

A canonical small divisor problem for the Nash-Moser method

BLAKE TEMPLE* AND ROBIN YOUNG*

In this note we prove a general elementary small divisor theorem for H^s norms of $N \times N$ matrices that provides a potentially useful estimate for expunging resonances in Nash-Moser Newton Iterations. The theorem requires compatibility conditions on the approximating matrices, and we investigate how the theorem can fail when the compatibility conditions are violated. This investigation suggests that establishing compatibility of the approximations, not the presence of small eigenvalues, is the main obstacle in obtaining small divisor theorems sufficient for expunging resonances in Nash-Moser Newton Iterations.

1. Introduction

We consider the problem of obtaining H^s estimates for the inverses of $N \times N$ complex matrices $A(\theta)$ depending on parameter θ in the compact interval

$$\theta \in I \equiv [\underline{\theta}, \bar{\theta}] \subset \mathbb{R}.$$

We assume that $A(\theta)$ depends differentiably on the parameter θ in the Euclidean L^2 norm $\|\cdot\|$ in the sense that there exists a constant $C_1 > 0$ such that, for every θ_0, θ in I ,

$$(1) \quad \|A(\theta) - A(\theta_0) - A'(\theta_0)(\theta - \theta_0)\| \leq C_1 |\theta - \theta_0|^2.$$

Assume also that $A'(\theta)$ is a bounded invertible matrix with forward and backward bounds independent of θ , that is, there exists C_2 independent of θ such that

$$(2) \quad \|A'(\theta)\| \leq C_2 \quad \text{and} \quad \|A'(\theta)^{-1}\| \leq C_2.$$

*Temple supported in part by NSF Applied Mathematics Grant Number DMS-040-6096. Young supported in part by NSF Applied Mathematics Grant Number DMS-010-4485.

This is indicative of a setting in which θ represents the period and $A(\theta)$ represents a finite dimensional approximation for the linearized operator that comes up in a Nash-Moser type Newton iteration. For example, if the problem is to construct periodic solutions to some evolution problem with inhomogeneous constraints, then (1) and (2) are natural bounds, applying in the case when the derivative of the linearized operator depends only on the invertible evolution, and not the period independent constraints. The connection to authors' program to prove existence of periodic solutions of compressible Euler is discussed in Section 5 below, but the simplicity of the assumptions (1) and (2) make for a canonical small divisor problem of mathematical interest in its own right, c.f. [4, 5].

The main issue in proving convergence of the Newton method is typically that $A(\theta)$ can fail to be invertible at *resonant* values of the period, and the central problem is then how to expunge a neighborhood of the resonant values of θ . At each stage of the method, one requires a so-called *small divisor theorem*, an estimate on inverses of $A(\theta)$ on the complement of the expunged set, such that the sum total of these expunged periods is of sufficiently small measure that periods remain in the limit. This strategy is motivated by the understanding that the quadratic convergence of Newton allows for weaker and weaker estimates on the inverses as the iteration proceeds, thereby making it sufficient to expunge a smaller set of periods at each step. This quadratic convergence is sufficient to handle growth rates in the norm of the inverses of order N^b for some constant b independent of N , (we refer to this as an *NM-small divisor estimate*), but not, for example, growth rates on the order of 2^N .

The purpose of this note is to isolate conditions on fixed $N \times N$ matrices $A \equiv A(\theta)$ sufficient to prove an *NM-small divisor estimate* for values of the angle in a neighborhood of θ , in the complement of a set of measure $O(1)N^{-a}$ for some constant $a > 1$ depending on b . The main point is that estimates for the smallest eigenvalue are natural to the problem, but then to obtain a *NM-small divisor theorem*, one must estimate the norm of A on the unit ball from below, (i.e., the norm of its inverse), in terms of the smallest eigenvalue. This is equivalent to estimating the smallest singular value in terms of the smallest eigenvalue, difficult to do directly. For this we show that one must impose compatibility conditions on the matrix

$$\tilde{A} \equiv AB^{-1}, \quad \text{where} \quad B(\theta) \equiv A'(\theta).$$

We note that all of our results below are easily modified if instead we set $\tilde{A} = B^{-1}A$.

We present examples that demonstrate how, in the absence of any compatibility conditions, the small divisor theorem can fail. We show that it can fail in essentially two different ways. First, in the case that \tilde{A} is diagonalizable, it can fail if the matrix of eigenvectors of \tilde{A} become singular too fast in the limit $N \rightarrow \infty$. And second, it can fail in a more subtle way if the nilpotent part of \tilde{A} dominates the small divisors of \tilde{A} in the limit $N \rightarrow \infty$, in a sense made precise in Section 3. We view this as a canonical example of how the presence of out of control nilpotents, however unlikely in a given problem, must be ruled out to conclude a useful theorem. A conclusion of this article, then, is that the main difficulty in obtaining an NM -small divisor estimate on the iterations $A(\theta)$ of a Newton method is not so much the presence of small eigenvalues in $A(\theta)$, but rather the problem of verifying the approximations satisfy the compatibility conditions. That is, (1) and (2) are not enough, but rather more subtle information about the structure of the matrices $\tilde{A}(\theta)$ at the points of resonance must be derived from the problem in order to make a small divisor theorem feasible.

To start, consider the simplest case in which \tilde{A} is symmetric. Recall that the L^2 -norm of a diagonal matrix is equal to the norm of its largest eigenvalue, and this extends directly to real symmetric matrices because they can be diagonalized by an orthogonal matrix. That is, if \tilde{A} is a fixed symmetric matrix, then

$$\tilde{A} = R\Lambda R^{-1},$$

where $\Lambda \equiv \text{diag}\{\lambda_i, i = 1, \dots, N\}$ are the eigenvalues of \tilde{A} and R is orthogonal with $\|R\| = \|R^{-1}\| = 1$, so

$$\|\tilde{A}\| \leq \|R\|\|\Lambda\|\|R^{-1}\| \leq \|\Lambda\| = |\lambda_{max}|,$$

where $\lambda_{max} = \max_i |\lambda_i|$. Fixing now the angle θ and matrix \tilde{A} , we have that for each real number ϵ , (we're thinking of ϵ as a (small) change in the angle θ in a sense made precise in Lemma 1 below), the eigenvalues of $\tilde{A} - \epsilon I$ are $\lambda_i - \epsilon$ and the eigenvalues of $(\tilde{A} - \epsilon I)^{-1}$ are $1/(\lambda_i - \epsilon)$. It follows that for \tilde{A} symmetric, we can estimate

$$\left\|(\tilde{A} - \epsilon I)^{-1}\right\| \leq \frac{1}{\min_i |\lambda_i - \epsilon|}.$$

To outline our basic line of argument, we now prove a general NM -small divisor estimate for matrices $A(\theta)$ and $B(\theta)$ in a general norm $\|\cdot\|$ assuming (1) and (2) hold, together with the assumption that there exists a constant

$C_0 \geq 1$ for which the critical estimate

$$(3) \quad \left\| (\tilde{A} - \epsilon I)^{-1} \right\| \leq \frac{C_0}{\min_i |\lambda_i - \epsilon|},$$

holds. In Section 2 we study the conditions under which (3) holds in the H^s norms. In Section 3 we show that a further bound on the matrix of eigenvectors of \tilde{A} as $N \rightarrow \infty$ is required to obtain (3) for general diagonalizable matrices \tilde{A} , and finally we show that a weaker version of (3) is required to obtain satisfying estimates in the case when \tilde{A} contains a nilpotent part as well. In Section 4 our examples show that these conditions are sharp. Finally, in Section 5, we briefly describe our program for proving the existence of periodic shock-free solutions to the compressible Euler equations, for which estimates such as these are critical.

Theorem 1. *Let $N \in \mathbb{N}$ and assume $A(\theta)$ and $B(\theta) \equiv A'(\theta)$ are $N \times N$ matrices satisfying (1) and (2) for $\theta \in I$. Set $\tilde{A} \equiv \tilde{A}(\theta) = AB^{-1}$ with eigenvalues $\lambda_i \equiv \lambda_i(\theta)$, and assume (3) holds for each $\theta \in I$. Then for every constant $b \geq 1$ there exists a constant K and set $\mathcal{N}(b, N) \subset I$ of θ -Lebesgue measure*

$$(4) \quad \mu \{ \mathcal{N}(b, N) \} \leq \frac{K}{N^{\frac{b}{2}-1}},$$

such that for every $\theta \in I \setminus \mathcal{N}(b, N)$, the matrix $A(\theta)$ is invertible with bound

$$(5) \quad \|A(\theta)^{-1}\| \leq N^b.$$

The constant K in the theorem can be taken to be

$$(6) \quad K = 2 C_0 C_2 \sqrt{C_1} |I|,$$

and the constants C_1, C_0 and C_2 can grow as fixed powers of N . In this case one should choose b large enough that (4) provides a (small) useful bound, [3, 1]. The proof of Theorem 1 is based on the following sharp lemma:

Lemma 1. *Let $N \in \mathbb{N}$, let A, B be real $N \times N$ matrices, and let $\epsilon \in I_\epsilon \equiv [\underline{\epsilon}, \bar{\epsilon}]$. Assume B is invertible with bound*

$$\|B^{-1}\| \leq C_2,$$

assume (3) holds with $\tilde{A} = AB^{-1}$, and let

$$K_0 = a C_0 C_2.$$

Then for each pair of fixed constants $a \geq 1$, $b \geq 1$, there exists a set $\mathcal{N} \subset I_\epsilon$ of Lebesgue measure

$$(7) \quad \mu\{\mathcal{N}\} \leq \frac{K_0}{N^{b-1}},$$

such that for all $\epsilon \in I_\epsilon \setminus \mathcal{N}$, $A - \epsilon B$ is invertible with bound

$$(8) \quad \|(A - \epsilon B)^{-1}\| \leq \frac{N^b}{a}.$$

Note that K_0 can depend on N if C_0 and C_2 depend on N , and the dependence is given according to (7).

Proof of Lemma 1. Letting $\tilde{A} = AB^{-1}$, write

$$\|(A - \epsilon B)^{-1}\| = \|B^{-1}(\tilde{A} - \epsilon I)^{-1}\| \leq C_2 \|(\tilde{A} - \epsilon I)^{-1}\|,$$

so it suffices to estimate $\|(\tilde{A} - \epsilon I)^{-1}\|$ off a set of small measure in ϵ . But by (3),

$$\|(\tilde{A} - \epsilon I)^{-1}\| \leq \frac{C_0}{\min_i |\lambda_i - \epsilon|}.$$

We conclude that

$$(9) \quad \|(A - \epsilon B)^{-1}\| \leq C_2 \|(\tilde{A} - \epsilon I)^{-1}\| \leq \frac{C_0 C_2}{\min_i |\lambda_i - \epsilon|}.$$

Next,

$$(10) \quad \frac{C_0 C_2}{\min_i |\lambda_i - \epsilon|} \leq \frac{N^b}{a}$$

if and only if

$$|\lambda_i - \epsilon| \geq \frac{a C_0 C_2}{N^b} \quad \text{for each } i = 1, \dots, N,$$

which holds outside a set $\mathcal{N} \subset I_\epsilon$ of measure

$$\mu(\mathcal{N}) \leq N \frac{a C_0 C_2}{N^b} = \frac{a C_0 C_2}{N^{b-1}},$$

which is (7). The lemma now follows from equations (9) and (10). \square

Taking $\epsilon \equiv \theta_0 - \theta$, Lemma 1 immediately implies the following corollary, to be used in the proof of Theorem 1.

Corollary 1. *Let $N \in \mathbb{N}$ and assume $A(\theta)$ are real $N \times N$ matrices satisfying (1), (2) and (3), and let*

$$(11) \quad K_0 = 2 C_0 C_2.$$

Then for every $b \geq 1$ and $\theta_0 \in I$ there exists a set $\mathcal{N}(\theta_0, b, N) \subset I$ with Lebesgue measure

$$\mu \{ \mathcal{N}(\theta_0, b, N) \} \leq \frac{K_0}{N^{b-1}},$$

such that for every $\theta \in I \setminus \mathcal{N}(\theta_0, b, N)$ we have

$$\| [A(\theta_0) + A'(\theta_0)(\theta - \theta_0)]^{-1} \| \leq \frac{N^b}{2}.$$

(More generally, the constant K_0 could depend on N if C_0 and C_2 depend on N .)

Proof of Theorem 1. Fix $b \geq 1$ and $N \in \mathbb{N}$, choose the largest integer

$$(12) \quad M_N \leq \sqrt{C_1} |I| N^{b/2}, \quad \text{and set} \quad \delta_N = \frac{|I|}{M_N} = \frac{1}{\sqrt{C_1} N^{b/2}}.$$

We partition the interval $I = [\underline{\theta}, \bar{\theta}]$ into M_N subintervals, as follows: define

$$\theta_k^N = \underline{\theta} + k\delta_N \quad \text{for} \quad k = 0, \dots, M_N,$$

and set

$$I_k^N = [\theta_{k-1}^N, \theta_k^N] \quad \text{for} \quad k = 1, \dots, M_N.$$

For each $k = 1, \dots, M_N$, we apply Corollary 1, to obtain a set $\mathcal{N}(\theta_k^N, b, N)$ of measure K_0/N^{b-1} , such that for $\theta \in I_k^N \cap I \setminus \mathcal{N}(\theta_k^N, b, N)$, and for any unit vector $\mathbf{x} \in \mathcal{S}^{N-1}$, we have

$$\begin{aligned} \|A(\theta)\mathbf{x}\| &= \| \{ A(\theta_k^N) + A'(\theta_k^N)(\theta - \theta_k^N) \\ &\quad + A(\theta) - A(\theta_k^N) - A'(\theta_k^N)(\theta - \theta_k^N) \} \mathbf{x} \| \\ &\geq \| A(\theta_k^N) + A'(\theta_k^N)(\theta - \theta_k^N) \| \\ &\quad - \| A(\theta) - A(\theta_k^N) - A'(\theta_k^N)(\theta - \theta_k^N) \| \\ &\geq \frac{2}{N^b} - C_1 (\delta_N)^2 \end{aligned}$$

$$\geq \frac{2}{N^b} - C_1 \frac{1}{C_1 N^b} \geq \frac{1}{N^b},$$

where we have used (1) and (12). It follows that

$$\|A(\theta)^{-1}\| = \sup_{\mathbf{x}} \frac{\|\mathbf{x}\|}{\|A(\theta)\mathbf{x}\|} \leq N^b,$$

which is (5). Since there are M_N intervals from which we remove the sets $\mathcal{N}(\theta_k^N, b, N)$ of measure K_0/N^{b-1} , it follows that (5) holds off the set

$$\mathcal{N}(b, N) = \bigcup_{k=1}^{M_N} \mathcal{N}(\theta_k^N, b, N),$$

of measure

$$\mu \{\mathcal{N}(b, N)\} \leq \sum_{k=1}^{M_N} \frac{K_0}{N^{b-1}} \leq K_0 \sqrt{C_1} |I| \frac{1}{N^{\frac{b}{2}-1}}.$$

Theorem 1 now follows with the choice

$$K = K_0 \sqrt{C_1} |I|,$$

which in light of (11) yields (6). □

2. H^s -Norms

Our main interest is when a vector v is defined by a Fourier series, and the L^2 norm satisfies

$$\|v\| = \sqrt{\sum |v_k|^2} = \|\hat{v}\|_{L^2},$$

where the function $\hat{v}(t)$ is the (inverse) Fourier transform of the vector v . More generally, we measure the energy of functions via the H^s norms,

$$E(\hat{v}) \approx \left\| \frac{d^s \hat{v}}{dt^s} \right\|_{L^2} = \|D^s v\| \equiv \|v\|_s,$$

where D is the differentiation matrix. That is, D is a diagonal matrix with entries corresponding to the derivatives of the associated Fourier mode: for

example, the entry corresponding to the k -th mode e^{ikt} would be ik . It is now easy to write down the induced H^s matrix norm, namely

$$(13) \quad \|A\|_s = \sup_{v \neq 0} \frac{\|Av\|_s}{\|v\|_s} = \sup_{v \neq 0} \frac{\|D^s A v\|}{\|D^s v\|} = \|D^s A D^{-s}\|.$$

Thus in the case when \tilde{A} satisfies (3) in the L^2 -norm, we can extend (8) to the case of the H^s -norm by the estimate

$$\begin{aligned} \left\| (\tilde{A} - \epsilon I)^{-1} \right\|_s &= \left\| D^s (\tilde{A} - \epsilon I)^{-1} D^{-s} \right\| \\ &\leq \|D^s\| \left\| (\tilde{A} - \epsilon I)^{-1} \right\| \|D^{-s}\| \\ &\leq \frac{\bar{C}_0}{\min_i |\lambda_i - \epsilon|}, \end{aligned}$$

which is (3) with $\bar{C}_0 = C_0 O(N^s)$ because

$$\|D^s\| = O(N^s) \quad \text{and} \quad \|D^{-s}\| = O(1).$$

We conclude:

Theorem 2. *If (1) and (2) hold in the H^s -norm, and estimate (3) holds with constant C_0 in the L^2 -norm, then (3) holds in the H^s -norm by increasing C_0 by a factor of N^s , and Theorem 1 then holds in the H^s -norm using this new value of C_0 .*

3. General conditions

It remains to find conditions on the matrix \tilde{A} , more general than symmetric, sufficient to conclude (3) in the L^2 -norm, and thereby the conditions for the validity of Theorem 1. The difficulty in obtaining estimate (3) in the L^2 -norm for general $N \times N$ matrices \tilde{A} is that estimate (3) is given in terms of the eigenvalues of \tilde{A} , but the L^2 -norm in general is not the norm of the largest eigenvalue, but rather the largest singular value [2]. These agree on symmetric matrices.

Recall the singular value decomposition: for any real matrix M , there are orthogonal matrices U and V , together with diagonal matrix $\Sigma = \text{diag}\{\sigma_j\}$ such that

$$M = U \Sigma V^T, \quad \text{with} \quad \sigma_1 \geq \dots \geq \sigma_N \geq 0,$$

so that $M v_j = \sigma_j u_j$ for each j , and $\{u_j\}$ and $\{v_j\}$ are orthonormal bases for \mathbb{R}^N . The matrix M is invertible if and only if $\sigma_N > 0$, and in this case,

$$M^{-1} = V \Sigma^{-1} U^T.$$

The squares of the singular values σ_j^2 are the eigenvalues of the symmetric matrix $M^T M$ (and of $M M^T$), and since U and V are orthogonal, the L^2 norms of M and M^{-1} are

$$\|M\| = \sigma_1 \quad \text{and} \quad \|M^{-1}\| = 1/\sigma_N.$$

For (3), we need an estimate of the singular values of M in terms of its eigenvalues, which is a delicate problem, [2]. For example, we have the block matrix identity

$$\begin{bmatrix} 0 & M^T \\ M & 0 \end{bmatrix} \begin{bmatrix} V & V \\ U & -U \end{bmatrix} = \begin{bmatrix} V & V \\ U & -U \end{bmatrix} \begin{bmatrix} \Sigma & 0 \\ 0 & -\Sigma \end{bmatrix},$$

so that the singular values are the eigenvalues of this extended symmetric matrix; it follows that $\|M\|$ and $1/\|M^{-1}\|$ are the largest and smallest absolute eigenvalues of this block matrix, respectively, but this matrix does not respect the decomposition in (13). We will see in the examples in the next section that there is no general estimate for the singular values of a matrix in terms of its eigenvalues when the matrix is not symmetric. Instead, we obtain general conditions for (3) by means of the semisimple-nilpotent decomposition of a matrix.

Recall that the semisimple-nilpotent decomposition, or Jordan-Chevalley decomposition, of a matrix A is a sum (or difference) $A = S - Z$, in which S is semisimple (i.e. diagonalizable), and Z is nilpotent (i.e. $Z^k = 0$ for some k), and S and Z commute. This is easily realized by the Jordan normal form, $A = R J R^{-1}$: S is obtained by using the diagonal entries of J , and Z by using the entries above the diagonal, and the decomposition is unique.

Suppose, then, that $\tilde{A} = A B^{-1}$ is a general, not necessarily symmetric matrix. Then \tilde{A} decomposes into a semisimple-nilpotent sum

$$\tilde{A} = S - Z, \quad \text{with} \quad Z S = S Z,$$

where S is diagonalizable with the same eigenvalues as \tilde{A} and Z is a nilpotent matrix that commutes with S . In this case, whenever \tilde{A} is nonsingular, so is S , and we can write

$$\tilde{A}^{-1} = [S(I - S^{-1}Z)]^{-1} = S^{-1} \sum_{k \geq 0} (S^{-1}Z)^k,$$

in which the sum is finite because $S^{-1}Z$ is nilpotent.

Consider first the simpler case when $Z = 0$ and $\tilde{A} = S$ is diagonalizable. In this case we can obtain estimate (3) under the further assumption that S be *bounded by symmetric* in the sense that it can be diagonalized by a matrix R of bounded norm, i.e.,

$$(14) \quad S = R\Lambda R^{-1}, \quad \text{with} \quad \|R\| \leq C_S, \quad \|R^{-1}\| \leq C_S,$$

for some constant $C_S \geq 1$; this means that the eigenvectors of S retain a bounded aspect ratio. Assuming (14), we can directly obtain (3) by

$$(15) \quad \|(S - \epsilon I)^{-1}\| = \|\{R(\Lambda - \epsilon I)R^{-1}\}^{-1}\| \leq \frac{C_S^2}{\min_i |\lambda_i - \epsilon|},$$

which is (3), with $C_0 \equiv C_S^2$.

It remains to consider the case when $\tilde{A} = S - Z$ and the nilpotent part of Z is nonzero. In an attempt to obtain conditions under which (3) holds, set $\tilde{A} = S - Z$ and using (14), write

$$\tilde{A} - \epsilon I = R(\Lambda - \epsilon I - \tilde{Z})R^{-1},$$

where we have set $\tilde{Z} = R^{-1}ZR$, and observe that

$$\Lambda\tilde{Z} = R^{-1}SZR = R^{-1}ZSR = \tilde{Z}\Lambda,$$

and \tilde{Z} is also nilpotent. It follows that

$$(16) \quad \begin{aligned} (\tilde{A} - \epsilon I)^{-1} &= R(\Lambda - \epsilon I)^{-1} \left(I - (\Lambda - \epsilon I)^{-1} \tilde{Z} \right)^{-1} R^{-1} \\ &= R(\Lambda - \epsilon I)^{-1} \left(\sum_k \{(\Lambda - \epsilon I)^{-1} \tilde{Z}\}^k \right) R^{-1}, \end{aligned}$$

where the sum is finite. Thus, taking norms, we obtain

$$\|(\tilde{A} - \epsilon I)^{-1}\| \leq C_S^2 \|(\Lambda - \epsilon I)^{-1}\| \left\| \sum_k \{(\Lambda - \epsilon I)^{-1} \tilde{Z}\}^k \right\|.$$

Using

$$\|(\tilde{\Lambda} - \epsilon I)^{-1}\| \leq \frac{C_S^2}{\min_i |\lambda_i - \epsilon|},$$

we obtain the estimate

$$(17) \quad \|(\tilde{A} - \epsilon I)^{-1}\| \leq \frac{C_S^2}{\min_i |\lambda_i - \epsilon|} \left\| \sum_k \{(\Lambda - \epsilon I)^{-1} \tilde{Z}\}^k \right\|,$$

where again the sum is finite because $(\Lambda - \epsilon I)^{-1} \tilde{Z}$ is nilpotent, and finite for almost every ϵ by (15). Estimate (17) displays the problem in obtaining (3) when $Z \neq 0$. Indeed, each term in the nilpotent sum is bounded by $\frac{C_S^2}{\min_i |\lambda_i - \epsilon|}$, but in the worst case, there can be N nonzero terms in the sum. In fact, when $S = 0$, (16) gives the exact expression

$$(18) \quad (Z - \epsilon I)^{-1} = \sum \epsilon^{-(k+1)} Z^k,$$

a finite sum that could be as large as ϵ^{-N} in the worst case, too large for an NM-small divisor estimate. (C.f. the examples in the next section.)

We show next that to get an NM-small divisor estimate, we can weaken (3) to the condition

$$(19) \quad \left\| (\tilde{A} - \epsilon I)^{-1} \right\| \leq \frac{C_0}{\min_i |\lambda_i - \epsilon|^p},$$

for some constants $C_0, p \geq 1$, but while p can be large, it must be finite. Thus by (17), a further condition on $\tilde{A} = S - Z$ must be imposed when $Z \neq 0$. This is recorded in the following theorem.

Theorem 3. *Assume that S and Z satisfy the compatibility conditions (14) together with the condition*

$$(20) \quad \left\| \left(\sum_k \{(\Lambda - \epsilon I)^{-1} \tilde{Z}\}^k \right) \right\| \leq \frac{C_Z}{\min_i |\lambda_i - \epsilon|^{p-1}},$$

for some constants $C_Z, p \geq 1$. Then (19) holds with $C_0 = C_S^2 C_Z$.

Proof. This follows directly upon substituting estimate (20) into (17). \square

It remains to show that an NM-small divisor theorem can be obtained when (3) is replaced by the weaker estimate (19), for some $p \geq 1$ independent of N . For this it is straightforward to modify Lemma 1 and Theorem 1 to get a NM-small divisor theorem, by simply replacing step (9) in the proof of Lemma 1 by

$$\left\| (A - \epsilon B)^{-1} \right\| \leq C_2 \left\| (\tilde{A} - \epsilon I)^{-1} \right\| \leq \frac{C_0 C_2}{\min_i |\lambda_i - \epsilon|^p},$$

and carrying on with the obvious p -dependent estimates from there yields the following p -dependent modifications of Lemma 1, Corollary 1 and Theorem 1. Details are omitted as the modifications to the p -dependent case are straightforward.

Lemma 2. *Let $N \in \mathbb{N}$, let A, B be real $N \times N$ matrices, assume B is invertible with bound*

$$\|B^{-1}\| \leq C_2,$$

and assume (19) holds with $\tilde{A} = AB^{-1}$. Let $\epsilon \in I_\epsilon \equiv [\underline{\epsilon}, \bar{\epsilon}]$, and let

$$K_0 = \{a C_0 C_2\}^{1/p}.$$

Then for each pair of constants $a \geq 1, b \geq 1$ there exists a set $\mathcal{N} \subset I_\epsilon$ of Lebesgue measure

$$\mu \{\mathcal{N}\} \leq \frac{K_0}{N^{\frac{b}{p}-1}},$$

such that for all $\epsilon \in I_\epsilon \setminus \mathcal{N}$, $A + \epsilon B$ is invertible with bound

$$\|(A - \epsilon B)^{-1}\| \leq \frac{N^b}{a}.$$

(Note again that K_0 can depend on N if C_0 and C_2 depend on N .)

Corollary 2. *Let $N \in \mathbb{N}$, assume $A(\theta)$ are real $N \times N$ matrices satisfying (1), (2) and (3), and let*

$$K_0 = \{2 C_0 C_2\}^{1/p}.$$

Then for every $b \geq 1$ and $\theta_0 \in I$ there exists a set $\mathcal{N}(\theta_0, b, N) \subset I$ with Lebesgue measure

$$\mu \{\mathcal{N}(\theta_0, b, N)\} \leq \frac{K_0}{N^{\frac{b}{p}-1}},$$

such that for every $\theta \in I \setminus \mathcal{N}(\theta_0, b, N)$ we have

$$\|[A(\theta_0) + A'(\theta_0)(\theta - \theta_0)]^{-1}\| \leq \frac{N^b}{2}.$$

(Again, the constant K_0 could depend on N .)

For the theorem, choose $M_N = \sqrt{C_1} |I| N^{\frac{b}{2p}}$ at stage (12) to obtain:

Theorem 4. *Let $N \in \mathbb{N}$, and assume $A(\theta)$ and $B(\theta)$ are $N \times N$ matrices satisfying (1) and (2) for $\theta \in I$. Set $\tilde{A} \equiv \tilde{A}(\theta) = AB^{-1}$ with eigenvalues $\lambda_i \equiv \lambda_i(\theta)$, assume (19) holds for each $\theta \in I$, and let*

$$K = \{2 C_0 C_2\}^{1/p} \sqrt{C_1} |I|.$$

Then for every constant $b \geq 1$ there exists a set $\mathcal{N}(b, N) \subset I$ of θ -Lebesgue measure

$$\mu \{ \mathcal{N}(b, N) \} \leq \frac{K}{N^{\frac{b}{2p}-1}},$$

such that for every $\theta \in I \setminus \mathcal{N}(b, N)$, the matrix $A(\theta)$ is invertible with bound

$$\|A(\theta)^{-1}\| \leq N^b.$$

(Again the estimates could be generalized somewhat by letting the constants C_1, C_0 and C_2 grow as appropriate fixed powers of N .)

4. Examples

We give simple examples demonstrating conditions (14) and (18), and calculating the associated norms. We also briefly discuss the effect of using H^s norms rather than L^2 norms.

First, consider the matrix A_1 with its given eigenvalue decomposition,

$$A_1 = \begin{pmatrix} 1 & 0 \\ 2z & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ z & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -z & 1 \end{pmatrix},$$

which clearly has eigenvalues $\lambda_{\pm} = \pm 1$, and whose singular values are easily found to be

$$\sigma_{\pm} = \sqrt{1 + 2z^2 \pm 2z\sqrt{1 + z^2}},$$

so that

$$\|A_1\| = \|A_1^{-1}\| = \sigma_+ = O(1 + |z|).$$

Thus for z small, the eigenvalues control the norm, but for z large, they do not, and we have $C_S = O(1 + |z|)$ in (14).

In order to understand the effect of the H^s norm, suppose that the second component represents a k -mode, which can be modelled by the differentiation matrix

$$D = \begin{pmatrix} 1 & 0 \\ 0 & k \end{pmatrix}.$$

Then according to (13), we have

$$\|A_1\|_s = \|D^s A_1 D^{-s}\| = \left\| \begin{pmatrix} 1 & 0 \\ 2zk^s & -1 \end{pmatrix} \right\| = O(1 + |zk^s|),$$

so that the s -norm of both A_1 and A_1^{-1} grow as N^s . Here, the effect of using the H^s -norm rather than the L^2 -norm is to align the eigenvectors of A_1 , which leads to growth of norms of both A_1 and A_1^{-1} , and introduces a factor of N^s into the constant C_S of (14).

Next, consider the matrix

$$A_2 = \begin{pmatrix} 1 & 2y \\ 2z & -1 \end{pmatrix}, \quad \text{with} \quad A_2^{-1} = \frac{1}{1 + 4zy} A_2,$$

which has eigenvalues $\lambda_{\pm} = \pm\sqrt{1 + 4zy}$ and singular values

$$\sigma_{\pm} = \sqrt{1 + 2(z^2 + y^2) \pm 2|z - y|\sqrt{1 + (z + y)^2}},$$

so that

$$\|A_2\| = \sigma_+ \quad \text{and} \quad \|A_2^{-1}\| = \frac{1}{\sigma_-}.$$

With D as above, we can again see the effect of using H^s -norms: we have

$$D^s A_2 D^{-s} = \begin{pmatrix} 1 & 2yk^{-s} \\ 2zk^s & -1 \end{pmatrix},$$

so that

$$\|A_2\|_s \approx \sqrt{1 + 4z^2 k^{2s}} = O(|z|k^s)$$

for large values of s . Thus, in general, the effect of using the H^s -norm is to attach more weight to the entries a_{ij} with $i > j$: this means that the H^s -norm takes into account the effects of low modes in the high modes, while attaching less significance to the effects of high modes on the lower modes. This is to be expected as smoothness of the operators in H^s should rule out any significant cascade of energy from high modes back into low modes.

Our final example shows that control of eigenvalues alone is not enough to control the inverse: consider the $N \times N$ matrix

$$A_3 = \begin{pmatrix} \epsilon & 0 & 0 & \dots & 0 \\ -1 & \epsilon & 0 & \dots & 0 \\ 0 & -1 & \epsilon & \dots & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & \dots & 0 & -1 & \epsilon \end{pmatrix},$$

of the form $A_3 = \epsilon I - Z$ as above. Then the only eigenvalue of A_3 is ϵ , and we calculate

$$A_3^{-1} = \begin{pmatrix} \epsilon^{-1} & 0 & 0 & \dots & 0 \\ \epsilon^{-2} & \epsilon^{-1} & 0 & \dots & 0 \\ \vdots & & \ddots & & \vdots \\ \epsilon^{-N} & & \dots & \epsilon^{-2} & \epsilon^{-1} \end{pmatrix},$$

whose norm is clearly $O(\epsilon^{-N})$, and so has faster than polynomial growth [2]. In particular, for A_3 , this effect is worse when measured in H^s norms. In this case, the smallest singular value of A_3 is $O(\epsilon^N)$, and the eigenvalues cannot be used to bound $\|A^{-1}\|$. It is clear that in this case (19) fails. This example demonstrates the advantage of using singular values rather than eigenvalues in controlling the matrix, and indicates that (14) and (19) do not impose conditions if the singular values can be controlled.

5. Application – periodic solutions of a PDE

In our main application, $A(\theta)$ is a finite dimensional linearized operator that appears in a Nash-Moser type Newton iteration for periodic solutions of the compressible Euler equations. Here θ is a Hopf bifurcation parameter related to the period of our solutions. In order to define the iteration and prove its convergence, it is important to expunge values of θ where $A(\theta)$ is *resonant* in the sense that it is not invertible.

The authors have recently developed a new strategy for proving the existence of periodic solutions of the compressible Euler equations, based on estimates such as (1), (2) and (5), to be outlined in a forthcoming paper. We anticipate that the theorems developed here will yield one of the estimates necessary for completion of the program.

Our matrix problem comes from the evolution of a linear PDE, an approximation to the linearized operator of a nonlinear evolution. Expressing the solution as Fourier series and performing a finite dimensional cut-off yields an $N \times N$ linear system of ODEs, and integrating this over a fixed interval of size θ yields our $N \times N$ matrix $A(\theta)$. Thus our initial data is approximated by an N -vector v , which evolves through θ to give an end state $A(\theta)v$. The k -th step of the Newton iteration requires solution of the equation

$$A_k(\theta)v_k = y_k,$$

with bounds, which is implied by a bound on the inverse of the matrix

$A_k(\theta)$. In the Nash-Moser iteration, the size N_k of the system varies with k , and at each step, we restrict the parameter θ to avoid resonances, at which the matrix (or its limit as $N \rightarrow \infty$) fails to be invertible or has small divisors.

Because smooth solutions to hyperbolic equations are stable in H^s , we expect that our assumed smoothness condition (1) will hold in the H^s norm. Also, the structure of our periodic solutions is such that $A'(\theta)$ is inhomogeneous linearized evolution of a (projected) hyperbolic system, so (2) should also hold. As in most applications of Nash-Moser iteration, the biggest apparent obstacle is to bound the inverse matrix $A(\theta)^{-1}$: we expect that the theorems in this paper will provide a means of doing this, for an appropriate set of θ .

We expect that the bounds we obtain here will be applicable to similar problems. Thus if the problem is to construct periodic solutions to some evolution problem with constraints, and a finite dimensional approximation is available, then these are natural bounds, applying in the case when the derivative of the linearized operator depends only on the invertible evolution, and not the period independent constraints, as in [4, 5].

References

- [1] W. Craig and G. Wayne, *Newton's method and periodic solutions of nonlinear wave equations*, Comm. Pure Appl. Math. **46** (1993), 1409–1498. [MR1239318](#)
- [2] G. H. Golub and J. H. Wilkinson, *Ill-conditioned eigensystems and the computation of the jordan canonical form*, SIAM Review **18** (1976), 578–619. [MR0413456](#)
- [3] J. Moser, *A new technique for the construction of solutions of nonlinear differential equations*, Proc. Nat. Acad. Sci. **47** (1961), 1824–1831. [MR0132859](#)
- [4] B. Temple and R. Young, *A paradigm for time-periodic sound wave propagation in the compressible Euler equations*, Methods and Appls of Analysis **16** (2009), no. 3, 341–364. [MR2650801](#)
- [5] B. Temple and R. Young, *Time-periodic linearized solutions of the compressible Euler equations and a problem of small divisors*, SIAM Journal of Math. Anal. **43** (2011), no. 1, 1–49. [MR2765682](#)

BLAKE TEMPLE

DEPARTMENT OF MATHEMATICS

UNIVERSITY OF CALIFORNIA

DAVIS, CA 95616

USA

E-mail address: temple@math.ucdavis.edu

ROBIN YOUNG

DEPARTMENT OF MATHEMATICS AND STATISTICS

UNIVERSITY OF MASSACHUSETTS

AMHERST, MA 01003

USA

E-mail address: young@math.umass.edu

RECEIVED SEPTEMBER 8, 2013