Recovery of a Sparse Signal When the Low Frequency Information is Missing

By

David L. Donoho

Philip B. Stark

Department of Statistics
University of California
Berkeley, CA 94720

Technical Report No. 179
June 1989

Research supported by NSF grant DMS 84-51753,
Postdoctoral Fellowship DMS 87-05843 and
by grants from Schlumberger-Doll Research from Western Geophysical

Department of Statistics
University of California
Berkeley, California

Recovery of a Sparse Signal when the Low Frequency Information is Missing.

David L. Donoho
Philip B. Stark
Department of Statistics
University of California
Berkeley, CA 94720

June 30, 1989

Abstract

We develop inequalities for the concentration of bandlimited functions to sets of small density. The L_2 inequality implies that wideband signals concentrated on sets of density $< 1/6$ can be reconstructed stably from low-cut filtered data containing noise, provided that the noise is small in L_2. The L_1 inequality implies that a bandlimited function corrupted by impulsive noise can be reconstructed perfectly, provided the noise is concentrated on a set of density $< 1/(2+\pi)$.

Key Words and Phrases. Large Sieve. Uncertainty Principle. Stable Recovery. Density. Bandlimited Functions. Entire Functions of Exponential Type.

Acknowledgements. This work was supported by National Science Foundation grant DMS 84-51753, Postdoctoral Fellowship DMS 87-05843, and by grants from Schlumberger-Doll Research and from Western Geophysical. The authors would like to thank Ytzhak Katznelson, Ken Larner and Ralphe Wiggins for helpful discussions.

1 Introduction

In an idealized model of exploration seismology one probes the Earth acoustically and interprets the reflected signal. For practical reasons, the source contains essentially no energy in the low frequency band 0-10Hz. However, for recovery of the acoustic impedance profile, the missing low frequency band is of considerable importance. Recently, researchers have developed nonlinear techniques which, in certain examples, are able to use bandlimited data to recover the original signals over the full range (0, Nyquist) (*Levy and Fullagar* [5], *Walker and Ulrych* [12]).

A naive application of information theory concepts might suggest that this is impossible: the missing frequency band was not measured, so information about it would seem to be unavailable. However, there is a piece of side information—the signal to be recovered is *sparse* (i.e. zero most of the time)—which all successful algorithms use in some way and which all the dramatic examples exhibit.

Santosa and Symes [11] came up with a mathematical proof that extreme sparsity allows recovery of the wideband signal in the noiseless case. *Donoho and Stark* [2] derived their result, as well as some stability results in the noisy case, from generalizations of the uncertainty principle of Fourier analysis. The basic phenomenon can be summarized as follows. *Suppose the set of missing frequencies is $W = [-\pi\Omega, \pi\Omega]$ and the set T supports the signal to be recovered. We do not know T a priori, but suppose we do know an a priori bound on the measure $|T|$. If this bound implies $\Omega|T| < 1/2$ then stable recovery of the full wideband signal is possible. In general the recovery method is nonlinear.*

This result, while it goes in the right direction, falls short of practical application. The sets T of interest in practice are those whose *density* is a priori small, but whose *total measure* $|T|$ is large. (Think of a long train of "events", each of brief duration, only a few events per unit time, but continuing persistently, so that the total duration of all events is great.)

In section 2 we show that when W is an interval of low frequencies $W = [-\pi\Omega, \pi\Omega]$, the density of T, rather than the total measure, controls the ability to stably recover the missing frequency band. Define the *maximum density* of T

$$\rho(T,\Omega) \equiv \Omega \sup_t \left| T \cap [t, t+\frac{1}{\Omega}] \right| .$$

We show that if $\rho(T,\Omega)$ is known a priori to be less than $1/6$, *i.e.* if T has small mass in every interval of length $1/\Omega$, then stable recovery is possible, with explicitly given bound on the noise amplification. The constants we get can certainly be improved (i.e. the $1/6$ could conceivably be replaced by something near to $1/2$), but at least the principle is established.

The result just stated depends on Theorems 1 and 2, which do not mention signal recovery at all. Instead, those theorems make statements about the fraction of a function's L_2 norm which can be concentrated on sets of low density. Such results have analogs in L_1-norm, which are developed in section 3.

The L_1 results also have applications in signal recovery. *Logan* [6] discovered an interesting phenomenon. Suppose we get noisy measurements of a bandlimited signal that is in L_1. We know nothing whatever about the noise, except that the measure of its support is small and it has finite L_1-norm. Then, Logan showed, we can recover the original bandlimited signal perfectly, without error, provided only that the support of the noise satisfies $|T|\Omega < 1/2$. The inequalities we develop in section 3 on L_1 concentration show that the same phenomenon occurs even if we only suppose that the density $\rho(T,\Omega) < 1/(2+\pi)$. In other words, the noise can be persistent, even supported on a set of infinite measure, yet the original bandlimited signal may be recovered perfectly. Again, the constant should be improved from $1/(2+\pi)$ to something larger, perhaps approaching $1/2$. (It cannot be better than $1/2$).

In section 4, we discuss versions of these results for discrete-time signals. In section 5 we discuss the possibility of improving the constants in these relations.

An interesting aspect of our approach is that we rely on versions of a family of inequalities developed in number theory which Montgomery [7] calls "the analytic principle of the Large Sieve". The large sieve is in some sense about the concentration of bandlimited functions to thin sets when the frequency space is a discrete set. As we show, a version of the large sieve due to Bombieri makes this connection transparent. Considerable work on best constants in the large sieve has been done. This makes our L_2 result on concentration of bandlimited functions, via the large sieve, somewhat sharper for our purposes than the classical work of Plancherel and Polya, Boas, and Nikolskii, which use complex variable techniques.

2 Concentration in L_2 Norm

We briefly discuss signal recovery. Suppose we observe the signal r defined by

$$r = (I - P_W)s + n$$

where r, s, and n are L_2 functions, I is the identity operator, and P_W is the operator that bandlimits to W. We would like to have measured s, but instead we get the noisy, data r missing frequencies in W. Suppose that $T = \text{supp}(s)$ belongs *a priori* to the class $\mathcal{T}$, and let $\mathcal{T}_2$ denote

$$\mathcal{T}_2 \equiv \{T_1 \cup T_2 : T_i \in \mathcal{T}\}.$$

Define

$$\lambda_0(T, \Omega) \equiv ||P_T P_W||,$$

where P_T is the operator $(P_T f) = f \cdot 1_{t \in T}$ that sets f to zero off T. *Donoho and Stark* [2] show that if

$$\Lambda(\mathcal{T}_2, \Omega) \equiv \sup_{T \in \mathcal{T}_2} \lambda_0(T, \Omega) < 1,$$

then stable recovery of s from r is possible—even though T is unknown (except for $T \in \mathcal{T}$)—with stability coefficient $2(1 - \Lambda)^{-1}$. The uncertainty principle they develop gives, in the notation of this paper,

$$\lambda_0(T, \Omega) \leq \sqrt{\Omega |T|};$$

therefore, if $\mathcal{T} = \{T : |T| < \frac{l^2}{\Omega}\}$, we have $\Lambda(\mathcal{T}_2, \Omega) \leq l < 1$, and the result mentioned in the introduction.

We now develop bounds on $\Lambda(\mathcal{T}, \Omega)$ when $W = [-\pi\Omega, \pi\Omega]$ is an interval of low frequencies, for sets T which are of low density. The ideas are clearest for thin, periodic sets.

Theorem 1 *Let T be a periodic set: $T = T + 1/\Omega$. Then*

$$\lambda_0(T, \Omega) = \sqrt{\rho(T, \Omega)}. \tag{1}$$

Proof. Let $B_2(\pi\Omega)$ be the set of bandlimited functions with finite L_2 norm, *i.e.*

$$B_2(\pi\Omega) = \left\{ f : f = \int_{-\pi\Omega}^{\pi\Omega} e^{i\omega t} \hat{f}(\omega) d\omega \text{ with } ||\hat{f}||_2 < \infty \right\}.$$

Now

$$\begin{aligned} ||P_T P_W|| &= \sup_{f \in L_2} \frac{||P_T P_W f||}{||f||} \\ &= \sup_{f \in B_2(\pi\Omega)} \frac{||P_T P_W f||}{||P_W f||} \quad (\text{as } ||f||_2 \geq ||P_W f||_2) \\ &= \sup_{f \in B_2(\pi\Omega)} \frac{||P_T f||}{||f||}. \end{aligned} \tag{2}$$

Now the rescaling

$$\begin{aligned} \tilde{f}(t) &= f\left(\frac{t}{\Omega}\right) \Omega^{1/2} \\ \tilde{T} &= T/\Omega \end{aligned}$$

maps $B_2(\pi\Omega)$ in a one-to-one fashion onto $B_2(\pi)$, preserves concentration:

$$\frac{\int_{\tilde{T}} |\tilde{f}|^2}{\int |\tilde{f}|^2} = \frac{\int_T |f|^2}{\int |f|^2},$$

and preserves density: $\rho(T, \Omega) = \rho(\tilde{T}, 1)$. Therefore if the result is true for $\Omega = 1$, it is true for all positive Ω.

For $f \in B_2(\pi)$, the *sampling theorem* says that

$$||f||^2 = \sum_{k=-\infty}^{\infty} f^2(k+h) \tag{3}$$

for any $h \in \mathbf{R}$. Thus

$$\begin{aligned} \int_T f^2 &= \sum_{k=-\infty}^{\infty} \int_{T \cap [k, k+1]} f^2(k+h) dh \\ &= \int_{T \cap [0, 0+1]} \sum_{k=-\infty}^{\infty} f^2(k+h) dh \quad (\text{Fubini}) \\ &= \int_{T \cap [0,1]} ||f||^2 dh \quad (\text{by } 3) \\ &= ||f||^2 \, |T \cap [0,1]| \, . \end{aligned} \tag{4}$$

Hence
$$\frac{||P_T f||^2}{||f||^2} = |T \cap [t, t+1]| = \rho(T, 1)$$
for every $f \in B_2(\pi)$, and this establishes the proof.

Theorem 1 gives simple examples of sets T of infinite measure on which no W-bandlimited function can be highly concentrated –simply pick the density of T small.

For aperiodic T, we rely on the large sieve. See the excellent article of *H.L. Mongtomery* [7] for a general survey, and, on page 562, reference to the following:

Bombieri's Large Sieve Inequality. *Let* $S(\alpha) = \sum_{k=m+1}^{n+1} a_k e^{2\pi i k\alpha}$ *be a trigonometric polynomial of degree* n *and period* 1. *Let* μ *be a measure supported on* $[0,1]$. *Then*

$$\int_0^1 |S(\alpha)|^2 d\mu \leq (n + 2\delta^{-1}) (\sup_\alpha \int_\alpha^{\alpha+\delta} d\mu) \sum_{k=m+1}^{m+n} |a_k|^2 \tag{5}$$

The large sieve may be viewed as a substitute for the sampling theorem (3) for non-equispaced sequences, as we explain later. With it, we prove

Theorem 2 *Let* T *be arbitrary measurable.*

$$\lambda_0(T, \Omega) \leq \sqrt{2 + 1/c}\,\sqrt{\rho(T, \Omega)} \tag{6}$$

Proof. As in Theorem 1, we may prove the result just in the case $\Omega = 1$. Assume that $f \in B_2(\pi)$ is not identically zero and has $\hat{f}(\omega)$ uniformly continuous on $[-\pi, \pi]$. We will show that

$$\frac{\int_T |f|^2}{\int |f|^2} \leq (2 + 1/c)\ \rho(T, c\Omega). \tag{7}$$

An easy approximation argument extends this inequality to all $f \in B_2(\pi)$, hence (6).

Define
$$f_N(t) = \frac{1}{2N} \sum_{k=-N}^{N} \hat{f}(\frac{\pi k}{N}) e^{i\pi \frac{k}{N} t}. \tag{8}$$

We claim that f_N is an excellent approximation to f on $[-N,N]$, and will show later that

$$\frac{\int_{T\cap[-N,N]}|f_N(t)|^2dt}{\int_{[-N,N]}|f_N(t)|^2dt} \to \frac{\int_T|f(t)|^2dt}{\int|f(t)|^2dt}. \tag{9}$$

Note that $f_N(t)$ is a trigonometric polynomial with period $2N$. Also, $S(\alpha) = f_N(-N+2N\alpha)$ is a trigonometric polynomial of degree $n = 2N+1$ and period 1. Put $A = (T+N)/2N \cap [0,1]$; then

$$\frac{\int_A|S(\alpha)|^2}{\int_0^1|S(\alpha)|^2} = \frac{\int_{T\cap[-N,N]}|f_N(t)|^2dt}{\int_{[-N,N]}|f_N(t)|^2dt}.$$

We use (5) to bound the left hand side. We have $S(\alpha) = \sum_{m+1}^{m+n} a_k e^{2\pi ik\alpha}$ with $a_k = \hat{f}(\frac{\pi k}{N})(-1)^k$, $m = -N-1$. Put $\mu(\alpha,\alpha+\delta] = |A\cap(\alpha,\alpha+\delta]|$. Bombieri's inequality gives

$$\int_A|S(\alpha)|^2 \le (n+2\delta^{-1})\,(\sup_\alpha|A\cap(\alpha,\alpha+\delta]|)\sum|a_k|^2.$$

By orthogonality of $e^{2\pi ik\alpha}$ and $e^{2\pi il\alpha}$ on $[0,1]$ when $k\neq l$, $\int_0^1|S(\alpha)|^2d\alpha = \sum|a_k|^2$. Hence

$$\frac{\int_A|S(\alpha)|^2}{\int|S(\alpha)|^2} \le (n+2\delta^{-1})\delta\rho(A,\delta^{-1}). \tag{10}$$

Put $c^{-1} = 2N\delta$, and note that

$$\rho(T\cap[-N,N],c) = \rho(A,\delta^{-1}).$$

Then $n\delta = (2N+1)\delta = c^{-1}+\delta$, and

$$(n+2\delta^{-1})\delta\rho(A,\delta^{-1}) \le (c^{-1}(1+\frac{1}{2N})+2)\rho(T,c). \tag{11}$$

The theorem follows from (10)-(11) and our claim (9).

To prove (9), note that $\hat{f}$ is uniformly continuous, so the Riemann Sum $\frac{1}{2N}\sum_{-N}^{N}\hat{f}(\frac{\pi k}{N})e^{i\pi\frac{k}{N}t}$ converges to the integral $\frac{1}{2\pi}\int_{-\pi}^{\pi}\hat{f}(\omega)e^{i\omega t}d\omega$, i.e.

$$f_N(t)\to f(t) \text{ for each } t. \tag{12}$$

Similarly, $|\hat{f}(\omega)|^2$ is uniformly continuous on $[-\pi,\pi]$ and so

$$\frac{1}{2N}\sum_{-N}^{N}|\hat{f}(\frac{\pi k}{N})|^2 \to \frac{1}{2\pi}\int_{-\pi}^{\pi}|\hat{f}(\omega)|^2d\omega \tag{13}$$

By Parseval's relation, this implies

$$\int_{-N}^{N} |f_N(t)|^2 dt \to \int_{-\infty}^{\infty} |f(t)|^2 dt \tag{14}$$

Define $g_N(t) = |f_N(t)|^2 1_{|t|\le N}$ and $g(t) = |f(t)|^2$. Then (12) says that $g_N(t) \to g(t)$ pointwise, and (14) says that $\int g_N \to \int g$. As $g_N \ge 0$ and $g \ge 0$, Sheffe's lemma says $\int_{-\infty}^{\infty} |g_N - g| \to 0$. In particular, $\int_T g_N \to \int_T g$. Combining this with (14) and the fact that g is not identically zero gives (9).

Putting $c = 1$ in (6) gives

Corollary 3 *Let $\mathcal{T}$ be the class of sets T with $\rho(T, \Omega) \le l < 1/6$. If it is known a priori that supp$(s) \in \mathcal{T}$ then stable recovery of s from $r = (I - P_W)s + n$ is possible. There is a nonlinear mapping $\tilde{s}(r)$ with*

$$||s - \tilde{s}(r)|| \le 2(1 - 6l)^{-1/2}||n|| \tag{15}$$

Putting $c > 1$ shows that stable recovery is possible with even higher density than $1/6$, provided the density is measured on sufficiently short intervals.

There are many ways to prove results analogous to Theorem 2, only with worse constants. Our early attempts in this direction used inequalities on entire functions developed by *Plancherel and Polya* [10] and by *Boas* [1]. These inequalities allow one to prove (6) in the case $c = 1$ with the constants 4.23794 and $(7 + \frac{4}{\pi^2})$ rather the constant $\sqrt{3}$. Our next attempts used an inequality of Duffin-Shaeffer from the theory of nonharmonic Fourier series, which allowed us to get the constant $1 + \pi$. We later learned of work by *Nikolskii* [8] which we used to give the same constant, $1 + \pi$, but by a different and perhaps simpler argument. Our best result not using the large sieve is based on an idea like that in section 3 below, which allows us to get the constant $1 + \frac{\pi}{\sqrt{3}}$. This is still considerably larger than $\sqrt{3}$ for our purposes.

The reader may wonder why the large sieve is useful here. Consider another version, see [7], Theorem 3

Standard large sieve inequality. *Let (α_i) be a sequence of points in $[0, 1]$ which are well spaced: $|\alpha_i - \alpha_j| > \delta$ for $i \ne j$. Let $S(\alpha)$ be a trigonometric polynomial of degree n and period 1. Then*

$$\sum |S(\alpha_i)|^2 \le (n - 1 + \delta^{-1}) \sum |a_k|^2. \tag{16}$$

This inequality has a long history - see Montgomery [7]. This particular form, with coefficient $n-1+\delta^{-1}$, is due to Selberg. It is closely related to Bombieri's inequality (5), and in fact easily implies

$$\int |S(\alpha)|^2 d\mu \leq 2(n-1+\delta^{-1})\,(\sup_\alpha \int_\alpha^{\alpha+\delta} d\mu) \sum |a_k|^2 \tag{17}$$

which is (5) to within a factor of 2.

Equations (16) and (17) allow us to see that Theorem 2 really does have the same logical structure as Theorem 1. If the α_i were equally spaced in (16), then we would have, by a discrete form of the sampling theorem (3),

$$\sum_{i=0}^{n-1} |S(\alpha_i)|^2 = n \sum_{k=m+1}^{m+n} |a_k|^2. \tag{18}$$

Hence the large sieve (16) says that the sampling theorem holds approximately, for well-spaced sampling points.

In fact, the derivation of (17) from (16) will seem familiar to the reader who has studied Theorem 1. Suppose that α_i is the point at which $|S(\alpha)|$ attains its maximum value in the interval $[\delta i, \delta(i+1)]$. Then

$$\begin{aligned} \int |S(\alpha)|^2 d\mu &\leq \sum_i |S(\alpha_i)|^2 \mu(\delta i, \delta(i+1)] & (19)\\ &\leq \left(\sum_i |S(\alpha_i)|^2\right) \max_i \mu(\delta i, \delta(i+1)] & (20) \end{aligned}$$

Now $\alpha_{i+2}-\alpha_i > \delta$ and so by (16)

$$\sum_{i\ even} |S(\alpha_i)|^2 \leq (n-1+\delta^{-1}) \sum |a_k|^2 \tag{21}$$

$$\sum_{i\ odd} |S(\alpha_i)|^2 \leq (n-1+\delta^{-1}) \sum |a_k|^2. \tag{22}$$

Combining (19) and (21)-(22) gives (17). Just as in Theorem 1, the key ideas are a 'sampling theorem' and the partitioning of the range into equal subintervals.

While similar in form to (5), (17) only gives $\lambda_0(T,\Omega) \leq \sqrt{2+2/c}\,\sqrt{\rho(T,c\Omega)}$. For moderate c this is better than any non-sieve result we have been able to get; e.g. for $c=1$ it gives the constant $2 < (1+\frac{\pi}{\sqrt{3}})$. But when $c=1$ (17) leads to a requirement of density below 1/8 for stable recovery, whereas the inequality (5) leads to 1/6.

3 Concentration in L_1 norm

We turn to results for L_1 concentration. Let $B_1(\pi\Omega)$ denote the entire functions of type $\pi\Omega$ that are in L_1 on the real axis. Then define

$$\mu_0(T,\Omega) \equiv \sup_{f\in B_1(\pi\Omega)} \frac{||P_T f||_1}{||f||_1}.$$

and also

$$M(\mathcal{T},\Omega) \equiv \sup_{t\in\mathcal{T}} \mu_0(T,\Omega)$$

We briefly describe an application of M. Suppose we observe a signal $s \in B_1(\pi\Omega)$ with impulsive noise, so that

$$r = s + n$$

and we know *a priori* that the noise is zero except on a set T which is unknown to us, but which is known to have small measure. Logan [6] showed that if $\Omega|T| < 1/2$ then s can be reconstructed *perfectly* from r provided only that $||n||_1 < \infty$. See [2] for extensions to the case where W is an arbitrary passband rather than an interval.

Actually, when W is an interval of length $2\pi\Omega$, Logan's phenomenon is implied by the inequality $M(\mathcal{T},\Omega) < 1/2$. Thus bounds on $M(\mathcal{T},\Omega)$ using $\rho(T,\Omega)$ would be a considerable improvement on the condition $\Omega|T| < 1/2$. We begin with a result for periodic sets.

Theorem 4 *Let $T = T + \frac{1}{\Omega}$. Then*

$$\mu_0(T,\Omega) \le 2\,\rho(T,\Omega) \tag{23}$$

Proof. Suppose we have a function $g \in L_1$ with the *reproducing property*: $g * f = f$, whenever $f \in B_1(\pi)$. Any function in L_1 with $\hat{g}(\omega) = 1$ for $\omega \in [-\pi,\pi]$ has the reproducing property. Then

$$\begin{aligned}
\int_T |f(t)|dt &= \int_T \left| \int g(t-u)f(u)du \right| dt \\
&\le \int |f(u)| \int_T |g(t-u)|dtdu \\
&\le ||f||_1 \sup_u \int_T |g(t-u)|dt.
\end{aligned}$$

Now

$$\begin{aligned}
\int_T |g(t-u)|dt &= \sum_k \int_{T\cap[k,k+1]} |g(t-u)|dt \\
&= \int_{T\cap[0,1]} \sum_k |g(k+h-u)|dh \\
&\le |T\cap[0,1]| \sup_h \sum_k |g(k+h-u)| \\
&= \rho(T,1)\, ||g||_{\infty,1},
\end{aligned}$$

where we define $||g||_{\infty,1} = \sup_h \sum_k |g(k+h)|$.

Therefore, (23) is implied by the assertion that for some g with the reproducing property, $||g||_{\infty,1} \le 2$. This is proved in the following lemma.

Lemma 5 *Let* $g(t) = 2\frac{\sin^2(\pi t)}{(\pi t)^2} - 2\frac{\sin^2(\frac{\pi t}{2})}{(\pi t)^2}$. *Then* g *has the reproducing property for* $B_1(\pi)$ *and* $||g||_{\infty,1} \le 2$.

Proof. Note that $\hat{g}(\omega) = \Delta_{2\pi}(\omega) - \Delta_\pi(\omega)$, where

$$\Delta_{2\pi}(\omega) \equiv (2 - \frac{|\omega|}{\pi})_+$$

$$\Delta_\pi(\omega) \equiv (1 - \frac{|\omega|}{\pi})_+$$

are triangular functions of heights 2 and 1, with supports $[-2\pi, 2\pi]$ and $[-\pi, \pi]$ respectively. (This may be checked by noting that each of these functions can be expressed as a convolution of boxcars,

$$\Delta_{2\pi}(\omega) = \frac{1}{\pi}\chi_\pi \star \chi_\pi$$

$$\Delta_\pi(\omega) = \frac{1}{\pi}\chi_{\pi/2} \star \chi_{\pi/2}$$

where $\chi_{\pi/2}(\omega) = I_{\{|\omega|\le\pi/2\}}$ *etc.* One then uses the formulas for the transform of a boxcar.)

From this representation, the reproducing property is immediate, since the graph of $\hat{g}$ is the difference of two triangles, and hence is trapezoidal, with height 1 on $[-\pi, \pi]$.

It remains to check the size of the norm.

$$\sum_k |g(k+h)| = \frac{2}{\pi^2}\sum \frac{1}{(k+h)^2} a_{k\,mod\,2},\, h$$

where $a_{k,h} = |sin^2(\pi(k+h)) - sin^2(\frac{\pi}{2}(k+h))|$. Using the identity

$$\frac{1}{\pi^2}\sum_{k=-\infty}^{\infty}(x-k)^{-2} = cosec^2(\pi x)$$

(Gradshteyn and Rhyzhik 1.422.4.), we have that

$$\begin{aligned}\frac{2}{\pi^2}\sum \frac{1}{(k+h)^2} a_{k\,mod\,2},\, h &= \frac{1}{2}\sum_{k=0}^{1} a_{k,h} cosec^2(\frac{\pi}{2}(-k-h)) \\ &= \frac{1}{2}\sum_{k=0}^{1} |\frac{sin^2(\pi(k+h))}{sin^2(\frac{\pi}{2}(k+h))} - 1|.\end{aligned}$$

Experimenting on the computer, the expression clearly has its maximum at $h = 0$ where it takes the value 2.

Theorem 4 shows that there are sets T of infinite measure and density nearly 1/4 for which Logan's Phenomenon occurs. Thus Logan's condition $\Omega|T| < 1/2$ does not exhaust the occasions when this interesting phenomenon happens.

For aperiodic T, we again turn to large sieve ideas. *Montgomery* [7] gives a large sieve inequality developed by P.X. Gallagher using the following Sobolev-Type inequalities.

Lemma 6 *Let f and f' be continuous on $[0,1]$. Then*

$$|f(t)| \leq \int_0^1 |f| + \int_0^1 |f'| \textit{ for all } t \in [0,1] \tag{24}$$

$$|f(1/2)| \leq \int_0^1 |f| + \frac{1}{2}\int_0^1 |f'| \tag{25}$$

The proof is based on the identity

$$f(t) = \int_0^1 f(u)du + \int_0^t uf\prime(u)du + \int_t^1 (1-u)f\prime(u)du. \tag{26}$$

These inequalities easily allow us to establish that, even if T is aperiodic,

$$\mu_0(T,\Omega) \leq (1+\pi)\rho(T,\Omega). \tag{27}$$

Let α_k be a maximizer of $|f(\alpha)|$ for $\alpha \in [k, k+1]$. Then $\int_T |f| \leq \sum |f(\alpha_k)||T \cap [k, k+1]| \leq \rho(T,1) \sum_k |f(\alpha_k)|$. By the lemma,

$$\begin{aligned} \sum_k |f(\alpha_k)| &\leq \sum_k \int_k^{k+1} |f| + \sum_k \int_k^{k+1} |f'| \\ &= ||f||_1 + ||f'||_1. \end{aligned}$$

S.M. Nikolskii [8] has shown that if $f \in B_1(\pi\Omega)$ then

$$||f'|| \leq \pi\Omega ||f||_1 \tag{28}$$

and, using this we get $\int_T |f| \leq \rho(T,1)(1+\pi)||f||_1$ – i.e. (27).

This simple proof gives a much better constant than an approach based on standard inequalities for entire functions. The results of Plancherel and Polya and of Boas, give the constants 6.18 and $7+\frac{4}{\pi^2}$. The particular constant in (27) could, however, be established using work of Nikolskii [8], who used an idea slightly different from (24).

Inspection of (24)-(25) might suggest that (24) gives away a factor of 2 on the $\int |f'|$ term, unnecessarily. By using (26) and an idea similar to the proof of Theorem 4, we are able to get the best constant known to us.

Theorem 7

$$\mu_0(T,\Omega) \leq (1+\frac{\pi}{2})\rho(T,\Omega) \tag{29}$$

Proof. As before, we assume $\Omega = 1$. Let $t \in [0,1]$; by (26), if $t \in [h, h+1)$,

$$f(t) = \int_h^{h+1} f(u)du + \int_h^{h+1} f'\zeta_{t,h}(u)du$$

and so if we average over $h \in [t-1, t)$ we get

$$f(t) = Ave_h \int_h^{h+1} f(u)du + Ave_h \int_h^{h+1} f'\zeta_{t,h}(u)du$$

Defining

$$\begin{aligned}\eta_1(t;u) &= \int_{t-1}^{t} 1_{[h,h+1]}(u)dh \\ \eta_2(t;u) &= \int_{t-1}^{t} \zeta_{t,h}(u)1_{[h,h+1]}(u)dh\end{aligned}$$

we get, by a calculation,

$$\begin{aligned}\eta_1(t;u) &= (1-|t-u|)_+ \\ \eta_2(t;u) &= (1-|t-u|)_+^2/2\end{aligned}$$

and the 'reproducing identity'

$$f(t) = \int \eta_1(t;u)f(u)du + \int \eta_2(t;u)f'(u)du.$$

Actually, using the formulas for η_1 and η_2 as definitions, we see that the identity is valid for all t.

We have

$$\begin{aligned}\int_T |f| &\leq \int_T |\int f(u)\eta_1(t;u)du|dt + \int_T |\int f'(u)\eta_2(t;u)du|dt \\ &\leq \int |f(u)| \int_T \eta_1(t;u)dtdu + \int |f'(u)| \int_T \eta_2(t;u)dtdu \\ &\leq ||f||_1 \sup_u \int_T \eta_1(t;u)dt + ||f'||_1 \sup_u \int_T \eta_2(t;u)dt \\ &\leq ||f||_1 \left(\sup_u \int_T \eta_1(t;u)dt + \pi \sup_u \int_T \eta_2(t;u)dt\right).\end{aligned}$$

In the last step we have used (28). The theorem now follows by applying the following lemma, with $d\mu(t) = 1_T dt$.

Lemma 8 *Suppose that μ is a positive sigma-finite measure, and that*

$$\sup_t \mu(t, t+1) \leq \rho. \tag{30}$$

Then

$$\int \eta_1(t;u)d\mu(t) \leq \rho \text{ for all } u \tag{31}$$

$$\int \eta_2(t;u)d\mu(t) \leq \rho/2 \text{ for all } u \tag{32}$$

Proof. The two statements being proved similarly, we prove the second one, which is more surprising. As $\eta_2(t;u)$ just depends on $t-u$, and as the condition on μ is translation invariant, it is enough to prove (32) for $u=1$. Put $\eta(t) = \eta_2(t;1)$ for short. Now η vanishes for t outside $[0,2]$, and is continuous. Therefore, integration by parts gives

$$\int_{-\infty}^{\infty} \eta d\mu = -\int_0^2 \mu d\eta$$

where we take $\mu(\alpha) \equiv \mu[0,\alpha]$. Now

$$d\eta(t) = \begin{cases} t\,dt & t \leq 1 \\ -(2-t)\,dt & t > 1 \end{cases}$$

so

$$\begin{aligned} \int_0^2 \eta d\mu &= \int_0^1 \mu(t)(-t)dt + \int_0^1 (1-t)\mu(t+1)dt \\ &= \int_0^1 (1-2t)\mu(t)dt + \int_0^1 (1-t)(\mu(t+1)-\mu(t))dt \end{aligned}$$

Now μ is nondecreasing. Hence

$$\begin{aligned} \int_0^1 (1-2t)\mu(t)dt &\leq \int_0^{1/2} (1-2t)\mu(1/2)dt + \int_{1/2}^1 (1-2t)\mu(1/2)dt \\ &= 0, \end{aligned}$$

and

$$\begin{aligned} \int \eta d\mu &\leq \int_0^1 (1-t)(\mu(t+1)-\mu(t))dt \\ &\leq \rho \int_0^1 (1-t)dt = \rho/2. \end{aligned}$$

Before proceeding, we mention that we only used the fact $f \in B_(\pi)$ in the above proof for the inequality $||f'||_1 \le \pi||f||_1$. If we had assumed instead that $f \in B_1(\pi\Omega)$ we would have arrived at the inequality $\mu_0(T,\Omega) \le (1 + \frac{\pi\Omega}{2})\rho(T,1)$. By a rescaling argument, this proves: **A Sieve-Like Inequality for** B_1. *Let* $f \in B_1(\pi\Omega)$. *Then*

$$\int |f| d\mu \le (c^{-1} + \pi\Omega/2) \left(\sup_t \int_t^{t+c} d\mu \right) ||f||_1. \tag{33}$$

The parallel with Bombieri's inequality (5) should be apparent, with c playing the same role as δ and $\pi\Omega/2$ the same role as n.

Corollary 9 *Suppose that* $r = s + n$, $s \in B_1(\pi\Omega)$, *and that supp*$(n) = T$ *has density* $\rho(T, c\Omega) < \frac{1}{2+\pi/c}$. *Then the minimum* L_1*-recovery technique recovers the signal* s *perfectly.*

Thus, fixing a condition on the density over intervals of length $(c\Omega)^{-1}$, with c large, we can get perfect recovery of the signal with noise densities approaching $1/2$ on short intervals.

4 Discrete Time

For some readers, such as those interested in the geophysical prospecting problem mentioned in the introduction, it might be useful to have analogs of the above results for discrete time. In this section we describe two variants.

4.1 The l_2 setting

Suppose that $r = (r_t, t = 0, \ldots, N-1)$ is a measured discrete signal. It contains noise and is missing low frequency information; thus $r = (I - P_W)s + n$ where $n = (n_t, t = 0, \ldots, N-1)$ is a noise sequence, and P_W is a circular bandlimiting operator, the matrix that operates as least squares projector onto the span of the sinusoids with frequencies in $W = \{\frac{2\pi j}{N} : \frac{-K}{2} \leq j \leq \frac{K}{2}\}$; see [2] for details.

Defining, in the natural way, $\lambda_0(T, K)$ and $\Lambda(\mathcal{T}, K)$, by arguments in [2] we have that $\Lambda(\mathcal{T}_2, K) < 1$ implies that stable recovery of s from r is possible. If, for this setting we define the discrete density

$$\rho(t, \Omega) = \Omega \sup_t \#(T \cap [t, t + \Omega^{-1}]) \tag{34}$$

then we can use the large sieve (5) to get

Theorem 10

$$\lambda_0(T, K) \leq \sqrt{3 + 1/K} \sqrt{\rho(T, \frac{K}{N})} \tag{35}$$

The argument is as follows. If $(x_t, t = 0, \ldots, N-1)$ is a discrete sequence bandlimited to W, then, for appropriate coefficients (a_k),

$$x_t = \sum_{k=-K/2}^{K/2} a_k e^{2\pi i k t/N}$$

Put $n = K + 1$, $m = -K/2 - 1$, etc. Then, with $S(\alpha)$ defined as in (5), $S(\frac{t}{N}) = x_t$. Let ν_α denote the unit Dirac mass at α, and put $\mu = \sum_{t \in T} \nu_{t/N}$. Then

$$\int |S(\alpha)|^2 d\mu = \sum_{t \in T} |x_t|^2$$

and

$$\sup_\alpha \int_\alpha^{\alpha+\delta} d\mu = \sup_t \#(T \cap [t, t + N\delta]).$$

By Parseval's relation for the finite discrete Fourier Transform, $\sum_0^{N-1} |x_t|^2 = N \sum_k |a_k|^2$. Using these with (5) gives

$$\frac{\sum_{t \in T} |x_t|^2}{\sum_0^{N-1} |x_t|^2} \leq (K + 1 + 2\delta^{-1}) \frac{1}{N} \sup_t \#(T \cap [t, t + N\delta])$$

Choosing $\delta^{-1} = K$ gives

$$\lambda_0(T, K) \leq (3 + 1/K)\, \rho(T, \frac{K}{N})$$

and completes the proof.

It follows that if $\mathcal{T}$ is the class of discrete sets with density $< l < 1/6$, stable recovery of the missing frequency band $j \in \{\frac{-K}{2}, \ldots, \frac{K}{2}\}$ is possible.

For this problem, of course, $\lambda_0(T, K)$ is the top singular value of a finite matrix, and the computer may be used to explore the sharpness of (35). Figure 1 displays the results of calculating λ_0 for a few hundred different combinations of T, K. The display shows that the inequality (35) is somewhat pessimistic.

4.2 The l_1 setting

For variety, we consider a different discrete time setting, with the time index being all integers rather than the integers mod N. Let $b_1(\pi\Omega)$ be the set of discrete bandlimited sequences: sequences in l_1 whose Fourier transform $\hat{X}(\omega) = \sum_{t=-\infty}^{\infty} x_t e^{-i\omega t}$ vanishes for ω outside of $[-\pi\Omega, \pi\Omega]$. Here we must have the bandlimit $\Omega < 1$; in fact $\Omega << 1/2$ in order for the setting to be interesting. We continue to use the definition (34) for the discrete density.

Theorem 11 *Pick $\delta > 0$ so that $\delta \geq 2\Omega$ and δ^{-1} is an integer. Then if* $(x_t) \in b_1(\pi\Omega)$

$$\frac{\sum_{t \in T} |x_t|}{\sum_{-\infty}^{\infty} |x_t|} \leq \frac{3}{2}(1 + \frac{\pi\delta}{4\Omega})\rho(T, Omega) \tag{36}$$

Proof. Under the hypotheses of the theorem, we may use Lemma 12 to show there exists $f \in B_1(\frac{\pi}{2})$ with $f(\delta t) = x_t$. Then

$$\sum_{-\infty}^{\infty} |x_t| = \sum |f(\delta t)| = \sum_{j=1}^{\delta^{-1}} \sum_{k=-\infty}^{\infty} |f(k + \delta j)|.$$

Now lemma 12 below establishes an analog (39) of the sampling theorem for l_1. Putting $h = \delta j$, we conclude that

$$\sum_{-\infty}^{\infty} |x_t| \geq \frac{2}{3}\delta^{-1} \int |f(t)| \tag{37}$$

Now

$$\sum |x_t| 1_T(t) = \sum |f(t\delta)| 1_T(t) = \int |f(u)| d\mu(u)$$

with $\mu[u, u + c] \equiv \#(\delta T \cap [u, u + c])$. Also,

$$\sup_u \int_u^{u+c} d\mu = (c/\delta)\rho(T, \delta/c).$$

Using the result (33) with (37) we get

$$\begin{aligned} \sum_{t \in T} |x_t| &= \int |f(u)| d\mu(u) \\ &\leq (c^{-1} + \frac{\pi}{4})(c/\delta)\rho(T, \delta/c) \int |f(t)| \\ &\leq \frac{3}{2}(1 + \frac{\pi c}{4})\rho(T, \delta/c) \sum_{-\infty}^{\infty} |x_t| \end{aligned}$$

The theorem follows upon setting $c = \delta/\Omega$.

It remains to prove the following:

Lemma 12 *If* $f \in B_1(\pi/2)$ *then*

$$2/3 \sum |f(k)| \leq \sum |f(k + h)| \leq 3/2 \sum |f(k)|, \quad h \in [-1, 1]. \tag{38}$$

$$2/3 ||f||_1 \leq \sum |f(k + h)| \leq 3/2 ||f||_1. \tag{39}$$

Proof. We only prove the first inequality; the second follows immediately. If $f \in B_1(\frac{\pi}{2})$, then $\hat{f}$ is a continuous function with support in $[-\pi/2, \pi/2]$. Note that $(f(k))$ is the sequence of Fourier coefficients of $\hat{f}$, viewed as a function on $[-\pi, \pi]$, and that $(f(k+h))$ is the sequence of Fourier coefficients of $e^{i\omega h}\hat{f}(\omega)$.

Define
$$G^{(h)}(\omega) \equiv e^{i\omega h} \max\left(0, \min\left(2 - \frac{2|\omega|}{\pi}, 1\right)\right).$$
This is equal to $e^{i\omega h}$ on $[-\pi/2, \pi/2]$ and goes to zero at $-\pi$ and π. Let $(g_k^{(h)})$ denote the Fourier coefficients of $G^{(h)}$. Then, as we will see below, $\sum |g_k^{(h)}| < \infty$, and so the identity

$$f(k+h) = \sum_m g_m^{(h)} f(k-m)$$

holds. This expresses $(f(k+h))$ as the convolution of $(g_k^{(h)})$ with $(f(k))$. Now l_1 is a convolution algebra, so

$$||f(k+h)||_{l_1} \leq ||g_k^{(h)}||_{l_1} ||f(k)||_{l_1}.$$

On the other hand,

$$f(k) = \sum_m g_m^{(-h)} f(k+h-m),$$

so
$$||f(k)||_{l_1} \leq ||g_k^{-h}||_{l_1} ||f(k+h)||_{l_1}.$$
The Lemma now follows from the inequality

$$||g_k^{-h}||_{l_1} \leq 3/2 \text{ for all } h \in [-1, 1]. \tag{40}$$

Note that the ordinary Fourier transform of $G^{(h)}(\omega)$ is just $g(t-h)$, where $g(t) = \frac{sin^2(\frac{\pi t}{2})}{(\pi t/2)^2} - \frac{sin^2(\frac{\pi t}{4})}{(\pi t/2)^2}$. Moreover, in the language of Theorem 4, (40) is equivalent to the statement that

$$||g||_{\infty,1} \leq 3/2. \tag{41}$$

This may be verified by computations completely analogous to those of Lemma 5.

The basic argument of the lemma goes back to Wiener's memoir on Tauberian theorems (*Wiener* [13]; see also *Katznelson* [4], section VIII, p.227). *Plancherel and Polya* [10], and *Boas* [1] have given explicit constants by Complex variable methods; however the value 3/2 obtained here is numerically somewhat better than these. For example, a direct adaptation of Wiener's argument, explained to us by Ytzhak Katznelson gives the constant 3; we have taken the approach a bit further to get the constant 3/2. Boas' argument gives the constant $7 + 4/\pi^2$.

5 Improvements

5.1 Increasing the Density Limit

While the inequalities we have given establish the basic principle that the density $\rho(T, \Omega)$, rather than $|T|\Omega$, is the controlling factor in signal recovery, we believe that much better results are possible. In the L_2 setting , we ought to have

$$\text{``}\sup_{\mathcal{T}_2} \rho(T, \Omega) < 1 \text{ implies } \sup_{\mathcal{T}_2} \lambda_0(T, \Omega) < 1\text{''}.$$

For the case T periodic, Theorem 1 establishes this. But Theorem 2 ceases being effective for $\sup_{\mathcal{T}_2} \rho(T, \Omega)$ greater than $1/3$. We indicate two reasons we believe such a better result should hold.

First, there is the example of the L_∞ norm.

Theorem 13 *Let $f \in B_\infty(\pi)$. In order that*

$$\sup_{t \in T^c} |f(t)| \leq \epsilon ||f||_\infty,$$

it is necessary that T contain an interval of length

$$\frac{2\sqrt{2}}{\pi}(1 - \epsilon)^{\frac{1}{2}}.$$

This implies that the L_∞ analog of λ_0 and μ_0:

$$\eta_0 \equiv 1 - \sup_{f \in B_\infty(\pi\Omega)} \frac{||P_{T^c} f||_\infty}{||f||_\infty}$$

satisfies

$$\eta_0 \leq \frac{\pi}{2\sqrt{2}} \sqrt{\rho(T, \Omega)},$$

which is effective for values of ρ as large as .81.

Proof of Theorem 13. Without loss of generality, suppose f attains its maximum in T at the point x, so $f(x) = ||f||_\infty$ and $f'(x) = 0$. Let $y \in T^c$.

$$f(y) = f(x) + \int_x^y \int_x^v f''(u) du dv.$$

Now

$$\left|\int_x^y \int_x^v f''(u)dudv\right| \leq ||f''||_\infty \frac{(y-x)^2}{2}.$$

But by Bernstein's inequality [4]

$$||f''||_\infty \leq \pi^2 ||f||_\infty.$$

Thus

$$\begin{aligned} f(y) &\geq f(x) - \frac{\pi^2 (y-x)^2}{2} ||f||_\infty \\ &= ||f||_\infty (1 - \frac{\pi^2 (y-x)^2}{2}). \end{aligned} \tag{42}$$

So in order that

$$\sup_{y \in T^c} |f(y)| \leq \epsilon ||f||_\infty,$$

we must have

$$|y-x| \geq \frac{1}{\pi}\sqrt{2(1-\epsilon)}$$

for every $y \in T^c$.

A second piece of evidence is the following result of *Paneyakh* [9] . If the *asymptotic density*

$$\tilde{\rho}(T) \equiv \overline{\lim_{n\to\infty}} \sup_t |T \cap [t, t+n]|/n$$

satisfies $\tilde{\rho}(T) < 1$ then $\lambda_0(T, \Omega) < 1$ for any $\Omega < \infty$. This result is too weak to imply the principle stated above, however: it is not uniform in T. That is, we do not know from this result that there exists a constant less than 1 bounding λ_0 for all T with density $< .95$, say.

5.2 Two Optimization Problems

Our method in section 3 can be abstracted as follows. If T is periodic, then

$$\mu_0(T, \Omega) \leq c_0 \rho(T, \Omega)$$

where c_0 solves the optimization problem

$$c_0 = \inf\{||g||_{\infty,1} : \; \hat{g}(\omega) = 1 \text{ for all } \omega \in [-\pi, \pi]\}.$$

Here $||\cdot||_{\infty,1}$ denotes the norm introduced in the proof of Theorem 4. Thus, c_0 is the minimal norm of any kernel g which is reproducing for $B_1(\pi)$, i.e. which satisfies $f * g = f$ for all f in $B_1(\pi)$.

If T is aperiodic, then

$$\mu(T,\Omega) \leq c_1 \rho(T,\Omega)$$

where c_1 solves the optimization problem

$$c_1 = \inf\{||g||_* : \ \hat{g}(\omega) = 1 \text{ for all } \omega \in [-\pi,\pi]\}.$$

Here $||\cdot||_*$ denotes what we call the *Dual-Density* norm

$$||g||_* \equiv \sup\{\int |g| d\mu : \ d\mu \geq 0 \ , \ \mu(t,t+1) \leq 1\}.$$

This is, roughly, the dual norm of the so-called "translation-bounded measures" norm; see *Fournier and Stewart* [3]. It is within a factor 2 of the norm

$$||g||_{1,\infty} = \sum_{k=-\infty}^{\infty} \sup_h |g(k+h)|.$$

Thus, c_1 represents the smallest Dual-Density norm of any reproducing kernel for $B_1(\pi)$.

Evidently, in section 3 we have simply exhibited two particular reproducing kernels – the trapezoid and $Ave_h\{\zeta_{t,h}(u)\}$ – and calculated constants for those kernels. The solution of the problems just mentioned here would lead to better constants.

References

[1] R.P. Jr. Boas. Entire functions bounded on a line. *Duke Math. J.*, 6:148–169, 1940.

[2] D.L. Donoho and P.B. Stark. Uncertainty principles and signal recovery. *SIAM J. Appl. Math.*, to appear, 1989.

[3] J.J.F. Fournier and J. Stewart. Amalgams of l^p and l_q. *Bulletin of the American Mathematical Society*, 13:1–22, 1985.

[4] Y. Katznelson. *An Introduction to Harmonic Analysis.* Dover, New York, 1968.

[5] S. Levy and P.K. Fullagar. Reconstruction of a sparse spike train from a portion of its spectrum and application to high-resolution deconvolution. *Geophysics*, 46:1235–1243, 1981.

[6] B.F. Logan. *Properties of High-Pass Signals.* PhD thesis, Columbia University, 1965.

[7] H.L. Montgomery. The analytic principle of the large sieve. *Bull. Am. Math. Soc.*, 84:547–567, 1978.

[8] S.M. Nikoĺskiĭ. Inequalitied for entire functions of finite degree and their application to the theory of differentiable functions of several variables. In *American Mathematical Society Translations Series 2, Vol. 80, Thirteen Papers on Functions of Real and Complex Variables*, pages 1–38, American Mathematical Society, Providence, R.I., 1969.

[9] B.P. Paneyakh. Some inequalities for functions of exponential type and a priori bounds for general differential operators. 75–114.

[10] M. Plancherel and G. Pólya. Fonctions entières et intégrales de fourier multiples. *Commentarii math. Helv.*, 10:110–163, 1938.

[11] F. Santosa and W.W. Symes. Linear inversion of band-limited reflection seismograms. *SIAM J. Sci. Stat. Comp.*, 7:1307–1330, 1986.

[12] C. Walker and T.J. Ulrych. Autoregressive recovery of the acoustic impedance. *Geophysics*, 48:1338–1350, 1983.

[13] N. Wiener. Tauberian theorems. *Ann. Math.*, 33:1–100, 1932.