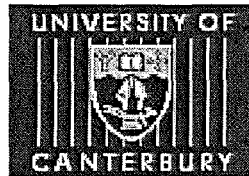


Department of Mathematics and Statistics



Constructive Spectral and Numerical Range Theory

A thesis submitted
in fulfilment of the requirements
of the degree for

Doctorate of Philosophy

at

The University of Canterbury

by

Robin Siale Havea

(*a.k.a.* Robin Tu'akilaumea)

Supervisor: Professor Douglas S. Bridges

April 20, 2001

QA

9.56

.H383

2001

Dedicated to my parents:

Siale Havea and 'Elenoa Tu'akilaumea

Abstract

Following an introductory chapter on constructive mathematics, Chapter 2 contains a detailed constructive analysis of the Toeplitz–Hausdorff Theorem on the convexity of the numerical range of an operator in a Hilbert space. It is shown that the results in the chapter are the best possible with constructive methods.

The rest of the thesis deals with the constructive theory of not-necessarily-commutative Banach algebras. Chapter 3 discusses the Spectral Mapping Theorem in that context, again showing that the results obtained are the best possible. Chapter 4 deals with the question, “Are positive integral powers of a hermitian element of a Banach algebra hermitian?”. A major problem that has to be overcome is to find the ‘right’ constructive definition of *hermitian*, since there is no guarantee in constructive mathematics that the state space of a Banach algebra is nonempty; this forces us to work with approximations to the state space, rather than the state space itself.

In the final chapter, these approximations are used to give careful estimates that lead us to a proof of Sinclair’s Theorem that the spectral radius of a hermitian element equals its norm.

The thesis has two appendices: one describing the axioms of intuitionistic first-order logic, and the other giving a proof of the Spectral Theorem for normal operators on a separable Hilbert space.

Acknowledgements

I am very grateful to my supervisor, Professor Douglas S. Bridges, for all the advice, criticism, and assistance since the first day of this research. There were ups and downs during the three years journey but your expertise, guidance, and patience constantly encouraged and kept me on the right track.

This research would not have been realised without the financial assistance of the Arthington-Davy Fund administered from Trinity College, Cambridge. In particular, I would like to thank Dr. Chris T. Morley.

I would like to express my gratitude to my parents for being there for me whenever I needed them, and, most of all, for always believing in me. I would not have been able to come this far without your support and prayers.

All my fellow postgraduates and colleagues in the department for sharing the facilities, coffee, and pizza.

Siofilisi Hingano, Dr. 'Okusitino Māhina, and Failo Tāufa for all your support.

Mālō and *'ofa atu* to all of my *famili* and *kainga*. Your combined support has been an encouraging factor in my career of being a student. See you soon.

My fellow compatriots at UoC: Siale Faitotonu, Mosese Funaki, Pauli Kautoke, Siaosi Liti, Savelio Mo'ale, Nonga Soakai, 'Ofa Taimani, Lisiate Takau, and Viliami Tu'ipulotu. A company that has nothing to do with my research but their friendships kept me sane throughout my UoC experience.

'Ofa Mafi, a very good (long lost) friend since my first day in GPS, and his wife Sarah for their friendship and support.

Last but not the least, I would like to thank the other 2/3 of my family: my wife, Sulieti, and son, Douglas. The seemingly endless hours of doing research took me away from home and I hardly had a chance to contribute to the daily domestic activities: Your patience, sacrifice, and support could never go unmentioned.

Contents

Abstract	iii
Acknowledgements	iv
Chapter 1 Introduction	
A Mathematical <i>Renaissance</i>	1
1.1 A brief historical development	1
1.2 Questioning tradition	3
1.3 Bishop's mathematics	6
1.4 Interpreting the connectives and quantifiers	9
1.5 Common sources of nonconstructivity	10
1.6 Examples	14
1.7 Short guide to the Thesis	18
Chapter 2 Convexity of the Numerical Range	20
2.1 Introduction	20
2.2 Analysis of the Halmos–de Boor proof	21
2.3 Finding a root of $p_c(t)$	23
2.4 The main results	28
2.5 A limiting example	33
Chapter 3 A Spectral Mapping Theorem	40
3.1 Introduction	40
3.2 The inclusion $p(\sigma(a)) \subset \sigma(p(a))$	41
3.3 The inclusion $\sigma(p(a)) \subset p(\sigma(a))$	44
3.4 A limiting example	50
Chapter 4 Powers of a Hermitian Element	52
4.1 Introduction	52

4.2	Preliminaries	52
4.3	Technical results	54
4.4	Extreme points of the state space	58
Chapter 5 Sinclair’s Theorem		66
5.1	Introduction	66
5.2	Preliminaries	66
5.3	A closer look at the spectral radius	67
5.4	Approximating the numerical range	69
5.5	Sinclair’s Theorem	76
Appendix A Intuitionistic Logic		82
Appendix B A Spectral Theorem		84
References		89
Symbols		95
Index		96

Chapter 1

Introduction

A Mathematical *Renaissance*

1.1 A brief historical development

The search for truth in science has always brought about differences in opinion among scientists. As a part of science, mathematics has its share of different views and philosophies about its foundations. For some philosophers, truth is absolute; for others, it is relative. One of the principle features of most sciences is objectivity. It is widely believed that mathematics is an objective affair. However, there are some that have treated mathematics as a matter of almost extreme subjectivism.

According to the *Platonist* philosophy of mathematics, mathematical realities exist as perfect Platonic forms, and every mathematical statement has an associated truth value even though we may never be able to determine that value within our preferred formal system. A good example of this is the Continuum Hypothesis relative to the formal system ZFC (Zermelo–Fraenkel set theory plus the Axiom of Choice).

A completely different philosophical approach is taken by *Logicism*, which re-

gards mathematics as simply a part, an offspring, of Logic. Later in this chapter we will discuss a different view where mathematics is thought to precede Logic!

One of the most popular philosophies, that marked a very important chapter in the history and philosophy of mathematics in the early twentieth century, was *Formalism*. The leading formalist of his time, David Hilbert (1862–1943) had a strong conviction that mathematics is nothing but a process of manipulating symbols according to a set of specified rules, and that a formal mathematical system is acceptable if it can be shown to be free from contradiction. A major drawback of this approach to mathematics is that meaning is redundant. In Bishop's words,

... Hilbert tried to show that it was all right to neglect computational meaning, because it could ultimately be recovered by an elaborate formal analysis of the techniques of proof. ([8, page 513])

However, Hilbert's program of proving the formal consistency of mathematics was doomed to failure and finally put to rest by the incompleteness theorem of the young Kurt Gödel (1906–1978). According to Davies,

In spite of its superficial plausibility, the formalist interpretation of mathematics received a severe blow in 1931. In that year the Princeton mathematician and logician Kurt Gödel proved a sweeping theorem to the effect that mathematical statements existed for which no systematic procedure could determine whether they are true or false. This was a no-go theorem with a vengeance, because it provided an irrefutable demonstration that something in mathematics is actually impossible, even in principle. The fact that there exist undecidable propositions in mathematics came a great shock, because it seemed to undermine the entire logical foundation of the subject. ([35, pages 100–101])

Despite the downfall of Formalism, it played an important role in the history of the foundation of mathematics, particularly in the period 1925–1930 where we have the peak of the “foundational crisis” (see [37, page 234]; an interesting exposition of this crisis can also be found in [68]). In the next section we discuss major issues in this foundational crisis in mathematics.

1.2 Questioning tradition

Those who were deeply engaged in mathematical foundational matters during the early years of the twentieth century were stunned to learn the new philosophy of mathematics (and life)—*Intuitionism* (INT)—propounded by the Dutch mathematician Luitzen Egbertus Jan Brouwer (1881–1966). This revolutionary view of mathematics was introduced in Brouwer’s doctoral dissertation [33] of 1907. To Brouwer, mathematics is nothing but a purely human intellectual affair; it is a free creation of the human mind. Basically, mathematics in Brouwer’s terms is based on intellectual constructions, and that is how mathematical objects are brought to life. They do not exist waiting to be found; their existence is realised by constructions performed by the human mind. This view contrasts sharply with the Platonic approach; as Dummett puts it,

... the platonistic picture is of a realm of mathematical reality, existing objectively and independently of our knowledge, which renders our statements true or false. On an intuitionistic view, on the other hand, the only thing which can make a mathematical statement true is a proof of the kind we can give: not, indeed, a proof in a formal system, but an intuitively acceptable proof, that is, a certain kind of mental construction. Thus, while, to a platonist, a mathematical theory relates to some external realm of abstract objects, to an intuitionist it relates to our own mental operations: mathematical objects themselves are mental constructions, that is, objects of thought not merely in the sense that they are thought about, but in the sense that, for them, esse est concipi. They exist only in virtue of our mathematical activity, which consists in mental operations, and have only those properties which they can be recognized by us as having. ([38, page 7])

An essential feature of INT is that mathematics precedes logic, which is the opposite of what the logicians believed. In fact, a consequence of Brouwer’s view of mathematics as a process of mental construction was his rejection of the logical *Principle of the Excluded Third*,

P or not P

—in other words, any mathematical statement is either true or false. One must not forget that this principle is one of the cornerstones of classical logic and hence of *classical mathematics* (**CLASS**). Brouwer believed that the Principle of the Excluded Third and even certain weak forms of it allow us to make moves and decisions that are highly nonconstructive in nature. This was a very practical and realistic observation.

Accepting Brouwer's views meant abandoning a substantial part of classical mathematics, at least in its standard form. From a formalist point of view, Brouwer's approach to mathematics was unacceptable, since the working formalist was uninterested in questions of meaning, and preferred to work in a formal system which apparently allowed one to prove more theorems.

Though Brouwer fiercely attacked CLASS, his philosophy of mathematics was hard to understand, and made little positive impact on the majority of mathematicians of his time. His most famous pupil, Arend Heyting (1898–1980), was able to make Brouwer's work more accessible to the nonspecialist by formalising the axioms of intuitionistic logic¹. It is important to point out that Brouwer's work on intuitionism paved the way for later development and progress in modern *constructive mathematics* (**CM**).

In [30], Bridges and Richman discuss various varieties of constructive mathematics.² Their experience of practising CM has led those authors to the informal, philosophy-free view that

CM is nothing but mathematics based on intuitionistic logic.

Based on Brouwer's view of mathematics as being an affair of the intellect, the varieties of CM are bound together by their strict notion and interpretation of

¹See Appendix A.

²Beeson in [3] also gives a thorough exposition of the various varieties of CM.

mathematical existence. Accordingly, when you prove that an object exists in CM, you have to show how to find, or at least how to approximate, it. In CLASS, one can easily get away with it by showing that a contradiction can arise when assuming nonexistence. Generally, to prove in CLASS that x exists, it is enough to show that $\neg x$ implies a contradiction.

Perhaps the most famous example of this classical approach to mathematical existence is Hilbert's existential proof of his famous Basis Theorem, where he established the existence of the basis by means of an indirect (contradiction) proof. His proof disappointed a number of mathematicians as they were expecting a proof which provided enough information from which one could actually construct the desired basis.

In CM, then,

existence means constructibility.

With this notion of mathematical existence in mind, one can easily see that to write a proof in CM is analogous to coding a computer program in some specified language. From a philosophical point of view, the treatment of existence in CM is realistic in the sense that existence is synonymous with constructibility/computability.

An interesting point about conceptual and constructive proofs has been made by Goodman (and others):

It is interesting to see that conceptual clarity and constructivity both lead to their own characteristic kinds of increased generality. A conceptual proof will often show that the hypothesis of the theorem can be weakened, and that therefore the theorem is true about more structures than we thought at first. A constructive proof, on the other hand, uses weaker logical and existential assumptions, and therefore may show that the theorem is true in more frameworks than we thought at first. ([41, page 138])

We end this section with a brief discussion of the Russian school of constructivism (**RUSS**) initiated by Andrei Andreyevich Markov (1903–1979) during the period 1948–1949. The program of the Russian school can be seen as a recursive-function-theoretic approach to mathematics: in essence, RUSS is recursive constructive mathematics. The notion of *Markov algorithm*—equivalent to that of a partial recursive function—provides the basic tool for the mathematical practice of RUSS. Its practitioners operate with a fixed programming language, and believe that a mathematical object exists if it can be produced by a Markov algorithm. Among the various varieties of constructive mathematics, RUSS is unique because it accepts

$$\forall x \in \mathbf{R} \, (\neg(x \leq 0) \Rightarrow x > 0),$$

a principle which is named after Markov. In section 1.5, we will see that this principle is considered dubious, even essentially nonconstructive, by most mathematicians from other schools of CM.

1.3 Bishop’s mathematics

Modern CM started when Bishop’s monograph *Foundations of Constructive Analysis* [5] appeared in 1967. This book brought about a different meaning to how we understand, appreciate, and look at mathematics both in CLASS and CM. Bishop single-handedly developed analysis in a fashion suitably approachable by anyone who is familiar with classical analysis. One is not required to understand Brouwer’s intuitionism when reading *Bishop’s mathematics* (**BISH**). Bishop did not commit to any of the special principles of INT or to the recursive framework of RUSS: he developed his mathematics based on a primitive, unspecified notion of *algorithm* and the Peano properties of the natural numbers, and adhered to the strict interpretation of ‘existence’ as ‘computability’.

In [7], Bishop addressed the requirement of ‘computability’ as the

Fundamental Constructivist Thesis: *Every (representation of) an integer can be converted in principle to decimal form by, a finite, purely routine, process.*

We take note of the careful usage of the phrase ‘in principle’ in the preceding quote. This also points to the fact that the efficiency or complexity of the ‘finite...routine’, or algorithm, is not part of the issue. The important point is our being able to produce a finite routine at least in principle. Of course, in a practical situation a computer programmer values the efficiency of an algorithm, but that is a different matter. The only thing that matters to Bishop is to be able to perform the computation in principle; in particular, he must know in advance that the computation will terminate, meeting its specification, at some stage.

Bishop showed by example that we can do mathematics constructively without distorting the spirit of analysis that an ordinary working mathematician of CLASS is accustomed to. As Richman writes:

Bishop showed that one could adopt a thoroughgoing point of view and still do mathematics as it is usually understood. He did this by appropriating standard mathematical symbolism to carry constructive meaning, rather than introducing a specialized notation, and by developing large areas of rather sophisticated mathematics in a constructive manner. His book can be appreciated by mathematicians unfamiliar with logic or recursive function theory, and avoids the more bizarre intuitionistic notions of choice sequence and bar induction. ([57, page 385])

There are several reasons why BISH may be preferable among varieties of CM, but we single out perhaps the most important one: BISH is compatible with CLASS, INT, and RUSS. A theorem in BISH is automatically a theorem in INT, RUSS, and CLASS, which can be regarded as models of BISH whose common core is BISH itself. This BISH has a remarkable richness of multiple interpretability.

We end this section with some remarks of Bishop concerning INT and RUSS. In his view:

Many mathematicians regard the theory of computation as a branch of recursive function theory. It is true that many constructivists, for instance the school of Markov in Russia, are recursivists. Brouwer, of course was not. The recursive constructivists seem to be motivated by the desire to avoid such vague terms as "rule" and "set". Their mathematics is forbiddingly involved and laborious, a great price to pay for the precision they hope to attain....In my opinion, the positive contributions of recursive function theory to both constructive mathematics and the more concrete aspects of the theory of computation are the construction of counterexamples, but here again impressions are somewhat misleading. The methods of Brouwer, now largely neglected, are more suitable for providing counterexamples in most cases of interest than are the methods of recursive function theory. ([8, page 514])

He went on to comment that:

The movement he [Brouwer] founded has long been dead, killed partly by compromises of Brouwer's disciples with the viewpoint of idealism, partly by extraneous peculiarities of Brouwer's system which made it vague and even ridiculous to practicing mathematicians, but chiefly by the failure of Brouwer and his followers to convince the mathematical public that abandonment of the idealistic viewpoint would not sterilize or cripple the development of mathematics. Brouwer and other constructivists were much more successful in their criticisms of classical mathematics than in their efforts to replace it with something better. ([5, page ix])

Brouwer was deeply engaged in foundational debates about mathematics and the creativity of the human mind. For him, mathematics only dealt with the objects that were given meaning by his philosophy of intuitionism. On the other hand, a practitioner of BISH finds that mathematics can be carried out with intuitionistic logic on any reasonably defined mathematical objects, not just some class of so-called constructive objects.

1.4 Interpreting the connectives and quantifiers

Since we will be doing mathematics based on intuitionistic logic, it is worth recalling the standard constructive interpretations of the logical quantifiers and connectives. It will become clear that some, if not most, of these constructive interpretations are very different from their classical interpretations.

In the following, P and Q are any given statements.

- \wedge (**and**) To prove $P \wedge Q$, we must have a proof of P and a proof of Q .
- \vee (**or**) To prove $P \vee Q$, we must have either a proof of P or a proof of Q .
- \Rightarrow (**implies**) To prove $P \Rightarrow Q$ means there is an algorithm that transforms a proof of P into a proof of Q .
- \neg (**not**) To prove $\neg P$, we must show that P implies a contradiction (such as $0 = 1$).
- \exists (**there exists**) To prove $\exists a P(a)$, we must compute a and demonstrate that $P(a)$ holds.
- \forall (**for all**) A proof of $\forall a \in A P(a)$ is an algorithm that, applied to each element a of A and to the data showing that a belongs to A , proves that $P(a)$ holds.

As we mentioned earlier, these constructive interpretations enabled Heyting to produce a complete list of the axioms of intuitionistic logic. In the next section, we shall see the impact of intuitionistic logic, which leads to the rejection of some trivial principles of classical mathematics and classical logic.

The interpretation of $P \wedge Q$ is similar to the classical treatment. Classically, to prove $P \vee Q$ it suffices to establish $\neg(\neg P \wedge \neg Q)$; but proving the latter is not enough to prove the former in CM. Why? Generally, in CM it is not possible to decide, from a proof of $\neg(\neg P \wedge \neg Q)$, which of the alternatives P, Q holds. The constructive

interpretation of \forall is well tied to the notion of decidability in CM; one of the main features of CM is being able to make decision and the constructive interpretation of \forall captures it all.

Finally, to prove $\exists x \in A P(x)$ in CLASS, it suffices to show that $\neg \forall x \neg P(x)$; classical existence is equivalent to the impossibility of nonexistence. In CM to prove $\exists x \in A P(x)$ we must construct an object ξ (at least in principle), show that ξ satisfies the conditions for membership of A , and then show that $P(\xi)$ holds.

1.5 Common sources of nonconstructivity

In this section, we discuss some of the common principles that bear the seeds of nonconstructivity in mathematics. Most of these principles are classically trivial. We will give some well known examples demonstrating how untrustworthy they are in CM. Upon rejecting these principles, it follows that any statement proved to be equivalent to, or implying, any of them is regarded as essentially nonconstructive.

The first principle is the

Law of Excluded Middle (LEM): For any given statement P , either P is true or P is false.

Bearing in mind the constructive interpretation of logic discussed in section 1.4, we see that the rejection of LEM is closely connected with the fact that the property

$$\forall n P(n) \vee \neg \forall n P(n)$$

need not hold in CM even when $P(n)$ is a decidable property of natural numbers n . Much of the reasoning we normally encounter in CLASS is based on LEM. Applications of LEM make life easier, especially in the case of existential proofs.

Consider the theorem:

There exist irrational numbers r and s such that r^s is rational.

Following a simple argument given by Bishop in [7, page 6], consider the real number $\sqrt{2}^{\sqrt{2}}$. By virtue of LEM, either $\sqrt{2}^{\sqrt{2}}$ is rational or $\sqrt{2}^{\sqrt{2}}$ is irrational. In the former case, if $\sqrt{2}^{\sqrt{2}}$ is rational, then simply take $r = s = \sqrt{2}$ and we are done. In the latter case, if $\sqrt{2}^{\sqrt{2}}$ is irrational, then take $r = \sqrt{2}^{\sqrt{2}}$ and $s = \sqrt{2}$. Then r and s are irrational and $r^s = \sqrt{2}^{\sqrt{2} \cdot \sqrt{2}} = 2$ which is rational. A closer look at this proof shows that under LEM we have proved there are numbers r and s satisfying the claim but there is no hint at all on how to actually finding the two numbers!

A *binary sequence* is a finite routine that assigns to each positive integer an element of $\{0, 1\}$. The next three principles deal with such sequences and are classically trivial and special cases of LEM.

Limited Principle of Omniscience (LPO): If (a_n) is a binary sequence, then either $a_n = 0$ for all n or else there exists n such that $a_n = 1$.

Weak LPO (WLPO): For any binary sequence (a_n) , either $a_n = 0$ for each n , or else it is impossible that $a_n = 0$ for each n .

Lesser LPO (LLPO): If (a_n) is a binary sequence containing at most one term equal to 1, then either $a_{2n} = 0$ for each n , or else $a_{2n+1} = 0$ for each n .

It appears that Brouwer was responsible for first drawing attention to the nonconstructive nature of LPO and LLPO ([67, page 6]) though under different names; we have chosen to use the names given by Bishop ([9, page 22]). None of these three omniscience principles can be derived within Heyting arithmetic, and each is false, even classically, in recursive mathematics ([30, Chapters 1,3,7]).

The following, due to Richman ([59, page 135, Section 3]), is a succession of ever weaker omniscience principles, LLPO being the special case $N = 2$.

LLPO_N ($N = 2, 3, \dots$): If $(a_n)_{n=0}^{\infty}$ is a binary sequence with at most one term equal to 1, then there exists j , where $0 \leq j \leq N - 1$, such that $a_{kN+j} = 0$ for all k .

For illustration, suppose that we can prove LPO constructively. Then, applying it to the binary sequence (a_n) defined by

$$a_n = \begin{cases} 0 & \text{if } 2k \text{ is a sum of two primes for } 2 \leq k \leq n+1, \\ 1 & \text{otherwise,} \end{cases}$$

we can either prove the Goldbach Conjecture—every even integer greater than 2 is a sum of two primes—or else find an even integer greater than 2 that is not a sum of two primes.

It is well known that each of the following statements is equivalent to LPO.

- $\forall x \in \mathbf{R} (x = 0 \vee x \neq 0)$.
- **Law of Trichotomy:** $\forall x \in \mathbf{R} (x < 0 \vee x = 0 \vee x > 0)$.
- **Least-upper-bound Principle:** *Each nonempty subset of \mathbf{R} that is bounded away has a least upper bound.*
- *Every real number is either rational or irrational.*

Similarly, each of the following is equivalent to LLPO.

- $\forall x \in \mathbf{R} (x \geq 0 \vee x \leq 0)$.
- *If $x, y \in \mathbf{R}$ and $xy = 0$, then $x = 0$ or $y = 0$.*
- **Intermediate Value Theorem:** *If $f : [0, 1] \rightarrow \mathbf{R}$ is a continuous function with $f(0) < 0 < f(1)$, then there exists $x \in (0, 1)$ such that $f(x) = 0$.*

Another, more controversial, omniscience principle is

Markov's Principle (MP) : If (a_n) is a binary sequence for which it is contradictory that all terms be 0, then there exists n such that $a_n = 1$.

Though accepted and freely used in RUSS, MP is considered a form of unbounded search and is considered highly nonconstructive in both INT and BISH.

CM is often identified, wrongly, with its rejection of the full-blooded

Axiom of Choice (AC): If A, B are sets, and S is a nonempty subset of $A \times B$ such that for each $a \in A$ there exists $b \in B$ with $(a, b) \in S$, then there exists a function $f : A \rightarrow B$ (called the *choice function*) such that $(a, f(a)) \in S$ for all $a \in A$.

From its introduction by Zermelo in 1908, AC was regarded with unease by many mathematicians; it was rejected outright by the intuitionists. Much later, in 1978, Goodman and Myhill in [43] showed that AC implies LEM. So AC cannot be used in any coherent constructive mathematics. However, most mathematicians in CM use the **Principle of Countable Choice**—the special case of AC in which A is the set \mathbb{N} of natural numbers—and the **Principle of Dependent Choice**:

If $a_0 \in A$, and if for each $a \in A$ there exists $a' \in A$ such that $P(a, a')$, then there exists a mapping $f : \mathbb{N} \rightarrow A$ such that $f(0) = a_0$ and $P(f(n), f(n+1))$ for each $n \in \mathbb{N}$.

A final comment: working in CM means working with fewer axioms and principles. This is not the case with CLASS. The rejection of many classically innocent, trivial principles brings extra challenge in CM. This does not mean that CM is an attempt to replace everything in CLASS but one can view CM as some sort of mathematical *revival* where numerical meaning and computation are highlighted and considered part of one's main focus. In the preface to his book [5, page x], Bishop wrote:

We are not contending that idealistic mathematics is worthless from the constructive point of view. This would be as silly as contending that unrigorous mathematics is worthless from the classical point of view. Every theorem proved with idealistic methods presents a challenge: to find a constructive version, and to give it a constructive proof.

1.6 Examples

A useful method for establishing the nonconstructivity of a given proposition P in CM is by means of examples in which we show that P entails, or is equivalent to, some highly nonconstructive principle. Such **Brouwerian counterexamples** serve as a good measure of what to expect and what we can hope for in CM. Further discussion of the role of Brouwerian counterexamples in CM can be found in [30, page 3].

Our first example is from the real number line. We show that the statement

Every real number is either rational or irrational.

entails LPO.

Brouwerian Example 1.6.1 Let $(a_n)_{n=0}^\infty$ be an increasing binary sequence, and define a real number by

$$x = \sum_{n=0}^{\infty} \frac{1 - a_n}{n!}.$$

Suppose that either x is rational or x is irrational. If x is rational, then $|x - e| > 0$, so there exists N such that

$$\sum_{n=0}^N \left(\frac{1}{n!} - \frac{1 - a_n}{n!} \right) > 0;$$

whence $a_n = 1$ for some $n \leq N$. If x is irrational, then, clearly, $a_n = 0$ for all n .

That is, either $a_n = 1$ for some n or $a_n = 0$ for all n which is LPO. ///

Later, in Chapter 2, we shall investigate the convexity of the *numerical range*

$$W(T) = \{ \langle Tx, x \rangle : x \in H, \|x\| = 1 \}$$

of a bounded linear operator T on a Hilbert space H . Our next example shows that the statement

The numerical range of a selfadjoint operator on a 2-dimensional complex Hilbert space is convex.

implies LLPO₃. First we need a lemma.

Lemma 1.6.2 *If λ, μ are complex numbers satisfying the equations*

$$4 \operatorname{Re} (\lambda \mu^*) + 2 |\mu|^2 = 1,$$

$$|\lambda|^2 + |\mu|^2 = 1,$$

then $|\lambda| \geq 1/5$ and $|\mu| \geq 1/5$.

Proof. If $|\mu| < 1/5$, then $|4 \operatorname{Re} (\lambda \mu^*)| < 4/5$ and so

$$4 \operatorname{Re} (\lambda \mu^*) + 2 |\mu|^2 < \frac{4}{5} + \frac{2}{25} < 1,$$

a contradiction; whence $|\mu| \geq 1/5$.

On the other hand, if $|\lambda| < 1/5$, then $|4 \operatorname{Re} (\lambda \mu^*)| < 4/5$ and so

$$\begin{aligned} 4 \operatorname{Re} (\lambda \mu^*) + 2 |\mu|^2 &> 2 (1 - |\lambda|^2) - \frac{4}{5} \\ &> 2 \left(\frac{24}{25} \right) - \frac{4}{5} \\ &> 1, \end{aligned}$$

a contradiction. Hence $|\lambda| \geq 1/5$. Q.E.D.

We now return to our Brouwerian counterexample.

Brouwerian Example 1.6.3 Let H be a 2-dimensional complex Hilbert space, x and y unit vectors in H , and (a_n) a binary sequence with at most one term equal to 1. Define

$$\xi = \sum_{n=1}^{\infty} \frac{a_{3n} + a_{3n+1} + a_{3n+2}}{n^2},$$

and define a selfadjoint operator T on H by

$$\begin{aligned} Ty &= \left(\sum_{n=1}^{\infty} a_{3n} i \xi + \sum_{n=1}^{\infty} a_{3n+1} i n \xi + \sum_{n=1}^{\infty} a_{3n+2} n \xi \right) x, \\ Tx &= \xi x + \left(- \sum_{n=1}^{\infty} a_{3n} i \xi - \sum_{n=1}^{\infty} a_{3n+1} i n \xi + \sum_{n=1}^{\infty} a_{3n+2} n \xi \right) y. \end{aligned}$$

For a unit vector $z = \lambda y + \mu x$ the equation

$$\langle Tz, z \rangle = \frac{1}{2} \langle Tx, x \rangle + \frac{1}{2} \langle Ty, y \rangle = \frac{1}{2} \xi \quad (1.1)$$

reduces to

$$|\mu|^2 \xi + 2 \operatorname{Re} \lambda \mu^* \left(\sum_{n=1}^{\infty} a_{3n} i \xi + \sum_{n=1}^{\infty} a_{3n+1} i n \xi + \sum_{n=1}^{\infty} a_{3n+2} n \xi \right) = \frac{1}{2} \xi.$$

If $a_{3N+1} = 1$, we must solve the equations

$$2|\mu|^2 - 4N \operatorname{Im} (\lambda \mu^*) = 1,$$

$$|\lambda|^2 + |\mu|^2 = 1.$$

We then have

$$|\operatorname{Im} (\lambda \mu^*)| = \frac{1}{4N} |1 - 2|\mu|^2| \leq \frac{3}{4N}.$$

If $a_{3N+2} = 1$, we must solve the equations

$$2|\mu|^2 + 4N \operatorname{Re} (\lambda \mu^*) = 1,$$

$$|\lambda|^2 + |\mu|^2 = 1.$$

In this case, we have

$$|\operatorname{Re} (\lambda \mu^*)| = \frac{1}{4N} |1 - 2|\mu|^2| \leq \frac{3}{4N}.$$

If $a_{3N} = 1$, we must solve the equations

$$2|\mu|^2 - 4 \operatorname{Im} (\lambda \mu^*) = 1,$$

$$|\lambda|^2 + |\mu|^2 = 1.$$

In this case we must have $|\lambda| \geq 1/5$ and $|\mu| \geq 1/5$, by Lemma 1.6.2 (applied with λ replaced by $-i\lambda$).

Now suppose that we have found a unit vector $z = \lambda y + \mu x$ satisfying equation (1.1). Either $|\lambda\mu^*| < 1/25$ or $\lambda\mu^* \neq 0$. In the first case we see from the foregoing arguments that $a_n = 0$ for each n . In the second case either $\operatorname{Re}(\lambda\mu^*) \neq 0$ or $\operatorname{Im}(\lambda\mu^*) \neq 0$. To handle the first of these alternatives, we choose ν such that $\frac{3}{4\nu} < |\operatorname{Re}(\lambda\mu^*)|$. If $a_{3n+1} = 1$ for some $n > \nu$, then $|\operatorname{Re}(\lambda\mu^*)| \leq \frac{3}{4n} < \frac{3}{4\nu}$, a contradiction. It follows that if $a_{3n+1} = 0$ for all $n \leq \nu$, then $a_{3n+1} = 0$ for all n . Hence

$$\forall n (a_{3n} = 0) \vee \forall n (a_{3n+1} = 0) \vee \forall n (a_{3n+2} = 0). \quad (1.2)$$

Finally, assuming that $\operatorname{Im}(\lambda\mu^*) = 0$, and choosing a positive integer ν such that $|\operatorname{Im}(\lambda\mu^*)| > \frac{3}{4\nu}$ we see that if $a_{3n+2} = 0$ for all $n \leq \nu$, then $a_{3n+2} = 0$ for all n ; whence statement (1.2) holds in this case also. This completes the proof that the statement about convexity of the numerical range of T implies LLPO_3 . ///

Now let B be a commutative Banach algebra, and B' its dual. A *character* of B is a bounded homomorphism of B onto \mathbb{C} , and that the *character space* (or *spectrum*) of B is the set

$$\Sigma_B = \{u \in B' : u(e) = 1, u(xy) = u(x)u(y) \text{ for all } x, y \in B\}.$$

Following [10, page 452] and [21, pages 150–151], our next example shows that the statement

The spectrum of every separable commutative unital Banach algebra is compact.

implies WLPO.

Brouwerian Example 1.6.4 Let $(a_n)_{n=0}^\infty$ be an increasing binary sequence. Let B be the algebra consisting of all sequences $\mathbf{x} = (x_n)_{n=0}^\infty$ of complex numbers for which

$$\|\mathbf{x}\| = \sum_{n=0}^{\infty} (1 - a_n) |x_n| \quad (1.3)$$

exists. We define the elements \mathbf{x} and $\mathbf{y} = (y_n)_{n=0}^\infty$ to be *equal* if $\|\mathbf{x} - \mathbf{y}\| = 0$. Then B is a Banach space equipped with norm given by (1.3). Moreover, if we define the product of any two elements \mathbf{x} and \mathbf{y} of B by

$$\mathbf{xy} = \left(\sum_{i=0}^n x_i y_{n-i} \right)_{n=0}^{\infty},$$

then B is a Banach algebra with identity $\mathbf{e} = (1, 0, 0, \dots)$. Let

$$\mathbf{z} = (1, 2^{-1}, 2^{-2}, 2^{-3}, \dots) \in B.$$

If $a_n = 1$ for some n , then the character space Σ_B of B consists of the single element $\mathbf{x} \mapsto x_0$. On the other hand, if $a_n = 0$ for all n , then to each complex number ξ with $|\xi| \leq 1$ there corresponds an element u_ξ of Σ_B defined by

$$u_\xi(\mathbf{x}) = \sum_{n=0}^{\infty} x_n \xi^n.$$

Suppose Σ_B is compact. Since the mapping $u \mapsto |u(\mathbf{z})|$ is uniformly continuous relative to the weak* topology on the unit ball of B' it maps Σ_B to a totally bounded subset of \mathbf{R} ; whence

$$R = \sup \{ |u(\mathbf{z})| : u \in \Sigma_B \}$$

exists. Either $R > 1$ or $R < 2$. In the first case, we have $a_n = 0$ for all n . In the second case, we cannot have $a_n = 0$ for all n . ///

1.7 Short guide to the Thesis

The analysis in the rest of this thesis is carried out within the framework of Bishop's constructive mathematics. Following this introductory chapter, there are four ad-

ditional chapters, two appendices, list of references, table of symbols, and an index for quick referencing. Throughout the thesis, we have tried to adopt most of the commonly used standard symbols of classical functional analysis.

Familiarity with standard classical Hilbert space theory is assumed throughout the whole of Chapter 2. Similarly, we assume familiarity with both standard (classical and constructive) Banach algebra theory in Chapters 3, 4, and 5.

Chapter 2 deals with an in-depth constructive analysis of a classical proof of the Toeplitz–Hausdorff Theorem.

Some of the standard notions in constructive Banach algebra theory which we need in the rest of the thesis are provided in Chapter 3. The highlight of the chapter is a constructive proof the Spectral Mapping Theorem.

In Chapter 4 we investigate the question of whether the square of a Hermitian element of a Banach algebra is Hermitian or not. Positive elements are also discussed.

Chapter 5 shows that, although we cannot necessarily construct elements of the numerical range of a Banach algebra, we can work successfully with approximations to the numerical range. A consequence of the approximation process is a constructive proof of Sinclair’s Theorem.

Chapter 2

Convexity of the Numerical Range

2.1 Introduction

Recall that the *numerical range* of an operator T on a Hilbert space H is the subset of the complex numbers given by

$$W(T) = \{ \langle Tx, x \rangle : x \in H, \|x\| = 1 \}.$$

The classical **Toeplitz–Hausdorff Theorem** says that:

If T is an operator on a Hilbert space, such that the adjoint T^ exists, then $W(T)$ is a convex subset of \mathbb{C} .*

Several proofs of this theorem are considered ‘computational’ ([46, pages 317–318]), but in fact none seems to fit the formal concept of being computational in CM. The main aim of this chapter is to find conditions which ensure constructively that $W(T)$ is convex. The following analysis is motivated by and based on the Halmos–de Boor proof on pages 317–318 of [46]. The final products are Theorems 2.4.3 and 2.4.5 which we show, by means of Brouwerian examples, to be the best we can hope for in a constructive setting.

The proof of our main result, Theorem 2.4.3, depends on our ability to decide that if $-1 \leq a \leq 1$, $0 \leq c \leq 1$, and $b \in \mathbf{R}$, then the quadratic equation

$$(2c(1-a) + 2b - 1)t^2 - 2(c(1-a) + b)t + c = 0, \quad (2.1)$$

has a root in $[0, 1]$. For convenience, we write

$$\begin{aligned} \alpha &= c(1-a) + b, \\ p_c(t) &= (2\alpha - 1)t^2 - 2\alpha t + c, \end{aligned}$$

so that equation (2.1) is just $p_c(t) = 0$. The quadratic formula gives

$$\begin{aligned} t_- &= \frac{\alpha - \sqrt{\alpha^2 - 2\alpha c + c}}{2\alpha - 1}, \\ t_+ &= \frac{\alpha + \sqrt{\alpha^2 - 2\alpha c + c}}{2\alpha - 1} \end{aligned}$$

as standard solutions to equation (2.1) in the case $2\alpha - 1 \neq 0$.

2.2 Analysis of the Halmos–de Boor proof

We now take a closer look at the classical Halmos–de Boor proof of the Toeplitz–Hausdorff theorem. This proof proceeds as follows.

Let x and y be unit vectors in H , and write

$$\xi = \langle Tx, x \rangle \quad \text{and} \quad \eta = \langle Ty, y \rangle.$$

We must show that the line segment, $t\xi + (1-t)\eta$ ($0 \leq t \leq 1$), joining ξ and η lies entirely inside $W(T)$. To this end, the proof is split into a number of steps as follows.

- (i) The problem is trivial when $\xi = \eta$. Assuming that $\xi \neq \eta$, reduce to the case where $\xi = 1$ and $\eta = 0$.

- (ii) Writing $T = A + iB$, where $A = \frac{1}{2}(T + T^*)$ and $B = \frac{1}{2i}(T - T^*)$, further reduce to the case where

$$\langle Bx, x \rangle = 0 = \langle By, y \rangle \quad \text{and} \quad \operatorname{Re} \langle Bx, y \rangle = 0.$$

- (iii) Writing

$$h(t) = tx + (1 - t)y \quad (0 \leq t \leq 1),$$

observe that $h(t)$ never vanishes since the vectors x and y are linearly independent.

- (iv) Expanding $\langle Bh(t), h(t) \rangle$, show that $\langle Th(t), h(t) \rangle$ is real for every $t \in [0, 1]$.

Hence the function

$$t \longmapsto \frac{\langle Th(t), h(t) \rangle}{\|h(t)\|^2}$$

is real-valued and continuous on $[0, 1]$. Since its values at 0 and 1 are, respectively, 0 and 1, we conclude from the Intermediate Value Theorem that the range of this function is $[0, 1]$.

A closer look at the foregoing steps reveals the following motivation for our constructive analysis in the sequel.

- In Step (i), we cannot hope constructively to be able to make the decision that for any two complex numbers ξ and η , either $\xi = \eta$ or $\xi \neq \eta$.
- The problem with Step (ii) occurs where a certain complex number z is multiplied by a complex number λ of unit modulus to obtain $\operatorname{Re}(\lambda z) = 0$; perhaps surprisingly, this cannot be done constructively for a general $z \in \mathbb{C}$.
- In Step (iii) we need to prove, for each $t \in [0, 1]$, not just that $\neg(h(t) = 0)$ but that $h(t) \neq 0$ in the stronger sense that $\|h(t)\| > 0$; note that this sense is stronger unless we are prepared to accept the constructively dubious MP.

- The biggest problem occurs in Step (iv) with the application of the Intermediate Value Theorem: the best conclusion we have, constructively, using the Halmos–de Boer argument as it stands, is that the range of the mapping

$$t \longmapsto \frac{\langle Th(t), h(t) \rangle}{\|h(t)\|^2}$$

is dense in $[0, 1]$, which only enables us to assert that the *closure* of the segment joining $\langle Tx, x \rangle$ and $\langle Ty, y \rangle$ lies in $W(T)$. For the conclusion of the classical Intermediate Value Theorem to hold constructively, the continuous function must satisfy one of a number of additional hypotheses, one of which is that the function be a polynomial.

Proposition 2.2.1 *If $f : [0, 1] \rightarrow \mathbf{R}$ is a polynomial function such that $f(0) < 0$ and $f(1) > 0$, then there exists $x \in [0, 1]$ such that $f(x) = 0$. ([10, page 63, Problem 17]).*

2.3 Finding a root of $p_c(t)$

The main emphasis of this section is on the process of locating a root of

$$p_c(t) = (2\alpha - 1)t^2 - 2\alpha t + c$$

in $[0, 1]$. A typical situation that we have to deal with below is one in which, for a certain parametrised family $(p_x)_{x \geq 0}$ of polynomials, we know that the equation $p_x(t)$ has a root if either $x > 0$ or $x = 0$. Since we cannot assume that for any $x \geq 0$ either $x > 0$ or $x = 0$, we compute a Cauchy sequence¹ of approximate solutions to the equation $p_x(t) = 0$, the limit of which is an exact solution.

¹The construction of the Cauchy sequence depends on the constructively valid proposition: *if a and b are any real numbers with $a < b$, then for each real number x , either $x > a$ or $x < b$* ([16, page 48, (4.9)(v)]). This technique is commonly used in CM, and is illustrated in the proof of Lemma 2.3.1.

Lemma 2.3.1 *Let $I = [0, 1]$, and let $(f_c)_{c \in I}$ be a family of mappings of I into \mathbf{R} such that $f_0(0) = 0$. Suppose that there exist strictly decreasing sequences $(\delta_n)_{n=1}^\infty$ and $(\varepsilon_n)_{n=1}^\infty$ in $(0, 1)$ converging to 0, such that if $0 < x \leq \delta_n$, then $f_x(t) = 0$ for some $t \in [0, \varepsilon_n]$. Then for each $c \in [0, \delta_1)$ there exists $t \in I$ such that $f_c(t) = 0$.*

Proof. Given $c \in [0, \delta_1)$, construct an increasing binary sequence (λ_n) such that

$$\lambda_n = 0 \Rightarrow c < \delta_{n+1},$$

$$\lambda_n = 1 \Rightarrow c > \delta_{n+2}.$$

Without loss of generality, we may assume that $\lambda_1 = 0$. If $\lambda_n = 0$, choose $t_n \in [0, \varepsilon_{n+1}]$ such that $f_{\delta_{n+1}}(t_n) = 0$. If $\lambda_n = 1 - \lambda_{n-1}$, then $\delta_{n+2} < c < \delta_n$, and so there exists $t_n \in [0, \varepsilon_n]$ such that $f_c(t_n) = 0$; in this case we set $t_k = t_n$ for all $k \geq n$. Then (t_n) is a Cauchy sequence in $[0, 1]$; in fact, $|t_m - t_n| \leq 2\varepsilon_n$ for all $n \geq 2$. Hence (t_n) converges to a limit $t_\infty \in [0, \varepsilon_1] \subset I$. If $f_c(t_\infty) \neq 0$, then $\lambda_n = 0$ for all n ; so $c = 0 = t_\infty$ and therefore $f_c(t_\infty) = f_0(0) = 0$, a contradiction. Hence, we conclude that $f_c(t_\infty) = 0$. Q.E.D.

At first glance, the next lemma seems trivial; but in fact that is not the case. If $p(x)$ is a monic quadratic polynomial with roots t_1 and t_2 given by the standard quadratic formula, then $\max\{t_1, t_2\}$ and $\min\{t_1, t_2\}$ are also solutions to $p(x) = 0$. The problem here is that we may be unable to decide whether $\max\{t_1, t_2\} = t_1$ or $\max\{t_1, t_2\} = t_2$.

Lemma 2.3.2 *Let $p(x)$ be a monic quadratic polynomial with real roots t_1 and t_2 given by the standard quadratic formula. Then $|t| \leq \max\{|t_1|, |t_2|\}$ for any real root t of p .*

Proof. Write $p(x) = x^2 + \beta x + \gamma$. Let t be a root of p , so that

$$p(x) = (x - t)(x - t')$$

where $-\beta = t + t'$ and $\gamma = tt'$. Suppose that $|t| > \max\{|t_1|, |t_2|\}$. Then since

$$(|t| + |t'|)^2 = \beta^2 + 2(|\gamma| - \gamma) = (|t_1| + |t_2|)^2,$$

we have

$$|t| + |t'| = |t_1| + |t_2| \leq 2 \max\{|t_1|, |t_2|\};$$

whence $|t'| < \max\{|t_1|, |t_2|\}$. It follows that $p(x)$ has three distinct roots: namely, t , t' , and at least one of t_1 and t_2 . This is absurd. Hence $|t| \leq \max\{|t_1|, |t_2|\}$. Q.E.D.

Lemma 2.3.3 *For each $\varepsilon > 0$ there exists $\delta > 0$ such that if $-1 \leq a \leq 1$, $|b| < \delta$, $0 \leq c < \delta$, and t is a real root of equation (2.1), then $|t| < \varepsilon$.*

Proof. Recall that $\alpha = c(1 - a) + b$. If $0 < \delta < 1$, $|b| < \delta$, and $0 \leq c < \delta$, then

$$|\alpha| < \delta(1 - a) + \delta \leq 3\delta$$

and therefore

$$\begin{aligned} |\alpha^2 - 2\alpha c + c| &\leq |\alpha|^2 + 2|\alpha|c + c \\ &< 9\delta^2 + 6\delta^2 + \delta \\ &< 16\delta. \end{aligned}$$

Furthermore, if $0 < \delta < 1/12$, then

$$|2\alpha - 1| \geq 1 - 2|\alpha| > \frac{1}{2},$$

and so the solutions t_- and t_+ of the quadratic equation (2.1) satisfy

$$\max\{|t_-|, |t_+|\} \leq 2(|\alpha| + |\alpha^2 - 2\alpha c + c|) < 2(3\delta + 16\delta) = 38\delta.$$

Applying Lemma 2.3.2, we see that $|t| < 38\delta$ for every real root t of equation (2.1).

Given $\varepsilon > 0$, we now need only take $\delta = \min\{1/12, \varepsilon/38\}$. Q.E.D.

Lemma 2.3.4 *Let $-1 \leq a \leq 1$ and $b < 0 \leq c < 1/4$. If $|b| > 3c/2$, then $0 < t_- < 1$.*

Proof. We first note that

$$2\alpha - 1 = 2c(1 - a) + 2b - 1 < 4c - 3c - 1 < -\frac{3}{4},$$

that

$$\alpha - c = c(1 - a) + b - c = -ac + b < c + b < 0,$$

and that

$$2\alpha - c = 2c(1 - a) + 2b - c < -3c + 4c - c = 0.$$

Since

$$\alpha^2 - 2\alpha c + c \geq \alpha^2 - 2\alpha c + c^2 = (\alpha - c)^2,$$

it follows that $\sqrt{\alpha^2 - 2\alpha c + c}$ is real, that

$$\alpha - \sqrt{\alpha^2 - 2\alpha c + c} \leq \alpha - (c - \alpha) = 2\alpha - c,$$

and therefore that

$$t_- = \frac{\alpha - \sqrt{\alpha^2 - 2\alpha c + c}}{2\alpha - 1} \geq \frac{2\alpha - c}{2\alpha - 1} > 0.$$

On the other hand, since $0 \leq c \leq 1/4$ and $2\alpha - 1 < 0$,

$$\begin{aligned} 0 &\leq \alpha^2 - 2\alpha c + c \\ &= \alpha^2 - 2c \left(\alpha - \frac{1}{2} \right) \\ &< \alpha^2 - \frac{1}{2} \left(\alpha - \frac{1}{2} \right) \\ &= \left(\alpha - \frac{1}{2} \right)^2 \end{aligned}$$

and therefore

$$\sqrt{\alpha^2 - 2\alpha c + c} < \frac{1}{2} - \alpha.$$

Hence

$$t_- < \frac{2\alpha - \frac{1}{2}}{2\alpha - 1}.$$

and therefore $t_- < 1$. Q.E.D.

Lemma 2.3.5 *If $-1 \leq a \leq 1$, $b > 0$, and $0 \leq c < 1$, then $p_c(t) = 0$ for some $t \in [0, 1)$.*

Proof. Since b is positive, so as α . Then for every $c \in (0, 1)$,

$$p_c(c^{1/2}) = (2\alpha - 1)c - 2\alpha c^{1/2} + c = 2\alpha(c - c^{1/2}) < 0.$$

But $p_c(0) = c > 0$; so by Proposition 2.2.1, there exists $t \in (0, c^{1/2})$ with $p_c(t) = 0$.

We now apply Lemma 2.3.1 to show that for every $c \in [0, 1)$ there exists $t \in [0, 1)$ such that $p_c(t) = 0$. Q.E.D.

Proposition 2.3.6 *Let $-1 \leq a \leq 1$, $b \in \mathbf{R}$, and $0 \leq c \leq 1$. Then $p_c(t) = 0$ for some t in $[0, 1]$.*

Proof. Given $c \in [0, 1]$, since $p_c(0) = c$ and $p_c(1) = c - 1$, we see from Proposition 2.2.1 that if $0 < c < 1$, then there exists $t \in (0, 1)$ such that $p_c(t) = 0$. Clearly, $p_c(t) = 0$ has a solution $t \in [0, 1]$ if $c = 0$ or $c = 1$. But what if, as can happen in the constructive context, c is close to, but not necessarily distinguishable from, one of the numbers 0, 1? We consider only the case where $0 \leq c < 1$, as the other case, $0 < c \leq 1$ is handled similarly.

In light of Lemmas 2.3.4 and 2.3.5, we need only deal with what happens when b is also near 0. To this end we use Lemma 2.3.3 to construct a decreasing sequence $(\delta)_{n=1}^\infty$ of positive numbers converging to 0, with $\delta_1 < 1/4$, such that if $|b| < \delta_n$ and

$0 \leq c < \delta_n$, then $|t| < 1/n$ for every solution t of equation (2.1). Given $c \in [0, 1)$, define an increasing binary sequence (λ_n) such that

$$\lambda_n = 0 \Rightarrow |b| < \delta_n \text{ and } c < \delta_{n+1},$$

$$\lambda_n = 1 \Rightarrow |b| > \delta_{n+1} \text{ or } c > \delta_{n+2}.$$

Without loss of generality, we may assume that $\lambda_1 = 0$. If $\lambda_n = 0$, set $t_n = 0$. If $\lambda_n = 1 - \lambda_{n-1}$, then we have three cases to deal with. In the first case, $b < -\delta_{n+1}$ and so, by Lemma 2.3.4, there exists $t_n \in (0, 1)$ such that $p_c(t_n) = 0$; in the second case, $b > \delta_{n+1}$ and so, by Lemma 2.3.5, there exists $t_n \in [0, 1)$ such that $p_c(t_n) = 0$; in the third case, $c > \delta_{n+2}$ and so, by the observation at the start of this proof, there exists a solution t_n of equation (2.1) in $(0, 1)$. In each of these three cases, we set $t_k = t_n$ for all $k \geq n$, and we note that $|t_n| < \frac{1}{n-1}$, since $\lambda_{n-1} = 0$ and therefore $\max\{|b|, c\} < \delta_{n-1}$. This completes the inductive construction of t_n . Since $|t_n| < \frac{1}{n-1}$ for each $n \geq 2$, $(t_n)_{n=1}^\infty$ is a Cauchy sequence and so converges to a limit $t_\infty \in [0, 1]$. If $p_c(t_\infty) \neq 0$, then $\lambda_n = 1$ for all n ; whence $c = 0 = t_\infty$ and $p_c(t_\infty) = p_0(0) = 0$, which is absurd. We conclude that $p_c(t_\infty) = 0$. Q.E.D.

2.4 The main results

We prove in this section our main results, Theorems 2.4.3 and 2.4.5; but first we need the following two lemmas.

Lemma 2.4.1 *Let x and y be unit vectors in H , and let*

$$h(t) = tx + (1-t)y \quad (0 \leq t \leq 1).$$

If $0 \leq t \leq \varepsilon < 1/2$, then $\|h(t)\| > 1 - 2\varepsilon$ and

$$\|\|h(t)\|^{-1} h(t) - y\| < \frac{4\varepsilon}{1 - 2\varepsilon}.$$

Proof. We have

$$\begin{aligned}
 \|h(t)\| &\geq (1-t)\|y\| - t\|x\| \\
 &\geq 1 - \varepsilon - \varepsilon \\
 &= 1 - 2\varepsilon.
 \end{aligned}$$

Since $\|h(t)\| \leq 1$, it follows that $0 \leq 1 - \|h(t)\| < 2\varepsilon$; whence

$$\begin{aligned}
 \|\|h(t)\|^{-1} h(t) - y\| &= \|h(t)\|^{-1} \|tx + (1-t)y - \|h(t)\| y\| \\
 &\leq \frac{1}{1-2\varepsilon} (t\|x-y\| + |1 - \|h(t)\|| \|y\|) \\
 &\leq \frac{1}{1-2\varepsilon} (2\varepsilon + 2\varepsilon) \\
 &= \frac{4\varepsilon}{1-2\varepsilon}.
 \end{aligned}$$

Q.E.D.

Recall that a mapping $f : X \rightarrow Y$ between metric spaces is *sequentially continuous* if, whenever (x_n) is a sequence converging to x in X , we have $f(x_n) \rightarrow f(x)$; and that if f is sequentially continuous and X is complete, then f is *strongly extensional* in the sense that $f(x) \neq f(y)$ implies that $x \neq y$ ([51, Theorem 1]).

Lemma 2.4.2 *Let T be a sequentially continuous operator on a Hilbert space H , let x and y be unit vectors in H such that $\langle Tx, x \rangle \neq \langle Ty, y \rangle$, and write*

$$h(t) = tx + (1-t)y \quad (0 \leq t \leq 1).$$

Then $h(t) \neq 0$ for every $t \in [0, 1]$. Moreover, if T is bounded, then $\inf_{t \in [0, 1]} \|h(t)\| > 0$.

Proof. Let I be the identity operator on H and replace T by $T - \langle Ty, y \rangle I$. We may assume that $\langle Ty, y \rangle = 0$. If $0 < t < 1$, then

$$\left\langle T \left(\frac{t}{1-t} x \right), \frac{t}{1-t} x \right\rangle = \frac{t^2}{(1-t)^2} > 0 = \langle T(-y), -y \rangle;$$

As the quadratic form induced by T is sequentially continuous, $\frac{t}{1-t}x \neq -y$ and therefore $tx + (1-t)y \neq 0$. If either $0 \leq t < 1/3$ or $2/3 < t \leq 1$, then Lemma 2.4.1

yields $\|h(t)\| > 1/3$. Putting these three possibilities for t together, we conclude that $h(t) \neq 0$ for all $t \in [0, 1]$.

Now consider the case where T is bounded and not just sequentially continuous. Choose $\delta > 0$ such that if $\|z\| \leq 1$, $\|z'\| \leq 1$, and $\|z - z'\| < \delta$, then $|\langle Tz, z \rangle - \langle Tz', z' \rangle| < 1/9$. If $t \in [1/4, 3/4]$, then

$$\left\langle T \left(\frac{t}{1-t} x \right), \frac{t}{1-t} x \right\rangle = \frac{t^2}{(1-t)^2} \geq \frac{1}{9},$$

so as $\langle T(-y), -y \rangle = 0$, we have $\left\| \frac{t}{1-t} x + y \right\| \geq \delta$ and therefore

$$\|h(t)\| \geq (1-t)\delta \geq \frac{\delta}{4}.$$

Taking this with the cases $0 \leq t < 1/3$ and $2/3 < t \leq 1$, we see that $\|h(t)\| \geq \min\{1/3, \delta/4\}$. Q.E.D.

We now prove our first main result.

Theorem 2.4.3 *Let T be a selfadjoint operator on a Hilbert space H , let x, y be unit vectors in H such that $\langle Tx, x \rangle \neq \langle Ty, y \rangle$, and let*

$$h(t) = tx + (1-t)y \quad (0 \leq t \leq 1).$$

The for each $c \in [0, 1]$ there exists $t \in [0, 1]$ such that

$$\langle T(\|h(t)\|^{-1} h(t)), \|h(t)\|^{-1} h(t) \rangle = c\langle Tx, x \rangle + (1-c)\langle Ty, y \rangle.$$

Proof. Let x and y be unit vectors in H and $0 \leq c \leq 1$. We need to find $t \in [0, 1]$ such that

$$\langle T(\|h(t)\|^{-1} h(t)), \|h(t)\|^{-1} h(t) \rangle = c\langle Tx, x \rangle + (1-c)\langle Ty, y \rangle. \quad (2.2)$$

Since $\langle Tx, x \rangle \neq \langle Ty, y \rangle$, we can find complex numbers λ and μ such that

$$\langle (\lambda T + \mu I)x, x \rangle = 1 \quad \text{and} \quad \langle (\lambda T + \mu I)y, y \rangle = 0;$$

in that case

$$c = \lambda (c \langle Tx, x \rangle + (1 - c) \langle Ty, y \rangle) + \mu,$$

so if there exists $t \in [0, 1]$ such that

$$\langle (\lambda T + \mu I) (\|h(t)\|^{-1} h(t)), (\|h(t)\|^{-1} h(t)) \rangle = c,$$

then equation (2.2) holds. Thus we need only consider the case where

$$\xi = \langle Tx, x \rangle = 1 \quad \text{and} \quad \eta = \langle Ty, y \rangle = 0.$$

For convenience, write $a = \operatorname{Re} \langle x, y \rangle$ and $b = \operatorname{Re} \langle Tx, y \rangle$. Using routine computation with inner products, we see that

$$c = \langle T (\|h(t)\|^{-1} h(t)), \|h(t)\|^{-1} h(t) \rangle = \|h(t)\|^{-2} \langle Th(t), h(t) \rangle$$

if and only if $p_c(t) = 0$. It follows from Proposition 2.3.6 that there exists $t \in [0, 1]$ such that $c \|h(t)\|^2 = \langle Th(t), h(t) \rangle$. Since T , being selfadjoint, is sequentially continuous ([26, Theorem 4]), it follows from this and Lemma 2.4.2 that equation (2.2) holds with $z = \|h(t)\|^{-1} h(t)$. Q.E.D.

Corollary 2.4.4 *Let T be a bounded operator with an adjoint T^* on a Hilbert space H , and x, y unit vectors such that $\langle Tx, x \rangle \neq \langle Ty, y \rangle$ and $\langle Tx, y \rangle \neq \langle T^*x, y \rangle$. Then for each $c \in [0, 1]$ there exists a unit vector $z \in H$ such that*

$$\langle Tz, z \rangle = c \langle Tx, x \rangle + (1 - c) \langle Ty, y \rangle.$$

Proof. Write $T = A + iB$ with $A = \frac{1}{2}(T + T^*)$ and $B = \frac{1}{2i}(T - T^*)$. Then $\langle Bx, y \rangle \neq 0$ so (as on page 22, Step (ii)) there exists a complex number γ such that $|\gamma| = 1$ and $\operatorname{Re} \langle B(\gamma x), y \rangle = 0$. Set

$$h_\gamma(t) = t(\gamma x) + (1 - t)y.$$

Then (as on page 317 of [46]) $\langle Bh_\gamma(t), h_\gamma(t) \rangle = 0$ for all t ; so (as on page 22, Step (iv)) $\langle Th_\gamma(t), h_\gamma(t) \rangle = \langle Ah_\gamma(t), h_\gamma(t) \rangle$. Also, $\langle Tx, x \rangle = \langle A(\gamma x), \gamma x \rangle$ and $\langle Ty, y \rangle = \langle Ay, y \rangle$. Applying Theorem 2.4.3 with A replacing T and γx replacing x , for each $c \in [0, 1]$ we obtain $t \in [0, 1]$ such that

$$\begin{aligned} \langle T(\|h_\gamma(t)\|^{-1} h_\gamma(t)), \|h_\gamma(t)\|^{-1} h_\gamma(t) \rangle &= \langle A(\|h_\gamma(t)\|^{-1} h_\gamma(t)), \|h_\gamma(t)\|^{-1} h_\gamma(t) \rangle \\ &= c \langle A(\gamma x), \gamma x \rangle + (1 - c) \langle Ay, y \rangle \\ &= c \langle Tx, x \rangle + (1 - c) \langle Ty, y \rangle. \end{aligned} \quad \text{Q.E.D.}$$

We now prove the second main result of this chapter.

Theorem 2.4.5 *If T is a bounded operator on a Hilbert space with an adjoint, then the closure of $W(T)$ is convex.*

Proof. Given that $\varepsilon > 0$ and $c \in [0, 1]$, we seek a unit vector z such that

$$|\langle Tz, z \rangle - c\langle Tx, x \rangle - (1 - c)\langle Ty, y \rangle| < \varepsilon. \quad (2.3)$$

If $|\langle Tx, x \rangle - \langle Ty, y \rangle| < \varepsilon$, we may take $z = x$. Thus we may assume that $\langle Tx, x \rangle \neq \langle Ty, y \rangle$.

We first consider the case where T is selfadjoint. Define a function $f : [0, 1] \rightarrow \mathbf{R}$ by

$$f(t) = |\langle T(\|h(t)\|^{-1} h(t)), \|h(t)\|^{-1} h(t) \rangle - c\langle Tx, x \rangle - (1 - c)\langle Ty, y \rangle|.$$

Since T is bounded, it is straightforward to show, using Lemma 2.4.2, that f is uniformly continuous on $[0, 1]$; so

$$m = \inf\{f(t) : 0 \leq t \leq 1\}$$

exists. It follows from Theorem 2.4.3 that if $m > 0$, then $\neg(\langle Tx, x \rangle \neq \langle Ty, y \rangle)$ and therefore $\langle Tx, x \rangle = \langle Ty, y \rangle$; whence, trivially, $m = 0$, a contradiction. We conclude that $m = 0$; whence equation (2.3) holds with $z = \|h(t)\|^{-1} h(t)$ for some $t \in [0, 1]$.

We now consider the general case. Write $T = A + iB$, where $A = \frac{1}{2}(T + T^*)$ and $B = \frac{1}{2i}(T - T^*)$ are bounded selfadjoint operators, and let $\varepsilon > 0$. Noting that

$$\langle Ax, x \rangle = \langle Tx, x \rangle \neq \langle Ty, y \rangle = \langle Ay, y \rangle$$

and that A is bounded, we see from Lemma 2.4.2 that

$$0 < r = \inf \{ \|h(t)\| : 0 \leq t \leq 1 \}.$$

Either $\operatorname{Re} \langle Bx, y \rangle \neq 0$ or $|\operatorname{Re} \langle Bx, y \rangle| < r^2 \varepsilon$. In the first case we apply Corollary 2.4.4 to obtain a unit vector z such that equation (2.2), and therefore equation (2.3), holds. In the second case we apply the first part of the proof with T replaced by A to obtain $t \in [0, 1]$ such that

$$|\langle A(\|h(t)\|^{-1}h(t)), \|h(t)\|^{-1}h(t) \rangle - c\langle Ax, x \rangle + (1-c)\langle Ay, y \rangle| < \frac{1}{2}\varepsilon.$$

Since

$$|\langle Bh(t), h(t) \rangle| = 2t(1-t) |\operatorname{Re} \langle Bx, y \rangle| \leq \frac{1}{2}r^2\varepsilon,$$

$\langle Tx, x \rangle = \langle Ax, x \rangle$, and $\langle Ty, y \rangle = \langle Ay, y \rangle$ it follows that

$$\begin{aligned} & |\langle T(\|h(t)\|^{-1}h(t)), \|h(t)\|^{-1}h(t) \rangle - c\langle Tx, x \rangle + (1-c)\langle Ty, y \rangle| \\ & \leq |\langle A(\|h(t)\|^{-1}h(t)), \|h(t)\|^{-1}h(t) \rangle - c\langle Ax, x \rangle + (1-c)\langle Ay, y \rangle| \\ & \quad + \|h(t)\|^{-2} |\langle Bh(t), h(t) \rangle| \\ & < \frac{1}{2}\varepsilon + r^{-2} \frac{1}{2}r^2\varepsilon \\ & = \varepsilon. \end{aligned}$$

Since ε is arbitrary, the required conclusion follows. Q.E.D.

2.5 A limiting example

Having proved our main results we now turn to answering some questions concerning the best we can hope for in a constructive context. In particular, is it possible to

remove (or at least weaken) some of the hypotheses of Theorems 2.4.3? Perhaps it is not necessary and does not bear any great practical value but it is interesting to know that under certain reasonable conditions we can remove the hypothesis that $\langle Tx, x \rangle \neq \langle Ty, y \rangle$ from Theorem 2.4.3. To show that we can do this, we need a lemma that examines the behaviour of the roots of $p_c(t)$ in $[0, 1]$ when b is large and positive.

Lemma 2.5.1 *For each positive integer n there exists $K_n > 0$, independent of the parameters a and c of p_c , such that if $b > K_n$, then $0 \leq t_- \leq 1/n$.*

Proof. Noting that $\alpha \geq b - 2$, we see that if b is large and positive, then so are α and $2\alpha - 1$, and also $\alpha^2 - 2\alpha c + c < \alpha^2$; whence

$$t_- = \frac{\alpha - \sqrt{\alpha^2 - 2\alpha c + c}}{2\alpha - 1} > \frac{\alpha - \alpha}{2\alpha - 1} = 0.$$

Furthermore, since for such b ,

$$\alpha^2 - 2\alpha c + c \geq \alpha^2 - 2\alpha c^{1/2} + c = (\alpha - c^{1/2})^2,$$

we have

$$t_- < \frac{\alpha - (\alpha - c^{1/2})}{2\alpha - 1} = \frac{c^{1/2}}{2\alpha - 1} \leq \frac{1}{2b - 5}.$$

Thus it is enough to set $K_n = \frac{n+5}{2}$. Q.E.D.

Returning to Theorem 2.4.3, we remove the assumption that $\langle Tx, x \rangle \neq \langle Ty, y \rangle$, and instead assume that $\langle Tx, y \rangle \neq \langle Ty, y \rangle \langle x, y \rangle$. For convenience, write

$$\delta = |\langle Tx, y \rangle - \langle Ty, y \rangle \langle x, y \rangle| > 0,$$

$\xi = \langle Tx, x \rangle$, and $\eta = \langle Ty, y \rangle$. If necessary replacing T by $T - \eta I$, we may assume that $\eta = 0$. Fix $c \in [0, 1]$. Take K_n as in Lemma 2.5.1, and construct an increasing binary sequence $(\lambda_n)_{n=1}^\infty$ such that

$$\lambda_n = 0 \Rightarrow |\xi| < K_n^{-1} \delta,$$

$$\lambda_n = 1 \Rightarrow |\xi| > K_{n+1}^{-1} \delta.$$

If $\lambda_2 = 1$, then we are back in the case already covered by Theorem 2.4.3; so we may assume that $\lambda_2 = 0$. If $\lambda_n = 0$, set $z_n = y$. If $\lambda_n = 1 - \lambda_{n-1}$, then $n \geq 3$ and we can choose $\gamma \in \mathbb{C}$ such that $|\gamma| = 1$ and

$$b = \operatorname{Re} (\xi^{-1} \langle T(\gamma x), y \rangle) = |\xi|^{-1} \delta > K_n.$$

Writing $h_\gamma(t) = t(\gamma x) + (1-t)y$, and applying Lemma 2.5.1 with $\xi^{-1}T$ replacing T and γx replacing x , we compute $t \in [0, 1/n]$ such that

$$z = \|h_1(t)\|^{-1} h_1(t)$$

satisfies

$$\langle Tz, z \rangle = c \langle T(\gamma x), \gamma x \rangle = c\xi.$$

Setting $z_k = z$ for all $k \geq n$, we see from Lemma 2.4.1 that

$$\|h(t)\| > 1 - \frac{2}{n} \geq \frac{1}{3}$$

and $\|z_k - y\| \leq \frac{4}{n-2}$. This completes the construction of a sequence (z_n) of unit vectors in H .

Since $\|z_m - z_n\| \leq \frac{8}{n-2}$ whenever $m \geq n \geq 3$, (z_n) is a Cauchy sequence and therefore converges to a unit vector $z_\infty \in H$. Suppose that $\langle Tz_\infty, z_\infty \rangle \neq c\xi$. Then $\lambda_n = 0$ for all n ; whence $\xi = 0$, $z_\infty = y$, and

$$\langle Tz_\infty, z_\infty \rangle = \langle Ty, y \rangle = 0 = c\xi,$$

a contradiction. It follows that $c\xi = \langle Tz_\infty, z_\infty \rangle \in W(T)$.

Let us now look at a consequence of the classical Intermediate Value Theorem (which, as we noted in Chapter 1, is equivalent to LLPO). Let T be a selfadjoint operator on a Hilbert space H , and let x and y be unit vectors in H . Write $\xi = \langle Tx, x \rangle$, $a = \operatorname{Re} \langle x, y \rangle$, $b = \operatorname{Re} \langle Tx, y \rangle$, and

$$h(t) = tx + (1-t)y \quad (0 \leq t \leq 1).$$

As before, we may assume that $\langle Ty, y \rangle = 0$. Routine computation with inner products shows that

$$\langle T(\|h(t)\|^{-1}h(t)), \|h(t)\|^{-1}h(t) \rangle = c\langle Tx, x \rangle + (1-c)\langle Ty, y \rangle \quad (2.4)$$

if and only if

$$p(t) = (2c(1-a)\xi + 2b - \xi)t^2 - (2c(1-a)\xi + 2b)t + c\xi = 0.$$

Since p is continuous and

$$p(0)p(1) = c(c-1)\xi^2 \leq 0,$$

the classical Intermediate Value Theorem implies that there exists $t \in [0, 1]$ such (2.4) holds. Thus under LLPO we can remove the hypothesis $\langle Tx, x \rangle \neq \langle Ty, y \rangle$ from Theorem 2.4.3.

Studying the proof of Theorem 2.4.5 reveals that the convexity of $W(T)$ can be established provided that for any complex numbers z we can compute a complex number γ with $|\gamma| = 1$ such that $\operatorname{Re} \gamma z = 0$. We can do this with the aid of LLPO as follows. Under LLPO, either $\operatorname{Re} z \geq 0$ or $\operatorname{Re} z \leq 0$. Thus the continuous function f defined on $[0, \pi]$ by $f(t) = \operatorname{Re}(e^{it}z)$ satisfies $f(0)f(\pi) \leq 0$. An application of the classical Intermediate Value Theorem shows that there exists $\tau \in [0, \pi]$ such that $f(\tau) = 0$; so $\operatorname{Re} \gamma z = 0$ where $\gamma = e^{i\tau}$.

We have already shown, in Brouwerian Example 1.6.2 of Chapter 1, that the convexity of the numerical range of a selfadjoint operator on a 2-dimensional Hilbert space implies LLPO_3 and so is essentially nonconstructive. To end the chapter, we now show that statement:

If T is a selfadjoint operator on a 2-dimensional complex Hilbert space H , if x, y are unit vectors in H such that $c\langle Tx, x \rangle + (1-c)\langle Ty, y \rangle$ belongs to $W(T)$ for each $c \in [0, 1]$, and if $h(t) = tx + (1-t)y$ for each $t \in [0, 1]$, then for each $c \in [0, 1]$ there exists $t \in [0, 1]$ such that

$$\langle T(\|h(t)\|^{-1}h(t)), \|h(t)\|^{-1}h(t) \rangle = c\langle Tx, x \rangle + (1-c)\langle Ty, y \rangle$$

implies LLPO.

Brouwerian Example 2.5.2 Let H be a 2-dimensional Hilbert space, and x, y orthogonal unit vectors in H . Given a binary sequence $(a_n)_{n=1}^{\infty}$ with at most one term equal to 1, define

$$\xi = \sum_{n=1}^{\infty} \frac{a_{2n} + a_{2n+1}}{n^2}.$$

With K_n as in Lemma 2.5.1, define a selfadjoint operator T on H by

$$\begin{aligned} Ty &= \left(\sum_{n=1}^{\infty} a_{2n} \xi^2 + \sum_{n=1}^{\infty} a_{2n+1} K_n \xi \right) x, \\ Tx &= \xi x + \left(\sum_{n=1}^{\infty} a_{2n} \xi^2 + \sum_{n=1}^{\infty} a_{2n+1} K_n \xi \right) y. \end{aligned}$$

Given $c \in [0, 1]$, consider the problem of finding a unit vector $z \in H$ such that

$$\langle Tz, z \rangle = c \langle Tx, x \rangle + c \langle Ty, y \rangle = c\xi. \quad (2.5)$$

Writing $z = \lambda y + \mu x$, with $|\lambda|^2 + |\mu|^2 = 1$, we see that equation (2.5) can be reduced to

$$|\mu|^2 \xi + 2 \operatorname{Re} \left(\lambda \mu^* \left(\sum_{n=1}^{\infty} a_{2n} \xi^2 + \sum_{n=1}^{\infty} a_{2n+1} K_n \xi \right) \right) = c\xi,$$

which is satisfied by taking $\lambda = \sqrt{1-c}$ and $\mu = i\sqrt{c}$. Thus $c \langle Tx, x \rangle + (1-c) \langle Ty, y \rangle$ belongs to $W(T)$.

Next, writing $b = \operatorname{Re} \langle Tx, y \rangle$ and taking $c = 1/2$, we see that equation (2.5) becomes

$$2bt^2 - (\xi + 2b)t + \frac{1}{2}\xi = 0,$$

which, if $\xi \neq 0$, can be rewritten as

$$\frac{2b}{\xi} t^2 - \left(1 + \frac{2b}{\xi} \right) t + \frac{1}{2} = 0 \quad (2.6)$$

If $a_{2N} = 1$, then equation (2.6) becomes

$$2\xi t^2 - (1 + 2\xi)t + \frac{1}{2} = 0,$$

whose solution in $[0, 1]$ is

$$t_- = \frac{1 + 2\xi - \sqrt{1 + 4\xi^2}}{4\xi}.$$

On the other hand, if $a_{2N+1} = 0$, then equation (2.6) becomes

$$2K_n t^2 - (1 + 2K_n)t + \frac{1}{2} = 0,$$

whose only solution in $[0, 1]$ is

$$t_- = \frac{1 + 2K_n - \sqrt{1 + 4K_n^2}}{4K_n}.$$

We proceed with our analysis as follows. First, noting that

$$\lim_{\xi \rightarrow 0^+} \frac{1 + 2\xi - \sqrt{1 + 4\xi^2}}{4\xi} = \lim_{\xi \rightarrow 0^+} \frac{1}{1 + 2\xi + \sqrt{1 + 4\xi^2}} = \frac{1}{2},$$

we find $r > 0$ such that if $|\xi| < r$, then

$$\frac{1 + 2\xi - \sqrt{1 + 4\xi^2}}{4\xi} \geq \frac{3}{8}.$$

Secondly, observing that

$$\lim_{n \rightarrow \infty} \frac{1 + 2K_n - \sqrt{1 + 4K_n^2}}{4K_n} = \lim_{n \rightarrow \infty} \frac{1}{1 + 2K_n + \sqrt{1 + 4K_n^2}} = 0,$$

so we compute a positive integer N such that

$$\frac{1 + 2K_n - \sqrt{1 + 4K_n^2}}{4K_n} < \frac{1}{8}$$

for all $n \geq N$. Now suppose that equation (2.4) has a solution $t = \tau \in [0, 1]$, and consider a_N . If $a_k = 1$ for some $k \leq N$, then either $a_n = 0$ for all even n or else $a_n = 0$ for all odd n ; so we may assume that $a_k = 0$ for all $k \leq N$; we may also

assume that $|\xi| < r$. Either $\tau > 1/8$ or $\tau < 3/8$. Consider the first case, and suppose that $a_{2n+1} = 1$ for some n with $2n + 1 > N$. Then $b = K_n \xi \neq 0$, so

$$\tau = \frac{1 + 2K_n - \sqrt{1 + 4K_n^2}}{4K_n} < \frac{1}{8},$$

a contradiction; hence $a_k = 0$ for all odd $k > N$ and therefore for all odd k . Now consider the case $\tau < 3/8$, and suppose that $a_{2n} = 1$ for some n with $2n > N$. Then $b = \xi^2 \neq 0$, and

$$t = \frac{1 + 2\xi - \sqrt{1 + 4\xi^2}}{4\xi} \geq \frac{3}{8},$$

since $|\xi| < r$; this contradiction ensures that $a_k = 0$ for all even $k > N$ and therefore for all even k . ///

Chapter 3

A Spectral Mapping Theorem

3.1 Introduction

Recall that a *unital Banach algebra* B is a complex algebra with a multiplicative identity e and a norm $\|\cdot\|$ that satisfy the following conditions:

- $\|e\| = 1$,
- B is a Banach space relative to the norm $\|\cdot\|$,
- $\|xy\| \leq \|x\|\|y\|$ for all x and y in B .

Much of the elementary classical theory of Banach algebras carries over virtually unchanged into the constructive setting. Nevertheless, there are substantial problems with even some of the elementary aspects of the theory, such as the compactness of the spectrum (see Brouwerian Example 1.6.4 of Chapter 1). Relatively little work has been carried out in constructive Banach algebra theory, despite Bishop's insightful and technically demanding developments in [5] and [10], and the recent work of Bridges in [18, 19, 21] and Bridges and the author in [25].

Before going further, we need to define some terms that will be used frequently. Throughout this chapter, we reserve the letter B to denote a unital complex Banach

algebra. For each $a \in B$ we write

$$R(a) = \{\lambda \in \mathbf{C} : a - \lambda e \text{ has two-sided inverse}\}$$

to denote the *resolvent* set of a , and

$$\sigma(a) = \{\lambda \in \mathbf{C} : \forall \lambda' \in R(a). (\lambda \neq \lambda')\}$$

for the *spectrum* of a .

Our main goal in this chapter is to analyse constructively the classical **Spectral Mapping Theorem**:

If p is a monic polynomial of degree at least 1, then $\sigma(p(a)) = p(\sigma(a))$.

A typical classical proof of this theorem, as in [52, page 82], relies on the decomposition of $p(z) - \lambda$, where $\lambda \in \mathbf{C}$, into linear factors, and the observation that $p(a) - \lambda e$ has no inverse if and only if at least one of its corresponding factors has no inverse. This type of proof does not work constructively, since we cannot be sure of deciding which of the factors has no inverse.¹ In this chapter, guided by the classical proof, we prove that $p(\sigma(a)) \subset \sigma(p(a))$, provide a constructive substitute for the opposite inclusion, and show that our result is the best possible.

3.2 The inclusion $p(\sigma(a)) \subset \sigma(p(a))$

Our proof of the first inclusion, $p(\sigma(a)) \subset \sigma(p(a))$, of the Spectral Mapping Theorem is based on some elementary constructive semigroup theory.

An *inequality* on a set X is a binary relation \neq on X with the following properties:

$$x \neq y \Rightarrow \neg(x = y),$$

$$x \neq y \Rightarrow y \neq x.$$

¹This is a simple consequence of the fact that, given two real numbers x and y whose product is 0, we may not be able to decide which of x or y equals 0 ([16, page 44]).

If also

$$\forall x, y \in X (x = y \vee x \neq y),$$

we call the inequality *discrete*. The *complement* of a subset Y of a set X with an inequality is the set

$$\sim Y = \{x \in X : \forall y \in Y (x \neq y)\}.$$

The standard inequality on a normed space X is given by

$$x \neq y \quad \text{if and only if} \quad \|x - y\| > 0.$$

In this case the statement $x \neq y$ is equivalent to $\neg(x = y)$ if and only if we assume MP. In the case where $X = \mathbf{R}$, it is shown in [10, page 26, (2.17)] that the standard inequality satisfies the special property

$$(x \neq y) \Rightarrow \forall x \in \mathbf{R} (x \neq z \vee z \neq y).$$

Let S be a semigroup with an identity e and an inequality \neq , and let $\text{inv}(S)$ be the set of invertible elements of S . The inequality on S is said to be *quasi-discrete* if

$$\forall x \in S (x \neq e \vee x \in \text{inv}(S)).$$

A discrete inequality on a semigroup S is clearly quasi-discrete. A more interesting example of a quasi-discrete inequality occurs when S is the multiplicative semigroup of a Banach algebra B with an identity e . In that case, for each $x \in B$ either $0 < \|e - x\|$ or $\|e - x\| < 1$, so either $x \neq e$ or x has a two-sided inverse.

We write $x \notin S$ to mean $\neg(x \in S)$. Furthermore, we assume that the operations of multiplication on the left and on the right are *strongly extensional*, in the sense that, for example, if $ax \neq bx$, then $a \neq b$.

Lemma 3.2.1 *Let S be a semigroup with identity e and a quasi-discrete inequality, and let a be an element of S such that $a \notin \text{inv}(S)$. Then*

$$\forall x \in S (ax \neq e \vee e \in \sim Sa)$$

and

$$\forall x \in S (xa \neq e \vee e \in \sim aS).$$

Proof. Fix $x \in S$. Either $ax \neq e$ or ax is invertible. In the latter case, a has right inverse $x(ax)^{-1}$. Moreover, for each $s \in S$, either $e \neq sa$ or sa is invertible. The second alternative implies that a has left inverse $(sa)^{-1}s$; whence a has inverses on either side and therefore a two-sided inverse, which contradicts the assumption that $a \notin \text{inv}(S)$. We conclude that if ax is invertible, then $e \in \sim Sa$.

The second part of the lemma is proved similarly. Q.E.D.

Proposition 3.2.2 *If $a \in \sim \text{inv}(B)$, then $ax \neq e$ for all $x \in B$.*

Proof. Either $ax \neq e$ or $\|e - ax\| < 1$. In the latter case, we have

$$e = ax((ax)^{-1}) = a(x(ax)^{-1}),$$

which implies that a is invertible, a contradiction. Q.E.D.

Proposition 3.2.3 *Let S be a semigroup with identity and a quasi-discrete inequality and, let a be an element of S such that $a \notin \text{inv}(S)$. Then*

$$aS \cap Sa \subset \sim \text{inv}(S).$$

Proof. Suppose that $a \notin \text{inv}(S)$ and $x \in aS \cap Sa$, and consider $y \in \sim \text{inv}(S)$. Then there exist $b, c \in S$ such that $ab = x = ca$. By Lemma 3.2.1, either $y^{-1}x = (y^{-1}c)a \neq e$ and therefore $x \neq y$, or else, as we may assume, $e \in \sim aS$. Then

$$e \neq a(by^{-1}) = (ab)y^{-1} = xy^{-1},$$

and so $y = ey \neq (xy^{-1})y = x$. Q.E.D.

Corollary 3.2.4 *If S is a semigroup with identity and a quasi-discrete inequality, and a is an element of S such that $a \notin \text{inv}(S)$, then $a \in \sim\text{inv}(S)$.*

Proof. Apply Proposition 3.2.3, noting that $a \in aS \cap Sa$. Q.E.D.

We are now in a position to prove the first half of the Spectral Mapping Theorem.

Theorem 3.2.5 *Let a be an element of unital Banach algebra, and p a nonconstant polynomial over \mathbf{C} . Then $p(\sigma(a)) \subset \sigma(p(a))$.*

Proof. Let $\gamma \in p(\sigma(a))$. Then there exists $\lambda \in \sigma(a)$ such that $\gamma = p(\lambda)$. If $p(z) = c_0 + \cdots + c_n z^n$, then

$$p(a) - p(\lambda) = c_1(a - \lambda e) + \cdots + c_n(a^n - \lambda^n e).$$

Note that for each k ,

$$\begin{aligned} a^k - \lambda^k e &= (a^{k-1} + a^{k-2}\lambda + \cdots + \lambda^{k-1})(a - \lambda e) \\ &= (a - \lambda e)(a^{k-1} + a^{k-2}\lambda + \cdots + \lambda^{k-1}). \end{aligned}$$

So $p(a) - p(\lambda) \in (a - \lambda e)B \cap B(a - \lambda e)$. Since $a - \lambda e \notin \text{inv}(B)$, it follows from Proposition 3.2.3 that $p(a) - p(\lambda) \in \sim\text{inv}(B)$ and hence that $p(\lambda) = \gamma \in \sigma(p(a))$. Q.E.D.

3.3 The inclusion $\sigma(p(a)) \subset p(\sigma(a))$

We now aim to show that the second inclusion of the Spectral Mapping Theorem holds constructively under some additional hypotheses on the element a of B .

We need a digression into subsets of \mathbf{R}^N . Recall that a subset S of a metric space (X, ρ) is:

- *located* if the *distance*

$$\rho(x, S) = \inf \{\rho(x, s) : s \in S\}$$

exists for each x in X ;

- ***totally bounded*** if for each $\varepsilon > 0$ there exists a finite ε -approximation Y to X , in the sense that Y is a subset of X such that for each x in X there exists y in Y with $\rho(x, y) < \varepsilon$;
- ***locally totally bounded*** if each bounded subset is contained in a totally bounded set in X ;
- ***compact*** if it is complete and totally bounded.

A compact set is locally totally bounded, and a locally totally bounded set is located. Hence compact sets are located. Furthermore, any located subset of a locally totally bounded set is locally totally bounded ([30, page 33, (4.11)]).

We write an inequality of the form $\rho(x, S) < r$ to mean that there exists $s \in S$ such that $\rho(x, s) < r$; in this usage we do not require the *distance expression* $\rho(x, S)$ to exist as a real number. Likewise, we write

- $\rho(x, S) \leq r$ to mean that $\rho(x, s) < r + \varepsilon$ for each $r > 0$ and each $s \in S$; and
- $\rho(x, S) > 0$ to mean that there exists $r > 0$ such that $\rho(x, s) \geq r$ for each $s \in S$.

With these interpretations, we define the ***metric complement of S in X*** to be

$$X - S = \{x \in X : \rho(x, S) > 0\}.$$

For convenience if ρ_1 and ρ_2 are distance expressions, we denote their (possibly fictitious) supremum by $\rho_1 \vee \rho_2$. This notation is intended to capture the equality

$$\{r \in \mathbf{R} : r > \rho_1 \vee \rho_2\} = \{r \in \mathbf{R} : r > \rho_1 \text{ and } r > \rho_2\}.$$

Let Ω be a subset of a metric space X . We say that

- a subset K of Ω is *well contained* in Ω , written $K \subset\subset \Omega$, if there exists $r > 0$ such that if $\rho(x, K) \leq r$, then $x \in \Omega$;
- K *approximates Ω internally to within $\varepsilon > 0$* if $K \subset\subset \Omega$ and $\rho(x, \partial\Omega) < \varepsilon$ for each $x \in \Omega - K$ (where, as usual, $\partial\Omega$ denotes the boundary $\Omega \cap \sim\Omega$ of Ω);
- Ω is *approximated internally by sets of type T* if for every $\varepsilon > 0$ it can be approximated internally by a set of type T ;
- Ω is *coherent* if $\Omega = -(\sim\Omega)$.

If Ω is approximated internally by sets of type T , we write K^ε to denote a set of type T that approximates Ω within ε . The most important types we deal with are *located* and *compact*.

In order to discuss the second inclusion of the Spectral Mapping Theorem, we state, without proof, a number of results from [32].

Proposition 3.3.1 *Let Ω be a subset of \mathbf{R}^N whose interior Ω° is the metric complement of a located set $L \subset \sim\Omega$. Then Ω is approximated internally by located subsets.*

Proposition 3.3.2 *Let Ω be a subset of \mathbf{R}^N that is approximated internally by located subsets. Then $\sim\Omega^\circ$ is located, and Ω° is coherent.*

A subset K of a Banach space has the *boundary crossing property* if $x \in K$ and $y \in \sim K$, then for each $\varepsilon > 0$ there exists a point $z \in K$ whose distance from the segment joining x and y is less than ε .

Proposition 3.3.3 *A located subset K of a Banach space has the boundary crossing property.*

Proof. Let x be an element of the Banach space. If K is located, then for each $\varepsilon > 0$ either $\rho(x, K) > 0$ or $\rho(x, K) < \varepsilon$. In the former case, $x \in -K \subset \sim K$. In the latter case, there exists $y \in K$ such that $\rho(x, y) < \varepsilon$. It follows that $K \cup -K$ is dense and so as $K \cup \sim K$. An application of Proposition 8 of [32] completes the proof. Q.E.D.

Proposition 3.3.4 *If Ω be a coherent subset of a Banach space such that both $\partial\Omega$ and $\sim\Omega$ are located, then Ω is located.*

Proof. Notice that $\sim\Omega \cup -(\sim\Omega)$ is dense. The coherence of Ω means $-(\sim\Omega) \subset \Omega$, so $\Omega \cup \sim\Omega$ is dense. The conclusion follows from this and Proposition 11 of [32]. Q.E.D.

We now apply the foregoing results of this section to the spectrum and resolvent set of an element of our unital Banach algebra.

Proposition 3.3.5 *The following are equivalent conditions on an element a of a unital Banach algebra B .*

- (i) $\sigma(a)$ is compact and $R(a)$ is coherent.
- (ii) $R(a)$ is approximated internally by located sets.

If either, and hence each, of these conditions holds, then a necessary and sufficient condition for $R(a)$ to be located is that $\partial\sigma(a)$ be located.

Proof. The equivalence of (i) and (ii) can be established by taking $\Omega = R(a)$ in Propositions 3.3.1 and 3.3.2, and note that if $\sigma(a)$ is located, then, being closed in \mathbf{R}^N , it is compact. Assume (i). Then as $\sigma(a)$ is located, $\sigma(a) \cup -\sigma(a)$ —that is, $\sim R(a) \cup R(a)$ —is dense in \mathbf{R}^N . Since $\partial\sigma(a) = \partial R(a)$, an application of Proposition 3.3.4 shows that $R(a)$ is located if and only if $\partial\sigma(a)$ is located. Q.E.D.

A **border** for a compact subset K of \mathbf{C} is a totally bounded subset Γ of K such that $\overline{B}(z, \rho(z, \Gamma)) \subset K$ for each $z \in K$. The importance of borders for us rests in the following result where

$$m(f, K) = \inf \{|f(z)| : z \in K\}.$$

Proposition 3.3.6 *Let K be a compact subset of \mathbf{C} , Γ be a border for K , and f a differentiable function on K such that $m(f, \Gamma) > m(f, K) = 0$. Then there exists $z \in K$ such that $f(z) = 0$ ([10, page 156, (5.8)]).*

Proposition 3.3.7 *Let Ω be an open subset of \mathbf{C} such that $\sim\Omega$ and $\partial\Omega$ are compact. Then $\partial\Omega$ is a border for $\sim\Omega$.*

Proof. Let $K = \sim\Omega$, fix $z_0 \in K$, and write $r = \rho(z_0, \partial\Omega)$. For each $t \in [0, 1]$ write

$$z_t = tz_0 + (1 - t)z,$$

the line segment joining z and z_0 . Assume $z \in \Omega$, and choose $\delta > 0$ such that $B(z, 3\delta) \subset \Omega$; then $r \geq 3\delta$. An application of Proposition 3.3.3 enables us to compute $\zeta \in \partial\Omega$ and $t \in [0, 1]$ such that $|\zeta - z_t| < \delta$. Since $\zeta \in \overline{K} = K$, we have

$$3\delta \leq \rho(z, K) \leq |\zeta - z| \leq |\zeta - z_t| + |z - z_t| < \delta + |z - z_t|.$$

Hence $|z - z_t| > 2\delta$ and therefore

$$\begin{aligned} |z_0 - \zeta| &\leq |z_0 - z_t| + |\zeta - z_t| \\ &= |z - z_0| - |z - z_t| + |\zeta - z_t| \\ &< r - 2\delta + \delta \\ &= r - \delta, \end{aligned}$$

a contradiction since $|z_0 - \zeta| \geq \rho(z_0, \partial\Omega) = r$. It follows that $\neg(z \in \Omega)$. Since Ω is open, we conclude that $z \in \sim\Omega$. Q.E.D.

Corollary 3.3.8 *Let a be an element of a unital Banach algebra B such that $\sigma(a)$ and $\partial\sigma(a)$ are compact. Then $\partial\sigma(a)$ is a border for $\sigma(a)$.*

Proof. Take $\Omega = R(a)$ in Proposition 3.3.7. Q.E.D.

Before we prove the main result of this section, we state the following well known results. The first of which is a constructive version of the **Fundamental Theorem of Algebra** ([10, page 156, (5.10)]).

Theorem 3.3.9 *If the polynomial $p(z) = a_0z^n + \cdots + a_n$ has degree at least k , then there exist complex numbers z_1, \dots, z_k and a polynomial q such that*

$$p(z) = (z - z_1) \cdots (z - z_k) q(z) \quad (z \in \mathbb{C}).$$

The next result is **Bishop's Lemma** ([10, page 92, (3.8)]), which is trivial in CLASS but very useful in constructive analysis.

Lemma 3.3.10 *Let S be a complete, located subset of a metric space X , and $x \in X$. Then there exists $s \in S$ such that if $\rho(x, s) > 0$, then $\rho(x, S) > 0$.*

Now we have our main result.

Theorem 3.3.11 *Let a be an element of a unital Banach algebra B such that $R(a)$ is approximated internally by located sets and $\partial\sigma(a)$ is located (and hence compact). Let p be a nonconstant monic polynomial over \mathbb{C} , and let λ be an element of $\sigma(p(a)) \cap \sim p(\partial\sigma(a))$. Then $\lambda \in p(\sigma(a))$.*

Proof. In light of Proposition 3.3.5, $\sigma(a)$ is compact and $R(a) = -\sigma(a)$. By Corollary 3.3.8, $\partial\sigma(a)$ is a compact border for $\sigma(a)$. If the degree of p is n , then, by the Fundamental Theorem of Algebra, there exist complex numbers $\lambda_1, \dots, \lambda_n$ such that

$$p(z) - \lambda = (z - \lambda_1) \cdots (z - \lambda_n) \quad (z \in \mathbb{C}).$$

For every $z \in \partial\sigma(a)$, since $p(z) \neq \lambda$, we have $\lambda_k \neq z$ ($1 \leq k \leq n$). By Bishop's Lemma, $\lambda_k \in \mathbf{C} - \partial\sigma(a)$ for each k ; whence $m(p - \lambda, \partial\sigma(a)) > 0$. On the other hand, if

$$\inf_{1 \leq k \leq n} \rho(\lambda_k, \sigma(a)) > 0,$$

then for each k , $\lambda_k \in -\sigma(a) = R(a)$ and so $a - \lambda_k e$ is invertible; whence $p(a) - \lambda e$ is invertible, a contradiction. Thus,

$$\inf_{1 \leq k \leq n} \rho(\lambda_k, \sigma(a)) = 0$$

and therefore $m(p - \lambda, \sigma(a)) = 0$. It now follows from Proposition 3.3.6 that there exists $\zeta \in \sigma(a)$ such that $p(\zeta) - \lambda = 0$. Hence $\lambda \in p(\sigma(a))$. Q.E.D.

The next corollaries follow trivially.

Corollary 3.3.12 *If $\lambda \in p(\partial\sigma(a))$, then $\lambda \in p(\sigma(a))$.*

Corollary 3.3.13 *If $R(a)$ is approximated internally by located sets and $\partial\sigma(a)$ is compact, then $\lambda \in p(\sigma(a))$ for each λ in the dense subset*

$$\sigma(p(a)) \cap (p(\partial\sigma(a)) \cup \sim p(\partial\sigma(a)))$$

of $\sigma(p(a))$.

3.4 A limiting example

We end this chapter with an example that our conditions for the inclusion $\sigma(p(a)) \subset p(\sigma(a))$ are the best we can hope for.

Brouwerian Example 3.4.1 Let B be a unital Banach algebra containing an element a whose resolvent set is the exterior of the closed unit disc D in \mathbf{C} and whose spectrum is that disc. Let $\zeta \in [-1/4, 1/4]$, and define

$$p(z) = (z + 1 - \zeta)(z - 1 - \zeta) \quad (z \in \mathbf{C}).$$

Suppose that $p(a)$ is invertible. Then both $a - (-1 + \zeta)e$ and $a - (1 + \zeta)e$ are invertible, so $-1 + \zeta \in -D$ and $1 + \zeta \in -D$; this implies that $\zeta < 0$ and $\zeta > 0$, which is absurd. Hence $p(a)$ is not invertible, and therefore $\neg(0 \in R(p(a)))$. Since $R(p(a))$ is open, we conclude that $0 \in \sim R(p(a)) = \sigma(p(a))$.

Now suppose that $p(\zeta_0) = 0$ for some $\zeta_0 \in D$. Either $\zeta_0 > -1/2$ or $\zeta_0 < 1/2$. In the first case, if $\zeta_0 \neq 1 + \zeta$, then the quadratic equation $p(z) = 0$ has three distinct roots, which is impossible; so $1 + \zeta = \zeta_0 \in [-1, 1]$ and therefore $\zeta \leq 0$. Similarly, in the second case, $-1 + \zeta = \zeta_0 \in [-1, 1]$ and so $\zeta \geq 0$. Thus the proposition,

If a is an element of a unital Banach algebra such that $R(a)$ is approximated internally by located sets and $\partial\sigma(a)$ is located (and hence compact), and if p is a nonconstant monic polynomial over \mathbf{C} , then $\sigma(p(a)) \subset p(\sigma(a))$.

entails

$$\forall x \in \mathbf{R} (x \geq 0 \vee x \leq 0),$$

an equivalent of LLPO ([30, pages 4 and 53–54]). ///

Chapter 4

Powers of a Hermitian Element

4.1 Introduction

Is the square of a Hermitian element is Hermitian? The answer is classically affirmative, but what is the constructive situation? In this chapter we answer this question by proving, eventually, that positive integral powers of a Hermitian element of a Banach algebra are indeed Hermitian. Our proof depends on an investigation of the character space, the state space, and extreme points.

4.2 Preliminaries

With reference to Chapter 7 of [10], we first recall some facts about the dual space X' of a normed linear space X .

An element f of X' is *normable* if its *norm*

$$\|f\| = \sup \{|f(x)| : \|x\| \leq 1\}$$

exists. If X' is separable, and $(x_n)_{n=1}^{\infty}$ is a dense sequence in the unit ball

$$X'_1 = \{f \in X' : \forall x \in X (|f(x)| \leq \|x\|)\}$$

of X' , then the weak* topology on X' is induced by the *double norm*, defined by

$$\|f\| = \sum_{n=1}^{\infty} 2^{-n} |f(x_n)| \quad (f \in X').$$

Double norms defined by different dense sequences in X are equivalent on X'_1 , and X'_1 is weak* compact. Moreover, for each $x \in X$ the mapping $f \mapsto f(x)$ is uniformly continuous on X'_1 with respect to the double norm.

In the remainder of this thesis, B will denote a complex unital Banach algebra with identity e .

The *state space* of B is the set

$$V_B = \{f \in B' : f(e) = 1 = \|f\|\}.$$

For each $t > 0$ the set

$$V_B^t = \{f \in B' : \|f\| \leq 1, |1 - f(e)| \leq t\}$$

is a t -approximation to V_B .

Proposition 4.2.1 *For all but countably many $t > 0$, V_B^t is a nonempty, weak* compact subset of B' .*

Proof. Since the mapping $f \mapsto |1 - f(e)|$ is uniformly continuous on B'_1 relative to the double norm, we see from Theorem 4.9 of [10, page 98] that for all but countably many $t > 0$, V^t is either empty or weak* compact. An application of Corollary 4.5 of [10, page 341] shows that for such t , V^t is nonempty and therefore weak* compact. Q.E.D.

We say that $t > 0$ is *admissible* if V_B^t is weak* compact. Note that

$$V_B = \bigcap \{V_B^t : t > 0 \text{ is admissible}\},$$

the intersection of a family of nonempty, weak* compact sets that is descending in the sense that if $0 < t' < t$, then $V_B^{t'} \subset V_B^t$. Being the intersection of a family of complete sets, V_B is complete relative to the double norm.

In the remainder of the thesis, when the context is clear we write Σ , V , and V^t to mean Σ_B , V_B , and V_B^t , respectively.

We introduce the following definitions.

- V is ***firm*** if it is compact and $\rho_w(V^t, V) \rightarrow 0$ as $t \rightarrow 0$, where ρ_w is the Hausdorff metric on the set of weak* compact subsets of B'_1 .
- An element x of B is ***Hermitian*** if for each $\varepsilon > 0$ there exists $t > 0$ such that $|\operatorname{Im} f(x)| < \varepsilon$ for all $f \in V^t$; we denote the set of all Hermitian elements of B by $\operatorname{Her}(B)$.
- An element x of B is ***positive*** if for each $\varepsilon > 0$ there exists $t > 0$ such that $\operatorname{Re} f(x) \geq -\varepsilon$ and $|\operatorname{Im} f(x)| < \varepsilon$ for all $f \in V^t$; we then write $x \geq 0$.
- An element f of B' is a ***positive linear functional*** if $f(x) \geq 0$ for each positive element x of B ; we then write $f \geq 0$.

Our main aim in this chapter is to prove the following.

Theorem 4.2.2 *Let a be a Hermitian element of a complex unital Banach algebra B that has firm state space. Then a^n is (i) Hermitian for each positive integer n and (ii) positive for each even positive integer n .*

But first we need some technical results which we establish in the next section.

4.3 Technical results

The following is a constructive version of the **Hahn–Banach Theorem** ([10, page 342]).

Theorem 4.3.1 *Let Y be a linear subset of a separable normed linear space X , and v a nonzero linear functional on Y whose kernel is located in X . Then for each $\varepsilon > 0$ there exists a normable linear functional u on X such that $u(y) = v(y)$ for all y in Y , and $\|u\| \leq \|v\| + \varepsilon$.*

Lemma 4.3.2 *Let E be a finite-dimensional subspace of a normed space X , and a a unit vector in E . Let $0 \leq t < 1$, and let f be a linear functional on E such that $\|f\| \leq 1$ and $|1 - f(a)| \leq t$. Then there exists a normable linear functional ϕ on X such that $\|\phi\| = 1$, $|1 - \phi(a)| \leq 2t$, and $|f(x) - \phi(x)| \leq t\|x\|$ for each $x \in E$.*

Proof. Since E is finite-dimensional, f has a norm, and so $\ker(f)$ is located in E ; whence $\ker(f)$ is finite-dimensional and therefore located in X . By the Hahn-Banach Theorem, there exists an extension $f^\#$ of f to an element of X' such that $1 \leq \|f^\#\| < 1 + t$. Let $\phi = \|f^\#\|^{-1} f^\#$. Then $\phi \in X'$, $\|\phi\| = 1$, and for each $x \in E$,

$$\begin{aligned} |f(x) - \phi(x)| &= \left| f(x) - \|f^\#\|^{-1} f^\#(x) \right| \\ &= \left| 1 - \|f^\#\|^{-1} \right| |f(x)| \\ &\leq \frac{t}{1+t} \|f\| \|x\| \\ &\leq t \|x\|. \end{aligned}$$

Moreover,

$$|1 - \phi(a)| \leq |1 - f(a)| + |f(a) - \phi(a)| \leq 2t \quad \text{Q.E.D.}$$

Lemma 4.3.3 *Suppose that the state space of B is firm. Let A be a Banach subalgebra of B , let $\{x_1, \dots, x_N\}$ be a finitely enumerable subset of A with $x_1 = e$, and let $\varepsilon > 0$. Then there exists an admissible $t > 0$ such that for each $f \in A'$ with $\|f\| \leq 1$ and $|1 - f(e)| \leq t$, there exists $g \in V_A$ with*

$$|f(x_k) - g(x_k)| \leq \varepsilon \quad (1 \leq k \leq N) \quad (4.1)$$

Proof. We first prove the result in the case where $\{x_1, \dots, x_N\}$ is a linearly independent subset spanning a finite-dimensional subspace E of A . Given $\varepsilon > 0$, choose $t \in (0, \varepsilon/2)$ such that

► V_B^t is weak* compact and

► for each $\phi \in V_B^{2t}$ there exists $v \in V_B$ with

$$|\phi(x_k) - v(x_k)| < \frac{\varepsilon}{2} \quad (1 \leq k \leq N). \quad (4.2)$$

Given $f \in A'$ with $\|f\| \leq 1$ and $|1 - f(e)| \leq t$, apply Lemma 4.3.2 to the restriction f^b of f to E , to construct $\phi \in V_B^{2t}$ such that $|f(x) - \phi(x)| \leq t\|x\|$ for each $x \in E$. Choose $v \in V_B$ such that (4.2) holds. The restriction g of v to A belongs to V_A , and for $1 \leq k \leq N$,

$$\begin{aligned} |f(x_k) - g(x_k)| &= |f^b(x_k) - v(x_k)| \\ &\leq |f(x_k) - \phi(x_k)| + |\phi(x_k) - v(x_k)| \\ &\leq t\|x_k\| + \frac{\varepsilon}{2} \\ &\leq \varepsilon, \end{aligned}$$

as we wanted.

It remains to remove the condition that $\{x_1, \dots, x_N\}$ be linearly independent. We proceed by induction, noting that the case $N = 1$ is dealt with by the work of the previous paragraph. Suppose that the desired conclusion holds for all sets of N vectors in A with $N < \nu$, and consider x_1, \dots, x_ν in A . We may assume that $\varepsilon < 1$. Rearranging the indices $1, \dots, \nu$ if necessary, we can find $m \leq \nu$ such that

► x_1, \dots, x_m are linearly independent and span an m -dimensional subspace Y of A , and

► $\rho(x_k, Y) < \varepsilon/4$ for $m+1 \leq k \leq \nu$.

Let

$$c = \sup \left\{ \sum_{j=1}^m |\lambda_j| : \left\| \sum_{j=1}^m \lambda_j x_j \right\| \leq 2 \right\}.$$

By the first part of the proof, we can find an admissible $t > 0$ such that for each $f \in V_A^t$ there exists $g \in V_A$ with

$$|f(x_k) - g(x_k)| < \frac{\varepsilon}{2c} \quad (1 \leq k \leq m).$$

Fix such f and g , consider k with $m+1 \leq k \leq \nu$, and choose $\lambda_1, \dots, \lambda_m \in \mathbb{C}$ such that

$$\left\| x_k - \sum_{j=1}^m \lambda_j x_j \right\| < \frac{\varepsilon}{4}.$$

Then $\left\| \sum_{j=1}^m \lambda_j x_j \right\| \leq 2$, so $\sum_{j=1}^m |\lambda_j| \leq c$. Hence

$$\begin{aligned} |f(x_k) - g(x_k)| &\leq \left| (f - g) \left(\sum_{j=1}^m \lambda_j x_j \right) \right| + \left| (f - g) \left(x_k - \sum_{j=1}^m \lambda_j x_j \right) \right| \\ &\leq \sum_{j=1}^m |\lambda_j| |f(x_j) - g(x_j)| + 2 \left\| x_k - \sum_{j=1}^m \lambda_j x_j \right\| \\ &\leq \sum_{j=1}^m |\lambda_j| \frac{\varepsilon}{2c} + \frac{\varepsilon}{2} \\ &\leq \varepsilon. \end{aligned}$$

Thus (4.1) holds for $N = \nu$, and our induction is complete. Q.E.D.

Lemma 4.3.4 *Let $(K_\lambda)_{\lambda \in L}$ be a nonempty family of totally bounded subsets of a metric space X , and let $K = \bigcap_{\lambda \in L} K_\lambda$. Suppose that for each $\varepsilon > 0$ there exists $\lambda \in L$ such that for each $x \in K_\lambda$ there exists $y \in K$ with $\|x - y\| < \varepsilon$. Then K is totally bounded. If also each K_λ is complete, then K is compact.*

Proof. Given $\varepsilon > 0$, choose $\lambda \in L$ as in the hypotheses. Let $\{x_1, \dots, x_N\}$ be a finite ε -approximation to K_λ , and for each n choose $y_n \in K$ such that $\|x_n - y_n\| < \varepsilon$. Let $y \in K \subset K_\lambda$. Then there exists n such that $\|y - x_n\| < \varepsilon$ and therefore

$$\|y - y_n\| \leq \|y - x_n\| + \|x_n - y_n\| < \varepsilon + \varepsilon = 2\varepsilon.$$

Thus $\{y_1, \dots, y_n\}$ is a 2ε -approximation to K . Since $\varepsilon > 0$ is arbitrary, K is totally bounded. If also each K_λ is complete, then K is an intersection of complete sets and so is complete; whence it is compact. Q.E.D.

Proposition 4.3.5 *If the state space of B is firm, then so is the state space of every separable Banach subalgebra of B .*

Proof. Let A be a separable Banach subalgebra of B , $(x_n)_{n=1}^\infty$ a dense sequence in the unit ball of A , and $\|\cdot\|$ the corresponding double norm on A' . Given $\varepsilon > 0$, choose N such that $\sum_{n=N+1}^\infty 2^{-n} < \varepsilon$. Using Lemma 4.3.3, choose $t > 0$ such that

- V_B^t and V_A^t are weak* compact,
- $\rho_w(V_B^t, V_B) < \varepsilon$, and
- for each $f \in V_A^t$ there exists $g \in V_A$ such that

$$|f(x_k) - g(x_k)| \leq \varepsilon \quad (1 \leq k \leq N). \quad (4.3)$$

Let $f \in V_A^t$, and choose $g \in V_A$ such that (4.3) holds. We have, in A'_1 ,

$$\begin{aligned} \|f - g\| &= \sum_{n=1}^\infty 2^{-n} |(f - g)(x_n)| \\ &= \sum_{n=1}^N 2^{-n} |f(x_n) - g(x_n)| + \sum_{n=N+1}^\infty 2^{-n} |f(x_n) - g(x_n)| \\ &\leq \sum_{n=1}^N 2^{-n} \varepsilon + 2 \sum_{n=N+1}^\infty 2^{-n} \\ &< 3\varepsilon. \end{aligned}$$

It follows from Lemma 4.3.4 that V_A is weak* compact. It is then clear from the foregoing that $\rho_w(V_A^t, V_A) \rightarrow 0$ as $t \rightarrow 0$. Q.E.D.

4.4 Extreme points of the state space

Let K be a convex subset of a normed space, and let $x_0 \in K$. We say that x_0 is

- a *classical extreme point* of K if

$$\forall x, y \in K \left(x_0 = \frac{1}{2}(x + y) \Rightarrow x = y = x_0 \right);$$

- an *extreme point* of K if

$$\forall \varepsilon > 0 \exists \delta > 0 \forall x, y \in K \left(\|x_0 - \frac{1}{2}(x + y)\| < \delta \Rightarrow \|x - y\| < \varepsilon \right).$$

An extreme point is a classical extreme point, and the converse holds classically.

The proof of the following result is very similar to that given in [50, page 38] for the special case where B is a Banach algebra of functions; we include it for the sake of completeness. In the following, recall that Σ_B is the character space of B .

Proposition 4.4.1 *Let A be a commutative, unital Banach algebra generated by Hermitian elements, and*

$$K^0 = \{f \in A' : f \geq 0, f(e) \leq 1\}.$$

Then every classical extreme point of K^0 is an element of Σ_B .

Proof. Let ϕ be a classical extreme point of K^0 . We want to show that $\phi(xy) = \phi(x)\phi(y)$ for all $x, y \in B$. Considering first the case where $y = e$, define an element ψ of A' by

$$\psi = (1 - \phi(e))\phi.$$

Then $\psi(x) = (1 - \phi(e))\phi(x) \geq 0$ for each positive $x \in B$. Also, for each $x \geq 0$,

$$(\phi + \psi)(x) = \phi(x) + (1 - \phi(e))\phi(x) = (2 - \phi(e))\phi(x) \geq 0$$

and

$$(\phi - \psi)(x) = \phi(x) - (1 - \phi(e))\phi(x) = \phi(e)\phi(x) \geq 0.$$

Thus $\psi \geq 0$ and $\phi \pm \psi \geq 0$. Furthermore,

$$(\phi + \psi)(e) = \phi(e) + (1 - \phi(e))\phi(e) \leq \phi(e) + (1 - \phi(e)) \leq 1$$

and

$$(\phi - \psi)(e) = \phi(e) - (1 - \phi(e))\phi(e) = [\phi(e)]^2 \leq 1.$$

Therefore $\phi \pm \psi \in K^0$, and so $\psi = 0$ since ϕ is a classical extreme point of K^0 .

Next consider the case where $0 \leq y \leq e$. Define $\psi \in A'$ by

$$\psi(x) = \phi(xy) - \phi(x)\phi(y).$$

Then

$$\begin{aligned} (\phi + \psi)(e) &= \phi(e) + \psi(e) \\ &= \phi(e) + \phi(y) - \phi(e)\phi(y) \\ &= \phi(e)(1 - \phi(y)) + \phi(y) \\ &\leq 1. \end{aligned}$$

Also, if $x \geq 0$, then

$$\begin{aligned} (\phi + \psi)(x) &= \phi(x) + \psi(x) \\ &= \phi(x) + \phi(xy) - \phi(x)\phi(y) \\ &= \phi(x)(1 - \phi(y)) + \phi(xy) \\ &\geq 0 \end{aligned}$$

and

$$\begin{aligned} (\phi - \psi)(x) &= \phi(x) - \psi(x) \\ &= \phi(x) - \phi(x)\phi(y) + \phi(xy) \\ &= \phi(x)[x(e - y)] + \phi(xy) \\ &\geq 0. \end{aligned}$$

Thus $\phi \pm \psi \in K^0$, and hence $\psi = 0$.

In the case where $y \in \text{Her}(B)$, there exists $s > 0$ such that

$$0 \leq sy + \frac{1}{2}e \leq e.$$

To see this, compute $s > 0$ such that $\|sy\| < 1/2$. Then for each $f \in V_A$ we have

$$|sf(y)| \leq s\|y\| \leq \frac{1}{2} = \frac{1}{2}f(e);$$

whence

$$0 \leq f\left(sy + \frac{1}{2}e\right) = sf(y) + \frac{1}{2}f(e) < f(e)$$

for each $f \in V_A$. Now, by the work in the preceding paragraph,

$$\begin{aligned} s\phi(xy) + t\phi(x) &= \phi(x(sy + te)) \\ &= \phi(x)\phi(sy + te) \\ &= \phi(x)(s\phi(y) + t\phi(e)) \\ &= s\phi(x)\phi(y) + t\phi(x). \end{aligned}$$

Hence $s\phi(xy) = s\phi(x)\phi(y)$ and therefore, as $s > 0$, $\phi(xy) = \phi(x)\phi(y)$. Taking the case $x = y$ and using induction, we now see that

$$\phi(y^n) = \phi(y)^n \quad (n \geq 1).$$

Hence

$$\phi(xy^n) = \phi(x)\phi(y)^n \quad (x \in A, y \in \text{Her}(A), n \geq 1). \quad (4.4)$$

We now consider the general case. Fix $x \in A, y \in A$, and $\varepsilon > 0$. By our hypotheses, A is generated by Hermitian elements. So we can find a complex polynomial $p = p(a_1, a_2, \dots, a_n)$ in Hermitian elements $a_1, \dots, a_n \in A$ such that $\|y - p\| < \varepsilon$.

By (4.4), we have $\phi(xp) = \phi(x)\phi(p)$. Hence

$$\begin{aligned}
 |\phi(xy) - \phi(x)\phi(y)| &\leq |\phi(xy - xp)| + |\phi(xp) - \phi(x)\phi(p)| \\
 &\quad + |\phi(x)\phi(p) - \phi(x)\phi(y)| \\
 &\leq \|xy - xp\| + 0 + |\phi(x)| |\phi(p - y)| \\
 &\leq \|x\| \|y - p\| + \|x\| \|p - y\| \\
 &= 2\|x\| \|y - p\| \\
 &\leq 2\|x\| \varepsilon.
 \end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, we conclude that $|\phi(xy) - \phi(x)\phi(y)| = 0$ and therefore that $\phi(xy) = \phi(x)\phi(y)$. Q.E.D.

Lemma 4.4.2 *For each $t \in (0, 1)$, if $0 < \alpha, \beta \leq 1$ and $1 - \frac{1}{2}(\alpha + \beta) < t/2$, then $\alpha > 1 - t$ and $\beta > 1 - t$.*

Proof. If $1 - \frac{1}{2}(\alpha + \beta) < t/2$, then

$$0 \leq \frac{1}{2}(1 - \alpha) + \frac{1}{2}(1 - \beta) < \frac{t}{2},$$

so both $\frac{1}{2}(1 - \alpha) < t/2$ and $\frac{1}{2}(1 - \beta) < t/2$. Hence $\alpha > 1 - t$ and $\beta > 1 - t$. Q.E.D.

Proposition 4.4.3 *If the state space V of B is firm, then every extreme point of V is a character of B .*

Proof. Let $\|\cdot\|$ be the double norm corresponding to a dense sequence $(x_n)_{n=1}^\infty$ in the unit ball of B with $x_1 = e$. Noting that $V \subset K^0$, we show that every extreme point of V is also one of K^0 . Accordingly, let f_0 be an extreme point of V , and let $\varepsilon > 0$. Choose $\delta_1 \in (0, \varepsilon)$ such that if $f, g \in V$ and $\|\frac{1}{2}(f + g) - f_0\| < \delta_1$, then $\|f - g\| < \varepsilon$. Then choose an admissible $t > 0$ such that $\rho_w(V^t, V) < \delta_1/2$. Finally, choose $\delta_2 > 0$ such that if $f, g \in B'$ and $\|f - g\| < \delta_2$, then $|f(e) - g(e)| < t/2$.

Now let

$$\delta = \min \left\{ \frac{1}{2} \delta_1, \delta_2 \right\},$$

and consider $f, g \in K^0$ with $\left\| \frac{1}{2} (f + g) - f_0 \right\| < \delta$. Since

$$\left| \frac{1}{2} (f + g)(e) - 1 \right| = \left| \frac{1}{2} (f + g)(e) - f_0(e) \right| < \frac{t}{2},$$

we have $|1 - f(e)| < t$ and $|1 - g(e)| < t$, by Lemma 4.4.2; whence $f, g \in V^t$, and therefore there exist $f', g' \in V$ such that $\|f - f'\| < \delta_1/2$ and $\|g - g'\| < \delta_1/2$. We now have

$$\begin{aligned} \left\| \frac{1}{2} (f' + g') - f_0 \right\| &\leq \left\| \frac{1}{2} (f + g) - f_0 \right\| + \frac{1}{2} \|f - f'\| + \frac{1}{2} \|g - g'\| \\ &< \frac{1}{2} \delta_1 + \frac{1}{4} \delta_1 + \frac{1}{4} \delta_1 \\ &= \delta_1. \end{aligned}$$

Hence $\|f' - g'\| < \varepsilon$, and therefore

$$\|f - g\| \leq \|f - f'\| + \|f' - g'\| + \|g - g'\| < \varepsilon + \delta_1 < 2\varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, this completes the proof that f_0 is an extreme point, and therefore a classical extreme point, of K^0 . By Proposition 4.4.1, f_0 is a character of B . Q.E.D.

We now state a constructive version of the **Krein–Milman Theorem** ([10, page 363, (7.5)]).

Theorem 4.4.4 *Let x be a point in a compact convex subset K of a separable normed space X over \mathbf{R} . Then for each $\varepsilon > 0$ there exist extreme points x_1, \dots, x_n of K , and nonnegative numbers c_1, \dots, c_n with $\sum_{i=1}^n c_i = 1$, such that*

$$\left\| x - \sum_{i=1}^n c_i x_i \right\| < \varepsilon.$$

Proposition 4.4.5 *If V is weak* compact, then every element of V is a convex combination of characters of B .*

Proof. It is easily shown that V is convex. An application of the Krein–Milman Theorem shows that V is the closed convex hull of its extreme points; so we can apply Proposition 4.4.3. Q.E.D.

Corollary 4.4.6 *If the state space of B is firm, then the character space of every separable commutative Banach subalgebra of B is nonempty.*

Proof. Let A be a separable commutative Banach subalgebra of B . Proposition 4.3.5 shows that V_A is firm; in particular, it is compact and so has extreme points. By Proposition 4.4.3, those extreme points are characters of A . Q.E.D.

One more lemma will take us to the proof of our main result.

Lemma 4.4.7 *Let V be firm. Then $a \in \text{Her}(B)$ if and only if $f(a) \in \mathbf{R}$ for each $f \in V$; and a is positive if and only if $f(a) \geq 0$ for each $f \in V$.*

Proof. We deal only with the criterion for positivity, since the Hermitian case is similar but simpler. If a is positive, then for each $\varepsilon > 0$ there exists an admissible $t > 0$ such that $\text{Re } g(a) \geq -\varepsilon$ and $|\text{Im } g(a)| < \varepsilon$ for all $g \in V^t$. If $f \in V$, then $f \in V^t$ and so $\text{Re } f(a) \geq -\varepsilon$ and $|\text{Im } f(a)| < \varepsilon$. Since $\varepsilon > 0$ is arbitrary, we conclude that $f(a) = \text{Re } f(a) \geq 0$.

Conversely, suppose that $f(a) \geq 0$ for each $f \in V$. Since there exist admissible numbers $t > 0$ such that $\rho_w(V, V^t)$ is arbitrarily small, we can choose an admissible t such that for each $g \in V^t$ there exists $f \in V$ with $|g(a) - f(a)| < \varepsilon$. It now follows that for each $g \in V^t$,

$$|\text{Im } g(a)| \leq |\text{Im } f(a)| + |g(a) - f(a)| \leq 0 + \varepsilon = \varepsilon$$

and

$$\text{Re } g(a) > \text{Re } f(a) - \varepsilon \geq 0 - \varepsilon = -\varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, we conclude that $a \geq 0$. Q.E.D.

At last we can give the **Proof of Theorem 4.2.2**. Let A be the (separable) closed subalgebra of B generated by the Hermitian element a and the identity e . By Proposition 4.3.5, the state space V_A of A is firm. It follows from Proposition 4.4.5 that for each $f \in V_A$ there exist characters u_1, \dots, u_m of A , and nonnegative numbers $\lambda_1, \dots, \lambda_m$, such that $\sum_{i=1}^m \lambda_i = 1$ and $\|f - \sum_{i=1}^m \lambda_i u_i\|$ is arbitrarily small. In particular, given a positive integer n and $\varepsilon > 0$, we can choose the u_i and λ_i such that $|(f - \sum_{i=1}^m \lambda_i u_i)(a^n)| < \varepsilon$. Lemma 4.4.7 shows that $u_i(a)^n \in \mathbf{R}$ for each i . Hence

$$\begin{aligned} |\operatorname{Im} f(a^n)| &\leq \left| f(a^n) - \sum_{i=1}^m \lambda_i u_i(a^n) \right| + \left| \operatorname{Im} \sum_{i=1}^m \lambda_i u_i(a^n) \right| \\ &\leq \varepsilon + \left| \operatorname{Im} \sum_{i=1}^m \lambda_i u_i(a)^n \right| \\ &= \varepsilon. \end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, $\operatorname{Im} f(a^n) = 0$ for each $f \in V$. We now see from Lemma 4.4.7 that a^n is Hermitian.

Moreover, if $a \geq 0$ and n is even, then, with f, u_i , and λ_i as above, we have

$$\begin{aligned} \operatorname{Re} f(a^n) &\geq \operatorname{Re} \sum_{i=1}^m \lambda_i u_i(a^n) - \left| f(a^n) - \sum_{i=1}^m \lambda_i u_i(a^n) \right| \\ &\geq \operatorname{Re} \sum_{i=1}^m \lambda_i u_i(a)^n - \varepsilon \\ &\geq -\varepsilon, \end{aligned}$$

the last step following from Lemma 4.4.7. Since $\varepsilon > 0$ is arbitrary, we have $\operatorname{Re} f(a^n) \geq 0$ for each $f \in V$; whence, again by Lemma 4.4.7, $a \geq 0$. Q.E.D.

Chapter 5

Sinclair's Theorem

5.1 Introduction

In this chapter we discuss the spectral radius, spectrum, and approximations to the numerical range of an element of a unital Banach algebra. Our work, which is based on the classical treatment on pages 52–57 of [14], culminates in a constructive proof of Sinclair's theorem on the spectral radius of a Hermitian element.

5.2 Preliminaries

Let $(x_n)_{n=1}^\infty$ be a dense sequence in our Banach algebra B , and $\|\cdot\|$ the corresponding double norm on B' . For all but countably many $t > 0$ the set

$$\Sigma^t = \{f \in B' : |f(x_i x_j) - f(x_i) f(x_j)| \leq t \ (1 \leq i, j \leq n), \ |1 - f(e)| \leq t\}$$

is (nonempty and) weak* compact ([10, Chapter 9, (1.3) and (2.7)]); furthermore,

$$\Sigma = \bigcap \{\Sigma^t : t > 0\}.$$

Since, as we showed by Brouwerian 1.6.4 in Chapter 1, Σ is not weak* compact in general, we choose to work with the t -approximation Σ^t of Σ for carefully chosen t .

We cannot assume in constructive mathematics that a closed ideal of a Banach algebra is contained in a maximal ideal, so we follow Bishop *et al.* ([10, page 453]) and work with finite linear combinations of given elements of B . The following two results illustrate the use of those approximations.

Proposition 5.2.1 *Let x_1, \dots, x_n be elements of B , and t, ε positive numbers such that*

$$|f(x_1)| + \dots + |f(x_n)| \geq \varepsilon \quad (f \in \Sigma^t).$$

Then there exist a positive number δ (depending on n, t, ε) and elements y_1, \dots, y_n of B such that $x_1 y_1 + \dots + x_n y_n = e$ ([10, page 459, (2.6)]).

We say that two sequences (x_n) and (y_n) are **equiconvergent** if for each term a_m of one sequence, and each $\varepsilon > 0$, there exists N such that $b_n < a_m + \varepsilon$ whenever b_n is a term of the other sequence with $n \geq N$.

Proposition 5.2.2 *Let $(t_n)_{n=1}^\infty$ be a strictly decreasing sequence of positive numbers converging to 0, and for each $x \in B$ and each n define*

$$\|x\|_{\Sigma_{t_n}} = \sup \{|f(x)| : f \in \Sigma_{t_n}\}.$$

Then the sequences $\left(\|x\|_{\Sigma_{t_n}}\right)_{n=1}^\infty$ and $\left(\|x^n\|^{1/n}\right)_{n=1}^\infty$ are equiconvergent ([10, page 460, (2.9)]).

5.3 A closer look at the spectral radius

We define the **spectral radius** of an element x of B to be

$$\rho(x) = \sup \{|\lambda| : \lambda \in \sigma(x)\},$$

where $\sigma(x)$ is the spectrum of x . The classical *spectral radius formula*,

$$\rho(x) = r(x) = \inf \|x^n\|^{1/n} = \lim_{n \rightarrow \infty} \|x^n\|^{1/n}$$

holds for any $x \in B$; but there is no guarantee that either the supremum $\rho(x)$ or the infimum $r(x)$ exists constructively. Nonetheless, we write $r(x) < c$ to signify that $\|x^n\|^{1/n} < c$ for all sufficiently large n , and $\rho(x) < c$ to signify that $|\lambda| < c$ for all $\lambda \in \sigma(x)$.

Proposition 5.3.1 *If $a \in B$ and $\lambda \in \sigma(a)$, then for each n ,*

$$|\lambda| \leq \|a^n\|^{1/n}.$$

Proof. Given any $\lambda \in \sigma(a)$, suppose that $|\lambda| > \|a\|$. Then $0 < \|a\lambda^{-1}\| < 1$; so $e - a\lambda^{-1}$ is invertible, and therefore

$$a - \lambda e = -\lambda(e - \lambda^{-1}a).$$

This shows that $\lambda \notin \sigma(a)$, a contradiction. Hence $|\lambda| \leq \|a\|$.

By Proposition 3.2.3, for each $n \geq 2$,

$$\begin{aligned} a^n - \lambda^n e &= (a - \lambda e)(a^{n-1} + \cdots + \lambda^{n-1}e) \\ &\in (a - \lambda e)B \cap B(a - \lambda e) \\ &\subset \sim\text{inv}(B), \end{aligned}$$

so $\lambda^n \in \sigma(a^n)$. It now follows from the preceding paragraph that $|\lambda^n| \leq \|a^n\|$ and therefore that $|\lambda| \leq \|a^n\|^{1/n}$. Q.E.D.

Proposition 5.3.2 *Suppose that $\rho(a)$ exists and that $R(a)$ is coherent, and let ζ, ζ_0 be complex numbers such that $|\zeta| > |\zeta_0| > \rho(a)$. Then $|\zeta| > \|a^n\|^{1/n}$ for all sufficiently large n .*

Proof. For each z with $|z| > \rho(a)$ we have

$$z \in -\sigma(a) = -\sim R(a);$$

but $R(a)$ is coherent, so $z \in R(a)$ and therefore $(a - ze)^{-1}$ exists. Following Rudin's proofs in [62, pages 354–355], for each z with $|z| > \|a\|$ we see that

$$(a - z)^{-1} = - \sum_{n=0}^{\infty} z^{-n-1} a^n$$

and that for each $u \in B'$,

$$u((a - ze)^{-1}) = - \sum_{n=0}^{\infty} z^{-n-1} u(a^n). \quad (5.1)$$

The uniqueness of the Laurent coefficients of the holomorphic function

$$z \mapsto u((a - ze)^{-1})$$

in the annulus $\{z : |z| > \rho(a)\}$ shows that the series on the right of equation (5.1) converges to $u((a - ze)^{-1})$ for all z with $|z| > \rho(a)$. Since $|\zeta_0| > \rho(a)$, there exists $\beta > 0$ such that for each n , $\|\zeta_0^{-n} a^n\| \leq \beta$ and therefore $\|a^n\| \leq \beta |\zeta_0|^n$. It follows that $\|a^n\|^{1/n} \leq \beta^{1/n} |\zeta_0|$. Choosing N such that $\beta^{1/N} < |\zeta/\zeta_0|$, we now see that $\|a^n\|^{1/n} < |\zeta|$ for all $n \geq N$. Q.E.D.

Proposition 5.3.3 *If $\rho(a)$ exists and $\{\zeta : |\zeta| > \rho(a)\} \subset R(a)$, then $r(a)$ exists.*

Proof. By Proposition 5.3.1, $\rho(a) \leq \|a^n\|^{1/n}$ for each n . Given $\varepsilon > 0$, we see from Proposition 5.3.2 that there exists n such that $\rho(a) + \varepsilon > \|a^n\|^{1/n}$. Since $\varepsilon > 0$ is arbitrary, $\inf \{\|a^n\|^{1/n} : n \geq 1\}$ exists and equals $\rho(a)$. Q.E.D.

The following corollaries are trivial.

Corollary 5.3.4 *If $\rho(a)$ exists and $R(a)$ is coherent, then $r(a)$ exists.*

Corollary 5.3.5 *If $\sigma(a)$ is compact and $R(a)$ is coherent, then $r(a)$ exists.*

5.4 Approximating the numerical range

We define the *numerical range* of an element x of B to be the set

$$V(x) = \{f(x) : f \in V\},$$

where V is the state space of B . Since the constructive Hahn–Banach Theorem produces only approximately norm-preserving extensions of elements of B' , we may

be unable to construct an element of $V(x)$. To overcome this difficulty, we introduce the sets

$$V^t(x) = \{f(x) : f \in V^t\},$$

where $t > 0$, as approximations to the numerical range of x ; we call $V^t(x)$ the t -approximation to $V(x)$. For each admissible $t > 0$ and each $x \in B$, the weak* uniform continuity of the mapping $f \mapsto f(x)$ on the weak* compact set V^t ensures that $V^t(x)$ is totally bounded. Also,

$$V(x) = \bigcap \{V^t(x) : t > 0 \text{ is admissible}\}.$$

In particular, $V(x)$ is weak* closed.

Proposition 5.4.1 *The following are equivalent conditions on a complex number λ and an element a of B .*

- (i) *There exists an admissible $t > 0$ such that $\rho(\lambda, V^t(a)) > 0$.*
- (ii) *There exists $z \in \mathbf{C}$ such that $|\lambda - z| > \|a - ze\|$.*

Proof. Assuming that $\rho(\lambda, V^t(a)) > 0$ for some admissible $t > 0$, first consider the case where $\rho(a, \mathbf{C}e) > 0$. Define a normable linear functional f_0 on the 2-dimensional space $\text{span}\{a, e\}$ by

$$f_0(\alpha a + \beta e) = \alpha\lambda + \beta \quad (\alpha, \beta \in \mathbf{C}).$$

Choosing $\varepsilon > 0$ such that

$$\frac{\varepsilon}{1+\varepsilon} (1 + |\lambda|) < \rho(\lambda, V^t(a)),$$

suppose that $\|f_0\| < 1 + \varepsilon/2$. By the Hahn-Banach Theorem, there exists a normable linear extension f of f_0 to B such that

$$f(e) = f_0(e) = 1,$$

$$f(a) = f_0(a) = \lambda,$$

and $\|f\| \leq 1 + \varepsilon$. Define $g = \|f\|^{-1}f$. Then $\|g\| = 1$ and

$$|1 - g(e)| = |1 - \|f\|^{-1}f(e)| = |1 - \|f\|^{-1}| \leq \frac{\varepsilon}{1 + \varepsilon} < t,$$

so $g \in V^t$. Moreover,

$$\begin{aligned} |\lambda - g(a)| &= |\lambda - \|f\|^{-1}f(a)| \\ &= |\lambda - \|f\|^{-1}\lambda| \\ &= |1 - \|f\|^{-1}||\lambda| \\ &\leq \frac{\varepsilon}{1 + \varepsilon}|\lambda| \\ &< \rho(\lambda, V^t(a)), \end{aligned}$$

a contradiction. Hence, in fact,

$$\|f_0\| > 1 + \frac{\varepsilon}{2} > 1,$$

so we can pick $\alpha, \beta \in \mathbf{C}$ such that $\beta \neq 0$ and $|f_0(\alpha e + \beta a)| > \|\alpha e + \beta a\|$. Then

$$\begin{aligned} \|a - (-\alpha\beta^{-1})e\| &= |\beta|^{-1}\|\alpha e + \beta a\| \\ &\leq |\beta|^{-1}|f_0(\alpha e + \beta a)| \\ &= |\beta|^{-1}|\alpha + \beta\lambda| \\ &= |\lambda - (-\alpha\beta^{-1})|, \end{aligned}$$

so (ii) holds with $z = -\alpha\beta^{-1}$.

We now remove the restriction that $\rho(a, \mathbf{C}e) > 0$. Let

$$0 < \varepsilon < \min \{1, \rho(\lambda, V^t(a))\}$$

and choose an admissible

$$t' < \min \left\{ t, \frac{\varepsilon}{3(1 + \|a\|)} \right\}.$$

Either $\rho(a, \mathbf{C}e) > 0$ and the preceding case applies, or else $\rho(a, \mathbf{C}e) < \varepsilon/3$. In the latter case, choose $z \in \mathbf{C}$ such that $\|a - ze\| < \varepsilon/3$. Then for all $g \in V^{t'}$ we have

$$\begin{aligned}
 \varepsilon &< |\lambda - g(a)| \\
 &\leq |\lambda - z| + |z - g(ze)| + |g(a - ze)| \\
 &\leq |\lambda - z| + |z| |1 - g(e)| + \|a - ze\| \\
 &< |\lambda - z| + (1 + \|a\|) |1 - g(e)| + \frac{\varepsilon}{3} \\
 &< |\lambda - z| + (1 + \|a\|) t' + \frac{\varepsilon}{3} \\
 &< |\lambda - z| + \frac{2\varepsilon}{3}.
 \end{aligned}$$

Hence

$$|\lambda - z| > \frac{\varepsilon}{3} > \|a - ze\|,$$

and again (ii) holds.

Conversely, assume that there exists $z \in \mathbf{C}$ such that $|\lambda - z| > \|a - ze\|$, and let

$$\delta = \frac{1}{2} (|\lambda - z| - \|a - ze\|) > 0.$$

Choose an admissible $t > 0$ such that $|z|t < \delta/2$. Then for all $f \in V^t$ we have

$$\begin{aligned}
 |f(a) - z| &< |f(a) - zf(e)| + |z - zf(e)| \\
 &\leq |f(a - ze)| + |z| |1 - f(e)| \\
 &\leq \|a - ze\| + |z|t \\
 &< |\lambda - z| - \delta + \frac{\delta}{2} \\
 &< |\lambda - z| - \frac{\delta}{2}.
 \end{aligned}$$

Hence

$$\begin{aligned}
 |\lambda - f(a)| &\geq |\lambda - z| - |f(a) - z| \\
 &> |\lambda - z| + \frac{\delta}{2} - |\lambda - z| \\
 &> \frac{\delta}{2}.
 \end{aligned}$$

It follows that $\rho(\lambda, V^t(a)) \geq \delta/2$. Q.E.D.

Proposition 5.4.2 *For each $t > 0$,*

$$V^t(a) \subset \bigcap_{z \in \mathbf{C}} \overline{B}(z, \|a - ze\| + t|z|).$$

Proof. If $z \in \mathbf{C}$, $t > 0$, and $f \in V^t$, then

$$|f(a) - z| \leq |f(a - ze)| + |z| |1 - f(e)| \leq \|a - ze\| + t|z| \quad \text{Q.E.D.}$$

Corollary 5.4.3 *For each element a of the Banach algebra B ,*

$$V(a) = \bigcap_{z \in \mathbf{C}} \overline{B}(z, \|a - ze\|).$$

Proof. Let $\lambda \in V(a)$. By Proposition 5.4.2, for each $z \in \mathbf{C}$ and each admissible $t > 0$,

$$|\lambda - z| \leq \|a - ze\| + t|z|;$$

so, letting $t \rightarrow 0$, we obtain $|\lambda - z| \leq \|a - ze\|$. Hence

$$V(a) \subset \bigcap_{z \in \mathbf{C}} \overline{B}(z, \|a - ze\|).$$

Conversely, if $\lambda \in \bigcap_{z \in \mathbf{C}} \overline{B}(z, \|a - ze\|)$, then by Proposition 5.4.1, for each admissible $t > 0$ we have $\lambda \in \overline{V^t(a)} = V^t(a)$; whence $\lambda \in V(a)$. Q.E.D.

Proposition 5.4.4 *If $a \in B$, then $\sigma(a) \subset V(a)$.*

Proof. Let $\lambda \in \sigma(a)$, and suppose that $\rho(\lambda, V^t(a)) > 0$ for some admissible $t > 0$. Then by Proposition 5.4.1, there exists $z \in \mathbf{C}$ such that $|z - \lambda| > \|ze - a\|$. Hence $e - (z - \lambda)^{-1}(ze - a)$ is an invertible element of B . Let b be its inverse; then

$$[e - (z - \lambda)^{-1}(ze - a)] b = e,$$

so

$$(z - \lambda)b - (ze - a)b = (z - \lambda)e.$$

Rearranging, we obtain

$$(a - \lambda e) [b(z - \lambda)^{-1}] = e,$$

so $\lambda \in R(a)$, a contradiction. Therefore $\rho(\lambda, V^t(a)) = 0$ for each admissible $t > 0$, and so $\lambda \in V(a)$. Q.E.D.

Given a unit vector x in B , for each $t > 0$ we define the set

$$V^{t,x} = \{f \in B' : \|f\| \leq 1, |1 - f(x)| \leq t\}.$$

For each $a \in B$, we then write

$$V^{t,x}(a) = \{f(ax) : f \in B', \|f\| \leq 1, |1 - f(x)| \leq t\}.$$

If $t > 0$ is admissible, then $V^{t,x}(a)$, being the range of the uniformly continuous mapping $f \mapsto f(ax)$ on the weak* compact set V^t , is totally bounded.

Proposition 5.4.5 *If $0 < t < 1/\sqrt{2}$ and $2t$ is admissible, $\|x\| = 1$, $a \in B$, and $\lambda \in V^{t,x}(a)$, then there exists $\lambda' \in V^{2t}(a)$ such that $|\lambda - \lambda'| \leq 3t\|a\|$.*

Proof. Fix $\lambda \in V^{t,x}(a)$, and choose $g \in V^{t,x}$ such that $\lambda = g(ax)$. Suppose, to begin with, that $\rho(a, Ce) > 0$, so that a and e span a 2-dimensional subspace B_0 of B . Define a normable linear functional f on B_0 by

$$f_0(y) = g(yx) \quad (y \in B_0).$$

If $y \in B_0$ and $\|y\| \leq 1$, then

$$|f_0(y)| = |g(yx)| \leq \|yx\| \leq \|y\|\|x\| = \|y\| \leq 1.$$

Since also

$$|1 - f_0(e)| = |1 - g(x)| \leq t,$$

it follows that $1 - t \leq \|f_0\| \leq 1$. By the Hahn-Banach Theorem, there exists a normable linear functional f on B such that $1 - t \leq \|f\| \leq 1 + t$, and such that $f(y) = f_0(y)$ for all $y \in B_0$. Now define $\phi \in B'$ by

$$\phi = \|f\|^{-1} f.$$

Then $\|\phi\| = 1$,

$$\phi(e) = \|f\|^{-1} f_0(e) = \|f\|^{-1} g(x),$$

and

$$\phi(a) = \|f\|^{-1} f_0(a) = \|f\|^{-1} \lambda.$$

Moreover,

$$\begin{aligned} |1 - \phi(e)| &= |1 - \|f\|^{-1} g(x)| \\ &\leq |1 - \|f\|^{-1}| + |(1 - g(x)) \|f\|^{-1}| \\ &\leq \left|1 - \frac{1}{1+t}\right| + |1 - g(x)| \frac{1}{1-t} \\ &\leq \frac{t}{1+t} + \frac{t}{1-t} \\ &= \frac{t}{1-t^2} \\ &< 2t, \end{aligned}$$

as $0 < t < 1/\sqrt{2}$. Hence $\phi(a) \in V^{2t}(a)$. Also,

$$\begin{aligned} |\lambda - \phi(a)| &= \|\phi(a) - \phi(a)\| \\ &\leq \|f\|^{-1} \|\phi(a) - \phi(a)\| \\ &\leq t\|a\|, \end{aligned}$$

so the proof in the case $\rho(a, \mathbf{C}e) > 0$ is complete.

Now consider the general case. Since $2t$ is admissible, $V^{2t}(a)$ is totally bounded and hence located in B . Either $\rho(\lambda, V^{2t}(a)) \leq 2t\|a\|$ or else, as we may assume,

$\rho(\lambda, V^{2t}(a)) > t\|a\|$. Then, by the first part of the proof, $\rho(a, \mathbf{C}e) = 0$, so $a = \alpha e$ and $|\alpha| = \|a\|$ for some $\alpha \in \mathbf{C}$. For any $f \in V^{2t}$ we have

$$\begin{aligned}
 |\lambda - f(a)| &= |\lambda - \alpha f(e)| \\
 &\leq |\lambda - \alpha| + |\alpha| |1 - f(e)| \\
 &\leq |g(\alpha x) - \alpha| + 2t\|a\| \\
 &\leq |\alpha| |g(x) - 1| + 2t\|a\| \\
 &\leq 3t\|a\|.
 \end{aligned}$$

Q.E.D.

5.5 Sinclair's Theorem

In this section we make use of our approximations to the numerical range to prove Sinclair's Theorem for Hermitian elements of a unital Banach algebra. Before we arrive at our main result, we dispose of some lemmas.

Lemma 5.5.1 *For each admissible $t > 0$ and each unit vector $x \in B$,*

$$\inf \{ \operatorname{Re} \lambda : \lambda \in V^t(a) \} \leq \|ax\|.$$

Proof. First observe that the infimum in question exists, since $V^t(a)$ is totally bounded. Choose an admissible ε such that

$$0 < \varepsilon < \min \left\{ \frac{1}{\sqrt{2}}, \frac{t}{2} \right\}.$$

By Proposition 5.4.5, for each $g \in V^{\varepsilon, x}(a)$ there exists $\lambda' \in V^t(a)$ such that $|g(ax) - \lambda'| \leq 3\varepsilon\|a\|$. Thus

$$\begin{aligned}
 \inf \{ \operatorname{Re} \lambda : \lambda \in V^t(a) \} &\leq \operatorname{Re} \lambda' \\
 &\leq \operatorname{Re} g(ax) + |g(ax) - \lambda'| \\
 &\leq |g(ax)| + 3\varepsilon\|a\| \\
 &\leq \|ax\| + 3\varepsilon\|a\|.
 \end{aligned}$$

Since ε is arbitrary, the required result follows. Q.E.D.

For convenience we define

$$\mu_t = \sup\{\operatorname{Re} \lambda : \lambda \in V^t(a)\}.$$

Lemma 5.5.2 *If $t > 0$ is admissible, α is positive number such that $\alpha\mu_t < 1$, and x is a unit vector in B , then*

$$1 - \alpha\mu_t \leq \|(e - \alpha a)x\|.$$

Proof. Given $\lambda \in V^t(e - \alpha a)$, choose $f \in V^t$ such that

$$\lambda = f(e - \alpha a) = f(e) - \alpha f(a).$$

Then

$$\operatorname{Re} [\lambda - (1 - \alpha f(a))] = \operatorname{Re} [(\lambda + \alpha f(a)) - 1] = \operatorname{Re} (f(e) - 1) \geq -t,$$

since $|1 - f(e)| \leq t$. Hence

$$\operatorname{Re} \lambda \geq \operatorname{Re} (1 - \alpha f(a)) - t \geq (1 - \alpha\mu_t) - t.$$

It follows from Lemma 5.5.1 that

$$1 - \alpha\mu_t - t \leq \inf \{\operatorname{Re} \lambda : \lambda \in V^t(e - \alpha a)\} \leq \|(e - \alpha a)x\|.$$

For each admissible $\varepsilon \in (0, t)$, since $\mu_\varepsilon \leq \mu_t$, we now have

$$1 - \alpha\mu_t \leq 1 - \alpha\mu_\varepsilon \leq \|(e - \alpha a)x\| + \varepsilon.$$

Since ε is arbitrary, the desired conclusion follows. Q.E.D.

Lemma 5.5.3 *If $t > 0$ is admissible, then*

$$\frac{1}{\alpha} \log \|\exp(\alpha a)\| \leq \mu_t$$

for each $\alpha > 0$.

Proof. First consider $\alpha > 0$ such that $\alpha\mu_t < 1$. Applying Lemma 5.5.2, for each $x \in B$ we have

$$(1 - \alpha\mu_t) \|x\| \leq \|(e - \alpha a)x\|.$$

Therefore, by induction,

$$(1 - \alpha\mu_t)^n \|x\| \leq \|(e - \alpha a)^n x\|$$

for each positive integer n .

Now, for any $\alpha > 0$ and for all sufficiently large n , we have $\frac{\alpha}{n}\mu_t < 1$; whence

$$(1 - \frac{\alpha}{n}\mu_t)^n \|x\| \leq \|(e - \frac{\alpha}{n}a)^n x\|.$$

Taking the limit as $n \rightarrow \infty$, we obtain

$$\exp(-\alpha\mu_t) \|x\| \leq \|\exp(-\alpha a)x\|.$$

In particular, the choice $x = \|\exp(\alpha a)\|$ yields

$$\|\exp(\alpha a)\| \leq \exp(\alpha\mu_t),$$

from which the desired inequality follows. Q.E.D.

Proposition 5.5.4 *If a is Hermitian, then $\|\exp(\pm i\alpha a)\| = 1$ for all $\alpha \in \mathbf{R}$.*

Proof. For the moment, take $\alpha > 0$. Given $\varepsilon > 0$, choose an admissible $t > 0$ such that $|\operatorname{Im} f(a)| < \varepsilon$ for each $f \in V^t$. Then as

$$\begin{aligned} \{\operatorname{Re} \lambda : \lambda \in V^t(ia)\} &= \{\operatorname{Re} f(ia) : f \in V^t\} \\ &= \{-\operatorname{Im} f(a) : f \in V^t\}, \end{aligned}$$

we have

$$\sup \{\operatorname{Re} \lambda : \lambda \in V^t(ia)\} \leq \sup \{|\operatorname{Im} f(a)| : f \in V^t\} \leq \varepsilon.$$

Now replace α by $i\alpha$ in Lemma 5.5.3, to obtain

$$\frac{1}{\alpha} \log \|\exp(i\alpha a)\| \leq \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary,

$$\frac{1}{\alpha} \log \|\exp(i\alpha a)\| \leq 0$$

and therefore

$$\|\exp(i\alpha a)\| \leq 1.$$

Since $-a$ is also Hermitian, it follows that

$$\|\exp(-i\alpha a)\| \leq 1.$$

Thus for each real $\alpha \neq 0$ we have $\|\exp(i\alpha a)\| \leq 1$; this inequality holds for every $\alpha \in \mathbf{R}$, by continuity of the exponential function on B . Since

$$1 = \|e\| \leq \|\exp(i\alpha a)\| \|\exp(-i\alpha a)\| \leq 1$$

we conclude that $\|\exp(\pm i\alpha a)\| = 1$. Q.E.D.

Lemma 5.5.5 *Let α be a positive number such that $\sigma(a) \subset (-\frac{\pi}{2} + \alpha, \frac{\pi}{2} - \alpha)$. Then $\sigma(\sin a) \subset B(0, 1)$.*

Proof. The hypotheses allow us to choose $r \in (0, 1)$ such that $\sin \sigma(a) \subset\subset B(0, r)$. Suppose that $\lambda \in \sigma(\sin a)$ and $|\lambda| > r$. Then $\rho(\lambda, \sin \sigma(a)) > 0$ and so the mapping

$$\zeta \longmapsto (\sin \zeta - \lambda)^{-1}$$

is holomorphic on some open set D well containing $\sigma(a)$. It follows by the holomorphic functional calculus (see [52, page 206, Theorem 3.3.5]) that $(\sin a - \lambda e)^{-1}$ exists as a two-sided inverse to $\sin a - \lambda e$. Hence $\lambda \notin \sigma(\sin a)$, a contradiction. Therefore $|\lambda| \leq r$, and so $\lambda \in B(0, r)$. Q.E.D.

Recall that if $|z| \leq 1$, then

$$\arcsin(z) = z + \frac{1}{2} \cdot \frac{1}{3} z^3 + \frac{1}{2} \cdot \frac{3}{4} \cdot \frac{1}{5} z^5 + \cdots = \sum_{n=1}^{\infty} c_n z^n,$$

say. The holomorphic functional calculus on page 206 of [52, Theorem 3.3.5], and the remark on page 56 of [14], together show that if $\sigma(a) \subset\subset B(0, 1)$, then

$$\arcsin a = \sum_{n=0}^{\infty} c_n a^n.$$

It follows from Lemma 5.5.5 that

$$\arcsin(\sin a) = \sum_{n=0}^{\infty} c_n (\sin a)^n.$$

Now $\arcsin(\sin \zeta) = \zeta$ for all $\zeta \in [-\frac{\pi}{2}, \frac{\pi}{2}]$, and therefore for all ζ in some open set well containing $\sigma(a)$. The functional calculus now ensures that $\arcsin(\sin a) = a$ for all $a \in B$.

Proposition 5.5.6 *Let a be a Hermitian and suppose that $\|a^n\|^{1/n} < \frac{\pi}{2}$ for some positive integer n . Then $\|a\| < \frac{\pi}{2}$.*

Proof. Note that $\sigma(a) \subset (-\frac{\pi}{2}, \frac{\pi}{2})$ and that, by Lemma 5.5.5, $\sigma(\sin a) \subset B(0, 1)$.

Taken with the remarks preceding this proposition, these facts show that

$$a = \arcsin(\sin a) = \sum_{n=1}^{\infty} c_n (\sin a)^n.$$

By Proposition 5.5.4, $\|\exp(\pm ia)\| = 1$ and hence $\|\sin a\| \leq 1$. Since c_n is positive for each n , we have

$$\|a\| \leq \sum_{n=1}^{\infty} c_n \|\sin a\|^n \leq \sum_{n=1}^{\infty} c_n = \sum_{n=1}^{\infty} c_n \left(\sin \frac{\pi}{2}\right)^n = \frac{\pi}{2}. \quad \text{Q.E.D.}$$

Lemma 5.5.7 *Let a be Hermitian and $t > 0$. If there exists n such that $\|a^n\|^{1/n} < t$, then $\|a\| < t$,*

Proof. If $\|a^n\|^{1/n} < t$, then

$$\left\| \left(\frac{\pi}{2t} a \right)^n \right\|^{1/n} = \frac{\pi}{2t} \|a^n\|^{1/n} < \frac{\pi}{2}.$$

Since $\frac{\pi}{2t}a$ is Hermitian, it follows from Proposition 5.5.6 that $\|\frac{\pi}{2t}a\| < \frac{\pi}{2}$. Therefore $\|a\| < t$ for each $t > 0$. Q.E.D.

Our last lemma is a curiosity.

Lemma 5.5.8 *Let (r_n) be a decreasing sequence of positive numbers and M a real number such that $M \geq r_1$. Suppose that for each $t > 0$, if $r_n < t$ then $M < t$. Then $r_n = M$ for all n .*

Proof. Let $\varepsilon > 0$ and take $t = M - \varepsilon$. If $r_n < M - \varepsilon$, then $M < M - \varepsilon$, a contradiction. Hence $r_n \geq M - \varepsilon$ for each n . Since ε is arbitrary, we conclude that $r_n \geq M$ for each n . But $r_n \leq r_1 \leq M$ for each n . Hence $r_n = M$ for each n . Q.E.D.

We now prove our version of **Sinclair's Theorem**.

Theorem 5.5.9 *If a is a Hermitian element of the unital Banach algebra B , then $\|a^n\|^{1/n} = \|a\|$ for every n .*

Proof. We first observe that $\left(\|a^n\|^{1/n}\right)_{n=1}^{\infty}$ is a decreasing sequence, and that $\|a\| \geq \|a^n\|^{1/n}$ for each n . By Lemma 5.5.7, $\|a\| < t$ whenever $\|a^n\|^{1/n} < t$ for each $t > 0$. The conclusion follows by taking $r_n = \|a^n\|^{1/n}$ and $M = \|a\|$ in Lemma 5.5.8. Q.E.D.

Appendix A

Intuitionistic Logic

In section 1.4, the constructive interpretations of the connectives and quantifiers are discussed. These interpretations led Heyting to abstracting and formalising the axioms of intuitionistic logic. The axioms of the *intuitionistic propositional calculus* were first described by Heyting in [47]; other works of Heyting in axiomatic methods and Intuitionism can be found in [48]. Working with a fixed first-order language \mathcal{L} , we adopt the primitive connectives \vee (*or*), \wedge (*and*), \Rightarrow (*implies*), \neg (*not*). We assume familiarity with basic notions of elementary classical logic.

Propositional Axioms

Note that by adding the Law of Excluded Middle, $p \vee \neg p$, to the above list we get classical logic.

1. $p \Rightarrow (p \wedge p)$.
2. $(p \wedge p) \Rightarrow (p \wedge p)$.
3. $(p \Rightarrow q) \Rightarrow (p \wedge r \Rightarrow q \wedge r)$.
4. $(p \Rightarrow q) \Rightarrow ((q \Rightarrow r) \Rightarrow (p \Rightarrow r))$.
5. $q \Rightarrow (p \Rightarrow q)$.

6. $(p \wedge (p \Rightarrow q)) \Rightarrow q$.
7. $p \Rightarrow (p \vee q)$.
8. $(p \vee q) \Rightarrow (q \vee p)$.
9. $((p \Rightarrow r) \wedge (q \Rightarrow r)) \Rightarrow (p \vee q \Rightarrow r)$.
10. $\neg p \Rightarrow (p \Rightarrow q)$.
11. $((p \Rightarrow q) \wedge (p \Rightarrow \neg q)) \Rightarrow \neg p$.

Predicate Axioms

The following axioms can be added to the foregoing list to obtain the axioms of *intuitionistic predicate calculus*. We adopt the usual meanings of \forall (*for all*) and \exists (*there exists*). Furthermore, we take $p[x/t]$ to mean the formula obtained on replacing every occurrence of x in p by t in accordance with standard conventions; see pages 57–67 of [11]. A generalisation of a formula p is any formula of the form $\forall x_1 \cdots x_n p$, where x_1, \dots, x_n are any variables (not necessarily distinct).

1. $\forall x(p \Rightarrow q) \Rightarrow (\forall xp \Rightarrow \forall xq)$.
2. $\forall x(p \Rightarrow q) \Rightarrow (\exists xp \Rightarrow \exists xq)$.
3. $p \Rightarrow \forall xp$ if x is not free in p .
4. $\exists xp \Rightarrow p$ if x is not free in p .
5. $\forall xp \Rightarrow p[x/t]$ if t is free for x in p .
6. $p[x/t] \Rightarrow \exists xp$ if t is free for x in p .
7. All generalisation of 1-6.

Appendix B

A Spectral Theorem

Throughout, X denotes the compact product metric space $\prod_{n=1}^{\infty}[-1, 1]$; π_n the projection map $(x_k) \mapsto x_n$ of X onto $[-1, 1]$; and P the real subalgebra of $C(X)$ generated by the functions π_n and the constant function 1. Note that P is dense in $C(X)$ ([10, page 375, (8.19)]).

Let $\mathcal{A} = (A_n)_{n=1}^{\infty}$ be a sequence of commuting selfadjoint operators on H , and let $P(\mathcal{A})$ be the real subalgebra of $\mathcal{B}(H)$ generated by the operators A_n and the identity operator I . Let $p \mapsto p(\mathcal{A})$ be the unique algebra homomorphism $\phi : P \rightarrow P(\mathcal{A})$ such that $\phi(1) = I$ and $\phi(\pi_n) = A_n$ for each n . The mapping $p \mapsto p(\mathcal{A})$ is called the *canonical homomorphism* of P into $P(\mathcal{A})$.

Given a complex polynomial

$$p(z_1, \dots, z_n) = \sum_{i_1, \dots, i_n=0}^N c_{i_1, \dots, i_n} z_1^{i_1} \dots z_n^{i_n}$$

of degree N in n variables, we define the corresponding polynomial $p(\mathcal{A})$ by

$$p(\mathcal{A}) = \sum_{i_1, \dots, i_n=0}^N c_{i_1, \dots, i_n} A_1^{i_1} \dots A_n^{i_n}.$$

Recall that a uniformly bounded sequence $(T_n)_{n=1}^{\infty}$ of operators on a Hilbert space H *converges strongly* to an operator T if $T_n x \rightarrow T x$ as $n \rightarrow \infty$ for each

$x \in H$. In that case, any common bound for the operators T_n is a bound for T . Additionally, if each T_n is selfadjoint, then so is T ([10, page 374]).

Given a positive measure μ on $C(X)$, we say that a mapping $f : X \rightarrow \mathbf{C}$ is μ -*integrable* if both $\operatorname{Re} f$ and $\operatorname{Im} f$ are μ -integrable; in which case we write

$$\mu(f) = \mu(\operatorname{Re} f) + i \mu(\operatorname{Im} f).$$

We then define notions like *measurable* and *convergence in measure* for complex-valued functions, and spaces like $L_\infty(\mu, \mathbf{C})$, analogously to their counterparts for real-valued functions.

We state without proof two theorems, the first of which is a constructive **Spectral Theorem for selfadjoint operators** ([5, 10]).

Theorem B.0.10 *Let $\mathcal{A} = (A_n)$ be a sequence of commuting selfadjoint operators, with uniform bound 1, on a separable Hilbert space H , let (e_n) be an orthonormal basis of H , and let μ be the complete extension of the positive measure on X that satisfies*

$$\mu(p) = \sum_{n=1}^{\infty} 2^{-n} \langle p(\mathcal{A})e_n, e_n \rangle \quad (\text{B.1})$$

for every $p \in P$. Then the canonical homomorphism of P into $P(\mathcal{A})$ extends to a bound-preserving homomorphism $\varphi \mapsto \varphi(\mathcal{A})$ of L_∞ onto an algebra of commuting selfadjoint operators on H , such that (B.1) holds for every $\varphi \in L_\infty$. Moreover, if $(\varphi_n)_{n=1}^\infty$ is a bounded sequence of elements of L_∞ which converges in measure to an element φ of L_∞ , then the sequence $(\varphi_n(\mathcal{A}))_{n=1}^\infty$ converges strongly to $\varphi(\mathcal{A})$.

The pair $(\mu, \varphi \mapsto \varphi(\mathcal{A}))$ in Theorem B.0.10 is called the *functional calculus* for \mathcal{A} relative to the orthonormal basis (e_n) .

The proof of the following **Fuglede–Putnam–Rosenblum Theorem** in [63, page 300] is essentially constructive as it stands.

Theorem B.0.11 *If M, N , and $T \in \mathcal{B}(H)$, T is normal, and*

$$MT = TN,$$

*then $M^*T = TN^*$.*

We now state and prove a constructive **Spectral Theorem** for normal operators.

Theorem B.0.12 *Let $\mathcal{N} = (T_n)_{n=1}^\infty$ be a sequence of commuting normal operators, with uniform bound 1, on a separable Hilbert space H . For each n write*

$$T_n = B_n + i C_n, \tag{B.2}$$

where

$$B_n = \frac{1}{2} (T_n + T_n^*) \quad \text{and} \quad C_n = \frac{1}{2i} (T_n - T_n^*).$$

Then

$$\mathcal{A} = (B_1, C_1, B_2, C_2, \dots)$$

is a sequence of commuting selfadjoint operators on H with uniform bound 1. Given an orthonormal basis (e_n) of H , let

$$(\mu, \varphi \mapsto \varphi(\mathcal{A}))$$

be the corresponding functional calculus for the family \mathcal{A} . Let $L_\infty(\mu, \mathbb{C})$ consist of all $f : X \rightarrow \mathbb{C}$ such that $\operatorname{Re} f$, and $\operatorname{Im} f$ belong to $L_\infty(\mu, \mathbb{C})$. Then

$$f \longmapsto f(\mathcal{N}) = \operatorname{Re} f(\mathcal{A}) + i \operatorname{Im} f(\mathcal{A}) \tag{B.3}$$

is a homomorphism of $L_\infty(\mu, \mathbb{C})$ onto a family of commuting normal operators in $\mathcal{B}(H)$. Moreover,

$$\mu(\operatorname{Re} f) + i \mu(\operatorname{Im} f) = \sum_{n=1}^{\infty} 2^{-n} \langle f(\mathcal{N}) e_n, e_n \rangle \quad (f \in L_\infty(\mu, \mathbb{C})).$$

Finally, if $(f_n)_{n=1}^\infty$ is a bounded sequence of elements of $L_\infty(\mu, \mathbb{C})$ that converges in measure to an element f of $L_\infty(\mu, \mathbb{C})$, then the sequence $(f_n(\mathcal{N}))_{n=1}^\infty$ converges strongly to $f(\mathcal{N})$ in $\mathcal{B}(H)$.

Proof. By the Fuglede–Putnam–Rosenblum Theorem, \mathcal{A} is a sequence of commuting selfadjoint operators; moreover, the operators in \mathcal{A} have common bound 1. Note that since the operators of the form $\varphi(\mathcal{A})$, with $\varphi \in L_\infty(\mu)$, commute with each other, so do the operators of the form $f(\mathcal{N})$ with $f \in L_\infty(\mu, \mathbb{C})$.

Let $f, g \in L_\infty(\mu, \mathbb{C})$. Using the definition (B.3) and applying the functional calculus for \mathcal{A} , we have

$$\begin{aligned} (f+g)(\mathcal{N}) &= \operatorname{Re}(f+g)(\mathcal{A}) + i \operatorname{Im}(f+g)(\mathcal{A}) \\ &= \operatorname{Re} f(\mathcal{A}) + \operatorname{Re} g(\mathcal{A}) + i (\operatorname{Im} f(\mathcal{A}) + \operatorname{Im} g(\mathcal{A})) \\ &= (\operatorname{Re} f(\mathcal{A}) + i \operatorname{Im} f(\mathcal{A})) + (\operatorname{Re} g(\mathcal{A}) + i \operatorname{Im} g(\mathcal{A})) \\ &= f(\mathcal{N}) + g(\mathcal{N}), \end{aligned}$$

and for any scalar α ,

$$\begin{aligned} (\alpha f)(\mathcal{N}) &= \operatorname{Re} \alpha f(\mathcal{A}) + i \operatorname{Im} \alpha f(\mathcal{A}) \\ &= \alpha \operatorname{Re} f(\mathcal{A}) + \alpha i \operatorname{Im} f(\mathcal{A}) \\ &= \alpha (\operatorname{Re} f(\mathcal{A}) + i \operatorname{Im} f(\mathcal{A})) \\ &= \alpha f(\mathcal{N}). \end{aligned}$$

Thus the mapping $f \mapsto f(\mathcal{N})$ is linear. Next we have

$$\begin{aligned}
 f(\mathcal{N})g(\mathcal{N}) &= (\operatorname{Re} f(\mathcal{A}) + i \operatorname{Im} f(\mathcal{A})) \times (\operatorname{Re} g(\mathcal{A}) + i \operatorname{Im} g(\mathcal{A})) \\
 &= (\operatorname{Re} f(\mathcal{A}) \operatorname{Re} g(\mathcal{A}) - \operatorname{Im} f(\mathcal{A}) \operatorname{Im} g(\mathcal{A})) \\
 &\quad + i (\operatorname{Re} f(\mathcal{A}) \operatorname{Im} g(\mathcal{A}) + \operatorname{Im} f(\mathcal{A}) \operatorname{Re} g(\mathcal{A})) \\
 &= ([\operatorname{Re} f \operatorname{Re} g](\mathcal{A}) - [\operatorname{Im} f \operatorname{Im} g](\mathcal{A})) \\
 &\quad + i ([\operatorname{Re} f \operatorname{Im} g](\mathcal{A}) + [\operatorname{Im} f \operatorname{Re} g](\mathcal{A})) \\
 &= [\operatorname{Re} f \operatorname{Re} g - \operatorname{Im} f \operatorname{Im} g](\mathcal{A}) + i [\operatorname{Re} f \operatorname{Im} g + \operatorname{Im} f \operatorname{Re} g](\mathcal{A}) \\
 &= (\operatorname{Re} fg)(\mathcal{A}) + i (\operatorname{Im} fg)(\mathcal{A}) \\
 &= (fg)(\mathcal{N}),
 \end{aligned}$$

so $f \mapsto f(\mathcal{N})$ is a homomorphism.

We compute

$$\begin{aligned}
 \mu(f) &= \sum_{n=1}^{\infty} 2^{-n} \langle f(\mathcal{N}) e_n, e_n \rangle \\
 &= \sum_{n=1}^{\infty} 2^{-n} \langle (\operatorname{Re} f(\mathcal{A}) + i \operatorname{Im} f(\mathcal{A})) e_n, e_n \rangle \\
 &= \sum_{n=1}^{\infty} 2^{-n} \langle \operatorname{Re} f(\mathcal{A}) e_n, e_n \rangle + i \sum_{n=1}^{\infty} 2^{-n} \langle \operatorname{Im} f(\mathcal{A}) e_n, e_n \rangle \\
 &= \mu(\operatorname{Re} f) + i \mu(\operatorname{Im} f).
 \end{aligned}$$

Finally, let $(f_n)_{n=1}^{\infty}$ be a bounded sequence in $L_{\infty}(\mu, \mathbf{C})$ that converges in measure to an element f of $L_{\infty}(\mu, \mathbf{C})$. Then the sequences $(\operatorname{Re} f_n)_{n=1}^{\infty}$ and $(\operatorname{Im} f_n)_{n=1}^{\infty}$ are bounded in $L_{\infty}(\mu, \mathbf{R})$, and converge in measure to $\operatorname{Re} f$ and $\operatorname{Im} f$, respectively. It follows from the spectral theorem for selfadjoint operators that the sequences $(\operatorname{Re} f_n(\mathcal{N}))_{n=1}^{\infty}$ and $(\operatorname{Im} f_n(\mathcal{N}))_{n=1}^{\infty}$ converge strongly to $\operatorname{Re} f(\mathcal{N})$ and $\operatorname{Im} f(\mathcal{N})$, respectively. Hence $(f_n(\mathcal{N}))$ converges strongly to $f(\mathcal{N})$. Q.E.D.

References

- [1] G. Bachman and L. Narici, *Functional Analysis*, Academic Press, New York, 1966.
- [2] M.J. Beeson, 'Formalizing constructive mathematics: why and how?', in *Constructive Mathematics* (Proceedings of Conference at New Mexico State University, 1980; Fred Richman, ed.), Lecture Notes in Mathematics **873**, Springer-Verlag, Berlin, 146–190, 1980.
- [3] _____ *Foundations of Constructive Mathematics*, Ergebnisse der Mathematik und ihrer Grenzgebiete **6**, Springer-Verlag, Berlin, 1985.
- [4] E.A. Bishop, 'Constructive methods in the theory of Banach algebras', in *Function Algebras* (Proceedings of International Symposium on Function Algebras, Tulane University, Tulane, 1965; F.T. Birtel, ed.), Scott-Foresman, Chicago, 343–345, 1966.
- [5] _____ *Foundations of Constructive Analysis*, McGraw-Hill, New York, 1967.
- [6] _____ 'Mathematics as a numerical language', in *Intuitionism and Proof Theory* (Proceedings of the Summer Conference at Buffalo, New York, 1968; A. Kino, J. Myhill, R.E. Vesley, eds.), North-Holland, Amsterdam, 53–71, 1970.
- [7] _____ 'Aspects of Constructivism', New Mexico State University, Las Cruces, NM, 1972.
- [8] _____ 'The crisis in contemporary mathematics', *Historia Math.* **2**, 507–517, 1975.
- [9] _____ 'Schizophrenia in Contemporary Mathematics', in *Errett Bishop: Reflections on Him and His Research* (Proceedings of Memorial Meeting for Erret Bishop, University of California, San Diego, 1983; M. Rosenblatt, ed.),

- Contemporary Mathematics **39**, American Math. Soc., Providence, RI, 1–32, 1985.
- [10] E.A. Bishop and D.S. Bridges, *Constructive Analysis*, Grundlehren der Mathematischen Wissenschaften **279**, Springer–Verlag, Berlin, 1985.
 - [11] J.L. Bell and M. Machover, *A Course in Mathematical Logic*, North–Holland, Amsterdam, 1977.
 - [12] F.F. Bonsall and J. Duncan, *Numerical Ranges of Operators on Normed Spaces and of Elements of Normed Algebras*, London Math. Soc. Lecture Notes **2**, Cambridge University Press, Cambridge, 1971.
 - [13] _____ *Numerical Ranges II*, London Mathematical Society Lecture Notes **10**, Cambridge University Press, Cambridge, 1973.
 - [14] _____ *Complete Normed Algebras*, Ergebnisse der Mathematik und ihrer Grenzgebiete **80**, Springer–Verlag, Berlin, 1973.
 - [15] D.S. Bridges, *Constructive Functional Analysis*, Research Notes in Mathematics **28**, Pitman, London, 1979.
 - [16] _____ ‘A constructive look at the real number line’, in *Real Numbers, Generalizations of the Reals, and Theories of Continua* (special issue of *Synthese*, P. Ehrlich, ed.), Kluwer Academic Publishers, Netherlands, 29–92, 1994.
 - [17] _____ *Foundations of Real and Abstract Analysis*, Graduate Texts in Mathematics **174**, Springer–Verlag, Berlin, 1998.
 - [18] _____ ‘A constructive look at ideals in a commutative Banach algebra’, preprint, University of Canterbury, Christchurch, New Zealand.
 - [19] _____ ‘Approximate characters of a Banach algebra’, preprint, University of Canterbury, Christchurch, New Zealand.
 - [20] _____ ‘Constructive mathematics: a foundation for computable analysis’, *Theoretical Computer Science* **219**, 95–109, 1999.
 - [21] _____ ‘Constructive methods in Banach algebra theory’, *Mathematica Japonica* **52**(1), 145–161, 2000.

- [22] D.S. Bridges and L.S. Dediú, 'Paradise lost or paradise regained?', *EATCS Bulletin* **63**, 141–155, 1997.
- [23] _____ 'Weak continuity properties in constructive analysis', *Logic Journal of the IGPL* **7**(3), 277–281, 1999.
- [24] D.S. Bridges and R.S. Havea, 'A constructive analysis of a proof that the numerical range is convex', *London Math. Soc. Journal of Computation and Mathematics* **3**, 191–206, 2000.
- [25] _____ 'A constructive version of the spectral mapping theorem', to appear in *Math. Logic Quarterly* **47**, 2001.
- [26] D.S. Bridges and H. Ishihara, 'Locating the range of an operator on a Hilbert space', *Bulletin of the London Math. Soc.* **24**(6), 599–605, 1992.
- [27] _____ 'Spectra of selfadjoint operators in constructive analysis', *Proc. Koninkl. Nederlandse Akad. Wet. (Indag. Math.)* **7**(1), 11–35, 1996.
- [28] _____ 'A definitive constructive open mapping theorem?', *Math. Logic Quarterly* **44**(4), 545–552, 1998.
- [29] D.S. Bridges and R. Mines, 'What is constructive mathematics?', *Math. Intelligencer* **6**(4), 403–410, 1984.
- [30] D.S. Bridges and F. Richman, *Varieties of Constructive Mathematics*, London Math. Soc. Lecture Notes **97**, Cambridge University Press, Cambridge, 1987.
- [31] D.S. Bridges, F. Richman and P.M. Schuster, 'Adjoint, absolute values and polar decompositions', *J. Operator Th.* **44**(2), 243–254, 2000.
- [32] D.S. Bridges, F. Richman and Y. Wang, 'Sets, complements and boundaries', *Proc. Koninkl. Nederlandse Akad. Wet. (Indag. Math.)* **7**(4), 425–445, 1996.
- [33] L.E.J. Brouwer, *Over der Grondslagen der Wiskunde*, Ph.D. thesis, University of Amsterdam, Netherlands, 1907.
- [34] _____ 'Historical background, principles and methods of intuitionism', *S. African J. Science* **49**, 139–146, 1952.
- [35] P. Davies, *The Mind of God*, Penguin Books, London, 1993.

- [36] J. Dieudonné, *Foundations of Modern Analysis*, Academic Press, New York, 1960.
- [37] _____ *Mathematics—The Music of Reason*, Springer-Verlag, Berlin, 1992.
- [38] M.A.E. Dummett, *Elements of Intuitionism*, Clarendon Press, Oxford, 1977; second edition, 2000.
- [39] N. Dunford and J.T. Schwartz, *Linear Operators Part II: Spectral Theory, Self Adjoint Operators in Hilbert Space*, Wiley Interscience, New York, 1963.
- [40] N.D. Goodman, 'Mathematics as an objective science', *Amer. Math. Monthly* **86**(7), 540–551, 1979.
- [41] _____ 'Reflections on Bishop's philosophy of mathematics', in *Constructive Mathematics* (Proceedings of Conference at New Mexico State University, 1980; Fred Richman, ed.), *Lecture Notes in Mathematics* **873**, Springer-Verlag, Berlin, 135–145, 1980; reprinted in *Math. Intelligencer* **5**(3), 61–68, 1983.
- [42] N.D. Goodman and J. Myhill, 'The formalization of Bishop's constructive mathematics', in *Toposes, Algebraic Geometry and Logic* (Conference at Dalhousie University, Halifax, 1971), *Lecture Notes in Mathematics* **274**, Springer-Verlag, Berlin, 83–96, 1972.
- [43] _____ 'Choice implies excluded middle', *Zeit. Math. Logik und Grundlagen der Math.* **24**, 461, 1978.
- [44] K.E. Gustafson and D.K.M. Rao, *Numerical Range*, Universitext, Springer, New York, 1997.
- [45] P.R. Halmos, *Introduction to Hilbert Space and the Theory of Spectral Multiplicity* (2nd Ed.), Chelsea, New York, 1957.
- [46] _____ *A Hilbert Space Problem Book*, *Graduate Texts in Mathematics* **19**, Springer-Verlag, New York, 1967.
- [47] A. Heyting, 'Die formalen Regeln der intuitionistischen Logik', *Sitzungsber. preuss. Akad. Wiss. Berlin*, 42–56, 1930.

- [48] _____ 'Axiomatic methods and Intuitionism', in *Essays on the Foundations of Mathematics* (dedicated to A.A. Fraenkel on his seventieth anniversary; Y. Bar-Hillel, ed.), Magnes Press, Jerusalem, 1961.
- [49] _____ *Intuitionism: An Introduction* (Third Edition), North-Holland, Amsterdam, 1971.
- [50] R.B. Holmes, *Geometric Functional Analysis and its Applications*, Graduate Texts in Mathematics **24**, Springer-Verlag, New York, 1975.
- [51] H. Ishihara, 'Continuity and nondiscontinuity in constructive mathematics', *J. Symbolic Logic* **56**(4), 1349–1354, 1991.
- [52] R.V. Kadison and J.R. Ringrose, *Fundamentals of the Theory of Operator Algebras: Elementary Theory (Volume 1)*, Graduate Studies in Mathematics **15**, Amer. Math. Soc., Providence, RI, 1997.
- [53] B.A. Kushner, *Lectures on constructive mathematical analysis*, Transl. Math. Monographs **60**, Amer. Math. Soc., Providence, RI, 1985.
- [54] _____ 'Markov's constructive analysis', *Theoretical Computer Science* **219**(1-2), 267–289, 1999.
- [55] B.V. Limaye, *Functional Analysis* (Second Edition), New Age International, New Delhi, 1996.
- [56] R. Mines, F. Richman and W. Ruitenburg, *A Course in Constructive Algebra*, Universitext, Springer-Verlag, New York, 1988.
- [57] F. Richman, 'Meaning and information in constructive mathematics', *Amer. Math Monthly* **89**(6), 385–388, 1982.
- [58] _____ 'Intuitionism as generalization', *Phil. Math.* **5**(1-2), 124–128, 1990.
- [59] _____ 'Polynomials and linear transformation', *Linear Algebra and Appl.* **131**, 131–137, 1990.
- [60] _____ 'Interview with a constructive mathematician', *Modern Logic* **6**(3), 247–271, 1996.
- [61] _____ 'Existence proofs', *Amer. Math. Monthly* **106**(4), 303–308, 1999.

- [62] W. Rudin, *Real and Complex Analysis*, McGraw–Hill, New York, 1970.
- [63] _____ *Functional Analysis*, McGraw–Hill, New York, 1973.
- [64] E. Schechter, ‘Constructivism is difficult’, *Amer. Math. Monthly* **108**, 50–54, 2001.
- [65] J.P. Seldin, ‘A constructive approach to classical mathematics’ in *Constructive Mathematics* (Proceedings of Conference at New Mexico State University, 1980; F. Richman, ed.), Lecture Notes in Mathematics **873**, Springer–Verlag, Berlin, 105–110, 1980.
- [66] A.S. Troelstra and D. van Dalen, *Constructivism in Mathematics* (2 volumes), North–Holland, Amsterdam, 1988.
- [67] D. van Dalen, *Brouwer’s Cambridge Lectures on Intuitionism*, Cambridge University Press, Cambridge, 1981.
- [68] _____ *Mystic, Geometer, and Intuitionist: The Life of L.E.J. Brouwer (Volume 1)*, Clarendon Press, Oxford, 1999.
- [69] K. Weihrauch, *Computability*, EATCS Monographs on Theoretical Computer Science **9**, Springer–Verlag, Berlin, 1987.
- [70] _____ ‘A foundation for computable analysis’, in *Combinatorics, Complexity, & Logic* (Proceedings of Conference in Auckland, 1996; D.S. Bridges, C.S. Calude, J. Gibbons, S. Reeves, I.H. Witten, eds.), Springer–Verlag, Singapore, 66–89, 1997.

Symbols

$\mathcal{B}(H)$	bounded linear operators on H
$\text{Her}(B)$	Hermitian elements of B
$\text{inv}(S)$	invertible elements of S
$R(x)$	resolvent of x
V_B	state space of B
V_B^t	t -approximation to V_B
$V_B(x)$	numerical range of x
$V_B^t(x)$	t -approximation to $V_B(x)$
$W(T)$	numerical range of T
X'	dual of X
∂S	boundary of S
Σ_B	character space (or spectrum) of B
Σ_B^t	t -approximation to Σ_B
$\sigma(x)$	spectrum of x
$\sigma_a(T)$	approximate point spectrum of T
$\tau(T)$	nonspectrum of T
$\ \cdot\ $	double norm

Index

- μ -integrable, 85
- admissibility, 53
- algorithm, 6
- Axiom of Choice, 13
- axioms of intuitionistic logic, 82
- binary sequence, 11
- BISH, 6
- Bishop's
 - Lemma, 49
 - mathematics, 6
- Bishop, E., 6
- boundary crossing property, 46
- Brouwer, L.E.J., 3
- Brouwerian counterexample, 14
- canonical homomorphism, 84
- character
 - of a Banach algebra, 17
 - space, 17
- CLASS, 4
- CM, 4
- compact sets, 45
- computability, 5
- constructibility, 5
- Continuum Hypothesis, 1
- de Boer, C., 20
- double norm, 53
- equiconvergent sequence, 67
- existence, mathematical, 5
- extreme points, 59
 - classical, 59
- firm set, 54
- Formalism, 2
- Fuglede-Putnam-Rosenblum Theorem, 85
- functional calculus, 85
- Fundamental
 - Constructivist Thesis, 7
 - Theorem of Algebra, 49
- Gödel, K., 2
- Goldbach Conjecture, 12
- Goodman, N., 13
- Hahn-Banach Theorem, 54
- Halmos, P., 20
- Hermitian element, 54
- Heyting, A., 4
- Hilbert, D., 2

- inequality, 41
 - discrete, 42
 - quasi-discrete, 42
 - standard, 42
- INT, 3
- Intermediate Value Theorem, 12
- Intuitionism, 3
- Intuitionistic Logic, 4
 - axioms of, 82
- Law
 - of Excluded Middle, 10
 - of Trichotomy, 12
- Least-upper-bound Principle, 12
- LEM, 10
- linear functional
 - norm of a , 52
 - normable, 52
 - positive, 54
- LLPO, 11
- $LLPO_N$, 11
- locally totally bounded sets, 45
- located sets, 44
- Logicism, 1
- LPO, 11
- Markov algorithm, 6
- Markov's Principle, 12
- Markov, A.A., 6
- MP, 12
- Myhill, J., 13
- norm, double, 53
- normable linear functional, 52
- numerical range
 - element of a Banach algebra, 69
 - operator, 14, 20
 - convexity of the, 14, 20
- omniscience principles, 12
- Platonism, 1, 3
- positive
 - element, 54
 - linear functional, 54
- Principle
 - Least-upper-bound, 12
 - of Countable Choice, 13
 - of Dependent Choice, 13
 - of the Excluded Third, 3
- recursive mathematics, 6
- resolvent, 41
- Richman, F., 7
- RUSS, 6
- Russian constructivism, 6
- sequence
 - binary, 11
 - equiconvergent, 67
- sequential continuity, 29
- set

- border, 48
- coherent, 46
- compact, 45
- complement of, 42
- firm, 54
- internally approximated, 46
- locally totally bounded, 45
- located, 44
- metric complement of, 45
- totally bounded, 45
- well contained, 46
- Spectral Mapping Theorem, 40
 - classical, 41
 - first half of, 41, 44
 - second half of, 44, 49
- spectral radius, 67
 - formula, 67
- Spectral Theorem
 - for normal operators, 86
 - selfadjoint operators, 85
- spectrum
 - of a Banach algebra, 17
 - of an element of a Banach algebra, 41
- state space, 53
- strong convergence, 84
- strongly extensional, 29
- Toeplitz–Hausdorff Theorem, 20
- totally bounded sets, 45
- WLPO, 11
- Zermelo, E., 13
- ZFC, 1