Chapter 2
Theoretical Foundations

In the following two chapters, we will first investigate a new "absolute value" norm, which we will then use to develop a general two-level multigrid convergence theory applicable to nonsymmetric problems. Both of the subsequent chapters will be very general, assuming only that we are interested in solving problems of the form

$$Au = f \tag{2.1}$$

with $A \in \mathscr{B}(\mathsf{V}, \mathsf{V}^*)$ being a bounded linear operator from a complex reflexive Banach space to the adjoint space. Additionally, the real part $H = \frac{1}{2}(A + A^*)$ of A will be assumed to be positive, which suffices, by the Lax-Milgram lemma, Theorem 2.1, to ensure the existence of $A^{-1} \in \mathscr{B}(\mathsf{V}^*, \mathsf{V})$, so that there is a unique solution u to (2.1) depending continuously on f.

Taking $\mathsf{V} = \mathbb{C}^n$ reduces this framework to the familiar matrix one, with (2.1) reducing to a system of linear equations governed by the matrix A. This and following chapters may be read with the assumption $\mathsf{V} = \mathbb{C}^n$ in mind, in which case one may ignore the functional analysis concepts that come up. As with (2.1), the equations should be readily interpretable in the matrix setting, as the notation matches standard matrix notation.

On the other hand, the above framework is general enough to encompass elliptic PDEs such as those encountered in the previous chapter. Hence, all of the concepts we consider will be applicable not just to discretized systems, but also to the underlying PDE as well. It is a feature of the finite element discretization that concepts will tend to have a compatible meaning whether we instantiate our framework to apply to the PDE or to its discretization. For example, $\|u\|_A^2$ may be, for a symmetric second-order elliptic problem, the energy of an arbitrary H^1 function in the one case, or the energy of some member of the finite-dimensional subspace of discretization functions in the other. We may thus draw insight as to the meaning of concepts in the discrete case from the infinite-dimensional one. This applies, for example, to the new norm we investigate in the next chapter.

The spaces V we have in mind (H^1 and $\mathbb{C}^n$) are in fact Hilbert spaces. Indeed, as we shall see, our premises on the operator $A \in \mathscr{B}(\mathsf{V}, \mathsf{V}^*)$ in (2.1), that it have positive

J. Lottes, *Towards Robust Algebraic Multigrid Methods for Nonsymmetric Problems*, Springer Theses, DOI 10.1007/978-3-319-56306-0_2

real part, ensures that V can always be equipped with a Hilbert space structure, with a norm equivalent to the Banach norm. In other words, our assumptions imply that V is topologically isomorphic to a Hilbert space. The reason we do not make the Hilbert space structure an assumption is that we regard the particular inner products (the Sobolev and Euclidean) and associated norms of these spaces as not being particularly fundamental. By assuming only a Banach space structure, we essentially make these inner products and their associated geometry unavailable to our developments. We will instead investigate a number of alternative Hilbert space structures, such as that given by the energy inner product, that are naturally associated with the given problem. The only aspect of the given spaces that we consider important is the topology, and this is captured by the Banach metric.

The reason for assuming a complex Banach space is that we will need to make use of the complex numbers almost immediately (so that the quadratic form is sensitive to the nonsymmetry of the operator), and it seems simplest to just treat the complex case in full at the outset. Instantiating results to the real case is usually trivial (e.g., ignoring complex conjugation), whereas generalizing in the other direction is less straight-forward.

In this chapter we will review the basic concepts of our abstract framework. For a functional analyst, all of the material should be review or at least very straight-forward, albeit, rendered in a matrix-like notation: favoring operators over forms, and denoting Hilbert spaces by the operator associated with the inner product. For a numerical analyst used to the matrix setting, the emphasis will be in focusing on Hilbert spaces other than the Euclidean one, and the notation we use should be fairly standard in this setting.

2.1 Sesquilinear Forms

The following definitions are standard and follow Kato [2]. Let V be a reflexive complex Banach space. The adjoint space V^* is defined to consist of bounded semilinear (alternatively conjugate- or antilinear) forms on V, as opposed to linear ones. We will reserve the notation $(\cdot, \cdot)$ for the duality pairing of V^* and V, so that, for any bounded semilinear form $f \in \mathsf{V}^*$ and vector $u \in \mathsf{V}$,

$$(f, v) := f[v] . \tag{2.2}$$

Note that the pairing $(\cdot, \cdot)$ is linear in the first argument and semilinear in the second. As above, we will consistently use the letters f and g to denote elements of an adjoint space, and the letters u, v, and w for elements of the main space. In the matrix setting, $\mathsf{V} = \mathbb{C}^n$, we identify V^* with $\mathbb{C}^n$ by letting $(\cdot, \cdot)$ be the usual Euclidean inner product.

Let $\mathfrak{a}$ be a bounded sesquilinear form on V. A sesquilinear form is defined to be linear in its first argument and semilinear in its second, matching the convention for the duality pairing. "Sesqui-" means "one and a half" and serves as a mnemonic for this convention. Taken together, these conventions allow the form $\mathfrak{a}$ to equivalently be considered as a bounded linear operator $A \in \mathscr{B}(\mathsf{V}, \mathsf{V}^*)$, defined by

$$(Au, v) = (Au)[v] = \mathfrak{a}[u, v] \ , \tag{2.3}$$

where $\mathscr{B}(\mathsf{V}, \mathsf{W})$ is the space of bounded linear operators from V to W. In the matrix setting $\mathscr{B}(\mathsf{V}, \mathsf{V}^*)$ is simply $\mathbb{C}^{n\times n}$, the space of all $n \times n$ matrices.

That boundedness of the form $\mathfrak{a}$ implies boundedness of the operator A is straightforward. If $\mathfrak{a}[u, v]$ has bound C, i.e., $|\mathfrak{a}[u, v]| \le C\|u\|_{\mathsf{V}}\|v\|_{\mathsf{V}}$, then so does A, since

$$\|Au\|_{\mathsf{V}^*} = \sup_{v\in\mathsf{V}\setminus\{0\}} \frac{|(Au, v)|}{\|v\|_{\mathsf{V}}} \le \sup_{v\in\mathsf{V}\setminus\{0\}} \frac{C\|u\|_{\mathsf{V}}\|v\|_{\mathsf{V}}}{\|v\|_{\mathsf{V}}} = C\|u\|_{\mathsf{V}} \ . \tag{2.4}$$

In the other direction, if $\|A\| \le C$, then

$$|\mathfrak{a}[u, v]| = |(Au, v)| \le \|Au\|_{\mathsf{V}^*}\|v\|_{\mathsf{V}} \le C\|u\|_{\mathsf{V}}\|v\|_{\mathsf{V}} \ . \tag{2.5}$$

That is, not only is A bounded, but it has the same bounding constant as $\mathfrak{a}$.

Conversely, if we are given an operator $A \in \mathscr{B}(\mathsf{V}^*, \mathsf{V})$, then (2.3) can be taken to define the sesquilinear form $\mathfrak{a}$, which by the above arguments is bounded with the same bounding constant as A. That is, we have established an isometric isomorphism between the space $\mathscr{B}(\mathsf{V}, \mathsf{V}^*)$ and the space of bounded sesquilinear forms on V. In light of this correspondence, we shall sometimes refer to an operator $A \in \mathscr{B}(\mathsf{V}, \mathsf{V}^*)$ as a bounded sesquilinear form, leaving the isomorphic mapping implied.

The sesquilinear form $\mathfrak{a}$ is associated with the quadratic form

$$\mathfrak{a}[u] := \mathfrak{a}[u, u] = (Au, u) \ . \tag{2.6}$$

In fact, the sesquilinear form is determined by the quadratic form via the polarization identity

$$\mathfrak{a}[u, v] = \frac{1}{4}(\mathfrak{a}[u + v] - \mathfrak{a}[u - v] + \mathrm{i}\mathfrak{a}[u + \mathrm{i}v] - \mathrm{i}\mathfrak{a}[u - \mathrm{i}v]) \ . \tag{2.7}$$

This one-to-one correspondence between sesquilinear and quadratic forms relies on the existence of the scalar i. That is, the correspondence would not hold if V were a real instead of complex Banach space. The adjoint form $\mathfrak{a}^*$ is defined by

$$\mathfrak{a}^*[u, v] := \overline{\mathfrak{a}[v, u]} = \overline{(Av, u)} = (A^*u, v) \ . \tag{2.8}$$

Note $A^* \in \mathscr{B}(\mathsf{V}^{**}, \mathsf{V}^*) = \mathscr{B}(\mathsf{V}, \mathsf{V}^*)$ as V was assumed to be reflexive. Thus, A^* is the operator associated with $\mathfrak{a}^*$ in exactly the same way as A is associated with $\mathfrak{a}$. In the matrix case, A^* is the conjugate transpose of A, as the notation suggests.

The form $\mathfrak{a}$ is *symmetric* when $\mathfrak{a}^* = \mathfrak{a}$, which is equivalent to A being equal to its adjoint ($A^* = A$). In this case the quadratic form $\mathfrak{a}[u]$ is real. In light of the correspondence (2.3), where we think of A as a sesquilinear form, we will call A symmetric when it is equal to its adjoint. There is a slight danger of confusion in the matrix setting, as the term "Hermitian" is more common, "symmetric" possibly referring to a matrix being equal to its regular (not conjugate) transpose. For a real matrix there is no distinction, and the condition $A^* = A$ (the "Hermitian" one) is the generalization of symmetry to the complex case more appropriate for our uses. Here, symmetric always means Hermitian.

The Cartesian decomposition of the quadratic form $\mathfrak{a}[u]$ is

$$\mathfrak{a}[u] = \mathfrak{h}[u] + \mathrm{i}\mathfrak{k}[u] , \qquad \mathfrak{h} := \frac{1}{2}(\mathfrak{a} + \mathfrak{a}^*) , \qquad \mathfrak{k} := \frac{1}{2\mathrm{i}}(\mathfrak{a} - \mathfrak{a}^*) . \tag{2.9}$$

Note that $\mathfrak{h}$ and $\mathfrak{k}$ are both symmetric, giving rise to real quadratic forms. The associated symmetric operators are

$$H_A := \frac{1}{2}(A + A^*) , \quad \text{and} \quad K_A := \frac{1}{2\mathrm{i}}(A - A^*) . \tag{2.10}$$

We will refer to these as the *real* and *imaginary* parts of A. Subsequently, we will usually shorten H_A and K_A to H and K. We will occasionally wish to refer to the real and imaginary parts of other operators. For example, note that

$$H_{A^*} = H_A , \qquad K_{A^*} = -K_A . \tag{2.11}$$

For the case of A being a real matrix, H is the symmetric part of A, while $\mathrm{i}K$ is the anti-symmetric part.

Table 2.1 summarizes the beginning of an analogy that arises between sesquilinear forms and the complex numbers, including the concepts introduced so far. In the analogy, the adjoint corresponds to complex conjugation, and "symmetric" corresponds to "real." Note that the second column reduces to the first in the scalar case $\mathsf{V} = \mathbb{C}$.

Table 2.1 Analogy between complex numbers and sesquilinear forms

$\mathbb{C}$	$\mathscr{B}(\mathsf{V}, \mathsf{V}^*)$
$\overline{z}$	A^*
$z \in \mathbb{R}$ if $z = \overline{z}$	$(Au, u) \in \mathbb{R}$ if $A = A^*$
$z = a + \mathrm{i}b$	$A = H + \mathrm{i}K$

2.2 Lax-Milgram Lemma

A sufficient condition for A^{-1} to exist, be continuous, and be defined on all of V^* is for $\mathfrak{a}$ to be coercive. This statement is a variant of the Lax-Milgram lemma, usually stated for a bilinear form on a real Hilbert space. The lemma will play a central role in our development, as its premises serve to define (nearly, see the corollary below) the class of problems we are interested in solving. It is easy to find the complex version of the lemma, for which the proof is essentially the same. Roşca appears to have been the first to consider its generalization to Banach spaces [3]. As it is difficult to find a version of the proof with both generalizations, and as it is relatively short, a proof is included here, for completeness. But note that considering a complex instead of a real space adds essentially no difficulty to the proof. This is the only result we will consider that involves topological notions, where the possibility of infinite dimensions introduces subtleties not present in the finite-dimensional case. As the result itself is (nearly) standard, the proof may be skipped, as none of the following material will require the concepts it uses.

Theorem 2.1 (Lax-Milgram) *Given $A \in \mathscr{B}(\mathsf{V}, \mathsf{V}^*)$ with V a reflexive complex Banach space, suppose*

$$|(Au, u)| \geq c\|u\|_{\mathsf{V}}^2 \tag{2.12}$$

for all $u \in \mathsf{V}$ and some $c > 0$. Then A is invertible and $A^{-1} \in \mathscr{B}(\mathsf{V}^*, \mathsf{V})$ with $\|A^{-1}\| \leq c^{-1}$.

Proof From the coercivity assumption and the generalized Cauchy-Schwarz inequality, we find that for all $u \in \mathsf{V}$,

$$c\|u\|_{\mathsf{V}}^2 \leq |(Au, u)| \leq \|Au\|_{\mathsf{V}^*}\|u\|_{\mathsf{V}} \tag{2.13}$$

so that

$$\|Au\|_{\mathsf{V}^*} \geq c\|u\|_{\mathsf{V}} \ . \tag{2.14}$$

We conclude [2, p. 146] that A has bounded inverse A^{-1} with domain $\mathsf{D}(A^{-1}) = \mathsf{R}(A)$, the range of A, and that $\|A^{-1}\| \leq c^{-1}$. We must yet show $\mathsf{R}(A) = \mathsf{V}^*$.

Because $|(A^*u, u)| = |\overline{(Au, u)}| = |(Au, u)|$, we see that A^* is also coercive on V, which by reflexivity is the whole domain of A^* ($\mathsf{D}(A^*) = \mathsf{V}^{**} = \mathsf{V}$). The above argument thus applies equally well to A^*, and we conclude in particular that the null space of A^* is trivial, $\mathsf{N}(A^*) = \{0\}$. Recall that the annihilator $\mathsf{S}^\perp$ of a set $\mathsf{S} \subseteq \mathsf{V}$ is the closed subspace of V^* consisting of those forms f for which $(f, u) = 0$ for all $u \in \mathsf{S}$, and that $\mathsf{S}^{\perp\perp}$ is the closure of the span of S [2, p. 136], provided V is reflexive as we have assumed. Since $\mathsf{N}(A^*) = \mathsf{R}(A)^\perp$ [2, p. 168], it follows that the closure of $\mathsf{R}(A)$ is

$$\mathsf{R}(A)^{\perp\perp} = \mathsf{N}(A^*)^\perp = \{0\}^\perp = \mathsf{V}^* \ . \tag{2.15}$$

If we can show that $\mathsf{R}(A)$ is closed, then it would follow that $\mathsf{R}(A) = \mathsf{V}^*$.

But it is a standard result that a bounded operator has closed range when it is bounded below. Here, the argument runs as follows. Let Au_n be an arbitrary Cauchy sequence in $\mathsf{R}(A)$ so that $\|Au_n - Au_m\|_{\mathsf{V}^*} \to 0$ as $n, m \to \infty$. From (2.14) we see that $\|u_n - u_m\|_{\mathsf{V}} \le c^{-1}\|Au_n - Au_m\|_{\mathsf{V}^*} \to 0$ so that u_n is a Cauchy sequence in V. By completeness, the sequence u_n has some limit u, and by continuity (boundedness of A) $Au_n \to Au$ in V^*. Thus $\mathsf{R}(A)$ is closed, and it follows that $\mathsf{R}(A) = \mathsf{V}^*$ and $A^{-1} \in \mathscr{B}(\mathsf{V}^*, \mathsf{V})$. □

We will actually be concerned with a slightly more restricted class of operators A, namely those with positive real part H.

Corollary 2.1 *If $A \in \mathscr{B}(\mathsf{V}, \mathsf{V}^*)$ has positive real part H, i.e.,*

$$\mathrm{Re}\,(Au, u) = (Hu, u) \ge c\|u\|_{\mathsf{V}}^2 \tag{2.16}$$

for all $u \in \mathsf{V}$ and some $c > 0$, then $A^{-1} \in \mathscr{B}(\mathsf{V}^*, \mathsf{V})$.

Proof $|(Au, u)| \ge \mathrm{Re}\,(Au, u)$. □

2.3 The “Energy” Norm and Dual Hilbert Space

A symmetric form $\mathfrak{h}$ has lower bound c, written $c \le \mathfrak{h}$, if

$$\mathfrak{h}[u] \ge c\|u\|_{\mathsf{V}}^2 \tag{2.17}$$

for all $u \in \mathsf{V}$. For the corresponding linear operator $H : \mathsf{V} \to \mathsf{V}^*$, defined by $(Hu, v) = \mathfrak{h}[u, v]$, we write $c \le H$. Assume $0 < c \le \mathfrak{h}$. Then the sesquilinear form $\mathfrak{h}$ is an inner product on V. Assuming $\mathfrak{h}$ (and thus H) is bounded, the induced norm

$$\|u\|_H^2 := \mathfrak{h}[u] = (Hu, u) \tag{2.18}$$

is equivalent to the Banach norm on V, as for all $u \in \mathsf{V}$,

$$c^{1/2}\|u\|_{\mathsf{V}} \le \|u\|_H \le \|H\|^{1/2}\|u\|_{\mathsf{V}}\ . \tag{2.19}$$

In particular, V is complete under the new metric and thus a Hilbert space. The Lax-Milgram lemma implies that H has inverse $H^{-1} \in \mathscr{B}(\mathsf{V}^*, \mathsf{V})$ with $\|H^{-1}\| \le c^{-1}$. The form defined by H^{-1} is also an inner product, as it is symmetric with positive lower bound, since, for all $f \in \mathsf{V}^*$,

$$(H^{-1}f, f) = \|H^{-1}f\|_H^2 \ge c\|H^{-1}f\|_{\mathsf{V}}^2 \ge \frac{c}{\|H\|^2}\|f\|_{\mathsf{V}^*}^2\ , \tag{2.20}$$

i.e., $c\|H\|^{-2} \le H^{-1}$. Note here that the duality pairing is of $\mathsf{V}^{**} = \mathsf{V}$ and V^*. The arguments above apply also to H^{-1}, and V^* is a Hilbert space under the inner product defined by H^{-1}. In this case the norm equivalence is given by

$$\frac{c^{1/2}}{\|H\|}\|f\|_{\mathsf{V}^*} \le \|f\|_{H^{-1}} \le c^{-1/2}\|f\|_{\mathsf{V}^*} \tag{2.21}$$

for all $f \in \mathsf{V}^*$. By construction, H is the isometric Riesz representation mapping between these dual Hilbert spaces.

$$\|Hu\|_{H^{-1}} = \|u\|_H \ , \qquad \|H^{-1}f\|_H = \|f\|_{H^{-1}} \ . \tag{2.22}$$

Versions of many familiar results involving the Banach norms arise with the Hilbert space norms replacing the Banach norms. We first consider the generalized Cauchy-Schwarz inequality $|(f, u)| \le \|f\|_{\mathsf{V}^*}\|u\|_{\mathsf{V}}$, which is little more than the definition of $\|f\|_{\mathsf{V}^*}$. We see that

$$|(f, u)| = |\mathfrak{h}[H^{-1}f, u]| \le \|H^{-1}f\|_H\|u\|_H = \|f\|_{H^{-1}}\|u\|_H \tag{2.23}$$

for all $u \in \mathsf{V}$, $f \in \mathsf{V}^*$, where we have used the Cauchy-Schwarz inequality for the inner product $\mathfrak{h}$. The Hilbert space norms have the following alternative characterizations.

$$\|f\|_{H^{-1}} = \sup_{\|u\|_H=1} |(f, u)| \ , \qquad \|u\|_H = \sup_{\|f\|_{H^{-1}}=1} |(f, u)| \ . \tag{2.24}$$

By the generalized Cauchy-Schwarz inequality (2.23), $\|f\|_{H^{-1}} \ge |(f, u)|$ for $\|u\|_H = 1$, with equality holding for $u = H^{-1}f/\|H^{-1}f\|_H$. This establishes the first identity; the second may be established by similar reasoning. The characterization of $\|f\|_{H^{-1}}$ in (2.24) is the definition of the norm of the adjoint space, and so we have just proved that these Hilbert spaces are dual, as previously asserted.

Suppose we have two reflexive complex Banach spaces, V_1 and V_2, and let $H_1 \in \mathscr{B}(\mathsf{V}_1, \mathsf{V}_1^*)$ and $H_2 \in \mathscr{B}(\mathsf{V}_2, \mathsf{V}_2^*)$ be positive symmetric operators. To refer explicitly to the norm of T as an operator between the Hilbert spaces associated with H_1 and H_2 we will use the notation

$$\|T\|_{H_2,H_1} := \sup_{u\in\mathsf{V}_1\setminus\{0\}} \frac{\|Tu\|_{H_2}}{\|u\|_{H_1}} \ , \tag{2.25}$$

which is the standard definition of the norm of an operator between two Banach spaces (here with the Hilbert space norms). When $H_1 = H_2 = H$, we will write simply $\|T\|_H$.

The norm of the adjoint of an operator is equal to the norm of the operator,

$$\|T^*\|_{H_1^{-1},H_2^{-1}} = \|T\|_{H_2,H_1} \ . \tag{2.26}$$

This is just the familiar identity $\|T^*\| = \|T\|$ [2, p. 154], but applied to T considered as an operator between the two Hilbert spaces, and taking account that T^* is an operator between the two dual spaces.

2.4 Sectorial Forms

Let us return our attention now to the nonsymmetric operator A, which we shall assume to have positive real part $H > 0$. If the imaginary part is bounded by

$$|\mathfrak{k}[u]| \leq (\tan\theta)\mathfrak{h}[u] \ , \tag{2.27}$$

for all $u \in \mathsf{V}$, then $\mathfrak{a}$ is said to be sectorial with vertex 0 and corresponding semi-angle θ (see Fig. 2.1). If A is bounded in addition to having positive real part, then $\mathfrak{a}$ is always sectorial, as we may take

$$\tan\theta = \|K\|_{H^{-1},H} \ . \tag{2.28}$$

This follows from the generalized Cauchy-Schwarz inequality,

$$|\mathfrak{k}[u]| = |(Ku, u)| \leq \|Ku\|_{H^{-1}} \|u\|_H \leq \|K\|_{H^{-1},H} \|u\|_H^2 = \|K\|_{H^{-1},H} \mathfrak{h}[u] \ . \tag{2.29}$$

We see that the angle θ, which falls in the range $0 \leq \theta < \pi/2$, is a direct measure of the "nonsymmetry" of A, being a measure of the size of the imaginary part relative to the real part. Note that $\theta = 0$ for positive symmetric A, as was the case for Poisson's equation as examined in Sect. 1.1.1, while for the advection-diffusion example of Sect. 1.2.1, we have on the other hand $\theta \to \pi/2$ as the diffusion coefficient $p \to 0$.

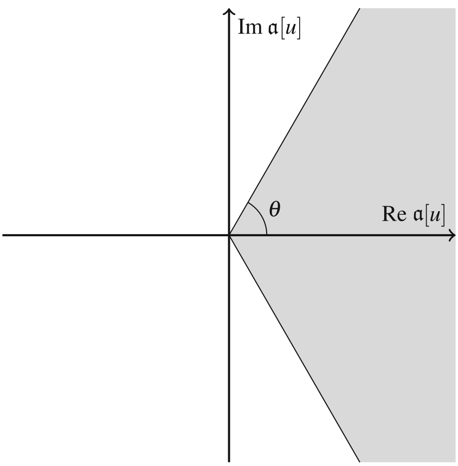


Fig. 2.1 Sectorial form with semi-angle θ. The quadratic form $\mathfrak{a}[u]$ takes on values in the *shaded sector* of the complex plane

It is easy to construct new Hilbert spaces that "pair" with those associated with H and H^{-1} to make A an isometry. Consider $\|Au\|_{H^{-1}}$.

$$\|Au\|_{H^{-1}}^2 = (H^{-1}Au, Au) = (A^*H^{-1}Au, u) \ . \tag{2.30}$$

Let us construct the norm $\|u\|_M = \|Au\|_{H^{-1}}$ by defining

$$M := A^*H^{-1}A = AH^{-1}A^* = H + KH^{-1}K \ , \tag{2.31}$$

which follows since $A = H + \mathrm{i}K$ while $A^* = H - \mathrm{i}K$. Note that $0 \le KH^{-1}K = K^*H^{-1}K$ so that $H \le M$, i.e., $(Hu, u) \le (Mu, u)$ for all $u \in \mathsf{V}$. As for H, we will occasionally use the notation M_A when we wish to make the dependence on A explicit. Here we remark that $M_A = M_{A^*}$. The Lax-Milgram lemma applies to M, so that $M^{-1} \in \mathscr{B}(\mathsf{V}^*, \mathsf{V})$. Indeed, M^{-1} is the real part of A^{-1}:

$$M^{-1} = A^{-1}HA^{-*} = \tfrac{1}{2}A^{-1}(A + A^*)A^{-*} = \tfrac{1}{2}(A^{-1} + A^{-*}) \ . \tag{2.32}$$

In other words,

$$H_{A^{-1}} = M_A^{-1} \ , \quad \text{and hence} \quad M_{A^{-1}} = H_A^{-1} \ . \tag{2.33}$$

By construction, we have the following identities, analogous to (2.22), which state that A is an isometry between the indicated Hilbert spaces.

$$\begin{aligned} \|Au\|_{H^{-1}} &= \|u\|_M \ , & \|A^{-1}f\|_M &= \|f\|_{H^{-1}} \ , \\ \|Au\|_{M^{-1}} &= \|u\|_H \ , & \|A^{-1}f\|_H &= \|f\|_{M^{-1}} \ . \end{aligned} \tag{2.34}$$

These identities remain valid when A is replaced by A^*. In any of these identities, note that H and M may be exchanged—they enjoy a kind of "duality," though not in the sense of the dual space. The sectorial semi-angle provides an equivalence between the two norms. We have

$$\|u\|_H \le \|u\|_M \le (\sec\theta)\|u\|_H \tag{2.35}$$

for any u. This follows from (2.28) and (2.31), noting that

$$\|u\|_M^2 - \|u\|_H^2 = \|Ku\|_{H^{-1}}^2 \le (\tan\theta)^2\|u\|_H^2 \tag{2.36}$$

and $\sec^2\theta = 1 + \tan^2\theta$. Dually, by taking $u = A^{-1}f$ we see that

$$\|f\|_{M^{-1}} \le \|f\|_{H^{-1}} \le (\sec\theta)\|f\|_{M^{-1}} \tag{2.37}$$

for all $f \in \mathsf{V}^*$.

By making use of the isometry identities (2.34) and the generalized Cauchy-Schwarz inequality (2.23), as well as the corresponding one for M, we see that

$$\begin{aligned} |(Au, v)| &\leq \|Au\|_{H^{-1}} \|v\|_H = \|u\|_M \|v\|_H \ , \\ |(Au, v)| &\leq \|Au\|_{M^{-1}} \|v\|_M = \|u\|_H \|v\|_M \ . \end{aligned} \tag{2.38}$$

Each member of this symmetric pair of inequalities resembles a Cauchy-Schwarz inequality, though two norms are involved and the form $\mathfrak{a}$ is not symmetric. The weaker result

$$|(Au, v)| \leq (\sec\theta) \|u\|_H \|v\|_H \tag{2.39}$$

then follows from (2.35). For the quadratic form, we have

$$\|u\|_H^2 \leq |(Au, u)| \leq \|u\|_H \|u\|_M \leq (\sec\theta) \|u\|_H^2 \ . \tag{2.40}$$

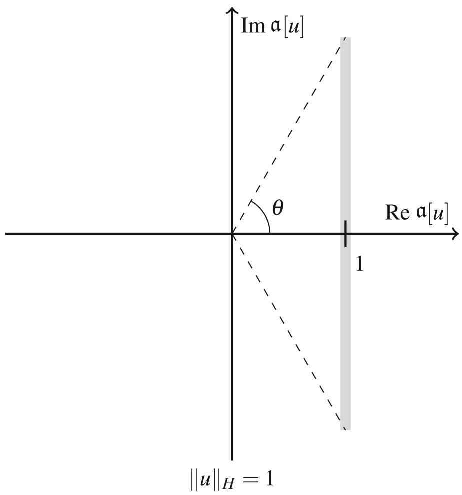
These inequalities bounding the quadratic form, apart from the intermediate bound involving M, can be found in Kato [2, p. 311], and are evident from studying Fig. 2.2. The dual versions of these inequalities are

$$|(f, A^{-1}g)| \leq \begin{matrix} \|f\|_{H^{-1}} \|g\|_{M^{-1}} \\ \|f\|_{M^{-1}} \|g\|_{H^{-1}} \end{matrix} \leq (\sec\theta) \|f\|_{M^{-1}} \|g\|_{M^{-1}} \ , \tag{2.41}$$

$$\|f\|_{M^{-1}}^2 \leq |(f, A^{-1}f)| \leq \|f\|_{M^{-1}} \|f\|_{H^{-1}} \leq (\sec\theta) \|f\|_{M^{-1}}^2 \ . \tag{2.42}$$

These follow from the former by taking $u = A^{-1}f$, $v = A^{-1}g$ and making use of the isometry identities (2.34).

Fig. 2.2 When u is normalized such that $\|u\|_H = 1$, the quadratic form $\mathfrak{a}[u]$ takes on values on the *shaded line segment* in the complex plane



2.5 On Energy

It is perhaps worth pausing to consider the physical significance of the abstract Hilbert spaces we have been considering. The norm $\|u\|_H$ is often called an "energy" norm. It is indeed sometimes the case that $\|u\|_H^2$ corresponds to a physical energy. For example, the Poisson equation (see Sect. 1.1.1) can be used to model the condition of equilibrium for a membrane subjected to some load. In this case $\|u\|_H^2$ is the potential energy of the membrane.

More generally, consider a time-dependent version of $Au = f$,

$$Q\dot{u} + Au = f \ , \tag{2.43}$$

where $Q \in \mathscr{B}(\mathsf{V}, \mathsf{V}^*)$ is nonnegative symmetric and $\dot{u} := du/dt$. For example, if we take Q to be the operator corresponding to the L^2 inner product, then for our Poisson equation example (Sect. 1.1.1) we recover the heat equation, while our steady-state advection-diffusion example (Sect. 1.2.1) becomes time-dependent. In either case, if we test (2.43) on u and take the real part, we find

$$\frac{d}{dt}\left(\tfrac{1}{2}\|u\|_Q^2\right) = -\|u\|_H^2 + \operatorname{Re}(f, u) \tag{2.44}$$

since $\frac{d}{dt}\|u\|_Q^2 = (Qu, \dot{u}) + (Q\dot{u}, u) = 2\operatorname{Re}(Q\dot{u}, u)$. As we have assumed $H > 0$, we can infer that the "energy" $\|u\|_Q^2$ cannot increase in the absence of a source (i.e., $f = 0$). We can also see that the imaginary part K of A leaves $\|u\|_Q^2$ invariant. This is an example of an "energy" method, a technique for analyzing PDEs (see e.g., Evans [1, pp. 42, 65, 86]). Despite the fact that the Navier-Stokes equations are a nonlinear and a vector version of the advection-diffusion example, an energy *inequality* similar to the above equality nevertheless holds. For that case, u is the fluid velocity vector, $\frac{1}{2}\|u\|_Q^2$ is the kinetic energy of the fluid, and $-\|u\|_H^2$ is the viscous dissipation of kinetic into thermal energy. That the energy norm does not depend on the imaginary part K physically corresponds to the fact that the convection transports kinetic energy without diminishing it.

The M norm does not appear to have any common name. It is the dual to the "energy" norm associated with A^{-1} (recall $M_A^{-1} = H_{A^{-1}}$), but there is no similar physical significance to attach to it, as, e.g., there is no physically significant analog of the time-dependent equation (2.43) for A^{-1}.

It is apparent that the energy norm $\|u\|_H^2$ is useful and has importance in certain contexts even for nonsymmetric problems. However, for other purposes it is inadequate. Returning to the problem we are actually concerned with, $Au = f$, one might like to use the Cauchy-Schwarz type inequality (2.39) or the bound on the quadratic form (2.40) in a convergence analysis of the multigrid method. Unfortunately, these bounds involve $\sec\theta$, and as $\sec\theta \to \infty$ as $\theta \to \pi/2$, they become useless for the highly nonsymmetric problems that interest us. We can see that this difficulty stems precisely from the fact that $\|u\|_H^2$ is completely independent of the imaginary (nonsymmetric) part of A, which we have just seen to be a useful and significant property in a different context.

In this chapter, we have reviewed the basic concepts of the abstract setting we have chosen for the analysis of multigrid methods applied to a nonsymmetric form defined by an operator A with positive real part H. We have also looked at some Hilbert spaces associated with A, such as that associated to H and the "energy" norm. One more such Hilbert space and norm will be investigated in the next chapter, one associated with a new notion of the absolute value $|A|$ of A, which, unlike H, does take account of the imaginary part K of A, and which will prove ideal for the multigrid analysis in the subsequent chapter.

References

1. Evans, L.C.: Partial Differential Equations. Graduate Studies in Mathematics, vol. 19. American Mathematical Society (1998)
2. Kato, T.: Perturbation Theory for Linear Operators. Springer, Heidelberg (1966)
3. Roşca, I.: Bilinear coercive and weakly coercive operators. An. Univ. Bucureşti Mat. (2), 183–188 (2002)

Springer
http://www.springer.com/978-3-319-56305-3

Towards Robust Algebraic Multigrid Methods for Nonsymmetric Problems
Lottes, J.
2017, X, 131 p. 21 illus., 15 illus. in color., Hardcover
ISBN: 978-3-319-56305-3