prop. 2.6.6. Reduce the general case to the separable case by considering the closure of $\mathrm{Span}\{\xi_n\}$. □

Problem 8.3. Let V be a dense linear subspace of $\mathcal{H}$. Let $T \in \mathfrak{L}(\mathcal{H}, \mathcal{K})$. Prove the equivalence of the following statements.

(1) T is compact.

(2) For each sequence (ξ_n) in $\overline{B}_V(0,1)$ converging weakly to 0, we have $\lim_n T\xi_n = 0$.

(3) For each sequence (ξ_n) in $\overline{B}_{\mathcal{H}}(0,1)$ converging weakly to 0, we have $\lim_n T\xi_n = 0$.

(4) For each sequence (ξ_n) in $\overline{B}_{\mathcal{H}}(0,1)$ converging weakly to some $\xi \in \overline{B}_{\mathcal{H}}(0,1)$, we have $\lim_n T\xi_n = T\xi$.

Problem 8.4. Let X, Y be compact Hausdorff spaces. Let ν be a finite Borel measure on Y. Let $K \in C(X \times Y)$. Prove that there is a well-defined compact map

$$T : C(Y) \to C(X) \qquad (Tf)(x) = \int_X K(x,y)f(y)d\nu(y)$$

(In particular, explain why Tf belongs to $C(X)$.)

Hint. Method 1: Show that T is approximable by using the Stone-Weierstrass theorem to prove that K lies in the closure of $\mathrm{Span}\{gh : g \in C(X), h \in C(Y)\}$.

Method 2: Apply the Arzelà-Ascoli Thm. 1.4.37. □

The following result is known as the **Schur test**.

Problem 8.5. Let μ and ν be σ-finite Borel measures on LCH spaces X and Y respectively. Let $K : X \times Y \to \mathbb{C}$ be a Borel function. Suppose there exist $a, b \in \mathbb{R}_{\geqslant 0}$ such that

$$\sup_{x \in X} \int_Y |K(x,y)| d\nu(y) \leqslant a \qquad \sup_{y \in Y} \int_X |K(x,y)| d\mu(x) \leqslant b$$

Prove that for each $f \in L^2(Y, \nu)$, the integral $\int_Y K(x,y)f(y)d\nu(y)$ can be defined for a.e. $x \in X$, and that we have a bounded linear map

$$T : L^2(Y,\nu) \to L^2(X,\mu) \qquad (Tf)(x) = \int_Y K(x,y)f(y)d\nu(y)$$

satisfying $\|T\| \leqslant \sqrt{ab}$.

Hint. Apply Cauchy-Schwarz to $\int_Y |K(x,y)|^{\frac{1}{2}} \cdot |K(x,y)|^{\frac{1}{2}}|f(y)|d\nu(y)$. □

The following problem establishes the compactness of the integral operators arising from the method of double layer potentials introduced in Subsec. 8.1.3 and 8.1.4.

Problem 8.6. Let X be a compact subset of $\mathbb{R}^{N-1}$ (where $N = \mathbb{Z}_+$). Assume that $K : X \times X \to \mathbb{C}$ is continuous outside the diagonal $\{(x,x) : x \in X\}$. Assume moreover that there exists $M \in \mathbb{R}_{\geqslant 0}$ such that

$$|K(x,y)| \leqslant M|x-y|^{2-N} \qquad \text{for each } x, y \in X$$

1. Prove that there is a well-defined compact linear operator $T : C(X) \to C(X)$ such that for each $f \in C(X)$ and $x \in X$ we have

$$(Tf)(x) = \int_X K(x,y)f(y)dm(y) \tag{8.31}$$

2. Prove that there is a well-defined compact linear operator $S : L^2(X,m) \to L^2(X,m)$ such that for each $f \in L^2(X,m)$, Eq. (8.31) holds (with T replaced by S) for a.e. $x \in X$.

Hint. 1. For each $\varepsilon > 0$, let $h_\varepsilon : \mathbb{R}_{\geqslant 0} \to [0,1]$ be a continuous function satisying $h_\varepsilon|_{[0,\varepsilon]} = 0$ and $h_\varepsilon|_{[2\varepsilon,+\infty)} = 1$. Let $T_\varepsilon : C(X) \to C(X)$ be defined by the integral kernel $K_\varepsilon(x,y) = K(x,y) \cdot h_\varepsilon(|x-y|)$, which is compact by Pb. 8.4. Show that $\lim_{\varepsilon\to 0} T_\varepsilon$ converges to in norm to T and apply Thm. 8.4.15.

2. Adapt the argument in Part 1 and use the Schur test (cf. Pb. 8.5). □

Recall Def. 8.6.4 for the meaning of ε-orthogonality.

Problem 8.7. Let $\mathcal{U}$ be a linear subspace of $\mathcal{V}$. Let $0 \leqslant \varepsilon < 1$.

1. Prove that the set of vectors of $\mathcal{V}$ that are ε-orthogonal to $\mathcal{U}$ is a closed subset of $\mathcal{V}$.

2. Prove that the only vector in $\mathcal{U}$ that is ε-orthogonal to $\mathcal{U}$ is the zero vector.

3. Assume that $\mathcal{U}$ is finite dimensional. Prove that for any $\xi \in \mathcal{V}$, there exists $\widetilde{\xi} \in \xi + \mathcal{U}$ that is 0-orthogonal to $\mathcal{U}$.

Problem 8.8. Let $T \in \mathfrak{L}(\mathcal{V})$ be a compact operator. Let $\lambda \in \mathbb{F}\backslash\{0\}$.

1. Prove that there exists $\delta > 0$ such that for each $\xi \in \mathcal{V}$,

$$\|(T-\lambda)\xi\| \geqslant \delta \cdot \mathrm{dist}(\xi, \mathrm{Ker}(T-\lambda)) \tag{8.32}$$

2. Assume that $\mathcal{V}$ is complete. Prove that for each $n \in \mathbb{Z}_+$, $\mathrm{Rng}((T-\lambda)^n)$ is a closed linear subspace of $\mathcal{V}$.

As we will see in Prop. 10.3.1, inequality (8.32) admits a topological interpretation: the operator $T-\lambda$ is an open map from $\mathcal{V}$ onto its range.

Proof. 1. Assume not. Prove that there is a sequence of unit vectors (ξ_n) that are 0-orthogonal [10] to $\mathrm{Ker}(T-\lambda)$ and satisfy $\lim_n(T-\lambda)\xi_n=0$. Then argue as in the proof of Thm. 8.5.4.

2. Reduce to the case $n=1$ by writing $(T-\lambda)^n$ as the sum of a compact operator and a non-zero scalar. Use the following Lem. 8.10.1. □

Lemma 8.10.1. *The normed vector space $\mathcal{V}$ is complete iff for each sequence $(v_n)_{n\in\mathbb{Z}_+}$ in $\mathcal{V}$ satisfying $\sum_n\|v_n\|<+\infty$, the series $\sum_n v_n$ converges.*

Proof. "$\Rightarrow$": This is obvious.

"$\Leftarrow$": Let (ξ_n) be a Cauchy sequence in $\mathcal{V}$. Then it has a subsequence (ξ_{n_k}) satisfying $\|\xi_{n_{k+1}}-\xi_{n_k}\|\leqslant 2^{-k}$ for each k. Set $v_k=\xi_{n_{k+1}}-\xi_{n_k}$. Since $\sum_k\|v_k\|<+\infty$, by assumption, the series $\sum_k v_k$ converges, and hence (ξ_{n_k}) converges to some $\xi\in\mathcal{V}$. One shows easily that $\lim_n\xi_n=\xi$. □

The following result, also obtained by Riesz in [Rie18], generalizes the Fredholm alternative.

Problem 8.9. Assume that $\mathcal{V}$ is complete. Let $T\in\mathfrak{L}(\mathcal{V})$ be compact and $\lambda\in\mathbb{F}\backslash\{0\}$.

1. Prove that there exists $\nu\in\mathbb{Z}_+$ such that for each $k\in\mathbb{N}$,

$$\mathrm{Ker}((T-\lambda)^\nu)=\mathrm{Ker}((T-\lambda)^{\nu+k})\qquad \mathrm{Rng}((T-\lambda)^\nu)=\mathrm{Rng}((T-\lambda)^{\nu+k})$$

2. For any ν as in Part 1, prove that $\mathcal{V}$ admits an (algebraic) direct sum

$$\mathcal{V}=\mathrm{Ker}((T-\lambda)^\nu)+\mathrm{Rng}((T-\lambda)^\nu)$$

   that is, any vector $\xi\in\mathcal{V}$ can be written uniquely as $\xi=\xi_1+\xi_2$ where $\xi_1\in\mathrm{Ker}((T-\lambda)^\nu)$ and $\xi_2\in\mathrm{Rng}((T-\lambda)^\nu)$.

3. Prove that $T-\lambda$ is a **Fredholm operator of index 0**; that is,

$$\dim\mathrm{Ker}(T-\lambda)=\dim\mathcal{V}/\mathrm{Rng}(T-\lambda)$$

Hint. 1. The first half has been proved in Prop. 8.7.3.

2. Prove the more general fact that if A is a linear operator on a vector space W satisfying $\mathrm{Ker}(A)=\mathrm{Ker}(A^2)$ and $\mathrm{Rng}(A)=\mathrm{Rng}(A^2)$, then W equals the (algebraic) direct sum of $\mathrm{Ker}(A)$ and $\mathrm{Rng}(A)$.

3. Show that $V_1:=\mathrm{Ker}((T-\lambda)^\nu)$ and $V_2=\mathrm{Rng}((T-\lambda)^\nu)$ are both T-invariant, and that $(T-\lambda)|_{V_2}:V_2\to V_2$ is bijective. Apply the rank-nullity theorem to $(T-\lambda)|_{V_1}$. □

[10] Indeed, $\frac{1}{n}$-orthogonality is sufficient.

Part 2 of the following problem generalizes the classical Poincaré inequality (8.30).

Problem 8.10. Let $T : \mathcal{H} \to \mathcal{K}$ be an unbounded closed operator. Assume that T has compact graph projection.

1. Prove that $\dim \mathrm{Ker}(T) < +\infty$.

2. (**Abstract Poincaré inequality**) Let $P \in \mathfrak{L}(\mathcal{H})$ be the projection onto $\mathrm{Ker}(T)$. Prove that there exists $\gamma \in \mathbb{R}_{>0}$ such that

$$\|T\xi\| \geqslant \gamma\|\xi - P\xi\| \qquad \text{for each } \xi \in \mathscr{D}(T) \tag{8.33}$$

3. Show that $\mathrm{Rng}(T)$ is a closed linear subspace of $\mathcal{K}$.

Hint for Part 2. Method 1. Note that $\mathrm{Ker}(T)$ is closed by Prop. 6.6.11. Assume that (8.33) does not holds for any γ. Show that there exists a sequence of unit vectors (ξ_n) in $\mathscr{D}(T) \cap \mathrm{Ker}(T)^\perp$ such that $\lim_n \|T\xi_n\| = 0$. Find a contradiction.

Method 2. Show that $\mathrm{Ker}(T) = \mathrm{Ker}(|T|)$ and $\|T\xi\| = \||T|\xi\|$ by using the polar decomposition of T. Then apply the Hilbert-Schmidt Thm. 8.8.7 to $|T|$. (This method actually implies that the smallest γ satisfying (8.33) equals the smallest non-zero eigenvalue of $|T|$.) □

The following problem is the unbounded-operator analogue of Pb. 8.9-3, generalizing the Fredholm alternative Thm. 8.8.11.

Problem 8.11. Let T be an unbounded operator on $\mathcal{H}$ with compact resolvent. Prove that

$$\dim \mathrm{Ker}(T) = \dim(\mathrm{Rng}T)^\perp \equiv \dim \mathrm{Ker}(T^*)$$

Problem 8.12. Let $T : \mathcal{H} \to \mathcal{K}$ and $S : \mathcal{H} \to \mathcal{M}$ be unbounded operators. Assume that S is T-bounded with bound $\alpha > 0$ (cf. Def. 7.9.1). Prove that if S has compact graph projection, then so does T.

Problem 8.13. Let $V : \mathbb{R}^N \to \mathbb{R}_{\geqslant 0}$ be a locally bounded (i.e., bounded on compact sets) Borel function satisfying $\lim_{|x|\to+\infty} V(x) = +\infty$. Let $\mathcal{H} = L^2(\mathbb{R}^N, m)$. Define an unbounded operator

$$T : \mathcal{H} \to \mathcal{H}^{\oplus(N+1)} \qquad \xi \mapsto \nabla\xi \oplus \sqrt{V}\xi$$

with domain $\mathscr{D}(T) = C_c^\infty(\mathbb{R}^N)$.

1. Prove that T is closable.

2. Prove that T has compact graph projection.

Hint. 1. Show that T^* is densely defined.

2. For each $r \geqslant 0$, choose $h_r \in C_c^\infty(\mathbb{R}^N)$ such that $0 \leqslant h_r \leqslant 1$, that $h_r(x) = 1$ if $|x| \leqslant r$, and $h_r(x) = 0$ if $|x| \geqslant r+1$. Let (ξ_n) be a sequence in $C_c^\infty(\mathbb{R}^N)$ such that $\sup_n \|T\xi_n\| < +\infty$. Using the Rellich compactness Thm. 8.9.1 (i.e. ∇ with domain $C_c^\infty(\Omega)$ has compact graph projection if $\Omega \subset \mathbb{R}^N$ is bounded and open), together with the diagonal method/Tychonoff theorem, find a subsequence (ξ_{n_k}) such that $\lim_k h_r\xi_{n_k}$ converges for each $r \in \mathbb{Z}_+$. Show that there exists $C_r \geqslant 0$ with $\lim_{r\to+\infty} C_r = +\infty$ such that

$$\|\sqrt{V}\xi\| \geqslant C_r\|(1-h_r)\xi\| \qquad \text{for each } \xi \in C_c^\infty(\mathbb{R}^N)$$

Conclude that $(\xi_{n_k})_k$ is a Cauchy sequence. □

Remark 8.10.2. Let T be as in Pb. 8.13. Let A be the positive operator on $\mathcal{H} = L^2(\mathbb{R}^N, m)$ defined by

$$A = -\Delta + \mathbf{M}_V \qquad \mathscr{D}(A) = C_c^\infty(\mathbb{R}^N)$$

Then T is an abstract gradient operator of A (cf. Def. 7.2.1). Therefore, by Def. 7.2.6, the Friedrichs extension A_{F} of A equals $\overline{T}^*\overline{T}$. Hence, by Prop. 8.8.2 and Thm. 8.8.10, the compactness of the graph projection of T (established in Pb. 8.13) is equivalent to A_{F} having compact resolvent.

8.A Hilbert-Schmidt theorem for compact normal operators

Let $\mathcal{H}$ be a Hilbert space. In this appendix section, we extend the Hilbert-Schmidt theorem to compact normal operators on $\mathcal{H}$, primarily for the sake of completeness. Since in most applications it suffices to treat compact self-adjoint operators (rather than compact normal ones)—see Rem. 8.8.8 for an exception—the reader may safely skip this section.

We begin with the following easy observation:

Lemma 8.A.1. *Let $T \in \mathfrak{L}(\mathcal{H})$ be normal. Let $\lambda_1, \lambda_2 \in \mathbb{C}$ be distinct. Then $\mathrm{Ker}(\lambda_1 - T) \perp \mathrm{Ker}(\lambda_2 - T)$.*

Proof. By Prop. 5.8.8, the projection onto $\mathrm{Ker}(\lambda_j - T)$ equals $\chi_{\{\lambda_j\}}(T)$. Since

$$\chi_{\{\lambda_1\}}(T)\chi_{\{\lambda_2\}}(T) = (\chi_{\{\lambda_1\}}\chi_{\{\lambda_2\}})(T) = 0$$

the two eigenspaces are orthogonal by Cor. 5.2.5. □

Theorem 8.A.2 (Hilbert-Schmidt theorem). *Let $T \in \mathfrak{L}(\mathcal{H})$. Then the following conditions (1) and (2) are equivalent.*

(1) T is a compact normal operator.

(2) *$\mathcal{H}$ admits an orthonormal basis $(e_1, e_2, \dots) \cup (f_j)_{j\in J}$, where the countable family $(e_1, e_2, \dots)$ is possibly finite, such that:*

(2a) *For each n, we have $Te_n = \lambda_n e_n$ for some $\lambda_n \in \mathbb{C}$, and $Tf_j = 0$ for each j.*

(2b) *If the sequence $(e_1, e_2, \dots)$ is infinite, then $\lim_{n\to\infty} \lambda_n = 0$.*

Proof. Step 1. The direction (2)⇒(1) can be proved exactly as in Thm. 8.3.1. Thus, we focus on proving (1)⇒(2). Assume that T is compact and normal. In this step, we show that $\mathrm{Sp}(T)\backslash\{0\}$ is a discrete set.

Suppose not. Then $\mathrm{Sp}(T)\backslash\{0\}$ has an accumulation point λ. Hence, then there exists a sequence (λ_n) in $\mathrm{Sp}(T)\backslash\{0, \lambda\}$ converging to λ. By removing repetitions, we may assume that the numbers λ_n are distinct. By Thm. 5.8.11, each λ_n is an approximate eigenvalue of T, and hence is an eigenvalue by Thm. 8.5.4 and the compactness of T. Choose a unit vector $\xi_n \in \mathrm{Ker}(T - \lambda_n)$. Then (ξ_n) is an orthonormal sequence by Lem. 8.A.1, and hence converges weakly to zero by Exp. 3.7.4. Therefore $\lim_n T\xi_n = 0$ by the complete continuity of T. But

$$\|T\xi_n\| = \|\lambda_n \xi_n\| = |\lambda_n| \to |\lambda| \neq 0$$

This gives a contradiction.[11]

Step 2. Since $\mathrm{Sp}(T)$ is a bounded set with no accumulation points outside 0, we can write $\mathrm{Sp}(T)$ as $\{\kappa_1, \kappa_2, \dots\}$ where (κ_n) is a possibly finite sequence of *distinct* complex numbers converging to 0 if the sequence is infinite.[12] Thus $1_{\mathrm{Sp}(T)} = \sum_n \chi_{\{\kappa_n\}} + \chi_{\{0\}}$, and hence the following identity holds in WOT

$$\mathbf{1}_{\mathcal{H}} = \sum_n P_n + Q \qquad \text{where } P_n = \chi_{\{\kappa_n\}}(T), Q = \chi_{\{0\}}(T)$$

because the Borel functional calculus of T is normal (cf. Thm. 5.7.13).

Since $P_1, P_2, \dots$ and Q are mutually orthogonal, by Rem. 5.10.6, the sum $\sum_n P_n + Q$ converges in SOT to the projection onto $\bigoplus_n \mathrm{Rng}(P_n) \oplus \mathrm{Rng}(Q)$. Therefore, we have

$$\mathcal{H} = \bigoplus_n \mathrm{Rng}(P_n) \oplus \mathrm{Rng}(Q)$$

By Prop. 5.8.8, this relation is equivalent to the orthogonal decomposition

$$\mathcal{H} = \bigoplus_n \mathrm{Ker}(T - \kappa_n) \oplus \mathrm{Ker}(T)$$

[11] Alternatively, since ω_T is completely continuous, we have $\lambda_n = \langle \xi_n | T\xi_n \rangle \to 0$, which is impossible.

[12] In the original statement of the theorem, $\lambda_1, \lambda_2, \dots$ are not assumed to be distinct; hence we use κ_n here.

Choosing an orthonormal basis for each $\mathrm{Ker}(T)$ and $\mathrm{Ker}(T-\kappa_n)$, we obtain an orthonormal basis of $\mathcal{H}$.

By Exp. 8.6.1, each $\mathrm{Ker}(T-\kappa_n)$ is finite dimensional. Therefore, the proof of the theorem is completed by taking $(f_j)_{j\in J}$ to be an orthonormal basis of $\mathrm{Ker}(T)$, and letting $(e_1, e_2, \dots)$ be the union of the orthonormal bases of all $\mathrm{Ker}(T-\kappa_n)$.[13] □

Exercise 8.A.3. In Thm. 8.A.2, show that $T = T^*$ iff $\lambda_n \in \mathbb{R}$ for all n; show that $T \geqslant 0$ iff $\lambda_n \in \mathbb{R}_{\geqslant 0}$ for all n.

[13] In particular, λ_n is the unique element in $\{\kappa_1, \kappa_2, \dots\}$ such that $e_n \in \mathrm{Ker}(T-\lambda_n)$.

9 Weak topology and the separation of convex sets

9.1 Indeterminate moment problems: the linear-extension approach

Let $\mathcal{V}$ be a normed vector space over $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$.

9.1.1 Indeterminate moment problems of bounded type

Recall that in a moment problem, one is given a sequence of scalars (c_n) and a sequence of functions (ξ_n), and one asks for the existence (and sometimes uniqueness) of a function f (or a measure) such that

$$\int \xi_n f = c_n \qquad \text{resp.} \qquad \int \xi_n df = c_n \tag{9.1}$$

for all n. There are two general types of moment problems: those of positive type and those of bounded type. The polynomial moment problem studied in Ch. 4 is of **positive type**; via the GNS construction (cf. Rem. 4.6.1), it is related to the spectral theory of bounded or unbounded Hermitian operators. Likewise, the trigonometric moment problems in Pb. 5.6 and 7.10 are of positive type and, through the GNS construction, correspond to the spectral theorem for strongly continuous unitary representations of the groups $\mathbb{Z}^N$ and $\mathbb{R}^N$, respectively.

Among these moment problems, the polynomial moment problems on the intervals $(-\infty, +\infty)$ and $[0, +\infty)$ (i.e., the Hamburger and Stieltjes moment problems, cf. Thm. 4.2.9) are special because their solutions need not be unique. This reflects the fact that a Hermitian operator may admit multiple self-adjoint extensions. Through the Cayley transform, the non-uniqueness of self-adjoint extensions corresponds to the non-uniqueness of extensions of unitary maps (cf. Thm. 6.3.5).

The topic of this chapter originates from the study of moment problems of bounded type, initiated by Riesz in [Rie10], which can be reformulated in terms of the dual spaces of normed vector spaces. In particular, non-uniqueness of solutions to such a moment problem corresponds exactly to the non-uniqueness of extensions of bounded linear functionals.

From a modern viewpoint, moment problems of **bounded type** can be formulated as follows. Let $(\xi_n)_{n\in\mathbb{Z}_+}$ be a sequence in $\mathcal{V}$. Let $(c_n)_{n\in\mathbb{Z}_+}$ be a sequence in $\mathbb{F}$. We seek necessary and sufficient conditions for the existence of some $\varphi \in \mathcal{V}^*$ such that

$$\langle \xi_n, \varphi \rangle = c_n \qquad \text{for all } n \tag{9.2}$$

Once this task is resolved, and once the dual space $\mathcal{V}^*$ is identified as a function space, the problem (9.2) immediately becomes a moment problem of the form

(9.1). Indeed, in the same paper [Rie10] in which Riesz studied the moment problem for $\mathcal{V} = L^p([a,b],m)$ (with $1 < p < +\infty$), he also identified the dual space $\mathcal{V}^*$ with $L^q([a,b],m)$ where $\frac{1}{p} + \frac{1}{q} = 1$, the earliest version of Thm. 1.6.16.

A necessary condition for (9.2) to have a solution is the existence of some $M \in \mathbb{R}_{\geqslant 0}$ such that

$$\Big|\sum_n a_n c_n\Big| \leqslant M \Big\|\sum_n a_n \xi_n\Big\| \qquad \text{for each } (a_n)_{n\in\mathbb{Z}_+} \in C_c(\mathbb{Z}_+,\mathbb{F}) \tag{9.3a}$$

where $C_c(\mathbb{Z}_+,\mathbb{F})$ is understood as the set of functions $\mathbb{Z}_+ \to \mathbb{F}$ with finite supports. Indeed, if $\varphi \in \mathcal{V}^*$ satisfies (9.2), then

$$\Big|\sum_n a_n c_n\Big| = \Big|\Big\langle \sum_n a_n \xi_n, \varphi\Big\rangle\Big| \leqslant \|\varphi\| \cdot \Big\|\sum_n a_n \xi_n\Big\|$$

So (9.3a) holds for any $M \geqslant \|\varphi\|$.

The difficulty lies in the implication (9.3a)⇒(9.2). From a modern perspective, (9.3a) is equivalent to that the linear functional

$$\varphi : \mathcal{U} \to \mathbb{F} \qquad \sum_n a_n \xi_n \mapsto \sum_n a_n c_n \qquad \text{is well-defined and bounded} \tag{9.3b}$$

where $\mathcal{U} = \mathrm{Span}\{\xi_n : n \in \mathbb{Z}_+\}$. Thus, solving (9.2) is equivalent to extending the linear functional φ in (9.3b) to a bounded linear functional on $\mathcal{V}$.

By Cor. 2.4.3, φ extends uniquely to a bounded functional on the closure $\overline{\mathcal{U}}$. Hence solutions of (9.2) exist and are unique precisely when (ξ_n) spans a dense linear subspace of $\mathcal{V}$; this is called the **determinate type**. When (ξ_n) does not span densely, extensions of $\varphi \in \overline{\mathcal{U}}^*$ to $\mathcal{V} \to \mathbb{F}$ may fail to be unique, and the moment problem is then said to be of **indeterminate type**.

When $\mathcal{V}$ is a Hilbert space, the implication (9.3b)⇒(9.2) is easy to verify. By the Riesz-Fréchet Thm. 3.5.3, the linear functional $\varphi \in \overline{\mathcal{U}}^*$ is represented by some $f \in \overline{\mathcal{U}}$, that is, $\varphi(u) = \langle f|u\rangle$ where $u \in \overline{\mathcal{U}}$. The extension $\varphi \in \mathcal{V}^*$ is then given by $\varphi(v) = \langle f|v\rangle$ for $v \in \mathcal{V}$. Thus, in the case of L^2-spaces, the moment problem (9.2) is completely solved. It was precisely in order to study moment problems in more general function spaces that Riesz, in [Rie10], introduced for the first time the now standard L^p-spaces $L^p([a,b],m)$, though only for $1 < p < +\infty$.

Although Riesz was a pioneer in applying the method of linear extensions to spectral theory in [Rie13], in his earlier work [Rie10] (and similarly in [Rie11], where he studied the moment problem of bounded type for $\mathcal{V} = C([a,b])$), he adopted the method of finite approximation. This aligns with the paradigm shift described in (5.19a). Contrary to the chronological order, we will first study the linear-extension approach in this section and postpone the finite-approximation method to Sec. 9.2.

9.1.2 The Helly-Hahn-Banach theorem

The Hahn-Banach theorem provides a complete solution to the moment problem of finite type for any normed vector space $\mathcal{V}$. Despite its name, the theorem was first proved by Helly in 1912 [Hel12] for $C([a,b],\mathbb{R})$; moreover, Helly's proof is essentially identical to the argument used in the general case of real normed vector spaces.[1] It was later rediscovered by Hahn in 1927 and again by Banach in 1929. Thus, although we follow the usual convention and call it the Hahn-Banach theorem, it would be more accurate to refer to it as the **Helly-Hahn-Banach** theorem.

Theorem 9.1.1 (Hahn-Banach theorem). *Let $\mathcal{U}$ be a linear subspace of $\mathcal{V}$. Then the following linear map is surjective*

$$\mathcal{V}^* \to \mathcal{U}^* \qquad \varphi \mapsto \varphi|_{\mathcal{U}}$$

The proof of Thm. 9.1.1 for the case $\mathbb{F} = \mathbb{R}$ relies on the following crucial lemma discovered by Helly in [Hel12].

Lemma 9.1.2. *Assume that $\mathbb{F} = \mathbb{R}$, and $\mathcal{U}$ is a linear subspace of $\mathcal{V}$. Let $\varphi \in \mathcal{U}^* = \mathfrak{L}(\mathcal{U},\mathbb{R})$ with operator norm $\|\varphi\| \leqslant 1$. Assume that $e \in \mathcal{V}\backslash\mathcal{U}$, and let $\widetilde{\mathcal{U}} = \mathcal{U} + \mathbb{R}e$. Then φ can be extended to a linear functional $\widetilde{\varphi} : \widetilde{\mathcal{U}} \to \mathbb{R}$ such that $\|\widetilde{\varphi}\| \leqslant 1$.*

Proof. Let $c \in \mathbb{R}$ whose value will be determined later. Since any vector in $\widetilde{\mathcal{U}}$ can be written uniquely as $\xi - \lambda e$ where $u \in \mathcal{U}$ and $\lambda \in \mathbb{R}$, we can define

$$\widetilde{\varphi} : \widetilde{\mathcal{U}} \to \mathbb{R} \qquad \widetilde{\varphi}(u - \lambda e) = \varphi(u) - \lambda c$$

Suppose we have $\|\widetilde{\varphi}\| \leqslant 1$, that is, $|\widetilde{\varphi}(\xi)| \leqslant \|\xi\|$ for each $\xi \in \widetilde{\mathcal{U}}$. It suffices to prove

$$\widetilde{\varphi}(\xi) \leqslant \|\xi\| \qquad \text{for each } \xi \in \widetilde{\mathcal{U}}$$

Then, we also have $\widetilde{\varphi}(-\xi) \leqslant \|-\xi\|$, and hence $|\widetilde{\varphi}(\xi)| \leqslant \|\xi\|$, finishing the proof.

Now, our goal is to prove $\varphi(u) - \lambda c \leqslant \|u - \lambda e\|$ for all $u \in \mathcal{U}, \lambda \in \mathbb{R}$. Clearly this is true when $\lambda = 0$. Thus, we may assume WLOG that $\lambda \neq 0$. Moreover, dividing both sides by $|\lambda|$, we see that it suffices to prove for each $u \in \mathcal{U}$ that

$$\varphi(u) - c \leqslant \|u - e\| \qquad \varphi(u) + c \leqslant \|u + e\|$$

Thus, the goal now is to prove that there exists $c \in \mathbb{R}$ such that

$$\varphi(u) - \|u - e\| \leqslant c \leqslant -\varphi(u) + \|u + e\| \qquad \text{for each } u \in \mathcal{U}$$

[1] The original paper [Hel12] is difficult to locate online; see [Hoc80] for an overview.

For that purpose, it suffices to prove

$$\sup_{u\in\mathcal{V}}\big(\varphi(u)-\|u-e\|\big)\leqslant\inf_{v\in\mathcal{V}}\big(-\varphi(v)+\|v+e\|\big) \tag{9.4}$$

namely, to prove that $\varphi(u)-\|u-e\|\leqslant-\varphi(v)+\|v+e\|$ for all $u,v\in\mathcal{V}$. Using $\|\varphi\|\leqslant 1$, we compute

$$\varphi(u)+\varphi(v)=\varphi(u+v)\leqslant\|(u-e)+(v+e)\|\leqslant\|u-e\|+\|v-e\|$$

This finishes the proof. □

Proof of Thm. 9.1.1 for $\mathbb{F}=\mathbb{R}$. Let us show that any $\varphi\in\mathcal{U}^*$ can be extended to a linear functional $\Phi:\mathcal{V}\to\mathbb{R}$ with $\|\Phi\|=\|\varphi\|$. Assume WLOG that $\varphi\neq 0$. By scaling φ, we assume further $\|\varphi\|=1$. Let

$$\begin{aligned}\mathscr{F}=\big\{(\mathcal{X},\Phi):&\mathcal{X}\text{ is a linear subspace of }\mathcal{V}\text{ and contains }\mathcal{U},\\ &\Phi\in\mathbb{R}^{\mathcal{X}}\text{ is linear and satisfies }\Phi|_{\mathcal{U}}=\varphi,\|\Phi\|=1\big\}\end{aligned}$$

Then $\mathscr{F}$ is nonempty since it contains $(\mathcal{U},\varphi)$. We view each $(\mathcal{X},\Phi)\in\mathscr{F}$ as a subset of $\mathbb{R}^{\mathcal{V}}$ so that $\mathscr{F}$ is a subset of $2^{\mathbb{R}^{\mathcal{V}}}$. Equip $\mathscr{F}$ with the partial order $\subset$. Then every totally ordered subset $\mathscr{P}$ of $\mathscr{F}$ has an upper bound in $\mathscr{F}$, defined by taking the union of all $(\mathcal{X},\Phi)\in\mathscr{P}$. Therefore, Zorn's lemma is available, which implies that $\mathscr{F}$ has a maximal element $(\mathcal{X},\Phi)$.

If $\mathcal{X}\neq\mathcal{V}$, we let $e\in\mathcal{V}\backslash\mathcal{X}$. Then by Lem. 9.1.2, Φ can be extended to $\widetilde{\Phi}\in\widetilde{\mathcal{X}}^*$ where $\widetilde{\mathcal{X}}=\mathcal{X}+\mathbb{R}e$, and $\|\widetilde{\Phi}\|=1$. So $(\widetilde{\mathcal{X}},\widetilde{\Phi})$ belongs to $\mathscr{F}$ and is strictly larger than $(\mathcal{X},\Phi)$, impossible. So $\mathcal{X}=\mathcal{V}$. □

Remark 9.1.3. When $\mathcal{V}$ is separable, Thm. 9.1.1 can be proved by induction instead of using Zorn's lemma. Indeed, as we will see in Subsec. 9.1.3, the reduction from the case $\mathbb{F}=\mathbb{C}$ to the case $\mathbb{F}=\mathbb{R}$ is fairly elementary and does not involve Zorn's lemma. Thus we may assume $\mathbb{F}=\mathbb{R}$.

Fix $\varphi\in\mathcal{U}^*$ with $\|\varphi\|=1$. Since $\mathcal{V}$ is separable, there exists a sequence $(e_n)_{n\in\mathbb{Z}_+}$ in $\mathcal{V}$ whose span is dense in $\mathcal{V}$. Let $\mathcal{U}_n=\mathcal{U}+\mathrm{Span}\{e_1,\dots,e_n\}$. By repeatedly applying Lem. 9.1.2, we obtain a sequence (φ_n) where each $\varphi_n:\mathcal{U}_n\to\mathbb{R}$ is a linear functional of norm 1 extending φ_{n-1} (with $\varphi_0=\varphi$). Hence φ extends to a linear functional $\Phi:\mathcal{U}_\infty\to\mathbb{R}$ of norm 1, where $\mathcal{U}_\infty=\mathcal{U}+\mathrm{Span}\{e_n:n\in\mathbb{Z}_+\}$. Since $\mathcal{U}_\infty$ is dense in $\mathcal{V}$, Cor. 2.4.3 allows us to extend Φ further to a norm-1 linear functional on all of $\mathcal{V}$.

Remark 9.1.4. The same argument used in the proof of the Hahn-Banach Thm. 9.1.1 yields the following stronger form of the Hahn-Banach theorem; see the second proof of Thm. 9.4.1 for an application.

Let V be an $\mathbb{R}$-vector space, let $p:V\to\mathbb{R}$ be a function satisfying

$$p(\lambda\xi)=\lambda p(\xi)\qquad p(\xi+\eta)\leqslant p(\xi)+p(\eta)$$

for each $\xi, \eta \in V$ and $\lambda > 0$. Let U be a linear subspace of V. Suppose that $\varphi : U \to \mathbb{R}$ is a linear functional satisfying

$$\varphi(\xi) \leqslant p(\xi) \tag{9.5}$$

for each $\xi \in U$. Then φ can be extended to a linear functional $\varphi : V \to \mathbb{R}$ satisfying (9.5) for each $\xi \in V$. □

9.1.3 The case $\mathbb{F} = \mathbb{C}$

The now-standard reduction of the complex Hahn-Banach theorem to the real case is due to Murray [Mur36] in 1936, many years after Helly's elegant proof of the real version in 1912. The long gap between these developments illustrates a familiar phenomenon in mathematics: once an idea is understood, the method appears obvious; but the truly difficult step is recognizing the direction in which one ought to think.

This reduction rests on a simple observation: Whether $\mathbb{F}$ is $\mathbb{R}$ or $\mathbb{C}$, linear functionals $\mathbb{F}^n \to \mathbb{F}$ can be identified with elements of $\mathbb{F}^n$. In particular, consider $\mathbb{C}^n$ equipped with the Euclidean (complex) inner product $\langle \cdot | \cdot \rangle$. Then the linear functionals $\mathbb{C}^n \to \mathbb{C}$ are precisely those of the form

$$\varphi : \mathbb{C}^n \to \mathbb{C} \qquad v \mapsto \langle \xi | v \rangle$$

for some $\xi \in \mathbb{C}^n$. On the other hand, $\mathbb{C}^n$ can also be viewed as the real Euclidean space $\mathbb{R}^{2n}$. Its real inner product, denoted by $\langle \cdot, \cdot \rangle_{\mathbb{R}}$, is easy to write:

$$\langle u, v \rangle_{\mathbb{R}} = \mathrm{Re}\langle u | v \rangle$$

Then the linear functionals $\mathbb{C}^n \simeq \mathbb{R}^{2n} \to \mathbb{R}$ are exactly those of the form

$$\psi : \mathbb{C}^n \to \mathbb{R} \qquad v \mapsto \langle \xi, v \rangle_{\mathbb{R}} = \mathrm{Re}\langle \xi | v \rangle$$

From the above discussion, we see that there is an $\mathbb{R}$-linear bijection between complex and real linear functionals on $\mathbb{C}^n$, related by

$$\psi(v) = \mathrm{Re}\varphi(v)$$

Since $\varphi(v) = \mathrm{Re}\varphi(v) + \mathbf{i}\mathrm{Im}\varphi(v)$ and $\mathrm{Im}\varphi(v) = -\mathrm{Re}(\mathbf{i}\varphi(v)) = -\mathrm{Re}\varphi(\mathbf{i}v)$, we see that φ can be recovered by ψ via

$$\varphi(v) = \mathrm{Re}\varphi(v) - \mathbf{i}\mathrm{Re}\varphi(\mathbf{i}v) \tag{9.6}$$

Moreover, since both φ and ψ correspond to the same vector $\xi \in \mathbb{C}^n$, their operator norms are both equal to $\|\xi\|$. This motivates the following property:

Proposition 9.1.5. *Let V be a $\mathbb{C}$-vector space. Then we have an $\mathbb{R}$-linear bijection*

$$\mathrm{Lin}(V,\mathbb{C}) \xrightarrow{\simeq} \mathrm{Lin}(V,\mathbb{R}) \qquad \varphi \mapsto \mathrm{Re}\varphi \tag{9.7}$$

(that is, $\varphi(v) = \mathrm{Re}\varphi(v)$ for each $v \in V$). Moreover, φ can be recovered from $\mathrm{Re}\varphi$ by the formula (9.6). In addition, if V is a $\mathbb{C}$-normed vector space, then

$$\|\varphi\| = \|\mathrm{Re}\varphi\| \qquad \text{for each } \varphi \in \mathrm{Lin}(V,\mathbb{C})$$

In particular, (9.7) restricts to an $\mathbb{R}$-linear isometric bijection

$$\mathfrak{L}(V,\mathbb{C}) \xrightarrow{\simeq} \mathfrak{L}(V,\mathbb{R}) \qquad \varphi \mapsto \mathrm{Re}\varphi$$

Proof. For each $\psi \in \mathrm{Lin}(V,\mathbb{R})$, the $\mathbb{R}$-linear map $\psi_{\mathbb{C}} : V \to \mathbb{C}$ defined by $\psi_{\mathbb{C}}(v) = \psi(v) - \mathbf{i}\psi(\mathbf{i}v)$ intertwines the scalar multiplication by $\mathbf{i}$, since

$$\mathbf{i}\psi_{\mathbb{C}}(v) = \mathbf{i}\psi(v) + \psi(\mathbf{i}v) = \psi_{\mathbb{C}}(\mathbf{i}v)$$

Therefore $\psi_{\mathbb{C}}$ is $\mathbb{C}$-linear. We thus have an $\mathbb{R}$-linear map

$$\mathrm{Lin}(V,\mathbb{R}) \to \mathrm{Lin}(V,\mathbb{C}) \qquad \psi \mapsto \psi_{\mathbb{C}}$$

It is the inverse map of (9.7), since one easily verifies that $(\mathrm{Re}\varphi)_{\mathbb{C}} = \varphi$ and $\mathrm{Re}(\psi_{\mathbb{C}}) = \psi$ for each $\varphi \in \mathrm{Lin}(V,\mathbb{C})$ and $\psi \in \mathrm{Lin}(V,\mathbb{R})$. This proves that (9.7) is a bijection.

It remains to compare the operator norms. Choose any $\varphi \in \mathrm{Lin}(V,\mathbb{C})$. Clearly $\|\mathrm{Re}\varphi\| \leqslant \|\varphi\|$. For each $v \in V$, by choosing $\lambda \in \mathbb{S}^1$ with $\lambda\varphi(v) \in \mathbb{R}_{\geqslant 0}$, we have

$$|\varphi(v)| = |\lambda\varphi(v)| = |\varphi(\lambda v)| = |\mathrm{Re}\varphi(\lambda v)| \leqslant \|\mathrm{Re}\varphi\| \cdot \|\lambda v\| = \|\mathrm{Re}\varphi\| \cdot \|v\|$$

Hence $\|\varphi\| \leqslant \|\mathrm{Re}\varphi\|$. $\square$

Proof of Thm. 9.1.1 for $\mathbb{F} = \mathbb{R}$. Assume $\mathbb{F} = \mathbb{C}$. We need to show that any $\varphi \in \mathcal{U}^*$ can be extended to a bounded linear functional Φ on $\mathcal{V}$ with operator norm $\|\Phi\| = \|\varphi\|$. By the real Hahn-Banach theorem, $\mathrm{Re}\varphi : \mathcal{U} \to \mathbb{R}$ (which has operator norm $\|\varphi\|$ by Prop. 9.1.5) can be extended to $\Psi \in \mathfrak{L}(\mathcal{V},\mathbb{R})$ with operator norm $\|\varphi\|$. By Prop. 9.1.5, we have $\Psi = \mathrm{Re}\Phi$ for some $\Phi \in \mathcal{V}^*$, and $\|\Phi\| = \|\varphi\|$. Since $\mathrm{Re}\Phi$ extends $\mathrm{Re}\varphi$, Φ extends φ. $\square$

9.2 Indeterminate moment problems: the finite-approximation approach

Let $\mathcal{V},\mathcal{W}$ be normed vector spaces over $\mathbb{F} \in \{\mathbb{R},\mathbb{C}\}$. Let $1 < p,q < +\infty$ with $p^{-1} + q^{-1} = 1$.

Definition 9.2.1. The **canonical map**

$$\mathbf{J}_{\mathcal{V}} : \mathcal{V} \to \mathcal{V}^{**} \qquad v \mapsto v^\sharp$$

is defined by setting

$$\boldsymbol{v}^\sharp : \mathcal{V}^* \to \mathbb{F} \qquad \varphi \mapsto \varphi(v)$$

This linear map has operator norm $\leqslant 1$, that is $\|v^\sharp\| \leqslant \|v\|$.

Proof. To see $\|v^\sharp\| \leqslant \|v\|$, we choose any $\varphi \in \mathcal{V}^*$ satisfying $\|\varphi\| \leqslant 1$, and compute that $|\langle v^\sharp, \varphi\rangle| = |\langle v, \varphi\rangle| \leqslant \|v\|$. □

9.2.1 Introduction

In this section, we study the finite-approximation approach to the moment problem introduced in Subsec. 9.1.1. This method was applied by Riesz in [Rie10] to the case $\mathcal{V} = L^p([a,b], m)$, and later in [Rie11] to the case $\mathcal{V} = C([a,b])$. Although it applies in a less general setting than the method of linear extension discussed in Sec. 9.1, several key concepts—such as quotient normed spaces and separation of convex sets—arise naturally within this framework. For this reason, we consider it worthwhile to examine this approach in detail.

Recall the setting of Subsec. 9.1.1: $(\xi_n)_{n\in\mathbb{Z}_+}$ is a sequence in a function space, $(c_n)_{n\in\mathbb{Z}_+}$ is a sequence in $\mathbb{F}$ satisfying (9.3a), which is a necessary condition for solvability of the moment problem, and we seek a function f such that

$$\int \xi_k f = c_k \qquad \text{for each } k \in \mathbb{Z}_+$$

We have already encountered the finite-approximation method in our treatment of polynomial moment problems in Sec. 4.2. The pattern here is similar: one first seeks approximate solutions—analogous to Padé approximation—a sequence of functions (f_n) satisfying

$$\int \xi_k f_n = c_k \qquad \text{for each } k \leqslant n$$

If one can show that (f_n) is uniformly bounded, then by a Banach-Alaoglu type argument the sequence (f_n) admits a subsequence converging weak-* to a solution of the moment problem.

The translation of the above argument into modern language is as follows. Let $\mathcal{U}$ be a linear subspace of $\mathcal{V}$. To simplify the discussion, we assume that $\mathcal{V}$ is separable. Let $\varphi \in \mathcal{U}^*$. Our goal is to show that φ can be extended to a bounded linear functional on $\mathcal{V}$. To that end, choose a sequence $(\xi_n)_{n\in\mathbb{Z}_+}$ whose span is dense in $\mathcal{U}$, and for each $n \in \mathbb{Z}_+$, define

$$\mathcal{U}_n = \mathrm{Span}\{\xi_1, \ldots \xi_n\}$$

Suppose that every bounded linear functional on $\mathcal{U}_n$ can be extended to a bounded linear functional on $\mathcal{U}$ without increasing the operator norm (or increasing it only by an arbitrarily small positive amount). Then, for each n, there exists $\varphi_n \in \mathcal{V}^*$ such that

$$\varphi_n|_{\mathcal{U}_n} = \varphi|_{\mathcal{U}_n}$$

and $\|\varphi_n\| = \|\varphi|_{\mathcal{U}_n}\|$, which in particular does not exceed $\|\varphi\|$. By the Banach-Alaoglu Thm. 2.6.5 (and Prop. 2.6.6), (φ_n) has a weak-* convergent subsequence whose limit we denote by $\widetilde{\varphi} \in \mathcal{V}^*$. Then

$$\widetilde{\varphi}|_{\mathcal{U}_n} = \varphi|_{\mathcal{U}_n}$$

for each n, and hence $\widetilde{\varphi}|_{\mathcal{U}} = \varphi$ by the density of $\bigcup_n \mathcal{U}_n$ in $\mathcal{U}$. Thus $\widetilde{\varphi}$ is a desired extension of φ.

The reader will immediately observe that the finite-approximation approach distinguishes between two kinds of moment problems: those of finite type and those of infinite type. In this framework, moment problems of finite type are straightforward to solve (at least when $\mathcal{V}$ is a concrete function space such as $L^p(X, \mu)$ or $C(X)$), while moment problems of infinite type are obtained as weak-* limits of solutions to the finite type ones.

As we shall see, the principal difficulty in the finite-approximation method lies in establishing an upper bound for the sequence (φ_n) of approximate solutions. That is, the real difficulty, at least when $\mathcal{V}$ is a common function space, lies in proving that any bounded linear functional defined on the finite-dimensional subspace $\mathcal{U}_n$ can be extended to a functional on $\mathcal{V}$ *without increasing the operator norm except by an arbitrarily small positive amount*. This is precisely where convexity enters the picture.

In the following subsections, we discuss, in turn, the existence of finite approximate solutions and the uniform boundedness of such solutions, within the abstract framework of normed vector spaces. However, we introduce one modification: we view the given sequence (ξ_n) as a sequence (φ_n) in the dual space $\mathcal{V}^*$ rather than in $\mathcal{V}$, and we look for solutions of the moment problem in $\mathcal{V}^{**}$. The advantage of this setting is that $\mathcal{V}^*$ already possesses a rich supply of bounded linear functionals, namely, those induced by the elements of $\mathcal{V}$.

For example, if (ξ_n) lies in $L^p(X, \mu)$ where (X, μ) is a σ-finite measure space, we take $\mathcal{V} = L^q(X, \mu)$ so that $\mathcal{V}^* = L^p(X, \mu)$ by Thm. 1.6.16. If (ξ_n) lies in $C(X)$ for a compact Hausdorff space X, we take $\mathcal{V} = C(X)^*$, and hence view $\xi_n \in C(X)^{**}$.[2] This is possible, since $C(X)$ embeds isometrically into $C(X)^{**}$, as the following observation shows:

[2]In this setting, a solution to the moment problem is a priori a bounded linear functional on $\mathcal{V}^* = C(X)^{**}$. By restricting to $C(X)$, we obtain a linear functional on $C(X)$, which corresponds to a complex Borel measure μ on X solving the original moment problem $\int_X \xi_k d\mu = c_k$.

Example 9.2.2. The canonical map $C(X) \to C(X)^{**}$ is an isometry. Indeed, if $f \in C(X)$, then $\|f^\sharp\| \leqslant \|f\|$ by Def. 9.2.1. Choose $x \in X$ such that $|f(x)| = \|f\|$, and let $\varphi \in C(X)^*$ be defined by $\varphi(g) = g(x)$ for each $g \in C(X)$. Then $\|\varphi\| = 1$, and hence

$$\|f^\sharp\| \geqslant |\langle f^\sharp, \varphi\rangle| = |\langle f, \varphi\rangle| = |f(x)| = \|f\|$$

This proves $\|f^\sharp\| = f$.

In the remainder of this section, the proofs will not rely on the Hahn-Banach theorem established in Sec. 9.1; in particular, they will not depend on the axiom of choice.

9.2.2 Quotient normed vector spaces and their dual spaces

To prepare for the finite-approximation approach to the moment problem in the general setting, we require several basic facts concerning quotient normed vector spaces.

Proposition 9.2.3. *Let $\mathcal{U}$ be a closed linear subspace of $\mathcal{V}$. Then, the quotient space $\mathcal{V}/\mathcal{U}$ has a norm defined by*

$$\|\xi + \mathcal{U}\| = \mathrm{dist}(\xi, \mathcal{U}) \equiv \inf_{u \in \mathcal{U}} \|\xi + u\| \qquad \text{for each } \xi \in \mathcal{V} \tag{9.8}$$

The normed vector space $\mathcal{V}/\mathcal{U}$ is called the **quotient normed vector space** (or simply **quotient space**) of $\mathcal{V}$ by $\mathcal{U}$. We will see in Cor. 10.3.3 that $\mathcal{V}/\mathcal{U}$ is complete if $\mathcal{V}$ is complete.

Proof. For each $\xi, \eta \in \mathcal{V}$, we compute that

$$\|\xi + \eta + \mathcal{U}\| \leqslant \|\xi + \eta + u + v\| \leqslant \|\xi + u\| + \|\eta + v\|$$

where $u, v \in \mathcal{U}$. Taking $\inf$ over $u, v \in \mathcal{U}$, we obtain $\|\xi + \eta + \mathcal{U}\| \leqslant \|\xi + \mathcal{U}\| + \|\eta + \mathcal{U}\|$. Similarly, for each $a \in \mathbb{F}$,

$$\|a\xi + \mathcal{U}\| \leqslant \|a\xi + au\| = |a| \cdot \|\xi + u\|$$

where $u \in \mathcal{U}$. Taking $\inf$ over $u \in \mathcal{U}$, we obtain $\|a\xi\| \leqslant |a| \cdot \|\xi\|$. In view of Rem. 2.3.2, we conclude that (9.8) defines a seminorm on $\mathcal{V}/\mathcal{U}$.

If $\|\xi + \mathcal{U}\| = 0$, then the distance of ξ to $\mathcal{U}$ is zero. Since $\mathcal{U}$ is closed, we conclude that $\xi \in \mathcal{U}$. Hence (9.8) is a norm on $\mathcal{V}/\mathcal{U}$. $\square$

Theorem 9.2.4. *Let $\mathcal{U}$ be a closed linear subspace of $\mathcal{V}$. Then we have a linear isometry*

$$\mathfrak{L}(\mathcal{V}/\mathcal{U}, \mathcal{W}) \to \mathfrak{L}(\mathcal{V}, \mathcal{W}) \qquad T \mapsto T \circ \pi_{\mathcal{U}} \tag{9.9}$$

where $\pi_{\mathcal{U}} : \mathcal{V} \to \mathcal{V}/\mathcal{U}$ is the quotient map $\xi \mapsto \xi + \mathcal{U}$. Moreover, the range of (9.9) consists precisely of $S \in \mathfrak{L}(\mathcal{V}, \mathcal{W})$ satisfying $S|_{\mathcal{U}} = 0$.

In particular, we have a linear isometry

$$(\mathcal{V}/\mathcal{U})^* \to \mathcal{V}^* \qquad \psi \mapsto \psi \circ \pi_{\mathcal{U}} \tag{9.10}$$

Proof. Write $S = T \circ \pi_{\mathcal{U}}$. Then, for each $\xi \in \mathcal{V}$, we have

$$T(\xi + \mathcal{U}) = S\xi$$

Since $\|\pi_{\mathcal{U}}\| \leqslant 1$, we have $\|S\| \leqslant \|T\| \cdot \|\pi_{\mathcal{U}}\| \leqslant \|T\|$.

By the Riesz lemma (Thm. 8.6.5), for each $\gamma > 1$, there exists $\xi' \in \xi + \mathcal{U}$ such that $\|\xi'\| \leqslant \gamma \mathrm{dist}(\xi, \mathcal{U}) \equiv \gamma\|\xi + \mathcal{U}\|$, and hence

$$\|T(\xi + \mathcal{U})\| = \|T(\xi' + \mathcal{U})\| = \|S\xi'\| \leqslant \|S\| \cdot \|\xi'\| = \gamma\|S\| \cdot \|\xi + \mathcal{U}\|$$

and hence $\|T\| \leqslant \gamma\|S\|$. Since γ is arbitrary, we conclude $\|T\| = \|S\|$.

Each linear map in the range of (9.9) clearly vanishes on $\mathcal{U}$. Conversely, assume that $S \in \mathfrak{L}(\mathcal{V}, \mathcal{W})$ vanishes on $\mathcal{U}$. Then S descends to a linear map $T : \mathcal{V} \to \mathcal{W}$. The computation in the previous paragraph shows $\|T\| \leqslant \gamma\|S\|$; in particular, $T \in \mathfrak{L}(\mathcal{V}/\mathcal{U}, \mathcal{W})$. This proves that S belongs to the range of (9.9). $\square$

Definition 9.2.5. For each $E \subset \mathcal{V}^*$ and $F \subset \mathcal{V}$, define the **pre-annihilator** $^{\perp}\boldsymbol{E}$ and the **annihilator** $\boldsymbol{F}^{\top}$ by[3]

$$\begin{aligned} ^{\perp}E &= \{\xi \in \mathcal{V} : \langle \xi, \varphi \rangle = 0 \text{ for each } \varphi \in E\} \\ F^{\top} &= \{\varphi \in \mathcal{V}^* : \langle \xi, \varphi \rangle = 0 \text{ for each } \xi \in F\} \end{aligned}$$

Then $^{\perp}E$ and $F^{\top}$ are linear subspaces of $\mathcal{V}$ and $\mathcal{V}^*$, respectively.

Corollary 9.2.6. *Let $\mathcal{U}$ be a closed linear subspace of $\mathcal{V}$. Then we have an isomorphism of normed vector spaces*

$$(\mathcal{V}/\mathcal{U})^* \xrightarrow{\simeq} \mathcal{U}^{\top} \qquad \psi \mapsto \psi \circ \pi_{\mathcal{U}}$$

where $\pi_{\mathcal{U}} : \mathcal{V} \to \mathcal{V}/\mathcal{U}$ is the quotient map $\xi \mapsto \xi + \mathcal{U}$.

Proof. By Thm. 9.2.4, the linear isometry $(\mathcal{V}/\mathcal{U})^* \to \mathcal{V}^*$ defined by $\psi \mapsto \psi \circ \pi_{\mathcal{U}}$ has range $\mathcal{U}^{\perp}$. $\square$

Corollary 9.2.7. *Let $\mathcal{X}$ be a linear subspace of $\mathcal{V}^*$. Let $\pi_{^{\perp}\mathcal{X}} : \mathcal{V} \to \mathcal{V}/{}^{\perp}\mathcal{X}$ be defined by $\xi \mapsto \xi + {}^{\perp}\mathcal{X}$. There we have an isomorphism of normed vector spaces*

$$(\mathcal{V}/{}^{\perp}\mathcal{X})^* \xrightarrow{\simeq} ({}^{\perp}\mathcal{X})^{\top} \qquad \psi \mapsto \psi \circ \pi_{^{\perp}\mathcal{X}} \tag{9.11}$$

where $({}^{\perp}\mathcal{X})^{\top}$ equals $\mathcal{X}$ when $\dim \mathcal{X} < +\infty$.

[3] Since $F^{\perp}$ is already reserved for the orthogonal complement in inner product spaces, we use the notation $F^{\top}$ instead.

Proof. By Cor. 9.2.6, (9.11) is an isomorphism of normed vector spaces. Clearly $\mathcal{X} \subset ({}^\perp\mathcal{X})^\top$. Let us assume $\dim \mathcal{X} < +\infty$ and prove $({}^\perp\mathcal{X})^\top \subset \mathcal{X}$.

Let $\varphi_1, \dots, \varphi_n$ be a basis of $\mathcal{X}$. Define a linear map

$$\Phi = (\varphi_1, \dots, \varphi_n) : \mathcal{V} \to \mathbb{F}^n$$

Then $\mathrm{Ker}(\Phi) = {}^\perp\mathcal{X}$. Moreover, Φ is surjective; otherwise, there exists a non-zero $(a_1, \dots, a_n) \in \mathbb{F}^n$ orthogonal to $\mathrm{Rng}(\Phi)$, which implies $a_1\varphi_1 + \cdots + a_n\varphi_n = 0$, contrary to the linear independence of $\varphi_1, \dots, \varphi_n$.

Now, let us prove that any $\psi \in ({}^\perp\mathcal{X})^\top$ belongs to $\mathcal{X}$. Note that $\psi : \mathcal{V} \to \mathbb{F}$ is a linear functional vanishing on ${}^\perp\mathcal{X} = \mathrm{Ker}(\Phi)$. Therefore, by the surjectivity of Φ, there is a unique linear map $\alpha : \mathbb{F}^n \to \mathbb{F}$ satisfying $\alpha \circ \Phi = \psi$. Set $a_j = \alpha(e_j)$ where e_j is the j-th standard coordinate of $\mathbb{F}^n$. Then $\psi = a_1\varphi_1 + \cdots + a_n\varphi_n$. This proves $\psi \in \mathcal{X}$, and hence $({}^\perp\mathcal{X})^\top \subset \mathcal{X}$. $\square$

9.2.3 Existence of finite approximate solutions to moment problems

Remark 9.2.8. If $\dim \mathcal{V} < +\infty$, then by Thm. 8.4.14, any linear isomorphism $\mathcal{V} \to \mathbb{F}^n$ is a homeomorphism. Therefore, every linear functional on $\mathcal{V}$ is bounded, i.e., $\mathrm{Lin}(\mathcal{V}, \mathbb{F}) = \mathcal{V}^*$.

The following theorem, due to Helly [Hel21], establishes the existence of finite approximate solutions to the moment problem.

Theorem 9.2.9. *Let $\mathcal{X}$ be a finite-dimensional linear subspace of $\mathcal{V}^*$. Then we have a linear bijection*

$$\mathcal{V}/{}^\perp\mathcal{X} \longrightarrow \mathcal{X}^* \qquad \xi + {}^\perp\mathcal{X} \mapsto \xi^\sharp|_\mathcal{X} \tag{9.12}$$

where $\xi^\sharp|_\mathcal{X} : \mathcal{X} \to \mathbb{F}$ is the restriction of $\xi^\sharp : \mathcal{V}^* \to \mathbb{F}$ to $\mathcal{X}$.

Proof. The linear map $\mathcal{V} \to \mathcal{X}^*$ defined by $\xi \mapsto \xi^\sharp|_\mathcal{X}$ has kernel ${}^\perp\mathcal{X}$, and hence descends to (9.12), which is an injective linear map. In particular, by Rem. 9.2.8,

$$\dim \mathcal{V}/{}^\perp\mathcal{X} \leqslant \dim \mathcal{X}^* = \dim \mathcal{X} < +\infty$$

By Cor. 9.2.7 (and also Rem. 9.2.8), $\dim \mathcal{X} = \dim \mathcal{V}/{}^\perp\mathcal{X}$. Therefore, the linear map (9.12) is bijective. $\square$

9.2.4 Uniformly bounded approximate solutions

Let us explain why the surjectivity of the map (9.12) ensures the existence of finite approximate solutions to the moment problem. Fix a sequence $(c_n)_{n\in\mathbb{Z}_+}$ in $\mathbb{F}$ and a sequence $(\varphi_n)_{n\in\mathbb{Z}_+}$ in $\mathcal{V}^*$ satisfying an analogue of (9.3a); that is, there exists $M \in \mathbb{R}_{\geqslant 0}$ such that

$$\Big|\sum_n a_n c_n\Big| \leqslant M \Big\|\sum_n a_n \varphi_n\Big\| \qquad \text{for each } (a_n) \in C_c(\mathbb{Z}_+, \mathbb{F})$$

Let $\mathcal{X}_n = \mathrm{Span}\{\varphi_1, \ldots, \varphi_n\}$. Then we have a linear functional

$$\Xi_n : \mathcal{X}_n \to \mathbb{F} \qquad a_1\varphi_1 + \cdots + a_n\varphi_n \mapsto a_1c_1 + \cdots + a_nc_n$$

with operator norm $\leqslant M$. By Thm. 9.2.9, there exists $\xi_n \in \mathcal{V}$ such that

$$\xi_n^\sharp|_{\mathcal{X}_n} = \Xi_n \tag{9.13}$$

Thus, (ξ_n) may be regarded as a sequence of approximate solutions to the moment problem.

However, we require the sequence (ξ_n) to be bounded so that $(\xi_n^\sharp)$ is bounded in $\mathcal{V}^{**}$. Therefore, by the Banach-Alaoglu Thm. 2.6.5, $(\xi_n^\sharp)$ admits a weak-* convergent subsequence/subnet with limit $\Xi \in \mathcal{V}^{**}$, satisfying

$$\langle \Xi, \varphi_n \rangle = a_n \qquad \text{for each } n$$

thus providing a solution of the moment problem.

The existence of a bounded sequence (ξ_n) follows from the fact that the linear isomorphism (9.12) is in fact an isometry (cf. Lem. 9.2.10). Indeed, this implies that for any ξ_n satisfying (9.13), we have $\|\xi_n + {}^\perp\mathcal{X}\| = \|\Xi_n\| \leqslant M$, and thus for some such ξ_n satisfying (9.13), we have

$$\|\xi_n\| \leqslant M + 1$$

Lemma 9.2.10. *In Thm. 9.2.9, the linear bijection $\mathcal{V}/{}^\perp\mathcal{X} \to \mathcal{X}^*$ (defined by (9.12)) is an isometry.*

Proof. Recall from Cor. 9.2.7 the isomorphism of normed vector spaces

$$\Phi : (\mathcal{V}/{}^\perp\mathcal{X})^* \xrightarrow{\simeq} \mathcal{X} \qquad \psi \mapsto \psi \circ \pi_{{}^\perp\mathcal{X}}$$

Therefore, its transpose is also an isomorphism of normed vector spaces

$$\Phi^\sharp : \mathcal{X}^* \xrightarrow{\simeq} (\mathcal{V}/{}^\perp\mathcal{X})^{**} \qquad \tau \mapsto \tau \circ \Phi$$

For each $\xi \in \mathcal{X}$, the restriction $\xi^\sharp|_{\mathcal{X}} : \mathcal{X} \to \mathbb{F}$ sends each $\varphi \in \mathcal{X}$ to $\langle \varphi, \xi \rangle$. Thus

$$\xi^\sharp|_{\mathcal{X}} \circ \Phi : (\mathcal{V}/{}^\perp\mathcal{X})^* \to \mathbb{F} \qquad \psi \mapsto \langle \psi \circ \pi_{{}^\perp\mathcal{X}}, \xi \rangle$$

Since $\langle \psi \circ \pi_{{}^\perp\mathcal{X}}, \xi \rangle = \psi \circ \pi_{{}^\perp\mathcal{X}}(\xi) = \psi(\xi + {}^\perp\mathcal{X}) = \langle \psi, (\xi + {}^\perp\mathcal{X})^\sharp \rangle$, we conclude that

$$\xi^\sharp|_{\mathcal{X}} \circ \Phi = (\xi + {}^\perp\mathcal{X})^\sharp$$

Consequently, we obtain the commutative diagram

$$\begin{tikzcd}
& \mathcal{V}/{}^\perp\mathcal{X} \arrow[dl, "(9.12)"'] \arrow[dr] & \\
\mathcal{X}^* \arrow[rr, "\Phi^\sharp", "\simeq"'] & & (\mathcal{V}/{}^\perp\mathcal{X})^{**}
\end{tikzcd}
\qquad
\begin{tikzcd}
& \xi + {}^\perp\mathcal{X} \arrow[dl, mapsto] \arrow[dr, mapsto] & \\
\xi^\sharp|_{\mathcal{X}} \arrow[rr, mapsto] & & (\xi + {}^\perp\mathcal{X})^\sharp
\end{tikzcd}$$

Therefore, to prove that (9.12) is an isometry, it suffices to show that the canonical map $\mathcal{V}/{}^\perp\mathcal{X} \to (\mathcal{V}/{}^\perp\mathcal{X})^{**}$ is an isomtry, cf. Thm 9.2.11. $\square$

Theorem 9.2.11. *Assume that* $\dim \mathcal{V} < +\infty$. *Then the canonical map* $\mathbf{J}_{\mathcal{V}} : \mathcal{V} \to \mathcal{V}^{**}$ *is an isomorphism of normed vector spaces.*

We defer the proof of Thm. 9.2.11 to Subsec. 9.3.3. Observe that the only nontrivial assertion in Thm. 9.2.11 is that the canonical map $\mathcal{V} \to \mathcal{V}^{**}$ is an isometry. Once this is known, the fact that $\dim \mathcal{V} < +\infty$, together with the injectivity of linear isometries, implies that the map $\mathcal{V} \to \mathcal{V}^{**}$ must be bijective.

As we shall see in Sec. 9.4, Thm. 9.2.11 follows easily from the Hahn-Banach Thm. 9.1.1. Indeed, the Hahn-Banach theorem proves a stronger statement: regardless of whether $\mathcal{V}$ is finite-dimensional or not, the canonical map $\mathcal{V} \to \mathcal{V}^{**}$ is always an isometry. However, the linear-extension proof of Hahn-Banach obscures the geometric relationship between Thm. 9.2.11 and the separation of convex sets, an aspect we will make explicit in Sec. 9.3.

A final remark: In Riesz's treatment of moment problems for $L^p([a,b],m)$ and $C([a,b])$, Lem. 9.2.10 was proved using concrete structural properties of these function spaces. For instance, when $\mathcal{V} = L^p([a,b],m)$, the space $\mathcal{X}$ is a finite-dimensional linear subspace of $\mathcal{V}^* = L^q([a,b],m)$. In that context, for each $\Lambda \in \mathcal{X}^*$, one can explicitly construct $f \in \mathcal{V}$ with $f^\sharp|_{\mathcal{X}} = \Lambda$, satisfying

$$\|f\| = \|f + {}^\perp\mathcal{X}\| = \|\Lambda\|$$

as follows (cf. [Rie10]). An element $g \in L^q([a,b],m)$ in the unit sphere of $\mathcal{X}$ can be chosen such that $\Lambda(g) = \|\Lambda\|$. Then

$$f = \|\Lambda\| \cdot |g|^{q-1}\overline{\mathrm{sgn} g}$$

where $\mathrm{sgn} g$ denotes $g/|g|$ wherever $g \neq 0$, and 0 otherwise.

9.3 Separation of convex sets: the finite-dimensional case

Let $\mathcal{V}$ be a finite-dimensional normed $\mathbb{F}$-vector space. The goal of this section is to prove Thm. 9.2.11: the canonical linear map

$$\mathcal{V} \to \mathcal{V}^{**} \qquad v \mapsto v^\sharp$$

is an isometry. This is the most crucial step in the finite-approximation approach to moment problems of bounded type. With Prop. 9.1.5, the complex case reduces easily to the real one. Therefore, unless otherwise stated, we shall assume that $\mathcal{V}$ is a normed $\mathbb{R}$-vector space.

9.3.1 Reduction to separation of convex sets

Choose $v \in \mathcal{V}$ with $\|v\| = 1$. Our aim is to show that $\|v^\sharp\| = 1$. By Def. 9.2.1, we already know that $\|v^\sharp\| \leqslant 1$, so it remains to prove $\|v^\sharp\| \geqslant 1$, i.e.,

$$\sup_{\varphi \in \overline{B}_{\mathcal{V}^*}(0,1)} |\varphi(v)| \geqslant 1$$

This would follow immediately if we can find $\varphi \in \mathrm{Lin}(\mathcal{V}, \mathbb{R})$ such that

$$\|\varphi\| \leqslant 1 \qquad \varphi(v) = 1$$

The inequality $\|\varphi\| \leqslant 1$ may be reformulated in a more geometrically transparent way:

Lemma 9.3.1. *Let $\mathcal{W}$ be a normed $\mathbb{R}$-vector space. Let $\varphi : \mathcal{W} \to \mathbb{R}$ be linear. Let $a, b \in \mathbb{R}_{>0}$. Then* [4]

$$\|\varphi\| \leqslant b/a \qquad \Longleftrightarrow \qquad B_{\mathcal{W}}(0, a) \subset \varphi^{-1}(-\infty, b]$$

Proof. The direction $\Rightarrow$ is obvious. Assume $B_{\mathcal{W}}(0, a) \subset \varphi^{-1}(-\infty, b]$. Then for each $v \in \overline{B}_{\mathcal{W}}(0, 1)$ and $0 < a' < a$, we have $\varphi(a'v) \leqslant b$, and hence $\varphi(v) \leqslant b/a'$. Since a' is arbitrary, we conclude $\varphi(v) \leqslant b/a$. Similarly $\varphi(-v) \leqslant b/a$, and hence $-b/a \leqslant \varphi(v) \leqslant b/a$. This proves $\|\varphi\| \leqslant b/a$. □

Thus our task reduces to finding $\varphi \in \mathrm{Lin}(\mathcal{V}, \mathbb{R})$ such that $v \in \varphi^{-1}(1)$ and $B_{\mathcal{V}}(0, 1) \subset \varphi^{-1}(-\infty, 1]$. Observe that there is a one-to-one correspondence between elements of $\varphi \in \mathcal{V}^*$ and hyperplanes $H \subset \mathcal{V}$ (that is, translates of linear subspaces of codimension 1) given by

$$H = \varphi^{-1}(1)$$

Therefore, it suffices to find a hyperplane H passing through v such that $B_{\mathcal{V}}(0, 1)$ lies entirely on one side of H, namely in the region $\{\varphi \leqslant 1\}$. It cannot lie completely in the region $\{\varphi \geqslant 1\}$, since 0 does not belong to that side.

In this way, the proof that the canonical map $\mathbf{J}_{\mathcal{V}} : \mathcal{V} \to \mathcal{V}^{**}$ is an isometry is reduced to establishing the geometric fact that $B_{\mathcal{V}}(0, 1)$ and any point $v \in \mathcal{V}$ with $\|v\| = 1$ can be separated by a hyperplane. Since $B_{\mathcal{V}}(0, 1)$ is convex, we will establish a stronger statement: any convex open subset $\Omega \subset \mathcal{V}$ and any point $v \in \mathcal{V} \backslash \Omega$ can be separated by a hyperplane.

9.3.2 Proof of separation of convex sets

Definition 9.3.2. Let W be an $\mathbb{F}$-vector space (viewed as an $\mathbb{R}$-vector space if $\mathbb{F} = \mathbb{C}$). For each $p, q \in W$, define the **intervals**

$$[p, q] \qquad (p, q) \qquad (p, q] \qquad [p, q)$$

to be the images of the sets $[0, 1]$, $(0, 1)$, $(0, 1]$, and $[0, 1)$ under the map

$$t \mapsto (1 - t)p + tq$$

[4] One may replace the open ball $B_{\mathcal{W}}(0, a)$ with the closed ball $\overline{B}_{\mathcal{W}}(0, a)$. We use the open ball here because, in the infinite-dimensional setting, the corresponding geometric statement is most easily established for open sets; see the proof of Thm. 9.4.1.

respectively. We also set

$$\boldsymbol{[p,\infty_q)} = \{(1-t)p + tq : t \in \mathbb{R}_{\geqslant 0}\}$$

We say that a subset $\Omega \subset W$ is **convex** if $[p,q] \subset \Omega$ whenever $p,q \in \Omega$.

Lemma 9.3.3. *Let $\mathcal{W}$ be a normed $\mathbb{F}$-vector space. Let Ω be an open convex subset of $\mathcal{W}$. Fix $p \in \Omega$. Suppose that $v \in \mathcal{W}$ is not a boundary point of $[p,\infty_v) \cap \Omega$, that is,*

$$[p,\infty_v] \cap \Omega = [p, p + r(v-p)) \qquad \textit{for some } 0 < r < 1$$

Then $v \notin \overline{\Omega}$.

Proof. By a translation we assume WLOG that $p = 0$. Choose $\varepsilon > 0$ such that $r(1+\varepsilon) < 1$. Then

$$v \notin [0, r(1+\varepsilon)v) = (1+\varepsilon)([0,\infty_v] \cap \Omega)$$

and hence $v \notin (1+\varepsilon)\Omega$. We claim that

$$v - \varepsilon\Omega = \{v - \varepsilon y : y \in \Omega\}$$

which is a neighborhood of v, is disjoint from Ω, proving that $v \notin \overline{\Omega}$.

Suppose that $v - \varepsilon\Omega$ and Ω has an intersection point x. Then $v = x + \varepsilon y$ with $x, y \in \Omega$. Since Ω is convex, we conclude that $(1+\varepsilon)^{-1}v \in \Omega$, and hence $v \in (1+\varepsilon)\Omega$. This gives a contradiction. □

Lemma 9.3.4. *Let $\mathcal{W}$ be a normed $\mathbb{R}$-vector space, and let $\varphi : \mathcal{W} \to \mathbb{R}$ be a non-zero bounded linear map. Then φ is an open map.*

See Cor. 10.3.4 for a higher dimensional generalization of this lemma.

Proof. Choose $e \in \mathcal{W}$ with $\varphi(e) = 1$. Then $\varphi(-e,e) = (-1,1)$. Set $r = \|e\|$. Then

$$\varphi(B_{\mathcal{W}}(0,r)) \supset (-1,1)$$

Therefore, the image of any non-empty open ball $B_{\mathcal{W}}(v,a)$ has $\varphi(v)$ as its interior point. It follows easily that φ is open. □

We are now ready to prove the main result of this section.

Theorem 9.3.5. *Let $\mathcal{V}$ be a finite-dimensional normed $\mathbb{R}$-vector space. Assume that Ω is an open convex subset of $\mathcal{V}$ containing 0. Let $v \in \mathcal{V}\backslash\Omega$. Then there exists $\varphi \in \mathcal{V}^*$ such that*

$$\Omega \subset \varphi^{-1}(-\infty,1) \qquad v \in \varphi^{-1}\{1\}$$

Proof. Step 1. By Lem. 9.3.4, it suffices to prove the weaker condition that $\Omega \subset \varphi^{-1}(-\infty, 1]$. By Thm. 8.4.14, we may assume that $\mathcal{V}$ is the Euclidean space $\mathbb{R}^n$. Choose $0 < r \leqslant 1$ such that

$$[0, \infty_v) \cap \Omega = [0, rv)$$

In this step, we consider the case $r < 1$. Then $v \notin \overline{\Omega}$ by Lem. 9.3.3. By Bolzano-Weierstrass, there exists $p \in \overline{\Omega}$ such that

$$\|v - p\| = \mathrm{dist}(v, \overline{\Omega})$$

Let H be the hyperplane passing through v and perpendicular to $[v, p]$.[5] Let $\varphi \in \mathcal{V}^*$ be the unique linear functional such that $\varphi^{-1}(1) = H$.

Let us prove $\varphi(\overline{\Omega}) \subset (-\infty, 1]$, noting that $\overline{\Omega}$ is convex. It is easy to prove the following fact:[6]

- If $x, y \in \mathbb{R}^n$ are lying on the two sides of H (that is, one is in $\varphi^{-1}(-\infty, 1)$, and the other one is in $\varphi^{-1}(1, +\infty)$), $u \in H$, and $[u, x]$ is perpendicular to H, then there exists $z \in [x, y]$ such that $\|u - z\| < \|u - x\|$.

Therefore $\varphi(p) \leqslant 1$; otherwise, setting $u = v, x = p, y = 0$, we obtain a point of $[0, p] \subset \overline{\Omega}$ whose distance to v is smaller than $\|v - p\|$, impossible. Moreover, $\varphi(p) \neq 1$; otherwise, $p \in H$, and hence $[v, p]$ is simultaneously lying inside H and perpendicular to H, contrary to $v - p \neq 0$ (because $v \notin \overline{\Omega}$ and $p \in \overline{\Omega}$). We conclude $\varphi(p) < 1$.

Consequently, for each $q \in \overline{\Omega}$ we have $\varphi(q) \leqslant 1$; otherwise, setting $u = v, x = p, y = q$, we obtain a point of $[p, q] \subset \overline{\Omega}$ whose distance to v is smaller than $\|v - p\|$, impossible. This finishes the proof that $\varphi(\overline{\Omega}) \subset (-\infty, 1]$.

Step 2. We treat the case $r = 1$, that is, $[0, \infty_v] \cap \Omega = [0, v)$. Let (v_k) be a sequence in $[0, \infty_v) \backslash [0, v]$ converging to v. By Step 2, there exists a sequence (φ_k) in $\mathcal{V}^*$ satisfying

$$\varphi_k(\Omega) \subset (-\infty, 1] \qquad \varphi_k(v_k) = 1$$

By Lem. 9.3.1, $\sup_k \|\varphi_k\| \leqslant 1$. Therefore, by Bolzano-Weierstrass, (φ_k) admits a subsequence converging to some $\varphi \in \mathcal{V}^*$, which clearly satisfies $\Omega \subset \varphi^{-1}(-\infty, 1]$ and $\varphi(v) = 1$. $\square$

[5]By saying H is perpendicular to an interval $[\alpha, \beta]$, we mean that $H - q \perp \beta - \alpha$ for some (and hence for all) $q \in H$.

[6]By an orthogonal transformation, we may assume that $u = 0$, $H = \{(0, a_2, \dots, a_n) : a_i \in \mathbb{R}\}$, $x = (a, 0, 0, \dots)$, and $y = (b, ?, \dots, ?)$ where $a > 0$ and $b < 0$. Then the cosine of the angle from $[x, u]$ to $[x, y]$ is positive, and hence $\|u - z\| < \|u - x\|$ for any $z \in (x, y)$ close to x.

9.3.3 Proof of Thm. 9.2.11

Let us prove Thm. 9.2.11, that is, if $\mathcal{V}$ is a finite-dimensional normed $\mathbb{F}$-vector space, then the canonical linear map $\mathcal{V} \to \mathcal{V}^{**}$ is an isomorphism of normed vector spaces.

Proof of Thm. 9.2.11. It suffices to prove that $\mathcal{V} \to \mathcal{V}^{**}$ is an isometry, since this will imply the injectivity, and hence the surjectivity because $\dim \mathcal{V} = \dim \mathcal{V}^{**} < +\infty$ (cf. Rem. 9.2.8). Choose any $v \in \mathcal{V}$. By the discussion at the beginning of Subsec. 9.3.1, it suffices to prove that for each $v \in \mathcal{V}$ with $\|v\| = 1$, there exists $\varphi \in \mathcal{V}^*$ such that $\|\varphi\| \leqslant 1$ and $\varphi(v) = 1$.

In the case $\mathbb{F} = \mathbb{R}$, by Thm. 9.3.5, there exists $\varphi \in \mathcal{V}^*$ such that $\varphi(B_{\mathcal{V}}(0,1)) \subset (-\infty, 1]$ (and hence $\|\varphi\| \leqslant 1$ by Lem. 9.3.1) and $\varphi(v) = 1$, finishing the proof.

In the case $\mathbb{F} = \mathbb{C}$, by viewing $\mathcal{V}$ as a normed $\mathbb{R}$-vector space, we obtain $\psi \in \mathfrak{L}(\mathcal{V}, \mathbb{R})$ such that $\|\psi\| \leqslant 1$ and $\psi(v) = 1$. By Prop. 9.3, there exists $\varphi \in \mathcal{V}^*$ such that $\mathrm{Re}\varphi = \psi$, $\|\varphi\| \leqslant 1$, and $\psi(v) = 1$. □

9.4 Separation of convex sets: the general case

Let $\mathcal{V}$ be a normed vector space over $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. In this section, we extend Thm. 9.3.5 to arbitrary normed vector spaces; see Thm. 9.4.1. In Sec. 9.5, we apply this extension to investigate the relationship between norm closure and weak closure.

9.4.1 Separation of an open convex set and a point

Theorem 9.4.1. *Assume that Ω is an open convex subset of $\mathcal{V}$ containing 0. Let $v \in \mathcal{V}\backslash\Omega$. Then there exists $\varphi \in \mathcal{V}^*$ such that*

$$\Omega \subset (\mathrm{Re}\varphi)^{-1}(-\infty, 1) \qquad v \in (\mathrm{Re}\varphi)^{-1}\{1\}$$

We present two proofs of Thm. 9.4.1. The first employs the method of finite approximation, reducing the general case to the finite-dimensional one (Thm. 9.3.5) via the Banach-Alaoglu theorem and the Hahn-Banach Thm. 9.1.1. The second proof, which is more standard in textbooks, gives a direct argument (without reduction to finite dimensions) by adapting the proof of the Hahn-Banach Thm. 9.1.1, replacing the norm of $\mathcal{V}$ with the Minkowski functional.

Although the second proof is more elegant, its geometric picture is less transparent (especially in finite dimensions); indeed, it is somewhat unnatural to interpret the separation of convex sets in $\mathbb{R}^n$ in terms of extending linear functionals. For this reason, we include both proofs.

First proof of Thm. 9.4.1. By Prop. 9.1.5, it suffices to assume $\mathbb{F} = \mathbb{R}$. Moreover, by Lem. 9.3.4, it suffices to prove

$$\Omega \subset (\mathrm{Re}\varphi)^{-1}(-\infty, 1] \qquad v \in (\mathrm{Re}\varphi)^{-1}\{1\} \tag{9.14}$$

Let $\mathscr{I}$ be the directed set of finite-dimensional linear subspaces of $\mathcal{V}$ ordered by the relation $\subset$. By Thm. 9.3.5, for each $U \in \mathscr{I}$ there exists $\varphi_U \in U^*$ satisfying

$$\Omega \cap U \subset \varphi_U^{-1}(-\infty, 1] \qquad v \in \varphi_U^{-1}\{1\}$$

Choose $a > 0$ such that $B_{\mathcal{V}}(0, a) \subset \Omega$. Then Lem. 9.3.1 implies $\|\varphi_U\| \leqslant 1/a$. By the Hahn-Banach Thm. 9.1.1, φ_U can be extended to a linear functional $\mathcal{V} \to \mathbb{R}$, still denoted by φ_U, such that $\|\varphi_U\| \leqslant 1/a$. In particular, $\sup_{U \in \mathscr{I}} \|\varphi_U\| < +\infty$. By the Banach-Alaoglu Thm. 2.6.5, the net $(\varphi_U)_{U \in \mathscr{I}}$ in $\mathcal{V}^*$ admits a subnet converging weak-* to some $\varphi \in \mathcal{V}$. Then φ satisfies the required property. □

Remark 9.4.2. When $\mathcal{V}$ is separable, the above proof can be adapted so that it does not rely on the axiom of choice. Assume again that $\mathbb{F} = \mathbb{R}$. Choose a densely-spanning sequence (v_n) in $\mathcal{V}$. Let $U_n = \mathrm{Span}\{v_1, \dots, v_n\}$. As in the above proof, we have a sequence (φ_n) in U_n^* with $\|\varphi_n\| \leqslant 1/a < +\infty$ such that

$$\Omega \cap U_n \subset \varphi_n^{-1}(-\infty, 1] \qquad v \in \varphi_n^{-1}\{1\}$$

By Hahn-Banach (which in this separable case does not rely on the axiom of choice, cf. Rem. 9.1.3), each φ_n can be extended to a linear functional on $\mathcal{V}$ without increasing the operator norm. By the second proof of the Banach-Alaoglu Thm. 2.6.5, (φ_n) admits a subsequence (φ_{n_k}) converging weak-* to some $\varphi \in \mathcal{V}^*$.

Clearly we have $\varphi(v) = 1$. For each $\xi \in \Omega$, choose a sequence (ξ_n) in $U_\infty = \bigcup_n U_n$ converging to ξ. Since Ω is open, by passing to a subsequence we assume $\xi_n \in \Omega$ for each n, and hence $\varphi(\xi_n) \leqslant 1$. Therefore $\varphi(\xi) \leqslant 1$. This proves (9.14). □

Second proof of Thm. 9.4.1. By Prop. 9.1.5, it suffices to assume $\mathbb{F} = \mathbb{R}$. Define the **Minkowski functional**

$$p : \mathcal{V} \to \mathbb{R}_{\geqslant 0} \qquad \xi \mapsto \inf\{r \in \mathbb{R}_{\geqslant 0} : \xi \in r\Omega\}$$

Note that $p(\xi) < +\infty$ because by choosing $a > 0$ such that $B_{\mathcal{V}}(0, a) \subset \Omega$, we have

$$\xi \in (\|\xi\| + 1)B_{\mathcal{V}}(0, 1) \subset a^{-1}(\|\xi\| + 1)\Omega$$

We claim that p satisfies that for each $\xi, \eta \in \mathcal{V}$ and $\lambda > 0$,

$$p(\lambda\xi) = \lambda p(\xi) \tag{9.15a}$$
$$p(\xi + \eta) \leqslant p(\xi) + p(\eta) \tag{9.15b}$$

The first relation is easy to verify. To verify the second one, choose any $r > p(\xi)$ and $d > p(\eta)$. Then $\xi \in r\Omega$ and $\eta \in d\eta$, and hence

$$\xi + \eta \in r\Omega + d\Omega = (r + d) \cdot \Big(\frac{r}{r + d}\Omega + \frac{d}{r + d}\Omega\Big) \subset (r + d)\Omega$$

where the last inclusion is due to the convexity of Ω. Thus $p(\xi + \eta) \leqslant r + d$. Since r, d are arbitrary, we obtain (9.15b).

Since (9.15) is true, and since $p(v) \geqslant 1$ (because $v \notin \Omega$), by Rem. 9.1.4, the linear functional $\varphi : \mathbb{R}v \to \mathbb{R}$ with $\varphi(v) = 1$ can be extended to a linear functional $\varphi : \mathcal{V} \to \mathbb{R}$ satisfying

$$\varphi(\xi) \leqslant p(\xi) \qquad \text{for all } \xi \in \mathcal{V}$$

In particular, for each $\xi \in \Omega$, we have $p(\xi) \leqslant 1$ and hence $\varphi(\xi) \leqslant 1$. This proves (9.14), finishing the proof. □

9.4.2 Separation of two convex sets

To apply Thm. 9.4.1 to the study of weak topologies in Subsec. 9.5, we need the following adaption.

Theorem 9.4.3. *Let A, B be disjoint convex subsets of $\mathcal{V}$, one of which is closed and the other one is compact. Then there exist $\varphi \in \mathcal{V}^*$ and $-\infty < a < b < +\infty$ such that*

$$(\mathrm{Re}\varphi)^{-1}(A) \subset (-\infty, a] \qquad (\mathrm{Re}\varphi)^{-1}(B) \subset [b, +\infty) \tag{9.16}$$

In Subsec. 9.5, we will use Thm. 9.4.3 only in the special case where one of the sets A, B consists of a single point.

Proof. Assume WLOG that A is closed and B is compact, and that (by Prop. 9.1.5) $\mathbb{F} = \mathbb{R}$. Moreover, we may assume that A, B are nonempty; otherwise, take $\varphi = 0$.

We first note that $\mathrm{dist}(A, B) > 0$. Indeed, since A is closed, $\mathrm{dist}(A, \xi) > 0$ for each $\xi \in \mathcal{V} \backslash A$; since B is compact, the continuous function

$$B \to \mathbb{R}_{>0} \qquad \xi \mapsto \mathrm{dist}(A, \xi)$$

has a (strictly positive) minimum.

Fix $0 < \varepsilon < \mathrm{dist}(A, B)$. Observe that if $C_1, C_2 \subset \mathcal{V}$ are convex, then $C_1 + C_2$ and $C_1 - C_2$ are convex. Therefore

$$C := B + B_{\mathcal{V}}(0, \varepsilon) = \{\xi + \eta : \xi \in B, \eta \in \mathcal{V}, \|\eta\| < \varepsilon\}$$

is a convex open set. Similarly, $A - C$ is a convex open set. Since $A \cap C = \emptyset$, we have $0 \notin A - C$. Therefore, by Thm. 9.4.1, there exists a nonzero $\varphi \in \mathcal{V}^*$ such that [7]

$$\varphi(A - C) < 0$$

[7] Choose $-v \in A - C$. Then $0 \in v + A - C$ and $v \notin v + A - C$. By Thm. 9.4.1, there exists $\varphi \in \mathcal{V}^*$ such that $\varphi(v + A - C) < 1$ and $\varphi(v) = 1$. Hence $\varphi \neq 0$ and $\varphi(A - C) < 0$.

That is, $\varphi(x) < \varphi(y)$ for each $x \in A, y \in C$. Thus, any $\varphi(y)$ is an upper bound of $\varphi(A)$. Therefore

$$a := \sup \varphi(A)$$

is a finite real number, and we have $\varphi(C) \subset [a, +\infty)$, and hence $\varphi(C) \subset (a, +\infty)$ by Lem. 9.3.4 and the fact that $\varphi \neq 0$. Since B is compact, the number

$$b := \inf \varphi(B)$$

lies in $\varphi(B)$, which implies $a < b$. $\square$

Remark 9.4.4. When $\mathcal{V}$ is a (complex) Hilbert space $\mathcal{H}$ and one of the sets A, B consists of a single point, Thm. 9.4.3 can be proved directly by adapting the argument of Thm. 9.3.5.

Indeed, by translation we may assume WLOG that $A = \{0\}$. By the following Thm. 9.4.5, there exists $p \in B$ such that $\|p\| = \mathrm{dist}(0, B)$. Using the convexity of B, it is not difficult to show that $\langle \xi, p \rangle_{\mathbb{R}} \geqslant \langle p, p \rangle_{\mathbb{R}}$ for each $\xi \in B$, where $\langle \cdot, \cdot \rangle_{\mathbb{R}} = \mathrm{Re}\langle \cdot | \cdot \rangle$ is the real inner product associated with $\mathcal{H}$. Thus Thm. 9.4.3 follows by taking $\varphi = \langle p | \cdot \rangle$, $a = 0$, and $b = \|p\|^2$. $\square$

In the above remark we used the existence part of the following property.

Theorem 9.4.5. *Let $\mathcal{H}$ be a Hilbert space, let A be a closed convex subset of $\mathcal{H}$, and let $v \in \mathcal{H} \backslash A$. Then there exists a unique $\xi \in A$ such that $\|v - \xi\| = \mathrm{dist}(v, A)$.*

Proof. The proof relies on the **parallelogram law**

$$2\|\xi\|^2 + 2\|\eta\|^2 = \|\xi + \eta\|^2 + \|\xi - \eta\|^2 \qquad \text{for each } \xi, \eta \in \mathcal{H} \tag{9.17}$$

which is easy to verify.

By translation, we assume WLOG $v = 0$. Let $D = \mathrm{dist}(0, A)$. If $\xi, \eta \in A$ both satisfy $\|\xi\| = \|\eta\| = D$, then, since $\frac{\xi+\eta}{2} \in A$ (because A is convex), we have $\|\xi + \eta\|^2 \geqslant 4D^2$, and hence $\|\xi - \eta\|^2 \leqslant 0$ by (9.17). This proves the uniqueness.

To prove the existence, choose a sequence (ξ_n) in A such that $\lim_n \|\xi_n\| = D$. For each $m, n \in \mathbb{Z}_+$, since A is convex, we have $\frac{\xi_m + \xi_n}{2} \in A$ and hence $\|\xi_m + \xi_n\|^2 \geqslant 4D^2$. Therefore, by (9.17),

$$\limsup_{m,n} \|\xi_m - \xi_n\|^2 \leqslant \limsup_{m,n} 2\|\xi_m\|^2 + \limsup_{m,n} 2\|\xi_n\|^2 - 4D^2 \leqslant 0$$

Therefore (ξ_n) is a Cauchy sequence, and hence converges to a vector $\xi \in A$ satisfying $\|\xi\| = D$. $\square$

9.5 Strong and weak closures of convex sets

Let $\mathcal{V}$ be a normed $\mathbb{F}$-vector space. In this section, we generalize Cor. 3.4.8 and 3.4.9 from the Hilbert space setting to the broader context of normed vector spaces.

Definition 9.5.1. The **weak topology** on $\mathcal{V}$ is defined to be the pullback of the weak-* topology of $\mathcal{V}^{**} = (\mathcal{V}^*)^*$ by the canonical map $\mathbf{J}_{\mathcal{V}} : \mathcal{V} \to \mathcal{V}^{**}$. Therefore, a net (ξ_α) in $\mathcal{V}$ converges weakly to $\xi \in \mathcal{V}$ iff

$$\lim_\alpha \langle \varphi, \xi_\alpha \rangle = \langle \varphi, \xi \rangle \qquad \text{for each } \varphi \in \mathcal{V}^*$$

Remark 9.5.2. In Def. 3.7.1 we introduced a notion of weak topology for inner product spaces. *That definition agrees with Def. 9.5.1 when $\mathcal{V}$ is a Hilbert space, or when $\mathcal{V}$ is an inner product space and we restrict attention to bounded subsets of $\mathcal{V}$.*

To see this, we view the inner product space $\mathcal{V}$ as a dense linear subspace of a Hilbert space $\mathcal{H}$ (cf. Pb. 3.2). By the Riesz-Fréchet Thm. 3.5.3 and the fact that $\mathcal{V}^* = \mathcal{H}^*$ (cf. Cor. 2.4.3), bounded linear functionals on $\mathcal{V}$ are precisely pairings with elements of $\mathcal{H}$. Therefore, if (ξ_α) is a net in $\mathcal{V}$ and $\xi \in \mathcal{V}$, then (ξ_α) converges weakly to v in the sense of Def. 9.5.1 iff

$$\lim_\alpha \langle \eta | \xi_\alpha \rangle = \langle \eta | \xi \rangle \qquad \text{for each } \eta \in \mathcal{H}$$

This is clearly equivalent to the weak convergence in Def. 3.7.1 when $\mathcal{V}$ is a Hilbert space (i.e. $\mathcal{V} = \mathcal{H}$), or—by Thm. 2.6.2—when $\sup_\alpha \|\xi_\alpha\| < +\infty$.

If $\mathcal{V}$ is incomplete and $\sup_\alpha \|\xi_\alpha\| = +\infty$, the two notions of weak convergence are inequivalent (cf. Pb. 9.1). Fortunately, we have never used the notion of weak topology and weak convergence from Def. 3.7.1 in such a situation. For the remainder of this course, the weak topology always refers to that of Def. 9.5.1. □

We are now ready to state the consequence of separating convex sets in the weak topology.

Corollary 9.5.3. *Let C be a convex subset of $\mathcal{V}$. Then the norm closure $\overline{C}$ and the weak closure $\overline{C}^{\mathrm{w}}$ of C coincide.*

Proof. If $\xi \in \overline{C}$, there is a sequence (ξ_n) in C converging to ξ, and hence converging weakly to ξ. Thus $\xi \in \overline{C}^{\mathrm{w}}$.

Conversely, suppose $\xi \notin \overline{C}$. Since $\overline{C}$ is convex, by Thm. 9.4.3 (or Rem. 9.4.4 when $\mathcal{V}$ is a Hilbert space), there exist $\varphi \in \mathcal{V}^*$ and real numbers $a < b$ such that $\mathrm{Re}\varphi(v) \leqslant a$ for all $v \in C$, and $\mathrm{Re}\varphi(\xi) \geqslant b$. Thus, every net (ξ_α) in C must satisfy $\mathrm{Re}\varphi(\xi_\alpha) \leqslant a$, and hence cannot converge weakly to ξ because $\lim_\alpha \mathrm{Re}\varphi(\xi_\alpha) \neq \mathrm{Re}\varphi(\xi)$. So $\xi \notin \overline{C}^{\mathrm{w}}$. □

Remark 9.5.4. In the case that $\mathcal{V}$ is a Hilbert space and C is a linear subspace, Cor. 9.5.3 follows directly from the relation $\overline{C} = C^{\perp\perp}$ (cf. Cor. 3.4.8). Indeed, since $C^{\perp\perp}$ is weakly closed, it contains the weak closure $\overline{C}^{\mathrm{w}}$ of C. Since $\overline{C} \subset \overline{C}^{\mathrm{w}}$, we have

$$\overline{C} \subset \overline{C}^{\mathrm{w}} \subset C^{\perp\perp} = \overline{C}$$

and hence $\overline{C} = \overline{C}^{\mathrm{w}}$.

Corollary 9.5.5. *For each $E \subset \mathcal{V}$, define the **convex hull***

$$\mathbf{conv}(\boldsymbol{E}) = \Big\{ \sum_{j=1}^{n} a_j \xi_j : n \in \mathbb{Z}_+, a_j \in \mathbb{R}_{\geqslant 0}, \xi_j \in E, \sum_j a_j = 1 \Big\}$$

which is the smallest convex set containing E. Suppose that there is a net in E converging weakly to $\xi \in \mathcal{V}$. Then there is a sequence in $\mathrm{conv}(E)$ converging (in norm) to ξ.

Any element belonging to $\mathrm{conv}(E)$ is called a **convex combination** of E.

Proof. By Cor. 9.5.3, we have $\xi \in \overline{E}^{\mathrm{w}} \subset \overline{\mathrm{conv}(E)}^{\mathrm{w}} = \overline{\mathrm{conv}(E)}$. □

Remark 9.5.6. Let (ξ_n) be a sequence in $\mathcal{V}$ converging weakly to ξ. By Cor. 9.5.5, there exists a sequence $(\eta_n)_{n\in\mathbb{Z}_+}$ of convex combinations of (ξ_n) converging to ξ in norm. In the case $\mathcal{V} = L^p(X,\mu)$ where (X,μ) is a measure space and $1 < p < +\infty$ (in particular, when $\mathcal{V}$ is a Hilbert space), the sequence $(\eta_k)_{k\in\mathbb{Z}_+}$ can be chosen more explicit: of the form

$$\eta_k = \frac{\xi_{n_1} + \cdots + \xi_{n_k}}{k}$$

for some subsequence (ξ_{n_k}) of (ξ_n). This is known as the **Banach-Saks theorem**. See [RF, Sec. 8.4] or [RN, Sec. 38] for the case $p = 2$, and the original paper of Banach and Saks [BS30] for general p.

Recall Def. 9.2.5 for the notion of ${}^\perp E$ and $F^\top$. With the aid of the following corollary, several density results previously proved for Hilbert spaces and L^2-spaces can be extended to more general function spaces; see Subsec. 9.8.1 for applications.

Corollary 9.5.7. *Let $\mathcal{U}$ be a linear subspace of $\mathcal{V}$. Then*

$$\overline{\mathcal{U}} = {}^\perp(\mathcal{U}^\top)$$

In particular, $\mathcal{U}$ is dense iff $\mathcal{U}^\top = 0$.

This provides an alternative proof that $\overline{\mathcal{U}} = \overline{\mathcal{U}}^{\mathrm{w}}$, because for any $E \subset \mathcal{V}$, the set ${}^\perp(E^\top)$ is weakly-closed, and hence $\overline{E} \subset \overline{E}^{\mathrm{w}} \subset {}^\perp(E^\top)$.

First proof. Clearly $\overline{\mathcal{U}} \subset {}^\perp(\mathcal{U}^\top)$. Conversely, suppose $\xi \in \mathcal{V} \backslash \overline{\mathcal{U}}$. By Thm. 9.4.3, there exist $\varphi \in \mathcal{V}^*$ and $-\infty < a < b < +\infty$ such that $\mathrm{Re}\varphi(\overline{\mathcal{U}}) \subset (-\infty, a]$ and $\mathrm{Re}\varphi(\xi) \geqslant b$. Since every non-zero real linear functional has range $\mathbb{R}$, we obtain $\mathrm{Re}\varphi|_{\mathcal{U}} = 0$, and hence $\varphi|_{\mathcal{U}} = 0$ by Prop. 9.1.5. So $a \geqslant 0$, and hence $\mathrm{Re}\varphi(\xi) > 0$. Thus $\varphi \in \mathcal{U}^\top$ but $\varphi(\xi) \neq 0$. Therefore $\xi \notin {}^\perp(\mathcal{U}^\top)$. The proof of $\overline{\mathcal{U}} = {}^\perp(\mathcal{U}^\top)$ is complete.

If $\mathcal{U}$ is dense, any bounded linear functional vanishing on $\mathcal{U}$ must vanish on its closure $\mathcal{V}$. Thus $\mathcal{U}^\top = 0$. Conversely, if $\mathcal{U}$ is not dense, then ${}^\perp(\mathcal{U}^\top) \neq \mathcal{V}$, and hence $\mathcal{U}^\top \neq 0$. □

Second proof. Clearly $\overline{\mathcal{U}} \subset ({}^\perp\mathcal{U})^\top$. Suppose that $\xi \in \mathcal{V} \backslash \overline{\mathcal{U}}$. Then $M := \mathrm{dist}(\xi, \mathcal{U}) > 0$. Therefore, for each $\eta \in \mathcal{U}$ and $a \in \mathbb{F}$ we have

$$\|a\xi + \eta\| \geqslant |a| \mathrm{dist}(\xi, \mathcal{U}) = |a| M$$

and hence the linear functional

$$\varphi : \mathbb{F}\xi + \mathcal{U} \qquad a\xi + \eta \mapsto a$$

is bounded with operator norm $\leqslant M^{-1}$. By the Hahn-Banach Thm. 9.1.1, φ can be extended to a linear functional on $\mathcal{V}$. So ξ does not vanish when evaluated with the element $\varphi \in \mathcal{U}^\top$. Hence $\xi \notin {}^\perp(\mathcal{U}^\top)$. □

9.6 Reflexive spaces: where weak and weak-* topologies coincide

Let $\mathcal{V}, \mathcal{W}$ be normed vector spaces over $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$.

9.6.1 Comparison of weak and weak-* topologies

In this course, we have encountered two topologies on function spaces that are weaker than the norm topology: the weak-* topology and the weak topology. The principal advantage of the weak-* topology lies in the compactness property provided by the Banach-Alaoglu Thm. 2.6.5. The weak topology, on the other hand, offers a strong connection between weak closures and norm closures, as demonstrated in Sec. 9.5. In addition, the weak topology has two further merits.

First, when referring to the weak-* topology on a function space, one must specify a predual, i.e., a space whose dual is the given function space; different choices of preduals yield different weak-* topologies.[8] In contrast, the weak topology involves no such ambiguity. Second, the weak topology enjoys the following generalization of Prop. 8.4.2:

[8]More precisely, if $\mathcal{W} = \mathcal{U}^* = \mathcal{V}^*$ and $\mathbf{J}_{\mathcal{U}}(\mathcal{U}) \neq \mathbf{J}_{\mathcal{V}}(\mathcal{V})$, then the weak-* topologies on $\mathcal{W}$ defined by $\mathcal{U}$ and $\mathcal{V}$ are distinct. See Pb. 9.1.

Proposition 9.6.1. *Let $T \in \mathfrak{L}(\mathcal{V}, \mathcal{W})$. Then $T : \mathcal{V} \to \mathcal{W}$ is continuous when both $\mathcal{V}$ and $\mathcal{W}$ are equipped with their weak topologies.*

Proof. Let (ξ_α) be a net in $\mathcal{V}$ converging weakly to ξ. Then for each $\psi \in \mathcal{W}^*$ we have

$$\lim_\alpha \langle \psi, T\xi_\alpha \rangle = \lim_\alpha \langle T^\sharp \psi, \xi_\alpha \rangle = \langle T^\sharp \psi, \xi \rangle = \langle \psi, T\xi \rangle$$

where $T^\sharp \psi := \psi \circ T$ has operator norm $\leqslant \|\psi\| \cdot \|T\|$ by Prop. 3.5.14. Therefore $T\xi_\alpha$ converges weakly to $T\xi$. □

In contrast, a bounded linear map $\mathcal{W}^* \to \mathcal{V}^*$ is not necessarily continuous when both spaces are equipped with their weak-* topologies, except when it arises as the transpose of a bounded linear operator $\mathcal{V} \to \mathcal{W}$. We take this opportunity to introduce the notion of transpose.

9.6.2 The transpose of a bounded linear operator

Definition 9.6.2. For each $T \in \mathfrak{L}(\mathcal{V}, \mathcal{W})$, define the **transpose**[9]

$$T^\sharp : \mathcal{W}^* \to \mathcal{V}^* \qquad \psi \mapsto \psi \circ T$$

By Prop. 3.5.14, we have $\|\psi \circ T\| \leqslant \|\psi\| \cdot \|T\|$ for each $\psi \in \mathcal{W}^*$, and hence $\|T^\sharp\| \leqslant \|T\|$.

The notation $T^\sharp$ is compatible with notation $v^\sharp$ for $v \in \mathcal{V}$ (cf. Def. 9.2.1) if we view v as the linear map $\mathbb{F} \to \mathcal{V}$ defined by $\lambda \mapsto \lambda v$, and if we identify $\mathbb{F}$ with $\mathbb{F}^*$ by sending 1 to $\mathrm{id}_{\mathbb{F}}$.

With the aid of the Hahn-Banach theorem, we will show below that in fact $\|T^\sharp\| = \|T\|$.

Theorem 9.6.3. *The canonical map $\mathbf{J}_{\mathcal{V}} : \mathcal{V} \to \mathcal{V}^{**}, v \mapsto v^\sharp$ is a linear isometry. More generally, the following map is a linear isometry*

$$\mathfrak{L}(\mathcal{V}, \mathcal{W}) \to \mathfrak{L}(\mathcal{W}^*, \mathcal{V}^*) \qquad T \mapsto T^\sharp$$

Proof. From Def. 9.2.1, we already know that $\|v^\sharp\| \leqslant \|v\|$ for each $v \in \mathcal{V}$. To show $\|v^\sharp\| \geqslant \|v\|$, assume WLOG that $v \neq 0$, and normalize so that $\|v\| = 1$. By the Hahn-Banach Thm. 9.1.1, the linear functional $\mathbb{F}v \to \mathbb{F}$ defined by $\lambda v \mapsto \lambda$ can be extended to a bounded linear functional $\varphi \in \mathcal{V}^*$ with $\|\varphi\| \leqslant 1$. We have thus obtained $\varphi \in \overline{B}_{\mathcal{V}^*}(0,1)$ such that $|\langle \varphi, v^\sharp \rangle| = 1$, which proves $\|v^\sharp\| \geqslant 1$. This finishes the proof that $\|v^\sharp\| = \|v\|$, i.e., that $\mathbf{J}_{\mathcal{V}}$ is an isometry.

Thus we may identify $\mathcal{V}$ with $\mathbf{J}_{\mathcal{V}}(\mathcal{V})$. Under this identification, it is easy to see that $T^{\sharp\sharp} : \mathcal{V}^{**} \to \mathcal{W}^{**}$ restricts to T, and hence $\|T\| \leqslant \|T^{\sharp\sharp}\|$. On the other hand,

[9]Many texts use the term adjoint here. We reserve adjoint for bounded (or unbounded) linear operators between Hilbert spaces, which is distinct from the transpose defined here.

by Def. 9.6.2, we have $\|T^\sharp\| \leqslant \|T\|$, and similarly $\|T^{\sharp\sharp}\| \leqslant \|T^\sharp\|$. Combining these inequalities together, we obtain $\|T\| = \|T^\sharp\|$.[10] □

Convention 9.6.4. In view of Thm. 9.6.3, we henceforth regard $\mathcal{V}$ as a normed linear subspace of $\mathcal{V}^{**}$ by identifying each $v \in \mathcal{V}$ with the linear functional $v^\sharp : \mathcal{V}^* \to \mathbb{F}$.

Corollary 9.6.5. *$\mathcal{V}^*$ separates points of $\mathcal{V}$. Consequently, the weak topology on $\mathcal{V}$ is Hausdorff.*

Proof. The injectivity of $\mathbf{J}_\mathcal{V}$, established in Thm. 9.6.3, is precisely the statement that $\mathcal{V}^*$ separates points of $\mathcal{V}$. Consequently, if (v_α) is a net in $\mathcal{V}$ converging weakly to both v and v', then $\langle v, \varphi\rangle = \lim_\alpha \langle v_\alpha, \varphi\rangle = \langle v', \varphi\rangle$ for each $\varphi \in \mathcal{V}^*$, and hence $v = v'$. □

9.6.3 Reflexive spaces

Roughly speaking, reflexive spaces are normed vector spaces whose weak and weak-* topologies coincide. Such spaces inherit the principal advantages of both topologies.

Definition 9.6.6. The normed vector space $\mathcal{V}$ is called **reflexive** if $\mathcal{V}^{**} = \mathbf{J}_\mathcal{V}(\mathcal{V})$; that is, if every bounded linear functional $\mathcal{V}^* \to \mathbb{F}$ is of the form $v^\sharp$ for some $v \in \mathcal{V}$. In view of Conv. 9.6.4, this means precisely that $\mathcal{V} = \mathcal{V}^{**}$.

Example 9.6.7. By Thm. 8.4.14, every finite-dimensional normed vector space is isomorphic to $\mathbb{F}^N$ for some N, and hence is reflexive.

Example 9.6.8. By the Riesz-Fréchet Thm. 3.5.3, every Hilbert space is reflexive.

Example 9.6.9. Let X be a set, and let $1 < p < +\infty$. Then $l^p(X, \mathbb{F})$ is reflexive by Thm. 2.8.7.

Example 9.6.10. Let (X, μ) be a σ-finite measure space, and let $1 < p < +\infty$. Then $L^p(X, \mu, \mathbb{F})$ is reflexive by Thm. 1.6.16.

We now illustrate how the advantage of the weak topology, as shown in Prop. 9.6.1, together with the compactness provided by the weak-* topology via the Banach-Alaoglu theorem, leads to the following generalization (Thm. 9.6.13) of Thm. 8.4.7 on the equivalence between completely continuous operators and compact operators.

[10] A more concrete argument: for any $v \in \overline{B}_\mathcal{V}(0,1)$, we have $\|Tv\| = \|(Tv)^\sharp\| = \sup |\langle (Tv)^\sharp, \psi\rangle| = \sup |\langle Tv, \psi\rangle| = \sup |\langle v, T^\sharp\psi\rangle| \leqslant \sup \|T^\sharp\psi\| \leqslant \|T^\sharp\|$ where the sup is over all $\psi \in \overline{B}_{\mathcal{W}^*}(0,1)$. Thus $\|T\| \leqslant \|T^\sharp\|$, and hence $\|T\| = \|T^\sharp\|$.

Definition 9.6.11. A linear map $T : \mathcal{V} \to \mathcal{W}$ is called **completely continuous** if T is bounded, and if the restriction

$$T|_{\overline{B}_{\mathcal{V}}(0,1)} : \overline{B}_{\mathcal{V}}(0,1) \to \mathcal{W}$$

is continuous when $\overline{B}_{\mathcal{V}}(0,1)$ is equipped with the weak topology and $\mathcal{W}$ is equipped with the norm topology.

Remark 9.6.12. In the case of Hilbert spaces, the boundedness requirement in the definition of complete continuity is redundant, due to the weak compactness of the closed unit ball. In general, however, $\overline{B}_{\mathcal{V}}(0,1)$ need not be weakly compact; indeed, as we will see shortly, reflexivity is equivalent to the weak compactness of $\overline{B}_{\mathcal{V}}(0,1)$.

Theorem 9.6.13. *Let $T : \mathcal{V} \to \mathcal{W}$ be a linear map. Assume that $\mathcal{V}$ is reflexive. Then the following statements are equivalent.*

(1) The bilinear form

$$\omega_T : \mathcal{W}^* \times \mathcal{V} \to \mathbb{F} \qquad (\psi, \xi) \mapsto \langle \psi, T\xi \rangle$$

*is **completely continuous**, that is, it restricts to a continuous map $\overline{B}_{\mathcal{W}^*}(0,1) \times \overline{B}_{\mathcal{V}}(0,1) \to \mathbb{F}$ when $\overline{B}_{\mathcal{W}^*}(0,1)$ is equipped with the weak-* topology, and $\overline{B}_{\mathcal{V}}(0,1)$ is equipped with the weak topology.*

(2) T is completely continuous.

(3) T is compact.

Proof. (3)⇒(2): The proof proceeds as in the implication (3)⇒(2) of Thm. 8.4.7, with Prop. 9.6.1 replacing Prop. 8.4.2.

(2)⇒(3): Since $\mathcal{V} = \mathcal{V}^{**}$, the weak topology on $\mathcal{V}$ coincides with the weak-* topology of $\mathcal{V}^{**}$. Therefore, $\overline{B}_{\mathcal{V}}(0,1)$ is weakly compact by Banach-Alaoglu.

(A more concrete argument: Suppose (v_α) is a net in $\overline{B}_{\mathcal{V}}(0,1)$. By Banach-Alaoglu, the net $(v_\alpha^\sharp)$ admits a subnet $(v_\beta^\sharp)$ converging weak-* to some $\psi \in \mathcal{V}^*$. Since $\mathcal{V}$ is reflexive, we have $\psi = v^\sharp$ for some $v \in \mathcal{V}$. Thus (v_β) converges weakly to v.)

The image of a compact space under a continuous map is compact. Thus $T(\overline{B}_{\mathcal{V}}(0,1))$ is compact, and hence T is compact.

(1)⇔(2): This follows from the following Lem. 9.6.14. □

Lemma 9.6.14. *Let $(\eta_\alpha)_{\alpha \in I}$ be a net in $\overline{B}_{\mathcal{W}}(0,1)$ converging weakly to $\eta \in \overline{B}_{\mathcal{W}}(0,1)$. Then the following are equivalent.*

(1) $(\eta_\alpha)_{\alpha \in I}$ converges to η.

(2) *For each net* $(\psi_\nu)_{\nu\in K}$ *in* $\overline{B}_{\mathcal{W}^*}(0,1)$ *converging weak-* to some* $\psi \in \overline{B}_{\mathcal{W}^*}(0,1)$ *we have* $\lim_{\alpha\in I,\nu\in K}\langle\psi_\alpha|\eta_\nu\rangle = \langle\psi|\eta\rangle$.

Proof. (1)⇒(2): This is proved in the same way as the implication (1)⇒(2) in Lem. 8.2.11.

¬(1)⇒¬(2): Assume that (η_α) does not converge to η. Then there exists $\varepsilon > 0$ such that $\{\alpha \in I : \|\eta_\alpha - \eta\| \geqslant \varepsilon\}$ is a cofinal subset of I. It follows that (η_α) admits a subnet $(\eta_\beta)_{\beta\in J}$ satisfying $\|\eta_\beta - \eta\| \geqslant \varepsilon$ for all β. By Thm. 9.6.3, $\|\eta_\beta^\sharp - \eta^\sharp\| \geqslant \varepsilon$. Therefore, there exist a net $(\psi_\beta)_{\beta\in J}$ in $\overline{B}_{\mathcal{W}^*}(0,1)$ such that

$$|\langle\psi_\beta, \eta_\beta - \eta\rangle| \geqslant \varepsilon/2$$

for all β. By Banach-Alaoglu, (ψ_β) admits a subnet $(\psi_\nu)_{\nu\in K}$ converging weak-* to some $\psi \in \overline{B}_{\mathcal{W}^*}(0,1)$. Thus $\lim_\nu\langle\psi_\nu,\eta\rangle = \langle\psi,\eta\rangle$, and hence $\lim_{\nu\in K}\langle\psi_\nu,\eta_\nu\rangle$ does not converge to $\langle\psi,\eta\rangle$ by the above inequality. It follows that $\lim_{\alpha\in I,\nu\in K}\langle\psi_\alpha|\eta_\nu\rangle$ does not converge to $\langle\psi|\eta\rangle$, and hence (2) is false. □

Corollary 9.6.15. *Assume that* $\mathcal{V}$ *and* $\mathcal{W}$ *are reflexive. Let* $T \in \mathfrak{L}(\mathcal{V},\mathcal{W})$. *Then* T *is compact iff* $T^\sharp : \mathcal{W}^* \to \mathcal{V}^*$ *is compact.*

Proof. It is easy to see that $\mathcal{V}^*$ and $\mathcal{W}^*$ are reflexive (cf. Cor. 9.7.7). By Thm. 9.6.13, the compactness of T and $T^\sharp$ are equivalent to the complete continuity of the bilinear forms

$$\begin{aligned} \omega_T &: \mathcal{W}^* \times \mathcal{V} \to \mathbb{F} \qquad (\psi,\xi) \mapsto \langle\psi, T\xi\rangle \\ \omega_{T^\sharp} &: \mathcal{V} \times \mathcal{W}^* \to \mathbb{F} \qquad (\xi,\psi) \mapsto \langle\xi, T^\sharp\psi\rangle \end{aligned}$$

respectively. These two bilinear forms coincide up to an exchange of the factors in the domain. Hence, the equivalence of the compactness of T and $T^\sharp$ follows. □

By adapting the above proof, one can show that Cor. 9.6.15 holds without assuming that $\mathcal{V}$ or $\mathcal{W}$ is reflexive; see Pb. 9.8.

9.6.4 Reflexivity and weak-compactness

As we observed in the proof of Thm. 9.6.13, the closed unit ball of a reflexive space is weakly compact. We now show that this weak compactness property completely characterizes reflexivity.

Theorem 9.6.16 (Goldstine theorem). $\overline{B}_{\mathcal{V}}(0,1)$ *is dense in* $\overline{B}_{\mathcal{V}^{**}}(0,1)$ *with respect to the weak-* topology on* $\overline{B}_{\mathcal{V}^{**}}(0,1)$. *Consequently,* $\mathcal{V}$ *is dense in* $\mathcal{V}^{**}$ *with respect to the weak-* topology on* $\mathcal{V}^{**}$.

Proof. Let $\psi \in \overline{B}_{\mathcal{V}^{**}}(0,1)$. Let $\mathscr{I}$ be the set of finite dimensional linear subspaces of $\mathcal{V}^*$, directed by $\subset$. We claim that for $\gamma > 1$ and $\mathcal{X} \in \mathscr{I}$, there exists $\xi_{\gamma,\mathcal{X}} \in \overline{B}_{\mathcal{V}^{**}}(0,\gamma)$ such that

$$\xi^\sharp_{\gamma,\mathcal{X}}|_{\mathcal{X}} = \psi|_{\mathcal{X}}$$

Assuming the claim, equip $\mathbb{R}_{>1}$ with an order in which values closer to 1 are larger. Then the net $(\xi^\sharp_{\gamma,\mathcal{X}})_{(\gamma,\mathcal{X})\in\mathbb{R}_{>1}\times\mathscr{I}}$ clearly converges weak-* to ψ. Equivalently, the net $(\xi_{\gamma,\mathcal{X}})$ converges weak-* to ψ. Hence $(\gamma^{-1}\xi_{\gamma,\mathcal{X}})$ is a net in $\overline{B}_{\mathcal{V}}(0,1)$ converging weak-* to ψ, finishing the proof.

Choose any $\gamma > 1$ and $\mathcal{X} \in \mathscr{I}$. By Thm. 9.2.9, there exists $\xi_{\mathcal{X}} \in \mathcal{V}$ such that $\xi^\sharp_{\mathcal{X}} = \psi|_{\mathcal{X}}$. By Lem. 9.2.10, we have

$$\|\xi_{\mathcal{X}} + {}^\perp\mathcal{X}\| = \|\psi|_{\mathcal{X}}\| \leqslant \|\psi\| \leqslant 1$$

By the definition of the quotient norm, there exists $u_{\gamma,\mathcal{X}} \in {}^\perp\mathcal{X}$ such that $\xi^\sharp_{\gamma,\mathcal{X}} := \xi_{\mathcal{X}} + u_{\gamma,\mathcal{X}}$ has norm $\leqslant \gamma$. This proves the claim. □

Corollary 9.6.17. *$\mathcal{V}$ is reflexive iff $\overline{B}_{\mathcal{V}}(0,1)$ is weakly compact.*

Proof. The implication $\Rightarrow$ is immediate, as explained in the proof of Thm. 9.6.13. Conversely, assume that $\overline{B}_{\mathcal{V}}(0,1)$ is weakly compact. Since the weak-* topology on $\overline{B}_{\mathcal{V}^{**}}(0,1)$ restricts to the weak topology on $\overline{B}_{\mathcal{V}}(0,1)$, it follows that $\overline{B}_{\mathcal{V}}(0,1)$ is weak-* closed in $\overline{B}_{\mathcal{V}^{**}}(0,1)$ (because every compact subset of a Hausdorff space is closed). Therefore, by Thm. 9.6.16, we obtain $\overline{B}_{\mathcal{V}}(0,1) = \overline{B}_{\mathcal{V}^{**}}(0,1)$ and hence $\mathcal{V} = \mathcal{V}^{**}$. □

9.7 Examples of non-reflexive spaces

Let $\mathcal{V}$ be a normed $\mathbb{F}$-vector space. In this section, we explore examples of non-reflexive spaces.

Note that reflexive spaces are dual spaces, and hence must be Banach spaces by Cor. 2.4.11. Therefore, any non-complete normed vector space cannot be reflexive. Thus, the only nontrivial task in searching for non-reflexive spaces is to identify non-reflexive Banach spaces.

Proposition 9.7.1. *Let X be a set, equipped with the discrete topology. Then we have an isomorphism of normed vector spaces*

$$l^1(X,\mathbb{F}) \xrightarrow{\simeq} C_0(X,\mathbb{F})^* \qquad f \mapsto \Big(g \mapsto \sum_X fg\Big)$$

Proof. This follows either from the Riesz-Markov theorem (Thm. 1.7.17 and Rem. 1.7.18) together with Thm. 2.10.6 on the norms of complex measures, or by a direct argument, which we leave to the reader as an exercise. □

Example 9.7.2. Let X be a set. Then $C_0(X,\mathbb{F})^* = l^1(X,\mathbb{F})$ by Prop. 9.7.1, and $l^1(X,\mathbb{F})^* = l^\infty(X,\mathbb{F})$ by Thm. 2.8.7. Therefore, if X is infinite, then $C_0(X,\mathbb{F}) \subsetneq l^\infty(X,\mathbb{F})$, so $C_0(X,\mathbb{F})$ is not reflexive.

Example 9.7.3. Let X be an LCH space that is not discrete, i.e., there exists $p \in X$ such that $\{p\}$ is not open in X. Then $C_0(X,\mathbb{F})$ is not reflexive. In particular, $C([0,1],\mathbb{F})$ is not reflexive.

Proof. By the Riesz-Markov Thm. 1.7.17, the dual space $C_0(X,\mathbb{F})^*$ can be identified with the space $\mathcal{RM}(X,\mathbb{F})$ of real or complex Radon measures on X, where the norm $\|\mu\|$ of each $\mu \in \mathcal{RM}(X,\mathbb{F})$ is described by Thm. 2.10.6. The linear functional

$$\psi : \mathcal{RM}(X,\mathbb{F}) \qquad \mu \mapsto \mu(\{p\}) \tag{9.18}$$

is clearly bounded, and hence belongs to $C_0(X,\mathbb{F})^{**}$. Let us prove that $\psi \notin C_0(X,\mathbb{F})$.

Suppose that ψ is represented by some $f \in C_0(X,\mathbb{F})$. Then for each $\mu \in \mathcal{RM}(X,\mathbb{F})$ we have

$$\int_X f d\mu = \mu(\{p\}) = \int_X \chi_{\{p\}} d\mu$$

By choosing μ to be Dirac measures on X, we see that $f = \chi_{\{p\}}$. However, since $\{p\}$ is not open, $\chi_{\{p\}}$ is not continuous. This yields a contradiction. $\square$

Unlike Exp. 9.7.2 and 9.7.3, where non-reflexivity was shown by explicitly constructing an element lying in the double dual but not in the original space, the fact that $l^\infty(\mathbb{Z})^* \supsetneq l^1(\mathbb{Z})$ (and hence $l^1(\mathbb{Z})$ is not reflexive) is inherently nonconstructive. The standard proof relies on applying the Hahn-Banach theorem to $l^\infty(\mathbb{Z})$ which is non-separable, and therefore ultimately depends on the axiom of choice. Remarkably, there exist axiomatic systems in which $l^1(\mathbb{Z})$ becomes reflexive; see [Väth98].

To see why $l^\infty(\mathbb{Z})^*$ is strictly larger than $l^1(\mathbb{Z})$, we must examine why the proof of $l^p(\mathbb{Z}) \simeq l^q(\mathbb{Z})^*$ in Thm. 2.8.7 breaks down in the case $p = 1, q = +\infty$. Recall the argument in Thm. 2.8.7: Given $\varphi \in l^q(\mathbb{Z})^*$, one defines $f(n) = \langle \varphi, \chi_{\{n\}} \rangle$ for $n \in \mathbb{Z}$, and verifies that $f \in l^p(\mathbb{Z})$. This step remains valid when $p = 1$. However, because the characteristic functions do not span a dense subspace of $l^\infty(\mathbb{Z})$, we cannot conclude that φ is represented by f. This observation leads to the following general theorem, where $\mathscr{X}$ may be taken as $\{\chi_n : n \in \mathbb{Z}\}$ when $\mathcal{V} = l^1(\mathbb{Z})$.

Theorem 9.7.4. *The following statements are equivalent.*

(1) $\mathcal{V}$ is not reflexive.

(2) There exists $\mathscr{X} \subset \mathcal{V}^*$ that separates points of $\mathcal{V}$ such that the (norm-)closure $\overline{\mathrm{Span}\mathscr{X}}$ is not equal to the whole dual space $\mathcal{V}^*$.

Proof. (2)⇒(1): Assume (2). Replacing $\mathcal{X}$ with $\overline{\mathrm{Span}\mathcal{X}}$, we assume that $\mathcal{X}$ is a proper closed linear subspace of $\mathcal{V}^*$. By Cor. 9.5.7, we have $\mathcal{X}^\top \neq 0$, that is, there exists a nonzero $\psi \in \mathcal{V}^{**}$ vanishing on $\mathcal{X}$. Then $\psi \notin \mathcal{V}$; otherwise, the fact that $\mathcal{X}$ separates points of $\mathcal{V}$ yields the contradiction $\psi = 0$. This proves (1).

(1)⇒(2): Assume (1). Choose $\psi \in \mathcal{V}^{**}\backslash\mathcal{V}$. Let $\mathcal{L} = \mathbb{F}\psi$. Then ${}^\perp\mathcal{L}$ is a closed linear subspace of $\mathcal{V}^*$. By Cor. 9.2.7, we have $({}^\perp\mathcal{L})^\top = \mathcal{L}$, and hence $({}^\perp\mathcal{L})^\top \cap \mathcal{V} = 0$. Thus ${}^\perp\mathcal{L}$ separates points of $\mathcal{V}$. Since $(\mathcal{V}^*)^\top = \{0_{\mathcal{V}^{**}}\} \neq \mathcal{L}$, we have ${}^\perp\mathcal{L} \neq \mathcal{V}^*$. □

Remark 9.7.5. From the proof of Thm. 9.7.4, we see that if $\mathcal{V}$ is not reflexive, then a proper closed linear subspace $\mathcal{X} \subsetneq \mathcal{V}^*$ separating points of $\mathcal{V}$ can be chosen to be $\mathcal{X} = {}^\perp(\mathbb{F}\psi)$ where $\psi \in \mathcal{V}^{**}\backslash\mathcal{V}$.

Therefore, if $\mathcal{V} = C_0(X, \mathbb{F})$ where X is a set, ψ can be choosen to be $1 \in l^\infty(X, \mathbb{F})$, and hence $\mathcal{X}$ consists of $f \in l^1(X, \mathbb{F})$ with $\sum_X f = 0$.

If $\mathcal{V} = C_0(X, \mathbb{F})$ where X is an LCH space with a non-isolated point $p \in X$, as seen in the proof of Exp. 9.7.3, ψ can be defined by (9.18), and hence $\mathcal{X}$ consists of all real or complex Radon measures μ satisfying $\mu(\{p\}) = 0$. □

Example 9.7.6. Let X be an infinite set. Then $l^1(X, \mathbb{F})$ is not reflexive.

Proof. Since $l^1(X, \mathbb{F})^* = l^\infty(X, \mathbb{F})$, and since $C_0(X, \mathbb{F})$ is a proper closed linear subspace of $l^\infty(X, \mathbb{F})$ separating points of $l^1(X, \mathbb{F})$, we conclude from Thm. 9.7.4 that $l^1(X, \mathbb{F})$ is not reflexive.

Alternatively, since $l^1(X, \mathbb{F}) = C_0(X, \mathbb{F})^*$ by Prop. 9.7.1, the non-reflexivity of $l^1(X, \mathbb{F})$ follows from the non-reflexivity of $C_0(X, \mathbb{F})$ and the following Cor. 9.7.7. (In fact, Cor. 9.7.7 as well as its proof is motivated by the argument in the first paragraph.) □

Corollary 9.7.7. *Assume that $\mathcal{V}$ is a Banach space. Then $\mathcal{V}$ is reflexive iff $\mathcal{V}^*$ is reflexive.*

Proof. If $\mathcal{V}$ is reflexive, then $\mathcal{V} = \mathcal{V}^{**}$, and hence $(\mathcal{V}^*)^{**} = (\mathcal{V}^{**})^* = \mathcal{V}^*$. So $\mathcal{V}^*$ is reflexive.

Conversely, assume that $\mathcal{V}$ is not reflexive. Then $\mathcal{V}$ is a proper closed linear subspace of $\mathcal{V}^{**}$, where closedness follows from the completeness of $\mathcal{V}$. Since $\mathcal{V}$ separates points of $\mathcal{V}^*$, by Thm. 9.7.4, $\mathcal{V}^*$ is not reflexive. □

Example 9.7.8. If X is an infinite set, then $l^\infty(X, \mathbb{F})$ is not reflexive.

Proof. This follows from Cor. 9.7.7, since $l^\infty(X, \mathbb{F}) = l^1(X, \mathbb{F})^*$ and $l^1(X, \mathbb{F})$ is not reflexive by Exp. 9.7.6. □

Example 9.7.9. $L^1([0,1], m)$ and $L^\infty([0,1], m)$ are not reflexive.

Proof. By Thm. 1.6.16, we have $L^\infty([0,1], m) = L^1([0,1], m)^*$. We view $C([0,1])$ as a closed linear subspace of $L^\infty([0,1], m)$, which is clearly proper. By Thm. 2.7.4, $C([0,1])$ is weak-* dense in $L^\infty([0,1], m)$, so it separates points of $L^1([0,1], m)$. It then follows from Thm. 9.7.4 that $L^1([0,1], m)$ is not reflexive. By Cor. 9.7.7, $L^\infty([0,1], m)$ is also not reflexive. □

Example 9.7.10. $L^1(\mathbb{R}^N, m)$ and $L^\infty(\mathbb{R}^N, m)$ are not reflexive.

Proof. The argument is the same as in the proof of Exp. 9.7.9, with the sole modification that $C([0,1])$ is replaced by $C_0(\mathbb{R}^N)$. □

9.8 Problems

Let $\mathcal{V}, \mathcal{W}$ be normed vector spaces over $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$.

Definition 9.8.1. Let $\mathcal{X} \subset \mathcal{V}^*$. The pullback of the product topology of $\mathbb{F}^{\mathcal{X}}$ by the map

$$\mathcal{V} \to \mathbb{F}^{\mathcal{X}} \qquad v \mapsto \Big(\varphi \in \mathcal{X} \mapsto \langle \varphi, v\rangle\Big) \tag{9.19}$$

is called the ***$\mathcal{X}$*-weak topology** (on $\mathcal{V}$) and denoted by $\boldsymbol{\sigma(\mathcal{V}, \mathcal{X})}$. In other words, it is the unique topology such that for any net (v_α) in $\mathcal{V}$ and $v \in \mathcal{V}$,

$$v_\alpha \xrightarrow{\sigma(\mathcal{V},\mathcal{X})} v \qquad \Longleftrightarrow \qquad \lim_\alpha \langle \varphi, v_\alpha\rangle = \langle \varphi, v\rangle \qquad \text{for each } \varphi \in \mathcal{X}$$

For example, the $\mathcal{V}^*$-weak topology of $\mathcal{V}$ is the usual weak topology, and the $\mathbf{J}_{\mathcal{V}}(\mathcal{V})$-weak topology of $\mathcal{V}^*$ (where $\mathbf{J}_{\mathcal{V}} : \mathcal{V} \to \mathcal{V}^{**}$ is the canonical map) is precisely the weak-* topology.

Remark 9.8.2. Let $\mathcal{X} \subset \mathcal{V}^*$. Then the injectivity of (9.19) is clearly equivalent to $\mathcal{X}$ separating points of $\mathcal{X}$ (cf. Def. 1.5.10). In that case, the $\mathcal{X}$-weak topology on $\mathcal{V}$ is Hausdorff.

Remark 9.8.3. From the definition of product topology in terms of a base (cf. Def. 1.4.12), it follows that a neighborhood basis of the $\mathcal{X}$-weak topology at $v \in \mathcal{V}$ may be taken to consist of the sets

$$U(v;\varphi_\bullet;\varepsilon_\bullet) \equiv U(v;\varphi_1,\dots,\varphi_n;\varepsilon_1,\dots,\varepsilon_n) = \{\xi \in \mathcal{V} : |\langle \varphi_j, \xi - v\rangle| < \varepsilon_j\}$$

where $n \in \mathbb{Z}_+$, $\varphi_1,\dots,\varphi_n \in \mathcal{X}$, and $\varepsilon_1,\dots,\varepsilon_n \in \mathbb{R}_{>0}$.

Problem 9.1. Let $\mathcal{X}$ and $\mathcal{Y}$ be *distinct* linear subspaces of $\mathcal{V}^*$.

1. Prove that $\sigma(\mathcal{V}, \mathcal{X}) \neq \sigma(\mathcal{V}, \mathcal{Y})$.

2. Conclude that if $\mathcal{U}$ is a proper dense linear subspace of $\mathcal{V}$, then the weak-* topologies on $\mathcal{U}^*$ and $\mathcal{V}^*$ do not coincide (even though the dual spaces $\mathcal{U}^*$ and $\mathcal{V}^*$ are equal by Cor. 2.4.3).

Hint for Part 1. Assume $\psi \in \mathcal{Y}\backslash\mathcal{X}$. Apply Thm. 9.2.9 to show that $U(0;\psi;1)$ does not contain any neighborhood of the form $U(0;\varphi_\bullet;\varepsilon_\bullet)$ where each $\varphi_j \in \mathcal{X}$. □

Problem 9.2. Let $\mathcal{X}$ be a linear subspace of $\mathcal{V}^*$. Let $K \subset \mathcal{X}$ be an algebraic basis (also called a **Hamel basis**) of $\mathcal{X}$, that is, K is linearly independent and $\mathcal{X} = \mathrm{Span}K$. Prove the equivalence of the following statements.

(1) K is a countable set.

(2) $\sigma(\mathcal{V}, \mathcal{X})$ is first-countable.

Moreover, if $\mathcal{X}$ separates points of $\mathcal{V}$, prove that these statements are equivalent to:

(3) $\sigma(\mathcal{V}, \mathcal{X})$ is metrizable.

Hint. (1)⇒(2), (3): Show that $\sigma(\mathcal{V}, K) = \sigma(\mathcal{V}, \mathcal{X})$. Then apply Prop. 1.4.16 and its first-countable analogue.

(2)⇒(1): Show that $\sigma(\mathcal{V}, \mathcal{X})$ has a countable neighborhood basis at 0 consisting of sets of the form $U(0; \varphi_\bullet; \varepsilon_\bullet)$ with $\varphi_j \in K$. Use Thm. 9.2.9 to show that $\mathcal{X}$ is spanned by all such φ_j. □

Remark 9.8.4. Suppose that $\mathcal{V}$ is an infinite-dimensional Banach space. As we shall see in Cor. 10.3.12, any algebraic basis of $\mathcal{V}$ is uncountable. Combined with Pb. 9.2, this implies that the weak-* topology on the whole dual space $\mathcal{V}^*$ is not first-countable. In contrast, if $\mathcal{V}$ is separable, then the weak-* topology on any bounded subset of $\mathcal{V}^*$ is metrizable (cf. Prop. 2.6.6).

9.8.1 Proving density via Hahn-Banach

Recall $e_t : \mathbb{R}^N \to \mathbb{C}$ defined by $e_t(x) = e^{\mathbf{i}tx} = e^{\mathbf{i}(t_1x_1+\cdots+t_Nx_N)}$. The following problem generalizes Prop. 7.7.1. We encourage the reader to consider how the proof of this problem is related to the proof of Prop. 7.7.1 (and thus to the proof of Thm. 7.5.4, on which Prop. 7.7.1 depends).

Problem 9.3. Let μ be a finite Borel measure on $\mathbb{R}^N$. Let $1 \leqslant p < +\infty$. Prove that $\mathrm{Span}\{e_t : t \in \mathbb{R}^N\}$ is dense in $L^p(\mathbb{R}^N, \mu)$.

Hint. By Cor. 9.5.7 and Thm. 1.6.16, it suffices to show that any $g \in L^q(\mathbb{R}^N, \mu)$ satisfying $\int g e_t d\mu = 0$ for all t is zero μ-a.e. Consider integrating the function $t \mapsto \int g e_t d\mu$ against some function $f \in C_c(\mathbb{R}^N)$ (or $f \in \mathcal{S}(\mathbb{R}^N)$), and use the properties of the Fourier transform developed in this course. (Thm. 1.7.11 might also be helpful.) □

Problem 9.4. Assume the setting of Thm. 7.7.4. Let $1 \leqslant p < +\infty$. Prove that $\mathbb{C}[x_1, \ldots, x_N]$ is dense in $L^p(\mathbb{R}^N, \mu)$.

Hint. Adapt the proof of Thm. 7.7.4. □

Problem 9.5. Let $\psi : \mathbb{R}^N \to \mathbb{C}$ be a continuous function for which there exist $C, r > 0$ such that

$$0 < |\psi(x)| < Ce^{-r(|x_1|+\cdots+|x_N|)} \qquad \text{for each } x \in \mathbb{R}^N$$

Prove that $\{p\psi : p \in \mathbb{C}[x_1, \ldots, x_N]\}$ is a dense linear subspace of $C_0(\mathbb{R}^N)$ (with respect to the l^∞-norm).

Hint. By Cor. 9.5.7 and Thm. 1.7.17, it suffices to show that any complex Borel measure μ on $\mathbb{R}^N$ satisfying $\int p\psi d\mu = 0$ for all $p \in \mathbb{C}[x_\bullet]$ is zero. Use (for example) the Radon-Nikodym Thm. 1.6.12 to write $d\mu = f d\nu$ where ν is a finite Borel measure on $\mathbb{R}^N$ and $f \in \mathscr{Bor}_{\mathrm{b}}(\mathbb{R}^N)$. Use Thm. 7.7.4 or Pb. 9.4 to show that $f = 0$ ν-a.e. $\square$

9.8.2 Transposes of compact operators

Fix $T \in \mathfrak{L}(\mathcal{V}, \mathcal{W})$. In this subsection, we examine Schauder's complement (cf. [Sch30]) to Riesz's theory of compact operators which we studied in Sec. 8.7 (and also in Pb. 8.8 and 8.9). Schauder's main contribution to what is now called the **Riesz-Schauder theory** was the incorporation of transpose operators into the framework.

Problem 9.6. Prove that $\mathrm{Ker}(T^\sharp) = \mathrm{Rng}(T)^\top$ and hence (by Cor. 9.5.7) $\overline{\mathrm{Rng}(T)} = {}^\perp\mathrm{Ker}(T^\sharp)$.

The following exercise is a variant of Prop. 9.6.1 as well as a generalization of Prop. 8.4.2.

Exercise 9.8.5. Prove that $T^\sharp : \mathcal{W}^* \to \mathcal{V}^*$ is continuous when both $\mathcal{W}^*$ and $\mathcal{V}^*$ are equipped with their weak-* topologies.

The following problem generalizes the equivalence between completely continuous operators and compact operators in the case of Hilbert spaces (cf. Thm. 8.4.7) and, more generally, reflexive spaces (cf. Thm. 9.6.13).

Problem 9.7. Prove that the following statements are equivalent.

(1) The restriction $T^\sharp|_{\overline{B}_{\mathcal{W}^*}(0,1)} : \overline{B}_{\mathcal{W}^*}(0,1) \to \mathcal{V}^*$ is continuous with respect to the weak-* topology on $\overline{B}_{\mathcal{W}^*}(0,1)$ and the norm topology on $\mathcal{V}^*$.

(2) $T^\sharp : \mathcal{W}^* \to \mathcal{V}^*$ is compact.

Problem 9.8. Assume that $T : \mathcal{V} \to \mathcal{W}$ is a compact operator. Prove that T is compact iff $T^\sharp : \mathcal{W}^* \to \mathcal{V}^*$ is compact.

Hint. It suffices assume that T is compact and prove that $T^\sharp$ is compact, since the converse follows by restricting $T^{\sharp\sharp}$ to T.

The strategy is similar to the proof of Cor. 9.6.15. To verify that $T^\sharp$ satisfies condition (1) of Pb. 9.7, choose a net (ψ_α) in $\overline{B}_{\mathcal{W}^*}(0,1)$ converging weak-* to ψ. Suppose that $\lim_\alpha \|T^\sharp\psi_\alpha - T^\sharp\psi\| \neq 0$. By passing to a subnet, assume $\|T^\sharp\psi_\alpha - T^\sharp\psi\| \geqslant \varepsilon$ for all α where $\varepsilon > 0$. Choose a net (ξ_α) in $\overline{B}_{\mathcal{V}}(0,1)$ such that $|\langle T^\sharp\psi_\alpha - T^\sharp\psi, \xi_\alpha\rangle| \geqslant \varepsilon/2$. Pick a convergent subnet of $(T\xi_\alpha)$ (cf. Def. 8.4.4-(2)) and find a contradiction. □

Theorem 9.8.6. *Assume that $T \in \mathfrak{L}(\mathcal{V})$ is a compact operator, $\mathcal{V}$ is a Banach space, and $\lambda \in \mathbb{F}\backslash\{0\}$. Then*

$$\dim \mathrm{Ker}(T-\lambda) = \dim \mathcal{V}/\mathrm{Rng}(T-\lambda) = \dim \mathrm{Ker}(T^\sharp-\lambda) = \dim \mathcal{V}^*/\mathrm{Rng}(T^\sharp-\lambda)$$

where each quantity is finite.

Proof. By Pb. 8.9, the first two quantities are equal and finite. Since $T^\sharp$ is compact (Pb. 9.8), the last two quantities are also equal and finite. By Cor. 9.2.7,

$$\dim \mathcal{V}/{}^\perp\mathrm{Ker}(T^\sharp-\lambda) = \dim \mathrm{Ker}(T^\sharp-\lambda)$$

By Pb. 8.8, $\mathrm{Rng}(T-\lambda)$ is closed. Therefore, by Pb. 9.6,

$$\dim \mathcal{V}/{}^\perp\mathrm{Ker}(T^\sharp-\lambda) = \dim \mathcal{V}/\mathrm{Rng}(T-\lambda)$$

This proves the equality of the second and the third quantities. □

10 Completeness as a domain property

10.1 Introduction

> The applications to Banach space theory are reasonable and powerful, but it seems difficult to obtain them with our bare hands. But we can with our Baire hands.
>
> B. Simon, [Sim-R, Sec. 5.4]

In this course, we have primarily used completeness as a property of the codomain. The role of completeness in this context has already been summarized in Subsec. 2.5.1. In this chapter, we investigate completeness as a property of the domain (or at least not solely of the codomain).

Two closely related classes of theorems are discussed in this chapter:

- The uniform boundedness principle.
- The open mapping theorem, the bounded inverse theorem, and the closed graph theorem (which are mutually equivalent).

Modern textbooks prove these results using the Baire category theorem. Since the Baire category theory—and its proof—differs markedly in flavor from the other methods of functional analysis (and from all techniques we have seen so far in the course), a functional-analytic proof that depends essentially on the theorems above generally cannot be easily replaced by methods that avoid them.

The uniform boundedness principle originated in Lebesgue's construction of continuous functions whose Fourier series diverge at a given point [Leb06, Leb09]. However, despite this origin in Fourier analysis, applications of the above theorems based on the Baire category theorem are typically abstract in nature, starting from general assumptions and yielding very general conclusions. When applied to topics encountered earlier in this course, the above theorems often play at most a psychological role, and even that role is difficult to identify without a retrospective perspective whose historical accuracy is itself questionable.

Despite this somewhat unusual situation, in Sec. 10.2 we will see that at least part of the above theorems and their treatments can, to some extent, be related to the methods and styles used in earlier chapters. To reveal this connection, we will not proceed via the Baire category theorem, but instead use the **gliding hump method**, which Lebesgue originally employed in his construction of continuous functions with divergent Fourier series. This method was also the standard approach to proving the uniform boundedness principle—even in the original approaches of Helly [Hel12] and Banach [Ban22]—until Saks suggested to Banach that it could be proved using the Baire category theorem.

10.2 Gliding hump method and the uniform boundedness principle

Let $\mathcal{V}, \mathcal{W}$ be normed vector spaces over $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$.

Theorem 10.2.1 (Uniform boundedness principle). *Assume that $\mathcal{V}$ is complete. Let $(T_\alpha)_{\alpha\in\mathscr{A}}$ be a family in $\mathfrak{L}(\mathcal{V}, \mathcal{W})$. Assume that*

$$\sup_{\alpha\in\mathscr{A}} \|T_\alpha\| = +\infty$$

Then there exists $\xi \in \mathcal{V}$ such that $\sup_\alpha \|T_\alpha \xi\| = +\infty$.

We defer the proof via the gliding hump method to the end of Subsec. 10.2.2, and the proof via the Baire category theorem to the end of Subsec. 10.3.3.

10.2.1 A toy model

The goal of this section is to explain how Thm. 10.2.1 can be proved using the gliding hump method. To illustrate this method, we first apply it to prove an example that can also be obtained as a consequence of Thm. 10.2.1.

Example 10.2.2. Let (X, μ) be a σ-finite measure space. Let $f : X \to \mathbb{C}$ be measurable. Let $1 < p, q < +\infty$ and $p^{-1} + q^{-1} = 1$. Suppose that $f\xi \in L^1(X, \mu)$ for each $\xi \in L^q(X, \mu)$. Then $\|f\|_{L^p(X,\mu)} < +\infty$.

First proof. In this proof, we apply the uniform bounded principle. Since X is σ-finite, we can write $X = \bigcup_n E_n$ where (E_n) is an increasing sequence of measurable sets and $\|f\chi_{E_n}\|_{L^p} < +\infty$.[1] Define

$$T_n : L^q(X, \mu) \to L^1(X, \mu) \qquad \xi \mapsto f\chi_{E_n} \cdot \xi \tag{10.1}$$

By Thm. 1.6.16, we have

$$\|T_n\| = \|f\chi_{E_n}\|_{L^p}$$

Suppose that $\|f\|_{L^p} = +\infty$. Then MCT implies $\sup_n \|T_n\| = +\infty$. By Thm. 10.2.1, there exists $\xi \in L^q(X, \mu)$ such that

$$\sup_n \|f\chi_{E_n}\xi\|_{L^1} = +\infty$$

By MCT, the LHS above equals $\|f\xi\|_{L^1}$. This gives a contradiction. □

[1] Write $X = \bigcup_n A_n$ where (A_n) is an increasing sequence of measurable sets with finite μ-measures, and let $E_n = A_n \cap \{x \in X : |f(x)| \leqslant n\}$.

Second proof. In this proof, we use the **gliding hump method**. As in the first proof, we write $X = \bigcup_n E_n$ where (E_n) is an increasing sequence of measurable sets and $\|f\chi_{E_n}\|_{L^p} < +\infty$. Suppose $\|f\|_{L^p} = +\infty$. Then

$$\sup_n \int_{E_n} |f|^p = +\infty$$

Fix an increasing $\rho : \mathbb{Z}_+ \to \mathbb{R}_{\geqslant 0}$ whose expression will be determined later. Choose a subsequence of (n_k) of (n) such that

$$\int_{F_k} |f|^p d\mu \geqslant \rho(k)^p \qquad \text{where } F_k = E_{n_{k+1}} \backslash E_{n_k} \tag{10.2}$$

By Thm. 1.6.16, there exists $\xi_k \in L^q(F_k, \mu)$ with non-zero L^q-norm such that $|\int_{F_k} f\xi_k|$ is close to $\|f\chi_{F_k}\|_{L^p} \cdot \|\xi_k\|_{L^q}$.[2] By scaling ξ_k, we obtain

$$\|\xi_k\|_{L^q} = 2^{-k} \qquad \int_{F_k} |f\xi_k| d\mu \geqslant 2^{-k-1} \|f\chi_{F_k}\|_{L^p} \tag{10.3}$$

Then $\xi = \sum_k \xi_k$ belongs to $L^q(X, \mu)$, and

$$\int_X |f\xi| d\mu \geqslant \sum_k 2^{-k-1} \|f\chi_{F_k}\|_{L^p} \geqslant \sum_k 2^{-k-1} \rho(k)$$

which is infinite if we choose $\rho(k) = 2^{k+1}$ at the beginning. □

10.2.2 The gliding hump method

Having seen an example of the gliding hump method, we return to the setting of Thm. 10.2.1 and explain this method in the general case. Thus, we are given an unbounded family (T_α) in $\mathfrak{L}(\mathcal{V}, \mathcal{W})$, where $\mathcal{V}$ is complete. Such a family necessarily contains an unbounded sequence. Therefore, we may assume that (T_α) is a sequence (T_n) with $\sup_n \|T_n\| = +\infty$.

The **gliding hump method** refers to the construction of a sequence (ξ_k) in $\mathcal{V}$ and a subsequence (T_{n_k}) of (T_n) satisfying the following conditions:

(a) $\sum_k \|\xi_k\| < +\infty$.

(b) $\sup_k \|T_{n_k}\xi_k\| = +\infty$.

(c) For each k, $\|T_{n_k}\xi_k\|$ is large compared to $\|T_{n_l}\xi_k\|$ for $l \neq k$.[3]

[2] By choosing ξ_n to be a scalar multiplication of $|f\chi_{F_n}|^{p-1}$, we indeed have $\int_{F_k} |f\xi_k| = \|f\chi_{F_k}\|_{L^p} \cdot \|\xi_k\|_{L^q}$, which is sufficient for the purpose of the proof.

[3] That is, the function $l \mapsto \|T_{n_k}\xi_l\|$ has a "hump" at $l = k$.

Condition (a) ensures that $\xi := \sum_k \xi_k$ converges in $\mathcal{V}$. Condition (c) guarantees that, for each k, the norm $\|T_{n_k}\xi\|$ does not differ significantly from $\|T_{n_k}\xi_k\|$. Together with (b), this implies that $\sup_k \|T_{n_k}\xi\| = +\infty$, completing the proof.

Example 10.2.3. In the proof of Exp. 10.2.2, T_n is given by (10.1), namely, the multiplication operator by $f\chi_{E_n}$. The sequence (ξ_n) is chosen as in (10.3), where each $\xi_k \in L^q(F_k, \mu)$. Condition (a) is then clearly satisfied. Moreover, since (n_k) is chosen to satisfy (10.2), we have

$$\|T_{n_k}\xi_k\| \geqslant 2^{-k-1}\rho(k)$$
$$\|T_{n_k}\xi_l\| = 0 \qquad \text{if } l > k$$

This verifies (b) (for a suitable choice of ρ) and one half of (c).

In the second proof of Exp. 10.2.2, the remaining half of condition (c)—namely, that $\|T_{n_k}\xi_k\|$ is large compared to $\|T_{n_l}\xi_k\|$ for $l < k$— is not needed, since the elements of (ξ_k) are mutually "orthogonal" (in the sense that their norms satisfy a Pythagorean-type identity). Moreover, this orthogonality is preserved under multiplication by functions. As a result, we have $\|T_{n_k}\xi_k\| \leqslant \|T_{n_k}\xi\|$. In the general case, however, such orthogonality is unavailable. Consequently, as we will see in the following proof, the full strength of condition (c) is required. □

Proof of Thm. 10.2.1 via the gliding hump method. Assume WLOG that (T_α) is a sequence $(T_n)_{n\in\mathbb{N}}$ in $\mathfrak{L}(\mathcal{V}, \mathcal{W})$ with $\sup_n \|T_n\| = +\infty$. We want to construct a subsequence $(T_{n_k})_{k\in\mathbb{N}}$ and a sequence (ξ_k) in $\mathcal{V}$ satisfying the following conditions:[4]

(1) $\|\xi_k\| \leqslant 2^{-k}$. (This corresponds to the previous condition (a).)

(2) $\|T_{n_k}\xi_l\| \leqslant 3^{-l}$ if $l > k$. (This corresponds to one half of condition (c).)

(3) $\|T_{n_k}\xi_k\| \geqslant k + \sum_{l<k} \|T_{n_k}\xi_l\|$. (This corresponds to condition (b) and the other half of condition (c).)

If such a construction is possible, the completeness of $\mathcal{V}$ implies that the series $\sum_k \xi_k$ converges to some $\xi \in \mathcal{V}$. Moreover, for each k,

$$\|T_{n_k}\xi\| \geqslant \|T_{n_k}\xi_k\| - \sum_{l<k}\|T_{n_k}\xi_l\| - \sum_{l>k}\|T_{n_k}\xi_l\| \geqslant k - \sum_{l>0} 3^{-l} = k - \frac{1}{2}$$

Hence $\sup_k \|T_{n_k}\xi\| = +\infty$, completing the proof.

We now carry out the construction inductively. Set $\xi_0 = 0$. Suppose $\xi_0, \dots, \xi_{k-1}$ and $n_0 < n_1 < \cdots < n_{k-1}$ have already been chosen. Choose $r_k > 0$ satisfying

$$r_k \leqslant 2^{-k} \qquad r_k \cdot \sup_{j<k} \|T_{n_j}\| \leqslant 3^{-k}$$

[4]The use of the different bounds 2^{-k} and 3^{-k} in conditions (1) and (2) below is merely to clarify, in the subsequent argument, which part of the construction is responsible for which condition.

Then conditions (1) and (2) are satisfied provided we choose $\xi_k \in \mathcal{V}$ with $\|\xi_k\| \leqslant r_k$ (and subsequently choose $\xi_{k+1}, \xi_{k+2}, \ldots$ so as to satisfy analogous bounds).

To construct ξ_k and T_{n_k} so that (3) also holds, consider the condition

$$\sup_{n\in\mathbb{N}} \|T_n\xi_l\| < +\infty \qquad \text{for each } l < k \tag{10.4}$$

If (10.4) holds, since (T_n) is unbounded, there exists $n_k > n_{k-1}$ such that

$$r_k\|T_{n_k}\| \geqslant 1 + k + \sum_{l<k} \sup_{n\in\mathbb{N}} \|T_n\xi_l\|$$

Thus, there exists $\xi_k \in \mathcal{V}$ with $\|\xi_k\| \leqslant r_k$ such that

$$\|T_{n_k}\xi_k\| \geqslant k + \sum_{l<k} \sup_{n\in\mathbb{N}} \|T_n\xi_l\|$$

and hence (3) holds. If (10.4) fails, then there is $l < k$ such that $\sup_n \|T_n\xi_l\| = +\infty$. In this case, there is no need to complete the inductive construction satisfying (1)–(3), since the proof of Thm. 10.2.1 is already completed by taking $\xi = \xi_l$. □

10.2.3 Continuous functions with divergent Fourier series

As mentioned in Sec. 10.1, Lebesgue's construction of a continuous function with divergent Fourier series marked the origin of the gliding hump method. In this subsection, we explain how his result can be viewed as a special case of the uniform boundedness principle.

Corollary 10.2.4. *Assume that $\mathcal{V}$ is complete. Suppose that $(T_n)_{n\in\mathbb{Z}_+}$ is a sequence in $\mathfrak{L}(\mathcal{V},\mathcal{W})$ converging in SOT. Then $\sup_n \|T_n\| < +\infty$.*

Proof. For each $\xi \in \mathcal{V}$, the sequence $(T_n\xi)$ converges; in particular, it is bounded.[5] Therefore, by Thm. 10.2.1, (T_n) is bounded. □

Example 10.2.5. Choose any $z \in \mathbb{S}^1 \simeq \mathbb{R}/2\pi\mathbb{Z}$. There exists $f \in C(\mathbb{S}^1)$ such that the series $\sum_{n\in\mathbb{Z}} \widehat{f}(n)z^n$ does not converge.

Here, $\widehat{f}(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-\mathbf{i}nt}dt$.

Proof. It suffices to treat the case $z = 1$; the general case follows by a rotation of $\mathbb{S}^1$. For each $n \in \mathbb{N}$, define the Dirichlet kernel

$$D_n(t) = \sum_{k=-n}^{n} e^{\mathbf{i}nt} = \frac{\sin((n+\frac{1}{2})t)}{\sin(\frac{1}{2}t)}$$

[5] In contrast, a convergent net need not be bounded.

Define a linear functional

$$\Lambda_n : C(\mathbb{S}^1) \to \mathbb{C} \qquad f \mapsto \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)D_n(t)dt$$

By Cor. 2.7.5, we have $\|\Lambda_n\| = \|D_n\|_{L^1(\mathbb{S}^1,\frac{\mathfrak{m}}{2\pi})}$. Suppose we can prove $\sup_n \|D_n\|_{L^1} = +\infty$. Then, by Cor. 10.2.4, there exists $f \in C(\mathbb{S}^1)$ such that

$$\sum_{n\in\mathbb{Z}} \widehat{f}(n) \equiv \lim_n \Lambda_n(f)$$

does not converge, finishing the proof.

Since $|\sin(t/2)| \leqslant t/2$, we have

$$\int_0^{\pi} \frac{1}{2}|D_n(t)|dt \geqslant \int_0^{\pi} t^{-1}\big|\sin(n+1/2)t\big|dt = \int_0^{(n+\frac{1}{2})\pi} t^{-1}|\sin t|dt$$

where the RHS diverges as $n \to +\infty$ because $\sum_{k\in\mathbb{Z}_+} 1/k$ diverges. This proves $\sup_n \|D_n\|_{L^1} = +\infty$. □

Remark 10.2.6. Cor. 10.2.4 does not hold in general for nets in place of sequences. To see this, let $\mathcal{V}$ any infinite-dimensional normed vector space, and $\mathcal{W} = \mathbb{F}$. For each $E \in \mathrm{fin}(2^{\mathcal{V}})$, choose a nonzero unit vector $\xi_E \in \mathcal{V}$ not inside $\mathrm{Span}E$. By Hahn-Banach, any (bounded) linear functional on $\xi_E + \mathrm{Span}E$ can be extended to a bounded linear functional on $\mathcal{V}$. Thus, for each $n \in \mathbb{Z}_+$, there exists $\varphi_{E,n} \in \mathcal{V}^*$ such that

$$\varphi_{E,n}|_{\mathrm{Span}E} = 0 \qquad \varphi_{E,n}(\xi_E) = n$$

Then $(\varphi_{E,n})_{(E,n)\in\mathrm{fin}(2^{\mathcal{V}})\times\mathbb{Z}_+}$ is a net in $\mathcal{V}^*$ converging pointwise to 0. However, this net is not bounded; in fact, it is not even eventually bounded (and hence contains no bounded subnet).

10.2.4 The closed graph theorem

By adapting the proofs of Exp. 10.2.2, one can show that if a measurable function $f : X \to \mathbb{C}$ does not belong to $L^\infty(X,\mu)$, then there exists $\xi \in L^2(X,\mu)$ such that $f\xi \notin L^2(X,\mu)$. This observation leads directly to the following complement to von Neumann's spectral theory of unbounded operators, which we prove using the gliding hump method and whose proof via the uniform boundedness principle is left to the reader.

Theorem 10.2.7 (Closed graph theorem). *Let $\mathcal{H},\mathcal{K}$ be Hilbert spaces, and let $T : \mathcal{H} \to \mathcal{K}$ be an unbounded closed operator. Suppose that $\mathscr{D}(T) = \mathcal{H}$. Then $T \in \mathfrak{L}(\mathcal{H},\mathcal{K})$.*

Proof. By the polar decomposition Thm. 7.1.11, it suffices to treat the case where $\mathcal{K} = \mathcal{H}$ and T is a positive closed operator on $\mathcal{H}$. (In particular, $\mathrm{Sp}(T) \subset \mathbb{R}_{\geqslant 0}$; cf. Prop. 7.1.2.) Therefore, by the spectral theory (Thm. 6.9.4 and 6.10.10), we may assume

$$\mathcal{H} = \bigoplus_{\alpha\in\mathscr{A}} L^2(\mathbb{R}_{\geqslant 0}, \mu_\alpha) \qquad T = \mathbf{M}_x$$

where each μ_α is a finite Borel measure on $\mathbb{R}_{\geqslant 0}$, and x is the coordinate function of $\mathbb{R}_{\geqslant 0}$. If we can prove the existence of some $R > 0$ such that $\mathrm{Supp}(\mu_\alpha) \subset \overline{B}_{\mathbb{R}_{\geqslant 0}}(0, R)$ for all α, then $T \in \mathfrak{L}(\mathcal{H})$ and $\|T\| \leqslant R$, finishing the proof.

Suppose that this is not the case. Fix an increasing $\rho : \mathbb{Z}_+ \to \mathbb{R}_{\geqslant 0}$ whose expression will be determined later. Then there exists a subsequence (n_k) of (n) satisfying the following conditions for each k:

- $n_k \geqslant \rho(k)$;
- there exists $\alpha_k \in \mathscr{A}$ such that $\mu_{\alpha_k}([n_k, n_{k+1})) > 0$.

Let ξ_k be a non-zero constant function on $[n_k, n_{k+1})$. By scaling ξ_k, we have

$$\|\xi_k\|_{L^2(\mu_{\alpha_k})} = 2^{-k} \qquad \|x\xi_k\|_{L^2(\mu_{\alpha_k})} \geqslant 2^{-k}\rho(k)$$

We view ξ_k and $x\xi_k$ as elements in $L^2(\mathbb{R}_{\geqslant 0}, \mu_{\alpha_k})$, and hence in $\mathcal{H}$. Then both (ξ_k) and $(x\xi_k)$ are orthogonal sequences in $\mathcal{H}$. Moreover, $\sum_k \xi_k$ converges to $\xi \in \mathcal{H}$ whose $L^2(\mathbb{R}_{\geqslant 0}, \mu_\alpha)$-component is given by the function

$$\xi(\alpha) = \sum_{\text{all } k \text{ such that } \alpha_k = \alpha} \xi_k$$

Therefore

$$\sum_{\alpha\in\mathscr{A}} \|x\xi\|^2_{L^2(\mu_\alpha)} = \sum_k \|x\xi_k\|^2_{L^2(\mu_{\alpha_k})} \geqslant \sum_k 2^{-2k}\rho(k)^2$$

where the RHS is infinite if we choose $\rho(k) = 2^k$ at the beginning. This proves $\xi \notin \mathscr{D}(\mathbf{M}_x)$, contrary to the assumption $\mathscr{D}(T) = \mathcal{H}$. $\square$

Corollary 10.2.8. *Suppose that $\mathcal{H}$ is a Hilbert space, and $T : \mathcal{H} \to \mathcal{H}$ is a linear map satisfying*

$$\langle \xi | A\eta \rangle = \langle A\xi | \eta \rangle \qquad \text{for each } \xi, \eta \in \mathcal{H}$$

Then $T \in \mathfrak{L}(\mathcal{H})$.

Proof. By assumption, T is a Hermitian operator, and hence is closable by Thm. 6.4.16. Since $T \subset \overline{T}$, and since $\mathscr{D}(T) = \mathcal{H}$, we must have $T = \overline{T}$. Therefore T is closed, and hence $T \in \mathfrak{L}(\mathcal{H})$ by Thm. 10.2.7. $\square$

The generalization of Thm. 10.2.7 to Banach spaces, which we discuss in Sec. 10.3, is much more difficult to prove using either the uniform boundedness principle or the gliding hump method. This difficulty necessitates the use of the Baire category theorem.

10.3 Baire's category theory and the open mapping / bounded inverse / closed graph theorems

Let $\mathcal{V},\mathcal{W}$ be normed vector spaces over $\mathbb{F}\in\{\mathbb{R},\mathbb{C}\}$. In this section, we introduce the second group of theorems mentioned in Sec. 10.1.

10.3.1 Bounded linear open maps

This is not the first time we have encountered non-injective open maps. The following proposition, together with Pb. 8.8, shows that if $T\in\mathfrak{L}(\mathcal{V})$ is compact and $\lambda\in\mathbb{F}\backslash\{0\}$, then, regardless of whether $T-\lambda$ is injective, the restriction $T-\lambda:\mathcal{V}\rightarrow\mathrm{Rng}(T-\lambda)$ is open.

Proposition 10.3.1. *Let $T\in\mathfrak{L}(\mathcal{V},\mathcal{W})$. Then the following are equivalent.*

*(1) The restriction $T:\mathcal{V}\rightarrow\mathrm{Rng}(T)$ is an **open map**, that is, if $\Omega\subset\mathcal{V}$ is open then $T(U)$ is open in $\mathrm{Rng}(T)$.*

(2) $0_{\mathcal{W}}$ is an interior point of $TB_{\mathcal{V}}(0,1)$ in $\mathrm{Rng}(T)$.

(3) There exists $\delta>0$ such that

$$\|T\xi\|\geqslant\delta\|\xi+\mathrm{Ker}(T)\|\qquad\text{for each }\xi\in\mathcal{V}\tag{10.5}$$

Recall from Prop. 9.2.3 that $\|\xi+\mathrm{Ker}(T)\|=\mathrm{dist}(\xi,\mathrm{Ker}(T))$ is the norm of the coset $\xi+\mathrm{Ker}(T)$ in $\mathcal{V}/\mathrm{Ker}(T)$.

Proof. Replacing $\mathcal{W}$ with $\mathrm{Rng}(T)$, we assume that T is surjective.

(1)⇒(2): This is obvious.

(2)⇒(1): Let $\Omega\subset\mathcal{V}$ be open. If $\xi\in\Omega$, let $\varepsilon>0$ such that $B_{\mathcal{V}}(\xi,\varepsilon)\subset\Omega$. By (2), there exists $\delta>0$ such that $TB_{\mathcal{V}}(0,\varepsilon)\supset B_{\mathcal{W}}(0,\delta)$, and hence

$$T\Omega\supset T\xi+TB_{\mathcal{V}}(0,\varepsilon)\supset T\xi+B_{\mathcal{W}}(0,\delta)$$

This proves that $T\xi$ is an interior point of $T\Omega$ in $\mathcal{W}$.

(2)⇒(3): Assume (2). Then there exists $\delta>0$ such that $B_{\mathcal{W}}(0,2\delta)\subset TB_{\mathcal{V}}(0,1)$, and hence $\overline{B}_{\mathcal{W}}(0,\delta)\subset T\overline{B}_{\mathcal{V}}(0,1)$. By linearity, for each $r\geqslant 0$ we have

$$\overline{B}_{\mathcal{W}}(0,r)\subset T\overline{B}_{\mathcal{V}}(0,\delta^{-1}r)$$

For each $\xi\in\mathcal{V}$, by letting $r=\|T\xi\|$, we have $T\xi\in\overline{B}_{\mathcal{W}}(0,r)$, and hence there exists $\widetilde{\xi}\in\overline{B}_{\mathcal{V}}(0,\delta^{-1}r)$ such that $T\xi=T\widetilde{\xi}$. Since $\xi-\widetilde{\xi}\in\mathrm{Ker}(T)$, we have

$$\|\xi+\mathrm{Ker}(T)\|\leqslant\|\widetilde{\xi}\|\leqslant\delta^{-1}r=\delta^{-1}\|T\xi\|$$

This proves (3).

(3)⇒(2): Assume (3). Then for each $\xi \in \mathcal{V}$ such that $\|T\xi\| < 1$, we have $\|\xi + \mathrm{Ker}(T)\| < \delta^{-1}$, and hence there exists $\widetilde{\xi} \in \xi + \mathrm{Ker}T$ such that $\|\widetilde{\xi}\| < \delta^{-1}$. Hence

$$T\xi = T\widetilde{\xi} \in TB_{\mathcal{V}}(0, \delta^{-1})$$

This proves $B_{\mathcal{W}}(0,1) \subset TB_{\mathcal{V}}(0, \delta^{-1})$. Thus (2) holds. □

The following property generalizes Lem. 5.10.21.

Proposition 10.3.2. *Let $T \in \mathfrak{L}(\mathcal{V}, \mathcal{W})$ be open. Assume that $\mathcal{V}$ is complete. Then $\mathrm{Rng}(T)$ is complete; in particular, it is closed in $\mathcal{W}$.*

Proof. By Lem. 8.10.1, it suffices to show that for each sequence (v_n) in $\mathcal{V}$ satisfying $\sum_n \|Tv_n\| < +\infty$, the series $\sum_n Tv_n$ converges. Since T is open, we may assume that T satisfies Condition (3) of Prop. 10.3.1. Thus, by adding a vector in $\mathrm{Ker}(T)$ to v_n, we assume that $\|Tv_n\| \geqslant \delta \|v_n\|$. Thus $\sum_n \|v_n\| < +\infty$, and hence $\sum_n v_n$ converges. By the continuity of T, the series $\sum_n Tv_n$ converges. □

Corollary 10.3.3. *Let $\mathcal{U}$ be a closed linear subspace of $\mathcal{V}$. Then the quotient map $\pi_{\mathcal{U}} : \mathcal{V} \to \mathcal{V}/\mathcal{U}$ is open. Moreover, if $\mathcal{V}$ is complete, then $\mathcal{V}/\mathcal{U}$ is complete.*

Proof. The surjective map $\pi_{\mathcal{U}}$ clearly satisfies Condition (3) of Prop. 10.3.1, and hence is open. If $\mathcal{U}$ is complete, the completeness of $\mathcal{V}/\mathcal{U}$ follows from Prop. 10.3.2. □

The following corollary generalizes Lem. 9.3.4.

Corollary 10.3.4. *Suppose that $T \in \mathfrak{L}(\mathcal{V}, \mathcal{W})$ is surjective and $\dim \mathcal{W} < +\infty$. Then T is open.*

Proof. Let $S : \mathcal{V}/\mathcal{U} \to \mathcal{W}$ be the unique linear map such that $S \circ \pi_{\mathcal{U}} = T$. Since S is a bijective linear map, by Thm. 8.4.14, S is a homeomorphism. By Cor. 10.3.3, $\pi_{\mathcal{U}}$ is open. Thus S is open. □

10.3.2 The open mapping theorem, the bounded inverse theorem, and the closed graph theorem

The open mapping theorem provides a converse to Cor. 10.3.2.

Theorem 10.3.5 (Open mapping theorem). *Let $T \in \mathfrak{L}(\mathcal{V}, \mathcal{W})$. Assume that $\mathcal{V}$ and $\mathrm{Rng}(T)$ are complete. Then the restriction map $T : \mathcal{V} \to \mathrm{Rng}(T)$ is open.*

Theorem 10.3.6 (Bounded inverse theorem). *Suppose that $T \in \mathfrak{L}(\mathcal{V}, \mathcal{W})$ is bijective, and that both $\mathcal{V}$ and $\mathcal{W}$ are complete. Then $T^{-1} \in \mathfrak{L}(\mathcal{W}, \mathcal{V})$.*

Definition 10.3.7. Equip the vector space $\mathcal{V} \oplus \mathcal{W}$ (cf. Def. 3.3.12 for the vector space structure) with the norm

$$\|v \oplus w\| = \sqrt{\|v\|^2 + \|w\|^2}$$

For each linear map $T : \mathcal{V} \to \mathcal{W}$, the **graph** $\mathscr{G}(T)$ is the linear subspace of $\mathcal{V} \oplus \mathcal{W}$ defined by

$$\mathscr{G}(T) = \{\xi \oplus T\xi : \xi \in \mathcal{V}\}$$

Theorem 10.3.8 (Closed graph theorem). *Let $T : \mathcal{V} \to \mathcal{W}$ be a linear map. Assume that $\mathcal{V}, \mathcal{W}$ are complete, and that T is closed in the sense that $\mathscr{G}(T)$ is closed in $\mathcal{V} \oplus \mathcal{W}$. Then $T \in \mathfrak{L}(\mathcal{V}, \mathcal{W})$.*

The proofs of these three theorems are deferred to Subsec. 10.3.4.

Remark 10.3.9. The above three theorems are equivalent. Therefore, the closed graph Thm. 10.2.7 for Hilbert spaces proved earlier already implies both the open mapping theorem and the bounded inverse theorem for Hilbert spaces.

Proof. Thm. 10.3.5 $\Rightarrow$ Thm. 10.3.6: This is obvious.

Thm. 10.3.6 $\Rightarrow$ Thm. 10.3.5: Suppose that T satisfies the assumption in the open mapping theorem. Replacing $\mathcal{W}$ with $\mathrm{Rng}(T)$, we assume that T is surjective and $\mathcal{W}$ is complete. Let $\mathcal{U} = \mathrm{Ker}(T)$, and let $S : \mathcal{V}/\mathcal{U} \to \mathcal{W}$ be the unique (automatically bijective) linear map such that $T = S \circ \pi_{\mathcal{U}}$. By Thm. 9.2.4, we have $S \in \mathfrak{L}(\mathcal{V}/\mathcal{U}, \mathcal{W})$. By Cor. 10.3.3, $\mathcal{V}/\mathcal{U}$ is complete. Therefore, by the bounded inverse theorem, S is a homeomorphism. Since $\pi_{\mathcal{U}}$ is open (Cor. 10.3.3), it follows that T is open.

Thm. 10.3.6 $\Rightarrow$ Thm. 10.3.8: Assume that T satisfies the assumption in the closed graph theorem. Since $\mathscr{G}(T)$ is a closed linear subspace of $\mathcal{V} \oplus \mathcal{W}$, it is complete. Therefore, by the bounded inverse theorem, the inverse of the graph projection

$$\Pi_T : \mathscr{G}(T) \to \mathcal{V} \qquad \xi \oplus T\xi \mapsto \xi$$

is a bounded linear map $\Pi_T^{-1} : \mathcal{V} \to \mathscr{G}(T)$. Since T equals the composition of $\Pi_T^{-1} : \mathcal{V} \to \mathscr{G}(T)$ and the projection $\mathscr{G}(T) \subset \mathcal{V} \oplus \mathcal{W} \to \mathcal{W}$ onto the $\mathcal{W}$-component, it follows that T is bounded.

Thm. 10.3.8 $\Rightarrow$ Thm. 10.3.6: Assume that T satisfies the assumption in the bounded inverse theorem. As in Prop. 6.4.8, one shows easily that T is closed. Thus T^{-1} is closed, because $\mathscr{G}(T^{-1})$ is the diagonal reflection of $\mathscr{G}(T)$. It follows from the closed graph theorem that T^{-1} is bounded. $\square$

Before proving the above three theorems, we briefly discuss possible applications of the open mapping theorem. At first glance, it may seem that when

$\mathcal{V}, \mathcal{W}$ are Banach spaces, one can use the theorem to establish the openness of $T \in \mathfrak{L}(\mathcal{V}, \mathcal{W})$ (or even the boundedness of T^{-1}) once it is known that T is surjective (or bijective). In practice, however, the open mapping theorem should more often be interpreted in the opposite direction: proving that T is surjective is typically no easier than proving that T is open (or that T^{-1} is bounded).

For example, as seen in the proof of the Fredholm alternative (Thm. 8.7.2), if $\mathcal{V}$ is a Banach space and $T \in \mathfrak{L}(\mathcal{V})$ is compact, and if $\lambda \in \mathbb{F}\backslash\{0\}$ satisfies $\mathrm{Ker}(T - \lambda) = 0$, then the method used to prove that $T - \lambda$ is surjective already yields the boundedness of $(T - \lambda)^{-1}$.

10.3.3 The Baire category theorem

As mentioned at the end of Subsec. 10.2.4, to establish the closed graph theorem (and hence the other two theorems in Subsec. 10.3.2), we need the Baire category theorem, which we introduce in the present subsection.

Note that if X is a topological space and $E \subset X$, then $\mathrm{Int}_X(E) = \emptyset$ iff $U \subsetneq E$ for each non-empty open set $U \subset X$.

Theorem 10.3.10 (Baire category theorem). *Let X be a complete metric space. Let $(E_n)_{n \in \mathbb{Z}_+}$ be a sequence of closed subsets of X such that $\mathrm{Int}_X(E_n) = \emptyset$ for each n. Then $\mathrm{Int}_X\big(\bigcup_n E_n\big) = \emptyset$. In particular, $\bigcup_n E_n \subsetneq X$.*

Proof. Let us prove $\mathrm{Int}_X\big(\bigcup_n E_n\big) = \emptyset$ by proving that any non-empty open set $U \subset X$ contains a point outside each E_n.

Let us show that there is a decreasing sequence of closed balls (B_n) in X satisfying the following conditions for each n.

- $B_n \subset U \backslash E_n$.
- The radius r_n of B_n satisfies $0 < r_n \leqslant 2^{-n}$.

Indeed, since $U \backslash E_1$ is a nonempty open subset of U, we simply choose B_1 to be any closed ball in $U \backslash E_1$ with radius $0 < r_1 \leqslant 2^{-1}$. Suppose $B_1, \dots, B_{n-1}$ have been constructed. Since $\mathrm{Int} B_{n-1} \backslash E_n$ is a nonempty open subset of $\mathrm{Int} B_{n-1}$, it contains a closed ball with radius $0 < r \leqslant 2^{-n}$. This finishes the inductive construction.

Choose any $x_n \in B_n$. Then (x_n) is a Cauchy sequence in X, and hence converges to some $x \in X$. For each $n \in \mathbb{Z}_+$, since $x_k \in B_n$ for each $k \geqslant n$, we have $x \in B_n$, and hence $x \in U \backslash E_n$. Thus $x \in U \backslash \bigcup_n E_n$. $\square$

When (E_n) is a finite sequence, a similar result holds without assuming the completeness:

Proposition 10.3.11. *Let X be a topological space and A, B are closed subsets of X satisfying $\mathrm{Int}_X(A) = \mathrm{Int}_X(B) = \emptyset$. Then $\mathrm{Int}_X(A \cup B) = \emptyset$.*

Proof. Let $U \subset X$ be open and non-empty. Since $\mathrm{Int}(A) = \emptyset$, the set $U \backslash A$ is (open and) nonempty. Since $\mathrm{Int}(B) = \emptyset$, the set $(U \backslash A) \backslash B = U \backslash (A \cup B)$ is nonempty. Thus U is not contained in $A \cup B$. □

Corollary 10.3.12. *Assume that $\mathcal{V}$ is complete and infinite-dimensional. Then any countable subset of $\mathcal{V}$ cannot span $\mathcal{V}$.*

First proof. Suppose that $(\xi_n)_{n\in\mathbb{Z}_+}$ is a sequence in $\mathcal{V}$. Let $E_n = \mathrm{Span}\{\xi_1, \dots, \xi_n\}$. Then E_n is a proper linear subspace of $\mathcal{V}$, and hence has empty interior (otherwise, by linearity, E_n equals $\mathcal{V}$). By Thm. 8.4.14, E_n is closed. Therefore, by Thm. 10.3.10, $\bigcup_n E_n \neq \mathcal{V}$. □

Second proof. Suppose that $(\xi_n)_{n\in\mathbb{Z}_+}$ is a linearly-independent sequence in $\mathcal{V}$ spanning $\mathcal{V}$. By Hahn-Banach, there exists a sequence (φ_n) in $\mathcal{V}^*$ such that

$$\varphi_n(\xi_1) = \cdots = \varphi_n(\xi_n) = 0 \qquad \varphi_n(\xi_{n+1}) = n\|\xi_{n+1}\|$$

Then (φ_n) converges weak-* to 0 and $\sup_n \|\varphi_n\| = +\infty$, contradicting the uniform boundedness principle (Thm. 10.2.1). □

We close this subsection with a proof of the uniform boundedness principle via the Baire category theorem.

Proof of Thm. 10.2.1 via the Baire category theorem. For each $n \in \mathbb{Z}_+$, the set

$$E_n = \big\{\xi \in \mathcal{V} : \sup_\alpha \|T_\alpha \xi\| \leqslant n\big\} = \bigcap_\alpha \{\xi \in \mathcal{V} : \|T_\alpha \xi\| \leqslant n\}$$

is closed in $\mathcal{V}$. Moreover, since T_α is not uniformly bounded on $B_\mathcal{V}(0,1)$, by linearity, it is not uniformly bounded on any non-empty open ball. Thus E_n has empty interior. By Thm. 10.3.10, there exists $\xi \in \mathcal{V}$ outside every E_n. Then $\sup_\alpha \|T_\alpha \xi\| = +\infty$. □

10.3.4 Proof of the main theorems

In this subsection, we prove the open mapping Thm. 10.3.5. Then, as noted in Rem. 10.3.9, the bounded inverse Thm. 10.3.6 and the closed graph Thm. 10.3.8 follow immediately.

Proof of the open mapping Thm. 10.3.5. Replacing $\mathcal{W}$ with $\overline{\mathrm{Rng}(T)}$, we assume that T is surjective and $\mathcal{W}$ is complete. Since $\mathcal{W} = \bigcup_{n\in\mathbb{Z}_+} \overline{T(B_\mathcal{V}(0,n))}$, by the Baire category Thm. 10.3.10 and the completeness of $\mathcal{W}$, there exists n such that $\overline{T(B_\mathcal{V}(0,n))}$ contains a non-empty open subset of $\mathcal{W}$. Since that open set intersects $\overline{T(B_\mathcal{V}(0,n))}$, it intersects $T(B_\mathcal{V}(0,n))$ at some point Tv (where $v \in B_\mathcal{V}(0,n)$). Thus Tv is an interior point of $\overline{T(B_\mathcal{V}(0,n))}$. So 0 is an interior point of $\overline{T(B_\mathcal{V}(0,n))} - Tv$, and hence an interior point of $\overline{T(B_{\mathcal{V}}(0,2n))}$. By linearity, we conclude that 0 is an interior point of $\overline{T(B_{\mathcal{V}}(0,1))}$.

Choose $\delta > 0$ such that $B_{\mathcal{W}}(0,\delta) \subset \overline{T(B_{\mathcal{V}}(0,1))}$ and hence

$$B_{\mathcal{W}}(0,r) \subset \overline{T(B_{\mathcal{V}}(0,\delta^{-1}r))} \qquad \text{for each } r \geqslant 0$$

by linearity. This implies the following statement:

- For each $\varepsilon > 0$ and $\psi \in \mathcal{W}$, there exists $\xi \in \mathcal{V}$ satisfying $\|\xi\| \leqslant \delta^{-1}\|\psi\|$ and $\|\psi - T\xi\| \leqslant \varepsilon$.

We shall use this observation to prove

$$B_{\mathcal{W}}(0,1) \subset T(B_{\mathcal{V}}(0,3\delta^{-1})) \tag{10.6}$$

Then the openness of T follows from Prop. 10.3.1.

Fix $\psi \in B_{\mathcal{W}}(0,1)$. We claim that there exists a sequence $(\xi_n)_{n\in\mathbb{Z}_+}$ in $\mathcal{V}$ satisfying

$$\delta\|\xi_{n+1}\| \leqslant \|\psi - T(\xi_1 + \cdots + \xi_n)\| \leqslant 2^{-n}$$

for each $n \in \mathbb{N}$. To see this, choose any $\xi_1 \in \mathcal{V}$ satisfying $\|\xi_1\| \leqslant \delta^{-1}\|\psi\|$ and $\|\psi - T\xi_1\| \leqslant 2^{-1}$. Suppose that $\xi_1,\dots,\xi_n$ have been constructed, we pick $\xi_{n+1} \in \mathcal{V}$ satisfying

$$\|\xi_{n+1}\| \leqslant \delta^{-1}\|\psi_n\| \qquad \|\psi_n - T\xi_{n+1}\| \leqslant \varepsilon$$

where $\psi_n = \psi - T(\xi_1 + \cdots + \xi_n)$. This finishes the inductive construction of the sequence (ξ_n).

Since $\sum_n \|\xi_n\| \leqslant 2\delta^{-1}$, by the completeness of $\mathcal{V}$, the series $\sum_n \xi_n$ converges to some $\xi \in \mathcal{V}$ satisfying $\|\xi\| \leqslant 2\delta^{-1}$. By the continuity of T, we have $\psi = T\xi$. This proves (10.6). □

A 依赖于选择公理的证明是否值得信任?

A.1

在本文中，我将从数学研究、数学教育、以及数学实践的立场（而非数学基础和数理逻辑的立场）出发，讨论如下问题：如果一个分析学定理的证明依赖于选择公理（或者它的等价形式：Zorn 引理），我们应当在何种程度上信任这个证明?——随着文本的展开，读者会意识到：本文的真正意图并不在于讨论选择公理本身，而是以选择公理为契机来谈论数学研究、数学教育、和数学实践。[1]

我想，大多数的数学学习者和数学科研人员接受选择公理，其最主要的原因便是：目前的主流学术圈承认选择公理，并且积极地允许选择公理渗透到成熟数学理论中的许多重要结论的证明当中。作为后果，数学工作者无论是出于人情（希望被学术共同体接纳）还是出于理论（自己通过长期科班教育所建立起来的数学知识体系，会随着拒斥选择公理而被釜底抽薪），都离不开选择公理。实际上，"出于理论"其实也是出于学术共同体，因为我们的数学科班教育是通过学术共同体里的人完成的。

这样看来，数学工作者接受选择公理的一个主要理由便是：他们和学术共同体之间的关系和纽带——或者说得更加冒犯一点，来自学术共同体的权威。我想，没有多少数学工作者认真钻研过选择公理背后的数理逻辑理由。实际上，数学的科班教育，本身是不鼓励学生在数学基础的层面花费太多精力的，因为其在实际科研上的回报率非常低。而"选择公理的表述本身看起来很自然"这一点也远远不是数学工作者接受选择公理的理由，因为许多数学证明诉诸的是 Zorn 引理这一选择公理的等价命题，而非选择公理本身——但 Zorn 引理却非常缺乏直观，更遑论由选择公理推出 Zorn 引理的证明。因为走通了选择公理推出 Zorn 引理的证明逻辑链，便相信选择公理/Zorn 引理在数学证明中的合法性——这不过是尊崇"逻辑链"的权威。众所周知，走通逻辑链距离真正理解一个证明相差非常遥远；而一般数学家不从事数理逻辑研究，从而不和数理逻辑学家这一共同体享有相同的学术取向和学术旨趣，又如何能真正理解这一证明背后的渊源和直觉呢?

既然学术共同体的传统和权威是让大部分数学工作者接受选择公理的原因，那我们便有了理由去拷问：是否我们应当满足于传统和权威? 首先我应当亮明我的拷问策略——它和本讲义中处理泛函分析中的若干概念和证明方法的拷问是源于相同的旨趣。

如读者所见，我们在本讲义中拷问过诸如紧算子、闭算子和算子闭包、无界自伴算子等一系列概念的合法性，我们也探讨过线性映射的几何视角、Cauchy 完备性、Riesz 表示定理作为谱定理证明之一环等等许多方法的合法性。特别地，我们展示了这一系列概念和方法在历史上的竞争对手——通过合适的呈现方式，这些竞争对手在现代的视角下仍然充满说服力。正是这种说服力，才有希望让读者明

[1]因此，本文并不遵从数学基础、集合论等相关主题的通常写作范式。希望读者不要因为本文标题而有所误解。

白：上述对概念和方法的拷问并不是没事找事和纯思辨性争论，也不是单纯地为反抗学术共同体权威而反抗——权威这个词本身是中性的，彻底地反抗权威只会让人们失去思考的立足点。

因此，我接下来打算回顾一下选择公理在本讲义所呈现的诸多结论和证明中是如何被使用的。我希望这种呈现所带来的不安感也能让读者明白：操心选择公理的合法性并非是纯思辨或者没事找事。

A.2

我将本讲义中使用选择公理的方式分为三类：

1. 这一类证明直接运用的是选择公理（而非 Zorn 引理）；然而，通过简单的改动，我们便能给出不依赖于选择公理的证明。

2. 这一类证明所探讨的对象是不假设可数性的（例如可分性、第二可数性、可度量性等等），对于这一类对象的证明必须依赖 Zorn 引理。然而，这一类对象的最重要的子类是具有可数性质的；对于这个子类，其证明可以规避 Zorn 引理（从而规避选择公理）。

3. 这一类证明所探讨的对象看起来具有可数性，但证明实际上依赖于选择公理或 Zorn 引理，因此也依赖于某些不具有可数性质的拓扑结构。并且，该类对象没有任何非平凡的子对象，其证明可以规避选择公理或 Zorn引理。

许多利用网来处理的点集拓扑结论都属于**第一类**。例如在命题1.2.17中，证明若 A 是拓扑空间 X 的子集，且 x 是 A 的闭包里的一个点，则存在 A 中的网 $(x_\alpha)_{\alpha\in\mathscr{I}}$ 收敛到 x. 我们的证明方法是：由假设，x 的每个邻域 U 都和 A 相交。因此，只需把指标集 $\mathscr{I}$ 取成所有 x 的邻域构成的集合 $\mathrm{Nbh}_X(x)$，并且对任何邻域 U **选择**一个 $U\cap A$ 中的元素 x_U，则 $(x_U)_{U\in\mathscr{I}}$ 是 A 中的网、且收敛到 x.

很容易把这一依赖于选择公理的证明改述成不依赖选择公理的形式——只需把指标集 $\mathscr{I}$ 取成

$$\{(U,y)\in\mathrm{Nbh}_X(x)\times A: y\in U\}$$
$$(U_1,y_1)\leqslant(U_2,y_2)\qquad\Longleftrightarrow\qquad U_1\supset U_2$$

并把网 $(x_{U,y})_{(U,y)\in\mathscr{I}}$ 设成 $x_{U,y}=y$ 便能完成证明。之所以我们在正文中给出上一段的证明，是因为该证明和 X是第一可数空间（例如度量空间）时构造 A 中收敛到 x 的序列相比，在形式上完全一致，因此极易为读者接受。

除了命题1.2.17以外，同一小章节中的命题1.2.16和定理1.2.18的证明皆属此类。实际上，这第一类证明遍布于本讲义的各个角落，它们对选择公理的依赖是无关痛痒的——不仅可以轻松地被替换成不依赖于选择公理的证明，也能够在具有可数性质的时候，轻松地被改编成适用于序列的形式。鉴于序列的构造依赖于

归纳法[2]，在第一类证明中，选择公理对归纳法的继承是直截了当的；因此，前者对后者之法理的继承也是直截了当的。

第三类结论的主要例子便是赋范线性空间 $l^1(\mathbb{Z}_+)$ 和 $L^1([0,1],m)$ 的非自反性。尽管这两个空间本身是可分的，但它们的对偶空间 $l^\infty(\mathbb{Z}_+)$ 和 $L^\infty([0,1],m)$ 不可分。非自反的证明依赖于 $l^\infty(\mathbb{Z}_+)$ 和 $L^\infty([0,1],m)$ 上的 Hahn-Banach 定理，而不可分空间的 Hahn-Banach 定理的证明依赖于选择公理。因此，$l^1(\mathbb{Z}_+)$ 和 $L^1([0,1],m)$ 的非自反性是依赖于选择公理的。在选择公理以外的适当的公理系统中，$l^1(\mathbb{Z}_+)$ 是自反的（见 [Väth98]）。

有足够的理由认为：可分的赋范空间才有"实在"的意义，因为有充足的分析学语境的赋范空间都是可分的（见下面对 Banach Alaoglu 定理的讨论）；而不可分的赋范空间更适合当作人造的空间，其引进目的在于简化（以研究可分赋范空间为目标的）理论。因此，$l^1(\mathbb{Z}_+)$ 和 $L^1([0,1],m)$ 的非自反性依赖于某些非常人为的数学构造，即它们的对偶空间的范数拓扑。[3]既然如此，拒绝" $l^1(\mathbb{Z}_+)$ 和 $L^1([0,1],m)$ 是非自反的"这一结论便没有那么不可想象了。

A.3

我们接下来谈论**第二类**，情况最为微妙的类别，其典型例子包括：

- Banach-Alaoglu 定理 2.6.5 的证明，其运用了（不可数乘积空间的）Tychonoff 定理 1.4.17，而后者依赖于 Zorn 引理。然而，当我们假设赋范空间是可分的话，证明只需要用到对角线法则这一可数版本的 Tychonoff 定理——见定理 2.6.5 的第二个证明——从而只依赖于归纳法而非 Zorn 引理或选择公理。

- Arzelà-Ascoli 定理 1.4.37 的证明运用了 Tychonoff 定理；而在空间 X 是可分的情况下只需要对角线法则，从而不依赖于 Zorn 引理或选择公理。

- Hahn-Banach 扩张定理的证明运用了 Zorn 引理；而在赋范空间是可分的时候，只需要归纳法即可。

我们以 Banach-Alaoglu 定理为例，来进行下面的讨论，因为该定理及其历史处于线性泛函分析的核心位置。

[2]如本文开头所说，本文采取的是数学实践而非数学基础的立场。因此，本文遵从许多数学学术写作和教材写作的惯例，不区分严格意义上的数学归纳法、可数选择公理、依赖选择公理之间的运用上的差异。

[3]注意，我并非意在表明 $l^\infty(\mathbb{Z}_+)$ 和 $L^\infty([0,1],m)$ 这两个空间本身是人造的；我所针对的是它们的范数拓扑。实际上，这两个空间的单位闭球的弱*拓扑是第二可数的；见命题 2.6.6。因此，它们的单位闭球上的弱*拓扑是非人为的。

Banach-Alaoglu 2.6.5 定理说的是：若 V 是数域 $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ 赋范线性空间，V^* 是其对偶空间，则 V^* 的单位闭球

$$B = \{\varphi \in V^* : \|\varphi\| \leqslant 1\}$$

在弱*拓扑下是紧的。我们在正文中已然见到的证明方法如下：任取 B 中的一个网 (φ_α); 通过运用 Tychonoff 定理，并且把每一个 φ_α 是做 V 上的函数（把 $v \in V$ 映射到 $\varphi(v)$），可知存在子网 (φ_{α_μ}) 作为 一族 V 上的函数逐点收敛到某个函数 $\varphi : V \to \mathbb{F}$。不难验证 φ 的线性性和有界性，从而得知 $\varphi \in B$. 这证明了 B 中的任意网都有弱*收敛子网；故 B是弱*紧的。

如正文中定理 2.6.5 的第二个证明所示，在额外假设 V 是可分的情况下，我们可以绕过 Tychonoff 定理，运用对**角线法**（见评论 1.4.18）来证明 B 的列紧性。[4] 方法如下：任取 V 的一个可数子集 E，满足 $\mathrm{Span}(E)$ 在V 中稠密。[5] 任取 B 中的序列 (φ_n)。由对角线法，存在子列 (φ_{n_k})，其作为 V 上的函数列限制在 E 上时是逐点收敛的，从而在$\mathrm{Span}E$上是逐点收敛的。因此，由一致有界性

$$\sup_n \|\varphi_n\| < +\infty \tag{A.1}$$

可知，(φ_{n_k}) 在整个空间 V 上逐点收敛（即弱*收敛）到某个有界线性泛函 $\varphi \in V^*$；见定理 2.4.6。

由如上的比较可知，相比于针对可分赋范空间的对角线法证明，针对一般赋范空间的 Tychonoff 定理证法少了重要一步：我们不需要定理 2.4.6来保证稠密张成的子集 E 上的收敛可以推出整个空间 V 的收敛；Tychonoff 定理（以及它背后的 Zorn 引理）一开始便能提供给我们整个空间上的收敛，而不需要像对角线法那样，以可数的稠密张成子集作为跳板。

到此，我们对依赖于选择公理的证明的拷问便有了立足点：用选择公理来换取证明中的一个关键的定理 2.4.6（稠密张成子集上的收敛 + 一致有界条件 (A.1) $\Rightarrow$ 整个空间上的收敛）的省略——**这笔交易是否符合道德？**[6]

A.4

接下来，我会给出两个拒绝这笔交易的理由。**第一个理由**是：它会使我们对泛函分析的理解变得贫瘠和破碎。为了解释这一点，我需要回到第 2.2节（它是本讲义真正意义上的开篇）中的一个论点：虽然（在满足一致有界 (A.1) 的情况

[4]不难证明 B 的弱*拓扑是可度量的，因此紧性和列紧性等价。见定理 2.6.5 的第二个证明。在本文中，我们忽略这一差异。

[5]实际上，这里只需把 E 取成 V 的可数稠密子集就行了。然而，为了和接下来的讨论对接，我们只要求 $\mathrm{Span}(E)$ 是稠密的。

[6]这里的"道德"一词当然应该以修辞和比喻的角度理解，但它并不是通常说的学术道德，而是本文接下来要给出的理由。

下）弱*收敛和矩收敛等价，但矩收敛却为弱*收敛提供了经典分析学中的逼近论语境，从而为对偶空间的研究赋予了逼近论内涵。

所谓矩收敛，便是一列向量在给定坐标系下，每个相应分量的收敛。在本文目前的语境下，一个可分赋范空间 V 的一个可数的稠密张成子集 E 便是一个坐标系，而每个线性泛函 $\varphi \in V^*$ 在 E 上的取值便是它在这个坐标系下的坐标值，也就是 φ 的**矩**。照此理解，**矩收敛**便是序列（或网）(φ_n) 限制在 E 上的逐点收敛。因此，弱*收敛和矩收敛的等价性，其实就是上述交易所交换掉的定理 2.4.6 所说的内容。

然而，泛函分析的早期发展以矩收敛而非弱*收敛为主导视角。这是因为矩收敛出现在各种各样的分析问题的语境下：

- 令 $V = L^2([-\pi,\pi],m)$，$E = \{e^{\mathrm{i}nx} : n \in \mathbb{Z}\}$。则坐标系 E 中的矩便是 Fourier级数。因此，一列 L^2 函数的矩收敛意味着其每一个Fourier系数的收敛。研究Fourier系数的收敛性和函数列的逐点收敛之间的联系，这是催生出以控制收敛定理为最大卖点的 Lebesgue 积分理论的最大推手。

- 令 $V = l^2(\mathbb{Z})$，$E = \{\chi_{\{n\}} : n \in \mathbb{Z}\}$。则一列 l^2 函数在坐标系 E 下的矩收敛意味着它作为 $\mathbb{Z}$ 上的函数列的逐点收敛。Hilbert 便是在这个语境下发现了 Hilbert 空间 $l^2(\mathbb{Z})$，意识到了 $l^2(\mathbb{Z})$ 的单位球中的矩收敛可以用来得到全连续二次型的对角化定理（即 Hilbert-Schmidt 定理 8.3.1）。特别地，Hilbert 所引入的“全连续二次型”这一概念的发现语境是非常自然的：它意味着二次型作为 $l^2(\mathbb{Z})$ 中向量的二次函数，当向量列满足一致有界条件 (A.1)，且在 $\mathbb{Z}$ 上逐点收敛（而非 l^2 收敛）时，二次型的取值收敛到相应的值。

- $V = C([a,b])$，$E = \{x^n : n \in \mathbb{N}\}$. 则 V 上的一列正线性泛函在坐标系 E 下的矩（即**多项式矩**）的收敛对应了发散级数的系数收敛，从而也和连分数的收敛方式联系起来。（实际上，紧区间 $[a,b]$ 能够替换成任意有界或无界的闭区间。）这一矩收敛对于 Stieltjes 发现以他本人命名的积分、以及对于之后 Hilbert 发现有界双线性型的谱定理和连续谱的概念，都至关重要。第 4 章和第 5.1 节对此有详细的讨论。

那么，放弃以定理 2.4.6——弱*收敛和矩收敛的等价性——来理解 Banach-Alaoglu 定理，意味着什么呢？这意味着我们错失了 Hilbert 发现 Hilbert 空间 $l^2(\mathbb{Z})$ 的单位闭球的弱/弱*紧性的自然语境。我们也错失了 Hilbert 通过用有限矩阵逼近无限矩阵来发现连续谱的自然语境——一方面，来自于有限矩阵对角化的离散谱**弱*收敛**到连续谱[7]，相应地，有限矩阵的谱分解“收敛到”无限矩阵的谱分解；另一方面，有限矩阵的预解式（resolvent）收敛到无限矩阵的预解式，对应着连分数的收敛过程，即连分数所代表的发散级数（其系数便是多项式矩）的 Páde 收敛。仍见第 4 章和第 5.1 节的讨论。

[7]严格来说：$[a,b]$ 上的有限支集测度（见引理 1.6.5）弱*收敛到有限 Borel 测度/Stieltjes 积分。

放弃以定理 2.4.6 来理解 Banach-Alaoglu 定理，也意味着将 Banach-Alaoglu 定理至于泛函分析的基本哲学之外。这个基本哲学便是：在算子范数（一致）有界的条件下，稠密线性子空间上的性质能够唯一扩张到整个空间上的性质。不仅定理 2.4.6 是这一哲学的一个基本表现形式，该定理证明所依赖的一个基本事实——稠密线性子空间上的有界线性映射能够唯一地扩张为全空间上的有界线性映射（定理 2.4.2）——则更是如此。定理 2.4.2 的运用遍布泛函分析的每个角落，甚至无界算子的闭包概念都是由这一定理经过逆 Cayley 变换的翻译而得到的（见第 6.3.3 小章节的讨论）。因此，放弃让定理 2.4.6 进入 Banach-Alaoglu 定理的证明，便会使得 Banach-Alaoglu 定理的理解成为一座孤岛。

A.5

我要给出的拒绝交易的**第二个理由**是：在 Banach-Alaoglu 定理的两个证明中，第一个证明（即运用 Tychonoff 定理的一般情形证明）的**真实性依赖于第二个证明**（即结合对角线法与定理 2.4.6 给出的可分空间证明）。

这第二个理由听起来似乎更有冒犯性：难道选择公理不足以支撑 Banach-Alaoglu 定理的正确性吗？难道我们该怀疑选择公理本身的可靠性，不顾数理逻辑学家对公理系统所做出的诸多深刻的研究吗？Gödel 证明了选择公理不会导致矛盾（假设 Zermelo-Fraenkel 公理系统无矛盾），这不足以说服我们信任依赖于选择公理的数学证明的可靠性吗?

读者应该注意到，我在措辞选择上区分了（数学证明语境下的）"正确性"和"真实性"。对于现代数学工作者而言，"一个数学证明是正确的"往往意味着证明里的逻辑链条是没有差错的，证明中从每一步推出下一步都是在逻辑上站得住脚的。而如果一个事先以为是正确的证明事后被发现是有错的，那么这种错误一定会作为逻辑链条里的某一步错误显现出来。而人们说服别人某个证明有错，最硬核的方式便是指出该证明里的哪一步在逻辑上犯了错。

然而，无论是日常生活、自然科学、还是数学历史上的绝大多数时间段，"逻辑上不出错"的正确性和"真实"、"真理"都不是一回事。甚至连"事实"都和"真"不是一回事。"真"永远是有价值取向的，从而既不是价值中立的事实（其扮演的是证据的作用），也不是价值无关的逻辑正确。对于本文的主题而言，"真"的概念中最核心的一项价值便是**对现象的解释力**。

对于物理和化学之类的典范自然科学而言，判断一套理论的真理性的主要依据便是对自然现象的解释力。尤其是当新理论还处在雏形阶段，没有成熟到给人类的实际生活带来应用的阶段，新理论和旧理论的论辩往往表现为针对同一物理/化学现象的解释力的较量。在自然科学之外，例如对于历史研究而言，解释力也是评判理论的核心考量——本讲义给出了关于泛函分析史的诸多观点；若读者能基于合适的历史文献，给出针对这些观点的更有解释力的反驳，那么读者给出的观点便是更为正确的。数学发展的大多数阶段亦是如此：在人们意识到欧几里得几

何并不对应真实的物理时空之前，人们对几何学的要求都不只是一套自洽的模型，而是符合物理现实的“真实”的几何，是一套能够参与解释物理现象的几何。

只有二十世纪以来，以集合论为各领域之共同基础的现代数学，才呈现出对“真”的别具一格的判断标准。现代数学的典型特征之一，便是划分出了评判数学工作的**两个世界——对错**的世界和**价值**的世界。在对错的世界里，人们谈论一项数学工作里的证明是否是正确的，是否有 gap；而在价值的世界里，人们谈论一项数学工作是否重要，它所采用的方法是否有趣、“自然”、“本质”。这种二分世界的最大优点在于，它圈出了一片领地，一片名为“对错的世界”的领地，意图把所有的争吵驱逐到领地外边去——二十世纪初的人们仍然对持续了几个世纪的微积分和函数论中的争吵记忆犹新，因此这种领地的划分便显得极为可贵。而在领地内，所有关于对错的争执，原则上都能通过使所有人都信服的流程得到判决。

毫无疑问，就评判数学工作而言，“对数学现象的解释力”是属于价值的世界的。一项数学工作可以是正确而平凡的；也可以是有一些小错或不太小的错误，却是极为深刻和创新的、或是采用了更为触及本质的方法或思想的。一种数学观点或数学方法是否“自然”、“本质”，归根到底是要被算作“a matter of taste”的，因为不同（学派）的人会给出不同的评价。“A matter of taste”本来也只是现代数学家为了保持圈子和睦和避免争执而采用的措辞，而并非真的是纯主观的东西。说一个数学方法“自然”和“本质”，不过是在说：针对很多的数学现象，这个方法能给出一套系统地、干净地、各环节之间有机联系的、对后续工作有启发性的处理方式——这不正是在说，这个方法对诸多数学现象具有很强的解释力么？

如果把现代数学和十八十九世纪的分析学状况相比，我们便会同意：就隔离出一片远离争执的净土这一目的而言，这种两个世界的划分是比较成功的。可这种成功也引起了人们的奢望——把数学工作和数学证明的真实性彻底封入对错的世界里，从而把“对数学现象的解释力”局限为只事关个人品味或学派风格的美学标准。然而，**价值的世界果真不会影响对错的世界吗？**

A.6

我想大家都会承认，如果一项数学工作是正确的，那它至少应当对正确的数学对象断言正确的结论。然而，判断所谈论之对象是否正确并非完全依据文章中给出的定义。实际上，数学发展的一个常见特征便是，在前沿领域中，不同的人或者不同的学派会给出表述完全不同的定义；而人们判断这些定义是否关乎相同的对象，往往不是靠给出两个定义等价的严格证明，而是靠判断：这两个数学工作中，是否涉及到**相同的数学现象**？

为了说明这一点，我们考察本讲义第 5 章所探讨过的谱理论的历史。“Hilbert 空间上的有界自伴算子的谱分解”这一定理往往被归功于 Hilbert。然而，Hilbert 对谱理论的表述（见定理 5.1.1）远远不是现代谱理论的表现形式。甚至，它要比现代表述来的弱，因而实用性也远远不如他之后的 F. Riesz 给出的谱理论。对

此，我在第 5.1 和 5.2 节有详细的论述。但这里我想说的重点在于：Hilbert 的谱定理 5.1.1 在历史上的下一个改良版是 Riesz 的谱定理 5.4.4；但这两个谱定理为何诉诸相同的对象，在当时是没有被严格证明的。因此，“Riesz 的谱定理推广了 Hilbert 的谱定理”这件事也是没有严格证明的。

Hilbert 谱定理和 Riesz 谱定理的核心差异在于对**预解式**之理解的差异。Riesz 版本的谱定理以函数演算为核心，而一个有界自伴算子 T 的预解式 $(\lambda - T)^{-1}$ 既可以通过逆算子来定义，也可以通过函数演算来表达。与之相对，Hilbert 的谱定理论述的不是函数演算，而是直接关于预解式的，其核心点在于给出预解式的 Stieltjes 积分表示——因此，Hilbert 谱定理所承上的是连分数和发散级数的积分表示。然而，Hilbert 对预解式的理解和 Riesz 的理解（也是我们现代人的理解）不同。Hilbert 采取的不是算子的视角，而是二次型和无穷矩阵的视角。相应的，Hilbert 并不把预解式理解为线性映射的逆，而是理解为有限矩阵的逆矩阵（通过行列式表达）在取极限之后的表达式；见 5.1 节的讨论。

Hilbert 的预解式不仅和后续 Riesz 的预解式不同，也和他之前的 Fredholm 的预解式不同。Fredholm 在积分方程的语境下引入预解式，他的预解式既不是矩阵或二次型，也不是线性映射，而是形如 $\Delta(\lambda, x, y)/\Delta(\lambda)$ 的积分核，其中分母是原积分方程之积分核的行列式，而分子是其余子式。见第 8.1.5 小章节的讨论。

那么，是什么让当时的学术共同体相信 Fredholm、Hilbert、以及 Riesz 在谱理论和预解式方面的工作具有继承关系的呢？这决不是靠严格证明各个定义之间的等价性来实现的。不只一百年前的数学不是这样，当今的数学仍然不是这样。无论什么年代，在一个数学方向中，不同的人或学派给出看起来截然不同的定义，却并不证明不同版本的定义之间的等价性，这都是极其常见的。判断不同的定义指涉相同的对象，其根本依据在于：不同的定义或者数学理论是否**涉及相同的数学现象**——所运用的方法、所能解决的问题、以及所遇到的瓶颈和限制……这些因素在不同的数学工作中是否有足够多的共性。就谱理论和预解式而言：积分方程以及其背后的 Dirichlet 问题与 Sturm-Liouville 问题、连分数与发散级数、矩问题……这些数学对象，以及由这些对象所延伸出的抽象概念、定理、研究方法、方法的威力与局限、定理的预言范围、定理之间的关系网、与邻近方向（例如积分论，Fourier分析）的关联等等，共同构成了谱理论早期发展阶段中的数学现象。

同样地，人们判断一项数学工作中的错误是否是容易修补的小错误，也是以数学现象为依据。在判断一个证明是否正确的时候，学术共同体中的检查者（例如审稿人）会问：这个证明是如何克服我在研究这类问题时遇到的难点的？为什么我的方法不行？前人的失败和该证明的成功是否有一个合乎常理的解释？该证明是否也可以解释该数学方向下的某些别的结论？这种解释多大程度地简化了前人的解释？该证明是否会导出不合常理的结论？……通过找到这些问题的回答，检查者拼出了所考察的数学证明的大致拼图。哪怕这个证明里有错，只要大致拼图

能够和检查者对数学现象的理解相吻合，那么这个错误便会被当作非本质的小错误。

我们已经看到，对于数学证明之正确性的评判标准中的两个重要要素

- 该证明是否关乎正确的数学对象；
- 该证明中的错误是小错误还是实质性的错误

都依赖于处在科研前线的数学工作者对数学现象的理解。并且，这种依赖是全局性的，不仅左右着检查者对该数学证明的整体框架之正确性的判断，也统筹支配着检查者对证明中的各项细节的检验。与此对照，对逻辑链条的验证往往是针对证明中的细节的，故而恰恰是局部的。在此，和理论物理的类比是有益的：量子场论中的数学推导向来都远未达到数学家对严格性的要求，却常能给出数学家没能发现的正确的数学预言。保证这种正确性的原因，同样是物理学家对相应的物理现象的深刻理解。

在一个科学共同体中，以现象为立足点，考察一套理论在整体框架层面的正确性——这不就是考察这套理论对现象的解释力么？那么，一个运用 Tychonoff 定理来偷懒的 Banach-Alaoglu 定理证明，一个错失了和矩问题、傅里叶级数、积分方程、无限矩阵的谱分解、连分数、发散级数等等数学现象的联系的定理证明，还能有多少真实性？与此相反，只要我们知道 Cantor 对角线法在上述现象中所能证明的具体的紧性定理，知道这些紧性定理能够解决的具体问题，那么从这些语境中提炼出来的抽象紧性定理——它叫 Banach-Alaoglu 定理还是叫 Cantor-Arzelà-Ascoli-Hilbert-Helly 定理都无所谓——是**不可能有实质错误的，因为它在数学现象中的根基是牢固的**。哪怕一种抽象定理（或者它的证明）有错，调整一下表述形式和推广策略即可。

A.7

我在上文中给出了两个拒绝交易的理由，并非是要鼓励大家把选择公理挡在泛函分析的门外，而是想要表达：如果说数学学习者未曾理解一个教科书中的定理便直接接纳它，这么做并非天经地义，那么在泛函分析中接纳选择公理也需要足够的理由。

在数学基础和数理逻辑的层面，人们对选择公理的正当性已有较多的讨论。本文意图从数学实践的立场来提供一个不一样的讨论角度。因此，我们也希望有一种来自于数学实践的理由来接纳选择公理。这样的理由会是什么呢？让我们来看看 Gödel 在 [Göd47] 中是怎么说的：

> Furthermore, however, even disregarding the intrinsic necessity of some new axiom... a decision about its truth is possible also in another way... that is, its fruitfulness in consequences and in particular

> in "verifiable" consequences, i.e., consequences demonstrable without the new axiom, whose proofs by means of the new axiom, however, are considerably simpler and easier to discover, and make it possible to condense into one proof many different proofs.

对于泛函分析来说，选择公理恰恰起到了 Gödel 所说的作用：它简化了许多特殊情形下的定理的证明，保障着泛函分析的理论体系不会随着系统化和抽象化而变得过于臃肿。这一点在对偶空间方面体现得尤为明显。赋范空间能典范嵌入它的二次对偶——光是这个依赖于选择公理的结论，便能给对偶空间理论带来极大的便利。$l^1(\mathbb{Z})$ 的非自反性虽然没有现实意义，却是一个精简的对偶空间理论所必然导出的结论。

而我们接受选择公理的这一理由，恰恰验证了我在上文中所说的第二个拒绝交易的理由：在泛函分析中，对于一个依赖于选择公理的定理证明，决定其真实性的—— 用上述引文的话来说就是 “a decision about its truth”——是不依赖于选择公理的定理证明。

可难道只有选择公理才面临过“一个证明还不够”的情形吗？或许，还有许多看起来无懈可击的证明，光凭自身也同样不足以确立定理的真实性？Hilbert 在引入 $l^2(\mathbb{Z})$ 上的全连续二次型的对角化理论来研究积分方程之前，已经在积分方程的具体设定中（而非 $l^2(\mathbb{Z})$ 这一更加代数的设定中）建立了对称核积分算子的对角化理论。用他在 [Hil06] 中的说法，$l^2(\mathbb{Z})$ 上的全连续二次型方法，是对之前结果的简化。或许，我们现代人已经当作典范的 Hilbert 空间方法，在当时仅凭自身也无法赋予积分方程对角化定理足够的真实性？如果没有 Fredholm 和 Hilbert 本人之前在更加具体的积分方程设定下的工作，如果 Hilbert 在一开始便用 $l^2(\mathbb{Z})$ 上的全连续二次型来研究积分方程，那么这一方法的合法性或许得历经坎坷，才能为学术共同体所接受？

数学中同一个定理被给予多个证明，是否只是为了更好地理解这一定理？如果没有多个证明的话，一个孤证是否足以建立起定理（以及该证明本身）的真实性？读者或许能看出我对这一问题的态度[8]，但我想是时候结束这篇文章了。

2025 年 12 月于北京

[8]读者若是有机会在科研工作中，用一个全新的办法证明一个结论；该方法和先前方法的共同起效范围如此之窄，以至于无论如何仔细地检查证明都无法获得安全感，除非能找到该结论主要情形的另一个证明......那么读者对这个问题或许会有一些特别的体会。

Index

References

[Apo] Apostol, T. M. (1974). Mathematical analysis. 2nd ed.

[BS30] Banach, S. and Saks, S., Sur la convergence dans les champs L^p. Stud. Math. 2, 51-57 (1930).

[Ban22] Banach, S., Sur les opérations dans les ensembles abstraits et leur application aux équations intégrales. Fundamenta Mathematicae 3, 133-181 (1922).

[Ban32] Banach, S., Theory of linear operators. (1932)

[Che73] Chernoff, Paul R., Essential self-adjointness of powers of generators of hyperbolic equations. J. Funct. Anal. 12, 401-414 (1973).

[DE-P] Deitmar, Anton and Echterhoff, Siegfried (2014). Principles of harmonic analysis. 2nd revised and updated ed.

[Dav] Davidson, K. R. (1996). C^*-algebras by example. RI: AMS, American Mathematical Society.

[Dir30] Dirac, P. A. M. (1958). The principles of quantum mechanics. 4th ed. First edition in 1930.

[Eva] Evans, L. C. Partial differential equations. 2nd ed. (2010)

[Fol-A] Folland, G.B. (2016). A course in abstract harmonic analysis. 2nd updated edition.

[Fol-P] Folland, G. B. (1995). Introduction to partial differential equations. 2nd ed.

[Fol-R] Folland, G. B. (1999). Real analysis: modern techniques and their applications. 2nd ed.

[Fre03] Fredholm, I. Sur une classe d'équations fonctionnelles. Acta Math. 27, 365-390 (1903).

[Fri34a] Friedrichs, Kurt. (1934) Spektraltheorie halbbeschränkter Operatoren und Anwendung auf die Spektralzerlegung von Differentialoperatoren. I. Math. Ann. 109, 465-487 (1934)

[Göd47] Gödel, K. (1947). What is Cantor's continuum problem?. The American Mathematical Monthly, 54(9), 515-525.

[Gray84] Gray, J. D. (1984). The shaping of the Riesz representation theorem: A chapter in the history of analysis. Archive for History of Exact Sciences, 31, 127-187.

[Gud74] Gudder, S. (1974). Inner product spaces. The American Mathematical Monthly, 81(1), 29-36.

[Gui-A] Bin Gui, Qiuzhen Lectures on Analysis. See https://binguimath.github.io/Pages/2023_Analysis.html

[Gui-S] Bin Gui, Spectral Theory for Strongly Commuting Normal Closed Operators. See the "Notes" section of https://binguimath.github.io

[Haw-L] Hawkins, T. (1979) Lebesgue's theory of integration: Its origins and development. Corrected reprint of the 2nd edition.

[Hel12] E. Helly, Über lineare Funktionaloperationen. Akademie der Wissenschaften in Wien, Mathematisch-Naturwissenschaftliche Klasse, Sitzungsberichte, Abteilung IIa, 121 (1912), pp. 265-297.

[Hel21] E. Helly, Über Systeme linearer Gleichungen mit unendlich vielen Unbekannten. Monatsh. Math. Phys. 31, 60-91 (1921).

[Hil06] Hilbert, D. (1906) Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen. The fourth part (Vierter Abschnitt, Theorie der quadratischen Formen mit unendlich vielen Variablen).

[Hoc80] Hochstadt, H. (1980). Eduard Helly, father of the Hahn-Banach theorem. The Mathematical Intelligencer, 2(3), 123-125.

[Jah] Jahnke, H. N. (2003). A history of analysis (No. 24). American Mathematical Soc..

[Lak63] Lakatos, I. (1963). Proofs and refutations.

[Lax] P. D. Lax. Functional analysis. (2002)

[Leb06] Lebesgue, H. Lecons sur les séries trigonométriques, professées au Collège de France. (1906)

[Leb09] Lebesgue, H. Sur les intégrales singulières. Ann. Fac. Sci. Univ. Toulouse (3) 1, 25-117 (1909).

[Mur36] Murray, F. J. Linear transformations in $\mathfrak{L}_p$, $p > 1$. Trans. Amer. math. Soc. 39, 83-100 (1936).

[vN27] von Neumann, J. Mathematische Begründung der Quantenmechanik. Nachr. Ges. Wiss. Göttingen, Math.-Phys. Kl. 1927, 1-57 (1927).

[vN29a] von Neumann, J. Allgemeine Eigenwerttheorie Hermitescher Funktionaloperatoren. Math. Ann. 102, 49-131 (1929).

[vN29b] von Neumann, J. Zur Algebra der Funktionaloperationen und Theorie der normalen Operatoren. Math. Ann. 102, 370-427 (1929).

[vN31] von Neumann, J. Über Funktionen von Funktionaloperatoren. Annals of Math. (2) 32, 191-226 (1931)

[vN32a] von Neumann, J. (1932). Mathematical foundations of quantum mechanics. English translation of von Neumann's book *Mathematische Grundlagen der Quantenmechanik*.

[vN32b] von Neumann, J. Über adjungierte Funktionaloperatoren. Ann. Math. (2) 33, 294-310 (1932).

[Pes] Pesin, I. N. (1970). Classical and modern integration theories.

[RF] H. Royden and H. Fitzpatrick, Real analysis. 4th ed. (2010)

[RN] F. Riesz and B. Sz.-Nagy, Functional analysis. Transl. from the 2nd French ed. by Leo F. Boron. (1956)

[RS-1] M. Reed and B. Simon, Methods of modern mathematical physics. I: Functional analysis. Rev. and enl. ed. (1982)

[Rie07a] Riesz, F. (1907). Über orthogonale Funktionensysteme. Nachr. Ges. Wiss. Göttingen, Math.-Phys. Kl. 1907, 116-122.

[Rie07b] Riesz, F. (1907). Sur une espèce de géométrie analytique des systèmes de fonctions sommables. C. R. Acad. Sci., Paris. 144, 1409-1411.

[Rie09] Riesz, F. (1909). Sur les opérations functionnelles linéaires. Gauthier-Vllars.

[Rie10] Riesz, F. (1910). Untersuchungen über systeme integrierbarer funktionen. Mathematische Annalen, 69(4), 449-497.

[Rie11] Riesz, F. (1911). Sur certains systémes singuliers d'équations intégrales. In Annales scientifiques de l'École Normale Supérieure (Vol. 28, pp. 33-62).

[Rie13] Riesz, F. (1913). Les systèmes d'équations linéaires à une infinite d'inconnues.

[Rie14] Riesz, F. (1914). Démonstration nouvelle d'un théorème concernant les opérations fonctionelles linéaires. Ann. Sci. Éc. Norm. Supér. (3).

[Rie18] Riesz, F. (1918). Über lineare funktionalgleichungen. Acta math, 41(1), 71-98.

[Rud-R] Rudin, W. (1987). Real and complex analysis. 3rd ed.

[Rud-F] W. Rudin, Functional analysis. 2nd ed. (1991)

[Sch30] Schauder, J. Über lineare, vollstetige Funktionaloperationen. Studia Mathematica, vol. 2, 183-196 (1930).

[Sch-U] Schmüdgen, K. Unbounded self-adjoint operators on Hilbert space. Grad. Texts Math. 265. (2012)

[Sim-R] Simon, Barry (2015). Real analysis. A comprehensive course in analysis, part 1.

[Sim-O] Simon, Barry (2015). Operator theory. A comprehensive course in analysis, part 4.

[Sti94] Stieltjes, T. J. (1894). Recherches sur les fractions continues. Toulouse Ann. VIII, J1-J122.

[Sti-C] Stieltjes. Œuvres complètes. Collected papers. Vol. II contains an English translation of [Sti94].

[Tay-1] Taylor, Michael E., Partial differential equations I. Basic theory. 3rd corrected and expanded edition (2023).

[Tay-2] Taylor, Michael E., Partial differential equations II. Qualitative studies of linear equations. 3rd corrected and expanded edition (2023).

[Xia] 夏道行，吴卓人，严绍宗，舒五昌. 实变函数论与泛函分析，下册，第二版修订本.

[Väth98] Väth, Martin. The dual space of L_∞ is L_1. Indag. Math., New Ser. 9, No. 4, 619-625 (1998).

[You10] Young, W. H. On a new method in the theory of integration. Proc. Lond. Math. Soc. (2) 9, 15–50 (1910).

[You13] Young, W. H. On the new theory of integration. Proc. R. Soc. Lond., Ser. A 88, 170–178 (1913).

[Yu] 于品 (2020). 数学分析之课程讲义. 2020.12.27版

[Zei] Zeidler, Eberhard. Applied functional analysis. Applications to mathematical physics. Vol. 1. (1995)