In: *Signal Processing for Multimedia, p. 35-54*
J. S. Byrnes (Ed.)
IOS Press, 1999

Mathematics for Multimedia Signal Processing II Discrete Finite Frames and Signal Reconstruction

Paulo J. S. G. Ferreira*
Universidade de Aveiro

Abstract. Certain signal reconstruction problems can be understood in terms of frames and redundant representations. The redundancy is useful because it leads to robust signal representations, that is, representations in which partial loss of data can be tolerated without misbehavior or adverse effects. This chapter begins by presenting a few engineering problems in which robust data representations play a central role. It turns out that these problems, which occur in signal processing, spectrum analysis, information theory, and fault-tolerant computing, are closely related or even equivalent. However, perhaps surprisingly, the connections between them have gone nearly unnoticed so far. Frames, and in particular discrete finite frames, provide one of the ways of understanding certain of these problems, including the important missing data problem. Some of the methods that can be used to recover from missing data errors are examined, emphasizing finite-dimensional theory because of its simplicity and practical usefulness, and interpreting the results in terms of discrete finite frames. The connection between the frame algorithm and a few other iterative reconstruction methods, such as POCS and the Papoulis-Gerchberg iteration, is detailed.

1 Introduction

Redundant data representations are useful in a variety of contexts. Their theoretical interest can hardly be denied — it is enough to consider the concept of "frame" and its role in mathematics and engineering. On the other hand, redundancy usually leads to robustness which, in turn, suggests several applications. Loosely speaking, a signal representation is robust when partial loss of data does not lead to misbehavior or severe adverse effects.

This chapter discusses frames and their usefulness in connection to certain problems that arise in interpolation, spectrum analysis, error-control coding, and fault-tolerant computing. We believe that the relations between these problems have gone nearly unnoticed so far. Some of the problems are shown to be equivalent in the following sense: if one of them can be solved using a certain algorithm, so can the other, using essentially the same algorithm. The specific problems that will be considered are (i) the band-limited missing data problem (ii) a nonlinear interpolation problem (iii) the problem of estimating a signal that is the superposition of a finite number of harmonics (iv) an error-control coding problem, formulated in the real field, and (v) certain techniques that occur in algorithm-based fault tolerant computing.

The advantages of studying the relations among these problems are clear. The techniques commonly used in one field can be imported to the others, the duplication

*The research that lead to this work was supported by the JNICT / FCT.

of research efforts is prevented, the overall degree of understanding of the problems increases, and new algorithms emerge as a result.

After discussing the above mentioned problems we will introduce the concept of frame. Frames were introduced by Duffin and Schaeffer [6] in 1952, in reference to nonharmonic Fourier series. The importance of the ideas underlying that paper was recognized by the scientific community, specially after the early 80's, when frames started what Daubechies called their "second career" [4].

Frames became one of the fundamental concepts in time-frequency and time-scale analysis and wavelet series. They are the central topic of many papers, and a variety of books carry discussions of frame theory. A clear introduction can be found, for example, in [48] or [3]. The recent issue of *The Journal of Fourier Analysis and Applications* dedicated to Richard J. Duffin offers many further examples of the mathematical importance of frames. Their significance in the context of signal analysis and engineering applications is also widely recognized.

Our motivation for introducing frames is the following: discrete finite frames provide one of the ways of approaching the missing data problem mentioned above. And the solution to this problem can be put to use when studying the other problems as well.

We will attempt to show how frames can be used to deal with missing data, emphasizing finite-dimensional theory because of its simplicity and practical usefulness. Robust signal representations are important for signal analysis and processing applications, and frames provide redundant representations. Redundancy is a necessary requirement for robustness, in the sense of information theory: if a data stream has no redundancy, error detection (not to mention error correction!) is impossible. Adding redundant data makes error detection possible (consider adding parity bits to a bit stream). Adding even more redundancy may allow for both error detection and correction.

This is often accomplished by error control coding in Galois (finite) fields. We will not follow this path: instead, we will work in the complex or real fields. This has advantages and disadvantages. On one hand, the block length restrictions that arise when dealing with finite fields can be circumvented (for example, the discrete Fourier transform and many other unitary transforms can be defined in $\mathbb{R}^N$, that is, for signals of N samples, for any $N > 1$). On the other hand, arithmetic in the real and complex fields is prone to round-off error: exact real arithmetic is technically impossible in practice.

Nevertheless, the approach used will hopefully bridge the gap between frames, signal reconstruction, and the problems mentioned initially. To emphasize the multiple connections between these topics we will consider a few of the iterative methods available for interpolating and extrapolating signals, which were used for signal reconstruction purposes before frames became popular alternatives: the Papoulis-Gerchberg iteration, alternating projections, projections onto convex sets (POCS), and more.

We will clarify the relation between the frame algorithm and these other methods, and show how the frame bounds relate to their numerical performance or convergence rate. At the end, we will hopefully have a better grasp of the mathematical foundations of finite-dimensional signal reconstruction and a few other seemingly unrelated problems.

2 Notation and preliminaries

2.1 The discrete Fourier transform

The Fourier matrix F is the $N \times N$ matrix with elements

$$F_{ij} = \frac{1}{\sqrt{N}} e^{-\mathtt{i}\frac{2\pi}{N} ij}.$$

The symbol $\mathtt{i}$ denotes the imaginary unit, not to be confused with i or j, which denote integers. The Fourier matrix is unitary, that is, its inverse is simply the Hermitian transpose:

$$F^H F = F F^H = I.$$

The discrete Fourier transform (DFT) of x is denoted by $\hat{x}$, and is defined by

$$\hat{x} \triangleq Fx.$$

Computing the discrete Fourier transform $\hat{x} = Fx$ seems to require $O(N^2)$ arithmetic operations, as any general matrix-vector product in $\mathbb{C}^N$. However, F is a very special matrix and the computation of the DFT (equivalently, of the matrix product $\hat{x} = Fx$) can be sped up considerably. The number of necessary arithmetic operations (flops) can often be made as low as $O(N \log_2 N)$, using the fast Fourier transform (FFT).

We say "often" because the computational load depends on the structure of the number N. When N is a highly composite number, such as a power of two, the computations are easier. For details and applications of the FFT, see, for example, [2, 42].

2.2 Circulants

A $N \times N$ matrix C is circulant if it has the following structure:

$$C = \begin{bmatrix} c_0 & c_1 & \cdots & c_{N-1} \\ c_{N-1} & c_0 & \cdots & c_{N-2} \\ c_{N-2} & c_{N-1} & \cdots & c_{N-3} \\ c_{N-3} & c_{N-2} & \cdots & c_{N-4} \\ \vdots & \vdots & \ddots & \vdots \\ c_1 & c_2 & \cdots & c_0 \end{bmatrix}.$$

Each row of C can be obtained from the preceding row by shifting its elements one position to the right and "wrapping around". This is called a circular shift, and the matrix-vector multiplication Cx is called circular convolution.

All circulants commute, a consequence of the decomposition

$$C = F^H \Lambda F$$

where Λ is diagonal. For our purposes, this is the most important fact concerning circulants. It asserts that the Fourier matrix F diagonalizes all circulants.

The eigenvalues of a circulant can be found rather easily, often in $O(N \log_2 N)$ flops when using the FFT algorithm, because they are determined by the N elements of the DFT of one of the rows or columns of the circulant. The FFT is also useful for computing circular convolutions, because

$$Cx = F^H \Lambda F x = F^H \Lambda \hat{x}.$$

The DFT vector $\hat{x}$ can often be computed in $O(N \log_2 N)$ flops, as well as the multiplication by F^H (the inversion of the DFT). The multiplication by the diagonal matrix Λ is a $O(N)$-flop task, and the conclusion is that circular convolution in $\mathbb{C}^N$ can be evaluated in $O(N \log_2 N)$ flops.

Circulants are a remarkably simple but interesting class of matrices. Many matrix problems, when formulated in terms of circulants, admit closed form solutions. For details concerning circulant matrices see [5].

2.3 A class of circulant matrices

We will need to consider the expansion of signals x of N samples, that is, vectors in $\mathbb{R}^N$ or $\mathbb{C}^N$, in terms of certain non-orthogonal vectors. To introduce these vectors, consider the $N \times N$ circulant matrix B given by

$$B \triangleq F^H \Lambda F,$$

where Λ is a $N \times N$ diagonal matrix. It is assumed that $\Lambda_{ii} = 1$ or $\Lambda_{ii} = 0$, that is, Λ is a zero-one diagonal matrix (and consequently $\Lambda^2 = \Lambda$). To avoid trivialities, it is preferable to exclude the possibility of having $\Lambda = 0$, the zero matrix, or $\Lambda = I$, the identity matrix. The matrix B is Hermitian, and satisfies

$$B^2 = B$$

because $\Lambda^2 = \Lambda$. Consider the set of vectors that satisfy the equation

$$x = Bx.$$

It is easily seen that this set is not empty and that it forms a subspace of $\mathbb{R}^N$.

The linear operator $x \to Bx$ can be regarded as a linear system or filter. Its action upon a signal x can be easily understood in terms of Λ and the Fourier transform $\hat{x}$ of x. The samples $\hat{x}_i$ such that $\Lambda_{ii} = 1$ are preserved by the system. The remaining samples $\hat{x}_j$, for which $\Lambda_{jj} = 0$, are set to zero.

By definition, the pass-band of B is the set P of integers for which $\Lambda_{ii} = 1$. The signals x that satisfy $x = Bx$ are those whose Fourier transform is supported in the pass-band of B. We will refer to them as band-limited signals.

The decomposition $B = F^H \Lambda F$ means that

$$\begin{aligned} B_{ij} &= \sum_{k=0}^{N-1} F_{ik}^H \Lambda_{kk} F_{kj} \\ &= \sum_{k \in P} F_{ik}^H F_{kj}, \end{aligned}$$

where P is the pass-band of B ($\Lambda_{kk} = 0$ for any integer $0 \leq k < N$ that does not belong to P). This shows that B can be written as $B = E^H E$, where the columns of the matrix E are the columns F_i of the Fourier matrix with $i \in P$.

The equation $x = Bx$ can also be written in terms of the columns $\{B_i\}_{0 \leq i < N}$ of B,

$$x = \sum_{i=0}^{N-1} x_i B_i.$$

B is idempotent, and $B^2 = B$ implies

$$\langle B_i, B_j \rangle = B_{ij},$$

and if x satisfies $x = Bx$ then

$$\langle x, B_i \rangle = x_i.$$

2.4 An example: low-pass signals

Let the main diagonal of Λ be given by

$$[1, \underbrace{1, \ldots, 1}_{M \text{ times}}, 0, \ldots, 0, \underbrace{1, \ldots, 1}_{M \text{ times}}].$$

Since $B = F^H \Lambda F$, this immediately translates into

$$\begin{aligned} B_{ij} &= \frac{1}{N} \sum_{k=-M}^{M} e^{-\mathrm{i}\frac{2\pi}{N}(i-j)k} \\ &= \frac{\sin[\pi(2M+1)(i-j)/N]}{N \sin[\pi(i-j)/N]}. \end{aligned}$$

In terms of the samples of x, $x = Bx$ becomes

$$x_i = \sum_{j=0}^{N-1} x_j \frac{\sin[\pi(2M+1)(i-j)/N]}{N \sin[\pi(i-j)/N]}.$$

In this case, the signals such that $x = Bx$ are called low-pass signals and B is a low-pass filter matrix. Note the similarity between the previous equation and

$$f(t) = \int_{-\infty}^{+\infty} f(x) \frac{\sin w(t-x)}{\pi(t-x)} \, dx$$

which is satisfied by any $f \in L_2(\mathbb{R})$ if and only if its Fourier transform vanishes almost everywhere outside $[-w, w]$. A signal belonging to $L_2(\mathbb{R})$ is band-limited if its Fourier transform vanishes outside a compact set. When the support is an interval of the form $[-w, w]$ the signal is called low-pass. Band-pass signals are characterized by a Fourier transform that vanishes outside $[-b, -a] \cup [a, b]$.

3 Robust data representations: some related problems

The missing data problem is the signal reconstruction problem around which this chapter turns. To relate it to the interpolation, extrapolation or prediction problems it is only necessary to consider specific distributions of the missing data (in the extrapolation problem the known data are contiguous, in the prediction problem the past of the signal is known). These problems are widely known in signal processing and in other fields (consider superresolution in optics, for example). The most common constraint imposed upon the signal is that of band-limitedness.

Problem 1 (missing data) *To determine a subset of the samples of a band-limited signal.*

Note that we are referring to finite-dimensional signals, and therefore assertions such as "x is band-limited" mean "there exists a low-pass circulant matrix B belonging to the class mentioned in section 2.3 and such that $x = Bx$". This will be tacitly assumed from now on.

One of the methods for solving this problem is the Papoulis-Gerchberg algorithm, in which the known time and frequency domain constraints are applied iteratively. Let J be the set of known samples, and let D be the diagonal matrix

$$D_{ij} \triangleq \begin{cases} 1, & \text{if } i = j \text{ and } i \in J, \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

The initialization step of the algorithm sets the unknown samples to some initial value. This yields the first approximation $x^{(0)}$. The first half of the nth step of the algorithm consists in filtering the result of the previous step, that is,

$$y = Bx^{(n-1)}. \tag{2}$$

This imposes the frequency-domain constraints (the result y is band-limited). But this filtering changes the values of the known samples. Therefore y can be improved simply by setting the samples y_i, $i \in J$, back to their known values. It can be readily verified that, given an arbitrary $a, b \in \mathbb{C}^N$, the operator P defined by

$$Pa \triangleq (I - D)a + Db \tag{3}$$

can be used to enforce these time-domain constraints. It is enough to note that $(Pa)_i$ and b_i will agree for all $i \in J$, a fact that follows at once from the definition of D in (1). Applying this principle to y leads to

$$x^{(n)} = (I - D)y + Dx^{(0)}. \tag{4}$$

This sets the samples y_i, $i \in J$, to their initial and correct values, completing the second half of step n. The two operations (2) and (4) can readily be combined in one equation,

$$x^{(n)} = (I - D)Bx^{(n-1)} + Dx^{(0)},$$

which describes the nth step of the algorithm. The process can be repeated until a satisfactory result is found.

For more concerning the Papoulis-Gerchberg iteration and similar methods, see, for example, [8, 22, 25, 26]. These methods and several others can be applied only if the positions of the unknown samples are known. In practice, this is not always the case, and the following problem arises.

Problem 2 *To simultaneously determine the number, positions and correct values of the incorrect samples of a band-limited signal corrupted by impulsive noise.*

This is a nonlinear problem. It can be approached in two steps: in the first step the number and positions of the incorrect samples are estimated. As soon as the positions of the incorrect samples have been estimated, the problem reduces to the missing data problem.

How can the corrupted samples be detected? One solution is as follows. Let

$$U = \{i_0, i_1, \dots i_{n-1}\}$$

denote the positions of the n incorrect samples of a band-limited signal $x \in \mathbb{R}^N$. Let y denote the observed signal, which coincides with x except for the samples whose indexes belong to U, and let $e = x - y$ be the error signal. It is convenient to denote the kth sample of e by $e(k)$. Observe that $e(k) = 0$ for all $k \notin U$. In practice, the cardinal of U is usually much smaller than N, that is, e is sparse. Consider the polynomial

$$P(z) = \sum_{i=0}^{n} h_i z^i,$$

defined by $h_n = 1$ and

$$P\left(e^{-\mathrm{i}\frac{2\pi}{N}i_m}\right) = 0, \tag{5}$$

for $m = 0, 1, \ldots, n-1$. Let

$$A = \{j_0, j_1, j_2, \ldots\}$$

be a set of distinct integers. Multiplying (5) by $e(i_m)e^{\mathrm{i}\frac{2\pi}{N}i_m j_\ell}$ and summing leads to

$$\sum_{k=0}^{n} h_k \sum_{m=0}^{n-1} e(i_m)e^{-\mathrm{i}\frac{2\pi}{N}i_m(k-j_\ell)} = 0,$$

which is equivalent to

$$\sum_{k=0}^{n} h_k\, \hat{e}(k - j_\ell) = 0,$$

where $\hat{e} = Fe$ is the DFT of e. Using the fact that $h_n = 1$, one obtains

$$\sum_{k=0}^{n-1} h_k\, \hat{e}(k - j_\ell) = -\hat{e}(n - j_\ell),$$

and proper choice of the j_ℓ lead to a set of linear equations $Th = b$ for the coefficients h_k of the polynomial P. The zeros of the polynomial, and consequently the position and the number of errors, can be easily determined using the FFT algorithm. For details, see [14, 44].

It is time to consider an apparently distinct problem: fault tolerant computing. Consider a parallel computer, perhaps composed of many individual processing units, each having a certain probability of failure. It is of course desirable that a failure in one of the processing units does not bring the whole system to a halt or crash. Instead, one would like to be able to continue the processing using only the operating units, and perhaps shutting down those that have failed. To do this it is clearly necessary to detect the fault, and then, if possible, to correct the data. One of the approaches to fault detection is algorithm-based, that is, the faults are detected by algorithms, or software, often using the known properties of the data being processed.

For example, denote by x a vector in $\mathbb{R}^N$, with N large. Assume that the DFT of x is being computed, possibly with the help of several processing units, and let y be the result of the computation. If y is indeed the DFT of x, then $\|x\| = \|y\|$, a condition that is easy to check. Similar norm checks can be made for example after completing each stage of the fast Fourier transform algorithm. They of course depend on the type of computation being performed, and cannot be applied without changes, say, when inverting a matrix.

However, there are also general methods for checking the accuracy of a result that are to a large extent independent of the type of data manipulation being carried out. Checksums provide the most obvious example.

Work along these lines has been carried out by several authors [1, 20, 28, 29]. The checksums can be computed modulo an integer (a parity bit is a checksum modulo two). In the real field, the checksums can be replaced by the average value of the data. Weighted checksums have also been used to extend the error-correcting capabilities. The weighted checksum problem is similar to the following general problem.

Problem 3 *Given a possibly corrupted subset of data samples, and a subset of the samples of a discrete orthonormal transform of the data (such as the DFT), determine the data.*

Note that the first sample of the DFT of a data vector $x \in \mathbb{R}^N$

$$\hat{x}_0 = \sum_{k=0}^{N-1} x_k$$

is proportional to the average of the data (the checksum). Any other DFT samples that might also be known constitute weighted checksums, that may help in determining the data after the occurrence of errors. Clearly, transforms other than the DFT might also be used for this purpose.

The next problem is a widely known spectrum analysis problem.

Problem 4 *Given a subset of the samples of a signal, which is known to consist of a linear combination of harmonics of unknown frequencies and amplitudes,*

$$x(t) = \sum_{k=1}^{r} a_k \, e^{-\mathrm{i}2\pi f_k t}$$

determine the signal.

This problem is very well-studied, and many methods (parametric and nonparametric) have been proposed to solve it. A comprehensive review of the spectrum analysis techniques up to 1981 can be found in [23].

We mention the method of Papoulis and Chamzas [32], for example, which is a modified Papoulis-Gerchberg iteration in which the filtering step is replaced by a non-linear operation: Fourier transforming, thresholding the spectrum, and inverse Fourier transforming.

The last problem that we consider is the following.

Problem 5 *The error-control problem in the real field: devise a coding procedure capable of locating and correcting up to a certain number of errors in a finite-length block of real numbers.*

One of the techniques that can be used is as follows. An initial block of k data words is padded with $n - k$ zeros, and an IDFT of n samples is taken. The n words are then transmitted, and at the receiving end a DFT of length n is computed. The $n - k$ words that were zero-padded (the syndrome) provide a window over the spectrum of the error signal. The problem is how to use this information to recover the error signal itself.

3.1 Connections between the problems

Some of the problems described above turn out to be closely related or even equivalent. Surprisingly, these connections do not seem to be widely recognized. An exception is perhaps problem 2 and 5 — see [27] and references therein.

Problem 2 and 3 are clearly equivalent. Let x be the data vector, and denote by $V = \{i_1, i_2, \ldots, i_p\}$ the indexes of the known samples of its DFT $\hat{x}$. In problem 3 we are given the possibly corrupted data vector x, and the set $\{\hat{x}(i_1), \hat{x}(i_2), \ldots, \hat{x}(i_p)\}$, whereas in problem 2 we are given just the corrupted data vector x. However, the hypothesis of x being band-limited implies that a known subset of the samples of $\hat{x}$ has zero value. This set plays the role of $\hat{x}(i_1), \hat{x}(i_2), \ldots, \hat{x}(i_p)$ in problem 3. The only difference is that, in the band-limited case, the known DFT samples are zero: $\hat{x}(i_k) = 0$ for all $i_k \in V$.

Problem 2 and 4 are the dual of each other (the term "dual" is used here in the sense explained in [11,12]). To understand why, denote by $U = \{i_1, i_2, \ldots i_r\}$ the positions of r incorrect samples of a band-limited signal x with a total of n samples. Let e be the error signal $e = x - y$, where y is the observed signal, which coincides with x except for the samples whose indexes belong to U. Thus, $e_k = 0$ for all $k \notin U$. Typically, r, the cardinal of U, is much less than n, that is, the error vector e is sparse.

Let x be band-limited, with p zero harmonics. Then, the DFT of y contains exactly p samples of the DFT of e. For example, if x is low-pass with $2m+1$ nonzero harmonics, then the samples $m+1$ through $n-m-1$ of y are equal to the corresponding samples of e. But then problem 2 can be rephrased as follows: given a subset of p samples of the DFT $\hat{e}$ of e, estimate e. Since $\hat{e}$ is given by

$$\hat{e}(i) = \sum_{k=1}^{r} e(i_k)\, e^{-\mathrm{i}\frac{2\pi}{n} i_k i},$$

this shows that the problem is equivalent to the problem 4 (set $t = i_k/n$), if the time and frequency domains are interchanged. We say that the two problems are the dual of each other.

Problem 2 and 5 are also closely related. The connection between problem 2 and certain topics in information theory (error-control codes in the real field) has been noted before, but does not seem to be widely known in the signal processing community. Specific algorithms have been suggested to solve this problem; we refer to the method described in [43], which is able to correct a single error, and to [14, 27, 44], for the correction of multiple errors.

3.2 The role of frames

The previous observations have several immediate consequences. Reference [41] offers one example: an iterative method to solve problem 2, based on a discrete-discrete version of the Papoulis-Chamzas nonlinear iteration [32], originally proposed to solve a problem similar to problem 4 but in L_2. Since problem 2 is equivalent to problems 3 and 5, the same algorithm can be applied to solve any of these problems.

Obviously, understanding the connections between these problems increases the impact of any study concerning one of them upon the others. We will concentrate, from this point onwards, on a particular signal reconstruction problem — the missing data problem (problem 1). Our plan of attack is the following: we will define discrete finite frames, and the problem of obtaining estimates for the frame bounds. After this we will consider the frame algorithm, and see how it can be used to iteratively solve the missing data problem. The convergence rate of the algorithm is determined by the values of the frame bounds.

Finally, we review a few other methods through which the missing data problem can be solved. Some of these have already been mentioned, in reference to some of the problems considered. These additional links between the frame method and other methods help in completing the picture that we have attempted to outline: from reconstruction problems to frames, and then from frames back to the reconstruction problems.

4 Signal representation and discrete finite frames

4.1 Discrete finite frames

A countable subset f_n ($n \in \mathbb{Z}$) in a separable Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and norm $\| \cdot \|$ is a frame if

$$\sum_{n \in \mathbb{Z}} | \langle g, f_n \rangle |^2 \asymp \|g\|^2$$

for every element g of that Hilbert space. The notation $\asymp$ is a concise way of expressing the fact that there exist constants $0 < \alpha \leq \beta$, independent of g, and such that

$$\alpha \|g\|^2 \leq \sum_{n \in \mathbb{Z}} | \langle g, f_n \rangle |^2 \leq \beta \|g\|^2.$$

Any two such constants α and β are called the frame bounds.

Frames come in diverse flavors. For example, if $\alpha = \beta$ the frame is said to be tight. We will be concerned with discrete finite frames, that is, sets of vectors

$$\{f_1, f_2, f_3, \ldots, f_n\},$$

where each of the $\{f_k\}_{1 \leq k < n}$ belongs to the vector space $\mathbb{R}^d$, regarded as a finite-dimensional Hilbert space when endowed with the usual inner product and norm. A recent introduction to discrete finite frames can be found in [33].

4.2 Eigenvalues and singular values

The eigenvalues of a $N \times N$ Hermitian matrix are real and we adopt the convention that they are labeled according to non-decreasing value,

$$\lambda_1 \leq \lambda_2 \ldots \leq \lambda_{N-1} \leq \lambda_N.$$

The smallest (largest) eigenvalue of a Hermitian matrix A is the solution to a certain constrained minimization (maximization) problem, namely

$$\lambda_1 = \min_{x \neq 0} \frac{x^H A x}{\|x\|^2},$$

$$\lambda_N = \max_{x \neq 0} \frac{x^H A x}{\|x\|^2}.$$

The transpose of a vector x is denoted by x^T, and the Hermitian transpose by x^H. The inequalities

$$\lambda_1 \|x\|^2 \leq x^H A x \leq \lambda_N \|x\|^2$$

hold for all $x \in \mathbb{C}^N$ and are sharp. For example, $v^H A v = \lambda_1 \|v\|^2$ when v is an eigenvector of A corresponding to the eigenvalue λ_1.

The singular values of a matrix A are eigenvalues of AA^H. They are of great importance in solving the numerical rank determination problem, and in many other problems and applications.

An up-to-date general reference on matrix analysis is [19]. Specific matrix algorithms (in a programming language similar to Matlab) and the theoretical background upon which they rest are detailed in [16]. Implementations in the C programming language are given, for example, in [34].

4.3 The Fourier transform domain counterpart

We have been dealing with the samples of x. It is certainly possible to consider the samples of its Fourier transform $\hat{x}$. For this purpose it is convenient to consider the set of all solutions $\hat{x}$ to

$$\hat{x} = T\hat{x} + \hat{v},$$

the $N \times N$ matrix T being given by

$$T \triangleq F \Gamma F^H,$$

where Γ is a $N \times N$ diagonal matrix satisfying the same restrictions that were imposed upon Λ above.

The role played by the vector $\hat{v}$ needs some explanation. The set of solutions of $x = Bx$ is interesting because many signals that occur in practice are approximately band-limited, that is, there is a subset of samples of its DFT that is zero or very nearly zero. In a sense, the spectrum of such signals is partially known and the known part happens to be zero. This is seldom the case in the time-domain. Assume that a subset of the time-domain samples of a signal is known. It can hardly be expected that all the known time-domain values are zero.

Multiplying the equation $\hat{x} = T\hat{x} + \hat{v}$ by F^H leads to

$$x = \Gamma x + v,$$

which shows that $x_i = v_i$ for all i such that $\Gamma_{ii} = 0$. Note the role of $\hat{v}$ and the meaning of v, which determines the known time-domain samples.

The properties of T are similar to those of B above. T is Hermitian, and $\Gamma^2 = \Gamma$ implies

$$T^2 = T.$$

The equation $\hat{x} = T\hat{x} + \hat{v}$ can be written in terms of the columns $\{T_i\}_{0 \leq i < N}$ of T,

$$\hat{x} = \sum_{i=0}^{N-1} \hat{x}_i T_i + \hat{v}.$$

Also,

$$\langle T_i, T_j \rangle = T_{ij},$$

and if $\hat{x} = T\hat{x} + \hat{v}$ then

$$\langle \hat{x}, T_i \rangle = \hat{x}_i - \hat{v}_i.$$

The interplay between time-domain and frequency-domain constraints, and its importance for signal reconstruction, is the subject of [12].

4.4 A discrete finite frame

Consider the set of vectors

$$X \triangleq \{B_i\}_{i\in J}$$

where J is an index set, that is, a subset of

$$E_N \triangleq \{0, 1, 2, \ldots, N-1\}.$$

Can X be a frame for the vector subspace formed by the set of all x that satisfy $x = Bx$?

There is a trivial necessary condition: the index set J must have at least d elements, where d is the dimension of the subspace of solutions to $x = Bx$. For low-pass signals with $2M + 1$ nonzero harmonics (see section 2.4) the necessary condition is, therefore, $\operatorname{card} J \geq 2M + 1$.

But we need more than a necessary condition, and so let us examine the frame condition. Fix a vector $x \in \mathbb{R}^N$ that satisfies $x = Bx$. Then,

$$\sum_{i\in J} |\langle x, B_i\rangle|^2 = \sum_{i\in J} |x_i|^2.$$

The following obvious inequality

$$\sum_{i\in J} |x_i|^2 \leq \|x\|^2$$

shows that $\beta = 1$ is an upper frame bound (although not the best possible bound, as we will shortly see).

The lower bound α requires some more work. The first question is the following: can we guarantee that, apart from the zero vector, there exists no vector x such that $x = Bx$ and $x_i = 0$ for all $i \in J$?

For low-pass signals, the answer is yes, and the proof is as follows [8]. The equation $x = Bx$ is equivalent to $x = F^H \Lambda F x$, which, in turn, means that $\hat{x} = \Lambda \hat{x}$. Therefore, $\hat{x}_i = 0$ for all $i \in \{k : \Lambda_{kk} = 0\}$.

On the other hand, $Dx = 0$ means that $x_i = 0$ for all $i \in J$. Putting together the two sets of conditions leads to

$$\sum_{j\in J} F_{ij} x_j = 0, \quad i \in \{k : \Lambda_{kk} = 0\}.$$

For low-pass signals with $2M + 1$ nonzero harmonics, the condition $\operatorname{card} J \geq 2M + 1$ and the linear independence of the columns of F yield $x = 0$.

It is possible to determine the best possible upper and lower bounds and at the same time exhibit the signals for which the bounds are attained. The key to this is the following question: among all signals satisfying $x = Bx$, which particular signals render

$$E(J, x) \triangleq \frac{\sum_{i\in J} |x_i|^2}{\|x\|^2}$$

maximum or minimum? Clearly, if the maximum is attained for a certain vector u, and the minimum for some other vector v, then

$$E(J, v) \leq E(J, x) \leq E(J, u),$$

and consequently

$$\sum_{i \in J} | \langle x, B_i \rangle |^2 \asymp \|x\|^2,$$

with the frame bounds $\alpha = E(J, v)$ and $\beta = E(J, u)$.

This is really an eigenvalue problem in disguise. To confirm, recall the definition of the $N \times N$ matrix in equation (1)

$$D_{ij} \triangleq \begin{cases} 1, & \text{if } i = j \text{ and } i \in J, \\ 0, & \text{otherwise.} \end{cases}$$

Clearly,

$$E(J, x) = \frac{\|Dx\|^2}{\|x\|^2},$$

and since $x = Bx$,

$$E(J, x) = \frac{\|DBx\|^2}{\|x\|^2}.$$

Now, both B and D are Hermitian, and so

$$\|DBx\|^2 = \langle DBx, DBx \rangle = x^H B^H D^H DBx = x^H BDBx,$$

since $D^2 = D$. This shows that the maximization and minimization problems of interest are

$$\max_{\substack{x = Bx \\ x \neq 0}} E(J, x) = \max_{\substack{x = Bx \\ x \neq 0}} \frac{x^H BDBx}{\|x\|^2},$$

$$\min_{\substack{x = Bx \\ x \neq 0}} E(J, x) = \min_{\substack{x = Bx \\ x \neq 0}} \frac{x^H BDBx}{\|x\|^2},$$

But BDB is Hermitian, and therefore the solution to these problems are, respectively, the largest eigenvalue of BDB and its smallest *nonzero* eigenvalue. If we agree to denote them by λ_{max} and λ_{min}, then

$$\sum_{i \in J} | \langle x, B_i \rangle |^2 \asymp \|x\|^2,$$

with the frame bounds

$$\alpha = \lambda_{\mathsf{min}}, \quad \beta = \lambda_{\mathsf{max}}.$$

The matrix BDB is nonnegative definite (the associated quadratic form can be written $x^H B^H D^H DBx = \|DBx\|^2 \geq 0$) and therefore its eigenvalues cannot be negative. But the eigenvectors corresponding to the zero eigenvalues can not be considered here, because they do not satisfy $x = Bx$. As we have seen, when card $J \geq 2M + 1$, $x = Bx$ and $Dx = 0$ imply $x = 0$. The eigenvectors that correspond to the zero eigenvalues are "high-pass" signals that satisfy $Bx = 0$.

The eigenvalues of BDB are singular values of the matrix DB. The analysis that has been made, and the frame bounds obtained, is related to the singular value decomposition of DB. But although DB is not Hermitian, its eigenvalues are real and in fact equal to those of BDB. This can be seen as follows.

Assume that $BDBv = \lambda v$. Left multiplication by DB leads to $(DB)^2 v = \lambda DBv$, which shows that λ is an eigenvalue of DB (it corresponds to the eigenvector DBv). Thus, every eigenvalue of BDB is also an eigenvalue of DB.

Assume now that $DBv = \lambda v$. Left multiplication by B leads to $BDBv = \lambda Bv$, which is equivalent to $BDBBv = \lambda Bv$. But this means that λ is also an eingenvalue of BDB, and thus every eigenvalue of DB is also an eigenvalue of BDB. Since the converse had already been shown, it follows that BDB, DB (as well as its transpose BD) have the same set of eigenvalues.

The eigenvectors of BDB that correspond to its nonzero eigenvalues are generalizations of the *periodic* discrete prolate spheroidal sequences (P-DPSS) [8, 21]. We say "generalization" only because we are interested in arbitrary sets J. The P-DPSS correspond to the contiguous case, in which J is a set of consecutive integers modulo N.

The eigenvector that corresponds to λ_{max} will be denoted by v_{max}, whereas v_{min} will denote the eigenvector corresponding to the smallest nonzero eigenvalue λ_{min}. Both satisfy $x = Bx$, that is, both are band-limited. An example is given in figures 1b and 1c. Note how the energy of v_{max} and v_{min} is concentrated inside or outside the "window" determined by D, respectively. We will not explore the orthogonality of the P-DPSS, although that opens interesting possibilities.

The quantity $E(J, x)$ can be interpreted as an energy distribution in J, and

$$E(J, x) = \frac{\|DBx\|^2}{\|x\|^2}$$

is maximum when $x = v_{\mathsf{max}}$, and minimum when $x = v_{\mathsf{min}}$. This happens for a certain, fixed J. What happens if J varies?

Roughly speaking, when J is contiguous and the signals are low-pass, the lower bound may turn out close to zero and the upper bound close to one. If J is more evenly distributed, this may not be so.

The figures 1a–c and 2a–c refer to two numerical examples. The figures 1a and 2a depict the nonzero eigenvalues of BDB for two different sets J (to keep the figures readable, N and M were given quite low values but the algorithms can certainly be used for much larger problems). The eigenvectors v_{max} and v_{min} are shown in figures 1b–c and 2b–c. Bear in mind that card J was kept constant in both examples, and note the effect of J on the frame bounds and the behavior of the eigenvectors. These are important points to understand the numerical stability of the reconstruction problem — see also [8], and [10, 13].

The importance of the prolate spheroidal wave functions in connection with time-frequency concentration, uncertainty and other related issues was stressed by Slepian, Landau and Pollak in an important series of papers, known as "the Bell papers". See, for example, [24, 37, 38].

5 The frame algorithm

The frame operator associated with the frame $X \triangleq \{B_i\}_{i \in J}$ is

$$Sf \triangleq \sum_{i \in J} \langle f, B_i \rangle \, B_i.$$

(a)

(b)

(c)

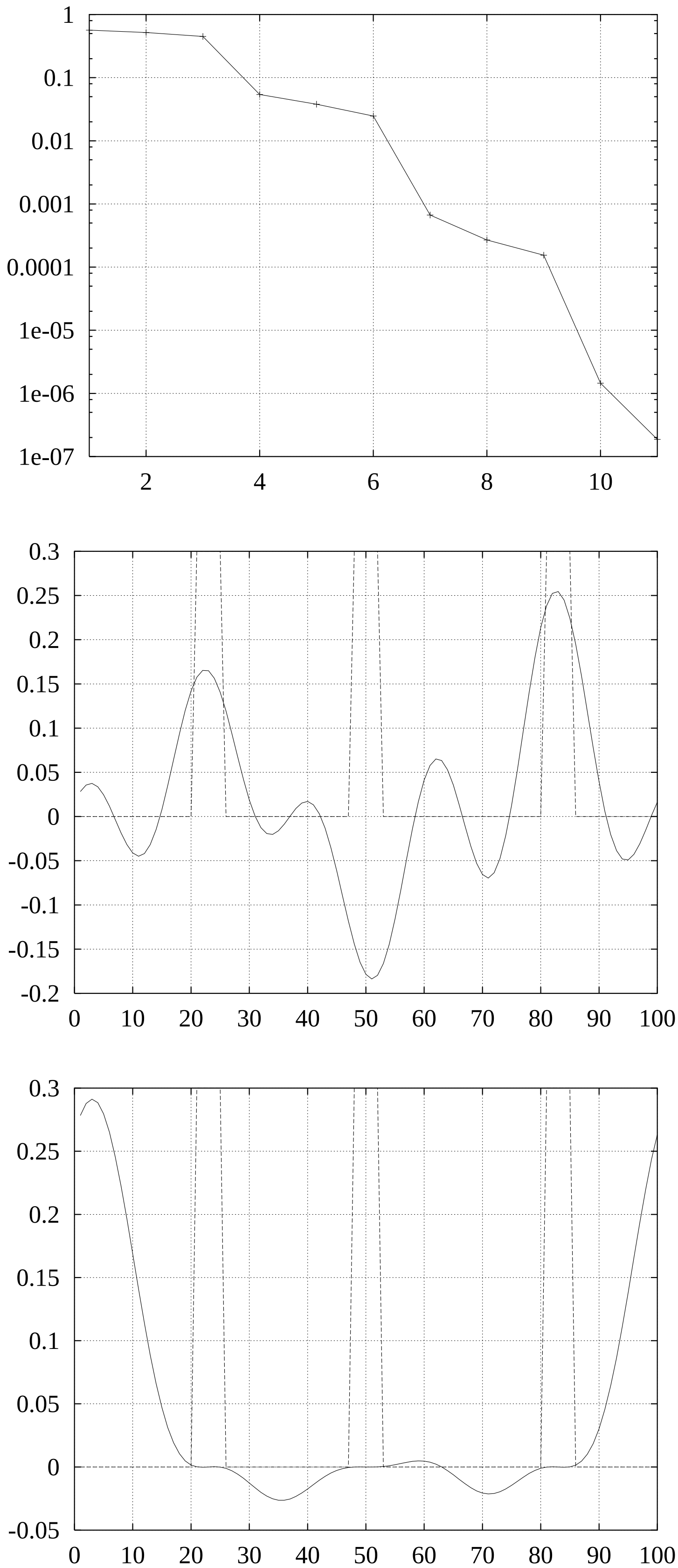

Figure 1: First example. (a) The nonzero eigenvalues of BDB. (b) The eigenvector v_{max} of BDB that corresponds to its largest eigenvalue λ_{max}. (c) The eigenvector v_{min} of BDB that corresponds to its smallest nonzero eigenvalue λ_{min}.

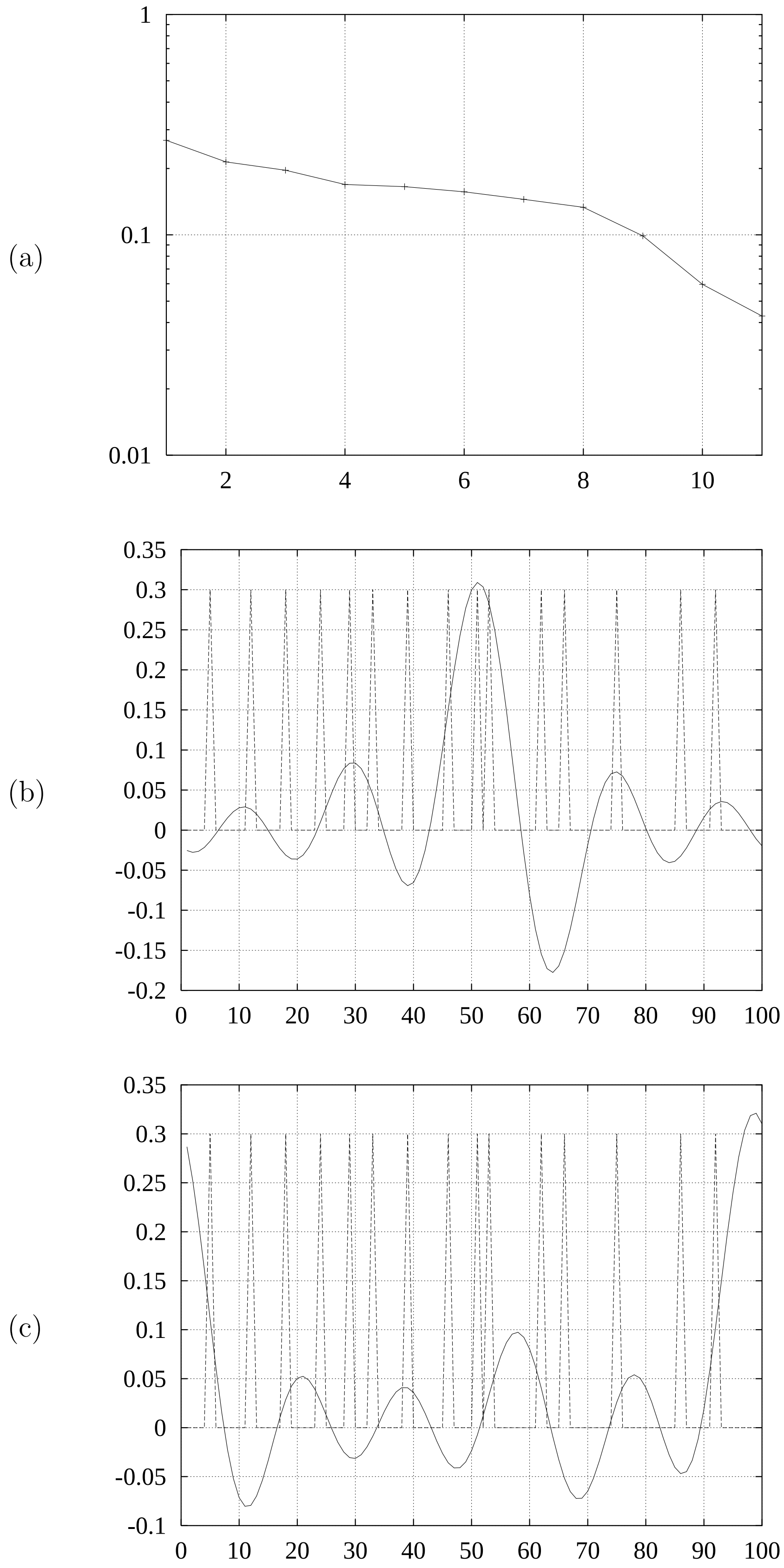

Figure 2: Second example. (a) The nonzero eigenvalues of BDB. (b) The eigenvector v_{max} of BDB that corresponds to its largest eigenvalue λ_{max}. (c) The eigenvector v_{min} of BDB that corresponds to its smallest nonzero eigenvalue λ_{min}.

The basic frame algorithm is the iteration

$$f^{(n)} = f^{(n-1)} + \mu S(f - f^{(n-1)}),$$

where μ is a real constant, given by

$$\mu = \frac{2}{\alpha + \beta},$$

and α and β are the frame bounds. The frame algorithm converges geometrically, at a rate given by

$$\left(\frac{\beta - \alpha}{\beta + \alpha}\right)^n.$$

This is the best possible rate if α and β are the best possible frame bounds. See, for example, [3, 6]. There are other variants and possibilities, but this is the simplest possibility. Acceleration of this basic iteration is discussed in [17].

We have seen that the frame bounds are $\alpha = \lambda_{\mathsf{min}} > 0$ and $\beta = \lambda_{\mathsf{max}} < 1$, and this determines μ. It is possible to avoid computing λ_{max} and λ_{min}, and use estimates for α and β. However, for missing data problems, computation of the bounds for several possible sets J, particularly those that are known to occur more often, can provide accurate information concerning the numerical difficulty of the reconstruction problem.

Recalling that $\langle f, B_i \rangle = f_i$ for any f satisfying $f = Bf$, we see that the frame operator can be computed using

$$Sf = \sum_{i \in J} f_i B_i.$$

The frame operator depends only on the samples f_i with $i \in J$. If these samples are known and all the others are unknown (say, because they were corrupted), Sf is still well defined, and the frame algorithm can then be used to recover the remaining samples.

The iteration can be written more explicitly as

$$f^{(n)} = f^{(n-1)} + \mu \sum_{j \in J} (f_j - f_j^{(n-1)}) B_j,$$

or in terms of the samples

$$f_i^{(n)} = f_i^{(n-1)} + \mu \sum_{j \in J} B_{ij} (f_j - f_j^{(n-1)}).$$

Introducing again the matrix D defined by (1), this becomes

$$f^{(n)} = f^{(n-1)} + \mu BD(f - f^{(n-1)}),$$

or, in slightly different form,

$$f^{(n)} = \mu BDf + (I - \mu BD) f^{(n-1)},$$

where I is the identity matrix. The matrix $I - \mu BD$ is called the iteration matrix.

The iteration converges if and only if the spectral radius of the iteration matrix $\rho(I - \mu BD)$ is below unity. This will be the case if

$$|1 - \mu \rho(BD)| < 1,$$

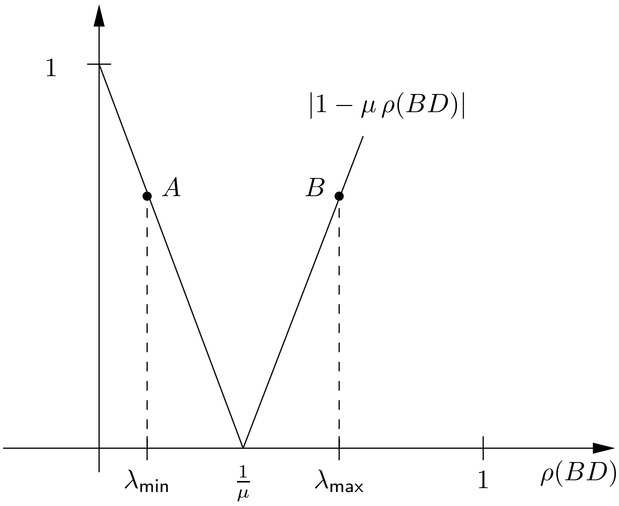

Figure 3: The optimum value of μ in the frame algorithm: if $1/\mu$ is not the average of λ_{max} and λ_{min}, either A or B will increase.

which shows that, for convergence,

$$\mu\rho(BD) < 2.$$

The optimum value of μ is indeed $\mu = 2/(\lambda_{\mathsf{max}} + \lambda_{\mathsf{min}})$ (refer to figure 3).

We will examine some of the connections between the frame algorithm and a few other methods in the next section.

6 The connections with other methods

6.1 Constrained restoration

The discrete finite frame algorithm can of course be obtained using methods that appear to bear no direct connection to frames, and that predate the widespread use of frames in the signal processing and engineering applications. The framework for constrained signal restoration discussed in [35] is one of these methods.

Assume that the signal f is distorted, the mathematical model for the distortion being an operator D. The result of applying D to f is available, but f itself is unknown: the signal restoration problem is an inverse problem.

The *a priori* knowledge concerning the signal f is expressed through one or more constraints. In this case, the relevant constraint is band-limiting, that is, one assumes that f satisfies an equation such as

$$f_i = \sum_{j=0}^{N-1} f_j \frac{\sin[\pi(2M+1)(i-j)/N]}{N\,\sin[\pi(i-j)/N]},$$

in the finite-dimensional vector case, or

$$f(t) = \int_{-\infty}^{+\infty} f(x) \frac{\sin w(t-x)}{\pi(t-x)}\,dx,$$

in $L_2(\mathbb{R})$. These are examples, but in the general case the equations will still be of the form $f = Bf$. The method starts from the identity

$$f = Bf + \mu(g - DBf),$$

where $g = Df = DBf$ is the distorted (observed) signal, and proceeds by iteration

$$f^{(n)} = Bf^{(n-1)} + \mu(g - DBf^{(n-1)}). \tag{6}$$

The convergence can be established using fixed-point theorems. Other known constraints (linear or nonlinear) can readily be incorporated in the basic iteration.

In the finite-dimensional setting, the distortion D can be thought of as multiplication by the matrix D defined by (1). The effect of band-limiting is to constraint the solution to the subspace of (say) low-pass signals, solutions to $x = Bx$, where B is the circulant matrix used before.

To compare the iteration (6) with the frame algorithm, note that applying B to both sides of (6) leads to

$$Bf^{(n)} = Bf^{(n-1)} + \mu(Bg - BDBf^{(n-1)}).$$

But $g = Df$, and so this turns out to be equivalent to

$$f^{(n)} = f^{(n-1)} + \mu BD(f - f^{(n-1)}),$$

which is the frame algorithm.

6.2 The Papoulis-Gerchberg iteration

Let $f = Bf$ and assume that some of the samples of f are known. Our task is to determine the remaining samples. Assume that the known samples are f_i, $i \in J$. With the help of the matrix D defined by (1), the given data can be written Df. As we have seen before, the Papoulis-Gerchberg iteration, originally introduced [15, 30, 31] as an extrapolation / superresolution method for $L_2(\mathbb{R})$ signals, can be used to approach this problem. The algorithm consists of two steps, one of which is band-limiting (application of the operator B). The other step enforces the time-domain knowledge, that is, it resets the known part of the signal to its true value.

In the finite-dimensional setting [8] band-limiting is multiplication by the circulant matrix B. Equation 3 shows how to insert the values of the known samples in a given vector x:

$$Px \triangleq (I - D)x + Df$$

Given the result of the iteration $n-1$, $f^{(n-1)}$, the Papoulis-Gerchberg iteration produces a new approximation $f^{(n)}$ to f according to the rule $f^{(n)} = PBf^{(n-1)}$, that is,

$$f^{(n)} = (I - D)Bf^{(n-1)} + Df.$$

Reversing the order of the two operations, that is, defining $f^{(n)}$ by $f^{(n)} = BPf^{(n-1)}$, or applying the band-limiting operator to both sides of the previous equation, leads to

$$f^{(n)} = B(I - D)f^{(n-1)} + BDf.$$

This is equivalent to

$$f^{(n)} = Bf^{(n-1)} + BD(f - f^{(n-1)}).$$

It is possible to replace $Bf^{(n-1)}$ by $f^{(n-1)}$, since the result of the iterations will be band-limited, if the initial vector was itself band-limited. Taking that step leads to the frame algorithm, with $\mu = 1$.

6.3 Alternating projections and POCS

The two operations upon which the Papoulis-Gerchberg method rests are projections onto convex sets. Band-limiting is a projection, and B is a projection matrix. Note that $B^2 = B$. In this case, the convex set is just the subspace of band-limited signals.

The operator P defined by (3) is also a projection. Note that $Pf = f$, and that $P^2 = P$. In this case, the convex set is the set of all signals x whose samples x_i, for $i \in J$, agree with the given data f_i, $i \in J$.

Hence, the Papoulis-Gerchberg iteration is an alternating projection method of the type discussed in [45]. This, in turn, is a special case of the POCS method [36, 46]. This opens new possibilities for incorporating nonlinear constraints in the problem (maximum or minimum amplitudes, positive or nonnegative character of the solution, and so on).

7 Comments and conclusion

We discussed a number of problems in signal reconstruction, error-control coding, fault-tolerant computing, and spectrum analysis. After examining the connections between these problems, we developed a tutorial exposition of frames in connection with the missing data problem. The frame bounds and the extremal signals of the restoration problem were related to the eigenvalues and eigenvectors of certain matrices. For extrapolation problems, the eigenvectors reduce to the periodic discrete prolate spheroidal wave sequences.

The finite-dimensional theory, despite its mathematical simplicity, is extremely useful for practical digital signal processing applications, which invariably involve a finite number of samples. It is also enlightening, because it exhibits many of the algebraic aspects that subsist in the more abstract and mathematically interesting settings, without the analytic subtleties that occur, for example, when dealing with limit processes and their interchange.

The frame algorithm is capable of acceptably good performance under certain conditions, but the convergence rate falls to very low values whenever $\beta/\alpha \gg 1$. These problems have been recognized, and there are several well understood ways of circumventing them, some of which are explored in the references that have already been given.

Among the possible solutions, we point out Chebyshev acceleration and conjugate gradient acceleration [17], and the adaptive weights method. See, for example, [7], which discusses several analytical and numerical aspects of the sampling problem. Another possibility is to interchange the time and frequency domains and try to solve for the DFT of f instead of solving for f itself (using a frame based on the columns of the matrix T, introduced in section 4.3, instead of a frame of column vectors of B). Yet another possibility is to reformulate the problem as a nonsingular set of linear equations. This is proposed, for example, in [18, 39] and [9]. The equations may then be solved using any of the standard methods, either iterative, noniterative, or semi-iterative [16, 40, 47]. Also in this case there is a choice, between the time-domain and the frequency-domain, as explained in [12]. The improvements that can be obtained using these techniques are substantial, often leading to performance that exceeds by orders of magnitude the simplest iterative methods.

References

[1] C. J. Anfinson and F. T. Luk. A linear algebraic model of algorithm-based fault tolerance. *IEEE Trans. Comput.*, 37(12):1599–1604, Dec. 1988.

[2] E. O. Brigham. *The Fast Fourier Transform and Its Applications.* Prentice-Hall, Englewood Cliffs, New Jersey, 1988.

[3] I. Daubechies. *Ten Lectures on Wavelets.* CBMS-NSF Regional Conference Series in Applied Mathematics. SIAM, Philadelphia, 1992.

[4] I. Daubechies. From the original framer to present-day time-frequency and time-scale frames. *The J. Fourier Anal. Appl.*, 3(5):485–486, 1997.

[5] P. J. Davis. *Circulant Matrices.* John Wiley & Sons, New York, 1979.

[6] R. J. Duffin and A. C. Schaeffer. A class of nonharmonic Fourier series. *Trans. Amer. Math. Soc.*, 72:341–366, 1952.

[7] H. G. Feichtinger and K. Gröchenig. Theory and practice of irregular sampling. In J. J. Benedetto and M. W. Frazier, editors, *Wavelets: Mathematics and Applications*, pages 305–363. CRC Press, Boca Raton, 1994.

[8] P. J. S. G. Ferreira. Interpolation and the discrete Papoulis-Gerchberg algorithm. *IEEE Trans. Signal Processing*, 42(10):2596–2606, Oct. 1994.

[9] P. J. S. G. Ferreira. Noniterative and faster iterative methods for interpolation and extrapolation. *IEEE Trans. Signal Processing*, 42(11):3278–3282, Nov. 1994.

[10] P. J. S. G. Ferreira. The stability of a procedure for the recovery of lost samples in band-limited signals. *Sig. Proc.*, 40(3):195–205, Dec. 1994.

[11] P. J. S. G. Ferreira. The duality of two recent image interpolation methods. In *Proceedings of the Third IEEE International Conference on Image Processing, ICIP-96*, pages 735–738, Lausanne, Switzerland, Sep. 1996.

[12] P. J. S. G. Ferreira. Interpolation in the time and frequency domains. *IEEE Sig. Proc. Letters*, 3(6):176–178, June 1996.

[13] P. J. S. G. Ferreira. The eigenvalues of matrices which occur in certain interpolation problems. *IEEE Trans. Signal Processing*, 45(8):2115–2120, Aug. 1997.

[14] P. J. S. G. Ferreira and J. M. N. Vieira. Detection and correction of missing samples. In *Proceedings of the 1997 Workshop on Sampling Theory and Applications*, pages 169–174, Aveiro, Portugal, June 1997.

[15] R. W. Gerchberg. Super-resolution through error energy reduction. *Optica Acta*, 21(9):709–720, 1974.

[16] G. H. Golub and C. F. Van Loan. *Matrix Computations.* The Johns Hopkins University Press, Baltimore, 1989.

[17] K. Gröchenig. Acceleration of the frame algorithm. *IEEE Trans. Signal Processing*, 41(12):3331–3340, Dec. 1993.

[18] K. Gröchenig. A discrete theory of irregular sampling. *Linear Algebra Appl.*, 193:129–150, 1993.

[19] R. A. Horn and C. R. Johnson. *Matrix Analysis.* Cambridge University Press, Cambridge, 1990.

[20] K.-H. Huang and J. A. Abraham. Algorithm-based fault tolerance for matrix operations. *IEEE Trans. Comput.*, 33(6):518–528, June 1984.

[21] A. K. Jain and S. Ranganath. Extrapolation algorithms for discrete signals with application in spectral estimation. *IEEE Trans. Acoust., Speech, Signal Processing*, 29(4):830–845, Aug. 1981.

[22] M. C. Jones. The discrete Gerchberg algorithm. *IEEE Trans. Acoust., Speech, Signal Processing*, 34(3):624–626, June 1986.

[23] S. M. Kay and S. L. Marple. Spectrum analysis — a modern perspective. *Proc. IEEE*, 69(11):1380–1419, Nov. 1981. Reprinted by IEEE Press, (Modern Spectrum Analysis II).

[24] H. J. Landau and H. O. Pollak. Prolate spheroidal wave functions, Fourier analysis and uncertainty – II. *Bell Syst. Tech. J.*, 40:65–84, Jan. 1961.

[25] F. A. Marvasti. An iterative method to compensate for the interpolation distortion. *IEEE Trans. Acoust., Speech, Signal Processing*, 37(10):1617–1621, Oct. 1989.

[26] F. A. Marvasti, C. Liu, and G. Adams. Analysis and recovery of multidimensional signals from irregular samples using nonlinear and iterative techniques. *Sig. Proc.*, 36:13–30, 1993.

[27] F. A. Marvasti and M. Nafie. Sampling theorem: A unified outlook on information theory, block and convolutional codes. *IEICE Trans. Fundam. Electron. Commun. Comput. Sci.*, E76–A(9):1383–1391, Sep. 1993.

[28] V. S. S. Nair and J. A. Abraham. Real-number codes for fault-tolerant matrix operations on processor arrays. *IEEE Trans. Comput.*, 39(4):426–435, Apr. 1990.

[29] A. L. Narasimha Reddy and P. Bannerjee. Algorithm-based fault detection for signal processing applications. *IEEE Trans. Comput.*, 39(10):1304–1308, Oct. 1990.

[30] A. Papoulis. A new algorithm in spectral analysis and band-limited extrapolation. *IEEE Trans. Circuits Syst.*, 22(9):735–742, Sep. 1975.

[31] A. Papoulis. *Signal Analysis.* McGraw-Hill International Editions, New York, 1987.

[32] A. Papoulis and C. Chamzas. Detection of hidden periodicities by adaptive extrapolation. *IEEE Trans. Acoust., Speech, Signal Processing*, 27(5):492–500, Oct. 1979.

[33] S.-C. Pei and M.-H. Yeh. An introduction to discrete finite frames. *IEEE Sig. Proc. Mag.*, 14(6):84–96, Nov. 1997.

[34] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. *Numerical Recipes in C.* Cambridge University Press, Cambridge, 1988.

[35] R. W. Schafer, R. M. Mersereau, and M. A. Richards. Constrained iterative restoration algorithms. *Proc. IEEE*, 69(4):432–450, Apr. 1981.

[36] M. I. Sezan and H. Stark. Image restoration by the method of convex projections: Part 2 — applications and numerical results. *IEEE Trans. Med. Imag.*, 1(2):95–101, Oct. 1982.

[37] D. Slepian. Prolate spheroidal wave functions, Fourier analysis and uncertainty — V: The discrete case. *Bell Syst. Tech. J.*, 57(5):1371–1429, May 1978.

[38] D. Slepian and H. O. Pollak. Prolate spheroidal wave functions, Fourier analysis and uncertainty — I. *Bell Syst. Tech. J.*, 40(1):43–63, Jan. 1961.

[39] T. Strohmer. On discrete band-limited signal extrapolation. In M. E. H. Ismail, M. Z. Nashed, A. I. Zayed, and A. F. Ghaleb, editors, *Mathematical Analysis, Wavelets, and Signal Processing*, volume 190 of *Contemporary Mathematics*, pages 323–337, Providence, Rhode Island, 1995. American Mathematical Society.

[40] R. S. Varga. *Matrix Iterative Analysis.* Prentice-Hall, Englewood Cliffs, New Jersey, 1962.

[41] J. M. N. Vieira and P. J. S. G. Ferreira. Interpolation, spectrum analysis, error-control coding, and fault-tolerant computing. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 97*, volume III, pages 1831–1834, Munich, Germany, Apr. 1997.

[42] J. S. Walker. *Fast Fourier Transforms.* Studies in Advanced Mathematics. CRC Press, Boca Raton, 1991.

[43] J. K. Wolf. Redundancy, the discrete Fourier transform, and impulse noise cancellation. *IEEE Trans. Commun.*, 31(3):458–461, Mar. 1983.

[44] C. K. W. Wong, F. Marvasti, and W. G. Chambers. Implementation of recovery of speech with impulsive noise on a DSP chip. *Electron. Letters*, 31(17):1412–1413, Aug. 1995.

[45] D. C. Youla. Generalized image restoration by the method of alternating orthogonal projections. *IEEE Trans. Circuits Syst.*, 25(9):694–702, Sep. 1978.

[46] D. C. Youla and H. Webb. Image restoration by the method of convex projections: Part 1 — theory. *IEEE Trans. Med. Imag.*, 1(2):81–94, Oct. 1982.

[47] D. M. Young. *Iterative Solution of Large Linear Systems.* Academic Press, Orlando, third edition, 1971.

[48] R. M. Young. *An Introduction to Nonharmonic Fourier Series.* Academic Press, New York, 1980.