



Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

SCIENCE @ DIRECT®

Advances in Applied Mathematics 33 (2004) 1–25

ADVANCES IN  
Applied  
Mathematics

[www.elsevier.com/locate/yaama](http://www.elsevier.com/locate/yaama)

# A geometric framework for sparse matrix problems

Gunnar Carlsson and Vin de Silva \*

*Department of Mathematics, Stanford, CA, USA*

Received 20 March 2002; accepted 21 August 2003

Available online 14 January 2004

---

## Abstract

In this paper, we set up a geometric framework for solving sparse matrix problems. We introduce *geometric sparseness*, a notion which applies to several well-known families of sparse matrix. Two algorithms are presented for solving geometrically-sparse matrix problems. These algorithms are inspired by techniques in classical algebraic topology, and involve the construction of a simplicial complex from certain data on the matrix. In both cases, large parts of the computation can be parallelised. © 2003 Elsevier Inc. All rights reserved.

*Keywords:* Sparse matrices; Block-triangular form; Parallel algorithms; Algebraic topology; Spectral sequences; Mayer–Vietoris

---

## 1. Introduction

The purpose of this paper is to introduce a general notion of structured sparseness for matrices, which we call *geometric sparseness*, and to construct algorithms for matrix procedures which are applicable to matrices satisfying this notion of sparseness. The algorithms we construct admit a great deal of parallelism, and so are particularly useful when doing parallel or distributed computing. We also believe that the geometric point of view is fruitful, and should be exploited.

To give an informal description of geometric sparseness, we suppose first that we are given a metric space  $X$ , which will act as a parameter space for describing the sparseness. Let  $M$  be an  $m \times n$  matrix, which we can regard as a linear map  $V \rightarrow W$  between vector spaces  $V = \mathbb{R}^n$  and  $W = \mathbb{R}^m$ . The columns and rows of  $M$  correspond to the standard basis elements of  $V$  and  $W$ , respectively.

---

\* Corresponding author.  
*E-mail address:* [silva@math.stanford.edu](mailto:silva@math.stanford.edu) (V. de Silva).

Now suppose that the columns and rows are labelled by elements of the metric space  $X$ . For  $\ell > 0$ , we say that the matrix  $M$  is  $\ell$ -sparse with respect to this labelling if the entry  $M_{ij}$  equals zero whenever their row and column labels are separated by a distance greater than  $\ell$  in  $X$ . Equivalently, this means that basis elements are sent by  $M$  to linear combinations of nearby basis elements—nearby in the sense that the corresponding labels in  $X$  lie within distance  $\ell$ . Here are some examples.

**Example 1.** Let  $X$  be the real line, and suppose that each row and column of  $M$  is labelled by its index. For any integer  $\ell$ , the matrix  $M$  is  $\ell$ -sparse if  $M_{ij} = 0$  whenever  $|i - j| > \ell$ . In other words,  $M$  is a *banded matrix*, with a band of width  $2\ell + 1$  about the diagonal (Fig. 2, left).

**Example 2.** Consider the problem of numerically solving a linear partial differential equation in the plane, say on the unit square. First create a grid of points in the unit square. This grid of points is regarded as a basis for the set of vectors of possible values of functions at the grid points. The partial differential operator  $D$  can be represented by a finite-difference approximation  $M = M_D$ , which is a linear transformation on this vector space of approximate functions. To compute a finite difference approximation to a partial derivative only requires the values at nearby grid points. Thus  $M$  is  $\ell$ -sparse for suitable  $\ell$ , with  $X$  taken to be the unit square.

Our original motivating example was the following.

**Example 3.** Suppose  $X$  is a simplicial complex embedded in a Euclidean space. For each  $k \geq 0$ , we let  $C_k(X)$  denote the vector space spanned by the  $k$ -simplices of  $X$ . In computing the homology groups of the simplicial complex  $X$ , we must evaluate the null space of the boundary linear transformation  $\partial_k : C_k(X) \rightarrow C_{k-1}(X)$ , which is defined by

$$\partial_k([v_0, \dots, v_k]) = \sum_{i=0}^k (-1)^i [v_0, \dots, v_{i-1}, \hat{v}_i, v_{i+1}, \dots, v_k].$$

If we now view each  $k$ -simplex (respectively  $(k - 1)$ -simplex) as a basis element for  $C_k(X)$  (respectively  $C_{k-1}(X)$ ) labelled in  $X$  by its barycenter, and if we let  $\ell$  denote the maximal diameter of any simplex of  $X$ , then the operator  $\partial_k$  is  $\ell$ -sparse.

As will be seen later, our methods return full circle to the idea of a simplicial complex; the general theme is to exploit the topological structure of the metric space  $X$ , and this often involves the consideration of a *nerve complex* denoted  $\mathcal{N}$ .

To construct an algorithm for Gauss or Gauss–Jordan elimination for an  $\ell$ -sparse matrix, the key input we require is a *covering* of  $X$  by subsets. For instance, the real line can be covered by intervals  $[i, i + 1]$ . The covering has to be chosen in a way which is adapted to the parameter  $\ell$ . The philosophy of our method is to solve subproblems depending on the basis elements which lie in particular subsets of the coverings; these problems can be solved independently (and hence concurrently) for each such subset. This gives a flexible

approach for creating concurrent algorithms for solving matrix problems in the presence of geometric sparseness.

A word about the origin of these ideas. In classical algebraic topology, one is often faced with the problem of computing the homology groups (or indeed any other algebraic invariant) of a topological space. A standard practice is to divide the problem into a collection of localised sub-problems where the answer is already known. One way to do this is to break up the space itself; for instance a decomposition  $X = U \cup V$  allows us to compute the homology  $H_*(X)$  in terms of  $H_*(U)$ ,  $H_*(V)$  and  $H_*(U \cap V)$ . This is the famous Mayer–Vietoris theorem; a more complicated Mayer–Vietoris *spectral sequence* applies when the decomposition involves several pieces  $U_1, \dots, U_p$  rather than just two. The theory of spectral sequences involves far-reaching generalisations of this idea; our two main algorithms are designed to mimic two specific spectral sequences from algebraic topology. The interested reader can consult [6] to learn about the subject in detail; background material in algebraic topology can be found in [1,5]. However, our treatment is self-contained and does not require familiarity with spectral sequences.

Finally, we wish to emphasise that the main point of this paper is to endorse a geometric point of view for solving sparse matrix problems. Our spectral sequence algorithms suggest ways in which geometric knowledge can be exploited. In one case, this leads to a strategy for generating block upper-triangular representations (quite different from the Dulmage–Mendelsohn decomposition [3]). Although this is of independent interest, what is more important is the idea that geometric information about the origins of a matrix should be retained as a guide to solving problems involving that matrix.

## 2. Geometry and sparseness

In this section we give a formal definition of geometric sparseness, and illustrate the geometric approach with several examples.

### 2.1. Basic examples

If  $M$  is an  $m \times n$  matrix, then the set of rows  $R$  can be identified with the set  $\{1, 2, \dots, m\}$  and the set of columns  $C$  can be identified with the set  $\{1, 2, \dots, n\}$ . Let  $\phi: R \rightarrow X$  and  $\psi: C \rightarrow X$  be mappings to some metric space  $X = (X, d)$ .

**Definition 4.** The matrix  $M$  is  $\ell$ -sparse with respect to  $(\phi, \psi, X)$  if  $M_{ij} = 0$  whenever  $d(\phi(i), \psi(j)) > \ell$ .

For convenience, we will often suppress some of the arguments  $\ell, \phi, \psi$  and  $X$ , or simply say that  $M$  is *geometrically sparse*.

**Example 5.** All matrices are  $\infty$ -sparse whatever  $X$  is; and 0-sparse whenever  $X$  is a one-point space.

Block-diagonal matrices are defined as follows. Suppose we are given a disjoint partition  $R = R_1 \sqcup \cdots \sqcup R_p$  of the set of rows of a matrix  $M$ , and a disjoint partition  $C = C_1 \sqcup \cdots \sqcup C_p$  of the set of columns. The block  $B_{ij}$  is defined to be the submatrix of  $M$  where the rows of  $R_i$  meet the columns of  $C_j$ . The matrix  $M$  is block-diagonal if  $B_{ij} = 0$  whenever  $i \neq j$ .

**Example 6.** A matrix  $M$  is 0-sparse if and only if it is block-diagonal with  $p$  blocks; where  $p$  is equal to the number of distinct points in  $\text{Im}(\phi) \cup \text{Im}(\psi)$ .

**Proof.** Given a 0-sparse matrix  $M$ , we can enumerate the distinct points in  $\text{Im}(\phi) \cup \text{Im}(\psi)$  as  $x_1, \dots, x_p$ . If we set  $R_i = \phi^{-1}(x_i)$  and  $C_i = \psi^{-1}(x_i)$  then  $M$  is block-diagonal. Indeed, if  $i \neq j$  then  $d(x_i, x_j) > 0$ , which implies that  $B_{ij} = 0$ . Conversely, if  $M$  is block-diagonal then take  $X = \{1, \dots, p\}$  and let  $\phi, \psi$  map each row or column to the index of the set  $R_i$  or  $C_j$  that contains it.  $\square$

In Example 5, the metric space  $X$  is essentially treated as an indiscrete space; whereas in Example 6 it is effectively a discrete space with  $p$  points. The more points, the better; since linear equations involving a block-diagonal matrix can be solved by treating each index  $i = 1, \dots, p$  separately, we recall:

**Remark 7.** The Gauss–Jordan elimination algorithm for solving matrix equations requires  $O(m^2n)$  operations for an  $m \times n$  matrix. If the matrix is block diagonal with (for instance) equal square blocks of size  $(m/p) \times (n/p)$ , then the total number of operations is  $O(m^2n/p^2)$ .

In general, the most natural space  $X$  for a given matrix  $M$  is neither indiscrete nor discrete but somewhere in between. Our constant goal is to find correspondingly improved linear algebra algorithms which take advantage of the topological structure of  $X$ .

We saw in the introduction that banded matrices are geometrically sparse, with  $X$  equal to the real line (or indeed any interval containing  $\text{Im}(\phi) \cup \text{Im}(\psi)$ ). This is the simplest 1-dimensional case. Conversely, if  $M$  is geometrically sparse with respect to the real line (or an interval), then  $M$  is block-banded. Figure 1 shows the non-zero entries of a typical matrix with sparse structure modelled on an interval. It is natural to reorder the rows and columns according to their position along the interval, as shown on the right. In this case if we decompose the interval into six subintervals, we obtain a  $6 \times 6$  block decomposition within which the matrix is banded of width 3.

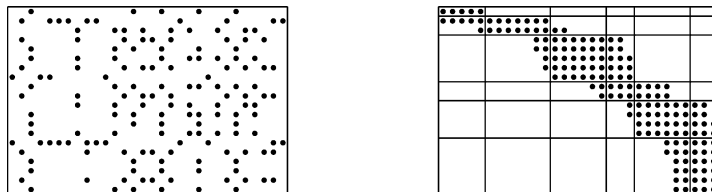


Fig. 1. A matrix with 1-dimensional sparse structure, before and after reordering the rows and columns.

We can combine Examples 1 and 2 from the introduction.

**Example 8.** The discrete Laplacian operator on the line is an example of an (infinite) banded matrix of width 3. Let  $\mathcal{H}$  be the vector space of functions  $\mathbb{Z} \rightarrow \mathbb{R}$ . The operator  $\Delta : \mathcal{H} \rightarrow \mathcal{H}$  defined by

$$\Delta f(i) = 2f(i) - (f(i-1) + f(i+1))$$

is the discrete Laplacian and has matrix

$$\Delta_{ij} = \begin{cases} 2 & \text{if } i = j, \\ -1 & \text{if } |i - j| = 1, \\ 0 & \text{otherwise.} \end{cases}$$

More generally, if  $f \in \mathcal{H}$  has finite support (say  $f(i) = 0$  for  $|i| > \ell$ ), then convolution with  $f$  is a linear operator whose matrix is banded of width  $2\ell + 1$ .

Finally, the product of two geometrically sparse matrices may also be geometrically sparse.

**Theorem 9.** If  $M_1$  is  $\ell_1$ -sparse with respect to  $(\phi, \chi, X)$  and  $M_2$  is  $\ell_2$ -sparse with respect to  $(\chi, \psi, X)$  then  $M_1 M_2$  is  $(\ell_1 + \ell_2)$ -sparse with respect to  $(\phi, \psi, X)$ .

**Proof.** The only way for  $[M_1 M_2]_{ij}$  to be non-zero is if for some  $k$  both  $[M_1]_{ik} \neq 0$  and  $[M_2]_{kj} \neq 0$ . Then  $d(\phi(i), \chi(k)) \leq \ell_1$  and  $d(\chi(k), \psi(j)) \leq \ell_2$  so by the triangle inequality  $d(\phi(i), \psi(j)) \leq \ell_1 + \ell_2$ .  $\square$

## 2.2. Examples with $X = S^1$

### 2.2.1. Cyclically banded matrices

Another family of geometrically sparse matrices has parameter space  $X = S^1$ , the circle. The most common examples are *cyclically banded* matrices (Fig. 2, right). These are  $n \times n$  matrices where the entry  $M_{ij}$  is non-zero only if  $i$  and  $j$  are at most  $\ell$  steps apart modulo  $n$  (so  $n$  is adjacent to 1, for example). Thus, in addition to the band around the main diagonal, there are some extra non-zero entries in the top-right and bottom-left corners.

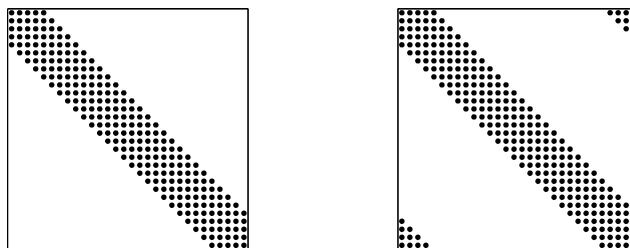


Fig. 2. A banded matrix ( $X = \text{interval}$ ) and a cyclically banded matrix ( $X = \text{circle}$ ).

Formally, let  $X$  be the unit circle in  $\mathbb{R}^2$ , and choose  $n$  equally spaced points on it:

$$x_k = [\cos(2\pi k/n), \sin(2\pi k/n)].$$

We identify the rows and columns of the  $n \times n$  matrix  $M$  with these points; so  $\phi(k) = \psi(k) = x_k$ . Setting  $\lambda = |x_1 - x_{1+\ell}|$ , the matrix  $M$  is  $\lambda$ -sparse if and only if it is cyclically banded in the sense described above. The width of the band is  $2\ell + 1$ .

This kind of matrix occurs in the numerical solution of linear differential equations on the circle; equivalently, on a real interval with periodic boundary conditions. As in the non-periodic case, a function  $f$  is represented by the sequence of values  $f_i = f(x_i)$ , where  $x_0, x_1, \dots, x_n$  is an evenly-spaced sequence of points covering the interval. Differential operators are approximated by linear maps, the simplest possibility being

$$\left[ \frac{df}{dx} \right]_i \approx \frac{f_{i+1} - f_{i-1}}{2\delta},$$

where  $\delta$  is the separation  $x_{i+1} - x_i$ . In the periodic case, we interpret all the formulas modulo  $n$ , so in particular  $f_n = f_0$ , and  $f_{n\pm k} = f_{\pm k}$  where required. The matrix corresponding to a differential operator is then an  $n \times n$  matrix, and is cyclically banded. The width of the band depends on the degree of the differential operator, and also on the quality of the approximation used. The approximate first-order derivative above has a cyclic band of width 3, but other approximations having wider bandwidth converge better with respect to the separation  $\delta$ .

There is a converse to the fact that cyclically banded matrices are sparse with  $X = S^1$  (compare the example in Fig. 1):

**Remark 10.** If a matrix is geometrically sparse with parameter space  $X = S^1$ , then it can be given a  $p \times p$  cyclically banded block decomposition, by dividing the circle into  $p$  equal arcs. Even without the blocks, this structure is evident after putting the rows and columns in the cyclic order defined by  $\phi$  and  $\psi$ .

A variation on the theme allows not just one cyclic band, but several regularly spaced bands (Fig. 3, left). Here we have a  $pn \times pn$  matrix, whose rows and columns are mapped to  $n$  equally spaced points  $x_1, \dots, x_n$  on the unit circle, in a  $p$ -to-1 fashion:

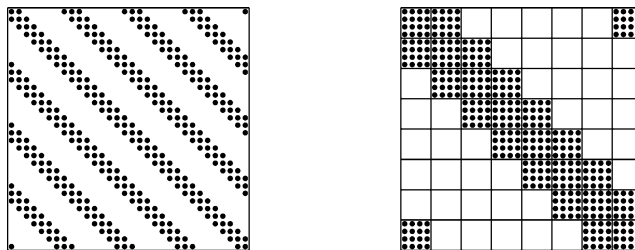


Fig. 3. Two views of a block-cyclic matrix ( $X = \text{circle}$ ).

$$\begin{aligned} \phi(k) &= \phi(k + n) = \dots = \phi(k + (p - 1)n) = x_k, \\ \psi(k) &= \psi(k + n) = \dots = \psi(k + (p - 1)n) = x_k. \end{aligned}$$

A  $\lambda$ -sparse matrix with  $\lambda = |p_1 - p_{1+\ell}|$  will then have  $p$  cyclic bands of width  $2\ell + 1$ .

The matrix in the figure has  $n = 8$ ,  $p = 4$  and  $\lambda = 1$ . For example, row 1 has non-zero entries in columns 1, 2, 8, 9, 10, 16, 17, 18, 24, 25, 26, 32; that is, at indices congruent to 0, 1, 2 modulo 8. This gives four cyclic bands of width 3.

On the other hand, if we order the rows and columns cyclically as suggested in Remark 10, the matrix takes the form shown in Fig. 3 (right). Here we have a block-structure with a single cyclic band of width 3. This explains the name “block-cyclic” in the caption.

### 2.2.2. Geometric transformations

We can interpret certain well-known matrix transformations in terms of geometric operations on the space  $X$ . For instance, suppose  $M$  is a cyclically banded matrix, and we wish to regard it as a banded matrix. In other words,  $X = S^1$  but we might prefer  $X = \text{interval}$ . Geometrically, this can be done by ‘squashing’ the circle, as shown in Fig. 4 (top). If we present the matrix by reordering the rows and columns as specified by their position on the interval, then it becomes a genuine banded matrix (Fig. 4, bottom). If the matrix has width  $2\ell + 1$  in its original cyclically banded form, then the new form is banded with width  $4\ell + 1$ .

### 2.3. Examples with $X = \mathbb{R}^2$

We now turn to examples where the sparse matrix has a two-dimensional geometric structure.

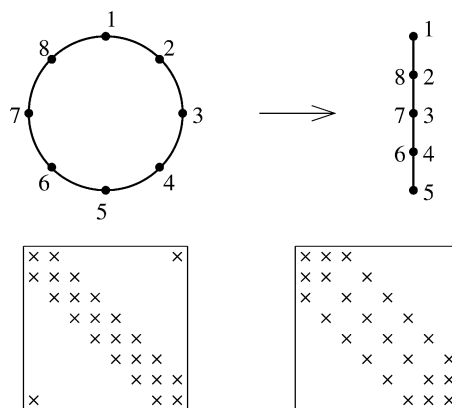


Fig. 4. Squashing a circle (here with 8 points) onto an interval, and the resulting matrix transformation.

### 2.3.1. Differential and difference operators

**Example 11.** The discrete Laplacian on the plane is an example of an infinite 1-sparse matrix with  $X = \mathbb{R}^2$ . Let  $\mathcal{H}$  be the vector space of functions  $\mathbb{Z}^2 \rightarrow \mathbb{R}$ , and let  $\Delta: \mathcal{H} \rightarrow \mathcal{H}$  be defined by the formula:

$$\Delta f(i, j) = 4f(i, j) - (f(i + 1, j) + f(i - 1, j) + f(i, j + 1) + f(i, j - 1)).$$

Then the matrix of  $\Delta$  is 1-sparse with respect to the natural labelling of its rows and columns by elements of  $\mathbb{Z}^2$ .

One can also consider the problem of finding approximate solutions to the continuous Laplace equation  $\Delta f = 0$ , say on the unit square  $X = [0, 1] \times [0, 1]$  in the plane. This time we choose a grid of points  $\{(x_i, y_j)\}$  where  $x_1, \dots, x_n$  and  $y_1, \dots, y_n$  are evenly spaced sequences of points covering the interval  $[0, 1]$ . The function  $f$  is represented by its values  $f_{i,j} = f(x_i, y_j)$  at the grid points; thus  $f$  is replaced by a vector of dimension  $n^2$ .

Second order partial derivatives may be represented by the finite difference quotients

$$\left[ \frac{\partial^2 f}{\partial x^2} \right]_{i,j} \approx \frac{f_{i+1,j} - 2f_{i,j} + f_{i-1,j}}{\delta^2},$$

$$\left[ \frac{\partial^2 f}{\partial y^2} \right]_{i,j} \approx \frac{f_{i,j+1} - 2f_{i,j} + f_{i,j-1}}{\delta^2},$$

where  $\delta$  is the separation  $x_{i+1} - x_i = y_{j+1} - y_j$ . The approximate Laplacian (also denoted  $\Delta$ ) is then the sum of these two terms, and is regarded as a linear map  $\mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n \times n}$ . Elements of the null space of  $\Delta$  are approximations to solutions of Laplace's equation.

Since  $[\Delta f]_{i,j}$  depends only on  $f_{i,j}$ ,  $f_{i \pm 1, j}$  and  $f_{i, j \pm 1}$ , the map  $\Delta$  is geometrically sparse. Specifically, let  $X$  be the unit square, and let

$$\phi(i, j) = \psi(i, j) = (x_i, y_j)$$

be the maps identifying the canonical basis vectors of  $\mathbb{R}^{n \times n}$  with grid points. With respect to these maps, the approximate Laplacian is  $\ell$ -sparse, where  $\ell = \delta$ . Just as before, more sophisticated approximations exist where the value of  $\ell$  is larger.

Note that in this case we do not have a simple structural description of the matrix (such as 'banded' or 'cyclically banded'). Since the space  $X$  is two-dimensional, there is no natural way of linearly ordering the rows and columns of  $\Delta$  to make the structure of the sparseness apparent.

The same idea applies to any linear partial differential operator over a two-dimensional domain. As the degree of the operator becomes larger, one must increase the sparseness parameter  $\ell$ . One can also consider periodic or doubly-periodic boundary conditions. In those cases, the correct parameter spaces are the cylinder  $X = \mathbb{R} \times S^1$  and the torus  $X = S^1 \times S^1$  respectively.

### 2.3.2. Kronecker products

The approximate Laplacian  $\Delta$  described above can be regarded as an algebraic combination of one-dimensional operators. To see this, let  $D$  denote the operator on  $\mathbb{R}^n$  specified by the formula:

$$[Df]_i = \frac{f_{i+1} - 2f_i + f_{i-1}}{\delta^2}.$$

Thus  $D$  is an approximation to the differential operator  $d^2/dx^2$  on a real interval, say  $[0, 1]$ ; here  $\mathbb{R}^n$  approximates the space of functions on the interval.

As above,  $\mathbb{R}^{n \times n}$  approximates the space of functions on the square  $[0, 1] \times [0, 1]$ . If we regard it as a tensor product

$$\mathbb{R}^{n \times n} = \mathbb{R}^n \otimes \mathbb{R}^n$$

then the approximate Laplacian can be expressed by the formula

$$\Delta = D \otimes 1_n + 1_n \otimes D$$

where  $1_n$  is the identity on  $\mathbb{R}^n$ . The left- and right-hand terms correspond to  $\partial^2/\partial x^2$  and  $\partial^2/\partial y^2$ , respectively. The tensor product of two matrices is sometimes known as a Kronecker product: the MATLAB expression `kron(A, B)` produces the block matrix whose  $(i, j)$ th block equals  $a_{ij}B$ .

Geometric sparseness of  $\Delta$  with respect to  $[0, 1] \times [0, 1]$  follows from geometric sparseness of  $D$  and  $1_n$  with respect to  $[0, 1]$ . We can state this as a general fact:

**Theorem 12.** *Let  $M_1$  be  $\ell_1$ -sparse with respect to the metric space  $(X_1, d_1)$ , and let  $M_2$  be  $\ell_2$ -sparse with respect to  $(X_2, d_2)$ . Then  $M_1 \otimes M_2$  is  $(\ell_1 + \ell_2)$ -sparse with respect to  $X_1 \times X_2$ , where the product metric is taken to be  $d = d_1 + d_2$ , that is:*

$$d((x_1, x_2), (y_1, y_2)) = d_1(x_1, y_1) + d_2(x_2, y_2).$$

**Proof.** The rows of  $M_1 \otimes M_2$  are labelled by pairs  $(i_1, i_2)$  where  $i_1$  is a row of  $M_1$  and  $i_2$  is a row of  $M_2$ . It is natural to define:

$$\phi(i_1, i_2) = (\phi_1(i_1), \phi_2(i_2)) \in X_1 \times X_2$$

and similarly

$$\psi(j_1, j_2) = (\psi_1(j_1), \psi_2(j_2)) \in X_1 \times X_2$$

for the columns. The entry in the  $(i_1, i_2)$ th row and  $(j_1, j_2)$ th column of  $M_1 \otimes M_2$  is non-zero only if the  $(i_1, j_1)$ th entry of  $M_1$  and the  $(i_2, j_2)$ th entry of  $M_2$  are both non-zero. If  $d((i_1, j_1), (i_2, j_2)) > \ell_1 + \ell_2$  then one of these terms must be zero.  $\square$

The choice of product metric is not important. If we replace  $d = d_1 + d_2$  by  $d = \max(d_1, d_2)$ , then  $M \otimes M_2$  is  $\max(\ell_1, \ell_2)$ -sparse, for instance.

### 3. Matrix coverings

The rows and columns of a geometrically sparse matrix can be organised in certain ways that are suitable for fast parallel computations. This process leads to a re-indexing of the rows and columns in terms of the cells of a particular simplicial complex  $N$ , whose geometry reflects that of the original indexing space  $X$ . The aim of this section is to set up these ideas, in theory and in practice.

#### 3.1. Coverings and the nerve complex

Let  $M$  be a matrix. A full covering of  $M$  is specified by the following data:

- (1) A covering of the set of rows  $R = R_1 \cup \dots \cup R_p$ , where the subsets  $R_i$  are not necessarily disjoint.
- (2) A covering of the set of columns  $C = C_1 \cup \dots \cup C_p$  where the subsets  $C_i$  are not necessarily disjoint.

It is called a *CR-covering* of  $M$  if the following condition is satisfied: if  $M_{ij} \neq 0$  then  $j \in C_k \Rightarrow i \in R_k$  for all  $k$ . Dually, it is an *RC-covering* of  $M$  if the following condition holds: if  $M_{ij} \neq 0$  then  $i \in R_k \Rightarrow j \in C_k$  for all  $k$ .

If  $M$  is regarded as a linear map between vector spaces  $V \rightarrow W$ , then row and column coverings can be regarded as decompositions of  $V$  and  $W$  into subspaces:

$$V = \sum_{i=1}^p V_i, \quad W = \sum_{i=1}^p W_i.$$

Here  $V_i$  is the subspace spanned by basis vectors corresponding to the columns included in  $C_i$ , and  $W_i$  is the subspace spanned by basis vectors corresponding to the rows included in  $R_i$ . The meaning of a *CR-covering* is precisely that  $M(V_i) \subseteq W_i$  for all  $i$ ; and of an *RC-covering* precisely that  $M^t(W_i) \subseteq V_i$  for all  $i$ .

Geometrically sparse matrices have natural *CR*- and *RC*-coverings, which depend on a choice of covering of the space  $X$  by pieces  $X_1, \dots, X_p$ . The following theorem makes this precise.

**Theorem 13.** *Suppose  $M$  is  $\ell$ -sparse with respect to some  $(\phi, \psi, X)$ . Let  $X = X_1 \cup \dots \cup X_p$  be any cover of  $X$ , and let  $Y_i = N(X_i, \ell)$  be the  $\ell$ -neighborhood<sup>1</sup> of  $X_i$  for  $i = 1, \dots, p$ . Then  $R_i = \phi^{-1}(Y_i)$ ,  $C_i = \psi^{-1}(X_i)$  defines a *CR-covering* of  $M$ . Similarly,  $R_i = \phi^{-1}(X_i)$ ,  $C_i = \psi^{-1}(Y_i)$  defines an *RC-covering* of  $M$ .*

**Proof.** We first look at the *CR* case, with  $C_i, R_i$  defined as stated. We must show that  $M(V_i) \subseteq W_i$  for every  $i$ . Let  $v$  be the basis vector of  $V$  corresponding to a particular column  $c \in C_i$ . We know that  $\phi(c) \in X_i$ . Metric sparseness tells us that  $M(v)$  is a linear

<sup>1</sup> The  $\ell$ -neighborhood is defined to be:  $N(A, \ell) = \{y \in X \mid \text{there exists } x \in A \text{ with } d(x, y) \leq \ell\}$ .

combination of basis vectors corresponding to rows  $r$  which are ‘close’ to  $c$ , in that  $d(\psi(r), \phi(c)) \leq \ell$ . This implies that  $\psi(r) \in Y_i$ , so these rows all belong to  $R_i$  as claimed. The  $RC$  case is dual to the  $CR$  case.  $\square$

Even without an explicit indexing space  $X$ , we can readily construct  $CR$ - and  $RC$ -coverings. For a  $CR$  covering, start with an arbitrary column covering  $C = C_1 \cup \dots \cup C_p$ . A compatible row covering  $R = R_1 \cup \dots \cup R_p$  is obtained in the following way. In  $R_i$  we include (indeed, we are forced to include) every row  $r$  which has a non-zero entry in any of the columns  $C_i$ . If  $R_i$  is defined to be exactly this set of rows, the result is a  $CR$ -covering, with each  $R_i$  of minimal size.

Similarly, one can start with a row covering and find a (unique minimal) column covering to make up an  $RC$ -covering. However, it is not generally possible to build a  $CR$ -covering from a given row covering; nor to build an  $RC$ -covering from a given column covering.

We return to geometry. Given a full covering of a matrix  $M$ , there is an associated simplicial complex  $\mathcal{N}$ . In the situation of Theorem 8, this complex may be regarded as a coarse approximation to the metric space  $X$ .

**Definition 14.** The *nerve complex* of a full covering is a simplicial complex  $\mathcal{N}$  with vertices  $\{1, \dots, p\}$ . The simplex  $\sigma = \llbracket i_0 \dots i_k \rrbracket$  belongs to  $\mathcal{N}$  if and only if  $R_{i_0} \cap \dots \cap R_{i_k} \neq \emptyset$  or  $C_{i_0} \cap \dots \cap C_{i_k} \neq \emptyset$ .

It is convenient to define subspaces

$$V_\sigma = V_{i_0} \cap \dots \cap V_{i_k}, \quad W_\sigma = W_{i_0} \cap \dots \cap W_{i_k}$$

for each simplex  $\sigma = \llbracket i_0 \dots i_k \rrbracket$  in  $\mathcal{N}$ . Whenever  $\sigma \geq \tau$  we have inclusion maps

$$I_{\sigma\tau} : V_\sigma \rightarrow V_\tau, \quad I_{\sigma\tau} : W_\sigma \rightarrow W_\tau$$

and projection maps

$$P_{\tau\sigma} : V_\tau \rightarrow V_\sigma, \quad P_{\tau\sigma} : W_\tau \rightarrow W_\sigma.$$

Note that  $I_{\sigma\tau}$  and  $P_{\tau\sigma}$  are adjoint. If  $\tau$  is empty (so  $V_\tau = V$ ,  $W_\tau = W$ ) we simply write  $I_\sigma$ ,  $P_\sigma$ . Finally, we define  $M_\sigma$  to be the submatrix of  $M$  mapping  $V_\sigma \rightarrow W_\sigma$ ; that is,  $M_\sigma = P_\sigma M I_\sigma$ .

**Lemma 15.** A full covering of a matrix  $M$  is a  $CR$ -covering if and only if  $M_\tau I_{\sigma\tau} = I_{\sigma\tau} M_\sigma$  for all  $\sigma, \tau$ . It is an  $RC$ -covering if and only if  $M_\sigma P_{\tau\sigma} = P_{\tau\sigma} M_\tau$  for all  $\sigma, \tau$ .

**Proof.** The original definition essentially characterises  $CR$ -coverings by the identity  $M I_i = I_i M_i$  for  $i = 1, \dots, p$ ; and  $RC$ -coverings by the identity  $M_i P_i = P_i M$  for  $i = 1, \dots, p$ . The more general identities involving  $\sigma, \tau$  are a consequence of these. We omit the details.  $\square$

### 3.2. Implementation

To make the constructions of the previous section more concrete, we now present a simple system for representing the relevant information computationally. We use logical vectors and matrices, whose entries consist of zeros and ones. A useful operation is  $A \mapsto A^\varepsilon$ , which replaces each non-zero entry of  $A$  by a 1. In MATLAB the expression  $(A \sim = 0)$  computes  $A^\varepsilon$ .

Let  $M$  be an  $m \times n$  matrix; the goal is to find a  $CR$ -covering of  $M$  and to determine the set of simplices of the nerve  $\mathcal{N}$ .

**Step 1** (*representing column and row coverings*). A column covering  $C = C_1 \cup \dots \cup C_p$  may be represented as an  $n \times p$  logical matrix  $D_c$  where the  $ij$ th entry equals 1 if and only if the  $i$ th column belongs to  $C_j$ . There are no all-zero rows in  $D_c$ . Similarly, a row covering may be represented as an  $m \times p$  logical matrix  $D_r$  with no all-zero rows.

**Step 2** (*building a  $CR$ -covering*). After choosing an  $n \times p$  column covering matrix  $D_c$ , the corresponding  $m \times p$  row covering matrix  $D_r$  is given by

$$D_r = (M^\varepsilon D_c)^\varepsilon$$

where the product is matrix multiplication. If there are any all-zero rows in  $D_r$ , they necessarily arise from all-zero rows in  $M$  itself. Those rows should be discarded at the outset.

**Step 3** (*constructing the simplicial complex  $\mathcal{N}$* ). The cells of  $\mathcal{N}$  correspond to distinct rows of  $D_c$  and  $D_r$ . A  $k$ -cell has exactly  $k + 1$  ones amongst the  $p$  entries. For example if  $p = 5$ , then  $[0\ 1\ 0\ 1\ 1]$  represents a triangle with edges  $[0\ 0\ 0\ 1\ 1]$ ,  $[0\ 1\ 0\ 0\ 1]$  and  $[0\ 1\ 0\ 1\ 0]$ . The MATLAB expression `unique([Dc;Dr], 'rows')` will produce a list of cells, by identifying the distinct expressions which occur in the rows of  $D_c$  and  $D_r$ .

Note that in Step 3 the list of cells is likely to be incomplete, in this sense: lower-dimensional faces of listed cells are not included if there is no row or column belonging to exactly the right combination of the  $R_i$  or the  $C_i$ . Since these inactive cells play no essential part in our main algorithms, there is in fact no need to enumerate them.

Furthermore, it is computationally inadvisable to list all the cells of  $\mathcal{N}$ , if there are any cells of comparatively high dimension—a single  $k$ -cell has  $2^{k+1} - 1$  proper faces. In contrast, the sorting procedure above lists at most  $n + m$  cells, precisely the essential ones.

## 4. The spectral sequence of a filtration

In this section we describe an approach to localising sparse matrix computations. Here the goal is to obtain a block-triangular form for a matrix which can be exploited to carry out a parallelized matrix reduction.

An important special feature is that the diagonal blocks of the triangular form are themselves block-diagonal. These sub-blocks account for most of the parallelism in the algorithm, and it is here that global calculations are broken up into independent, local calculations.

As before, the procedure is derived from a construction in algebraic topology. Let  $X$  be a topological space which can be expressed as a nested sequence of subspaces

$$\emptyset = X_{-1} \subset X_0 \subset X_1 \subset \cdots \subset X.$$

This is known as a *filtration* of  $X$ . There is a spectral sequence expressing the homology  $H_*(X)$  of  $X$  in terms of the relative homology groups  $H_*(X_{i+1}, X_i)$ . In matrix terms, this corresponds to a particular sequential procedure for reducing a block-triangular matrix.

#### 4.1. Filtered matrices

Let  $M$  be a matrix regarded as a linear map  $V \rightarrow W$ . We say that  $M$  is a *filtered matrix* if both  $V$  and  $W$  are filtered by subspaces

$$\begin{aligned} V &= V^0 \supseteq V^1 \supseteq \cdots \supseteq V^p = 0, \\ W &= W^0 \supseteq W^1 \supseteq \cdots \supseteq W^p = 0 \end{aligned}$$

and  $M(V^i) \subseteq W^i$  for all  $i = 0, \dots, p$ . In other words,  $M$  may be represented by a block-triangular structure:

$$M = \begin{bmatrix} M_{p,p} & M_{p,p-1} & \cdots & M_{p,1} \\ & M_{p-1,p-1} & \cdots & M_{p-1,1} \\ & & \ddots & \vdots \\ & & & M_{1,1} \end{bmatrix}.$$

Here the blocks are the naturally induced maps

$$M_{k,l} : (V^{l-1}/V^l) \longrightarrow (W^{k-1}/W^k)$$

between quotients of successive terms in each filtration.

A matrix is *strongly filtered* if the subspaces  $V^i, W^i$  are generated by sets of vectors drawn from the canonical bases of  $V$  and  $W$ . This implies that the block-triangular form above is realised by permuting the rows and columns of  $M$ .

The main contribution of this section is to show that one can construct strong filtrations on geometrically sparse matrices in a natural way.



#### 4.1.1. Comparison with the Dulmage–Mendelsohn decomposition

Our block-triangular form is quite different from the Dulmage–Mendelsohn decomposition (DM), which also puts sparse matrices into block-triangular form (see [3]; or [7] for a more recent account). The two approaches are tailored to different kinds of matrices. DM tends to be useful for matrices which are much sparser than those we consider in this paper. It works by comparing different maximal pairings in the row-column incidence graph; if these are plentiful then most of the matrix will tend to clump into one very large block, so very little is gained.

Another difference is that DM is canonical and valid for general sparse matrices; whereas our version involves the choice of a covering and is special to geometrically sparse matrices.

#### 4.2. The spectral sequence

There is a rich literature of techniques (cf. [2]) for reducing a sparse matrix to echelon form (or finding an  $LU$ -decomposition). In this section we present the spectral sequence approach to this task. It supports our belief that algebraic topology can shed light on matrix algebra. (See the end of this subsection for a converse.) We make no claim that the algorithm in its current form is optimised to compete with state-of-the-art sparse methods.

Let  $M$  be a matrix with a filtration, not necessarily arising from a  $CR$ -covering. For simplicity, we assume that the goal is to put  $M$  in echelon form. It is possible to turn this into an  $LU$ -decomposition, by careful book-keeping [4]; a  $QR$ -decomposition can be obtained with suitable modifications. These topics belong to a future paper.

Using elementary row operations on  $M$ , we want a new matrix  $N$  for which:

- (1) The leftmost non-zero element in each row of  $N$  is equal to 1 (unless the row is all zero). This is called the pivot element of that row.
- (2) No two pivot elements occupy the same column.

Given the above, we can order the rows of  $N$  by position of pivot (leftmost first); all-zero rows coming last. This recovers the traditional notion of echelon form, if we need it.

**Definition 19.** If  $M$  is a matrix, let  $A(M)$  denote the indices of the non-zero rows of  $M$ , and let  $Z(M)$  denote the indices of the all-zero rows of  $M$ . If  $I$  is a set of indices, then let  $M[I, :]$  be the submatrix of  $M$  consisting of the rows indicated by  $I$ .

We assume that we have some definite procedure in mind (such as Gauss–Jordan elimination) for reducing a matrix to echelon form. The spectral sequence makes repeated use of such a procedure, on matrices much smaller than  $M$  itself.

The matrix  $M$  is reduced in the following way.

- (1) For each  $k = 1, \dots, p$ , reduce the matrix  $M_{k,k}$  to an echelon form matrix  $N_{k,k}^1$ . At the same time, apply the same row operations to the matrices  $M_{k,k-1}, \dots, M_{k,1}$  to obtain matrices  $N_{k,k-1}^1, \dots, N_{k,1}^1$ .

- (2) For all  $k$  let  $A_k^1 = A(N_{k,k}^1)$  and let  $Z_k^1 = Z(N_{k,k}^1)$ . For  $i \leq k$  the submatrices  $N_{k,i}^1[A_k^1, :]$  are now part of the final answer, and will undergo no further changes.
- (3) The submatrices  $N_{k,k}^1[Z_k^1, :]$  are now all zero, and will be left alone.
- (4) Further processing is required for the submatrices  $N_{k,i}^1[Z_k^1, :]$  for  $i < k$ . Define new matrices

$$M_{k-1,i}^1 = \begin{bmatrix} N_{k-1,i}^1[A_{k-1}^1, :] \\ N_{k,i}^1[Z_k^1, :] \end{bmatrix}$$

for  $k = 2, \dots, p$  and  $i = 1, \dots, k - 1$ . Put these together to form a new matrix:

$$M^1 = \begin{bmatrix} M_{p-1,p-1}^1 & M_{p-1,p-2}^1 & \cdots & M_{p-1,1}^1 \\ & M_{p-2,p-2}^1 & \cdots & M_{p-2,1}^1 \\ & & \ddots & \vdots \\ & & & M_{1,1}^1 \end{bmatrix}.$$

In effect, we have omitted the first few rows (which are in echelon form); omitted the first few columns (which are all zero from the second block down); and rearranged the remaining rows into a matrix with  $(p - 1) \times (p - 1)$  blocks.

- (5) Repeat the procedure on the new matrix  $M^1$  to obtain  $M^2$ . Make sure to leave the submatrices  $N_{k-1,i}^1[A_{k-1}^1, :]$  unchanged (the diagonal terms  $N_{k-1,k-1}^1[A_{k-1}^1, :]$  are already in echelon form).
- (6) Continuing the iteration  $M^3, \dots, M^p$ , we find that every row is accounted for. The final answer takes the form:

$$N = \begin{bmatrix} N_{p,p}^1[A_p^1, :] & N_{p,p-1}^1[A_p^1, :] & \cdots & N_{p,1}^1[A_p^1, :] \\ & N_{p-1,p-1}^2[A_{p-1}^2, :] & \cdots & N_{p-1,1}^2[A_{p-1}^2, :] \\ & & \ddots & \vdots \\ & & & N_{1,1}^p[A_1^p, :] \\ & & & N_{1,1}^p[Z_1^p, :] \end{bmatrix}.$$

The very last block  $N_{1,1}^p[Z_1^p, :]$  is by definition all zero.

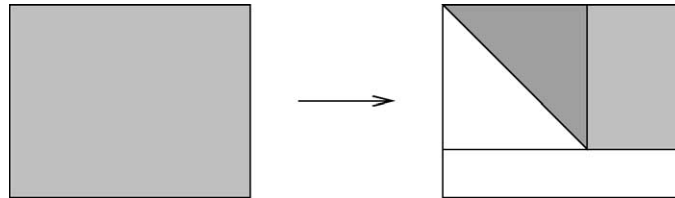


Fig. 5. Schematic representation of Gaussian elimination.

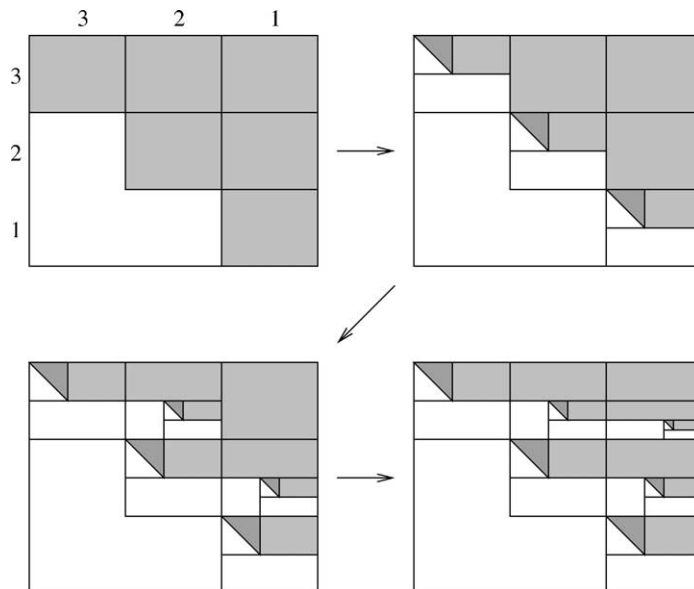


Fig. 6. The spectral sequence for a  $3 \times 3$  block-triangular matrix.

It may be helpful to give a pictorial description. Figure 5 is a schematic representation of the basic Gaussian elimination step. The dark grey triangle represents an invertible upper-triangular matrix with diagonal entries equal to 1. A column permutation may be necessary to see the matrix in exactly this form. Depending on the shape and rank of the matrix, the white rectangle (all zero) or the light grey rectangle may be missing.

Figure 6 shows the three iterations required to reduce a  $3 \times 3$  block triangular-matrix. We refer to the blocks as  $M_{ij}$  throughout (even as their entries are change); note the non-standard labelling of rows and columns.

First, we reduce the diagonal blocks  $M_{33}$ ,  $M_{22}$  and  $M_{11}$ , duplicating the row operations on the off-diagonal blocks. For the second iteration, consider  $M_{32}$ . We act on the lower ‘half’ of this block, that is on the rows whose  $M_{33}$  part is entirely zero. After using the dark grey triangle of the  $M_{22}$  block to clear the entries above it, we reduce what is left of  $M_{32}$  to echelon form (duplicating these operations on  $M_{31}$ ). We use  $M_{11}$  to do the same to  $M_{21}$ . The third iteration involves the bottom strip of  $M_{31}$ : we use the dark grey triangles of  $M_{11}$  and then  $M_{21}$  to clear the entries above them, and then reduce the remaining thin sliver to echelon form.

This pictorial description differs from the formal version in that we omit the row rearrangements of step (4), for clarity. Figure 7 shows the final result, if we perform all the rearrangements.

The authors have found that Fig. 6 gives a useful way of conceptualising spectral sequences, which are notoriously difficult to understand at first. This is the promised converse to the assertion at the beginning of this section; matrix algebra sheds light on algebraic topology. We thank Marco Grandis for pointing out that *Zeeman diagrams* [8] were devised to fulfill a similar role, in a more formal context.

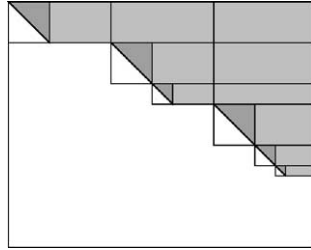


Fig. 7. Final echelon form, after rearranging the rows.

#### 4.2.1. Remarks on parallelisability

There are two degrees of parallelism in the algorithm. In step (1) (and its iterates), each index  $k = 1, \dots, p$  can be treated separately. Much more importantly, Proposition 18 tells us that each block  $M_{k,k}$  is a direct sum of sub-blocks  $M_{\sigma,\sigma}$ . Thus

$$[M_{k,k} \quad M_{k,k-1} \quad \cdots \quad M_{k,1}]$$

can be treated as

$$[M_{\sigma,\sigma} \quad M_{\sigma,k-1} \quad \cdots \quad M_{\sigma,1}]$$

for each  $\sigma$  separately.

This means that the calculations in the first iteration of step (1) can be carried out in parallel for all  $\sigma$  and all  $k$ . On subsequent iterations it is still possible to exploit the simplicial cell structure to a limited extent. However, the predominant goal is to cover as much of the calculation as possible on the first iteration. See Section 4.4.

### 4.3. Examples of block-triangularization

#### 4.3.1. A cyclically banded matrix

Consider a  $12 \times 12$  cyclically banded matrix  $M$  with bandwidth 3 (so  $\ell = 1$ ). Regarding  $M$  as a linear map  $V \rightarrow W$ , we write  $v_0, \dots, v_{11}$  and  $w_0, \dots, w_{11}$  for the standard bases of  $V$  and  $W$  respectively. The parameter space  $X$  is a circle, and if

$$\phi(k) = \psi(k) = x_k := [\cos(\pi k/6), \sin(\pi k/6)]$$

then the matrix is  $\lambda$ -sparse, where  $\lambda = |x_k - x_{k+1}|$ .

Figure 8 shows the construction of a  $CR$ -covering. On the left are the twelve points  $x_0, \dots, x_{11}$ . The circle is then covered by three sets as shown in the middle picture. This gives a column covering:

$$C_1 = \{0, 1, 2, 3, 4\}, \quad C_2 = \{4, 5, 6, 7, 8\}, \quad C_3 = \{8, 9, 10, 11, 0\}.$$

The three sets are then thickened by  $\lambda$  to give the covering shown on the right. This gives the row covering:

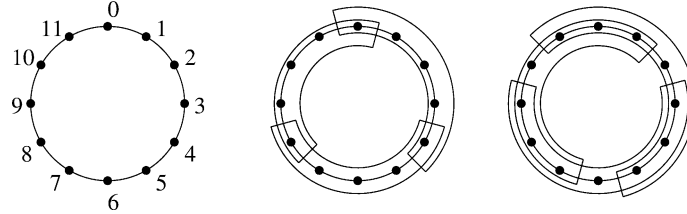


Fig. 8. Covering a circle.

$$R_1 = \{11, 0, 1, 2, 3, 4, 5\}, \quad R_2 = \{3, 4, 5, 6, 7, 8, 9\}, \quad R_3 = \{7, 8, 9, 10, 11, 0, 1\}.$$

The simplicial complex  $\mathcal{N}$  consists of the three vertices  $[[1], [2], [3]]$  and the three edges  $[[12], [23], [31]]$  of a triangle. The interior  $[[123]]$  is missing because  $C_1 \cap C_2 \cap C_3$  and  $R_1 \cap R_2 \cap R_3$  are both empty. It is now easy to identify the various subspaces  $V_\sigma, W_\sigma$ . For example,  $V_{12} = \text{Span}\{v_4\}$  since  $C_1 \cap C_2 = \{4\}$ , and  $W_{12} = \text{Span}\{w_3, w_4, w_5\}$  since  $R_1 \cap R_2 = \{3, 4, 5\}$ .

To determine the filtrations, note that 0, 4 and 8 are the only indices which belong to exactly two of the sets  $C_i$ . All other indices belong to exactly one set. Thus:

$$V^0 = V, \quad V^1 = \text{Span}\{v_0, v_4, v_8\}, \quad V^2 = \{0\}.$$

The indices  $\{0, 1, 3, 4, 5, 7, 8, 9, 11\}$  are the ones belonging to exactly two of the  $R_i$ , so we have:

$$W^0 = W, \quad W^1 = \text{Span}\{w_0, w_1, w_3, w_4, w_5, w_7, w_8, w_9, w_{11}\}, \quad W^2 = \{0\}.$$

Figure 9 shows the result of reordering the bases of  $V$  and  $W$  using the simplicial complex structure. There are  $2 \times 2$  blocks, and the diagonal blocks each break into three sub-blocks; these sub-blocks correspond to the six faces of the simplicial complex  $\mathcal{N}$ . For example, the sub-block  $\{w_3, w_4, w_5\} \times \{v_4\}$  corresponds to the edge  $[[12]]$ .

Note that the sub-blocks in the top left block  $M_{22}$  are all overdetermined (tall and thin); whereas the subblocks in the bottom right block  $M_{11}$  are underdetermined (short and wide).

	$v_0$	$v_4$	$v_8$	$v_1$	$v_2$	$v_3$	$v_5$	$v_6$	$v_7$	$v_9$	$v_{10}$	$v_{11}$
$w_{11}$	$\times$										$\times$	$\times$
$w_0$	$\times$			$\times$								$\times$
$w_1$	$\times$			$\times$	$\times$							
$w_3$		$\times$			$\times$	$\times$						
$w_4$		$\times$				$\times$	$\times$					
$w_5$		$\times$					$\times$	$\times$				
$w_7$			$\times$					$\times$	$\times$			
$w_8$			$\times$						$\times$	$\times$		
$w_9$			$\times$							$\times$	$\times$	
$w_2$				$\times$	$\times$	$\times$						
$w_6$							$\times$	$\times$	$\times$			
$w_{10}$										$\times$	$\times$	$\times$

Fig. 9. The block-triangular form of  $M$ .

It is not hard to see that this follows from the way in which the  $R_i$  are constructed as thickened versions of the  $C_i$ .

#### 4.3.2. Discrete differential operators in the plane

Consider the earlier example where  $M$  is an approximation to a partial differential operator on the square  $[0, 1] \times [0, 1]$ . If we approximate functions by their values on an  $n \times n$  grid, then  $M$  is an  $n^2 \times n^2$  matrix. We label the rows and columns by points in an integer  $n \times n$  lattice.

Coverings of the plane give *CR*-coverings for the matrix  $M$ . Figure 10 shows a close-up of one possibility, a regular grid of equal overlapping squares. The covering used for the columns is shown on the left, the central square highlighted with a bold outline. On the right is the thickened covering (with larger squares) used for the rows.

In these coverings every point belongs to one, two or four squares (as indicated by the grey shading), so the nerve  $\mathcal{N}$  is 3-dimensional. A modification of this covering will give a 2-dimensional nerve, though that confers no particular advantage.

If the squares for the column covering have edge  $d$  and overlap  $r$ , then the squares for the row covering will have edge  $d + 2\ell$  and overlap  $r + 2\ell$ . We assume that  $2r + 2\ell < d$  so that the overlaps are exactly as shown. The space  $V^4$  is spanned by basis vectors in the small dark squares, so it breaks into pieces of dimension  $r^2$  approximately, one piece for each square. Similarly,  $W^4$  is a direct sum of pieces of dimension  $(r + 2\ell)^2$ . Thus the top left-hand block  $M_{44}$  in the upper-triangular form splits into sub-blocks of size  $(r + 2\ell)^2 \times r^2$ .

In the same way, we can determine the sizes of the other sub-blocks. Block  $M_{33}$  is empty, but block  $M_{22}$  splits into subblocks of size  $(r + 2\ell)(d - 2r - 2\ell) \times r(d - 2r)$ , and block  $M_{11}$  has sub-blocks of size  $(d - 2r - 2\ell)^2 \times (d - 2r)^2$ . These estimates are relevant to any detailed performance analysis.

#### 4.4. Load balancing

It is important to choose the covering of  $X$  carefully, because the performance of the algorithm depends on it. It is unlikely that there is a general answer to the question of finding the ‘best’ covering: special account of the geometry of  $X$  needs to be taken in each case. However, certain general points can be made. We illustrate this by considering

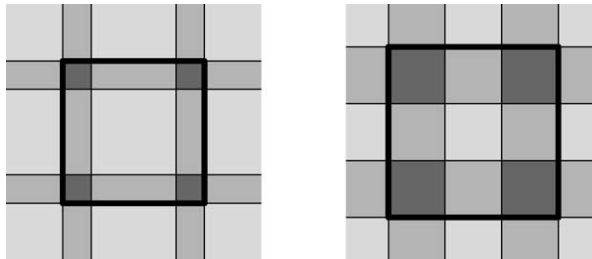


Fig. 10. Covering the plane with squares.

a family of coverings for a cyclically banded matrix (so  $X = S^1$ ). The analysis is intended to be suggestive rather than exhaustive; detailed results belong to a future paper.

Suppose that  $M$  is an  $n \times n$  matrix, and that it is cyclically banded with width  $2\ell + 1$ . The rows and columns of  $M$  are labelled by points on a circle of circumference  $n$ . We can cover the circle by  $k$  regularly-spaced intervals of equal length  $(n/k) + r$ . This gives a column covering; the corresponding row covering has intervals of length  $(n/k) + r + 2\ell$ .

If we suppose that  $r + 2\ell < (n/k)$  then no row or column can belong to more than two of the covering sets, so we have a  $2 \times 2$  block-triangular form. It is easy to check that  $M_{22}$  has  $k$  sub-blocks of size  $(r + 2\ell) \times r$ , and  $M_{11}$  has  $k$  sub-blocks of size  $((n/k) - r - 2\ell) \times ((n/k) - r)$ . There are  $2k$  sub-blocks in  $M_{21}$ , each of size  $(r + 2\ell) \times ((n/k) - r)$ .

We can now begin the reduction process. Each sub-block in  $M_{22}$  can be reduced independently, with the calculation carried over to two of the sub-blocks in  $M_{21}$ . This gives  $k$  independent costs of

$$C_{22} = (r + 2\ell)^2(r + 2((n/k) - r))$$

operations. Meanwhile, the  $k$  sub-blocks of  $M_{11}$  are reduced independently, each at a cost of

$$C_{11} = ((n/k) - r - 2\ell)^2((n/k) - r).$$

That completes the first iteration.

The complexity of the second iteration depends on the number of pivots found so far: if there are many pivots, then less of the matrix remains to be reduced.

**Hypothesis 20** (*local full-rank hypothesis*). At each stage during the reduction, the blocks being reduced have full rank.

This hypothesis leads to best-case estimates for a given matrix reduction; without it, the complexity estimate quickly becomes intractable. It is not unreasonable as an approximation to the true behaviour, since a generic matrix has full rank if its sparse structure has full rank. In the conclusion to this paper, we discuss the local full-rank hypothesis further.

The second iteration begins by using the pivots of  $M_{11}$  to clear parts of  $M_{21}$ . Assuming the local full-rank hypothesis, we can estimate this cost to be  $k$  independent instances of:

$$D_{21} = 2(2\ell)^2((n/k) - 2r - 2\ell).$$

Finally, we reduce the remaining  $2k\ell \times 2k\ell$  part of the matrix. This requires

$$C_{21} = (2k\ell)^3$$

operations.

If we have  $k$  separate processors, then the total time cost is  $C_{22} + C_{11} + D_{21} + C_{21}$ ; with one processor the time cost is  $k(C_{22} + C_{11} + D_{21}) + C_{21}$ . To first approximation

we assume that  $\ell \ll (n/k)$ ; for otherwise the  $C_{21}$  term is comparable to  $n^3$ , dominating the calculation. If we make that assumption then  $C_{22}$  and  $C_{11}$  are approximated by cubic functions in  $r$  and  $(n/k)$ . One can calculate that  $C_{22} + C_{11}$  is minimised when  $r = c_1(n/k)$ , where  $c_1 \approx 0.39$ . Then  $k(C_{22} + C_{11}) = c_2 n^3/k^2$ , where  $c_2 \approx 0.47$ .

We can now choose  $k$  by balancing the term  $k(C_{22} + C_{11})$  which improves quadratically in  $k$ , with the term  $C_{21}$  which gets cubically worse. (The  $D_{21}$  term is dominated.)

Although this is only a rough calculation, the choice of  $k$  will typically have this flavor. Too small, and the first-iteration sub-blocks are large; too large, and it leaves much left to be done in the second-iteration. We also point out that the optimal value of  $r$  gives sub-blocks in  $M_{22}$  and  $M_{11}$  which are about the same size. This is a load-balancing phenomenon, and entirely to be expected.

Finally, we have only covered the  $2 \times 2$  case, imposed by the assumption  $r + 2\ell < (n/k)$ . If  $r$  is larger, then the covering intervals overlap with greater multiplicity, and there are more blocks and more iterations. A full analysis would include all these possibilities.

## 5. The Mayer–Vietoris spectral sequence

In this section we describe a different procedure inspired by the Mayer–Vietoris spectral sequence in algebraic topology. The spectral sequence is an iterative algorithm for determining the homology groups  $H_*(X)$  of a space  $X$ , given a covering  $X = U_1 \cup \dots \cup U_p$ , in terms of the homology groups  $H_*(U_{i_0} \cap \dots \cap U_{i_k})$  of finite intersections of the  $U_i$ .

We begin with a lemma in homological algebra. Recall that an “exact sequence” is one where the image of one map is equal to the kernel of the next one.

**Lemma 21.** *Let  $V$  be a vector space decomposed as described after Definition 14. Then there are exact sequences*

$$\dots \longleftarrow V^{(1)} \longleftarrow V^{(0)} \longleftarrow V \longleftarrow 0$$

and dually

$$\dots \longrightarrow V^{(1)} \longrightarrow V^{(0)} \longrightarrow V \longrightarrow 0$$

where  $V^{(k)} = \bigoplus_{\dim(\sigma)=k} V_\sigma$ .

Contrast the definition of  $V^{(k)}$  with the definition of  $V^k$  in Proposition 17.

**Proof.** The horizontal maps  $V^{(k)} \leftarrow V^{(k-1)}$  are defined this way: if  $\sigma = \llbracket i_0 \dots i_k \rrbracket$  is a  $k$ -simplex and  $\tau = \llbracket i_0 \dots \hat{i}_d \dots i_k \rrbracket$  is a face, then the component  $V_\sigma \leftarrow V_\tau$  is the map  $(-1)^d P_{\tau\sigma}$ ; if  $\tau$  is any other  $(k-1)$ -simplex, then the component  $V_\sigma \leftarrow V_\tau$  is zero.

The rest is standard homological algebra. The result is well illustrated by considering what happens at  $V^{(0)}$ . Note that an element of  $V^{(0)}$  is specified by a sequence of vectors  $v_i \in V_i$ . The map to  $V^{(1)}$  computes the difference between the projections of  $v_i$  and  $v_j$  on the common subspace  $V_{ij}$ , for each pair  $i, j$ . If the  $v_i$  are actually projections of a

vector  $v \in V$ , then these differences are necessarily zero. Conversely, if the differences are all zero, then one can write down a vector  $v$  which projects to  $v_i$  on each  $V_i$ , since the ‘overlaps’ agree. The general proof is a more complicated version of this; we omit the details.

The second sequence is dual to the first one, so that result follows immediately from the first. The horizontal maps are defined in terms of the inclusions  $I_{\sigma\tau}$  (recall that  $I_{\sigma\tau}$  is the adjoint of  $P_{\tau\sigma}$ ).  $\square$

**Lemma 22.** *Let  $M$  be a matrix with an RC-covering. Then there is a commutative diagram of exact sequences:*

$$\begin{array}{ccccccc} \cdots & \leftarrow & W^{(1)} & \leftarrow & W^{(0)} & \leftarrow & W & \leftarrow & 0 \\ & & \uparrow & & \uparrow & & \uparrow & & \\ \cdots & \leftarrow & V^{(1)} & \leftarrow & V^{(0)} & \leftarrow & V & \leftarrow & 0 \end{array}$$

where the vertical maps are the direct sums  $M^{(k)} = \bigoplus_{\sigma} M_{\sigma}$ .

**Proof.** Given Lemma 21, it remains to check that the diagram commutes. Let  $\sigma = \llbracket i_0 \dots i_k \rrbracket$  be a  $k$ -simplex and let  $\tau = \llbracket i_0 \dots \hat{i}_d \dots i_k \rrbracket$  be a face. We must check that  $M_{\sigma} \circ (-1)^d P_{\tau\sigma} = (-1)^d P_{\tau\sigma} \circ M_{\tau}$ . This is given by Lemma 15.  $\square$

Since the diagram in Lemma 22 commutes, it follows that the horizontal map  $V^{(0)} \rightarrow V^{(1)}$  restricts to a map  $\text{Ker}(M^{(0)}) \rightarrow \text{Ker}(M^{(1)})$ . It is then a short step to the following theorem.

**Theorem 23.** *There is a natural isomorphism*

$$\text{Ker}(M) \cong \text{Ker}[\text{Ker}(M^{(0)}) \rightarrow \text{Ker}(M^{(1)})]$$

induced by the horizontal map  $V \rightarrow V^{(0)}$ .

This gives us an algorithm for finding the kernel of a matrix. Let  $M$  be a (geometrically sparse) matrix with an RC-covering and nerve complex  $\mathcal{N}$ .

We will solve the equation

$$M\vec{x} = \vec{0}$$

in three steps. The first two steps are local in character, involving very many small, independent calculations which can be carried out in parallel. The inevitable global calculation which follows in step (3) cannot be made parallel, but is substantially simpler than the original problem.

- (1) For each vertex of  $\mathcal{N}$ , that is for  $i = 1, \dots, p$ , find a basis of solutions to the equation  $M_i \vec{x}_i = \vec{0}$ .

- (2) For each 1-simplex  $[[ij]] \in N$  find a basis of solutions to the equation  $M_{ij}\vec{x}_{ij} = \vec{0}$ .  
 (3) Determine the kernel of the natural map  $\text{Ker}(M^{(0)}) \rightarrow \text{Ker}(M^{(1)})$ . (This map can be expressed as a matrix in terms of the bases found in steps (1) and (2)).

Steps (1) and (2) may be carried out in parallel highly efficiently, given good row and column coverings. The bottleneck is the third stage: although the matrix in step (3) is usually significantly smaller than  $M$  itself, it may still be quite large.

The method allows us to compute the kernel of a matrix with an  $RC$ -covering, or the cokernel of a matrix with a  $CR$ -covering, by transposing. However:

**Remark 24.** If  $M$  is given a  $CR$ -covering, then the correct diagram is:

$$\begin{array}{ccccccc} \cdots & \longrightarrow & V^{(1)} & \longrightarrow & V^{(0)} & \longrightarrow & V \longrightarrow 0 \\ & & \downarrow & & \downarrow & & \downarrow \\ \cdots & \longrightarrow & W^{(1)} & \longrightarrow & W^{(0)} & \longrightarrow & W \longrightarrow 0 \end{array}$$

with the same vertical maps  $M^{(k)}$ .

In this case there still is a Mayer–Vietoris calculation for  $\text{Ker}(M)$ , but it is much more complicated. In effect, we are computing  $H_1$  of a homology spectral sequence; earlier we were computing  $H^0$  of a cohomology spectral sequence. Spectral sequence computations in dimension 1 typically require many more algebraic steps than in dimension 0.

## 6. Conclusions

It seems clear to the authors that there is much scope for exploiting geometric sparseness in matrix computations. In particular, it is natural to look for new local–global algorithms. There are typically two stages to these algorithms: first carry out a large number of independent, small local calculations, then finish with a global calculation. The global calculation should be a much simplified version of the original problem; the simplification being made possible by the local calculations. The structure of the calculation will be regulated by the geometry of the parameter space  $X$ , and of the nerves  $\mathcal{N}$  of suitable coverings.

The Mayer–Vietoris algorithm follows this pattern exactly. In the filtration spectral sequence, the local stage takes place in the sub-block computations of the first iteration. The remaining iterations constitute the global assembly process. There may be ways of taking account of the sub-block structure even after the first iteration; if so, then the local–global distinction becomes blurred.

The local full-rank hypothesis of Section 4.4 is analogous to the concept of ellipticity in partial differential equations. An operator is elliptic if it is ‘more or less’ locally invertible. For example the operator  $d/dx$  has ‘inverse’  $\int^x \cdot dx + c$ , with an undetermined additive constant  $c$ . The space of solutions to an elliptic differential equation typically measures *discrete* global information: the existence of local solutions carries essentially

no information (since the operator is locally invertible), so the interest resides in the construction of global solutions. The de Rham theorem in differential geometry [5] is a classic example of this. In the proof, a potentially infinite-dimensional problem is broken down into an explicit local calculation (the Poincaré lemma) together with a combinatorial patchwork process for assembling global results from local ones.

Without ellipticity, there is no hope of establishing any such result. In the case of our spectral sequence algorithm, the analogous task is to replace a big global calculation with a small global calculation, by first doing lots of small local calculations. As we saw, best performance occurs when the matrix is locally full-rank, which is when the local reduction process has the greatest power. Without the assumption, the matrix reduction might proceed extremely slowly; it is not hard to see that this can happen in practice.

We hope that other authors will take up the challenge of finding new local–global algorithms in the same spirit.

### Acknowledgments

This work was supported by NSF grant DMS-0101364. A preliminary version of the filtration spectral sequence algorithm was implemented by Erik Carlsson and Peter Lee during summer 2001. We would also like to thank Lek-Heng Lim for many useful discussions. He is also collaborating with us on a detailed analysis of the performance and stability of the algorithm. The results will appear in a forthcoming paper. We also wish to thank Persi Diaconis, Susan Holmes, Marco Grandis, and Joshua Tenenbaum, for various discussions and comments.

### References

- [1] G. Bredon, *Topology and Geometry*, Springer-Verlag, New York, 1993.
- [2] I.S. Duff, A.M. Erisman, J.K. Reid, *Direct Methods for Sparse Matrices*, Oxford Univ. Press, Oxford, 1986.
- [3] A.L. Dulmage, N.S. Mendelsohn, Two algorithms for bipartite graphs, *J. Soc. Indust. Appl. Math.* 11 (1963) 183–194.
- [4] G. Golub, C. Van Loan, *Matrix Computations*, third ed., Johns Hopkins Univ. Press, Baltimore, 1996.
- [5] W.S. Massey, *A Basic Course in Algebraic Topology*, Springer-Verlag, New York, 1991.
- [6] J. McCleary, *A User's Guide to Spectral Sequences*, Publish or Perish, Wilmington, DE, 1985.
- [7] A. Pothen, C.-J. Fan, Computing the block triangular form of a sparse matrix, *ACM Trans. Math. Softw.* 16 (1990) 303–324.
- [8] E.C. Zeeman, On the filtered differential group, *Ann. of Math.* (2) 66 (1957) 557–585.