Nonlinear approximation theory on finite groups

K.-L. Kueh*
T. Olson†
D. Rockmore‡
K.-S. Tan§

November 21, 1999

Abstract

Motivated by problems in signal recovery, we will investigate the distribution of the energy of the Fourier transform of a positive function on a finite group. In particular, we are able to bound from below the fraction of energy contained in various subsets of the Fourier transform of a positive function defined on a finite group. Applications to signal recovery for positive functions, as well as partial spectral analysis for data on finite groups are also presented.

1 Introduction

1.1 History

The problem of approximating functions of $L^2(\mathbf{R})$ or $L^2(S^1)$ by a subsum of their Fourier expansion has a long history. In 1911, a fundamental theorem in approximation theory was published by Jackson [11, 12], relating the smoothness of a function to the rate of decay of its Fourier coefficients. While volumes of work have been written on similar topics in approximation theory improving the scope and quantitative nature of this theorem, this basic result remains at the heart of linear approximation theory and can be stated simply.

Theorem 1.1 (Jackson's Theorem, see [11, 12, 1]) *Let $f \in L^2(\mathbf{R})$ and let $S_n(f)$ be the n^{th} partial sum of the Fourier series. Then f has k continuous derivatives if and only if*

$$\|f - S_n(f)\| = O(n^{-k}).$$

While these sorts of linear approximation theorems are very well known, the analogous nonlinear theory in which positivity is added to the description of f is less well-studied. One result which characterizes the Fourier transform of such functions is the following.

Theorem 1.2 (Bochner's Theorem, see eg. [19], p. 330) *Let $f \in L^2(\mathbf{R})$ and say f is a function of positive type if the matrix $\{f(\lambda_i - \lambda_j)\}_{i,j}$ is positive for any $\lambda_1, \ldots, \lambda_n$ in $\mathbf{R}^n$. The cone of functions of positive type is exactly the Fourier transform of the positive functions.*

*Institute of Mathematics, Academia Sinica, Nankang, Taipei, Taiwan. Email: maklk@ccvax.sinica.edu.tw
†Department of Mathematics, Univ. of Florida, Gainesville, FL. Supported by NSF. Email: olson@math.ufl.edu
‡Department of Mathematics, Dartmouth College, Hanover, NH 03755. Supported by NSF. Email: rockmore@cs.dartmouth.edu
§Dept. of Mathematics, National Taiwan University, Taipei 106, Taiwan. Email: tan@math.ntu.edu.tw

Bochner's theorem gives us an accurate characterization of the image cone, i.e., the Fourier transform of the positive cone. Nevertheless, the characterization of positive functions according to the decay rate or concentration of the energy of their Fourier transforms lacks a complete Jackson-type theorem. An early result in this direction is due to Erdös and Fuchs [9].

Theorem 1.3 (Erdös and Fuchs [9]) *Suppose that* $\phi(z) = \sum_{n=0}^{\infty} b_n z^n$ *is convergent for* $|z| < 1$ *and suppose that all* b_n *are non-negative real numbers. Then for* $0 < \alpha \leq \pi$, $z = re^{i\theta}$ *and* $0 < r < 1$,

$$\frac{1}{2\alpha} \int_{-\alpha}^{\alpha} |\phi(z)|^2 d\theta \geq \frac{1}{6\pi} \int_{-\pi}^{\pi} |\phi(z)|^2 d\theta. \tag{1}$$

In the limiting case where $r = 1$ (assuming convergence), if we let $\alpha = \pi/2$, then we obtain

$$\int_{-\pi/2}^{\pi/2} |\phi(z)|^2 d\theta \geq \frac{1}{6} \int_{-\pi}^{\pi} |\phi(z)|^2 d\theta = \frac{1}{6} \|\phi\|^2. \tag{2}$$

If we view $\phi(z)$ as the Fourier transform of a function $b \in L^1(\mathbf{Z})$, then (2) can be reformulated as the statement that at least $\frac{1}{6}$ of the squared energy of such a function (nonnegative, with support on the nonnegative integers), and correspondingly, at least $\frac{1}{\sqrt{6}} \approx .41$ of the energy of this positive definite function, must be contained in the "low frequency" component of the Fourier Transform. Thus, such a positive function cannot consist only of high frequency information.

Various efforts have been directed at improving the bound (2). In particular, both Shapiro [21] and Logan [14], have shown that the constant $\frac{1}{6}$ can be improved to $\frac{1}{4}$. Their result is sharp in that it cannot be improved uniformly for all values of α in (1).

In this paper we examine finite analogues of bounds with the form (2) to study the energy content of different subsets of the frequency domain for positive functions on finite groups. The case of cyclic groups is of particular interest for digital signal processing and for this we obtain a bound of 1/4, effectively transferring the Logan-Shapiro bound of $\frac{1}{4}$ to the discrete setting. However, our main result (Theorem 2.3) is of a general form stated in terms of the representation theory of finite groups. Thus, this sort of finite Paley-Weiner analysis has implications even for noncommutative groups and thereby extends to the general spectral analysis framework originally proposed by Diaconis [5] for studying data on groups.

1.2 Motivation

Our motivation comes from digital signal processing, and the problem of signal recovery from limited Fourier transform data. For finite, discrete positive functions, a class of functions defined by a nonlinear constraint, it appears that the stability of signal recovery is tied to the ratio of low frequency energy to total signal energy. A linear analysis of the problem would tie stability to the condition number, a quantity which pertains when the class of object functions is a linear subspace.

By signal recovery, we mean the problem of inverting convolution operators which may in general be hard to invert. The extreme case of this is that of an operator L of the form

$$Ls = F^{-1} K F s \tag{3}$$

where F is the Fourier transform, K is a diagonal operator consisting of 0's and 1's, and s is the signal that we would like to recover. This operator models the problem of recovering a signal from partial knowledge of its Fourier transform. The regions in which K is 1 represent the "known" or measured Fourier data, and the regions in which K is 0 represent the "unknown" or unmeasured data. This type of operator is also referred to as a *band-pass* operator. More moderate examples are of the form outlined above, but with K a diagonal operator with some "large" values, and some "small" values. For example, a Gaussian blurring operator would fall into this class.

Band-pass operators are impossible to invert without a priori conditions on the object function s. For example, if s has compact support Ω, and we consider the restricted band-pass operator on $L^2(\Omega)$, then the problem becomes one of analytic continuation. The Fourier transform of a compactly supported function is an entire function, and therefore limited knowledge of the Fourier transform of s, or a small region in which K is 1, is sufficient to uniquely determine s. Unfortunately, this is a classically ill-conditioned problem [6].

Nevertheless, while in the general situation the usual linear analysis (i.e., according to condition number) may be appropriate, it seems that when the class of object functions is defined by some sort of nonlinear constraint, the linear analysis may not give a good indication of the stability of the associated inverse problem.

For example, Donoho and Stark consider the effect of sparsity on the recovery of discrete, finite signals. They proved that if a discrete function is non-zero on N points, then its Fourier transform cannot vanish on N consecutive points [8]. This implies that for certain finite signals, partial knowledge of its Fourier transform is sufficient for recovery. Related papers show that sparsity of the signal will guarantee stability of these types of inverse problems [6, 7].

We are interested in the effect of the nonlinear constraint of positivity on inverse problems. This is an important class of functions which arise naturally in a variety of areas including image processing and density estimation. Our belief is that in this situation, signal recovery is tied to the proportion of energy contained in the low frequencies. We now give some indication of this link.

Let us suppose that $s \geq 0$. With this extra information, bounding the error in the recovery of s from some noisy data $Ls + n$ may now be stated as solving the constrained optimization problem:

$$\begin{cases} \text{Minimize:} & ||Lx - (Ls + n)|| \\ \text{Subject to:} & x \geq 0. \end{cases} \tag{4}$$

It is enlightening to consider this problem from a different perspective. Since the linear image of a cone is once again a cone, we consider the range cone $\mathcal{R} = \mathcal{R}_L = \{Lx | x \geq 0\}$ and denote by P the projection onto this cone, i.e.,

$$Pz = \text{argmin}_{y \in \mathcal{R}} ||y - z||. \tag{5}$$

It follows that the solution to the constrained optimization problem (4) is $Lx = P(Ls + n)$, with error

$$E(s, n) = s - x = s - L^{-1}(P(Ls + n)) = L^{-1}(Ls - P(Ls + n)) = L^{-1}(Pn), \tag{6}$$

since $PLs = Ls$. Note that because the projection P depends in a nonlinear way on s, we should not expect sharp bounds for (6) which are independent of s. However, because P is a projection, it follows that the error is bounded by

$$\max_{||n|| \leq \epsilon, Ls+n \in \mathcal{R}} ||L^{-1}n||. \tag{7}$$

While it seems difficult to bound (7) sharply for all s, it appears that this is a measure of how expansive L^{-1} is on the range cone $\mathcal{R}_L$ rather than on the entire space, or that the nonlinear bound

$$NLB = \max_{||x||=1, x \in \mathcal{R}} ||L^{-1}x|| \tag{8}$$

would be an appropriate measure of stability. This is the bound which we are investigating for the operator which maps a discrete function to a truncated version of its Fourier series. Beyond this general bound, we believe that the measurement of the ratio for each function $x \in \mathcal{R}$,

$$NLB(x) = \frac{||L^{-1}x||}{||x||}$$

will give insight into the stability of recovering s, where $Ls = x$. These insights are not the normal ones usually expected from linear stability theory.

To understand how NLB is related to Theorem 2.5 let $x \in \mathcal{R}$ and $L = F^{-1}KF$, where K is the diagonal operator which is 1 on the low frequency data, and 0 on the high frequency data. By Parseval's equality, $\|x\| = 1$ implies that the energy contained in the low frequency information is exactly 1, and $\|L^{-1}x\|$ represents the energy of the total signal. By Theorem 2.5 the ratio of the low frequency energy to the total energy, is least $\frac{1}{2}$. This then implies that $NLB(x)$, which is the reciprocal of this, is at most 2. Note that $NLB(x)$ does not bound the true error (7), which is the worst possible error that can occur for the constrained optimization problem.

Linear inversion theory (illustrated by the condition number and other linear stability measures) suggests that inputs of relatively small energy, which produce outputs with relatively large energy, will have unstable recovery. The general rule of thumb is that the noise will be magnified by the ratio of the energy of the input to that of the output. However, in the nonlinear case which we address, the opposite seems to be the case. If we try to recover a small amount of high frequency energy from a function which is mostly low frequency energy, our conjecture is that the recovery problem is not stable. This is the case when $NLB(x)$ is small, or near 1. On the other hand, if $NLB(x)$ is large, or near 2, then it should be possible to stably recover the large high frequency component of the function from the low frequency component. This is exactly the opposite of what one would expect from linear stability analysis.

Our explanation of this phenomenon comes from a link between the ratio $NLB(x)$ and sparsity (Theorem 3.2), which indicates that large $NLB(x)$ is related to signal sparsity. In [6] Donoho shows that sparse functions admit stable recovery of their high frequency components from low frequency information. In addition, the general requirement for a function to approach the maximum NLB is that it will need some isolated high frequency components, with low energy components separating them.

Further appreciation for the difference between the error (7) and the linear error $\|L^{-1}n\|$ is gained by remembering that high-dimensional cones have nearly no interior, and infinite-dimensional cones have no interior. Therefore the fundamental linear idea of measuring the maximal deformation of an open ball of noise, does not necessarily make sense in the nonlinear setting, since this noise ball may almost surely lie outside of the range cone. In some sense, the measurement of a normalized function as in NLB, or the ratio (23) measure whether a function lives in the interior of the cone, or near the boundary of the cone. The nonlinearities presented by the boundary of the cone will be much stronger for a function whose NLB is large.

1.3 Organization

The rest of the paper is organized as follows: After some preliminaries on the representation theory of finite groups, we prove our main result (Theorem 2.3), bounding from below the percentage of energy contained in certain subsets of representations (Fourier space). Section 3 contains our applications. In Section 3.1 we discuss in a bit more detail the relation to positivity constrained inverse problems. This includes some numerical experiments which support the claim that by taking positivity into account, a wider class of functions admit stable recovery. In Sections 3.2 and 3.3 we consider some potential applications to partial spectral analysis. Finally in Section 5 we close with a list of possible directions for future research.

Acknowledgement. Special thanks are due to Jong-Shi Pang of The Johns Hopkins University for developing the optimization software used in the numerical examples of Section 3.1.2.

2 Main result

We prove a general result, bounding from below the fraction of energy within particular subsets of the frequency domain, i.e., the length of the projection of an arbitrary positive function onto a set of irreducible matrix elements spanning a union of irreducible representation spaces. In the classical abelian case, this is simply the subspace spanned by certain subsets of vectors of regularly sampled exponentials. Necessarily we will require some of the rudiments of group representation theory (see [20] for a basic reference). We begin with some notation and terminology (Section 2.1) and then in stages prove the main result (Theorem 2.3) in Section 2.2. We postpone the applications to Section 3.

2.1 Preliminaries

Let G be a finite group and let $L^2(G) = \{f \mid f : G \longrightarrow \mathbf{C}\}$ be the vector space of $\mathbf{C}$-valued functions on G with the usual inner product

$$\langle f, g\rangle = \frac{1}{|G|}\sum_{x\in G} f(x)\overline{g(x)} \tag{9}$$

for any $f, g \in L^2(G)$.

Let $\rho : G \longrightarrow GL(r, \mathbf{C})$ be a unitary irreducible representation of G. Thus, $r = r_\rho = deg(\rho)$. Recall that any irreducible representation is equivalent (attainable by a change of basis) to one which is unitary.

We denote the i, j matrix element associated with ρ as ρ_{ij}. Then for every i, j with $1 \le i, j \le r$, $\rho_{ij} \in L(G)$, and (cf. [20], p. 14),

$$\langle \rho_{ij}, \rho_{kl}\rangle = \frac{1}{r}\delta_{i,k}\cdot\delta_{j,l}, \text{ for all } i, j, k, l = 1, \ldots, r. \tag{10}$$

Let

$$L_\rho = \mathrm{Span}\{\rho_{ij} \mid\ i, j = 1, ..., r\},$$

then by (10) the functions

$$\tilde{\rho}_{ij} = \sqrt{r}\rho_{ij}$$

as $1 \le i, j \le r_\rho$ is an orthonormal basis of L_ρ. The subspace L_ρ is an isotypic subspace of $L^2(G)$, but is not irreducible unless $r = 1$. For $r > 1$, L_ρ is a direct sum of r irreducible subspaces. $L^1_\rho, \ldots, L^r_\rho$ each of dimension r, defined by

$$L^i_\rho = \mathrm{Span}\{\rho_{i1}, \ldots, \rho_{ir}\}.$$

Example: In the abelian case, each $r = 1$. In particular for a cyclic group the isotypics are naturally indexed by the integers $0, \ldots, N-1$ (for $G \cong \mathbf{Z}/N\mathbf{Z}$, so $|G| = N$) and

$$L_j = \mathrm{Span}\{(\omega^0, \ldots, \omega^{jk}, \ldots, \omega^{j(N-1)})\}$$

for $\omega = e^{2\pi i/N}$.

In this way, from a complete set of irreducible matrix elements we obtain an orthonormal basis for $L^2(G)$. The orthogonality of the ρ_{ij} (for distinct ρ) or equivalently, the $\tilde{\rho}_{ij}$ gives a direct sum decomposition

$$L^2(G) = \oplus_\rho L_\rho$$

where ρ runs through a complete set of inequivalent irreducible representations for G.

A *Fourier expansion* for $f \in L^2(G)$, is simply the expansion of f in terms of the orthonormalized irreducible matrix elements,

$$f = \sum_{\rho,i,j} \hat{f}_{\rho,i,j} \tilde{\rho}_{ij} \tag{11}$$

where $\hat{f}_{\rho,i,j} \in \mathbf{C}$ is the associated *Fourier coefficient.*

For $f \in L^2(G)$, let $|\hat{f}|^2_\rho$ denote the energy content of f at the irreducible representation ρ. Thus,

$$|\hat{f}|^2_\rho = \sum_{1 \le i,j \le r} |\hat{f}_{\rho,i,j}|^2.$$

Then by the orthonormality of the basis

$$\|f\|^2 = |\hat{f}|^2 = \sum_\rho |\hat{f}|^2_\rho.$$

This is the natural generalization of the power spectrum of a positive function.

Remark. Efficient algorithms for the computation of the Fourier coefficients (11) go by the name of *fast Fourier transforms* or *FFTs.* The well-known Cooley-Tukey FFT [3] and its many variants accomplish this task for any finite abelian group. Techniques are also available for many families of noncommutative groups. See [17] for a current survey. The usual convention is for the Fourier transform of a function $f \in L^2(G)$ at a given irreducible matrix ρ_{ij} element to be defined as the sum

$$\widehat{f}(\rho_{ij}) = \sum_{x \in G} f(x) \rho_{ij}(x)$$

and the Fourier transform at ρ is the entire matrix

$$\widehat{f}(\rho) = \sum_{x \in G} f(x) \rho(x).$$

If we define the function $f^{\#}(x) = f(x^{-1})$ then

$$\hat{f}_{\rho,i,j} = \sqrt{r}\,\widehat{f^{\#}}(\rho_{ji}) \tag{12}$$

where we use the fact that for a unitary representation $\rho_{ij}(x^{-1}) = \overline{\rho_{ji}(x)}$.

2.2 A general bound

For all $x \in G$ define

$$f^c(x) = \sum_{z \in G} f(z) \cdot \overline{f(zx^{-1})} \tag{13}$$

Notice that if G is abelian, then (13) takes the form of $\sum_{z \in G} f(z)\overline{f(z-x)}$ which is the autocorrelation of f. Thus, equation (13) is a natural generalization of the autocorrelation of f under the definition of convolution for $L^2(G)$ given by

$$f \star g(x) = \sum_{z \in G} f(z) g(z^{-1}x). \tag{14}$$

To begin, we consider the Fourier expansion of f^c. By matrix multiplication,

$$\tilde{\rho}_{kl}(zx^{-1}) = \frac{1}{\sqrt{r}} \sum_{m=1}^{r} \tilde{\rho}_{km}(z) \cdot \tilde{\rho}_{ml}(x^{-1}),$$

and by unitarity

$$\tilde{\rho}_{ml}(x^{-1}) = \overline{\tilde{\rho}_{lm}(x)}.$$

Thus, we have the following expansion:

$$\begin{aligned}
f^c(x) &= \sum_{z \in G} \sum_{\rho,i,j,\xi,k,l} \hat{f}_{\rho,i,j} \cdot \overline{\hat{f}_{\xi,k,l}} \cdot \tilde{\rho}_{ij}(z) \cdot \overline{\tilde{\xi}_{kl}}(zx^{-1}) \\
&= \sum_{z \in G} \sum_{\rho,i,j,\xi,k,l,m} \hat{f}_{\rho,i,j} \cdot \overline{\hat{f}_{\xi,k,l}} \cdot \frac{1}{\sqrt{r_\xi}} \tilde{\rho}_{ij}(z) \cdot \overline{\tilde{\xi}_{km}(z)} \cdot \tilde{\xi}_{lm}(x) \\
&= \sum_{\rho,j,l} \frac{|G|}{\sqrt{r_\rho}} \cdot \left(\sum_i \hat{f}_{\rho,i,j} \cdot \overline{\hat{f}_{\rho,i,l}} \right) \tilde{\rho}_{lj}(x) \\
&= |G| \sum_{\rho,j,l} \sum_i \hat{f}_{\rho,i,j} \cdot \overline{\hat{f}_{\rho,i,l}} \rho_{lj}(x). \qquad (15)
\end{aligned}$$

Let

$$\chi_\rho = \sum_i \rho_{i,i}$$

be the character (trace operator) associated to ρ. This function is independent of choice of basis. Notice that the orthogonality of the irreducible matrix elements (10) implies orthonormality of the characters. In particular, if η is any representation of G, then for any irreducible representation ρ, the inner product $\langle \chi_\rho, \chi_\eta \rangle$ gives the number of direct summands equivalent to ρ which occur in any irreducible decomposition of η.

The following Lemma is a direct consequence of the above expansion of f^c and the orthogonality relations (10).

Lemma 2.1 *For $f \in L^2(G)$ and ρ an irreducible representation of G*

$$\langle f^c, \chi_\rho \rangle = \frac{|G|}{r} |\hat{f}|_\rho^2.$$

For an irreducible representation ρ, define

$$\psi_\rho = r_\rho \chi_\rho.$$

This is the character of a direct sum of $r = r_\rho$ copies of ρ. For a finite set of irreducible representations S, define

$$\psi_S = \sum_{\rho \in S} \psi_\rho. \qquad (16)$$

Example. Let S denote a complete set of inequivalent irreducible representations of G. Recall that the regular representation of G contains any $\rho \in S$ exactly r_ρ times. Thus in this case, ψ_S is exactly the character of the regular representation.

For any representation ρ, let ρ' denote the corresponding contragredient representation. That is,

$$\rho'(x) = \rho(x^{-1})^t \qquad (17)$$

where the superscript t denotes transpose. Notice that $\chi_{\rho'} = \overline{\chi_\rho}$.

Recall that for any representations ρ_1 and ρ_2, their tensor product is defined by $\rho_1 \otimes \rho_2(x) = \rho_1(x) \otimes \rho_2(x)$. It follows that the associated relation on characters is $\chi_{\rho_1 \otimes \rho_2} = \chi_{\rho_1} \cdot \chi_{\rho_2}$. In general, the tensor product of irreducible representations is not irreducible. In particular, the orthonormality of the characters implies that for an irreducible representation ρ, $\rho \otimes \rho'$ contains the trivial representation exactly once, so that if ρ is multidimensional, $\rho \otimes \rho'$ is reducible.

For any subset of distinct irreducible representations S, consider the irreducible decomposition of $(\sum_{\rho \in S} \rho) \otimes (\sum_{\rho \in S} \rho)'$. If we let χ_S denote the character of $\sum_{\rho \in S} \rho$, then on the level of characters the irreducible decomposition becomes

$$\chi_S \cdot \overline{\chi_S} = \sum_{\xi} n_\xi \chi_\xi,$$

where ξ varies over all irreducible representations of G and $n_\xi \geq 0$.

Given S, define

$$T(S) = \{\xi \mid n_\xi \neq 0\}.$$

We also define

$$deg(S) = \sum_{\rho \in S} r_\rho^2.$$

Again, if S is a complete set of irreducible representations for G, then $deg(S) = |G|$.

Lemma 2.2 *Let S be a set of inequivalent irreducible representations. Then in the above notation,*

$$Max\{n_\xi \mid \xi \in T(S)\} = deg(S).$$

Proof: Let χ_{reg} denote the character of the regular representation of G. Then $\chi_{reg} = \sum_\rho r_\rho \chi_\rho$. By orthonormality of characters, for every ρ, we have

$$\chi_\rho \cdot \chi_{reg} = r_\rho \chi_{reg}.$$

Therefore, given $\rho_1 \in S$, then as ρ_2 runs through S, a given ξ can occur in all the $\rho_1 \otimes \rho_2'$ at most r_ξ times. Thus, for every ξ,

$$n_\xi \leq deg(S)$$

and

$$Max\{n_\xi \mid \xi \in T(S)\} \leq deg(S).$$

On the other hand, let ξ_0 denote the trivial representation. Then for any $\rho_1, \rho_2 \in S$, ξ_0 occurs in $\rho_1 \otimes \rho_2'$ if and only if $\rho_1 = \rho_2$, and in this case it only occurs once. Since each representation ρ occurs with multiplicity r (equal to its degree) in ψ_S, we get

$$n_{\xi_0} = deg(S),$$

so that

$$Max\{n_\xi \mid \xi \in T(S)\} \geq deg(S).$$

♠

We are interested in the case of nonnegative functions. We will make use of the following obvious facts, that if f is nonnegative, then f^c is nonnegative, and for any f, $f \cdot \overline{f}$ is nonnegative.

We are now ready to prove our main theorem.

Theorem 2.3 *Let S be a subset of inequivalent irreducible representations of G and let $f \in L^2(G)$ be a positive function, then*

$$\sum_{\rho \in T(S)} |\hat{f}|_\rho^2 \geq \frac{deg(S)}{|G|} |\hat{f}|^2$$

or equivalently

$$\frac{\sum_{\rho \in T(S)} |\hat{f}|^2_\rho}{|\hat{f}|^2} \geq \frac{deg(S)}{|G|}. \tag{18}$$

In addition, the lower bound in (18) is sharp in the sense that there are subgroups S of groups G that come arbitrarily close to satisfying the equality in (18).

Proof of the inequality: Since f^c and $\psi_S \cdot \overline{\psi_S}$ are both positive, we have

$$\begin{aligned} \langle f^c, \psi_S \cdot \overline{\psi_S} \rangle &\geq f^c(id)\psi_S(id)^2 \\ &= |\hat{f}|^2 deg(S)^2. \end{aligned} \tag{19}$$

Note that in the first line of (19) is the only place where positivity is used.

Expanding out the lefthand side of (19) yields

$$\langle f^c, \sum_{\xi \in T} n_\xi \psi_\xi \rangle,$$

and by Lemma 2.1 this is equal to

$$\sum_{\xi \in T} n_\xi |\hat{f}|^2_\xi. \tag{20}$$

By Lemma 2.2 the sum (20) is at most

$$deg(S) \sum_{\xi \in T} |\hat{f}|^2_\xi.$$

♠

It is straightforward to adapt this bound to any *homogeneous space* for G. This is the case in which G acts transitively as a group of permutations on a set set X, so that the vector space of functions on X, denoted $L^2(X)$ is naturally identified with the subspace $L^2(G/H) < L^2(G)$ of right H-invariant functions on G, for a subgroup $H < G$ which fixes a chosen basepoint $x_0 \in X$.

In this case, given $f \in L^2(G/H)$, identify f with the function on G whose value on an element of a coset xH is simply $\frac{1}{H} f(x)$. Note that this is an orthogonal embedding of $L^2(G/H)$ into $L^2(G)$. With these conventions, we easily have the following.

Theorem 2.4 *Let X be a homogeneous space for G with $H < G$ the stabilizer subgroup of some fixed basepoint $x_0 \in X$. With all the notation above, let S be a subset of inequivalent irreducible representations of G and let $f \in L^2(G/H)$ be a positive function, then*

$$\sum_{\rho \in T(S)} |\hat{f}|^2_\rho \geq \frac{deg(S)}{|X|} |\hat{f}|^2.$$

2.3 Cyclic groups

Of interest for digital signal processing is the case of cyclic groups (i.e., discretized circles) and their products, which make up the class of all abelian groups. For their independent interest we restate these results here. The result for cyclic groups reobtains the results of Logan [14] and Shapiro[21].

Theorem 2.5 *Let $G = \mathbf{Z}/N\mathbf{Z}, n < N/2$ and $f \geq 0$. Then*

$$\sum_{-n \leq j \leq n} |\hat{f}(j)|^2 \geq \frac{n}{N} \|\hat{f}\|^2,$$

or

$$\frac{\sum_{-n \leq j \leq n} |\hat{f}(j)|^2}{\|\hat{f}\|^2} \geq \frac{n}{N} \tag{21}$$

or

$$\frac{\|f\|_n}{\|f\|} \geq \sqrt{\frac{n}{N}} \tag{22}$$

where $\|f\|_n$ denotes the projection of f onto the frequencies between $-n$ and n.

Proof: The irreducible representations of G are indexed by the integers modulo N. We take as our representative set $\{-N/2, -N/2+1, \ldots, N/2-1\}$. Let

$$S = \{-n/2, \ldots, 0, \ldots, n/2-1\}.$$

Thus, $deg(S) = n$, $\rho'_k = \rho_{-k}$ and

$$T(S) = \{k - l| - n/2 \leq k, l < n/2\} = \{m| - n \leq M < n\}.$$

Since all representations are one-dimensional, we have that the trivial representation (ρ_0) occurs n times. Hence the result.

♠

Of particular interest is the energy content of the low frequency part of the spectrum. This is defined as the range from $-N/4$ to $N/4$.

Corollary 2.6 *For $f : \mathbf{Z}/N\mathbf{Z} \longrightarrow \mathbf{C}$, $f \geq 0$,*

$$\sum_{-N/4 \leq j \leq N/4} |\hat{f}(j)|^2 \geq \frac{1}{4} \|f\|^2,$$

or

$$\frac{\sum_{-N/4 \leq j \leq N/4} |\hat{f}(j)|^2}{\|f\|^2} \geq 1/4 \tag{23}$$

or

$$\frac{\|f\|_{\frac{N}{4}}}{\|f\|} \geq \frac{1}{2}. \tag{24}$$

The result for cyclic groups (Theorem 2.5) can be easily extended to any abelian group (which can always be written as a direct product of cyclic groups). In this case we are able to bound from below the frequency content of rectangular regions about the trivial representation.

Theorem 2.7 *Let $N_1, \ldots, N_m > 2$ and $f : \mathbf{Z}/N_1\mathbf{Z} \times \cdots \times \mathbf{Z}/N_m\mathbf{Z} \longrightarrow \mathbf{C}$, $f \geq 0$. For $n_i < N_i/2$,*

$$\sum_{-n_1 \leq j_1 \leq n_1} \cdots \sum_{-n_m \leq j_m \leq n_m} |\hat{f}(j_1, \ldots, j_m)|^2 \geq \frac{n_1 \cdots n_m}{N_1 \cdots N_m} \|\hat{f}\|^2.$$

Notice that Theorem 2.7 excludes the case of all $N_i = 2$. The result here is is slightly different, since in this case, for every representation χ, we have $\chi \otimes \overline{\chi} = 1$. In addition, this case is of independent interest for its relation to the analysis of 2^k factorial designs. This is taken up in Section 3.2.

2.4 Sharpness of the bound

We will now establish that Theorems 2.3, and 2.5 are sharp, in the sense that one cannot uniformly increase the lower bounds for the ratios (18) and (21). This is analogous to the tightness of the 1/4 bound provided by Shapiro and Logan. We will begin with some notation. Let us restrict ourselves to the cyclic group, and consider its subgroups. Let us assume that $N = k \cdot m$, and define the following notation,

$$\mathbf{III}^k(i) = \begin{cases} 1 & \text{if } i = q * k \text{ for some } q \\ 0 & \text{otherwise.} \end{cases}$$

Thus $\mathbf{III}^k$ is the characteristic function on a subgroup of the cyclic group. An interesting an useful fact is that the Fourier transform of $\mathbf{III}^k$ is a multiple of the indicator on the dual subgroup, $\sqrt{\frac{m}{k}}\mathbf{III}^m$. We will use these functions to show that we asymptotically achieve the bounds in Theorems 2.3 and 2.5.

Proof of Sharpness for Theorems 2.3, and 2.5. Let us recall the bound in Theorem 2.5

$$\frac{\sum_{-n \le j \le n} |\hat{f}(j)|^2}{\|\hat{f}\|^2} \ge \frac{n}{N}. \tag{25}$$

We will show that if $N = 4 * m$, and $n = m - 1$, the left and right hand sides will be equal in the limit as $m \to \infty$. We now consider the function $\frac{2}{\sqrt{m}}\mathbf{III}^4$, which has the function $\mathbf{III}^m$ as its Fourier transform. Note that $\mathbf{III}^m$ consists of exactly 4 equally spaced "spikes" of height 1. If $n = m - 1$, the right of (25) is given by

$$\frac{m-1}{N} = \frac{m-1}{4 * m}.$$

As $m \to \infty$ this approaches 1/4, as $\frac{m-1}{m} \to 1$.

The left side of (25) is equal to 1/4, for all m, since only one of the 4 "spikes" will be included in the upper sum, while the lower norm consists of the sum of all 4 spikes.

This demonstrates that there are subgroups and groups which come arbitrarily close to satisfying the bounds of Theorem 2.3, and Theorem 2.5.

♠

Comment: The above examples can be generalized in the following way. Letting $N = 6*m$, $n = m-1$, and our function $f = \sqrt{\frac{6}{m}}\mathbf{III}^6$, we will get the Fourier transform $\mathbf{III}^6$. The left hand side of repeat will be 1/6, for all m, and the right hand side will approach 1/6 as $m \to \infty$. This type of construction will work any time $N = 2 * q * m$, showing that equality can be approached whenever the right hand is of the form $\frac{m-1}{2qm} \approx \frac{1}{2q}$.

3 Applications

3.1 Stability of Positivity Constrained Inverse Problems

3.1.1 Background

As discussed in Section 1, a linear stability analysis of an operator of the form (3) would amount to the computation of the singular vectors of the operator, along with the singular values of the operator. This would produce an error ellipse around the desired signal to be recovered. The corresponding nonlinear optimization problem (4), in which positivity is taken into consideration, would project this error ellipse onto the cone of positive functions. The calculation of the projection of this error ellipse onto the cone of positive functions for an arbitrary positive signal s, and operator L would solve the stability question for these types of problems.

As mentioned, our approach is to treat the recovery of positive functions as a constrained optimization problem. As a first step towards understanding the stability of this approach, in this section we present some numerical experiments in which we show the results of this approach for several basic signals with respect to varying SNRs. In each case we attempt to invert low frequency band-pass operators. These are Monte Carlo in the sense of the varying additive noise.

Our inspiration comes from the work of Donoho and Stark [8, 6] which shows that if signals are sparse enough, then recovery of the original signal from the low frequency portion of the signal is possible. To be specific, assume that we have a function on $\mathbf{Z}$, and its corresponding Fourier transform on $[-\pi, \pi]$. If the Fourier transform of that function is known on $[-\frac{\pi}{2}, \frac{\pi}{2}]$, Donoho proves that stable recovery is possible if there is only 1 nonzero value on any 8 consecutive integers. We present numerical evidence that with the additional constraint of positivity, an even more general class of functions may admit stable recovery.

Our experience is that whenever a function approaches the bound of (23), then recovery of the positivity constrained inverse problem (4) will be very different from the corresponding linear inverse problem, in that it will be more stable. This fuzzily defined class of signals contains the positive signals within the class considered in [8, 6] for which stable recovery is shown, but is more general. On the other hand [6] does not assume positivity, and hence also shows stability for some non-positive sparse signals, a class which we do not consider.

The following two theorems show some link between sparsity, positivity and the bound (23). We anticipate that there is room for improvement of these results.

If a positive function s can be stably recovered from its low frequency Fourier components, then it must follow that there is no high frequency perturbation p_h, which can be added to s such that the sum $s + s_h$ is also positive. This has been proven this for certain classes of sparse functions [6]. We believe that this is also essentially true for any function which approaches the equality in (23). To this end we present the following theorem.

Theorem 3.1 *Let a positive function $s = s_l + s_h$ achieve the equality in (23) in Theorem 2.5, where s_l is its low frequency component (known) and s_h is its high frequency (unknown) component. Then there is no high frequency perturbation p_h in the orthogonal complement of s_h such that $s + p_h$ is positive.*

Proof: If such a p_h did exist, $s + p_h$ would violate inequality (23) in Theorem 2.5, since its energy ratio would be strictly less than that of s, which achieves equality in (23).

♠

Theorem 3.1 implies that in the M-dimensional high frequency space, there is an $(M-1)$-dimensional subspace from which no perturbation p_h can be added to s and have s remain positive, unless a corresponding change in s_h is made. Thus, in essence, the perturbations which will allow s to remain positive are nearly one-dimensional.

The relationship between sparse functions and the bound (2.5), seems to be that sparse functions generally approach the bound. As a first step we show that any function which satisfies equality of the bound in (2.5) must be zero in at least 2 places.

Theorem 3.2 *If s is a positive function on $\mathbf{Z}/2n\mathbf{Z}$ which satisfies the equality of the bound in Theorem 2.5, then s must be zero in at least 2 places.*

Proof: Suppose that s is zero at only one integer i. Choose the perturbation function p to be $p = [1, -1, 1, -1, 1, -1....]$. The Fourier transform of p is only non-zero at the highest frequency component of $\hat{p}$. If p is positive at i, where s is zero, then we can add a multiple of p to s while keeping s positive, i.e. $s_p = s + \min_{j \neq i}(s_j)p \geq 0$. If p is negative at i then we can similarly subtract a multiple of p. Since s_p has the same amount of "low frequency" energy as s, and more "high frequency" information it follows that the ratio (21) will be smaller for s_p than for s. But we have assumed that s achieved the equality, which implies that s_p would violate inequality (23). Thus s must have a least two zeros.

♠

Theorem 3.2 connects the energy criterion (21) and sparsity conditions. In fact, we believe that the connection between sparsity and the energy criterion is much stronger than this.

3.1.2 Monte Carlo studies of positivity constrained inverse problems

We now examine a few enlightening Monte Carlo examples. In our experiments we solve the problem

$$\begin{cases} \text{Minimize:} & ||Lx - (Ls + n)|| \\ \text{Subject to:} & x \geq 0. \end{cases}$$

for several particular signals s. We study the mean values of the error $\|x - s\|_2$ for various realizations of n, where n is uncorrelated Gaussian noise of a certain power. To compute the Monte Carlo examples, we utilize an interior point method for solving the optimization problem (4) with software written by J.-S. Pang.

We will refer to the noise amplification level (NAL) in the following examples. By this we mean that the mean error at a given noise level ϵ is bounded in the manner

$$E(\|x - s\|) \approx NAL\|\epsilon\|.$$

We say that the error scales linearly with the noise if such a bound exists.

We study the recovery of the high frequency component of functions which have significant high frequency content, using only their low-frequency components, and the fact that they are compactly supported. Some of the functions which we will study will nearly achieve the lower bound (23), while others will have significant high frequency energy as measured by the ratio (23). Our principal thesis is that a function with significant high frequency content, i.e. which nearly achieves the equality in (23), will actually be easier to recover from its low frequency component than a function which is mainly low frequency, with a small amount of high frequency information.

To appreciate how specialized the functions are which approach this bound let us consider a function which is generated from uncorrelated Gaussian white noise. We begin by generating 32 outcomes of Gaussian white noise. The energy distribution of white noise is constant across frequency, and as expected, the ratio (23) is .51. This white noise outcome is not positive, however, so we add a constant to it to make it positive, i.e. $g = noise + |\min(noise)|$. Now the energy distribution of g is .91, nearly 1 (the obvious upper bound). This function is illustrated in Figure 1.

Similarly we consider the function $f = sin(16 * x) + 1$, displayed below in Figure 1. This function consists entirely of a pure high frequency component, and the smallest constant which will make the function positive. The ratio for this function is still only .66, which is far from our lower bound of .25. This function is also illustrated in Figure 1.

In all of the following examples, we are working with discrete signals of length 32. We would like to explore 3 questions:

1. When we recover the functions from their low frequency components, do the estimation errors scale linearly with the noise level?

Figure 1: The above two functions are used to illustrate that it is very difficult to approach the lower bound for the energy ratio (23). The first function which was generated from white noise, has an energy bound of .90. The second function consisting of a pure high frequency component plus a constant to make it positive. Both of these functions would appear to have significant high frequency components, but they do not approach the lower bound .25.

2. If the noise level scales linearly, what is the NAL, and does there seem to be a relation to the ratio (23), or sparsity?

3. If the noise level doesn't scale linearly, is there a noise level at which interesting and somewhat reliable results can be attained?

In the first of our examples we study the inversion of the operator $L = F^{-1}KF$ where K is 1 for the lowest half of the frequency domain and 0 on the upper half.

Example 1: Our signal is a positive and sparse function s which coincides in spirit with sparsity constraint in [6], but has more non-zero points than the results of [6] allow, i.e. one non-zero point for every 4 integers as opposed to one for every 8 integers as in [6]. The signal s of this example is also has a relatively low energy bound (23), with a value of .40.

The compact support of the signals will be contained for all of our examples on 32 points of $\mathbf{Z}$. For this signal, the function is also sparse, being non-zero only on every fourth point, having 4 non-zero points. The results are striking, and are illustrated in Figure 2. We begin with an l^2 signal to noise ratio of 100, i.e.

$$SNR = \frac{\|signal\|_2}{\|noise\|_2} = 100,$$

and uncorrelated Gaussian noise. We solve the above optimization problem for 1000 different realizations of uncorrelated Gaussian noise and the average relative mean error was .083. This means that the noise is magnified by a factor of only 8, even though the operator is very ill-conditioned (10^{15}). The standard linear assumption is that the noise will be magnified by a factor on the order of the condition number, so it is apparent that the nonlinearity has greatly affected the outcome of this experiment.

We then increase the SNR to 10^3, and the corresponding mean error from 1000 outcomes is .0083, which once again illustrates a noise magnification of 8. Thus the noise magnification seems to be relatively constant throughout noise levels, or the errors scale linearly with the noise level. This is surprising, and illustrates that positivity and sparsity create a very strong constraint. When we increased the SNR again to 10^4, the mean error was .00085, which illustrates a slightly larger noise magnification but is very close to that demonstrated at lower levels. Thus the inversion problem seems to scale linearly, with a ratio of approximately 8.

Figure 2: An outcome of the Monte Carlo study is shown above. The positive function is the original signal s. The oscillatory function is the low frequency portion of the spiky signal, or Ls, where L is the bandlimiting operator, and the spikes are separated by 3 zeros. The recovered function x agrees with the original function to a high degree.

Example 2: To illustrate the necessity of some sparsity, we then increase the number of non-zero points to 1 in every 3. This function is shown in Figure 3. The energy ratio (23) for this function is also 0.40. For the first example we are able to use a signal to noise ratio of 100 to get an error of 0.08.

The noise magnification error appears to be linear for this problem also, but the constant seemed to be much larger. The respective mean errors from 1000 trials at a SNRs of $10^3, 10^4, 10^5$, and 10^6 are $.310, .04, .004$, and .0004. Thus the errors scaled linearly with the noise, but the average noise magnification was approximately $.04/.0001 = 400$ for each experiment. This is compared to an average noise magnification of only 8 for the spikes which occurred on only every fourth point. Thus the denser the spikes, the more difficult the reconstruction.

This problem once again demonstrates linear scaling of the noise level, but with a higher NAL. This might be expected in light of [6], since it is locally less sparse. This demonstrates the importance of both the energy ratio and sparsity.

Figure 3: An outcome of the Monte Carlo study is shown above. The positive function is the original signal s. The oscillatory function is the low frequency portion of the spiky signal, or Ls, where L is the bandlimiting operator. The spikes are separated by 2 zeros. This function appears to be "stably" recoverable, but with noise magnification errors which are much higher than the example where the spikes are separated by 3 zeros.

Example 3: We will now investigate whether or not a non-sparse signal can be recovered reliably, if it has a energy ratio (23) which is comparable to the examples above. We used the signal in Example 1, but made it non-sparse by adding a small constant to it. While this makes it non-sparse, its energy ratio (23) only increases slightly (from .40 to .44). The results are very interesting, both numerically and visually.

Numerically the mean errors from 1000 trials at SNRs of $10^2, 10^3$, and 10^4, are respectively .31, .21, and .17. This indicates that the noise amplification did not scale linearly, but rather began at 31, increased to 210, and finished at 1700.

Nevertheless, by all appearances, Figure 4 shows that the spikes are approximated rather well. The error is primarily in approximating the low constant value. Thus the interesting content of the signal is still recovered.

Figure 4: The above example was tested to see if essential sparsity, i.e. a sparse signal plus a small positive constant, shares the results of linear noise amplification with the sparse signals. The results suggest that the noise amplification does not scale linearly with the noise, even with only this small constant added to the sparse function. The visual results, however, suggest that most of the error which does not scale was can be attributed to the small constant value. Useful visual results were obtained at all SNR levels.

Example 4: In this example we investigate the stability of recovering a function s which is not sparse in the sense of [6], and whose energy ratio is not as low as the other examples. The outcome of one recovery is illustrated in Figure 5. Although this function is not sparse in the sense defined in [6], there are large, well-distributed regions which are zero in the function. In addition the energy ratio (23) is .80, further from the bound .25 than the spike examples which give .4. There is still significant high frequency energy in the function, as is illustrated by the difference between the low frequency component and the function in Figure 5.

We begin with an SNR of 100, and the recovery is once again relatively stable, although not as stable as the more sparse examples. The outcome of the Monte Carlo study shows a relative mean error of .35 at SNR = 100, which implies that the noise is magnified by a factor of 35. While this error is still quantitatively large, examination of the results in Figure 5 show that the qualitative nature of the signal is recovered relatively well. When we increase the SNR to 10^3 the relative mean error decreased to .23. The noise magnification factor is now a very large 230. We then increased the SNR to 10^4 and found that the relative mean error only decreases to .16, and the magnification factor increases to 1600. A final test at a SNR of 10^5 shows a decrease in the average error to .12, which implies that the noise magnification is a whopping 12000.

This experiment indicates that the recovery of functions which are not sparse in the sense of [6], and not close to the error bound (23), appear to give rise to inverse problems in which the errors do not depend linearly upon the noise level. At the same time, recovery via the solution of the related optimization problem still yields useful results.

Example 4 and Example 3, yield comparable results, even though Example 3 has a much lower energy ratio (.4 versus .8). While this may seem to contradict the thesis that the energy ratio is an indicator of stability, we must point out that Example 3, has no non-zero terms, while in Example 4 there are significant regions in which the signal is zero. We will now examine the results from this signal when a small constant is added to it, as in Example 3.

Figure 5: We conducted a Monte Carlo test on above signal, which is not sparse in the sense of [6]. The results show that although the errors do not scale linearly with the noise level, useful results were obtained with reasonable SNR levels. This is reflected in the recovery of the major "spikes" were recovered, even though the small constant value was not recovered well at all.

Example 5: After adding a small constant to the signal of Example 4, we ran the experiment at SNRs of $10^2, 10^3$, and 10^4 and obtain average errors of .58,.38, and .32. The energy ratio increases from .95 to .96. We now recognize that the average error of recovering this signal is .32 at an SNR of 10^5, while the average error of recovering the signal in Example 3 was .31 at an SNR of 100. Thus the signal in Example 3 seems much easier to recover. This is as predicted, since the signal in Example 3 was non-sparse but had a lower energy ratio than this signal, which is non-sparse with a high energy ratio.

Discussion The examples above give us some insight into our three questions. We will now restate the questions, and conjecture answers.

1. *When we recover the functions from their low frequency components, do the estimation errors scale linearly with the noise level?* The first two examples are sparse, and seem to scale linearly; the second two examples are not sparse, and correspondingly, the errors do not scale linearly with the noise. Thus it would seem that sparsity will be required for linear scaling of errors. This sparsity requirement is less restrictive than that of [6].

2. *If the noise level scales linearly, what is the NAL, and does there seem to be a relation to the ratio (23), or sparsity?* In the first two cases, the example with non-zero points separated by 3 zeros has a NAL of 8, and that with only 2 zeros separating nonzero points has a NAL of 40. Thus the level of sparsity seems to effect the NAL.

3. *If the noise level doesn't scale linearly, is there a noise level at which interesting and somewhat reliable results can be attained?* The last three examples suggest that interesting and useful

results can be obtained, even if the noise levels don't scale linearly. The signal in Example 3 is non-sparse, but with a low energy level. This signal seemed to be recoverable with approximately the same accuracy as the signal of Example 4, which has significant areas which are zero, but has a much higher energy ratio than that of Example 3. Example 5 has a high energy ratio (23), and no zero terms, and as a result was the most difficult to recover.

The conclusion of these numerical studies suggests that the energy ratio of a signal affects the ability to recover the signal from its low frequency components. This effect is the opposite of what might be expected from typical linear inverse theory, however. The more high frequency signal that there is to be recovered, the more stable the recovery will be. In addition to the energy ratio, sparsity plays a big role in the stability of these inverse problems.

3.2 Partial factorial analysis

A 2^k factorial design is a particular designed experiment. In such a situation, the experimenter is interested in the effect of k different factors on the yield of a particular process. Each of the factors may be set at a high or low level, resulting in 2^k different experimental conditions. A typical sort of example (cf. [2], Chapter) is a wheat farmer interested in the effects of combinations of weed killer, sunlight and fertilizer on her plants. We imagine that each of these three factors can be administered at either a high or low level. Each combination corresponds naturally to an element of $(\mathbf{Z}/2\mathbf{Z})^3$. The experiment is carried out at each of the eight combinations some number of times (the replication number) and the plant height is measured. The average yield at each combination then gives an function in $L^2((\mathbf{Z}/2\mathbf{Z})^3)$. Standard "factor analysis" of such data is precisely the corresponding Fourier analysis. The projections may be grouped in a hierarchical manner. The "main effects" are determined as the difference in average yield obtained by holding a given factor high and then averaging over all possible combinations of the remaining factors and then subtracting the average obtained by fixing the same factor to be low. "Two factor interactions" measure analogously the interactions of factors taken two at a time, similarly for higher order interactions. These projections also have interpretations as least square estimates for the coefficients in various models for such data.

In general there are $\binom{k}{m}$ m-factor interactions with $m = 0$ denoting the overall average yield (the projection onto the trivial representation) and $m = 1$ giving the main effects. Again, since for each irreducible representation χ of $(\mathbf{Z}/2\mathbf{Z})^k$ we have $\chi \otimes \overline{\chi} = 1$, we have the following specialization of Theorem 2.3:

Theorem 3.3 *Let $f : (\mathbf{Z}/2\mathbf{Z})^n \longrightarrow \mathbf{C}$, $f \geq 0$, be data from a 2^k factorial design. Let $0 \leq m \leq k$. Let $\|f|_m^2$ denote the energy contained up to and including the m^{th} order interactions. Then*

$$\|f\|_m^2 \geq \frac{1}{\sum_{l=0}^{m} \binom{k}{l}} \|f\|^2.$$

3.3 Partial spectral analysis of ranked data

The previous examples are both classical applications of Fourier spectral analysis on abelian groups. From a group theoretic point of view, these are situations in which we are interested in a decomposition of the data vector as a sum of projections onto eigenvectors for a natural symmetry group for the data. For an abelian group, this is equivalent to the projection onto the subspaces spanned by vectors consisting of the different irreducible characters, or equivalently, the distinct irreducible representations. As Diaconis first pointed out (cf. [5] for many references as well as [18]) in cases in which the domain of the data has a nonabelian symmetry group, this sort of generalized spectral analysis approach still makes sense.

A favorite example is ranked data. Respondents are given a list of n items (eg. movies, restaurants) and asked to rank them in decreasing order of preference. In this way, each respondent picks a permutation of the original list, i.e., an element of the symmetric group on n elements, denoted S_n. After all respondents are surveyed, the resulting data is an integer-valued function on S_n, whose value at a permutation π is the number of respondents choosing ranking π.

Spectral analysis proceeds by computing the projections of the data onto the various isotypic subspaces of the regular representation. The irreducible representations of S_n are indexed by the proper partitions of n.[1] For each such partition λ we have a permutation action of S_n on the ordered partitions of the set $\{1, \ldots n\}$ into blocks of sizes $\lambda_1, \ldots, \lambda_k$ respectively with

$$\pi(\{a_1, \ldots, a_{\lambda_1}\}, \ldots, \{a_{n-\lambda_k}, \ldots, a_n\}) = (\{\pi(a_1), \ldots, \pi(a_{\lambda_1})\}, \ldots, \{\pi(a_{n-\lambda_k}), \ldots, \pi(a_n)\}).$$

Each such partition of the set $\{1, \ldots, n\}$ is called a *tabloid of shape* λ. These representations are in general reducible, but contain a uniquely determined irreducible subspace (Specht module) which is the irreducible representation associated to the partition λ. The books [13, 10] are tremendous resources for the subject.

To give some flavor of the spectral analysis, the partition $(n-1, 1)$ corresponds to the representation which computes the "first-order effect" of how popular, or unpopular a particular choice is (the number of times item j was ranked in position k); the partition $(n-2, 1, 1)$ computes an "ordered second-order effect" of how often a given pair i, j were ranked in positions l and m respectively; the partition $(n-2, 2)$ computes an "unordered second-order effect" of how often a given pair i, j were ranked in positions l and m, with no concern for order. The interpretation of other partitions generalizes this sort of approach. The paper [4] works out a complete example, using the the real data of an American Psychological Association vote.

In analogy with the previous example of factorial designs, it is natural to ask for the percentage of energy contained in the Fourier transforms at the representations whose partitions have a bounded number of parts. The number of parts can be thought of as the number of coalitions that we take into account. A careful application of Theorem 2.3 requires a detailed understanding of the decomposition of the tensor products of irreducible representations. This is a difficult problem in general (see [10], page 61, ex. 4.51) , but we are able to obtain some bound by making use of the following simple standard result.

Theorem 3.4 *[see [13], page 97, Lemma 2.9.16] For any partition λ of n, let η_λ denote the permutation representation of S_n on tabloids of shape λ. Then*

$$\eta_\alpha \otimes \eta_\beta = \sum_\lambda \eta_\lambda$$

where $\lambda = (\lambda_{ij})$ is such that the λ_{ij} are nonnegative integers with

$$\sum_i \lambda_{ij} = \alpha_j \quad \textit{and} \quad \sum_j \lambda_{ij} = \beta_i.$$

To apply Theorem 3.4 we use two simple facts (see eg. [13]). The first is that for each irreducible representation ρ_λ of S_n, the corresponding contragredient representation is equivalent. This follows directly from the fact that the characters are all real-valued. The second fact is that the permutation representations η_λ only contain those irreducibles ρ_α for those partitions α that majorize λ. These two facts (along with Theorem 3.4) allow us to compute $T(S)$ for some special cases.

[1] A proper partition $\lambda = (\lambda_1, \ldots, \lambda_k)$ of n is such that (1) $\sum_i \lambda_i = n$ and (2) $\lambda_1 \geq \cdots \geq \lambda_k > 0$.

Theorem 3.5 *Let d^2 divide n and let S be the set of all partitions of n with at most d parts. Then $T(S)$ is the set of all partitions with at most d^2 parts.*

To give one simple application of this, suppose that n is divisible by 4. Then S is the set of all partitions of at most two parts, and $T(S)$ is all partitions with at most 4 parts. Also,

$$deg(S) = 1 + \sum_{k=1}^{n/2} \left(\binom{n}{k} - \binom{n}{k-1} \right)^2 .$$

This gives us a bound on the information in coalitions into at most 4 groups.

4 Band-limited functions on compact groups

The technique of proof of Theorem 2.3 can be extended to the realm of compact groups, provided we limit ourselves to the band-limited case. In fact, the original problem can be recast in this light. Let our data vector $f = (f(0), \dots, f(2N-1))$ be viewed as equispaced samples of a function $f : S^1 \longrightarrow \mathbf{C}$, so that $f(k) = f(\frac{k}{2N-1})$ (with the circle identified with the unit interval). In general, the Fourier expansion of f is given by

$$f(t) = \sum_{l \in \mathbf{Z}} \widehat{f}(l) e^{-2\pi i l t} \tag{26}$$

where

$$\widehat{f}(l) = \int_0^1 f(t) e^{2\pi i l t} dt. \tag{27}$$

Suppose now that $\widehat{f}(l) = 0$ for $|l| \geq N$. In this case we say that f is *band-limited* with bandwidth N and there is a *quadrature rule* or *sampling theory* for f. That is to say that the Fourier coefficients of any band-limited function can be computed from only a finite set of samples,

$$\widehat{f}(l) = \sum_{k=0}^{2N-2} \frac{1}{2N-1} f\left(\frac{k}{2N-1}\right) e^{-2\pi i k l/(2N-1)} \tag{28}$$

where the factor $\frac{1}{2N-1}$ should be viewed as a (constant) weight function with support at the equispaced points $\{\frac{k}{2N-1}\}_{k=0}^{2N-2}$.

Thus, Theorem 2.5 can also be recast as a statement about the energy in the low frequency of a band-limited nonnegative function on the circle.

In fact, recent work of D. Maslen [15, 16] derives a theory for any compact group completely analogous to the case of band-limited functions on the circle. We now give a simplified version of his work and show how we may use the sampling theory for band-limited functions on a compact group to derive a generalization of Theorem 2.3.

Recall that the representation theory of compact groups differs little from that of finite groups with the caveat that inner products now become integrals instead of summations. In particular, the irreducible representations of a compact group G are all finite dimensional, and any square-integrable function f (with respect to Haar measure) has an expansion in terms of irreducible matrix elements

$$f = \sum_{\lambda \in \Lambda} \sum_{j,k=1}^{d_\lambda} \widehat{f}(\lambda)_{jk} T^\lambda_{jk} \tag{29}$$

where T^λ denotes an irreducible representation of degree $d_\lambda < \infty$, Λ is a countable set and the implied convergence is in the mean. The Fourier coefficients $\{\widehat{f}(\lambda)_{jk}\}$ are computed by

$$\widehat{f}(\lambda)_{jk} = d_\lambda \langle f, T^\lambda_{jk} \rangle = d_\lambda \int_G f(x) T^\lambda_{jk}(x) dx \tag{30}$$

where dx denotes (the translation-invariant) Haar measure. Convolution for functions in $L^2(G)$ is as in the finite case with

$$f \star h(y) = \int_G f(yx^{-1}) h(x) dx \tag{31}$$

with

$$\widehat{f \star h}(\lambda) = \widehat{f}(\lambda)\widehat{h}(\lambda) \tag{32}$$

where $\widehat{f}(\lambda)$ denotes the entire matrix of Fourier coefficients for representation T^λ. The book [22] is an excellent reference for the theory and includes many, many examples.

There is a natural definition of band-limited in the compact case, simply corresponding to those functions whose Fourier expansion has ony a finite number of terms. However, a general notion of low or high frequency is a bit more problemmatic. We present here the simplest version of Maslen's theory, whose full exposition requires some familiarity with the theory of filtered algebras (cf. [16] for details).

For our purposes, the following notion of band-limit (see [17]) will suffice:

Definition 4.1 *Let $\mathcal{R}$ denote a complete set of irreducible representations of a compact group G. Suppose that*

$$\mathcal{R} = \cup_{b \geq 0} \mathcal{R}_b \tag{33}$$

such that

- *[1] $|\mathcal{R}_b| < \infty$ for all $b \geq 0$;*
- *[2] $b_1 \leq b_2$ implies that $\mathcal{R}_{b_1} \subseteq \mathcal{R}_{b_2}$;*
- *[3] $\mathcal{R}_{b_1} \otimes \mathcal{R}_{b_2} \subseteq span_{\mathbf{Z}} \mathcal{R}_{b_1+b_2}$.*

We call the decomposition (33) a system of band-limits *on G.*

Let $f \in L^2(G)$ and suppose that $\{\mathcal{R}_b\}_{b \geq 0}$ is a system of band-limits on G. Then we say f is band-limited with band-limit b *if $\widehat{f}(T^\lambda_{jk}) = 0$ for all $\lambda \notin \mathcal{R}_b$.*

Examples. 1. $G = S^1$. Then let $\mathcal{R}_b = \{\chi_j : |j| \leq b\}$ where $\chi_j(z) = z^j$. Then $\chi_j \otimes \chi_k = \chi_{j+k}$ and the corresponding notion of band-limited (as per Definition 4.1) coincides with the usual notion.

2. $G = SO(3)$. In this case the irreducible representations of G are indexed by the nonnegative integers with V_λ the unique irreducible of dimension $2\lambda + 1$ (see eg. [22], Chapter IX). Let $\mathcal{R}_b = \{V_\lambda : \lambda \leq b\}$. Using the well-known Clebsch-Gordon relations

$$V_{\lambda_1} \otimes V_{\lambda_2} = \sum_{j=|\lambda_1 - \lambda_2|}^{\lambda_1+\lambda_2} V_j \tag{34}$$

we see that this is a system of band-limits for $SO(3)$. Notice that when restricted to the quotient $S^2 \cong SO(3)/SO(2)$, band-limits are described in terms of the highest order spherical harmonics that appear in a given expansion.

For our purposes, the importance of developing the band-limited theory is that in this case there exists a *sampling theory* or *quadrature rule* that allows the Fourier coefficients to be computed exactly as finite sums.

Theorem 4.2 ([15]) *Let G be compact with a system of band-limits $\{\mathcal{R}_b\}_b$. For any band-limit b, there exists a finite set of points $X_b \subset G$ such that for any function $f \in L^2(G)$ of band-limit b,*

$$\widehat{f}(T^\lambda_{jk}) = \sum_{x \in X_b} f(x) T^\lambda_{jk}(x) w(x) \tag{35}$$

for all $\lambda \in \mathcal{R}_b$ and some weight function w on X_b.

We can now state our final theorem.

Theorem 4.3 ([15]) *Let G be compact with a system of band-limits $\{\mathcal{R}_b\}_b$. Furthermore, assume that $\mathcal{R}_b$ is closed under taking contragredients. (In particular, notice that this is true for the examples of S^1 and $SO(3)$.) Let $f \in L^2(G)$ be nonnegative and of band-limit B and let $b \leq B/2$. Then*

$$\frac{\|f\|_b^2}{\|f\|^2} \geq \frac{\left[\sum_{\lambda \in \mathcal{R}_b} d_\lambda\right]^2}{M(S)} w_{max} \tag{36}$$

where $\|f\|_b$ is the length of the projection of f onto the irreducible matrix elements in $\mathcal{R}_b$, $M(S)$ is the maximum multiplicity that occurs in the irreducible decomposition of $\mathcal{R}_b \otimes \mathcal{R}_b$ and w_{max} is the maximum weight among $\{w(x)\}_{x \in X}$.

Proof: By restricting ourselves to the band-limited case, we can adapt the proof of Theorem 2.3 as follows.

Define f^c as $f^c = \tilde{f} \star f$ where $\tilde{f}(x) = f(x^{-1})$. By definition,

$$\begin{aligned} f^c(z) &= \int_G \tilde{f}(zx^{-1}) f(x) dx \\ &= \int_G f(xz^{-1}) f(x) dx. \end{aligned} \tag{37}$$

By (31)

$$\begin{aligned} \widehat{f^c}(\alpha) &= \widehat{\tilde{f}}(\alpha) \widehat{f}(\alpha) \\ &= \widehat{f}(\alpha)^\dagger \widehat{f}(\alpha) \end{aligned} \tag{38}$$

so that

$$\begin{aligned} \langle f^c, \chi_\alpha \rangle &= \langle f^c, T^\alpha_{1,1} + \cdots + T^\alpha_{d_\alpha, d_\alpha} \rangle \\ &= \sum_{j,k} |\widehat{f}(\alpha)_{j,k}|^2 \cdot \tfrac{1}{d_\alpha} \\ &= \|f\|_\alpha^2. \end{aligned} \tag{39}$$

As before we now consider $\langle f^c, \chi_b \otimes \overline{\chi_b} \rangle$ where $\chi_b = \sum_{\alpha \in \mathcal{R}_b}$. Suppose that

$$\chi_b \otimes \overline{\chi_b} = \sum_{\alpha \in \mathcal{R}_b} n_\alpha \chi_\alpha.$$

Thus, on the one hand

$$\begin{aligned} \langle f^c, \chi_b \otimes \overline{\chi_b} \rangle &= \langle f^c, \textstyle\sum_{\alpha \in \mathcal{R}_b} n_\alpha \chi_\alpha \rangle \\ &\leq (\max_\alpha n_\alpha) \textstyle\sum_{\alpha \in \mathcal{R}_{2b}} \|f\|_\alpha^2. \end{aligned}$$

On the other hand,

$$\begin{aligned} \langle f^c, \chi_b \otimes \overline{\chi_b} \rangle &= \int_G f^c(x) |\chi_b(x)|^2 dx \\ &= \int_G f^c(x_0^{-1}x) |\chi_b(x_0^{-1}x)|^2 dx \\ &= \textstyle\sum_{x \in X} f^c(x_0^{-1}x) |\chi_b(x_0^{-1}x)|^2 w(x) \\ &\geq f^c(\mathrm{id}) |\chi_b(\mathrm{id})|^2 w(x_0) \end{aligned}$$

where

$$w(x_0) = max_{x \in X} w(x).$$

Noting that $f^c(\mathrm{id}) = \|f\|_2^2$ and $|\chi_b(\mathrm{id})|^2 = \left(\sum_{\alpha \in \mathcal{R}_b} d_\alpha\right)^2$ concludes the proof.

♠

5 Final remarks

There remain several open questions suggested by this work:
1. **Further extensions.** Recall that the original Erdos theorem can be interpreted as a statement about the spectrum of a certain class of nonnegative functions on $\mathbf{Z}$. It would be interesting if similar theorems could be obtained for other discrete groups, noncompact groups, or even if a fuller theory for compact groups could be developed.

2. **Tight bounds.** Can examples akin to that constructed in Section 2.4 be constructed for general noncommutative groups to show the bounds of Theorem 2.3 to be sharp?

3. **Sparsity-energy concentration relation.** Can Theorem 3.2 be improved upon? Is there a deeper connection between sparsity and energy concentration?

References

[1] E. W. Cheney, *Introduction to approximation theory*, Second Edition, Chelsea Publishing, NY 1982.

[2] G. Constantine, *Combinatorial theory and experimental design*, Wiley, NY 1985.

[3] J. W. Cooley and J. W. Tukey, An algorithm for machine calculation of complex Fourier series, *Math. Comp.*, **19** (1965), pp. 297–301.

[4] P. Diaconis, A generalization of spectral analysis with applications to ranked data, *Ann. Statistics* **17** (1989), pp. 949–979.

[5] P. Diaconis, *Group representations in probability and statistics*, IMS, Hayward, CA, 1988.

[6] D. L. Donoho, Superresolution via sparsity constraints, *SIAM Jour. Math. Anal.*, **23**(5), (1992), pp. 1309–1331.

[7] D. L. Donoho and B. F. Logan, Signal recovery and the large sieve, *SIAM Jour. Appl. Math.*, **52**(2) (1992), pp. 577–591.

[8] D. L. Donoho, and P. B. Stark, Uncertainty principles and signal recovery, *SIAM Jour. Appl. Math.*, **49**(3) (1989), pp. 906–931.

[9] P. Erdös and W. H. J. Fuchs, On a problem in additive number theory, *Jour. London Math. Soc.*, **31** (1956), pp. 67–73.

[10] W. Fulton and J. Harris, *Representation theory*, Springer-Verlag, NY 1991.

[11] D. Jackson, *Über die Genauigkeit det Annäherung Stetiger Functionen Durch Ganze Rational Functionen Gegenbenen Grades und Trigonometrische Summen Gegebener Ordnung*, Dissertation, Göttigen, 1911.

[12] D. Jackson, On approximation by trigonometric sums and polynomials, *Trans. of the Amer. Math. Soc.*, **13** (1912), pp. 491–515.

[13] G. D. James and A. Kerber, *The representation theory of the symmetric group*, Addison-Wesley Publishing Co., MA 1981.

[14] B. F. Logan, An interference problem for exponentials, *Mich. Math. Jour.*, **35** (1988), pp. 369–401.

[15] D. Maslen, Efficient computation of Fourier transforms on compact groups *J. Fourier Anal. Appl.* **4**(1) (1998).

[16] D. Maslen, Sampling of functions and sections for compact groups, Technical Report PMA-TR99-193, Department of Mathematics, Dartmouth College, Hanover, NH, October, (1999).

[17] D. Maslen and D. Rockmore, Generalized FFTs, in *DIMACS Series in Disc. Math. and Theor. Comp. Sci.*, Volume 28, L. Finkelstein and W. Kantor (eds.), 1997, pp. 183–237.

[18] D. Maslen and D. Rockmore, Some applications of generalized FFTs, in *DIMACS Series in Disc. Math. and Theor. Comp. Sci.*, Volume 28, L. Finkelstein and W. Kantor (eds.), 1997, pp. 329–369.

[19] M. Reed and B. Simon, *Methods of mathematical physics I, functional analysis*, Academic Press, New York 1980.

[20] J. P. Serre, *Linear representations of finite groups*, Springer-Verlag, New York 1977.

[21] H. S. Shapiro, Majorant problems for Fourier coefficients, *Quart. Jour. Math.*, Oxford Series, **26**(2) (1975), pp. 9–18.

[22] N. J. Vilenkin, *Special functions and the theory of group representations,* Translations of Mathematical Monographs, Vol. 22, American Mathematical Society, Providence, RI 1968.