Discrete Irregular Sampling With Larger Gaps

Haja N. Razafinjatovo
Department of Mathematics
James Madison University
122 Burruss Hall
Harrisonburg, Virginia 22807

Submitted by Michael Neumann

ABSTRACT

Discrete irregular sampling is the problem of recovering a band-limited discrete signal from its nonuniformly sampled values. We introduce a new method for discrete irregular sampling. It features a new sampling geometry where the sampling points are chosen to come in pairs. A frequency domain formulation reduces the problem to a system of linear equations for which the matrix is of Toeplitz type and is positive definite as soon as there are enough samples. An explicit estimate of the condition number of that matrix is given under some condition on the maximum sampling gap. This condition allows gaps larger than the Nyquist gap. We recombine the new method and the conjugate method to deduce an efficient iterative reconstruction scheme for band-limited discrete signals. Numerical results are provided to illustrate the efficiency of the method. © Elsevier Science Inc., 1997

1. INTRODUCTION

The numerical aspect of irregular sampling has drawn much attention in the last decade mainly because of its importance in signal analysis in general and also because of the successful attempts to use different techniques, such as the theory of frames, to tackle the problem. Sampling of band-limited signals have been considered the most (see [1–9]) since many practical problems deal with them in many scientific domains such as signal and image...
For band-limited discrete periodic signals the reconstruction problem amounts to solving a linear system of equations. Recently very efficient numerical methods have been developed by H. Feichtinger and K. Gröchenig [4]. These methods work best if the largest gap in the sampling sequence is comparable to the Nyquist gap or less. However, for many applications, signal reconstruction from sampling sequences with gaps larger than Nyquist are very common. In the presence of larger gaps, the above two authors used preconditioners to speed up the convergence since the condition number of the original matrix increases. In the worst situation when the size of the spectrum is comparable to the number of samples and large gaps occur, huge condition number can then be registered (as large as $10^{12}$–$10^{15}$).

In this article we present a new method for sampling band-limited periodic discrete signals. The method allows larger gaps in the sampling set, yet keeps the system fairly stable. A key to the better conditioning of the method is its sampling geometry. Instead of just any nonuniformly spaced discrete sequence, the sampling points are chosen to come in pairs. As a result of this geometry, the analysis of the problem in the frequency domain gives rise to a system of linear equation with a matrix of the form $C + D\hat{C}D$, where $C$ and $\hat{C}$ are Toeplitz matrices and $D$ is a diagonal matrix, the size of which are independent of the number of samples. This matrix of the system is positive definite if and only if there are at least $M + 1$ pairs of samples, with $M$ being the size of the spectrum of the signal; hence the equations could be solved directly by a matrix inversion. However, for large $M$ it is much more efficient to use the conjugate gradient method to find an approximate solution. Moreover, the matrix structure allows reduction of further computational costs.

The new model was inspired by recent results on the theory of irregular sampling with derivatives of band-limited functions on $\mathbb{R}$ [10]. There, the theory of frames was used to deduce the reconstruction scheme. For a numerical implementation of that result we were led to consider a new discrete model. This model was initially regarded as a discretization of the continuous case of irregular sampling with derivatives, but in fact it turned out to provide a better discrete method for the reconstruction from samples only. For more on frames and their relevance to irregular sampling refer to [11–13].

This article is organized as follows: Section 2.1 gives a definition of band-limited discrete signals and standard formulas of their DFT and IDFT. The new method is then developed in Section 2.2. We explain first the principal ideas and then formulate the main theorem, which gives an estimate on the condition number of the system matrix under a largest gap condition. Finally, we present the weighted difference-conjugate gradient algorithm.
Section 3 gives some numerical results where the performance of the new method is tested by a numerical comparison with two other methods for discrete nonuniform sampling.

2. A NEW METHOD FOR DISCRETE IRREGULAR SAMPLING

2.1. Band-Limited Discrete Signals

Consider a discrete signal $s$ of length $N$, $s(0)$, $s(1)$, ..., $s(N-1)$, and extend it to all of $\mathbb{Z}$ by periodicity, i.e., $s(j) = s(j + lN)$ for $j = 0, 1, \ldots, N - 1$ and $l \in \mathbb{Z}$. By $l^2(\mathbb{Z}_N)$ we denote the space of such signals equipped with the norm

$$||s|| = \left( \sum_{n=0}^{N-1} |s(n)|^2 \right)^{1/2}.$$

The discrete Fourier transform (DFT) $\hat{s}$ of $s$ is defined by

$$\hat{s}(k) = \frac{1}{N} \sum_{n=0}^{N-1} s(n) e^{-2\pi i kn / N}$$

for $k = 0, 1, \ldots, N - 1$.

For an integer $M$ such that $0 < M < N/2$, define the space of band-limited discrete signals $B_M$ by

$$B_M = \{ s \in l^2(\mathbb{Z}_N) \mid \hat{s}(k) = 0, |k| > M \}.$$

For $s \in B_M$, the inverse discrete Fourier transform (IDFT) is given by

$$s(n) = \frac{1}{N} \sum_{k=-M}^{M} \hat{s}(k) e^{2\pi i kn / N}$$

for $n = 0, 1, \ldots, N - 1$. Note that $||s|| = ||\hat{s}||$.

2.2. The Weighted Difference Method

In a discrete irregular sampling problem the goal is to recover a discrete signal $s \in l^2(\mathbb{Z}_N)$ from the knowledge of some of its values at irregularly spaced sampling points. In general if the largest gap in the sampling sequence
is smaller than the Nyquist gap $N/(2M + 1)$ and the problem is overdetermined then the reconstruction amounts to solving a well-conditioned Toeplitz system. However, if there are larger gaps, which is the case in many applications, then the reconstruction becomes less stable.

The following method is introduced to improve the stability of the reconstruction in the presence of gaps larger than Nyquist. We deal with a new system that is not Toeplitz any more but still involves Toeplitz matrices.

For this purpose consider a sampling sequence where each point comes in pairs, i.e., a subset

$$\{n_1, n_1 + 1, n_2, n_2 + 1, \ldots, n_r, n_r + 1\}$$

of \{0, 1, \ldots, N - 1\} with $n_j + 1 < n_{j+1}$. For $s \in B_M$, suppose that we are given the samples

$$\{s(n_j), s(n_j + 1)\}_{j=1}^r.$$

Let $\Delta s(n_j) = s(n_j + 1) - s(n_j)$ and let $\sin c_M$ be defined by

$$\sin c_M(j) = \frac{1}{N} \sum_{k=-M}^{M} e^{2\pi i k j/N} = \frac{\sin\{(2M + 1)\pi j/N\}}{N \sin\{\pi j/N\}}. \quad (2)$$

Also set

$$\Delta L_n \sin c_M = L_{n+1} \sin c_M - L_n \sin c_M,$$

where $L$ is the translation operator $L_j s(n) = s(n - j)$. Then, for $s \in B_M$, $s(n) = \langle s, L_n \sin c_M \rangle$.

Set $l_j = [(n_{j+1} + n_j + 1)/2]$, $j = 0, 1, \ldots, r$, where $[x]$ is the largest integer $\leq x \in \mathbb{R}$. Here $n_0 = n_r + 1 - N$ and $n_{r+1} = n_1 + N$ by periodicity.

The central idea of the method is the following: From the given samples \{$s(n_j), s(n_j + 1)\}_{j=1}^r$, define the operator $S: B_M \to B_M$ by

$$Ss = \sum_{j=1}^{r} \left\{ \mu_j s(n_j) L_n \sin c_M + \widetilde{\mu}_j \Delta s(n_j) \Delta L_n \sin c_M \right\}. \quad (3)$$
with
\[ \mu_j = l_j - l_{j-1} \quad \text{and} \quad \widetilde{\mu}_j = \sum_{k = l_{j-1} + 1}^{l_j} (k - n_j)^2 \]
for \( j = 1, \ldots, r \).

Then the following proposition gives the relation between \( \tilde{S}s \) and \( \tilde{s} \) and describes the matrix of our new method:

**Proposition 2.** For \( s \in B_m \) let \( \tilde{s} = (\tilde{s}(k))_{|k| \leq M} \) and \( \tilde{S}s = (\tilde{S}s(k))_{|k| \leq M} \). Then we have
\[ \tilde{S}s = (C + D^* \tilde{C}D) \tilde{s}, \]
where \( C \) and \( \tilde{C} \) are Hermitian Toeplitz matrices given by
\[ c_{lk} = \frac{1}{N} \sum_{j=1}^{r} \mu_j e^{-2\pi i n_j(l-k)/N}, \]
\[ \tilde{c}_{lk} = \frac{1}{N} \sum_{j=1}^{r} \tilde{\mu}_j e^{-2\pi i n_j(l-k)/N} \]
for \(|l|, |k| \leq M\), and \( D = \text{diag}(e^{2\pi ik/N} - 1), |k| \leq M \). Moreover, \( \tilde{S}s \) can be evaluated from the given data as follows:
\[ \tilde{S}s(k) = N^{-1/2} \sum_{j=1}^{r} \mu_j s(n_j) e^{-2\pi i n_j / N} \]
\[ + N^{-1/2}(e^{-2\pi ik/N} - 1) \sum_{j=1}^{r} \tilde{\mu}_j \Delta s(n_j) e^{2\pi i n_j / N}. \]

**Proof.** Using Eqs. (1) and (2), straightforward computations show that
\[ \tilde{S}s(l) = \frac{1}{N} \sum_{k=-M}^{M} \left( \sum_{j=1}^{r} \mu_j e^{-2\pi in_j(l-k)/N} \right) \tilde{s}(k) \]
\[ + \frac{1}{N}(e^{-2\pi il/N} - 1) \sum_{k=-M}^{M} \left( \sum_{j=1}^{r} \tilde{\mu}_j e^{-2\pi in_j(l-k)/N} \right)(e^{2\pi ik/N} - 1) \tilde{s}(k) \]
thus obtaining (5). The expression (6) for \( \tilde{S}s \) is obtained in a similar way.
Corollary 1. For \( n = 0, 1, \ldots, N - 1 \), define \( u, v, u_b, \) and \( v_b \) by

\[
u_{\mu}(n) = \begin{cases} \mu_j / N & \text{if } n = n_j \\ 0 & \text{else} \end{cases} \quad v_{\mu}(n) = \begin{cases} \overline{\mu_j} / N & \text{if } n = n_j \\ 0 & \text{else} \end{cases}
\]

and

\[
u_b(n) = \begin{cases} \frac{1}{\sqrt{N}} \mu_j s(n_j) & \text{if } n = n_j \\ 0 & \text{else} \end{cases} \quad v_b(n) = \begin{cases} \frac{1}{\sqrt{N}} \overline{\mu} \overline{s}(n_j) & \text{if } n = n_j \\ 0 & \text{else} \end{cases}
\]

Then the entries \( c_0, c_1, \ldots, c_{2M} \) and \( \overline{c_0}, \overline{c_1}, \ldots, \overline{c_{2M}} \) of the matrices \( C \) and \( \overline{C} \) can be computed using DFT by

\[c_k = \overline{u_{\mu}}(k) \quad \text{and} \quad \overline{c_k} = \overline{v_{\mu}}(k) \quad \text{for } k = 0, 1, \ldots, 2M.
\]

Similarly the two summations in (6) are given respectively by \( \overline{u_b}(k) \) and \( \overline{v_b}(k) \) for \( |k| \leq M \).

Proof. These follow immediately from the definition of the DFT. ■

Proposition 1 states that our problem now turns into solving system (5) for \( \hat{s} \). Of course a direct solution of (5) is obtained if the matrix \( t_{\mu} = C + D^* \overline{C}D \) is invertible. Here are necessary and sufficient conditions for that to happen:

Proposition 2. Let \( T_{\mu} \) be as defined in Proposition 1 and \( r \) be the number of pairs of sampling points. Then the following statements are equivalent:

(i) \( r \geq M + 1 \)
(ii) \( T_{\mu} \) is invertible
(iii) \( T_{\mu} \) is positive definite.

Thus if \( r \geq M + 1 \) then \( \hat{s} = T_{\mu}^{-1} \overline{s}s \) and \( s \) is recovered by IDFT.

Proof. • (i) \( \rightarrow \) (ii). If \( T_{\mu} \) is not invertible, then there exists a nonzero vector \( a = (a_k)_{|k| \leq M} \) of \( \mathbb{C}^{2M + 1} \) and a nonzero signal \( s(n) = (1/\sqrt{N}) \)
\[ \sum_{k=-M}^{M} a_k e^{2\pi i k n / N} \] in \( B_M \) such that \( T_\mu a = 0 \). In particular we have
\[ 0 = \langle T_\mu a, a \rangle = \langle S s, s \rangle = \sum_{j=1}^{r} \left\{ \mu_j |s(n_j)|^2 + \widetilde{\mu_j} |\Delta s(n_j)|^2 \right\}. \tag{7} \]
which implies that \( s(n_j) = s(n_j + 1) = 0 \) for \( j = 1, \ldots, r \) since \( \mu_j, \widetilde{\mu_j} > 0 \). But then (1) together with the condition \( r > M + 1 \) implies that \( a_k = 0 \), \( |k| \leq M \) since the \((2M + 1) \times 2r\) matrix with entries \( \{ e^{2\pi i k n_j / N}, e^{2\pi i k (n_j + 1) / N} \} \) is a Vandermonde matrix which always has full rank whenever \( 2r \geq 2M + 1 \). This is a contradiction.

- (ii) \( \Rightarrow \) (iii). Let \( a = (a_k)_{|k| \leq M} \) be a nonzero vector of \( \mathbb{C}^{2M+1} \) and let \( s(n) = N^{-1/2} \sum_{k=-M}^{M} a_k e^{2\pi i k n / N} \). Then,
\[ \langle T_\mu a, a \rangle = \langle S s, s \rangle = \sum_{j=1}^{r} \left\{ \mu_j |s(n_j)|^2 + \widetilde{\mu_j} |\Delta s(n_j)|^2 \right\} \geq 0, \]
which shows that \( T_\mu \) is positive semidefinite. Invertibility thus implies that \( T_\mu \) is positive definite.

- (iii) \( \Rightarrow \) (ii). If \( T_\mu \) is positive definite, then all of its eigenvalues are positive reals; thus \( T_\mu \) is invertible.

- (ii) \( \Rightarrow \) (i). Suppose that \( r \leq M \) and consider the following system of equations:
\[ 0 = N^{-1/2} \sum_{k=-M}^{M} a_k e^{2\pi i k n_j / N} \]
\[ 0 = N^{-1/2} \sum_{k=-M}^{M} a_k e^{2\pi i k (n_j + 1) / N} . \]
Since there are \( 2r \) \((\leq 2M)\) equations and \( 2M + 1 \) unknowns \( a_k \)'s there must exist a nontrivial solution \( a = (a_k) \in \mathbb{C}^{2M+1} \). Define \( s \in B_M \) by \( s(n) = N^{-1/2} \sum_{k=-M}^{M} a_k e^{2\pi i k n / N} \). Then \( T_\mu s = T_\mu a = 0 \) by (6). Thus \( T_\mu \) is not invertible.

The proposition is proved.

REMARKS. 1. Clearly the proof of (i) \( \Rightarrow \) (ii) works for any choice of weights \( \mu_j, \widetilde{\mu_j} > 0 \). However, as we see later, the choice of the weights defined in (4) keeps the system fairly stable even in the presence of larger gaps.

2. An important advantage of the above frequency domain formulation of the problem is that the size of the underlined matrix \( T_\mu \) is independent of the number of samples. It only depends on the size of the spectrum \( M \).
3. The Toeplitz structure of $C$ and $\tilde{C}$, along with Corollary 1, plays an important role for numerical implementation (see Section 3 for more details).

4. By Eq. (11), another way of finding $\hat{s}(k)$ from the knowledge of $\{s(n_j), s(n_j + 1)\}$ is to solve the system of linear equations:

$$s(n_j) = N^{-1/2} \sum_{k=-M}^{M} \hat{s}(k) e^{2\pi i k n_j / N}$$

$$s(n_j + 1) = N^{-1/2} \sum_{k=-M}^{M} \hat{s}(k) e^{2\pi i k (n_j + 1) / N}$$

for $j = 1, \ldots, r$. If $r \geq M + 1$, then any $2M + 1$ of the $2r$ equations suffice to determine the unique solution since the underlying matrix is a $(2M + 1) \times (2M + 1)$ Vandermonde matrix $V$ with distinct nodes taken from $\{e^{2 \pi i n_j / N}, e^{2 \pi i (n_j + 1) / N}\}$. However, in most practical problems $N$ is large so that $|e^{2 \pi i n_j / N} - e^{2 \pi i (n_j + 1) / N}| \approx 0$; hence estimates in [14–16] imply that the condition number of $V$ is very large. Thus in practical situations a better conditioned method is needed.

The following lemma defines a quantity needed for our main theorem, and is used in its proof:

**Lemma 1.** Let $x_1, x_2, \ldots, x_n$ be $n$ complex numbers and set $\Delta^2 x_k = x_{k+2} - 2x_{k+1} + x_k$. If $x_1 = x_2 = 0$, then

$$\sum_{k=1}^{n} |x_k|^2 \leq K(n) \sum_{k=1}^{n-2} |\Delta^2 x_k|^2, \quad n \geq 3,$$

where

$$K(n) = \frac{n(n-2)(n-1)^2}{12}.$$

**Proof.** For $k \geq 3$, write $x_k = \sum_{j=1}^{k-2} (k-j-1) \Delta^2 x_j$. Hence by Cauchy-Schwarz,

$$|x_k|^2 \leq \sum_{j=1}^{k-2} (k-j-1)^2 \sum_{j=1}^{k-2} |\Delta^2 x_j|^2.$$
Now sum both sides from \( k = 3 \) to \( n \). Then \( K(n) \) is the largest coefficient of the \( |\Delta^2 x_k|^2 \), \( k = 1, \ldots, n - 2 \). This is the coefficient of \( |\Delta^2 x_1|^2 \) and is given by

\[
K(n) = 1^2 + (1^2 + 2^2) + \cdots + (1^2 + 2^2 + \cdots + (n - 2)^2)
\]

\[
= (n - 2)1^2 + (n - 3)2^2 + \cdots + (n - 2)^2 = \sum_{k=1}^{n-2} (n - k - 1)k^2.
\]

But for any positive integer \( p \) we have that \( \sum_{k=1}^{p} k^2 = p(p + 1)(2p + 1)/6 \) and \( \sum_{k=1}^{p} k^3 = p^2(p + 1)^2/4 \). Thus,

\[
K(n) = (n - 1)\sum_{k=1}^{n-2} k^2 - \sum_{k=1}^{n-2} k^3
\]

\[
= \frac{2(n - 1)^2(n - 2)(2n - 3) - 3(n - 2)^2(n - 1)^2}{12}
\]

\[
= \frac{(n - 1)^2(n - 2)[4n - 6 - 3n + 6]}{12} = \frac{n(n - 2)(n - 1)^3}{12},
\]

and the lemma is proved.

Here is our main theorem. It gives estimates for the smallest and largest eigenvalues \( \lambda \) and \( \Lambda \), and the condition number of \( T_\mu \).

**Theorem 1.** Set \( d = \max_{j=0, \ldots, r} (n_{j+1} - n_j) \).

\[
\mu_j = l_j - l_{j-1} \quad \text{and} \quad \tilde{\mu}_j = \sum_{k=l_{j-1}+1}^{l_j} (k - n_j)^2.
\]

If \( r \geq M + 1 \), then

\[
\alpha \|s\|^2 \leq \sum_{j=1}^{r} \left( \mu_j |s(n_j)|^2 + \tilde{\mu}_j |\Delta s(n_j)|^2 \right) \leq \beta \|s\|^2
\]

(8)
for all \( s \in B_M \) whenever

\[
\left[ d/s \right] \sqrt{\left[ d/2 \right]^2} - 1 < \frac{\sqrt{3}}{2} \left( \sin \frac{\pi M}{N} \right)^{-2},
\]

(9)

with

\[
\alpha = 2^{-1}(1 - \kappa_d^2)^2, \quad \beta = 2(2 - \sqrt{3})^{-1}(1 + \kappa_d^2)^2,
\]

(10)

where \( \kappa_d = (2/\sqrt{3})\left[ d/2 \right] \sqrt{\left[ d/2 \right]^2 - 1} \sin^2(\pi M/N) \). It follows from (8) that \( \alpha < \lambda \) and \( \beta \geq \Lambda \), thus

\[
\text{cond}(T_\mu) \leq \frac{\beta}{\alpha}.
\]

(11)

Three other lemmas are also needed for the proof of the theorem.

**Lemma 2.** Let \( B \) be a Banach space and \( A \) a bounded operator on \( B \) such that \( \|\text{Id} - A\|_{\text{op}} < 1 \), where \( \| \cdot \|_{\text{op}} \) denotes the operator norm on \( B \). Then \( A \) has a bounded inverse on \( B \) and \( \| A^{-1} \|_{\text{op}} \leq (1 - \|\text{Id} - A\|_{\text{op}})^{-1} \).

**Lemma 3.** Let \( s \in B_M \). Set \( \Delta s(n) = s(n+1) - s(n) \) (and by periodicity \( \Delta s(N-1) = s(0) - s(N-1) \)). Then we have:

\[
\|\Delta s\| \leq 2\sin\frac{\pi M}{N} \|s\|.
\]

This is the discrete version of the well-known Bernstein inequality.

**Proof.** By definition

\[
\widehat{\Delta s}(k) = N^{-1/2} \sum_{n=0}^{N-1} [s(n+1) - s(n)] e^{-2\pi i kn/N} = \hat{s}(k)(e^{2\pi ik/N} - 1).
\]
Hence,
\[ \| \Delta s \| = \left( \sum_{k = -M}^{M} | \hat{s}(k)|^2 \left| e^{2\pi ik/N} - 1 \right|^2 \right)^{1/2} \leq \max_{|k| \leq M} \left| e^{2\pi ik/N} - 1 \right| \| \hat{s} \| \]
\[ \leq 2\sin \frac{\pi M}{N} \| \hat{s} \| \]
since for \( |k| \leq M \) we have \( |e^{2\pi ik/N} - 1| = 2\sin(\pi k/N) \leq 2\sin(\pi M/N) \). 

**Lemma 4.** Let \( v_j = \sum_{i=j-1}^{j}(n-i) \). Then
\[ \frac{v_j^2}{\mu_j \mu_j} \leq \frac{3}{4} \]
for all \( j \).

**Proof.** Recall that \( \mu_j = l_j - l_{j-1} \), and \( \tilde{\mu}_j = \sum_{k=j-1}^{j}(k-n_j)^2 \). Now, let \( b_j = l_{j-1} + 1 - n_j \). Then
\[ v_j = \sum_{k=0}^{\mu_j-1} (b_j + k) = \mu_j \left( b_j + \frac{\mu_j - 1}{2} \right) \]
and
\[ \tilde{\mu}_j = \sum_{k=0}^{\mu_j-1} (b_j + k)^2 + \mu_j \left( b_j^2 + (\mu_j - 1)b_j + \frac{(\mu_j - 1)(2\mu_j - 1)}{6} \right) \].

Hence,
\[ \frac{v_j^2}{\mu_j \tilde{\mu}_j} = \frac{(b_j + (\mu_j - 1)/2)^2}{(b_j^2 + (\mu_j - 1)b_j + (\mu_j - 1)(2\mu_j - 1)/6)} \].

Two cases: if \( b_j + (\mu_j - 1)/2 = 0 \), then the claim is obviously true; otherwise we have
\[ \frac{v_j^2}{\mu_j \tilde{\mu}_j} = \frac{1}{1 + \frac{1}{3}(\mu_j - 1)(\mu_j + 1)/(2b_j + \mu_j - 1)^2} \]
in which case \( (\mu_j - 1)(\mu_j + 1)/(2b_j + \mu_j - 1)^2 \geq 1 \) will prove the claim.
Now, write $\mu_j - 1 = t_j - n_j + n_j - l_{j-1} - 1$ and set $s_j = l_j - n_j$, $t_j = n_j - l_{j-1} = -b_j$. Then

$$ \frac{(\mu_j - 1)(\mu_j + 1)}{(2b_j + \mu_j - 1)^2} = \frac{(s_j + t_j)(s_j + t_j + 2)}{(s_j - t_j)^2} = \frac{(s_j - t_j)^2 + 4s_jt_j + 2(s_j + t_j)}{(s_j - t_j)^2} \geq 1 $$

since $s_j, t_j \geq 0$. Thus, our claim follows.

Proof of Theorem 1. Let $\chi_j$ be the characteristic function of $\{l_{j-1} + 1, l_{j-1} + 2, \ldots, l_j\}$ for $j = 1, \ldots, r$. Then $\sum_{j=1}^r \chi_j(n) = 1$ for all $n$. Denote by $P$ the orthogonal projection onto $B_M$ and define the operator $A$ of $B_M$ by

$$ A\chi_j(n) = P\left( \sum_{j=1}^r \{ s(n_j) + (n - n_j)\Delta s(n_j) \} \chi_j(n) \right). $$

We first estimate $\|A\|$, where $A$ denotes the identity operator. For that we have

$$ \|s - As\|^2 \leq \sum_{n=0}^{N-1} \sum_{j=1}^r |s(n) - s(n_j) - (n - n_j)\Delta s(n_j)|^2 \chi_j(n) $n=0, j=1 \leq \sum_{j=1}^r \sum_{n=l_{j-1}+1}^{l_j} |s(n) - s(n_j) - (n - n_j)\Delta s(n_j)|^2. \quad (12) $$

Let $t(n) = s(n) - s(n_j) - (n - n_j)\Delta s(n_j)$. Since $t(n_j) = t(n_j + 1) = 0$, we can write the second summation in (12) as $\sum_{n=l_{j-1}+1}^{n_j+1} + \sum_{n=n_j+1}^{l_j}$. Since $\Delta^2 t = \Delta^2 s$, Lemma 1 yields

$$ \sum_{n=l_{j-1}+1}^{n_j+1} |t(n)|^2 \leq K(n_j - l_{j-1} + 1) \sum_{n=l_{j-1}+1}^{n_j+1} |\Delta^2 s(n)|^2 $$

\[ \sum_{n=l_{j-1}+1}^{n_j+1} \sum_{n=l_{j-1}+1}^{n_j+1} |t(n)|^2 \leq K(n_j - l_{j-1} + 1) \sum_{n=l_{j-1}+1}^{n_j+1} |\Delta^2 s(n)|^2 \]
and
\[ \sum_{n=n_j}^{l_j} |t(n)|^2 \leq K(l_j - n_j + 1) \sum_{n=n_j}^{l_j-2} |\Delta^2 s(n)|^2. \]

But
\[ \max\{n_j - l_{j-1} + 1, l_j - n_j + 1\} \leq \lceil d/s \rceil + 1, \]
and \( K \) is a nondecreasing function; thus
\[ \|s - As\|^2 \leq K\left(\lceil d/2 \rceil + 1\right)\|\Delta^2 s\|^2 \leq K\left(\lceil d/2 \rceil + 1\right) \left(2 \sin \frac{\pi M}{N}\right)^4 \|s\|^2, \]
where the last inequality followed from Lemma 3.

Therefore, if \( \kappa_d = 4 \sqrt{K\left(\lceil d/2 \rceil + 1\right) \sin^2(\pi M/N)} < 1 \), then \( A \) is invertible by Lemma 2.

Now, to show (8) we proceed as follows:
\[
\|s\|^2 = \|A^{-1}As\|^2 \leq \|A^{-1}\|^2 \|As\|^2 \leq (1 - \|\text{Id} - A\|)^{-2} \|As\|^2
\]

But,
\[
\|As\|^2 \leq 2 \sum_{j=1}^{r} \mu_j |s(n_j)|^2 + \tilde{\mu_j} |\Delta s(n_j)|^2;
\]
thus the left inequality is proved.

For the right inequality, note first that the following was implicitly obtained from the estimation of \( \|s - As\| \):
\[
\left\| \sum_{j=1}^{r} \left( s - s(n_j) - (\cdot - n_j) \Delta s(n_j) \right) \chi_j \right\| \leq \kappa_d \|s\|. 
\]
(13)

Next, we claim that for \( j \in \mathbb{Z} \),
\[
\mu_j |s(n_j)|^2 + \tilde{\mu}_j |\Delta s(n_j)|^2 \leq \frac{1}{2 - \sqrt{3}} \sum_{n=l_{j-1} + 1}^{l_j} |s(n_j) + (n - n_j) \Delta s(n_j)|^2.
\]
For, denote bu $U_j$ and $V_j$ the matrices corresponding to the quadratic forms defined by the left and right sides of the above inequality, that is,

$$U_j = \begin{bmatrix} \mu_j & 0 \\ 0 & \tilde{\mu}_j \end{bmatrix} \quad \text{and} \quad V_j = \begin{bmatrix} \mu_j & \nu_j \\ \nu_j & \tilde{\mu}_j \end{bmatrix},$$

where $\nu_j = \sum_{i=l_{j-1}+1}^{l_j} (n - n_j)$. We would like to find a constant $k$, independent of $j$ such that the matrix

$$\begin{bmatrix} (k - 1)\mu_j & k\nu_j \\ k\nu_j & (k - 1)\tilde{\mu}_j \end{bmatrix}$$

is positive semidefinite. Hence the best $k$ should satisfy the equation

$$(k - 1)^2 \mu_j\tilde{\mu}_j - k^2\nu_j^2 = 0,$$

which has the solutions

$$\left(1 - \sqrt{\nu_j^2/\mu_j\tilde{\mu}_j}\right)^{-1}, \quad \left(1 + \sqrt{\nu_j^2/\mu_j\tilde{\mu}_j}\right)^{-1}.$$

Since $k \geq 1$ is necessary and since $\nu_j^2/\mu_j\tilde{\mu}_j \leq 3/4$ by Lemma 4, we can choose $k = 2/(2 - \sqrt{3})$ and our claim is proved. Thus,

$$\sum_{j=1}^{r} \left( \mu_j|s(n_j)|^2 + \tilde{\mu}_j|\Delta s(n_j)|^2 \right)$$

$$\leq \frac{2}{2 - \sqrt{3}} \sum_{j=1}^{r} \sum_{n=l_{j-1}+1}^{l_j} |s(n_j) + (n - n_j)\Delta s(n_j)|^2$$

$$- \frac{2}{2 - \sqrt{3}} \left\| \sum_{j=1}^{r} (s(n_j) + (\cdot - n_j)\Delta s(n_j))\chi_j \right\|^2$$

$$\leq \frac{2}{2 - \sqrt{3}} \left( \left\| \sum_{j=1}^{r} (s - s(n_j) - (\cdot - n_j)\Delta s(n_j))\chi_j \right\| + \|s\| \right)^2.$$

The right inequality in (8) now follows from (13).
Finally, the estimate (11) follows readily from the fact that

$$\langle \mathbf{T}_\mu \hat{s}, \hat{s} \rangle = \langle \mathbf{s}, \mathbf{s} \rangle = \sum_{j=1}^{r} \left( \mu_j |s(n_j)|^2 + \tilde{\mu}_j \Delta s(n_j) |^2 \right).$$  \hspace{1cm} (14)

The theorem is proved.

**REMARKS.**

1. The definition of the operator $S$ in (3) and the choice of the weights $\mu_j$ and $\tilde{\mu}_j$ in (4) are now justified by (14) and the estimate on the condition number of $T_\mu$.

2. Condition (9) essentially states that

$$d^2 \leq \frac{N^2}{4} \left( \frac{N}{\pi M} \right)^2$$

or

$$d < \frac{\sqrt{2\sqrt{3}} \cdot N}{\pi M}.$$ 

On the other hand the Nyquist condition states that $d < N/(2M + 1)$. But

$$\frac{\left( \sqrt{2\sqrt{3}} / \pi \right) (N/M)}{N/(2M + 1)} = \frac{2\sqrt{2\sqrt{3}}}{\pi} \left( 1 + \frac{1}{2M} \right) > 1.18$$

for all $M$; i.e., condition (9) allows gaps larger than the Nyquist gap. This is why the outlined method works better with larger gaps than previous methods (see Section 3 for some numerical comparisons).

Since inverting a general matrix is very expensive, we must use the special structure of $T_\mu$ to obtain an efficient reconstruction algorithm. As $T_\mu$ is positive definite, it is obvious that one can employ the conjugate gradient method to find good approximate solutions by iteration.

**THEOREM 2** (Weighted Difference-Conjugate Gradient Algorithm). Given the samples $\{s(n_j), s(n_j + 1)\}_{j=1}^{r}$, where $r > M + 1$, $M$ being the size of the spectrum and $0 \leq n_1 < n_1 + 1 < n_2 < n_2 + 1 < \cdots < n_r < n_r + 1 < N$. Set $n_0 = n_r + 1 - N$ and $n_{r+1} = n_1 + N$. Let the matrix $T_\mu = C + \cdots$. 
$D^*\tilde{\mathcal{G}}D$ be defined as in Proposition 1. Compute

$$b_k = N^{-1/2} \sum_{j=1}^{r} \mu_j s(n_j) e^{-2\pi i kn_j/N}$$

$$+ N^{-1/2} \left( e^{-2\pi i k/N} - 1 \right) \sum_{j=1}^{r} \tilde{\mu}_j \Delta s(n_j) e^{-2\pi i kn_j/N},$$

for $|k| \leq M$ and set $r_0 = q_0 = b \in \mathbb{C}^{2M+1}$, $a_0 = 0$. Compute iteratively for $n \geq 1$

$$a_n = a_{n-1} + \frac{\langle r_{n-1}, q_{n-1} \rangle}{\langle T_{\mu} q_{n-1}, q_{n-1} \rangle} q_{n-1},$$

$$r_n = r_{n-1} - \frac{\langle r_{n-1}, q_{n-1} \rangle}{\langle T_{\mu} q_{n-1}, q_{n-1} \rangle} T_{\mu} q_{n-1},$$

and

$$q_n = \frac{\langle r_n, T_{\mu} q_{n-1} \rangle}{\langle T_{\mu} q_{n-1}, q_{n-1} \rangle} q_{n-1}.$$

Then $a_n$ converges to the solution $\hat{s} \in \mathbb{C}^{2M+1}$ of the equation $T_{\mu} \hat{s} = b$ in at most $2M + 1$ iterations. The reconstruction of $s \in B_M$ is then given by $s(n) = N^{-1/2} \sum_{k=-M}^{M} \hat{s}(k) e^{2\pi i kn/N}$. Furthermore, if $\kappa_\delta < 1$ then the following error estimate holds for $n < 2M + 1$, and $t_n(m) = N^{-1/2} \sum_{k=-M}^{M} a_n(k) e^{2\pi i km/N}$:

$$\|s - t_n\|_T \leq 2 \left( \frac{\sqrt{\beta} - \sqrt{\alpha}}{\sqrt{\beta} + \sqrt{\alpha}} \right) \|s\|_T,$$

where $\|x\|_T = \langle x, T_{\mu} x \rangle^{1/2} = \sum_{j=1}^{r} (\mu_j |x(n_j)|^2 + \tilde{\mu}_j |\Delta x(n_j)|^2)$ is the $T_{\mu}$-norm of $x$.

The error estimate (15) follows from the fact that $\alpha \leq \lambda$ and $\Lambda \leq \beta$. Refer to [17, 18] for more on the conjugate gradient method.
Remark. In general the convergence of the CG is faster than 
\((\sqrt{\beta} - \sqrt{\alpha}) / (\sqrt{\beta} + \sqrt{\alpha})^n\) because \(\alpha\) and \(\beta\) are just estimates for the extreme eigenvalues and because the convergence also depends on the distribution of the singular values of \(T_w\).

The above method, in particular the choice of the sampling geometry, was inspired by a recent result in the theory of irregular sampling with derivatives of band-limited signals. We refer to [10] for the precise details.

3. IMPLEMENTATION AND NUMERICAL RESULTS

In this section we compare the numerical performance of the new method to previously known methods. For comparison we chose the adaptive weights Toeplitz (AWT) method developed by H. Feichtinger and K. Gröchenig [4] and the frame method or the standard method (ST). The frame method was used by Benedetto [11], Analoui and Marvasti [8] and others. The ideas go back to Duffin and Schaffer and seem to occur implicitly in several algorithms in the engineering literature. Both methods reduce the discrete problem to a Toeplitz system when developed in the frequency domain, a technique introduced by Gröchenig [6]. Those Toeplitz matrices are given below. Strohmer [19] did an extensive comparison of the numerical performance of these two methods. In general the results show that the AWT method is much more efficient than the standard method. For further details refer to [19, 20].

Now, suppose we are given pairs of sampling points

\(\{n_1, n_1 + 1, n_2, n_2 + 1, \ldots, n_r, n_r + 1\}\),

a subset of \(\{0, 1, \ldots, N - 1\}\). Note that this particular sampling set is just a highly nonuniformly spaced sampling points. For ease of notation, let us rename them as

\(\{x_1 = n_1, x_2 = n_1 + 1, x_3 = n_2, x_4, \ldots, x_{2r-1}, x_{2r}\}\).

Then the first method, AWT, generates a Toeplitz matrix \(T_w\) defined by

\[
(T_w)_{ik} = \frac{1}{N} \sum_{j=1}^{2r} w_j e^{2\pi i x_j (l-k)/N} \quad \text{for } |l|, |k| \leq M,
\]
where the weights are given by \( w_j = (x_{j+1} - x_{j-1})/2 \) for \( j = 1, \ldots, 2r \), with \( x_0 = x_{2r} - N \) and \( x_{2r+1} = x_1 + N \) by periodicity.

For the second method, ST, the derived Toeplitz matrix \( T \) is given by

\[
T_{lk} = \frac{1}{N} \sum_{j=1}^{2r} e^{-2\pi i j (l-k)/N} \quad \text{for } |l|, |k| \leq M.
\]

We compared the performance of these two methods to the one just developed in Section 2.2, namely the weighted difference method (WD), for which the matrix \( T_{\mu} \) is once again given by

\[
T_{\mu} = C + D^*\hat{C}D
\]
as described in Proposition 1.

An important numerical aspect of the Toeplitz structure is the fact that only one row or column is needed to be evaluated and all the other entries can be obtained from that vector. Moreover, the simple observation in Corollary 1 helps to compute the entries of all the Toeplitz matrices faster. Furthermore, the multiplication of a vector \( a \in \mathbb{C}^n \) with an \( n \times n \) Toeplitz matrix \( T \) can be done quickly. In fact \( T \) can be augmented to a \((2n-1) \times (2n-1)\) circulant matrix \( \tilde{T} \) and zeros are added to \( a \) to get a vector \( \tilde{b} \in \mathbb{C}^{2n-1} \), then \( \tilde{Tb} \) is identical to \( t * b \), the discrete convolution of the first column \( t \) or \( f \) with \( b \) and \( Ta \) can be obtained from the first \( n \) coordinates of \( \tilde{Tb} \). But we know that \( \tilde{t} \ast \tilde{b}(k) = \hat{t}(k)\hat{b}(k) \); hence \( t \ast b \) can be evaluated by inverse Fourier transform (see, e.g., [21]).

Now here are some numerical results. First we compare the condition numbers of the matrices generated by the three methods. Figure 1 shows the condition numbers (shown in logarithmic scale for clarity) as a function of the spectrum size \( M \). Here the condition number is the ratio of the largest and the smallest singular values of the matrices. In the experiment we choose a fixed sampling set consisting of \( r = 65 \) number of disjoint pairs of positive integers \( N = 512 \). Then we vary the values of the spectrum \( M \) from 1 to \( r - 1 \). The condition numbers of the three matrices \( T \), \( T_w \), and \( T_{\mu} \) are then computed for each \( M \). For this special sampling set Nyquist condition \( \max(n_{j+1} - n_j) \leq N/(2M + 1) \) is satisfied for \( M \leq 20 \). As \( M \) increases we expect those condition numbers to increase also. Furthermore, note that the theoretical conditions for positive definiteness of all the matrices are satisfied, namely \( 2r \geq 2M + 1 \) for \( T \) and \( T_w \) and \( r \geq M + 1 \) for \( T_{\mu} \). The result of our first experiment shows that as \( M \) gets closer to \( r - 1 \), the condition number of \( T_{\mu} \) is in general smaller. The difference is significant for a large range of
values of $M$, in particular when gaps larger than Nyquist occur. In [4] the authors claim that the particular choice of the weights $w_j$ in $T_w$ is optimal in the sense of minimizing its condition number. We can safely argue that for our sampling geometry, which is nothing but a highly irregular sampling sequence, and in the presence of larger gaps, the new method is much more stable.

In Fig. 2, we show the result of the following experiment: We consider a computer-generated real signal of length $n = 4096$. The number of sampling pairs is $r = 329$ (i.e., a total of 658 samples) and the size of the spectrum of the signal is $M = 88$, chosen so that Nyquist condition is satisfied. Then for the first 50 iterations the rate of convergence of the CG algorithm applied to each system is compared by computing the errors term $\|x - x_n\|_2/\|x\|_2$. The result shows that for the given sampling geometry the new method still performs a little better under Nyquist condition.

Figure 3 shows another comparison of the speed of convergence of the CG algorithm. A signal with a spectrum size $M = 200$ is sampled at the same sampling sequence as before. This means that now many gaps that are twice as large as the Nyquist gap occur. Clearly the convergence of the new method
is now by far faster than the others. Note that the reconstruction using WD-CG was completed after about 45 iterations while the other two methods needed at least 230 iterations to achieve an error of the order $10^{-13}$. This result emphasizes once again the better conditioning of $T_\mu$ in the presence of large gaps.

If we use the same techniques for matrix vector multiplication on both Toeplitz matrices in $T_\mu$, and on $T$ or $T_w$, it is clear that for a given number of iterations, the WD-CG algorithm requires at least twice the amount of computations used by either one of the other methods. However, this is compensated by its much faster rate of convergence.

All experiments were conducted on a SUN workstation and 486-based PCs and MATLAB was used for the implementation. The M-files that implement the WD-CG algorithm have been developed to extend to the MATLAB-toolbox for irregular sampling which was developed by the Numerical Harmonic Analysis Group of the University of Vienna, Austria, led by H. Feichtinger.
I thank Professor Karlheinz Gröchenig for his valuable suggestions that helped improving the paper.

REFERENCES


Received 13 December 1994; final manuscript accepted 7 August 1995