ALGEBRAIC FREDHOLM THEORY

JOEL H. SHAPIRO

Abstract. These notes develop the purely algebraic parts of the basic theory of Fredholm operators. The setting is vector spaces over arbitrary fields, and linear transformations between these spaces. No topology is assumed for the vector spaces and consequently no continuity is assumed for the linear transformations. Within this reduced setting the notion of "Fredholm transformation" is studied, emphasizing invertibility properties, expression of these properties in terms of Fredholm index, invariance of these properties under finite-rank perturbations, and implications for the study of spectra. The exposition culminates in a Fredholm-inspired analysis of the spectra of multiplication operators on spaces of holomorphic functions.

1. Linear Spaces

Throughout these notes the setting is an arbitrary, possibly infinite dimensional, vector space—devoid of topology. In this initial section we'll discuss the fundamental notions of dimension, codimension, and invertibility, emphasizing the computation and estimation of codimension.

Notation and terminology

In what follows, "vector space" will mean "vector space over a field $\mathbb{F}$," henceforth called "the scalar field," or just "the scalars." Since there is no topology in sight, there will be no discussion of continuity. Symbols X, Y, and Z, possibly with subscripts, will always denote vector spaces over $\mathbb{F}$. The collection of linear transformations $X \to Y$ will be denoted by $\mathscr{L}(X, Y)$, or, if $X = Y$, just by $\mathscr{L}(X)$. The sub-collection of finite-rank transformations (those linear transformations T for which $\operatorname{ran} T := T(X)$ is finite dimensional) will be denoted by $\mathscr{F}(X, Y)$, or, if $X = Y$, just by $\mathscr{F}(X)$.

Basis and Dimension

Recall that a *basis* for a vector space is a maximal linearly independent set, i.e., a linearly independent set whose span is the whole space. Every vector in

January 27, 2011.

the space is thus uniquely expressible as a (finite) linear combination of basis vectors. If a linearly independent set does not span the whole space, vectors can always be appended (often thanks to the Axiom of Choice) to extend the original set to a basis (see , e.g., [2, §1.11, and §1.14 Exercise 2] or [3, Theorem 1.1(i), page 3]).

The *dimension* of a vector space is the cardinality of a basis; this number is independent of the basis (see, e.g., [3, Theorem 1.1(ii)]). Here we will think of the cardinality of an infinite set as just "∞", ignoring any subtleties attached to infinite cardinal numbers.[1] In particular, dimensions of vector spaces will be either finite or "infinite". Thus, under the usual conventions for addition in the set $\{\text{Non-Negative Integers}\} \cup \{\infty\}$, the following fundamental result is true even if some of the dimensions involved are infinite.

Theorem 1.1 (The "Fundamental Theorem of Linear Algebra"). *If* $T \in \mathscr{L}(X, Y)$ *then* $\dim X = \dim \ker T + \dim \operatorname{ran} T$*.*

Proof. Fix K a basis for $\ker T$ and add a linearly independent set J of vectors to extend this basis to one for X. Since T is injective on the span of J, it's enough to prove that T takes J to a basis for $\operatorname{ran} T$. Clearly $T(J) \subset \operatorname{ran} T$. To see $T(J)$ spans $\operatorname{ran} T$, write an arbitrary vector in X as a linear combination of the vectors in $K \cup J$ and note that upon applying T to this vector, only the terms involving $T(J)$ survive. To see that $T(J)$ is linearly independent, suppose that the zero-vector of Y is a linear combination of vectors in $T(J)$, say $\sum c_k T(v_k) = 0$, where the vectors v_k lie in J, the c_k's are scalars, and the sum extends over just finitely many indices k. Let $x = \sum c_k v_k$ so that $Tx = 0$. Thus $x \in (\operatorname{span} J) \cap \ker T = \{0\}$, hence by the linear independence of the set J, all the $c_k' s$ are zero. □

Perhaps the best known application of Theorem 1.1 is the following result, which we will encounter in increasing generality throughout the sequel.

Corollary 1.2 (Fredholm Alternative I). *Suppose X is a finite dimensional vector space and $T \in \mathscr{L}(X)$. Then T is injective if and only if it is surjective.*

Proof. T is injective if and only if $\ker T = \{0\}$. By Theorem 1.1 this is equivalent to "$\dim \operatorname{ran} T = \dim X$." Since $\operatorname{ran} T$ is a subspace of X, and X is finite dimensional, this last statement is equivalent to "$\operatorname{ran} T = X$." □

[1]This informality comes at a price; not all vector spaces of "dimension ∞" are isomorphic.

Examples 1.3 (Shift transformations)**.** Here are two examples of linear transformations that illustrate many of the phenomena that we are going to study. Right now they show us that Corollary 1.2 is not true in infinite dimensional spaces. For our vector space we'll take the collection ω of all scalar sequences, i.e., all functions from the natural numbers into the scalar field.

(a) *The Forward Shift.* This is the transformation $S \in \mathscr{L}(\omega)$ defined for $n \in \mathbb{N}$ and $x \in \omega$ by: $Sx(n) = 0$ if $n = 1$, and $= x(n-1)$ if $n > 1$. That is:

$$Sx = (0, x(1),\ x(2), \dots) \qquad (x \in \omega).$$

Thus S shifts each sequence one unit to the right, placing a zero in the newly emptied first position. S is an injective linear transformation with range consisting of those sequences with first coordinate zero, and so is not surjective.

(b) *The Backward Shift.* This is the transformation $B \in \mathscr{L}(\omega)$ defined by $Bx(n) = x(n+1)$, i.e.,

$$Bx = (x(2), x(3),\ \dots) \qquad (x \in \omega).$$

Thus B is a surjective linear transformation with kernel equal to those sequences that are zero except possibly in the first coordinate; it is surjective but not injective.

Here are two further applications of the Fundamental Theorem of Linear Algebra that will prove useful later on.

Corollary 1.4. *Suppose $T \in \mathscr{L}(X, Y)$ has finite dimensional kernel, and V is a finite dimensional subspace of Y. Then $T^{-1}(V)$ is finite dimensional.*

Proof. Let T_0 denote the restriction of T to $T^{-1}(V)$. Then $\operatorname{ran} T_0 = \operatorname{ran} T \cap V$ and $\ker T_0 = \ker T$ (since $\ker T = T^{-1}\{0\} \subset T^{-1}(V)$). By Theorem 1.1,

$$\dim T^{-1}(V) \;=\; \dim\ker T_0 + \dim\operatorname{ran} T_0 \;=\; \dim\ker T + \dim(V \cap \operatorname{ran} T).$$

Since both summands on the right are finite, so is $\dim T^{-1}(V)$. □

Theorem 1.5 (Kernel of a product)**.** *Suppose $T \in \mathscr{L}(X, Y)$ and $S \in \mathscr{L}(Y, Z)$. Then*

$$\dim(\ker ST) = \dim(\ker T) + \dim(\ker S \cap \operatorname{ran} T) \tag{1}$$

Proof. The argument depends on these two facts about $\ker ST$:

(a) $\ker T \subset \ker ST$, and
(b) $\ker ST = T^{-1}(\ker S)$.

Let T_0 denote the restriction of T to $\ker ST$. By (a) above, $\ker T_0 = \ker T$, so upon applying Theorem 1.1 to $T_0 : \ker ST \to Z$ we obtain:

$$\dim(\ker ST) \;=\; \dim(\ker T_0) + \dim(\operatorname{ran} T_0) \;=\; \dim(\ker T) + \dim(T(\ker ST))$$

from which follows (1), because

$$T(\ker ST) = T(T^{-1}(\ker S)) = \ker S \cap \operatorname{ran} T\,,$$

where the first equality follows from (b) above. □

Complements and Codimension

Definition 1.6 (Complements and direct sums)**.** Subspaces X_1 and X_2 of a vector space X are called *complementary* if $X = X_1 + X_2$ and $X_1 \cap X_2 = \{0\}$. When this happens we write $X = X_1 \oplus X_2$, call X_1 and X_2 (algebraic) *complements* of each other, and call X the (algebraic) *direct sum* of X_1 and X_2.

Here is a reinterpretation, in the language of complements, of the fact that any linearly independent set of vectors can be extended to a basis.

Proposition 1.7 (Complementary extension)**.** *If X is a vector space with subspaces M and N such that $M \cap N = \{0\}$, then $X = M \oplus \tilde{N}$ for some subspace $\tilde{N} \supset N$.*

Proof. If $M + N = X$ there is nothing to prove, so suppose this is not the case. Take bases B_M for M and B_N for N and note that, since $M \cap N = \{0\}$, the set $B_M \cup B_N$ is linearly independent. Thus we may choose a linearly independent set K of vectors so that $B_M \cup B_N \cup K$ is a basis for X. Then $\tilde{N} := \operatorname{span}(B_N \cup K)$ contains N, and when summed with M produces X. □

"External" direct sums. It is easy to check that $X = X_1 \oplus X_2$ if and only if each $x \in X$ is uniquely represented as a sum $x = x_1 + x_2$ where $x_j \in X_j$ $(j = 1, 2)$. Thus the map that associates x with the ordered pair (x_1, x_2) is an isomorphism taking X onto the cartesian product $X_1 \times X_2$; a vector space when endowed with coordinatewise operations.

Conversely, we can think of a cartesian product $X = X_1 \times X_2$ of vector spaces (over the same field) as the direct sum $X_1 \oplus X_2$ by identifying X_1 with the subspace $X_1 \times \{0\}$ and X_2 with $\{0\} \times X_2$. From now on we'll use the notation $X_1 \oplus X_2$ to denote either the previously defined "internal" direct sum of subspaces, or the just-defined "external" cartesian product of vector spaces, relying on context to distinguish the two notions.

Definition 1.8 (Codimension)**.** If M is any proper subspace of X, then whenever we complete a basis of M to one for X, the span of the added vectors will be a complement for M, and any such complement can be viewed as arising in this way. Thus the dimensions of all the complements of M are the same (possibly "infinity"). This common dimension is called the *codimension of* M in X, denoted by $\operatorname{codim}(M, X)$. If there is no confusion about the ambient space X, we'll just write $\operatorname{codim}(M)$ and call it the "codimension of M." In case $M = X$ we define $\operatorname{codim}(M, X) = 0$.

Exercise 1.9. Suppose $F \in \mathscr{L}(X, Y)$ has finite rank, and that $\dim X = \infty$. Show that $\ker F$ has infinite codimension in X. [2]

Throughout these notes it will be crucial for us to be able to compute, or at least to estimate, codimension. As a start, observe from the definition that if $X = M \oplus N$ then $\operatorname{codim} M = \dim N$, even if "$\infty$" is allowed; this yields the following *transitivity property* of codimension, which holds even if some of the codimensions involved are infinite:

If M and N are subspaces of a vector space X and $M \subset N \subset X$, then

$$\operatorname{codim}(M, X) = \operatorname{codim}(M, N) + \operatorname{codim}(N, X). \tag{2}$$

In particular:

Corollary 1.10 (Transitivity of finite codimension)**.** *If M has finite codimension in N and N has finite codimension in X then M has finite codimension in X.*

It will be useful to have several different ways to get a handle on codimension. Here is an estimate that involves sums that are not necessarily direct.

Proposition 1.11 (Codimension and sums)**.** *Suppose M and N are subspaces of X for which $X = M + N$. Then* $\operatorname{codim} M \leq \dim N$

Proof. Let B_0 be a basis for $M \cap N$ and extend by vectors B_1 to a basis for N. Let $N_1 = \operatorname{span} B_1$. Then N_1 is a subspace of N with $N_1 \cap M = N_1 \cap (M \cap N) = \{0\}$ (by the linear independence of $B_1 \cup B_2$). Thus $X = M \oplus N_1$, so $\operatorname{codim} M = \dim N_1 \leq \dim N$. □

Corollary 1.12. *Suppose M is a subspace of X and $T \in \mathscr{L}(X, Y)$. Then* $\operatorname{codim}(T(M), T(X)) \leq \operatorname{codim}(M, X)$

[2] For selected exercises: see Appendix B for hints and Appendix C for solutions.

Proof. We have $X = M \oplus N$ with $\operatorname{codim} M = \dim N$. Thus $T(X) = T(M) + T(N)$, so by Proposition 1.11: $\operatorname{codim}(T(M), T(X)) \leq \dim T(N) \leq \dim N$. □

Codimension also admits a characterization through *quotient spaces.*

Definition 1.13 (Quotient space)**.** If M is a subspace of the vector space X the *quotient* of X by M, written X/M, is the collection of cosets

$$x + M := \{x + m :\in M\}.$$

The quotient space X/M, when endowed with the algebraic operations it inherits from X, namely

$$(x + M) + (y + M) := (x + y) + M \quad \text{and} \quad \alpha(x + M) := (\alpha x) + M$$

for $x, y \in M$ and α a scalar, becomes a vector space over the original scalar field. The *quotient map* $Q : x \to x + M$ is then a linear transformation taking X onto X/M, and having kernel M. Upon applying Theorem 1.1 (the Fundamental Theorem of Linear Algebra) to this map we obtain the promised quotient-space characterization of codimension:

Proposition 1.14 (Codimension via quotients)**.** *If M is a subspace of a vector space X, then* $\operatorname{codim} M = \dim X/M$.

Proposition 1.14 provides an easy proof of a codimension calculation that we'll need in §4.

Corollary 1.15. *Suppose the vector space X is the (not necessarily direct) sum of two subspaces M and N. Then* $\operatorname{codim} M = \operatorname{codim}(M \cap N, N)$.

Proof. According to Proposition 1.14, another way to phrase the desired conclusion is: "$\dim X/M = \dim N/(M \cap N)$". To prove this we need only find an isomorphism between the featured quotient spaces. An obvious candidate springs to mind: Since $X = M + N$, each coset of X modulo M has the form $n + M$ for some vector $n \in N$. I leave it to you to check that the map $T : n + M \to n + (M \cap N)$ is "well-defined", linear, and takes the quotient space X/M onto $N/(M \cap N)$. That T is injective follows from the fact that if $T(n + M) := n + (M \cap N)$ is the zero-vector of $X/(M \cap N)$, i.e., the coset $(M \cap N)$, then $n \in M$, hence the original coset $n + M$ is just M, the zero-vector of X/M. □

Exercise 1.16 (Alternate proof)**.** Prove Corollary 1.15 by starting with a basis for $M \cap N$ and extending it two ways: to a basis for M and to a basis for N.

We will meet up with quotient spaces in a different context when in §4 we study the invertibility properties of Fredholm transformations.

Direct sums of linear transformations

Definition 1.17. Suppose X_1, X_2 are complementary subspaces of X, and Y_1, Y_2 are complementary subspaces of Y. Then for $T_j \in \mathscr{L}(X_j, Y_j)$ with $j = 1, 2$ we can define a linear transformation $T : X \to Y$ by $T(x_1 + x_2) = T_1x_1 + T_2x_2$. That this definition is unambiguous follows from the fact that each vector in X is uniquely the sum of a vector in X_1 and a vector in X_2. We write $T = T_1 \oplus T_2$ and refer to T as the *direct sum* of T_1 and T_2.

The notion of "external direct sum" of vector spaces (the cartesian product, endowed with coordinatewise algebraic operations) gives rise to a companion direct sum construction for linear transformations. If $T_j \in \mathscr{L}(X_j, Y_j)$ for $j = 1, 2$, then we define $T_1 \oplus T_2$ on the external direct sum $X_1 \oplus X_2$ by

$$(T_1 \oplus T_2)(x_1, x_2) = T_1x_1 + T_2x_2 \qquad (x_1 \in X_1, x_2 \in X_2).$$

As in the case of direct sums of vector spaces, we will rely on context to distinguish between the "external" and "internal" notions of direct sum for linear transformations.

Invertibility

We all know that a linear transformation $T : X \to Y$ is invertible (meaning: there exists a linear transformation $S : Y \to X$ such that $ST = I_X$ and $TS = I_Y$) if and only if T is both injective and surjective. Here is a result that generalizes this in two different ways.

Theorem 1.18 (One-sided Invertibility)**.** *$T \in \mathscr{L}(X, Y)$ is:*

(a) *Injective if and only if it is left-invertible (i.e., there exists $S \in \mathscr{L}(Y, X)$ such that $ST = I_X$).*
(b) *Surjective if and only if it is right-invertible (i.e., there exists $S \in \mathscr{L}(Y, X)$ such that $TS = I_Y$).*

Proof. That left-invertibility (resp. right-invertibility) implies injectivity (resp. surjectivity) is trivial. The converses require a little work.

(a) Suppose T is injective, so it is an invertible map when viewed as a transformation from X onto $\operatorname{ran} T$. Let $S_0 : \operatorname{ran} T \to X$ be the inverse of this linear transformation. Let N be a subspace of Y complimentary to $\operatorname{ran} T$, so that $Y = \operatorname{ran} T \oplus N$. Define $S : Y \to X$ by setting S equal to S_0 on $\operatorname{ran} T$ and (for example) equal to zero on N. Then S is a linear transformation $Y \to X$ with $ST = I_X$.

(b) Suppose T is surjective, i.e., $\operatorname{ran} T = Y$. Choose a subspace M of X complementary to $\ker T$; then $T|_M$ is an isomorphism of M onto Y. Let S denote the inverse of this isomorphism, but now regarded as a map taking Y into X. Then $S : Y \to X$ is a linear map with $TS = I_Y$. □

2. Linear Functionals

Here we'll continue the theme of finding different ways to compute codimension, this time via linear functionals.

A *linear functional* is a scalar-valued linear transformation on a vector space. The collection of all linear functionals on the vector space X is called *the dual space of* X, denoted herein by X'. Clearly X', endowed with pointwise operations, is itself a vector space over the original scalar field.

An important class of linear functionals arises whenever we encounter a basis. In order to cleanly generalize notions like "n-tuple" and "sequence" we use the terminology "indexed set" $\{v_a : a \in A\}$ to refer to a function v defined on a set A, where v_a denotes the value of v at $a \in A$.

Definition 2.1 (Coordinate functionals)**.** Recall that a (possibly infinite) indexed set of vectors $\{e_a : a \in A\}$ is a *basis* for a vector space X if and only if: for every $x \in X$ there is a unique indexed set $\{\Lambda_a(x) : a \in A\}$ of scalars, all but a finite number of which are zero, such that $x = \sum_{a \in A} \Lambda_a(x) e_a$. Thus for each $a \in A$ we have a function $\Lambda_a : X \to \{\text{scalars}\}$, and it's easily checked that this function is linear. The linear functionals $\{\Lambda_a : a \in A\}$ are called the *coordinate functionals* of the basis $\{e_a : a \in A\}$.

Definition 2.2 (Biorthogonality)**.** Indexed sets $\{e_a\}$ of vectors and $\{\Lambda_a\}$ of linear functionals are said to be *biorthogonal* whenever $\Lambda_a(e_b) = 0$ if $a \neq b$, and $= 1$ if $a = b$.

In the above definition we also say the vectors are biorthogonal to the functionals, and vice versa.

It's easy to check that whenever $\{e_a\}$ and $\{\Lambda_a\}$ are biorthogonal, then both are linearly independent. In particular, the coordinate functionals for a basis form a linearly independent set in the dual space.

An important property of linear functionals is their ability to separate points from subspaces.

Theorem 2.3 (Separation Theorem)**.** *Suppose M is a proper subspace of a vector space X, and $x \in X$ does not belong to M. Then there exists a linear functional Λ on X that vanishes on M, but not at x.*

Proof. Let M_1 be the linear span of M and x, so $M_1 = M \oplus \operatorname{span}\{x\}$. Define Λ on M_1 by

$$\Lambda(m + \lambda x) = \lambda \qquad (m \in M, \gamma \in \mathbb{F}).$$

Thus $\Lambda \equiv 0$ on M and $\Lambda(x) = 1$. If $M_1 = X$ we are done. Otherwise let N be a subspace of X complementary to M_1 (see Definition 1.6) and extend Λ to all of X, keeping the same name, by defining it to be zero, for example,[3] on N. □

The latter part of the argument above proves that:

> *Any linear functional on defined on a subspace can be extended to the whole space.*

Thus, for linear functionals the notions of separation and extension are inextricably entwined.

We study next the connection between kernels of linear functionals and codimension of subspaces. The first of these shows that nontrivial linear functionals have kernels of codimension one.

Proposition 2.4. *Suppose Λ is a linear functional on X, and $e \in X$ with $\Lambda(e) \neq 0$. Then $X = \ker \Lambda \oplus \operatorname{span}\{e\}$.*

Proof. By hypothesis $\ker \Lambda \cap \operatorname{span}\{e\} = \{0\}$. If $x \in X$ then

$$x - \frac{\Lambda(x)}{\Lambda(e)}\, e \in \ker \Lambda,$$

hence $x \in \ker \Lambda + \operatorname{span}\{e\}$. Thus $X = \ker \Lambda \oplus \operatorname{span}\{e\}$. □

Corollary 2.5 (Kernel containment I)**.** *Suppose Λ and Λ_1 are linear functionals on X. Then $\ker \Lambda \supset \ker \Lambda_1$ if and only if $\Lambda = c\Lambda_1$ for some scalar c.*

---

[3]... or, equally well, any other linear functional.

Proof. If Λ is a scalar multiple of Λ_1 then it's clear that the kernel of Λ contains that of Λ_1, so the issue is to prove the converse.

Suppose $\ker \Lambda \supset \ker \Lambda_1$. If Λ_1 is the zero-functional, then both kernels equal X, so both functionals are identically zero and the result is trivially true. Suppose, then, that Λ_1 is not identically zero, i.e., there exists $e \in X$ with $\Lambda_1(e) \neq 0$. Then according to Proposition 2.4, $X = \ker \Lambda_1 \oplus \operatorname{span}\{e\}$. We may, upon properly scaling e, assume that $\Lambda_1(e) = 1$. Thus if $x \in X$ we have $x = x_1 + \Lambda_1(x)e$, where $x_1 \in \ker \Lambda_1$. By our hypothesis on the containment of kernels, $x_1 \in \ker \Lambda$, so $\Lambda(x) = \Lambda_1(x)\Lambda(e)$. Since this is true for all $x \in X$ we have shown that $\Lambda = \Lambda(e)\Lambda_1$, hence the desired result holds with $c = \Lambda(e)$. $\square$

The next two results give significant generalizations of Corollary 2.5; the first leads to a short proof of the second.

Proposition 2.6 (Kernel containment II). *Suppose $T \in \mathscr{L}(X, Y)$ and $\Lambda \in Y'$. Then* $\ker T \subset \ker \Lambda$ *if and only if $\Lambda = \tilde{\Lambda} \circ T$ for some $\tilde{\Lambda} \in Y'$.*

Proof. If $\Lambda = \tilde{\Lambda} \circ T$ then the containment of kernels is obvious. For the converse, note that the hypothesis on kernel containment insures that the equation

$$\tilde{\Lambda}(Tx) := \Lambda(x) \qquad (x \in X)$$

defines a linear functional on $\operatorname{ran} T$ (the point being that if $Tx_1 = Tx_2$ then $x_1 - x_2 \in \ker T \subset \ker \Lambda$, and so $\Lambda(x_1) = \Lambda(x_2)$, i.e., the value of $\tilde{\Lambda}(Tx)$ depends only on Tx, and not on x.) If $\operatorname{ran} T = Y$ we are done. Otherwise choose a subspace N complementary to $\operatorname{ran} T$ and extend $\tilde{\Lambda}$, as in the proof of Theorem 2.3, to a linear functional on all of Y. The resulting functional, which we'll still call $\tilde{\Lambda}$, does the job! $\square$

Theorem 2.7 (Kernel containment III). *Let $\mathscr{E} = \{\Lambda_j : 1 \leq j \leq n\}$ be a finite subset of of X'. Then $\Lambda \in X'$ lies in* $\operatorname{span} \mathscr{E}$ *if and only if* $\ker \Lambda \supset \cap_{j=1}^{n} \ker \Lambda_j$.

Proof. It's clear that if Λ is a linear combination of the Λ_j's then its kernel contains the intersection of the kernels of the Λ_j's.

For the converse, suppose $\ker \Lambda \supset \cap_{j=1}^{n} \ker \Lambda_j$. Let Y be the direct sum of n copies of the scalar field, and define $T \in \mathscr{L}(X, Y)$ by

$$Tx = (\Lambda_1(x), \Lambda_2(x), \ldots \Lambda_n(x)) = \sum_{j=1}^{n} \Lambda_j(x) e_j \qquad (x \in X),$$

where, in the last equality, e_j denotes the j-th standard basis vector for Y, i.e., the vector whose j-th coordinate is 1, with all other coordinates equal to zero. Since the kernel of T is just the intersection of the kernels of the Λ_j's, our hypothesis on the containment of kernels simply asserts that $\ker \Lambda \supset \ker T$, whereupon Proposition 2.6 provides $\tilde{\Lambda} \in Y'$ such that $\Lambda = \tilde{\Lambda} \circ T$. Thus for each $x \in X$:

$$\Lambda x = \tilde{\Lambda}(Tx) \;=\; \tilde{\Lambda}\left(\sum_j \Lambda_j(x) e_j\right) = \sum_j \Lambda_j(x)\tilde{\Lambda}(e_j) = \left(\sum_j c_j \Lambda_j\right)(x)$$

where $c_j = \tilde{\Lambda}(e_j)$. Thus $\Lambda = \sum_j c_j \Lambda_j$, i.e., Λ lies in the span of the Λ_j's. □

Exercise 2.8 (Counterexample for $n = \infty$). Show that the nontrivial part of Theorem 2.7 cannot be extended to the case of infinitely many linear functionals Λ_n.

Suggestion: Work on the space ω of all scalar sequences, and for $n \in \mathbb{N}$ let Λ_n be the linear functional of "evaluation at n" (or the "n-th coordinate functional" if you prefer to think of sequences as lists): $\Lambda_n(x) = x(n)$ for $x \in \omega$.

Theorem 2.7 can be rephrased like this:

> *A finite set of linear functionals is linearly independent if and only if the kernel of no one of these functionals contains the intersection of the kernels of the others.*

Equivalently, $\{\Lambda_1, \ldots \Lambda_n\} \subset X'$ is linearly independent if and only if: for each j between 1 and n there exists a vector e_j such that $\Lambda_j(e_j) = 1$ and $\Lambda_j(e_k) = 0$ if $j \neq k$ $(1 \leq k \leq n)$, i.e., the n-tuple of vectors $(e_1 \ldots e_n)$ is *biorthogonal* to the n-tuple of Λ's (see Definition 2.2). The language of biorthogonality allows a further rephrasing of Theorem 2.7:

Corollary 2.9 (Independence and biorthogonality). *An n-tuple of linear functionals is linearly independent if and only if it has a biorthogonal n-tuple of vectors.*

Given a biorthogonal system of vectors $\{e_\alpha : \alpha \in A\}$ and linear functionals $\{\Lambda_\alpha : \alpha \in A\}$, we can reverse roles and regard the vectors as functionals on the functionals, e.g., $e_\alpha(\Lambda_\beta) := \Lambda_\beta(e_\alpha)$. Thus Corollary 2.9 implies that not only are the functionals linearly independent, so are the vectors (a fact which, as we noted on page 9, is easy to prove directly).

The results established so far set the stage for our main result on the computation of codimension via linear functionals.

Theorem 2.10 (Kernels and codimension)**.** *A subspace M of the vector space X has codimension n (possibly $= \infty$) if and only if M is the intersection of the kernels of a linearly independent set of n linear functionals on X.*

Proof. We'll consider finite and infinite codimension separately.

CASE I: $n < \infty$. Suppose first that $\{\Lambda_1, \Lambda_2, \dots \Lambda_n\}$ is a linearly independent subset of X', and $M := \cap_{j=1}^n \ker \Lambda_j$. By Corollary 2.9 the n-tuple $(\Lambda_1, \dots \Lambda_n)$ has a biorthogonal n-tuple of vectors $(e_1, \dots e_n)$ which, as we have noted previously, forms a linearly independent subset of X. Thus the linear span E of these vectors is a subspace of X having dimension n. Suppose $x \in E \cap M$. Then $x = \sum_{j=1}^n \Lambda_j(x) e_j$ since $x \in E$, but $\Lambda_j(x) = 0$ for all j since $x \in M$. Thus $x = 0$, i.e., E has trivial intersection with M. That $X = M + E$ follows immediately from the fact that, for each $x \in X$, the vector $x - \sum_{j=1}^n \Lambda_j(x) e_j$ lies in M. Thus $X = M \oplus E$, hence $\operatorname{codim} M = \dim E = n$, as desired.[4]

Conversely, suppose M is any subspace of X having codimension n. Then there is a subspace E of X having dimension n for which $X = M \oplus E$. Let $\{e_1, e_2, \dots e_n\}$ be a basis for E, and let $\{\Lambda_1, \Lambda_2, \dots \Lambda_n\}$ be the coordinate functionals for this basis, i.e.,

$$x = \sum_{j=1}^{n} \Lambda_j(x) e_j \qquad (x \in E).$$

Extend each Λ_j to a linear functional on X (keeping the same name) by defining it to be the zero-functional on M. The collection of extended Λ_j's inherits the linear independence of the original one and, by the way we extended the Λ_j's, the subspace $K := \cap_{j=1}^n \ker \Lambda_j$ contains M. In fact, K is *equal to* M. To see why, suppose $x \in K$, so $x = m + e$ where $m \in M$ and $e \in E$. Now for any index j, the definition of K and the fact that $M \subset K$ imply that $0 = \Lambda_j(x) = \Lambda_j(e)$. Thus $e = 0$ (since the Λ_j's are coordinate functionals for a basis of N), hence $x \in M$. Thus K both contains and is contained in M, i.e., $K = M$.

CASE II: $n = \infty$. Suppose M is the intersection of the kernels of an infinite linearly independent subset $\Lambda_a : a \in A\}$ of X'. Given $n \in \mathbb{N}$ choose a subset A_n of the index set A having cardinality n. Then the intersection of the kernels

[4]See Exercise 2.11 below for more on the expression $\sum_{j=1}^n \Lambda_j(x) e_j$ which appeared in this argument.

of the functionals Λ_a for $a \in A_n$ contains M, and by Case I has codimension n. Thus $\operatorname{codim} M \geq n$. Since n is an arbitrary positive integer, $\operatorname{codim} M = \infty$.

Conversely, suppose $\operatorname{codim} M = \infty$. Then M has an infinite dimensional complement N, which has a basis $\{e_a : a \in A\}$ for some infinite index set A. Now proceed as in Case I: extend the coordinate functionals $\{\Lambda_\alpha : \alpha \in A\}$ for this basis by defining each extension to be zero on M, observe that the extended functionals are linearly independent, and that M lies in the intersection K of their kernels. The same argument used in Case I (whose transposition to the case $n = \infty$ I leave to the reader) now implies that $K = M$, and completes the proof. □

Exercise 2.11 (Projections). Suppose M is a subspace of a vector space X and $P \in \mathscr{L}(X)$ is the identity operator on M, and the zero-operator on some complementary subspace E of M. We call P the *projection of X onto M along E.*

(a) Make some sketches in $\mathbb{R}^2$ to illustrate this situation.
(b) Suppose M and E are subspaces of the vector space X with $X = M \oplus E$, so that each $x \in X$ has the unique representation $x = e + m$ with $e \in E$ and $m \in M$. Show that the mapping $P : x \to e$ is the projection of X onto E along M.
(c) Show that $P \in \mathscr{L}(X)$ is a projection (onto its range, along its kernel) if and only if $P^2 = P$.
(d) Suppose $X = M \oplus E$ and that $\{e_\alpha : \alpha \in A\}$ is a basis for E, with coordinate functionals $\{\Lambda_\alpha \alpha \in A\}$. As in the proof of Case I of Theorem 2.10, extend each coordinate functional to X by defining it to be zero on M. Show that the map $P : x \to \sum_\alpha \Lambda_\alpha(x) e_\alpha$ is the projection of X onto E along M.

Polars

For a subset S of the vector space X, the *polar* of S, denoted S°, is the set of linear functionals on X that take the value zero at each point of S. One checks easily that the polar of any subset of X is a subspace of X', and that the polar of a set coincides with the polar of its linear span.

Theorem 2.12 (Polars and codimension). *Suppose X is a vector space and M a subspace of X. Then* $\operatorname{codim} M = \dim M^\circ$.

Proof. First suppose $\operatorname{codim} M = n < \infty$. Then by Theorem 2.10, $M = \cap_{j=1}^{n} \ker \Lambda_j$, where the set of functionals $\{\Lambda_j\}_1^n$ is linearly independent in X'. Since each Λ_j annihilates M, it lies in M°. Thus $M^\circ \supset \operatorname{span}\{\Lambda_1, \Lambda_2, \ldots \Lambda_n\}$. On the other hand, if $\Lambda \in M^\circ$ then $\ker \Lambda$ contains in M, so Λ, by Theorem 2.7, is a linear combination of the Λ_j's. Thus $\operatorname{span}\{\Lambda_1, \Lambda_2, \ldots \Lambda_n\}$ contains—and is therefore equal to—M°, hence $\dim M^\circ = n$, as desired.

If, on the other hand, $\operatorname{codim} M = \infty$, then Theorem 2.10 guarantees that M is the intersection of the kernels of an infinite set of linearly independent linear functionals on X. Since each of these belongs to M° we see that M° has infinite dimension. □

Is there a similar connection between the codimension of M° in X' and the dimension of M? The next result provides the key.

Proposition 2.13 (Dual of a direct sum)**.** *If M and N are subspaces of X with $X = M \oplus N$, then $X' = M^\circ \oplus N^\circ$.*

Proof. For $\Lambda \in X'$, let Λ_1 be the linear functional obtained by restricting Λ to M, then extending this restriction to X by defining it to be zero on N. Define Λ_2 by similarly extending the restriction of Λ to N. Thus $\Lambda_1 \in N^\circ$, $\Lambda_2 \in M^\circ$, and $\Lambda = \Lambda_1 + \Lambda_2$. Consequently $X' = M^\circ + N^\circ$.

As for the "directness" of this sum, note that if $\Lambda \in M^\circ \cap N^\circ$ then Λ is identically zero on both M and N, hence is identically zero on all of X. Thus $M^\circ \cap N^\circ = \{0\}$. □

Corollary 2.14 (Polars and codimension)**.** *If M is a subspace of a vector space X, then* $\dim M = \operatorname{codim} M^\circ$.

Proof. Choose a complementary subspace N of M, so that $X = M \oplus N$ and $\dim M = \operatorname{codim} N$. Theorem 2.12 asserts that $\operatorname{codim} N = \dim N^\circ$, while Proposition 2.13 provides the decomposition $X' = M^\circ \oplus N^\circ$, from which it follows that $\dim N^\circ = \operatorname{codim} M^\circ$.

In summary: $\dim M = \operatorname{codim} N = \dim N^\circ = \operatorname{codim} M^\circ$, as promised. □

Adjoints

Proposition 2.6 suggests that for $T \in \mathscr{L}(X, Y)$ we might profitably study the map T' defined by $T'\Lambda = \Lambda \circ T$ for $\Lambda \in Y'$. Clearly T', called the *adjoint* of T, is a linear transformation from Y' to X'.

Furthermore, in Proposition 2.6 the statement $\ker T \subset \ker \Lambda$ asserts that the linear functional Λ belongs to the polar of $\ker T$. Part (a) of the next result simply rephrases Proposition 2.6 in the language of adjoints and polars.

Theorem 2.15 (Polars and adjoints). *For $T \in \mathscr{L}(X, Y)$:*

(a) $(\ker T)^\circ = \operatorname{ran} T'$
(b) $(\operatorname{ran} T)^\circ = \ker T'$

Proof. (a) As noted above, the statement $\ker T \subset \ker \Lambda$ translates to: "$\Lambda \in (\ker T)^\circ$", and the equivalent (according to Proposition 2.6) statement "$\Lambda = T \circ \tilde{\Lambda}$" becomes "$\Lambda = T'(\tilde{\Lambda})$".

(b) The statement "$\Lambda \in (\operatorname{ran} T)^\circ$" means that Λ takes the value zero on $\operatorname{ran} T$, i.e., that $T'(\Lambda)(x) := \Lambda(Tx) = 0$ for every $x \in X$, i.e., that $T'\Lambda = 0$. Thus $\Lambda \in (\operatorname{ran} T)^\circ$ if and only if $\Lambda \in \ker T'$. □

Exercises 2.16. Suppose $T : X \to Y$ is a linear transformation.

(a) Show that T is invertible if and only if $T' : Y' \to X'$ is invertible.
(b) Examine the connection between left and right invertibility for T and for T'.

3. Stabilization

In this section we'll characterize the important class of linear transformations on a vector space that have the form "isomorphism $\oplus$ nilpotent."

For $T \in \mathscr{L}(X)$ and n a non-negative integer let $R_n := \operatorname{ran} T^n$ and $K_n := \ker T^n$. Then the "range sequence" (R_n) is a decreasing sequence of subspaces of X:

$$X = R_0 \supset R_1 \supset R_2 \supset \cdots \tag{3}$$

while the "kernel sequence" (K_n) is an increasing sequence of subspaces:

$$\{0\} = K_0 \subset K_1 \subset K_2 \subset \cdots \tag{4}$$

More generally let's say that say a monotonic sequence (E_n) of sets:

- Is *stable at index* N if $E_N = E_{N+1} = \cdots$,
- *Stabilizes at index* N if N is the smallest index at which it is stable, and just plain ...
- *Stabilizes* if it stabilizes at some index.

Furthermore, let's say that a transformation $T \in \mathscr{L}(X)$ is *stable* if both its range and kernel sequences stabilize. Clearly any linear transformation on a finite dimensional vector space is stable. On the other hand, the forward and backward shifts on the sequence space ω are *not* stable.

Examples of stable transformations that will play important roles in our later work are the nilpotent ones (some positive power is zero) and transformations of the form "finite-rank plus identity," the so-called *finite-rank perturbations of the identity.*

Proposition 3.1. *Nilpotent transformations are stable.*

Proof. Suppose $T \in \mathscr{L}(X)$ and $T^N = 0$ for some $N \in \mathbb{N}$. Then for all integers $n \geq N$ we have $K_N = X$ and $R_N = \{0\}$. □

Proposition 3.2. *Finite rank perturbations of the identity are stable*

Proof. Suppose $T = I + F$ where I is the identity transformation and F is a finite rank transformation, both acting on X. Then $\ker T \subset \operatorname{ran} F$ and, since $T = I$ on $\ker F$, we have $\operatorname{ran} T \supset \ker F$.

Now for each $n \in \mathbb{N}$ the Binomial Theorem guarantees that $T^n = I + G_n F = I + F G_n$ where $G_n \in \mathscr{L}(X)$ commutes with F. Upon applying the result of the above paragraph, with T^n in place of T, we see that $K_n \subset \operatorname{ran} F G_n \subset \operatorname{ran} F$ and $R_n \supset \ker G_n F \supset \ker F$. Since $\operatorname{ran} F$ is finite dimensional and $\ker F$ finite codimensional, both the (decreasing) range sequence and the (increasing) kernel sequence of T must stabilize. □

Corollary 3.3. *Any direct sum of a nilpotent linear transformation and an isomorphism is stable.*

Proof. Suppose $V \in \mathscr{L}(X)$ is an isomorphism of X onto itself, and $N \in \mathscr{L}(Y)$ is nilpotent, say of order ν. Let $T = V \oplus N \in \mathscr{L}(X \oplus Y)$. Then (adopting the "external" point of view) for all $n \geq \nu$ we have $\operatorname{ran} T^\nu = X \times \{0\}$ and $\ker T^\nu = \{0\} \times Y$. □

The following remarkable structure theorem provides the converse to Corollary 3.3; it shows that stability is, in fact, *equivalent* to "isomorphism ⊕ nilpotent."

Theorem 3.4 (The Stabilization Theorem)**.** *Suppose that for $T \in \mathscr{L}(X)$ is stable. Then:*

(a) *Both the range and kernel sequences stabilize at the same index—call it* ν.
(b) $X = R_\nu \oplus K_\nu$.
(c) $T = T_1 \oplus T_2$, *where* $T_1 := T|_{R_\nu}$ *is an isomorphism of* R_ν *into itself, and* $T_2 := T|_{K_\nu}$ *is nilpotent of order* ν.

To say a linear transformation is *nilpotent of order* ν means that $T^\nu = 0$, but $T^{\nu-1} \neq 0$.

The proof is best broken into a number of steps, some of which are interesting results in their own right. In the first two steps we do not assume the full stability of T.

Step I: *Stabilization somewhere.* It's obvious from the definition $R_n := T^n(X)$ that if $R_n = R_{n+1}$ then $R_{n+1} = R_{n+2}$. Thus, if the range sequence stabilizes at all, it stabilizes at the first place where there is equality between two successive terms.

A similar result for the kernel sequence follows from the fact that, in general, $\ker(ST) = T^{-1}(\ker S)$. In particular, $K_{n+1} = T^{-1}(K_n)$ for each n, hence

$$K_{n+1} = K_n \implies T^{-1}(K_{n+1}) = T^{-1}(K_n) \quad \text{i.e.,} \quad K_{n+2} = K_{n+1}\,.$$

Thus—just as for the range sequence—if the kernel sequence stabilizes at all, it stabilizes at the first place where two successive terms are equal. □

Step II: *Equivalents to stabilization.* For $n = 0, 1, 2, \ldots,$

(a) $K_n = K_{n+1} \iff K_n \cap R_n = \{0\}$
(b) $R_n = R_{n+1} \iff K_n + R_n = X$.

Proof. (i) Suppose $K_n \cap R_n = \{0\}$; we wish to show that $K_n = K_{n+1}$. Since $K_n \subset K_{n+1}$ we need only show the opposite containment. For this, suppose $x \in K_{n+1}$, so $T^{n+1}x = 0$. Then

$$T^n x \in R_n \cap K_1 \subset R_n \cap K_n = \{0\},$$

so $x \in K_n$.

Suppose, conversely, that $K_n = K_{n+1}$. Then if $x \in K_n \cap R_n$ we have $T^n x = 0$ and $x = T^n x'$ for some $x' \in X$. Thus x' belongs to K_{2n} which, by Step I, is equal to K_n. So $x = T^n x' = 0$, hence $K_n \cap R_n = \{0\}$, as desired.

(ii) Suppose $X = R_n + K_n$. Upon applying T^n to both sides of this equation we obtain $R_n := T^n(X) = T^n(R_n)$, that is, $R_n = R_{2n}$. But $R_{2n} \subset R_{n+1} \subset R_n$. Trivially $R_n \subset R_{n+1}$, so $R_n = R_{n+1}$ as desired.

Suppose, conversely, that $R_n = R_{n+1}$. Then by Step I, $R_n = R_{2n}$, so given $x \in X$ we have $T^n x = T^{2n} x'$ for some $x' \in X$, whereupon $T^n(x - T^n x') = 0$, i.e., $x - T^n x' \in K_n$. Consequently $x \in K_n + \{T^n x'\} \subset K_n + R_n$. Thus $X = K_n + R_n$, as desired. □

The next step shows that if the range sequence stabilizes from the very beginning, then so does the kernel sequence, and *vice versa.*

Step III: *The "Fredholm Alternative-II"* (cf. Corollary 1.2).

(i) If T is surjective and its kernel sequence stabilizes, then T is injective.
(ii) If T is injective and its range sequence stabilizes, then T is surjective.

Proof. (i) Suppose T is surjective and (K_n) is stable at index N, i.e., $K_N = K_{N+1} = \cdots$. By surjectivity, the range sequence is stable at index zero, i.e., $R_n = X$ for all n, so by Step II(ii), $\{0\} = K_N \cap R_N = K_N$. Therefore T^N is injective, hence so is T.

(ii) Suppose T is injective, and that the range sequence is stable at index N. By Step II(i) we have $X = R_N + K_N$. But $K_N = \{0\}$ since T, and hence T^N is injective. Thus $X = R_N$, i.e., T^N is surjective, and therefore so is T. □

Step IV: *Completion of proof of Part (a).* (i) Suppose the range sequence is stable at index N and the kernel sequence stabilizes. Then $T_N := T^N|_{R_N}$ maps R_N onto itself, hence by "Fredholm Alternative-II" (Step III above) it is injective, i.e.,

$$\{0\} = \ker T_N = \ker T^N \cap R_N = K_N \cap R_N.$$

Thus by Step II(i) the kernel sequence is stable at index N. Thus the kernel sequence stabilizes no later than the range sequence.

(ii) Suppose the kernel sequence is stable at index N and the range sequence is stable somewhere, say at index M, so by Step II(ii) $X = K_M + R_M$. By part (i) above, $N \leq M$, so

$$X = K_M + R_M = K_N + R_M \subset K_N + R_N \subset X,$$

so there is equality throughout, i.e., $X = K_N + R_N$. STEP II(i) now guarantees that $R_N = R_{N+1}$, i.e., the range sequence can stabilize no later than the kernel sequence.

(iii) Summarizing: If both the range and kernel sequences stabilize, then the kernel sequence can stabilize no later than the range sequence (part (i) above), and the range sequence can stabilize no later than the kernel sequence (part(ii)). Hence they must both stabilize at the same index. This completes the proof of Part (a) of the Theorem. □

STEP V: *Direct sum decomposition.* The fact that $X = R_\nu \oplus K_\nu$ now follows immediately from STEP II.

STEP VI: *Isomorphism plus nilpotent.* By STEP IV both the range and kernel sequences for T stabilize at the same index ν, and by STEP V we have $X = R_\nu \oplus K_\nu$. Since T_1, the restriction of T to R_ν, maps R_ν *onto* itself and, since $\ker T \subset K_\nu$, T_1 is also injective, hence invertible on R_ν. As for T_2, the restriction of T to K_ν, the fact that $K_\nu = \{0\}$ means that $T_2^\nu = 0$, while the strict containment of $K_{\nu-1}$ in K_ν means that $T_2^{\nu-1} \neq 0$. Thus T_2 is nilpotent of index ν, and T is the direct sum of T_1 and T_2. This completes the proof of Theorem 3.4. □

For a linear transformation T on a finite dimensional vector space the range and kernel sequences must stabilize. In this situation the above Corollary provides the first step in the Jordan decomposition of T. Here one works over an algebraically closed field, say the complex numbers, and given an eigenvalue λ for T performs the decomposition of Theorem 3.4(c) on the transformation $T - \lambda I$. Having done that, one moves on to the isomorphic part of that operator, showing that its underlying subspace is also T-invariant. Then the process is repeated for the next eigenvalue, and proceeds until T is decomposed into the direct sum of operators of the form "nilpotent plus eigenvalue times identity". The rest of the proof of the Jordan decomposition then involves decomposing the nilpotent parts into cyclic sub-parts, see [1, Chapter 8] for example.

4. Fredholm Transformations

In this section we develop the basic theory of Fredholm linear transformations with emphasis on the connection between Fredholmness and invertibility, which we eventually encapsulate in the notion of Fredholm index.

Definition 4.1 (Fredholm transformations)**.** By a *Fredholm transformation* we mean a linear transformation between vector spaces X and Y whose kernel is finite dimensional and whose range has finite codimension in Y. The class of Fredholm transformations from X into Y will be denoted $\Phi(X, Y)$, with $\Phi(X, X)$ abbreviated to $\Phi(X)$.

Examples 4.2 (Important Fredholm transformations)**.** Clearly every linear transformation between finite dimensional vector spaces is Fredholm, as is every *invertible* linear transformation between any two vector spaces. The zero transformation is Fredholm if and only if the underlying space is finite dimensional.

Here are three further important examples.

(a) *The Forward Shift.* This is the transformation S defined in §1.3(a) on the vector space ω of all scalar sequences; it has trivial kernel and range of codimension one, and so is Fredholm.

(b) *The Backward Shift.* This is the transformation B defined on ω in §1.3(b); it has range ω and kernel of dimension one, hence is Fredholm.

(c) *Finite rank perturbations of the identity.* In the proof of Proposition 3.2 we observed that if $T \in \mathscr{L}(X)$ has the form $I + F$ where I is the identity on X and $F \in \mathscr{L}(X)$ has finite rank, then: $\operatorname{ran} T$ contains $\ker F$, so has finite codimension, and $\ker T$ is contained in $\operatorname{ran} F$, and so has finite dimension. Thus T is Fredholm.

Exercise 4.3. Characterize those pairs of vector spaces (X, Y) for which $\Phi(X, Y)$ is non-empty.

Exercise 4.4 (Direct sums)**.** Suppose $T_j \in \mathscr{L}(X_j, Y_j)$ for $j = 1, 2$. Let $X = X_1 \oplus X_2$ as in §1.17. Show that T is Fredholm if and only if both T_1 and T_2 are Fredholm.

Theorem 4.5 (The Product Theorem)**.** *If* $T \in \Phi(X, Y)$ *and* $S \in \Phi(Y, Z)$, *then* $ST \in \Phi(X, Z)$.

Proof. We need to show that $ST : X \to Z$ has finite dimensional kernel and finite codimensional range. The finite dimensionality of the kernel follows from Theorem 1.5. For the finite codimensionality of the range, note that since $\operatorname{ran} T$ has finite codimension in X, Corollary 1.12 (with $Y = \operatorname{ran} S$) guarantees that $\operatorname{ran} ST = S(\operatorname{ran} T)$ has finite codimension in $S(X) = \operatorname{ran} S$. Since $\operatorname{ran} S$

has finite codimension in Z, the finite codimension of $\operatorname{ran} ST$ in Z follows by Corollary 1.10 (transitivity). □

Here are a couple of exercises that explore the possibility of a converse to the Product Theorem.

Exercise 4.6 (Fredholm Roots)**.** Suppose $T \in \mathscr{L}(X)$, $k \in \mathbb{N}$, and $T^k \in \Phi(X)$. Show that $T \in \Phi(X)$.

Exercise 4.7 (Fredholm Factors)**.** Suppose $S, T \in \mathscr{L}(X)$ with $ST \in \Phi(X)$. Must S and T be Fredholm?

Invertibility

Consider once again the forward shift S and the backward shift B (see Examples 1.3 and 4.2), both acting on the space ω of all sequences of field elements. Note that $BS = I_\omega$ while SB differs from I_ω by a rank-one transformation. Thus S and B are, in some sense, "invertible modulo finite rank transformations." The next two theorems, which make this notion precise, assert that such "almost-invertibility" is synonymous with being Fredholm.

Theorem 4.8 (First Invertibility Theorem)**.** *$T \in \mathscr{L}(X, Y)$ is Fredholm if and only if there exists a linear transformation $S : Y \to X$ such that both $ST - I_X$ and $TS - I_Y$ have finite rank.*

Proof. (a) Suppose T is Fredholm. Let X_1 be a complement for $\ker T$ in X, and Y_0 a complement for $\operatorname{ran} T$ in Y. Because T is Fredholm, $\ker T$ is finite dimensional and $\operatorname{ran} T$ finite codimensional. The diagram below summarizes the situation; here T_1, the restriction of T to X_1, is an isomorphism taking X_1 onto $\operatorname{ran} T$.

$$
\begin{array}{cccl}
X & \xrightarrow{T} & Y & \quad\text{Fredholm} \\
\| & & \| & \\
\ker T & \xrightarrow{0} & Y_0 & \quad\text{Finite dim'l} \\
\bigoplus & & \bigoplus & \\
X_1 & \xrightarrow{T_1} & \operatorname{ran} T & \quad\text{Isomorphism}
\end{array}
$$

The Fundamental Fredholm Diagram

Define $S : Y \to X$ as follows: $S = T_1^{-1}$ on $\operatorname{ran} T$ and $S = 0$ on Y_0. Then one checks that ST is the identity map on the subspace X_1, and $ST \equiv 0$ on

$\ker T$. In other words, ST is the "projection of X onto X_1 along $\ker T$." Thus $F_1 := I_X - ST$ is the projection of X onto $\ker T$ along X_1; it is a finite rank transformation, so $ST = I_X - F_1$, a finite rank perturbation of the identity.

Similarly: TS coincides with I_Y on $\operatorname{ran} T$, and is $\equiv 0$ on Y_0, so TS is the projection of Y onto $\operatorname{ran} T$ along Y_0. Thus $F_0 := I_Y - TS$ is the projection of Y onto Y_0 along $\operatorname{ran} T$, another finite rank transformation, which reveals $TS = I_Y - F_0$ to be a finite rank perturbation of the identity.

(b) For the converse we recycle the idea used in Proposition 3.2 and Example 4.2(c). Suppose there exists $S \in \mathscr{L}(Y, X)$ such that $ST - I_X = F_1$ and $TS - I_X = F_0$, where F_0 and F_1 are finite-rank operators. Then $\ker T \subset \operatorname{ran} F_1$ so $\ker T$ has finite dimension, and $\operatorname{ran} T \supset \ker F_0$ so (by Theorem 1.1) $\operatorname{ran} T$ has finite codimension. Thus T is Fredholm. □

Now that the Theorem is proved, note that it guarantees that the transformation S in its hypothesis is necessarily Fredholm. More importantly, it guarantees that "Fredholmness is stable under finite rank perturbation:"

Corollary 4.9 (Finite-rank Perturbation Theorem I). *If $T \in \Phi(X, Y)$ and $F \in \mathscr{F}(X, Y)$ then $T + F \in \Phi(X, Y)$.*

Proof. We know from Theorem 4.8 that there exists $S \in \mathscr{L}(Y, X)$ that is an "almost-inverse" of T in the sense that there exist finite rank linear transformations $F_1 : X \to X$, and $F_2 : Y \to Y$ such that $ST - I_X = F_1$ and $TS - I_Y = F_2$. The idea here is to show that S is also an "almost inverse" for $T + F$. Indeed,

$$S(T + F) - I_X = ST - I_X + SF = F_1 + SF,$$

and

$$(T + F)S - I_Y = TS - I_Y + FS = F_2 + FS,$$

Since both SF and FS have finite rank, so do $S(T+F)-I_X$ and $(T+F)S-I_Y$, hence $T + F$ is Fredholm by the converse implication of Theorem 4.8. □

Quotient Algebra Interpretation. The vector space $\mathscr{L}(X)$ of all linear transformations on X is an algebra over the scalar field, and the subspace $\mathscr{F}(X)$ of finite rank transformations is a two-sided ideal in that algebra. It follows that the quotient space $\mathscr{L}(X)/\mathscr{F}(X)$ is an algebra over the same field. In this setting the First Invertibility Theorem asserts that:

> *A linear transformation on X is Fredholm if and only if its coset in $\mathscr{L}(X)/\mathscr{F}(X)$ is invertible.*

By paying a little more attention to the proof of the First Invertibility Theorem we can considerably refine that result.

Theorem 4.10 (Second Invertibility Theorem). *Suppose* $T \in \Phi(X, Y)$.

(a) *If* $\dim\ker T = \operatorname{codim}\operatorname{ran} T$*, then some finite rank perturbation of* T *is invertible*
(b) *If* $\dim\ker T < \operatorname{codim}\operatorname{ran} T$ *then some finite rank perturbation of* T *is left-invertible*
(c) *If* $\dim\ker T > \operatorname{codim}\operatorname{ran} T$ *then some finite rank perturbation of* T *is right-invertible*

Proof. Referring to the "Fundamental Fredholm Diagram" on page 21:

(a) If $\dim\ker T = \operatorname{codim}\operatorname{ran} T$ then, in the diagram, $\ker T$ and $\dim Y_0$ both have the same (finite) dimension, so there is an isomorphism F taking $\ker T$ onto Y_0. Extend F to all of X (keeping the same name) by setting $F \equiv 0$ on X_1. Thus $F \in \mathscr{F}(X, Y)$, and $\tilde{T} := T + F$ is a linear transformation from X to Y that is equal to T on X_1, and to F on $\ker T$. Consequently $\tilde{T}$ is an injective linear map taking X onto Y, so it is invertible.

(b) Suppose $\dim\ker T < \operatorname{codim}\operatorname{ran} T$, i.e., $\dim\ker T < \dim Y_0$. Then there is an injective linear map $F : \ker T \to Y_0$. Define $\tilde{T}$ as before, and observe that it is an injective linear map $X \to Y$. Therefore, by Theorem 1.18(a), $\tilde{T}$ has a left-inverse.

(c) Suppose $\dim\ker T > \operatorname{codim}\operatorname{ran} T$, i.e., $\dim\ker T > \dim Y_0$. Then there is a finite rank transformation F taking $\ker T$ *onto* Y_0, so upon defining $\tilde{T} : X \to Y$ as before we see that $\tilde{T}$ maps X onto Y. Thus, by Theorem 1.18(b), $\tilde{T}$ has a right-inverse. □

These arguments emphasize the fundamental property of Fredholm transformations: *their properties depend ultimately on finite dimensional phenomena.* The Second Invertibility Theorem shows that, for a Fredholm transformation T, the difference between the dimension of its kernel and the codimension of its range gives a precise description of the "invertibility properties" of T. Let's now formalize this difference and study its properties.

Fredholm Index

Definition 4.11 (Index). For $T \in \Phi(X, Y)$ the *(Fredholm) index* of T is:

$$i(T) := \dim\ker T - \operatorname{codim}\operatorname{ran} T.$$

Before proceeding further, let's take a look at some examples.

Exercise 4.12 (First examples)**.** Compute the Fredholm index for each of our three primary examples: invertible transformations, and the forward and backward shifts of in §1.3 and §4.2.

Theorem 4.13. *Every stable*[5] *Fredholm transformation has index zero.*

Proof. According to Theorem 3.4, if T is stable then it has the form $V \oplus N$ where V is an isomorphism and N is nilpotent. Since T is Fredholm, its nilpotent summand is Fredholm, so must live on a finite dimensional subspace (cf. Exercise 4.24), and must therefore have index zero. The isomorphic summand is, of course, Fredholm of index zero, hence $i(T) = 0$ by Exercise 4.17. □

Every finite-rank perturbation of the identity is Fredholm (Examples 4.2(c)) and stable (Theorem 3.2). Thus by Theorem 4.13:

Corollary 4.14. *Every finite rank perturbation of the identity has index zero.*

Exercise 4.15 (Range of index function)**.** Show that if X is an infinite dimensional vector space then for every integer n there exists a Fredholm transformation T on X with $i(T) = n$.

Exercise 4.16 (Index of adjoint)**.** If $T \in \mathscr{L}(X, Y)$ is Fredholm, show that the adjoint $T' \in \mathscr{L}(X', Y')$ is also Fredholm, with $i(T') = -i(T)$.

Exercise 4.17 (Index of direct sum)**.** Continuing Exercise 4.4: Show that if linear transformations $T_j : X_j \to Y_j$ are Fredholm for $j = 1, 2$, then $T_1 \oplus T_2$ (which is Fredholm by Exercise 4.4) has index $i(T_1) + i(T_2)$.

The notion of index allows a succinct restatement the Second Invertibility Theorem of the last section.

Theorem 4.18. *Suppose* $T \in \Phi(X, Y)$.

(a) *If* $i(T) = 0$ *then some finite rank perturbation of* T *is invertible.*
(b) *If* $i(T) < 0$ *then some finite rank perturbation of* T *is left invertible.*
(c) *If* $i(T) > 0$ *then some finite rank perturbation of* T *is right invertible.*

It is natural to ask if, for each of the statements of Theorem 4.18, the *converse* is true. For example the converse of part (a) would assert that if $T + F$ is

[5]see page 15

invertible for some finite rank transformation F, then $i(T) = 0$. Since $T + F$ invertible implies $i(T + F) = 0$, the desired converse would be trivially true if we could prove that the Fredholm index is invariant under perturbation by a finite rank operator. In view of Exercises 4.19 and 4.20 below this would also prove the converses of parts (b) and (c).

The desired invariance of the Fredholm index is indeed true. Its proof, which takes some work, will be be taken up later in this section.

Here are some exercises, the first two of which generalize the fact that invertible transformations are Fredholm with index zero. The third is a result that will be needed when we prove the invariance of Fredholm index under finite-rank perturbation.

Exercise 4.19 (Left invertibility)**.** Suppose $T \in \Phi(X, Y)$ is left invertible, i.e, $ST = I_X$ for some $S \in \mathscr{L}(Y, X)$. Show that $i(T) \leq 0$, that $S \in \Phi(Y, X)$, and that $i(S) = -i(T)$.

Exercise 4.20 (Right invertibility)**.** State and prove the corresponding results for right invertibility.

Corollary 4.21. *For each $T \in \Phi(X, Y)$ there exists $S \in \Phi(Y, X)$ such that $i(S) = -i(T)$*

Every linear transformation between finite dimensional vector spaces is Fredholm. What is the index of such a map?

Theorem 4.22 (Index in finite dimensions)**.** *If X and Y are finite dimensional and $T \in \mathscr{L}(X, Y)$, then $i(T) = \dim X - \dim Y$.*

Proof. By Theorem 1.1, the "Fundamental Theorem of Linear Algebra," for any linear transformation T from X to Y, with both spaces finite dimensional:

$$\dim X = \dim \ker T + \dim \operatorname{ran} T = \dim \ker T + \dim Y - \operatorname{codim} \operatorname{ran} T ,$$

where the last equality requires the finite dimensionality of Y. The finite dimensionality of X now allows us to write this last equation as

$$\dim X = i(T) + \dim Y ,$$

and then to obtain the desired result upon solving for $i(T)$. □

Note that this proof shows that Theorem 4.22 is simply a restatement of the finite dimensional version of the Fundamental Theorem of Linear Algebra.

Corollary 4.23. *If* $\dim X < \infty$ *then every* $T \in \mathscr{L}(X)$ *is Fredholm of index zero.*

Exercise 4.24 (Nilpotent Fredholm transformations)**.** Show that a vector space X supports a *nilpotent* Fredholm transformation if and only if $\dim X < \infty$.

5. The Multiplication Theorem

Here we establish, via two different proofs, the multiplicative property of the index (the index of a product is the sum of the indices), and use it to establish the invariance of index under finite rank perturbation.

According to Theorem 4.22, the index of a linear transformation between finite dimensional spaces does not depend on the transformation. In this sense the notion of index seems trivial in the finite dimensional setting. Not so: precisely this case lies at the core of our next result, the crucial *Multiplication Theorem* for the Fredholm index.

Theorem 5.1 (The Multiplication Theorem)**.** *If* $T \in \Phi(X,Y)$ *and* $S \in \Phi(Y,Z)$ *then* $ST \in \Phi(X,Y)$ *and* $i(ST) = i(S) + i(T)$*.*

Before moving on to the proof, I invite you to work out a few exercises. The first illustrates the theorem in action for shift operators, the second asks you to prove the finite dimensional case, and the third provides an amusing application.

Exercise 5.2 (Index of a shift)**.** Suppose S and B are the forward and backward shifts of of §1.3. Show that $i(S^n) = -n$, $i(B^n) = n$, and $i(S^nB^n) = i(B^nS^n) = 0$.

Exercise 5.3 (Multiplication theorem in finite dimensions)**.** Prove Theorem 5.1 for finite dimensional spaces X, Y, and Z.

Exercise 5.4 (Square roots ... or not)**.** Let S and B be, respectively, the forward and backward shifts on the vector space ω of all scalar sequences, as discussed in Examples 1.3 and 4.2. Show that neither S nor B has a square root; e.g. there exists no $T \in \mathscr{L}(\omega)$ with $T^2 = S$. What about higher roots?

Regarding the statement of the Multiplication Theorem, we have already proved that products of Fredholm transformations are Fredholm (Theorem 4.5), so what's at stake here is the formula for the index of a product, and this is important enough to deserve more than one proof. Here I'll present two. The first one, due to Donald Sarason [5], illustrates dramatically that everything

Fredholm is, at its core, finite dimensional. The second proof, which is a less dramatic brute force calculation, illustrates the same principle less elegantly, but with more precision.

Sarason's proof of the Multiplication theorem

The idea of this proof is to reduce the general case to the finite dimensional one (Exercise 5.3 above) by filling in the bottom two rows of the diagram below, where the subscripted transformations are restrictions of the unsubscripted ones.

$$
\begin{array}{ccccccl}
X & \xrightarrow{T} & Y & \xrightarrow{S} & Z & & \text{Fredholm} \\
\| & & \| & & \| & & \\
X_0 & \xrightarrow{T_0} & Y_0 & \xrightarrow{S_0} & Z_0 & & \text{Finite dim'l} \\
\bigoplus & & \bigoplus & & \bigoplus & & \\
X_1 & \xrightarrow{T_1} & Y_1 & \xrightarrow{S_1} & Z_1 & & \text{Isomorphisms}
\end{array}
$$

The Product Diagram

The desired result will then follow quickly from the finite dimensional version (Exercise 5.3) and the "Direct Sum Theorem" of Exercise 4.17, since all the transformations on the bottom row are (surjective) isomorphisms, and so are Fredholm with index zero.

Our strategy for filling in the bottom two rows of the *Product Diagram* involves common sense and perhaps a bit of luck.

Common sense suggests that to make $S_1 T_1$ injective we should choose $X_0 := \ker ST$ (finite dimensional by §1.5 above), and then take for X_1 any subspace complementary to X_0. This produces the first column of the *Fundamental Diagram,* where T_0 and T_1 are the restrictions of T to X_0 and X_1 respectively.

As for the second column note that, since $X_0 = \ker ST \supset \ker T$, the map $T_1 := T|_{X_1}$ is one to one. Thus it makes sense to take $Y_1 := T(X_1)$, so that T_1 an isomorphism of X_1 onto Y_1. That Y_1 has finite codimension in Y follows from the "transitivity of codimension" (see §1.8 above). More precisely: X_1 has, by its definition, finite codimension in X, so $Y_1 = T(X_1)$ has finite codimension in $T(X)$, and $T(X)$ has, since T is Fredholm, finite codimension in Y.

Now we must hope that S_1, which is going to be the restriction of S to Y_1, is one-to one, i.e., that $Y_1 \cap \ker S = \{0\}$. This is easy to check: Fix $y_1 \in Y_1 \cap \ker S$, so $y_1 = T(x_1)$ for some $x_1 \in X_1$, and also $Sy_1 = 0$. Thus $x_1 \in \ker ST = X_0$, so $x_1 \in X_0 \cap X_1 = \{0\}$, hence $x_1 = 0$, and therefore $y_1 = Sx_1 = 0$.

Thus it's natural to define $Z_1 := S(Y_1)$, so $S_1 := S|_{Y_1}$ is an isomorphism of Y_1 onto Z_1 and, again by the transitivity of codimension, Z_1 has finite codimension in Z.

So far we have produced these components of the *Product Diagram:*

$$X_0 = \ker ST, \quad T_0 = T|_{X_0},$$
$$X_1 = \text{any complement of } X_0, \quad T_1 = T|_{X_1},$$
$$Y_1 = T(X_1), \quad S_1 = S|_{Y_1}, \quad \text{and} \quad Z_1 = S(Y_1).$$

It remains to find Y_0 and Z_0.

Since $Y_1 \cap \ker S = \{0\}$, "complimentary extension" (Proposition 1.7) guarantees that Y_1 has a complement Y_0 (necessarily of finite dimension) that contains $\ker S$.

To finish the argument we need to be able to choose Z_0 complementary to Z_1 in such a way that $Z_0 \supset S(Y_0)$. We can do this, again by "complementary extension," if we can show that $S(Y_0) \cap Z_0 = \{0\}$. Indeed we can: Suppose $z_1 \in S(Y_0) \cap Z_1 = S(Y_0) \cap S(Y_1)$. Then $z_1 = S(y_1)$ for some $y_1 \in Y_1$ and also $z_1 = S(y_0)$ for some $y_0 \in Y_0$ hence $y_1 - y_0 \in \ker S \subset Y_0$. Thus y_1 lies in Y_0 so $\in Y_1 \cap Y_0 = \{0\}$. Thus $y_1 = 0$, so $z_1 = S(y_1) = 0$, as desired.

This completes the *Product Diagram,* and with it the proof of the multiplication theorem. □

The table below summarizes the proof, with the boxes numbered to indicate the flow of the argument. In this table the notation $M \perp N$ for subspaces M and N of a vector space means $M \cap N = \{0\}$, and $M = N^{\perp}$ means that $X = M \oplus N$.[6]

[6] In contrast to the situation in an inner product space, neither of these notations assumes any uniqueness for M.

	X	Y	Z
Finite dim'l	① $X_0 = \ker ST$	⑤ $Y_0 = Y_1^\perp,\ Y_0 \supset \ker S$	⑥ $Z_0 = Z_1^\perp,\ Z_0 \supset S(Y_0)$
Isomorphic	② $X_1 = X_0^\perp$	③ $Y_1 = T(X_1) \perp \ker S$	④ $Z_1 = S(Y_1)$

Multiplication theorem: summary of proof.

Brute force proof of the Multiplication theorem

As above, let X, Y, and Z be vector spaces, with $T : X \to Y$ and $S : Y \to Z$ linear transformations. For this proof we need explicit formulae for $\dim\ker ST$ and $\operatorname{codim}\operatorname{ran} ST$. The first of these we have already found; it is equation (1) in §1.5. The second goes like this[7]:

Proposition 5.5 (Codimension of the range of a product)**.**

(5) $$\operatorname{codim}(\operatorname{ran} ST) = \operatorname{codim}(\operatorname{ran} S) + \operatorname{codim}(\operatorname{ran} T + \ker S)$$

Proof. Since $\operatorname{ran} ST = S(\operatorname{ran} T) \subset \operatorname{ran} S$ the transitivity of codimension (§1.8(b) above) tells us that

$$\operatorname{codim}(\operatorname{ran} ST) = \operatorname{codim}(\operatorname{ran} ST, \operatorname{ran} S) + \operatorname{codim}(\operatorname{ran} S)$$

Thus the desired result is equivalent to:

(6) $$\operatorname{codim}(\operatorname{ran} T + \ker S) = \operatorname{codim}(\operatorname{ran} ST, \operatorname{ran} S)\,.$$

To prove this, let N be a subspace of Y complementary to $\operatorname{ran} T + \ker S$, so that

(7) $$\operatorname{codim}(\operatorname{ran} T + \ker S) = \dim N$$

Now $Y = (\operatorname{ran} T + \ker S) \oplus N$, hence upon applying S to both sides of this equation we obtain

(8) $$\operatorname{ran} S = \operatorname{ran} ST + S(N)$$

Claim: $\operatorname{ran} ST \cap S(N) = \{0\}$.

[7]Recall that, with the usual notion of addition in $\mathbb{Z} \cup \{\infty\}$, equation (1) holds even if some or all of the quantities involved are infinite. The same will be true for the formulae developed below in §5.6–5.7

Proof of Claim. Suppose $z \in \operatorname{ran} ST \cap S(N)$. We wish to show that $z = 0$. By hypothesis, $z = STx = Ny$ for some $x \in X$ and $y \in N$. Thus $Tx - y \in \ker S$, hence

$$y \in (\ker S + Tx) \cap N \subset (\ker S + \operatorname{ran} T) \cap N = \{0\},$$

i.e., $y = 0$. Thus $z = Ny = 0$, as desired. □

The Claim, along with (8), implies

$$\operatorname{codim}(\operatorname{ran} ST, \operatorname{ran} S) = \dim S(N) = \dim(N), \tag{9}$$

the last equality following from the injectivity of S on N which, in turn, arises from the fact that $\ker S$ is complementary to N. Equation (6) then follows from (7) and (9), and—as pointed out above—this establishes the desired result (5). □

Suppose for a moment that the transformations S and T of Proposition 5.5 are Fredholm. The calculation of the product index $i(ST)$, upon employing equation (1) of §1.5, starts out like this:

$$\begin{aligned} i(ST) &:= \dim\ker ST - \operatorname{codim}\operatorname{ran} ST \\ &= \dim\ker T + \dim(\ker S \cap \operatorname{ran} T) - \operatorname{codim}\operatorname{ran} ST\ . \end{aligned}$$

Upon substituting (5) in this equation we obtain

$$\begin{aligned} i(ST) &= \dim\ker T + \dim(\ker S \cap \operatorname{ran} T) \\ &\qquad - \operatorname{codim}(\operatorname{ran} S) - \operatorname{codim}(\operatorname{ran} T + \ker S) \end{aligned} \tag{10}$$

Thus further progress requires information about the codimension of a sum of subspaces. We already have a beginning in Corollary 1.15, which asserts that if M and N are subspaces of X whose sum is X then the codimension of M in X equals the codimension of $M \cap N$ in N. In our application we won't assume that the subspaces sum to X, but will replace X by that sum, thus obtaining:

Lemma 5.6 (Codimension of a sum I)**.** *If M and N are subspaces of a vector space then*

$$\operatorname{codim}(M, M + N) = \operatorname{codim}(M \cap N, N). \tag{11}$$

It's now just a short step to obtain the result we really need:

Theorem 5.7 (Codimension of a sum II)**.** *If M and N are subspaces of a vector space, then*

$$\operatorname{codim}(M + N) + \dim N = \operatorname{codim} M + \dim(M \cap N) \tag{12}$$

Proof. We have

$$\begin{aligned}\operatorname{codim} M &= \operatorname{codim}(M, M+N) + \operatorname{codim}(M+N)\\ &= \operatorname{codim}(M \cap N, N) + \operatorname{codim}(M+N) \quad \text{[by (11)]},\end{aligned}$$

and the desired result follows upon adding $\dim(M \cap N)$ to both sides. □

Completion of "brute force" proof. Now we're assuming that S and T are Fredholm transformations, so all relevant dimensions and codimensions are finite, hence subtraction is allowed. From equation (12) with $M = \operatorname{ran} T$ and $N = \ker S$ we obtain

$$\operatorname{codim}(\operatorname{ran} T + \ker S) = \operatorname{codim} \operatorname{ran} T + \dim(\operatorname{ran} T \cap \ker S) - \dim \ker S$$

Substitute this into the right-hand side of (10) to obtain:

$$\begin{aligned} i(ST) &= \dim \ker T + \dim(\ker S \cap \operatorname{ran} T) - \operatorname{codim} \operatorname{ran} S\\ &\qquad - \operatorname{codim} \operatorname{ran} T - \dim(\operatorname{ran} T \cap \ker S) + \dim \ker S\\ &= i(T) + i(S),\end{aligned}$$

the final equality reflecting the cancellation of both terms involving intersection of $\operatorname{ran} T$ and $\ker S$. □

Invariance II

We showed in Theorem 4.9 that the notion of "Fredholmness" is invariant under finite-rank perturbation. Now, thanks to the multiplication theorem, we can now prove that the *Fredholm index* is invariant under finite-rank perturbation. More precisely:

Theorem 5.8 (Finite-rank perturbation theorem II)**.** *If $T \in \Phi(X, Y)$ and $F \in \mathscr{F}(X, Y)$ then $i(T + F) = i(T)$.*

One might hope that the proof of Theorem 4.9 might provide the precision necessary to prove Theorem 5.8. Alas, more work is required. The desired result turns out to be a consequence of the following special case, which is itself both a generalization of Corollary 4.14, and the converse of Theorem 4.18(a).

Lemma 5.9 (Finite rank perturbation of an invertible)**.** *Suppose V is an isomorphism of X onto Y and $F \in \mathscr{F}(X, Y)$. Then $i(V + F) = 0$.*

Proof. Given V and F as in the statement of the Lemma, we have

$$i(V+F) = i(V(I_X + V^{-1}F)) = i(V(I_X + F_1)),$$

where $F_1 = V^{-1}F$ is a finite rank operator $X \to X$ which, by Corollary 4.14, is Fredholm of index zero. This, along with the Multiplication Theorem and the fact that invertibles are Fredholm of index zero, yields

$$i(V+F) = i(V) + i(I_X + F_1) = 0 + 0 = 0,$$

as desired. □

Proof of Theorem 5.8. Suppose $T \in \Phi(X,Y)$ and $F \in \mathscr{F}(X,Y)$. By Exercise 4.21 there exists $S \in \Phi(Y,X)$ with $i(S) = -i(T)$. By the Multiplication Theorem, $i(ST) = 0$, so by Theorem 4.18(a) there is a finite rank transformation $F_1 : X \to X$ and an invertible transformation $V : X \to X$ such that $ST = V + F_1$. Thus

$$S(T+F) = ST + SF = V + F_1 + SF = ST + F_2$$

with $F_2 \in \mathscr{F}(X)$.

By Lemma 5.9, $i(ST + F_2) = 0$, so this, along with the Multiplication Theorem and our choice of S yields:

$$0 = i(S(T+F)) = i(S)i(T+F) = -i(T)i(T+F)$$

hence $i(T+F) = i(T)$, as desired. □

6. Spectra

In this section we'll apply Fredholm theory to study spectra of linear transformations. The *spectrum* of $T \in \mathscr{L}(X)$ is the set $\sigma(T)$ of scalars λ for which the transformation $T - \lambda I$ fails to be invertible. Recall that for finite dimensional vector spaces the spectrum is nonempty whenever the underlying field is algebraically complete (e.g., the field of complex numbers), but otherwise may be empty; nontrivial rotations of $\mathbb{R}^2$ have no eigenvalues, hence empty spectrum. The next exercise shows that the spectrum may be empty even if the underlying field *is* algebraically complete.

Exercise 6.1 (Empty spectrum)**.** Let $\omega_0(\mathbb{Z})$ denote the space of two-sided scalar sequences with all but a finite number of entries equal to zero. Alternatively, $\omega_0(\mathbb{Z})$ is the set of functions that take the integers into the scalars and have finite support. Let B be the *backward shift* on $\omega_0(\mathbb{Z})$: the transformation that shifts each sequence one unit to the left. More precisely,

$$Bx(n) = x(n+1) \qquad (x \in \omega_0(\mathbb{Z}), n \in \mathbb{Z}).$$

Show that the spectrum of B is empty.

Spectral points: First classification

The spectrum of T can be viewed as the union of two (not necessarily disjoint) subsets:

- The *point spectrum,* or eigenvalues $\sigma_p(T)$ consisting of those scalars λ for which $T - \lambda I$ fails to be injective, and
- The *compression spectrum* consisting of those scalars λ for which $T-\lambda I$ fails to be surjective.[8]

Suppose, for example, that $\dim X < \infty$. Then every linear transformation on X is Fredholm of index zero (by Theorem 4.22, for example), hence—as proven in any beginning course in linear algebra—a linear transformation T on X is injective if and only if it is surjective. Thus $\sigma(T) = \sigma_p(T) = \sigma_c(T)$.

Exercises 6.2 (Spectra of adjoints)**.** For $T \in \mathscr{L}(X)$:

(a) Show that $\sigma(T) = \sigma(T')$.
(b) Examine the relationship between the point and compression spectra of T and T'.

Exercises 6.3 (Spectra of shifts)**.** Determine the spectrum, point spectrum, and compression spectrum for:

(a) The forward shift on the vector space ω of scalar sequences.
(b) The backward shift on ω
(c) The forward shift on the vector space ω_0 consisting of scalar sequences with at most infinitely many nonzero entries.
(d) The backward shift on ω_0.

Exercise 6.4 (The spectral mapping theorem.)**.** Suppose X is a vector space over $\mathbb{C}$ and $T \in \mathscr{L}(X)$. Let f be a polynomial with complex coefficients. Show that $\sigma(f(T)) = f(\sigma(T))$. Does your proof work for fields more general than $\mathbb{C}$?

[8]Remember: this is the *algebraic* compression spectrum. In the setting of *complex Banach spaces,* the compression spectrum is usually taken to be the set of complex numbers λ for which $T - \lambda I$ is not bounded below.

Fredholm decomposition of the spectrum

Our theory of Fredholm transformations suggests a second classification of spectral points—this time into two *disjoint* subsets. For $T \in \mathscr{L}(X)$ we'll say that $\lambda \in \sigma(X)$ is in:

(a) The *Fredholm spectrum* $\sigma_f(T)$ if $T - \lambda I$ is a (non-invertible) Fredholm transformation, and
(b) The *essential spectrum* $\sigma_e(T)$ if $T - \lambda I$ is not Fredholm.

Thus $\sigma(T)$ is the disjoint union of $\sigma_e(T)$ and $\sigma_f(T)$. We'll further classify each point $\lambda \in \sigma_f(T)$ as having *index* equal to the index of $T - \lambda I$.

In case X is finite dimensional, note that $\sigma(T) = \sigma_f(T)$, and $\sigma_e(T)$ is empty.

Exercise 6.5 (Empty essential spectrum)**.** Give an example of a linear transformation whose essential spectrum is not empty.

Exercises 6.6 (Fredholm decomposition of spectrum)**.** Determine the Fredholm and essential spectra of the operators of Exercises 6.3.

Recall the notation $\mathscr{F}(X)$ for the collection of finite rank linear transformations on the vector space X, i.e., those $T \in \mathscr{L}(X)$ for which $\dim \operatorname{ran} T < \infty$. Recall from elementary linear algebra that if $\lambda_1, \lambda_2, \ldots, \lambda_n$ is a set of distinct eigenvalues for $T \in \mathscr{L}(X)$, and x_i is an eigenvector for λ_i, then the set of vectors $\{x_1, x_2, \ldots x_n\}$ is linearly independent.

Theorem 6.7 (Finite-rank spectra)**.** *If $F \in \mathscr{F}(X)$ then $\sigma(F)$ has finitely many points; each nonzero one of which is a Fredholm point of index zero.*

Proof. Any finite rank perturbation of the identity transformation is Fredholm of index zero (Theorem 5.8, or more specifically, Lemma 5.9, or even more specifically, Corollary 4.14), so for $F \in \mathscr{F}(X)$ and λ a non-zero scalar, $F - \lambda I$ is injective if and only if surjective, hence non-invertible if and only if an eigenvalue. Thus $\sigma(T)\backslash\{0\}$ consists entirely of eigenvalues, which are Fredholm points of index zero. By the discussion preceding the statement of the Theorem, there can be at most $\dim \operatorname{ran} T < \infty$ of such eigenvalues. □

Exercise 6.8 (Number of eigenvalues)**.** For an eigenvalue λ of a linear transformation T let's call $\dim \ker(T - \lambda I)$ the (geometric) *multiplicity* of λ. Suppose $T \in \mathscr{F}(X)$. Show that the number of eigenvalues of T, *with each counted as many times as its multiplicity,* is $\leq \dim \operatorname{ran} T$.

7. Multiplication operators

In this final section we'll connect Fredholm theory with classical function theory by exploring the Fredholm properties of an important class of naturally occurring examples: multiplication operators on spaces of holomorphic functions. Our study will culminate in a connection between Fredholm index and the notion of winding number.

Definition 7.1 (Holomorphic functions)**.** By a *plane domain* we'll mean a subset of the complex plane that is nonempty, open, and connected. Henceforth the generic plane domain will be denoted by the symbol "G". A complex valued function is said to be *holomorphic* (or *analytic*) on G if it is complex-differentiable at each point of G. The collection of holomorphic functions on G, henceforth denoted by $\mathscr{H}(G)$, is easily seen to be a complex vector space under pointwise operations.

Definition 7.2 (Multiplication operators)**.** The complex vector space $\mathscr{H}(G)$ is closed under pointwise multiplication, and so is an *algebra* over the field of complex numbers. As such it has many interesting properties, but we will be concerned here only with one: Each $\varphi \in \mathscr{H}(G)$ induces on $\mathscr{H}(G)$ a *multiplication operator* M_φ, defined as follows:

$$M_\varphi f(z) = \varphi(z)f(z) \qquad (f \in \mathscr{H}(G), z \in G).$$

Clearly M_φ is a linear transformation taking $\mathscr{H}(G)$ into itself, the various properties of which are somehow coded into the behavior of the holomorphic function φ. It is the goal of this section to unscramble this code.

Invertibility of multiplication operators

Fredholm properties are intimately involved with invertibility, so it makes sense to try to classify those $\varphi \in \mathscr{H}(G)$ for which M_φ is invertible on $\mathscr{H}(G)$.

First of all, note that (unless $\varphi \equiv 0$) M_φ *is injective*. This is a reflection of one of the fundamental properties of holomorphic functions, the *Identity Theorem* [6, VII.14, page 89]:

> *If two functions holomorphic on G agree a sequence that has a limit point in G, then those functions agree on all of G.*

In particular, if two functions holomorphic on G agree on a nonvoid open subset of G then they agree everywhere. Suppose now, for $\varphi \in \mathscr{H}(G)\backslash\{0\}$, we have $M_\varphi f = M_\varphi g$ for two functions $f, g \in \mathscr{H}(G)$. Then $f = g$ on the set $\{z \in G : \varphi(z) \neq 0\}$, which is nonempty (since we're assuming φ is not the zero-function) and open (since holomorphic functions are continuous). Thus, by the Uniqueness Theorem, $f = g$ at every point of G, so M_φ is injective.

The question of invertibility, therefore, boils down to that of surjectiveness: when is $\operatorname{ran} M_\varphi = \mathscr{H}(G)$?

Suppose, for example, that $0 \notin \varphi(G)$, i.e., that φ never takes the value 0 in G. Then $1/\varphi$ is holomorphic in G (another important property of holomorphic functions), so clearly M_φ is invertible with inverse $M_{1/\varphi}$.

Conversely suppose $\operatorname{ran} M_\varphi = \mathscr{H}(G)$. Then, since the constant function 1 is in $\mathscr{H}(G)$, there exists $\psi \in \mathscr{H}(G)$ such that $\varphi\psi = M_\varphi\psi = 1$. Thus $0 \notin \varphi(G)$ (so, by the way, $1/\varphi \in \mathscr{H}(G)$ and $M_\varphi^{-1} = M_{1/\varphi}$).

With this we have proved:

Theorem 7.3 (Multiplier invertibility). *Suppose $\varphi \in \mathscr{H}(G)\backslash\{0\}$. Then M_φ is injective; it is invertible on $\mathscr{H}(G)$ if and only if $0 \notin \varphi(G)$, in which case $M_\varphi^{-1} = M_{1/\varphi}$*

Note that for $\lambda \in \mathbb{C}$ we have $M_\varphi - \lambda I = M_{\varphi-\lambda}$, so $\lambda \in \sigma(M_\varphi)$ if and only if $\varphi - \lambda$ takes the value zero somewhere on G. That is:

Corollary 7.4 (Multiplier spectrum). *For $\varphi \in \mathscr{H}(G)$, the spectrum of M_φ is $\varphi(G)$.*

Fredholm properties of multiplication operators

The Fredholm properties of multiplication operators on $\mathscr{H}(G)$ reflect another fundamental property of holomorphic functions: the special nature of their zeros. One of the cornerstones of the theory of holomorphic functions is the fact that each function $f \in \mathscr{H}(G)$ is infinitely differentiable on G (a property *not* possessed by infinitely real-differentiable functions). Thus for each $a \in G$ the function f has a formal Taylor series $\sum_{n=0}^{\infty} f^{(n)}(z-a)^n$. Remarkably, this series converges to f in a neighborhood of a (another property not possessed by infinitely real-differentiable functions).[9] This power series representation quickly establishes the following fundamental result (see [4, Theorem 10.18, pp. 208–209] or [6, VII.13, pp. 87–88] for example):

[9] In fact, the series converges to f in the largest open disc in G with center at a.

Theorem 7.5 (Zeros of holomorphic functions)**.** *Suppose $\varphi \in \mathscr{H}(G)$, $a \in G$, and $\varphi(a) = 0$. Then there is a non-negative integer ν for which*

$$\varphi(z) = (z-a)^{\nu}\varphi_1(z) \qquad (z \in G),$$

where $g_1 \in \mathscr{H}(G)$ and $g_1(a) \neq 0$.

The integer $\nu = \nu(a) = \nu_{\varphi}(a)$ is called the *order,* or *multiplicity,* of the zero a of φ. Let's extend the function ν_{φ} to the entire complex plane by defining its value to be zero off the set of zeros of φ. Thus $\nu_{\varphi}(a) = 0$ except for at most a countable set of points of G.

Here is the main result of this section; note how its first part generalizes Theorem 7.3 above.

Theorem 7.6 (Fredholm multipliers)**.** *For $\varphi \in \mathscr{H}(G)$:*

(a) *M_{φ} is Fredholm if and only if φ has at most finitely many zeros in G.*
(b) *Each $\lambda \in \sigma(M_{\varphi}) = \varphi(G)$ is a Fredholm point of index -$\nu_{\varphi}(\lambda)$.*

We often think of $\nu_{\varphi}(a)$ as counting the "number of zeros φ has at a," or "the number of times f covers zero by a," and interpret $\nu_{\varphi}(a) = 0$ to mean "f is not zero at a." With this convention part (b) of the above result can be rephrased:

> *If $\varphi \in \mathscr{H}(G)$ has only finitely many zeros in G, then M_{φ} is a Fredholm transformation on $\mathscr{H}(G)$ whose index is minus the number of zeros (counting multiplicity) that φ has in G.*

The proof of Theorem 7.6 will consist of several subsidiary results, each of some interest in its own right. The first of these states that $\operatorname{ran} M_{\varphi}$ consists of all those functions holomorphic on G whose zero-set, counting multiplicities, contains that of φ.

Lemma 7.7 (Multiplier ranges)**.** $\operatorname{ran} M_{\varphi} = \{f \in \mathscr{H}(G) : \nu_f(a) \geq \nu_{\varphi}(a) \quad \forall a \in G\}$.

Proof. Since $\operatorname{ran} M_{\varphi} = \varphi\mathscr{H}(G)$, it's clear that, counting multiplicity, f has at least as many zeros as φ, i.e., that $\nu_f \geq \nu_{\varphi}$ on G. Conversely, if $\nu_f \geq \nu_{\varphi}$ on G, then $h := f/\varphi$ is holomorphic on G, except possibly for the zeros of φ, where h, initially holomorphic on G, except possibly at the zeros of φ, is extendable to be holomorphic on all of G (by Theorem 7.5). Thus $f = \varphi h \in \varphi\mathscr{H}(G)$. $\square$

As a consequence of Theorem 7.5 we can use derivatives to determine the order of a zero of a holomorphic function. To make this more precise, let's employ the notation $f^{(k)}$ for the k-th derivative of the function f, setting $f^{(0)}$ equal to f. Then (see [6, §VII.13] for example):

> *Suppose f is holomorphic in a neighborhood of a point $a \in \mathbb{C}$, and $f(a) = 0$. Then $\nu_f(a)$ is the smallest positive integer k for which $f^{(k)}(a) \neq 0$.*

For example, $f(z) = z^3$ has a zero of order 3 at the origin, a point at which both it and its first two derivatives vanish, but at which its third derivative does not.

As a consequence of this expression of zero-multiplicity in terms of derivatives, we can use linear functionals to characterize $\operatorname{ran} M_\varphi$. For each $a \in G$ and k a non-negative integer, let's define a linear functional $\Lambda_a^{(k)}$ on $\mathscr{H}(G)$ by:

$$\Lambda_a^{(k)} f = f^{(k)}(a)\,.$$

Then Lemma 7.7 can be rephrased succinctly as:

$$\operatorname{ran} M_\varphi = \bigcap \{\ker \Lambda_a^{(k)} : a \in G, k < \nu_\varphi(a)\} \tag{13}$$

where the only terms of the intersection on the right that are not the whole space $\mathscr{H}(G)$ are the ones corresponding to points a in the zero-set of φ.

Thus, in view of the injectivity of M_φ, to prove Theorem 7.6 we need only be able to calculate the codimension of an intersection of kernels of linear functionals, and for this we can use Theorem 2.7—once we have established the linear independence of the functionals involved.

Lemma 7.8 (Independence of derivatives)**.** *The set of linear functionals*

$$\{\Lambda_a^{(k)} : a \in G,\ k = 0, 1, 2, \ldots\}$$

is linearly independent in the dual space of $\mathscr{H}(G)$.

Proof. Fix a finite subset A of G and a positive integer K. It's enough to prove the independence of the finite set

$$\{\Lambda_a^{(k)} : a \in A,\ 0 \leq k \leq K\}$$

of linear functionals on $\mathscr{H}(G)$. To this end, suppose the linear combination

$$\Lambda := \sum \{c_{a,k} \Lambda_a^{(k)} : a \in A,\ 0 \leq k \leq K\} \tag{14}$$

is identically zero on $\mathscr{H}(G)$. We want to show that all the coefficients $c_{a,k}$ are zero.

The key is that for each $\lambda \in \mathbb{C}$, the linear functional L annihilates the entire function $z \to \exp(\lambda z)$, resulting in the equations

$$0 = \sum_{a \in A} p_a(\lambda) e^{\lambda a} \qquad \forall\, \lambda \in \mathbb{C} \tag{15}$$

where

$$p_a(\lambda) := \sum_{k=0}^{K} c_{a,k} \lambda^k \,.$$

If A is the singleton $\{a\}$ then the polynomial p_a is identically zero on $\mathbb{C}$, so all its coefficients are zero, as desired.

Otherwise, let b denote the element of A whose modulus is maximum (if there is more than one such element, pick any one). Let $\sigma = \overline{b}/|b|$, so that $\sigma b = |b|$. Then for $\lambda = \sigma t$ with $t > 0$ we have

$$0 = \sum_{a \in A\setminus\{b\}} p_a(\sigma t) e^{\sigma a t} + p_b(\sigma t) e^{|b|t} \qquad \forall t > 0. \tag{16}$$

On the right-hand side of this equation the sum over $a \in A\setminus\{b\}$ is $o(e^{|b|t})$ as $t \to +\infty$. But this right-hand side has to vanish for all $t > 0$, and since no cancellation can be expected from the sum, both the sum and the last term must vanish for all $t > 0$. Vanishing of the last term on the right means that the polynomial $t \to p_b(\sigma t)$ is identically zero, hence all its coefficients are zero. Vanishing of the sum on the right means that we may repeat the argument, replacing A by $A\setminus\{b\}$ to get another of the polynomials p_a, and hence its coefficients, identically zero. Upon repeating the argument a finite number of times we conclude that all the coefficients $c_{a,k}$ are zero, as desired. □

Completion of the proof of Theorem 7.6. We are assuming that φ is not identically zero, in which case M_φ is injective (Theorem 7.3). Thus M_φ is Fredholm if and only if $\operatorname{ran} M_\varphi$ has finite codimension in $\mathscr{H}(G)$.

If φ has at most finitely many zeros in G, then by Lemma 7.8 and Theorem 2.7 the codimension of $\operatorname{ran} M_\varphi$ in $\mathscr{H}(G)$ is finite, and equal to the number of these zeros, counting multiplicities, i.e., to $\sum_{a \in G} \nu_\varphi(a)$.

If φ has infinitely many zeros in G, then given a positive integer n let A_n be a subset of these zeros having cardinality n (we are not counting multiplicity here). Let Λ_a be the linear functional of evaluation at $a \in A_n$. Then

$$\operatorname{ran} M_\varphi \subset \bigcap_{a \in A_n} \ker \Lambda_a \,.$$

Since the set of functionals $\{\Lambda_a : a \in_n\}$ is linearly independent (Lemma 7.8), the codimension of the intersection on the right-hand side of this display is n. Thus $\operatorname{ran} M_\varphi$ has codimension at least n. Since n is arbitrary, $\operatorname{ran} M_\varphi$ has infinite codimension in $\mathscr{H}(G)$. □

Index and winding number

For $f \in \mathscr{H}(G)$ we have interpreted the order $\nu_f(a)$ of the zero f has at $a \in G$ as the "number of times f covers the origin by the point a." Now it's time to shift the emphasis from the point that does the covering to the point that gets covered. To this end let's define for $\lambda \in \mathbb{C}$

$$\mu_f(\lambda) := \sum_{a \in f^{-1}(\lambda)} \nu_{f-\lambda}(a) \, ,$$

a quantity which should, in accordance with our covering interpretation of $\nu_{f-\lambda}$, be interpreted as "the number of times λ is covered by f, or more succinctly, the "multiplicity of the value λ" (with $\mu_f(\lambda) = 0$ if and only $\lambda \notin f(G)$).

With these conventions we can restate that last part of Theorem 7.6 as follows:

> *$\lambda \in \mathbb{C}$ is a Fredholm point of M_φ if and only if $\mu_\varphi(\lambda) < \infty$, in which case $i(M_\varphi - \lambda I) = -\mu_\varphi(\lambda)$.*

For yet another interpretation another interpretation of the Fredholm index of $M_\varphi - \lambda I$ we can turn to the *Argument Principle* (see [4, Theorem 10.43, pp. 225-6] or [6, X.11, pp. 137-8] for example). For simplicity let's restrict to a special case: Let $\mathbb{U}$ denote the open unit disc of the complex plane, and suppose φ is holomorphic on a neighborhood of the closure of $\mathbb{U}$. Then φ can have only finitely many zeros in $\mathbb{U}$, so M_φ, acting on $\mathscr{H}(\mathbb{U})$, has spectrum $\varphi(\mathbb{U})$, every point of which is a Fredholm point (i.e., $M_\varphi - \lambda I$ is a Fredholm transformation of $\mathscr{H}(\mathbb{U})$). For this situation the Argument Principle states that

> *For each $\lambda \in \mathbb{C} \backslash \varphi(\partial \mathbb{U})$, the multiplicity $M_\varphi(\lambda)$ is the number of times the closed curve $\varphi(\partial \mathbb{U})$ winds counter-clockwise around the point λ.*

For example, if $\varphi(z) = z^2$ then φ winds the unit circle twice around itself, which, according to the Argument Principle, reflects the fact that φ covers each point of the unit disc twice (with the origin being covered twice thanks to our interpretation of "multiplicity two").

When applied to the Fredholm theory of the operator M_φ, the argument principle allows this reinterpretation of Theorem 7.6(b):

Theorem 7.9 (Fredholm points of multipliers)**.** *Suppose φ is holomorphic in a neighborhood of the closed unit disc and $\lambda \in \mathbb{C}\backslash\varphi(\partial\mathbb{U})$. Then the transformation $M_\varphi - \lambda I$ has Fredholm index equal to "minus the number of times $\varphi(\partial\mathbb{U})$ wraps counter-clockwise around the point λ."*

For a nontrivial example consider the function $\varphi(z) = z^2 + z$. The mapping properties of φ are best understood by writing

$$\varphi(z) = \left(z + \frac{1}{2}\right)^2 - \frac{1}{4} \ ,$$

whereupon $\varphi(\mathbb{U}) = \sigma(M_\varphi)$ is revealed to be the region inside the outer cardioid-shaped curve in Figure 1.

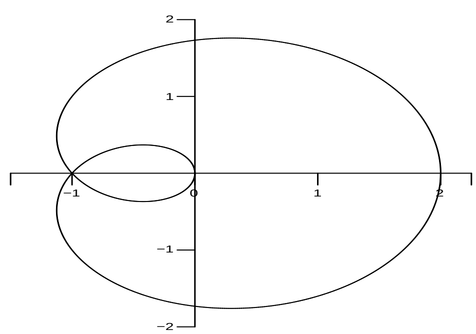

FIGURE 1. $\varphi(\partial\mathbb{U})$ for $\varphi(z) = z^2 + z$

In this picture the boundary of $\varphi(\mathbb{U})$ is the image of the larger arc of $\partial\mathbb{U}$ that lies between the points $e^{\pm i2\pi/3}$, both of which φ maps to -1. The boundary of the small interior loop is the image of the smaller arc of $\partial\mathbb{U}$ between those same two points. The domain $\varphi(\mathbb{U})$ that lies outside this smaller loop is singly covered by φ, while the points inside the inner loop are doubly covered. Thus $i(M_\varphi - \lambda I)$ is equal to: *zero* for λ outside $\varphi(\mathbb{U})$ and on its outer boundary (the cardioid), *one* between the outer boundary and the small loop, including that loop, and *two* inside the small loop.

References

[1] Sheldon Axler, *Linear Algebra Done Right,* second ed., Springer 2002.
[2] Nelson Dunford and J. T. Schwartz, *Linear Operators, Part I,* Wiley Interscience 1963.
[3] John L. Kelley, Isaac Namioka, et al. *Linear Topological Spaces,* van Nostrand 1963.
[4] Walter Rudin, *Real and Complex Analysis,* third ed., McGraw-Hill 1987.
[5] Donald Sarason, *The multiplication theorem for Fredholm operators,* American Math. Monthly 94 (1987) 68-70.
[6] Donald Sarason, *Complex Function Theory,* American Mathematical Society 2007.

Appendix A. Hints for selected exercises

Exercise 1.9: Fundamental Theorem of Linear Algebra.

Exercise1.16: Fix a basis B_0 for $M \cap N$ and extend B_0 in two "directions": first by a set B_1 of vectors (necessarily from $N \backslash M$) to form a basis for N, and then by adding a set B_2 of vectors (necessarily from $M \backslash N$) to form a basis for M.

Exercise 2.8: Continuing the *Suggestion:* It's clear that $\cap_n \ker \Lambda_n = \{0\}$, so *any* linear functional on ω will have kernel containing that intersection. So the trick is to find one that isn't in the span of the Λ_n's.

Exercise 4.3: Dimension.

Exercise 4.6: Note that $\ker T^k \supset \ker T$ and $\operatorname{ran} T^k \subset \operatorname{ran} T$.

Exercise 4.7: Look at a $T \in \mathscr{L}(\omega)$ defined by $Tx = (x(1), 0, x(2), 0, \cdots)$ for $x = (\xi_1, \xi_2, \cdots) \in \omega$. Alternatively, take T to be an appropriate "vector-valued" shift.

Exercise 5.3: Use Theorem 4.22

Exercise 5.4: You may wish to review Exercise 4.6.

Exercise 4.19: Being left invertible, T is injective. Show that, additionally, $ST = I_X$ implies also that $Y = \ker S \oplus \operatorname{ran} T$.

Exercise 4.20: Being right invertible, T is surjective. Show that, additionally, $X = \ker T \oplus \operatorname{ran} S$.

Exercise 4.15: The various "inverses" (two-sided, left, right) produced in Theorem 4.18 for appropriate finite rank perturbations of T are Fredholm with index $-i(T)$.

Exercise 6.1: Clearly B is invertible, so $0 \notin \sigma(B)$. For $\lambda \neq 0$ show that the equation

$$S_\lambda = \sum_{n=0}^{\infty} \lambda^{-(n+1)} B^n$$

defines—notwithstanding absence of topology—an inverse on $\omega_0(\mathbb{Z})$, for $B - \lambda I$.

Exercises 6.2: Cf. Exercises 2.16.

Exercises 6.3: (a) For $S : \omega \to \omega$, $\sigma_p(S)$ is empty, $\sigma(S) = \sigma_c(S) = \{0\}$.

(b) For $B : \omega \to \omega$: $\sigma(B) = \sigma_p(B)$ is the entire scalar field, while $\sigma_c(B)$ is empty.

(c) Same answers as for part (a).

(d) For $B : \omega_0 \to \omega_0$: As in part (b), $\sigma_c(B)$ is empty, and $\sigma(B) = \sigma_p(B)$, but unlike part (b), $\sigma(B) = \sigma_p(B) = \{0\}$.

Exercise 6.6: Build on the results of Exercise 6.3.

Exercise 6.8: Suppose $\{\lambda_1, \lambda_2, \ldots \lambda_n\}$ are the distinct eigenvalues of our finite rank transformation T. Let $E_j = \ker(T - \lambda I)$, a subspace of dimension $d_j < \infty$. Thus $d_1 + d_2 + \cdots + d_n$ is the number of eigenvalues of T, "counting multiplicities."

Let $\mathscr{E}_j$ be a basis for E_j, and $\mathscr{E}$ the union of all these bases. Show that $\mathscr{E}$ is a linearly independent set, and so a basis for the subspace $E_1 + E_2 + \cdots E_n$. Consequence: number of eigenvalues of T, counting multiplicities equals the dimension of E, which is $\leq \dim \operatorname{ran} T$ (since $E \subset \operatorname{ran} T$).

Appendix B. Solutions to selected exercises

Exercise 1.16: Continuing the *Hint:* Let $V_j = \operatorname{span} B_j$ for $j = 1, 2$. Now $V_1 \cap M = \{0\}$, else some nontrivial linear combination from B_1 would belong to M, and so would also be a nontrivial linear combination of vectors from B_0, thus contradicting the linear independence of $B_0 \cup B_1$. Similarly $V_2 \cap N = \{0\}$.

Thus we have $N = (M \cap N) \oplus V_1$, so

$$(*) \qquad \operatorname{codim}(M \cap N, N) = \dim V_1 \,,$$

and since $N \cap V_2 = \{0\}$ we also have

$$(**) \quad N + M = N \oplus V_2 = (M \cap N) \oplus V_1 \oplus V_2 \,.$$

But by definition, $M = (M \cap N) \oplus V_2$, so (**) implies $\operatorname{codim}(M, M+N) = \dim V_1$, which, along with (*) establishes (11).

Exercise 2.8: Continuing the *Hint:* For $n \in \mathbb{N}$ let $e_n \in \omega$ be defined by $e_n(k) = 0$ if $k \neq n$ and $= 1$ if $k = n$. Note that the lists of vectors and functionals (e_n) and (Λ_n) are *biorthogonal* in the sense that $\Lambda_n(e_k) = 0$ if $n \neq k$ and $= 1$ if $n = k$. Furthermore the linear span of the vectors $\{e_n : n \in \mathbb{N}\}$ is the subspace ω_0 consisting of sequences whose entries are zero except for at finitely many. Since ω_0 is a proper subspace of ω there is a nontrivial linear functional Λ on ω that vanishes on ω_0 (see §2.3).

Claim: $\Lambda \notin \operatorname{span}\{\Lambda_n : n \in \mathbb{N}\}$.

Proof of Claim. Suppose for the moment that Λ is *any* linear functional in the span of the Λ_n's. Then $\Lambda = \sum_{n=1}^{N} c_n \Lambda_n$ for some $N \in \mathbb{N}$, whereupon for each $k \in \mathbb{N}$:

$$0 = \Lambda(e_k) = \sum_{n=1}^{N} c_n \Lambda_n(e_k) = c_k .$$

Thus Λ is the zero-functional on ω. But *our* Λ was constructed to *not* be the zero-functional, so it can't lie in the span of the Λ_n's, and therefore provides the desired counterexample. □

Exercise 4.19: We have $ST = I_X$, which implies that T is injective, i.e., $\ker T = \{0\}$. Thus $i(T) \leq 0$. As for S, note that it is surjective, so its Fredholmness, along with the desired result for its index ($i(S) = -i(T)$) will follow from:

Claim: $Y = \ker S \oplus \operatorname{ran} T$.

Proof of Claim. Indeed, $ST = I_X$ implies that S is injective on $\operatorname{ran} T$, hence $\ker S \cap \operatorname{ran} T = \{0\}$. Next, observe that for $y \in Y$ we have $y - TSy \in \ker S$, so

$$y \in T(Sy) + \ker S \subset \operatorname{ran} T + \ker S,$$

as desired. □

Exercise 4.20: $TS = I_Y$ implies that T is surjective, so $\operatorname{codim} \operatorname{ran} T = 0$, hence $i(T) \geq 0$. As for S, note that it is injective, so its Fredholmness, along with the desired result for its index, will follow from:

Claim: $X = \ker T \oplus \operatorname{ran} S$.

Proof of Claim. Note first that if $x \in \ker T \cap \operatorname{ran} S$ then $Tx = 0$ and $x = Sy$ for some $y \in Y$. Thus $0 = Tx = TSy = y$ hence x (which $= Sy) = 0$.

Next, suppose $x \in X$ and observe that $x - STx \in \ker T$, hence $x \in S(Tx) + \ker T \subset \operatorname{ran} S + \ker T$. Thus $X = \ker T + \operatorname{ran} S$, which completes the proof. □

Exercises 6.3: (a) $S : \omega \to \omega$. $\sigma_p(S)$ is empty, since $Sx = \lambda x$ implies that $\lambda \neq 0$ and that $S^n x = \lambda^n x$ for $n \in \mathbb{N}$. Thus, for every $n \in \mathbb{N}$ the first n coordinates of $S^n x$ vanish, hence the same must be true of x and so $x = 0$.

As for compression spectrum, since S is not surjective, $0 \in \sigma_c(S)$. In fact, $\sigma_c(S) = \{0\}$, i.e., $S - \lambda I$ is surjective for all scalars $\lambda \neq 0$. To see this, fix $y = \eta_n) \in \omega$ and solve $y = (S - \lambda I)x$ for $x = (\xi_n) \in \omega$. This vector equation is equivalent to the infinitely many scalar equations $\xi_n = \xi_{n-1} - \eta_n)/\lambda$ for $n = 1, 2, \ldots$, where in case $n = 1$ we set $\xi_0 = 0$. The (unique) solution is

$$\xi_n = -\left(\frac{\eta_1}{\lambda^n} + \frac{\eta_2}{\lambda^{n-1}} + \cdots + \frac{\eta_n}{\lambda}\right),$$

which one checks does give the desired preimage for y.

Summary: $\sigma_p(S)$ is empty, while $\sigma(S) = \sigma_c(S) = \{0\}$

(b) $B : \omega \to \omega$: One checks easily that $\ker B$ consists of all vectors whose coordinates vanish, except for the first one. So $0 \in \sigma_p(B)$. Furthermore, if λ is a nonzero scalar, and $x = (1, \lambda, \lambda^2, \ldots)$, then $Bx = \lambda x$, so $\lambda \in \sigma_p(B)$. Thus $\sigma_p(B) =$ all scalars.

As for compression spectrum—there isn't any! Clearly B itself is surjective. As for $B-\lambda I$ with λ a nonzero scalar: fix $y \in \omega$ and choose $z \in \omega$ (by surjectivity of B) so that $Bz = -y/\lambda$. Then by the surjectivity of $S-\lambda^{-1}I$ on ω (see part of (a) above, with λ replaced by its reciprocal) choose $x \in \omega$ so that $(S - \lambda^{-1}I)x = z$. Apply B to both sides of this equation, noting that $BS = I$; we obtain $(I-\lambda^{-1}B)x = Bz$ which, upon multiplying both sides by $-\lambda$, becomes $(B - \lambda I)x = -\lambda Bz = y$, as desired.

Thus: $\sigma_p(B) = \sigma(B) =$ all scalars, while $\sigma_c(B)$ is empty.

(c) $S : \omega_0 \to \omega_0$. By part (a), $S - \lambda I$ is injective for all scalars λ, hence just as in the ω setting, $\sigma_p(S)$ is empty. Also, as in part (a), S is not surjective on ω_0, so $0 \in \sigma_c(S)$. But unlike the situation of part (a), there's nothing more in the compression spectrum!

Indeed, suppose λ is a nonzero scalar. Let e_1 be the vector in ω_0 with 1 in the first coordinate and zeros elsewhere. Then by part (a) above, the equation $e_1 = (S - \lambda I)x$ has, in ω, the *unique* solution

$$x = -(\frac{1}{\lambda}, \frac{1}{\lambda^2}, \ldots)$$

and this vector does not belong to ω_0. Thus $S - \lambda I$ is surjective on ω_0 for *no* scalar λ.

In summary: $\sigma(S) = \sigma_c(S) = \{0\}$, while $\sigma_p(S)$ is empty.

(d) Check that the operator $S_\lambda =$ defined in the hint for Exercise 6.1 gives an inverse on ω_0 for $B - \lambda I$ when $\lambda \neq 0$. Thus $\sigma(B) \subset \{0\}$. Since B has nontrivial kernel on ω_0 we have $0 \in \sigma_p(B)$. Thus on ω_0: $\sigma(B) = \sigma_p(B) = \{0\}$. Since $BS = I$ on ω_0 (S being the "forward shift"), B is surjective, hence $\sigma_c(B)$ is empty.

Exercise 6.6: (a) For S on ω, $\sigma(S) = \{0\}$: From Exercise 6.3(a) $\operatorname{ran} S =$ all scalar sequences with first coordinate zero. Since S is injective, 0 is a Fredholm point of index -1. The essential spectrum is empty.

(b) For B on ω: Fix a scalar λ. By Exercise 6.3(b), λ is an eigenvalue and $T - \lambda I$ is surjective. Now if you solve the equation $Tx = \lambda x$ for $x \in \omega$, you see quickly that the only solutions are constant multiples of the eigenvector produced in Exercise 6.3(b), i.e., $dim \ker(B - \lambda I) = 1$. Thus each scalar is a Fredholm point of index $+1$; the essential spectrum is empty.

(c) For S on ω_0: Same as part (a).

(d) For B on ω_0: From Exercise 6.3(d), $\sigma(B) = \{0\}$. We already know that 0 is a Fredholm point of index $+1$, thus the essential spectrum of B on ω_0 is empty.

Exercise 6.8: Continuing from Hint A, We have a basis $\mathscr{E}_j = \{e_{j,k}\}_{k=1}^{d_j}$ for $\ker(T - \lambda_j I)$, and desire to show that the union $\mathscr{E}$ of all these bases is linearly independent. Suppose some linear combination $\sum a_{j,k} e_{j,k}$ is zero. Split this sum up into sums first on k and then on j, thus obtaining vectors $x_j \in E_j$ ($j = 1, 2, \ldots n$) that sum to zero. Any x_j that isn't zero is an eigenvector for λ_j, and we know from basic linear algebra that any collection of eigenvectors, each corresponding to a different eigenvalue, is linear independent. Thus each of the x_j's must be zero. Since each $\mathscr{E}_j$ is linearly independent, this guarantees that, for each j, each of the coefficients $a_{j,k}$ is zero ($k = 1, 2, \ldots d_j$). This completes the proof.

Fariborz Maseeh Department of Mathematics and Statistics, Portland State University, Portland OR 97207

E-mail address: `shapiroj@pdx.edu`