

MIT Open Access Articles

Three steps on an open road

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

Citation: Strang, Gilbert. "Three steps on an open road." *Inverse Problems and Imaging* 7, no. 3 (August 2013): 961-966. © American Institute of Mathematical Sciences

As Published: <http://dx.doi.org/10.3934/ipi.2013.7.961>

Publisher: American Institute of Mathematical Sciences

Persistent URL: <http://hdl.handle.net/1721.1/81407>

Version: Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

Terms of Use: Article is made available in accordance with the publisher's policy and may be subject to US copyright law. Please refer to the publisher's site for terms of use.



THREE STEPS ON AN OPEN ROAD

GILBERT STRANG

Massachusetts Institute of Technology
Cambridge, MA 02139, USA

ABSTRACT. This note describes three recent factorizations of banded invertible infinite matrices

1. If A has a *banded inverse* : $A=BC$ with block–diagonal factors B and C .
2. *Permutations* factor into a shift times $N < 2w$ tridiagonal permutations.
3. $A = LPU$ with lower triangular L , permutation P , upper triangular U .

We include examples and references and outlines of proofs.

This note describes three small steps in the factorization of banded matrices. It is written to encourage others to go further in this direction (and related directions). At some point the problems will become deeper and more difficult, especially for doubly infinite matrices. Our main point is that matrices need not be Toeplitz or block Toeplitz for progress to be possible.

An important theory is already established [2, 9, 10, 13-16] for non-Toeplitz “band-dominated operators”. The Fredholm index plays a key role, and the second small step below (taken jointly with Marko Lindner) computes that index in the special case of permutation matrices.

Recall that banded Toeplitz matrices lead to Laurent polynomials. If the scalars or matrices $a_{-w}, \dots, a_0, \dots, a_w$ lie along the diagonals, the polynomial is $A(z) = \sum a_k z^k$ and the bandwidth is w . The required index is in this case a winding number of $\det A(z)$. Factorization into $A_+(z)A_-(z)$ is a classical problem solved by Plemelj [12] and Gohberg [6-7]. (Always there are new questions.) A non-Toeplitz matrix is less beautiful and its factorization is less constructive, but it is not hopeless.

Step 1. Banded matrices with banded inverses

A banded permutation matrix, or more generally a banded orthogonal matrix, will have a banded inverse. In those cases A^{-1} is just A^T . Another special case is a *block-diagonal matrix*, because the inverse is also block-diagonal. We just invert each block. Our goal is to describe all matrices with this unusual property: banded with banded inverse.

These matrices form a group. If B and C have bandwidth w , their product BC has bandwidth $2w$. If B^{-1} and C^{-1} are banded, $C^{-1}B^{-1}$ is also banded. In finite dimensions the group contains all invertible matrices—not exciting. In infinite dimensions it appears to be new. Our description of these matrices will take the form of a factorization:

2010 *Mathematics Subject Classification*. Primary: 58F15, 58F17; Secondary: 53C35.

Key words and phrases. Banded matrix, banded inverse, doubly infinite matrix, elimination, factorization.

The author is supported by NSF grant 1023152 and a MathWorks Professorship at MIT.

Theorem 0.1. *If A and A^{-1} have bandwidth w , then $A = BC$ with block-diagonal B and C . The blocks have size $2w$ and the blocks in C are offset by w rows from the blocks in B .*

That word “offset” is the key. Let me illustrate the case $w = 1$ by an example. Offset 2 by 2 blocks in B and C produce a pair of *singular blocks* side by side in $BC = A$:

$$\begin{bmatrix} \bullet & & & \\ & \begin{bmatrix} a & b \\ c & d \end{bmatrix} & & \\ & & \bullet & \end{bmatrix} \begin{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} & \\ & \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} \bullet & \bullet & & \\ \begin{bmatrix} 3a & 4a \\ 3c & 4c \end{bmatrix} & \begin{bmatrix} 5b & 6b \\ 5d & 6d \end{bmatrix} & & \\ & & \bullet & \bullet \end{bmatrix}$$

Columns 2 and 3 of B are multiplying rows 2 and 3 of C . On the right side, the singular blocks of $A = BC$ ensure (by Asplund’s condition below) that A^{-1} is banded.

The key to a banded inverse was provided by Asplund [1]. It is a “rank condition” on all the left (lower) and right (upper) submatrices of A . Later results and later proofs are summarized in [21]. We express this rank condition in the form we need for the factorization:

If A^{-1} has bandwidth w , then the left half L and the right half R of any $2w$ consecutive rows of A have rank $\leq w$.

The left-right break is centered on the main diagonal, as in this case with $w = 2$ and four rows :

	x	x	X	x	x	
		x	x	X	x	x
L has rank 2		x	x	X	x	x
			x	x	X	x

R has rank 2

Asplund’s statement was that all left and right submatrices of A have rank $\leq w$. This becomes an equality here, since $2w$ rows of an invertible matrix must altogether have rank $2w$.

Notice the important case $w = 1$, when the left and right submatrices reach as far as the main diagonal but not beyond. For the inverse of a tridiagonal matrix ($w = 1$), all of those left and right submatrices have rank ≤ 1 .

Without giving details, we can sketch the proof of Theorem 1. Suppose A and A^{-1} have bandwidth $w = 2$, with the four rows displayed above. Row elimination will reduce row 3 and row 4 to zero in the left half L . This is left multiplication by a block of size 4. Column elimination will reduce columns 3 and 4 to zero in the right half R . This is right multiplication by an *offset block* of size 4.

If we watch carefully the effect on the next set of four rows, the elimination process can do more. The whole matrix can be reduced to the identity matrix. (Details are in [17-18].) Then the blocks B_i and C_i that invert row elimination on the left and column elimination on the right produce a block-diagonal factorization $A = BC$.

Step 2. Banded doubly infinite permutation matrices.

Permutation matrices have a single 1 in each row and column. All other entries are 0. The matrices are invertible and $P^{-1} = P^T$. But the doubly infinite case allows an anomaly, illustrated by the lower triangular shift with $S_{ij} = 1$ when $j = i - 1$. The inverse S^T is *upper* triangular. A matrix needs to be centered on its main diagonal (here it is the diagonal of 1's) before a normal factorization. Expressed differently, we must allow a power S^- as a factor of A in the doubly infinite case.

This integer κ identifies the main diagonal of P . It is κ diagonals above the main diagonal. Thus $\kappa = -1$ for the shift matrix $P = S$. That number κ is the **Fredholm index** of the matrix P_+ , formed from rows $i \geq 0$ and columns $j \geq 0$ of P . The index is the difference between the dimension of the nullspace and the codimension of the range :

$$\kappa(P) = \text{index}(P_+) = \text{dimension of } N(P_+) - \text{dimension of } N(P_+^T).$$

Two properties of this “plus-index” are especially useful :

- i. If $A = BC$ (banded matrices) then $\kappa(A) = \kappa(B) + \kappa(C)$
- ii. Changing finitely many entries of a matrix does not change κ .

Property (ii) allows $\kappa(A)$ to be computed from *any* of the lower right singly infinite submatrices, formed from rows $i \geq N$ and columns $j \geq N$. We just change N rows and columns to agree with the identity matrix. (When A is band-dominated—the limit of banded matrices—the plus-index is still independent of N . A stronger form of property (ii) allows us to change A by a compact matrix—the limit of matrices with finitely many nonzeros.)

This index is a crucial number for banded doubly infinite matrices. The joint paper [10] with Marko Lindner computes κ for any permutation matrix, and we will repeat this in Theorem 2. For more general matrices this index can be a difficult computation “at infinity”, but for permutations we may again look only at $2w$ consecutive rows.

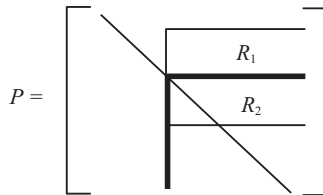
Theorem 0.2. *Suppose a permutation P has bandwidth w . Then $\kappa + w$ is equal to the number n of 1's in the right half R of any $2w$ consecutive rows.*

Example. The shift matrix S (with $w = 1$) has all zeros in its right half :

$$\begin{array}{ccc|cc} L & 1 & \mathbf{0} & 0 & \\ & & 1 & \mathbf{0} & 0 \end{array} R$$

Then $\kappa + w = n$ becomes $\kappa + 1 = 0$, yielding the correct value $\kappa = -1$.

We prove Theorem 2 by separating R into R_1 and R_2 with w rows each. The top half R_1 is outside P_+ and the lower half R_2 is inside :



Each 1 in R_1 means that the column beneath it in P_+ contains all zeros. And a zero column of P_+ corresponds to a vector $(\dots, 0, 1, 0, \dots)$ in its nullspace :

$$n_1 = \text{number of 1's in } R_1 = \text{dimension of } N(P_+)$$

Similarly a zero row in R_2 corresponds to a vector $(\dots, 0, 1, 0, \dots)$ in the nullspace of P_+^T . If R_2 has n_2 1's then it has $w - n_2$ zero rows :

$$w - n_2 = \text{dimension of } N(P_+^T).$$

Subtraction gives $n_1 + n_2 - w$ as the plus-index κ . Thus $\kappa + w = n_1 + n_2 = n$, the number of 1's in R .

Notice that $\kappa = 0$ (centered matrix) corresponds to $n = w$ (the count of 1's in R , which equals the rank). This is exactly the Asplund rank condition in Theorem 1, for the special case of permutation matrices.

Lindner has extended Asplund's theorem to a large class of *centered banded doubly infinite matrices* in [10]. We feel sure that the same ideas will apply to banded orthogonal matrices. The plus-index κ can be found from any $2w$ consecutive rows, by subtracting w from the rank n of the right half R .

The statement and proof of Theorem 1 continue to hold when the permutation matrix is infinite. It factors into $P = BC$. But Panova's neat proof [11] of our conjecture provides a more precise factorization of P :

$$P = S^- P_1 \cdots P_N \text{ with tridiagonal permutation matrices } P_i \text{ and } N < 2w.$$

From block-diagonal factors B and C we could always reach tridiagonal factors. But we might need $N = O(w^2)$ factors. The precision of Panova's result gives the best possible bound $2w - 1$, for permutation matrices. Note that a tridiagonal permutation matrix allows a set of transpositions of neighbors (2 by 2 permutations in a block-diagonal matrix).

We believe that Theorem 1 ($A = BC$) continues to hold for centered banded doubly infinite matrices with banded inverses. In this direction, which is still to complete, we took step 3.

Step 3. Triangular factorization $A = LPU$ of banded invertible infinite matrices.

Gaussian elimination subtracts multiples of rows from lower rows to reach an upper triangular matrix U . To recover the original matrix A —to add back in reverse order—put the multipliers into a lower triangular matrix L with 1's on the main diagonal. Then $A = LU$ is a perfect expression of elimination.

Row exchanges are required when one or more zeros appear in the pivot positions. Then permutation matrices P_{ij} are mixed with the row operation matrices L_{ij} . Naturally we want to separate the P_{ij} from the L_{ij} , and we have two good choices :

1. Imagine that we do all the row exchanges first, and factor the correctly ordered matrix into LU . In reverse order the exchanges come last, to give $A = PL_1U$.
2. Imagine that in elimination on each successive column $k = 1, \dots, n$, we always choose the first available nonzero entry as the pivot. The pivot row $i(k)$ is not moved into row k of the upper triangular U , until elimination is complete.

The elimination steps are still lower triangular, but they produce a different L_2 . After elimination, the pivot rows are moved by P into their correct order $1, \dots, n$ in the upper triangular U . Thus $A = L_2PU$.

The matrices P and U are the same in PL_1U and L_2PU , provided we exchange rows only when necessary. The lower triangular L_1 and L_2 are connected by $PL_1 = L_2P$.

All good linear algebra software would also exchange rows to avoid small pivots—this is not necessary algebraically, but essential in practice. The row permutation matrix P_1 goes into $A = P_1 L_1 U_1$ with no problem. But the “Bruhat form” $L_2 P_2 U_2$ with P_2 in the middle is *not possible* with those extra row exchanges, since this form has a unique P [4, 5, 20]. Thus algebraists prefer a form which numerical analysts never consider.

Now we come to $A = LPU$ for *doubly infinite matrices* (the new step). The problem is that A has no “first entry” where elimination can begin. If we start with a_{11} and go forward, then infinitely many earlier rows are untouched. A less constructive approach is needed.

Singly infinite matrices will present the same problem of getting started, if we happen to want $A = UL$. Wiener and Hopf observed that if A is Toeplitz, and its triangular factors U and L are also to be Toeplitz, then UL provides the correct order. For this we want to eliminate *upwards*, but an infinite matrix has no last entry. The way out for a Toeplitz matrix is to factor the associated polynomial $\sum a_j z^j$, but our problem is not Toeplitz.

Theorem 0.3. *An invertible banded doubly infinite matrix can be factored into $A = LPU$.*

Proof. Elimination proceeds a column at a time, using the pivot row $i(k)$ to produce zeros in column k below the pivot. In our case, that elimination step is decided by the matrix $C(k)$ formed from columns $\leq k$ and rows $\leq k + w$. Below $C(k)$ the banded matrix A has all zeros. Since A is invertible, the columns of $C(k)$ are independent and the nullspace of $C(k)$ is the zero vector. \square

Suppose d rows of $C(k)$ are dependent on earlier rows. Those must be among rows $k - w, \dots, k + w$. (All earlier rows of A are zero beyond column k . They are independent by invertibility.) Thus the Fredholm index of $C(k)$ is $0 - d$. We can prove that this number d is the same for all k .

$C(k + 1)$ includes the new column $k + 1$ and the new row $k + 1 + w$, but still d rows are dependent on earlier rows. Among the rows $k - w, \dots, k + w + 1$, one row must have become newly independent when the new column $k + 1$ was included in C . **That row will be the pivot row $i(k + 1)$.**

Example. The shift matrix S has $w = 1$, so each $C(k)$ includes column k and row $k + 1$. The 1's are on the diagonal of $C(k)$. Its index is 0 and then $d = 0$ (no dependent rows). The pivot row is $i(k) = k + 1$, and the permutation P is exactly S . No actual elimination steps are required, and LPU is just I times S times I .

The reverse shift also has $w = 1$. Now $C(k)$ has all zeros in rows k and $k + 1$, so there are $d = 2$ dependent rows. The pivot row is $i(k) = k - 1$, since it became nonzero and independent when column k was included. LPU is now I times S^T times I .

The LPU factorization could begin by finding the dependent rows in $C(0)$. Those are d rows among the final $2w$ rows of this singly infinite submatrix (and then d is determined). This step requires the solution of infinite linear systems—it is a slow start compared to the usual elimination on a finite matrix.

Ultimately a good factorization will take advantage of the bandedness of L, P , and U (they have bandwidth $\leq 2w$). In any case it will take a long time...

REFERENCES

- [1] E. Asplund, *Inverses of matrices $\{a_{ij}\}$ which satisfy $a_{ij} = 0$ for $j > i + p$* , Math. Scand., **7** (1959), 57–60.
- [2] S. N. Chandler-Wilde and M. Lindner, *Limit Operators, Collective Compactness, and the Spectral Theory of Infinite Matrices*, Amer. Math. Society Memoirs, **210** (2011).
- [3] C. de Boor, *What is the main diagonal of a banded matrix?*, in “Quantitative Approximation” (eds. R. A. DeVore and K. Scherer), Academic Press, 1980.
- [4] L. Elsner, *On some algebraic problems in connection with general eigenvalue algorithms*, Lin. Alg. Appl., **26** (1979), 123–138.
- [5] I. Gohberg and S. Goldberg, *Finite dimensional Wiener-Hopf equations and factorizations of matrices*, Lin. Alg. Appl., **48** (1982), 219–236.
- [6] I. Gohberg, S. Goldberg and M. A. Kaashoek, “Basic Classes of Linear Operators,” Birkhäuser Verlag, Basel, 2003.
- [7] I. Gohberg, M. Kaashoek and I. Spitkovsky, *An overview of matrix factorization theory and operator application*, in “Factorization and Integrable Systems” (Faro, 2000), Operator Th. Adv. Appl., **141**, Birkhäuser, Basel, (2003), 1–102.
- [8] L. Yu. Kolotilina and A. Yu. Yeremin, *Bruhat decomposition and solution of sparse linear algebra systems*, Soviet J. Numer. Anal. Math. Modelling, **2** (1987), 421–436.
- [9] M. Lindner, “Infinite Matrices and Their Finite Sections. An Introduction to the Limit Operator Method,” Frontiers in Mathematics, Birkhäuser Verlag, Basel, 2006.
- [10] M. Lindner and G. Strang, *The main diagonal of a permutation matrix*, Linear Algebra and Its Applications, **439** (2013), 524–537.
- [11] G. Panova, *Factorization of banded permutations*, Proceedings Amer. Math. Soc., **140** (2012), 3805–3812.
- [12] J. Plemelj, *Riemannsche Funktionenschar mit gegebener Monodromiegruppe*, Monat. Math. Phys., **19** (1908), 211–245.
- [13] V. S. Rabinovich, S. Roch and J. Roe, *Fredholm indices of band-dominated operators*, Integral Eqns. Oper. Th., **49** (2004), 221–238.
- [14] V. S. Rabinovich, S. Roch and B. Silbermann, “Limit Operators and Their Applications in Operator Theory,” Operator Theory: Advances and Applications, **150**, Birkhäuser Verlag, Basel, 2004.
- [15] V. S. Rabinovich, S. Roch and B. Silbermann, *The finite section approach to the index formula for band-dominated operators*, Operator Theory, **187** (2008), 185–193.
- [16] S. Roch, *Finite sections of band-dominated operators*, AMS Memoirs, **191** (2008).
- [17] G. Strang, *Banded matrices with banded inverses and $A = LPU$* , in “Fifth International Congress of Chinese Mathematicians. Part 1, 2,” AMS/IP Studies in Advanced Mathematics, **51**, American Math. Society, Providence, RI, (2012), 771–784.
- [18] G. Strang, *Fast transforms: Banded matrices with banded inverses*, Proc. Natl. Acad. Sci., **107** (2010), 12413–12416.
- [19] G. Strang, *Groups of banded matrices with banded inverses*, Proceedings Amer. Math. Soc., **139** (2011), 4255–4264.
- [20] G. Strang, *The algebra of elimination*, manuscript, 2012.
- [21] G. Strang and Tri-Dung Nguyen, *The interplay of ranks of submatrices*, SIAM Review, **46** (2004), 637–646.

Received October 2012; revised December 2012.

E-mail address: gs@math.mit.edu