Direct Construction of Superoscillations

Dae Gwan Lee and Paulo Jorge S. G. Ferreira

***Abstract*—Oscillations of a bandlimited signal at a rate faster than its maximum frequency are called "superoscillations" and have been found useful e.g., in connection with superresolution and superdirectivity. We consider signals of fixed bandwidth and with a finite or infinite number of samples at the Nyquist rate, which are regarded as the adjustable signal parameters. We show that this class of signals can be made to superoscillate by prescribing its values on an arbitrarily fine and possibly nonuniform grid. The superoscillations can be made to occur at a large distance from the nonzero samples of the signal. We give necessary and sufficient conditions for the problem to have a solution, in terms of the nature of the two sets involved in the problem. Since the number of constraints can in general be different from the number of signal parameters, the problem can be exactly determined, underdetermined or overdetermined. We describe the solutions in each of these situations. The connection with oversampling and variational formulations is also discussed.**

***Index Terms*—Superoscillations, signal design, algorithms, Hilbert space, interpolation, nonuniform sampling, sampling methods, matrices, numerical stability.**

I. INTRODUCTION

A. Statement of the Problem

SUPEROSCILLATIONS occur when a bandlimited signal oscillates faster than its maximum frequency. This interesting phenomenon has already found a number of applications, e.g., in quantum mechanics, superdirectivity and superresolution. In fact, it has been argued that [1] "super-oscillation-based imaging has unbeatable advantages over other technologies". The basic principle is that a superposition of Fourier components with a bounded spectrum e.g., $|\lambda_k| < 1$ can still approximate an oscillation of frequency higher than the bandlimit:

$$\sum c_k e^{i2\pi\lambda_k t} \to e^{i2\pi a t}.$$

Manuscript received March 05, 2014; revised April 23, 2014; accepted April 24, 2014. Date of publication April 30, 2014; date of current version May 19, 2014. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Antonio Napolitano. The work of D. G. Lee is partially supported by the National Research Foundation of Korea (NRF)(2012R1A1A2038650). The work of P. Ferreira was partially supported by the European Fund for Regional Development (FEDER) through the Operational Program Competitiveness Factors (COMPETE) and by the Portuguese Foundation for Science and Technology (FCT), in the context of projects FCOMP-01-0124-FEDER-022682 (FCT reference PEst-C/EEI/UI0127/2011) and Incentivo/EEI/UI0127/2013.

D. G. Lee is with the Department of Mathematical Sciences, KAIST, Daejeon 305-701, South Korea.

P. J. S. G. Ferreira is with DETI/IEETA, Universidade de Aveiro, Aveiro 3810-193, Portugal (e-mail: pjf@ua.pt).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2014.2321119

There are several ways of obtaining superoscillations. They range from simple zero replacement to variational approaches that seek to minimize the energy of the signal while forcing it to attain certain amplitudes and/or derivatives. A brief account of these approaches and some of their applications is given in Section I.B.

Any discussion of superoscillations must be carried out with a reference bandlimit or rate in mind (sometimes, it is more appropriate to consider a reference scale, as argued in [2]). The reference rate can be normalized without loss of generality. In this paper, to simplify the notation, we restrict our attention to $PW_{1/2}$ (the Hilbert space of square-integrable functions bandlimited to $1/2$ Hz). We consider bandlimited functions whose amplitudes are determined by coefficients a_k:

$$f(t) = \sum_{k \in I} a_k \operatorname{sinc}(t-k)$$

where $I \subseteq \mathbb{Z}$ is a (possibly infinite) fixed index set and

$$\operatorname{sinc} t = \frac{\sin(\pi t)}{\pi t}.$$

The cardinal property of $\operatorname{sinc} t$ (i.e., $\operatorname{sinc} n = \delta_{n,0}$ for $n \in \mathbb{Z}$, where $\delta_{k,\ell}$ is the Kronecker delta) implies that $f(k) = a_k$ for any $k \in I$ and $f(k) = 0$ otherwise. This tends to concentrate the energy of f mainly on the region containing the index set I.

We consider the problem of building superoscillations controlled by their behavior at the scale T (with $T < 1$). That is, we want to find signals of the form

$$f(t) = \sum_{k \in I} f(k) \operatorname{sinc}(t-k) \tag{1}$$

that satisfy a set of constraints

$$f(kT) = c_k, \quad k \in J, \tag{2}$$

where J is a finite index set.

Any signal f of the form (1) is completely determined by its samples $\mathbf{f} = \{f(k)\}_{k \in I}$. It will be convenient to distinguish the sampling grid determined by I from others; when necessary, we will refer to it as the reference grid. Note that it is denser when I consists of contiguous integers.

The conditions (2) constrain the behavior of f on a grid determined by T, possibly much finer than the reference grid. Thus, the constraints (2) may force f to oscillate at a much higher frequency than its bandlimit $1/2$ Hz. Furthermore, the sets I and J can be contained in intervals that are far apart, a possibility that reduces the apparent influence of the constraints on the coefficients of f. These issues raise the question of whether the conditions (2) can be realized at all.

The case $|I| = |J|$ was considered in [3], using oversampled sampling expansions. Its starting point is a pair of relations that specify the mapping between the samples taken at the reference rate and the samples at scale T:

$$f(\ell T) = \sum_{k=-\infty}^{\infty} f(k)\mathrm{sinc}(\ell T - k),$$
$$f(\ell) = T \sum_{k=-\infty}^{\infty} f(kT)\mathrm{sinc}(\ell - kT).$$

In the present paper we take a different and more general approach, which is directly based on the (2). In matrix form, they can be written as

$$A\,\mathbf{f} = \mathbf{c} \tag{3}$$

where

$$A = [\mathrm{sinc}(jT - k)]_{j\in J, k\in I}, \tag{4}$$

and

$$\mathbf{f} = \{f(k)\}_{k\in I}, \quad \mathbf{c} = \{c_k\}_{k\in J}.$$

The differences between our approach and the one in [3] are as follows. First, we do not restrict ourselves to $|I| = |J|$, that is, to square matrices A. We also consider the underdetermined and overdetermined cases. Second, we consider the more general case

$$f(t_k) = c_k, \quad k = 1, 2, \ldots, K,$$

in which the constraints are given on irregularly spaced points $t_1 < t_2 < \cdots < t_K$, so that the matrix A becomes

$$A = [\mathrm{sinc}(t_j - k)]_{t_j\in J, k\in I}. \tag{5}$$

Third, we allow the cardinality of I to be infinite, unlike [3], which considers only finite sets. Of course, when $|I| = \infty$, A should be interpreted as the linear operator from $\ell^2(I)$ into $\mathbb{R}^{|J|}$ with the given matrix representation.

We determine necessary and sufficient conditions for A to have full rank. The conditions depend on the nature of the sets I and J. More precisely, we find that the number of integer elements of J and whether they belong to I play a crucial role in the matter. Of course, if A is square we obtain necessary and sufficient conditions for the nonsingularity of A.

We derive expressions for the energy of the signal and compare it with the variational approach [4]. This allows us to establish a connection between the two seemingly different approaches to the construction of superoscillations: the variational approach to minimum energy constrained interpolation, based on the Euler-Lagrange equation, and the direct approach studied in the present paper, based on the direct specification of the signal value at a different scale.

B. Related Work

Superoscillations are a key component of Aharonov's weak measurement formalism [5], [6]. Aharonov's introduced the function

$$f_n(t, a) = (\cos(2\pi t/n) + ia\sin(2\pi t/n))^n, \tag{6}$$

where $a > 1$. In an interval sufficiently close to the origin, one has $f_n(x, a) \approx e^{i2\pi at}$, i.e., the function oscillates at a frequency that can greatly exceed the bandlimit [5].

Superoscillations can be generated in a very simple way by zero manipulation, and some of the early works that came close to explicit formulations employed that technique. These works form, so to speak, the pre-history of superoscillations. Bond and Cahn [7] were among the first to explore the representation and manipulation of signals through their zeros, as discussed in [8]. They solved the problem of synthesizing a bandlimited function with a given set of zeros and also observed the impact of the distribution of the zeros on the amplitude of the generated bandlimited signal. Landau investigated how large a bandlimited function of unit energy could become when it is constrained by having more zeros than the Nyquist rate allows [9]. The earliest work of which we are aware that uses a variational approach to find the bandlimited signal of minimum energy that interpolates N given data is [10]. Requicha ([11], p. 319) gave a zero-replacement procedure by means of which a bandlimited function such as $\sin kx/x$ can be locally modified to acquire the zeros of another function, possibly of higher bandwidth. He mentioned that the phenomenon was "counter-intuitive" and noted that "abrupt changes in the local density of zeros cause extreme fluctuations in the functions' moduli". Requicha's zero-replacement idea seems to have been rediscovered later [12], this time already in connection with superoscillations.

Turning now to the period that followed the explicit introduction of superoscillations by Aharonov *et al.*, we find Berry's work [13]. Using integral representations, he investigated the amplitude of superoscillating signals in the region of normal oscillation and found it to be exponentially larger than the superoscillations. The energy cost of superoscillations was considered in [4], [14], which discuss the dynamical range and energy required by superoscillating signals as a function of the superoscillation's frequency, number, and maximum derivative. These papers also discuss some of the implications of superoscillating signals, in the context of information theory and time-frequency analysis, and establish that the required energy grows exponentially with the number of superoscillations, and polynomially with the reciprocal of the bandwidth or the reciprocal of the period of superoscillation. The article [4] also shows that there is no contradiction between Shannon's capacity formula and superoscillating signals, and explains the role that the amplitude and energy of such signals play in the matter. Aharonov *et al.* [15] discuss some of the mathematical properties of superoscillations, and in particular the approximation properties of the sequence (6). They show that for $|x| \leq M$, where M is a fixed real number, the sequence converges uniformly to e^{iax}, the frequency of which can be arbitrarily large.

The problem of optimizing superoscillatory signals was considered in [16]. The authors maximize the superoscillation yield, that is, the ratio of the energy in the superoscillations to the total energy of the signal, given the range and frequency of the superoscillations. The constrained optimization leads to a generalized eigenvalue problem, which has to be solved numerically. The work [17] introduces a periodicity measure and applies it to yield-optimizsed superoscillating signals, and [18] investigates the stability of superoscillations, that is, the extent to which they are affected by small deviations in the Fourier coefficients.

The work [2] addresses superoscillations from a different angle, emphasizing scale rather than frequency and discussing the extent to which an arbitrarily narrow pulse can be constructed by linearly combining arbitrarily wider pulses. Interestingly, Aharonov *et al.* ([19], p. 2967) had pointed out long ago that a superposition of Gaussians centred between -1 and 1 could yield a Gaussian centred at 3. However, in [2] the pulses are in fact of different width, and the matter is treated from a very different viewpoint.

Beethoven's ninth symphony, mentioned in [13] as the target of an experiment in superoscillations, is featured in [20], where superoscillations are applied to the problem of transplanckian frequencies in black hole radiation. Berry [21] discusses superoscillations in the context of a quantum billiards problem. The article [22] considers superoscillations in quantum mechanical wave functions, and some unusual associated phenomena of interest from the viewpoint of thermodynamics, information theory and measurement theory.

The interest in superoscillating functions is relatively recent but a number of applications have already been described. A brief overview can be found in [8]. The article [23] discusses optical superresolution without evanescent waves and [24] proposes an array of nanoholes in a metal screen to focus light into subwavelength spots in the far-field, the formation of which is related to superoscillations, without contributions from evanescent fields. The article [25] discusses approaches capable of beating the diffraction limit and [26] proposes a solution that is also free from evanescent fields. An optical mask is used to create superoscillations by constructive interference of waves, leading to a subwavelength focus. The authors also demonstrate that the mask can be used also as a superresolution imaging device. The paper [1] reports subwavelength resolution down to $\lambda/6$ and concludes that "super-oscillation-based imaging has unbeatable advantages over other technologies".

The method introduced in [27], [28] explores the relation between superdirectivity and superoscillation and leads to subwavelength focusing schemes in free space and within a waveguide. The authors demonstrate subwavelength focusing down to 0.6 times the diffraction limit, five wavelengths away from the source. The work [29] demonstrates a superoscillatory sub-wavelength focus in a waveguide environment. The authors claim the formation of a focus at 75% the spatial width of the diffraction limited sinc pulse, 4.8 wavelengths away from the source distributions.

A function and its Fourier transform cannot both be sharply localized, but the work [30] tries to get around this. The authors seek to arbitrarily compress a temporal pulse and report the design of a class of superoscillatory electromagnetic waveforms for which the sideband amplitudes, and hence the sensitivity, can be regulated. They claim a pulse compression improvement of 47% beyond the Fourier transform limit.

The article [31] argues that random functions, defined as superpositions of plane waves with random complex amplitudes and directions, have regions that are naturally superoscillatory. It also derives the joint probability density function for the intensity and phase gradients of isotropic complex random waves in any number of dimensions. The connections between information theory and spectral geometry are used in [32] to obtain results on a quantum gravity motivated natural ultraviolet cutoff which describes an upper bound on the spatial density of information. The article [33] deals with superoscillations in monochromatic waves in several dimensions. Berry [34] shows that waves involving Bessel functions can oscillate faster than their bandlimited Fourier transforms suggest, with the superoscillations being fastest near phase singularities. Other applications to physics include [35], [36] and [37], this last one on backflow, a phenomenon related to superoscillation.

II. Results

The nature of the (3) suggests a least-squares approach to the problem of constructing superoscillations. The fact that the minimum-norm least-squares solution can be conveniently expressed by means of the pseudo-inverse of A, denoted by $A^\dagger$, suggests the consideration of equations such as $\mathbf{f} = A^\dagger \mathbf{c}$. However, the matter is not as simple as that.

The construction of superoscillations is in general a numerically difficult problem, and there is no *a priori* reason to believe that $A^\dagger \mathbf{c}$ can be found without difficulties.

The first obstacle is the rank of A. The computation of the pseudo-inverse $A^\dagger$ is an ill-posed problem when the matrix A does not have full rank. If the matrix A, of size $m \times n$, does not have full rank, then it is possible to change its rank by an arbitrarily small perturbation. It is shown in [38] that for any real c and $\epsilon > 0$, there exists a matrix E with norm smaller than ϵ and such that

$$\|(A+E)^\dagger - A^\dagger\| \geq c.$$

Because of this, our first task is to give a necessary and sufficient condition for A to have full rank.

Understanding the conditions under which A is of full rank is important in computing $A^\dagger$ in a stable way and in understanding when matrices such as AA^T and A^TA are invertible, but we remark that in practice the computation of $A^\dagger$ is best done by means of the QR factorization or the singular value decomposition, without explicitly finding matrices such as $(AA^T)^{-1}$ or $(A^TA)^{-1}$ [38], [39].

A. The Rank of A

We start with a lemma that involves two sets $\tilde{I}$ and $\tilde{J}$. The former contains only integers and the latter only non-integers. More precisely, fix an integer $K \geq 1$; then $\tilde{I} = \{n_k\}_{k=1}^K \subset \mathbb{Z}$, where $n_1 < n_2 < \cdots < n_K$, and $\tilde{J} = \{t_k\}_{k=1}^K \subset \mathbb{R}\backslash\mathbb{Z}$, where $t_1 < t_2 < \cdots < t_K$.

Lemma 1: The matrix

$$\tilde{A} := [\operatorname{sinc}(t_k - n_\ell)]_{k,\ell=1}^K$$

is invertible.

Proof: Observe that

$$\begin{aligned}\operatorname{sinc}(t_k - n_\ell) &= \frac{\sin[\pi(t_k - n_\ell)]}{\pi(t_k - n_\ell)} \\ &= (-1)^{n_\ell}\frac{\sin(\pi t_k)}{\pi(t_k - n_\ell)} \neq 0, \quad t_k \in \tilde{J}, n_\ell \in \tilde{I},\end{aligned}$$

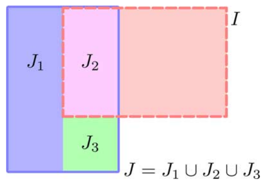

Fig. 1. The sets used in the proof of Theorem 1.

since no t_k is an integer. This means that the determinant of $\tilde{A}$ can be expressed in the following form:

$$\begin{aligned}|\tilde{A}| &= \left| \left[(-1)^{n_\ell} \frac{\sin(\pi t_k)}{\pi(t_k - n_\ell)} \right]_{k,\ell=1}^{K} \right| \\ &= \left(\frac{1}{\pi^K} \prod_{k=1}^{K} \sin(\pi t_k) \right) \cdot \left| \left[(-1)^{n_\ell} \frac{1}{t_k - n_\ell} \right]_{k,\ell=1}^{K} \right| \\ &= \left(\frac{1}{\pi^K} \prod_{k=1}^{K} \sin(\pi t_k) \right) \cdot \left(\prod_{\ell=1}^{K} (-1)^{n_\ell} \right) \\ &\quad \cdot \left| \left[\frac{1}{t_k - n_\ell} \right]_{k,\ell=1}^{K} \right|.\end{aligned}$$

Using a known fact (see [40], p. 92, Problem 3)

$$\left| \left[\frac{1}{a_k + b_\ell} \right]_{k,\ell=1}^{K} \right| = \frac{\prod_{k>\ell}(a_k - a_\ell)(b_k - b_\ell)}{\prod_{k,\ell}(a_k + b_\ell)},$$

we have

$$\left| \left[\frac{1}{t_k - n_\ell} \right]_{k,\ell=1}^{K} \right| = \frac{\prod_{k>\ell}(t_k - t_\ell)(n_\ell - n_k)}{\prod_{k,\ell}(t_k - n_\ell)} \neq 0$$

so that $|\tilde{A}| \neq 0$. ∎

We now consider the rank of A, the matrix given by (5). Before going on it is convenient to partition the elements of $J = \{t_k\}$ in three sets (cf. Fig. 1):

- The set J_1, which contains the non-integer elements of J.
- The set J_2, which contains the integer elements of J that are also in I.
- The set J_3, which is the set of integer elements of J that are not in I.

Some of these sets may be empty. For example, if J consists of numbers of the form mT, with $m \neq 0$ and T irrational, then J_2 and J_3 are empty.

By definition, the set J_1 contains no integers. Hence, by Lemma 1, the square submatrix A_1 formed with the rows of A that correspond to J_1 is invertible.

The set J_2 contains the common elements of J and I, if any. When it is nonempty, it determines a principal submatrix of A that can be written $A_2 := A(J_2, J_2)$.

We are now ready to show the following theorem.

Theorem 1: Let $A = [\mathrm{sinc}(t_m - n)]_{t_m \in J, n \in I}$ where $I \subseteq \mathbb{Z}$ and $|J| < \infty$.

a) Assume that $|I| \geq |J|$. Then A is of full rank if J_3 is empty, i.e., there are no integer elements of J that are not in I.

b) Assume that $|I| \leq |J|$. Then A is of full rank if and only if the number of integer elements of J that are not in I satisfies $|J_3| \leq |J| - |I|$.

Remark 1: Both conditions agree if $|I| = |J|$. Furthermore, they are equivalent to the following single statement: A is of full rank if and only if

$$|J_3| \leq \max(0, |J| - |I|).$$

Proof: Note that A is a matrix of size $|J| \times |I|$, where $|J|$ is finite but $|I|$ may be infinite. To show that A is of full rank, we only need to find a full size square submatrix of A which is invertible. Thus, if $L := \min\{|I|, |J|\}$, it is enough to find an $L \times L$ submatrix of A which is invertible.

(a) Assume that $J_3 = \emptyset$, so that $J = J_1 \cup J_2$. Let A_2 be the principal submatrix of A associated with J_2. Since each $t_k \in J_2$ matches one element of I, denoted by $n_{k'}$, we have

$$(A_2)_{k\ell} = \mathrm{sinc}(t_k - n_\ell) = \mathrm{sinc}(n_{k'} - n_\ell) = \delta_{k',\ell},$$

a matrix that has unit determinant. As a result, the determinants of A and A_1 have the same absolute value, where A_1 is a submatrix of A associated with J_1. By definition, J_1 contains no integers. Thus, by Lemma 1, A is invertible.

Conversely, assume that $|I| \geq |J|$ and there exists some t_0 in J such that $t_0 \in \mathbb{Z} \backslash I$. Note that since $\mathrm{sinc}(t_0 - n) = 0$ for any $n \in I$, the row of A associated with t_0 is zero. Therefore the rank of A is less than $|J|$, that is, A is not of full rank.

(b) Assume that $|I| \leq |J|$ and $|J_3| \leq |J| + |I|$. Consider submatrices of A associated with J_1 and J_2. As in the previous case, the matrix associated with J_2 is a permutation of the identity and the other matrix is nonsingular, by Lemma 1. As a result, A is nonsingular. The result remains clearly valid if J_1 or J_2 are empty.

Conversely, assume that $|I| \leq |J|$ and $|J_3| > |J| - |I|$. This means that there is at least one integer element of J that is not an element of I. To each of these elements there corresponds a zero row of A. As a result, there are more than $(|J| - |I|)$ zero rows in A, and hence A cannot have full rank. This completes the proof. ∎

When A is given by (4), we have the following result.

Corollary 1: When $|I| \geq |J|$ and J contains only multiples of T, A is of full rank if and only if one of the following conditions hold:

i) T is an irrational number;

ii) T is rational, but mT is not integer for any $m \in J$;

iii) T is rational, and mT is integer for some $m \in J$, but those integers also belong to I.

This is a direct consequence of Theorem 1.

B. The Equation $A\mathbf{f} = \mathbf{c}$

The results given so far fully characterize the rank of A in terms of the sets I and J. Having settled that, we may turn to the equation $A\mathbf{f} = \mathbf{c}$.

We want to address the problem in the three possible scenarios: overdetermined, exactly determined, and underdetermined. As a result, the equation may have zero, one or infinitely many solutions. To treat the three situations in a unified way, we replace the original equation $A\,\mathbf{f} = \mathbf{c}$ with the generalized problem

$$\min_{\mathbf{f}} \|A\,\mathbf{f} - \mathbf{c}\|^2. \tag{7}$$

The solution of this problem and the solution of $A\,\mathbf{f} = \mathbf{c}$ clearly coincide whenever the latter exists. However, the generalized problem always has a solution even if there is no $\mathbf{f}$ such that $A\,\mathbf{f} = \mathbf{c}$.

If there are multiple solutions, we agree to select the solution of least norm. This convention enables us to speak of the generalized solution of $A\,\mathbf{f} = \mathbf{c}$ without ambiguity: it means the minimum-norm least-squares solution to $Af = c$.

Theorem 2: Let $|J_3| \le \max(0, |J| - |I|)$. Then $A^\dagger$, the pseudo-inverse of A, is well defined, and the generalized solution $\mathbf{f}_0$ of $A\mathbf{f} = \mathbf{c}$ exists and is given by $\mathbf{f}_0 = A^\dagger\mathbf{c}$, with

$$A^\dagger = \begin{cases} A^{-1}, & \text{if } |I| = |J|, \\ A^T(AA^T)^{-1}, & \text{if } |I| > |J|, \\ (A^TA)^{-1}A^T, & \text{if } |I| < |J|. \end{cases}$$

Proof: By Theorem 1, A has full rank. For an arbitrary matrix A, $\operatorname{rank} A = \operatorname{rank} A^T = \operatorname{rank} A^TA = \operatorname{rank} AA^T$. Thus, $(A^TA)^{-1}$ and $(AA^T)^{-1}$ exist. The pseudo-inverse ([41], p. 207) can then be written in the given form. It is well-defined since the required inverse matrices all exist. ■

When the problem is exactly determined, i.e., $|I| = |J|$, the solution is $\mathbf{f} = A^{-1}\mathbf{c}$ and it determines a function that interpolates the given $\mathbf{c}$ at the given t_k.

In the underdetermined case, $|I| > |J|$, there are infinitely many solutions to $A\,\mathbf{f} = \mathbf{c}$. In this case, the theorem yields the solution with minimum norm $\|\mathbf{f}\|$, that is, $\mathbf{f} = A^T(AA^T)^{-1}\mathbf{c}$ (see e.g., [41]). Again, the solution determines a function that interpolates the given data.

When $|I| < |J|$ the problem is overdetermined. In this case, A is injective but not surjective so that $A\,\mathbf{f} = \mathbf{c}$ is solvable if and only if $\mathbf{c} \in R(A)$, a condition that is unlikely to hold in practice. The solution of the generalized problem (7) leads to $\mathbf{f} = (A^TA)^{-1}A^T\mathbf{c}$, i.e., the unique $\mathbf{f}$ that minimizes $\|A\mathbf{f} - \mathbf{c}\|^2$ (see e.g., [41]). This solution does not satisfy $A\,\mathbf{f} = \mathbf{c}$ in general. However, it is the natural choice because it is consistent with the exactly determined and underdetermined cases, in the sense that

$$A^{-1}\mathbf{c} = (A^TA)^{-1}A^T\mathbf{c} = A^T(AA^T)^{-1}\mathbf{c} \quad \text{if } |I| = |J|.$$

In practice, however, the computation of $A^\dagger$ is best done by means of the QR factorization or the singular value decomposition, without explicitly finding matrices such as $(AA^T)^{-1}$ or $(A^TA)^{-1}$ [38], [39], which have a quadratically worst condition number.

C. Energy Minimization

The choice given in Theorem 2 for $|I| > |J|$ singles out one possibility among infinitely many. We will see now that it does this by minimizing a certain quantity, and that this quantity is related to the energy or squared L^2 norm of the signal.

Observe that for any signal f of the form (1),

$$\begin{aligned}\|f\|^2 = \|\hat{f}\|^2 &= \int_{-1/2}^{1/2} \left| \sum_{k\in I} f(k) e^{-i2\pi k\omega} \right|^2 d\omega \\ &= \sum_{k\in I} |f(k)|^2 = \|\mathbf{f}\|^2\end{aligned}$$

where $\mathbf{f} = \{f(k)\}_{k\in I}$. Therefore among all $\mathbf{f}$ that satisfy the constraint $A\,\mathbf{f} = \mathbf{c}$, the choice $\mathbf{f} = A^T(AA^T)^{-1}\mathbf{c}$ gives the signal f with minimum energy

$$\begin{aligned}E = \mathbf{f}^T\mathbf{f} &= \mathbf{c}^T(AA^T)^{-1}AA^T(AA^T)^{-1}\mathbf{c} \\ &= \mathbf{c}^T(AA^T)^{-1}\mathbf{c}.\end{aligned}$$

Assume now that $I = \mathbb{Z}$, in which case the set of signals of the form (1) coincides with the space $PW_{1/2}$. Let J contain only numbers of the form $t_m = mT$. Then,

$$\begin{aligned}S := AA^T &= \left[\sum_{k\in\mathbb{Z}} \operatorname{sinc}(mT - k)\operatorname{sinc}(nT - k) \right]_{m\in J, n\in J} \\ &= [\operatorname{sinc}((m-n)T)]_{m\in J, n\in J}.\end{aligned}$$

To see this, apply the sampling theorem to the function $\operatorname{sinc}(\cdot - mT) \in PW_{1/2}$. Then

$$\operatorname{sinc}(t - mT) = \sum_{k\in\mathbb{Z}} \operatorname{sinc}(k - mT)\operatorname{sinc}(t - k)$$

and the formula follows by setting $t = nT$. Therefore

$$\begin{aligned}\mathbf{f} &= [\operatorname{sinc}(mT - n)]^T_{m\in J, n\in I} S^{-1}\mathbf{c} \\ &= [\operatorname{sinc}(m - nT)]_{m\in I, n\in J} S^{-1}\mathbf{c}.\end{aligned} \tag{8}$$

Recall that among all signals in $PW_{1/2}$, the signal of minimum energy which satisfy the constraints (2) is given as (see Theorem 2 in [4])

$$\tilde{f}(t) = \sum_{k\in J} x_k \operatorname{sinc}(t - kT)$$

where $\mathbf{x} = S^{-1}\mathbf{c}$. Note that setting $t = m \in I$, gives

$$[\tilde{f}(m)]_{m\in I} = [\operatorname{sinc}(m - nT)]_{m\in I, n\in J}\,\mathbf{x}. \tag{9}$$

From (8) and (9), we have $\mathbf{f} = [\tilde{f}(m)]_{m\in I}$ so that $f = \tilde{f}$ by the sampling theorem. Note also that the energy formulas for f and $\tilde{f}$ coincide.

The reason for this interesting result is that when $I = \mathbb{Z}$, the set of signals of the form (1) is the entire space $PW_{1/2}$, according to the sampling theorem. This shows the close connection between two different approaches to the construction of superoscillations, one that relies on the variational approach and the explicit minimization of the energy, and the present one, based on the direct specification of values at different scales.

D. Alternative View of Energy Minimization

The bandlimited signal of minimum energy that satisfies the constraints

$$f(t_k) = c_k, \quad 1 \le k \le N,$$

can be found using the variational approach (cf. Theorem 2 in [4]). The method can also be used when there are other types of constraints (for example, on the derivative [42]). The solution is obtained by combining the Euler-Lagrange equation with the appropriate constraints and Lagrange multipliers. Here we derive a method that does not use Lagrange multipliers. By a general version of the sampling theorem, any $f \in PW_{\mu/2}$ can be represented as

$$f(t) = \sum_{k\in\mathbb{Z}} f(\frac{k}{\mu})\,\mathrm{sinc}(\mu t - k). \tag{10}$$

Then $\|f\|^2 = \|\hat{f}\|^2 = \int_{-\mu/2}^{\mu/2} |\frac{1}{\mu}\sum_{k\in I} f(\frac{k}{\mu}) e^{-i2\pi\frac{k}{\mu}\omega}|^2 d\omega = \frac{1}{\mu}\sum_{k\in I} |f(\frac{k}{\mu})|^2 = \frac{1}{\mu}\|\mathbf{f}\|^2$ where $\mathbf{f} = \{f(\frac{k}{\mu})\}_{k\in\mathbb{Z}}$. Therefore the problem can be rewritten in terms of $\mathbf{f}$ as

$$\min \|\mathbf{f}\| \text{subject to } A\,\mathbf{f} = \mathbf{c}$$

where $A = [\mathrm{sinc}(\mu t_j - k)]_{j\in J, k\in I}$ and $\mathbf{c} = \{c_k\}_{1\le k\le N}$. The solution is $\mathbf{f} = A^T(AA^T)^{-1}\mathbf{c}$ with energy $E = \frac{1}{\mu}\mathbf{f}^T\mathbf{f} = \frac{1}{\mu}\mathbf{c}^T(AA^T)^{-1}\mathbf{c}$. By arguments similar to those used above, we have $S := AA^T = [\sum_{k\in\mathbb{Z}} \mathrm{sinc}(\mu t_m - k)\,\mathrm{sinc}(\mu t_n - k)]_{1\le m\le N, 1\le n\le N} = [\mathrm{sinc}(\mu(t_m - t_n))]_{1\le m\le N, 1\le n\le N}$. Then $\mathbf{f} = A^T\mathbf{x}$ where $\mathbf{x} = S^{-1}\mathbf{c}$ and this is the same as

$$f\left(\frac{k}{\mu}\right) = \sum_{i=1}^{N} x_i\,\mathrm{sinc}\left(\mu\left(\frac{k}{\mu} - t_i\right)\right), \quad k \in \mathbb{Z}.$$

Since $\tilde{f}(t) := \sum_{i=1}^{N} x_i\,\mathrm{sinc}(\mu(t - t_i)) \in PW_{\mu/2}$ and due to the unique reconstruction property of the sampling expansion (10) from samples, we have $f(t) = \tilde{f}(t)$. Note that the Lagrange multipliers method yields the same solution $\tilde{f}(t)$.

E. A Note on a Previous Result

We now discuss the connection with the solution obtained in [3], where it is shown that if $I| = J| < \infty$, $\mathbf{f} = \{f(k)\}_{k\in I}$ satisfies (1) and (2) if

$$\tilde{A}\,\mathbf{f} = \mathbf{g} \tag{11}$$

where $\tilde{A} = [T\sum_{k\in J}\mathrm{sinc}(kT - m)\,\mathrm{sinc}(kT - n)]_{m\in I, n\in I}$, $\mathbf{g} = \{g_n\}_{n\in I}$, and $g_n = T\sum_{k\in J} f(kT)\,\mathrm{sinc}(n - kT)$. It is easy to check that

$$\tilde{A} = TA^TA \quad \text{and} \quad \mathbf{g} = TA^T\mathbf{c}$$

where $A = [\mathrm{sinc}(jT - k)]_{j\in J, k\in I}$, so (11) is equivalent to the normal equations

$$A^TA\,\mathbf{f} = A^T\mathbf{c}, \tag{12}$$

which can of course also be obtained by applying A^T to both sides of $A\,\mathbf{f} = \mathbf{c}$. Notice that $\mathbf{f}$ of (12) reduces to $A^{-1}\mathbf{c}$ if $|I| = |J|$, and to the least-squares solution of $A\,\mathbf{f} = \mathbf{c}$ if $|I| < |J|$.

It is clear that A^TA is (symmetric) positive semi-definite. For any $\mathbf{z}$ in $\mathbb{R}^{|I|}$ (or in $\ell^2(I)$ if $|I| = \infty$), $\mathbf{z}^TA^TA\mathbf{z} = 0$ if and only if $A\mathbf{z} = 0$, so the invertibility (or positive definiteness) of A^TA is equivalent to the injectivity of A. Therefore A^TA is invertible (or positive definite) if and only if $|I| \le |J|$ and A has full rank.

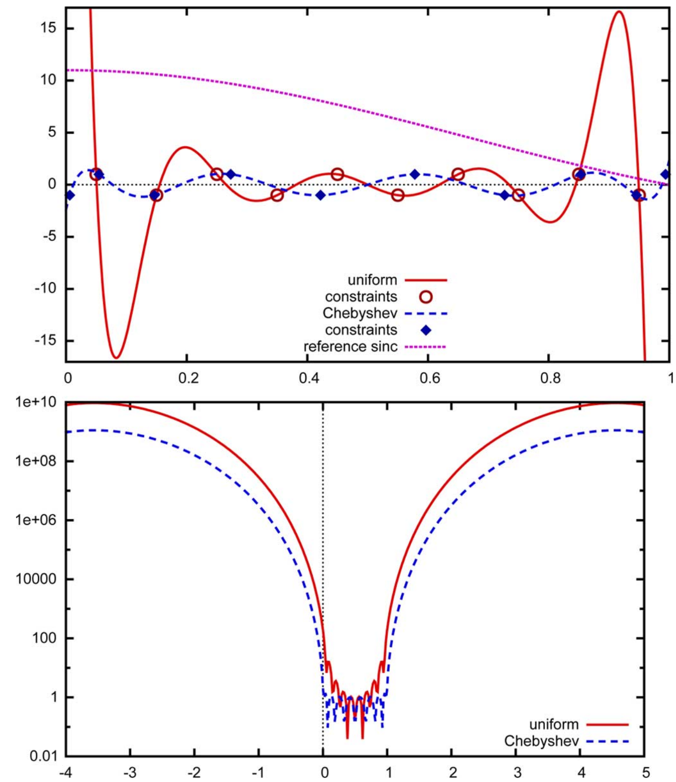

Fig. 2. Superoscillations at about 10 times the Nyquist rate. Two examples are shown: one is based on an uniform 10-point grid with spacing $1/10$, the other on a 10-point Chebyshev grid in $[0, 1]$. Top: the superoscillating region. Bottom: absolute value of the signal outside the superoscillating region.

It is claimed in [3] that $\tilde{A} = TA^TA$ is positive definite, but the strict positivity is not established. The full treatment of the matter is given by Theorem 1.

F. Examples

The following examples complement the results already given and raise some additional points.

Example 1: This example demonstrates the impact that the grid defined by J can have. We selected $I = \{-4, -3, \ldots, 4, 5\}$, thus $|I| = 10$. The two sets used for J also had cardinal 10, but different distributions. In one case the points were uniformly spaced in the interval $[0, 1]$, and in the other we considered the Chebyshev nodes on the same interval:

$$t_k = \frac{1 + \cos\frac{\pi(2k+1)}{2n}}{2}, \quad k = 0, 1, 2, \ldots, n-1.$$

Fig. 2 shows the differences between the two cases. The prescribed signal values were ± 1. The superoscillations built with the help of the uniform grid vary in amplitude, with peaks that attain values as large as 15 times the magnitude of the prescribed points. This behavior is not present in the superoscillations obtained with the Chebyshev nodes. The maximum amplitude outside the superoscillating segment is also smaller in the Chebyshev case, by one order of magnitude. The associated energy cost is smaller by a factor of over 70.

Example 2: This example shows the effect of the redundancy in the norm of the solution. By redundancy we mean

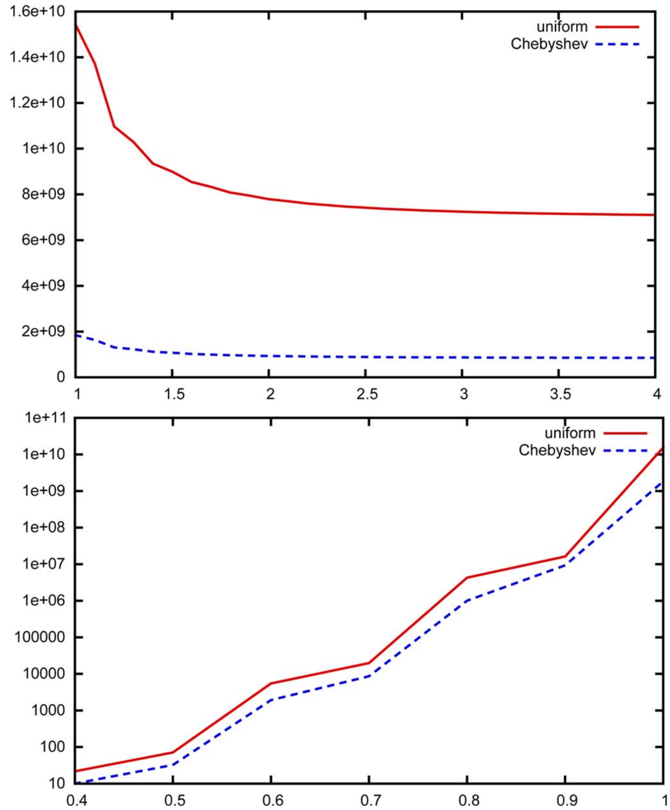

Fig. 3. Norm of the solution versus redundancy $r = |I|/|J|$, for superoscillations at 10 times the Nyquist rate, for 10-point uniformly spaced nodes and 10-point Chebyshev nodes. Top: increasingly underdetermined problems. The norm decreases rapidly as the redundancy increases, but there are no significant improvements as r increases past 4 or 5. Bottom: in the overdetermined case there are more prescribed conditions than adjustable signal parameters. The norm of the solution decreases very rapidly as the number of signal parameters decreases, but the best-effort solution cannot satisfy all the constraints.

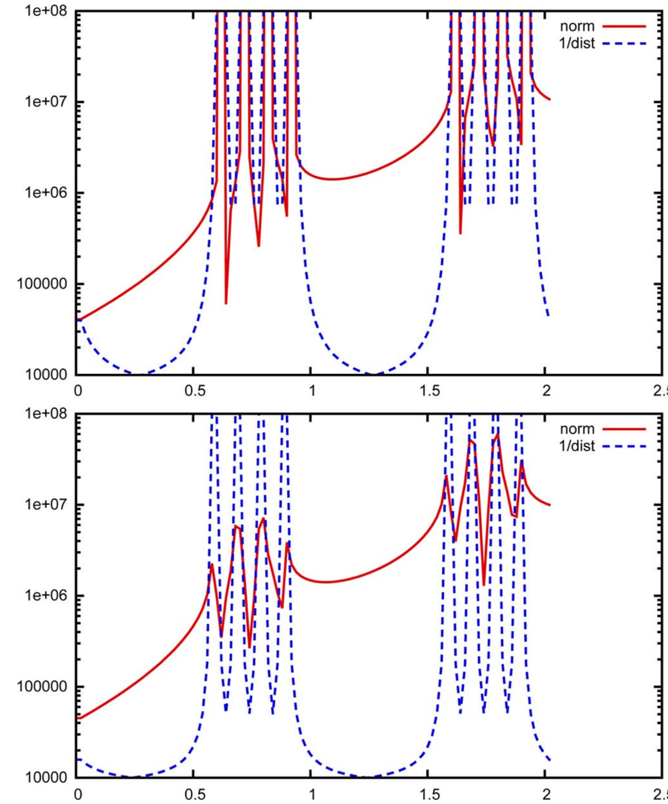

Fig. 4. Norm of the solution as a function of the distance between I and J, for superoscillations at 10 times the Nyquist rate. Top: as the uniformly spaced nodes in J are translated away from the fixed grid I they assume integer values, violating the necessary and sufficient condition for A to have full rank. The reciprocal of the smallest distance between J and the integers is shown for comparison. Bottom: J is moved away from I by amounts that are not a multiple of the grid spacing. Hence, no element of J ever assumes an integer value and the norm of the solution behaves more regularly.

$r = |I|/|J|$. We have $r < 1$ for overdetermined problems, in which there are fewer signal parameters than prescribed signal values, and $r > 1$ for underdetermined problems, which correspond to the opposite situation. The case $r = 1$ corresponds to a square matrix. For $r \geq 1$, Fig. 3 shows that the norm of the solution decreases as r increases, but reaches a stable value quickly. For $r \leq 1$, the norm decreases very rapidly with r, but the signal has not enough parameters to meet all the constraints.

Example 3: It shows how the norm of the solution evolves as J and I move away from each other, and the behavior as any of the elements of J approach an integer. The superoscillations created have a frequency equal to 10 times the Nyquist rate. The problem is underdetermined, with $|I|=3|J|$ (thus, redundancy 3). Initially, the smallest element of J is $t_1 = 0.1$ and the largest element of I is zero, so that the two sets are separated by 0.1. Then J is translated away from I, adding up to 2 to the separation.

The problem becomes numerically harder, since the superoscillations have to be created by cancellation of the sinc tails, and they get progressively weaker as J is moved away from I. Furthermore, when one of the translates of J intersects the integers, the necessary and sufficient condition for full rank given in Theorem 1 is violated. This is easily seen in the figure, which also shows the reciprocal of the smallest distance between J and the integers.

Randomly perturbing the elements of J or translating them by amounts that avoid the integers leads to more regular behavior, as shown in Fig. 4 (bottom).

Example 4: Example 3 suggests that as I and J move away from each other the norm of the solution tends to increase. Perhaps surprisingly, if the set I is not contiguous the results may vary.

Fig. 5 shows the behavior of the norm for several possible sets I. The set J was fixed for all experiments, to create superoscillations at 10 times the Nyquist rate in the interval $[0, 1]$. The sets I were generated by means of the transformation $aI - b$, with $a = 1, 2, 3, 4$ and $b = 0, 1, 2, 3, 4, 5$.

The largest integer in I is zero, so that all the integers in it lie to the left of the superoscillations in the unit interval. As b increases and the set is translated to the left and away from the superoscillations, the norm of the solution increases, as in Example 3. However, for example, $2I - b$ leads to smaller norms than $I - b$, for all values of $b \geq 2$, despite containing elements that are further from the superoscillations, by a factor of two.

Example 5: Recall from (1) and (3) that we are building superoscillations by constraining the coefficients $\mathbf{f}$ in

$$f(t) = \sum_{k \in I} f(k) \operatorname{sinc}(t - k)$$

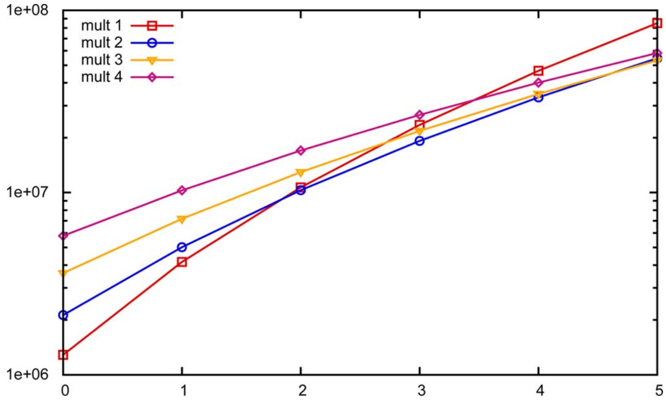

Fig. 5. Norm of signals containing superoscillations at 10 times the Nyquist rate, for sets of the form $aI - b$. Each curve corresponds to a different value of the multiplier a, with $a = 1, 2, 3, 4$. The offset b (horizontal axis) varies between 0 and 5. It shows that the set I that leads to a solution with the least norm is not necessarily contiguous, nor closest to the data J.

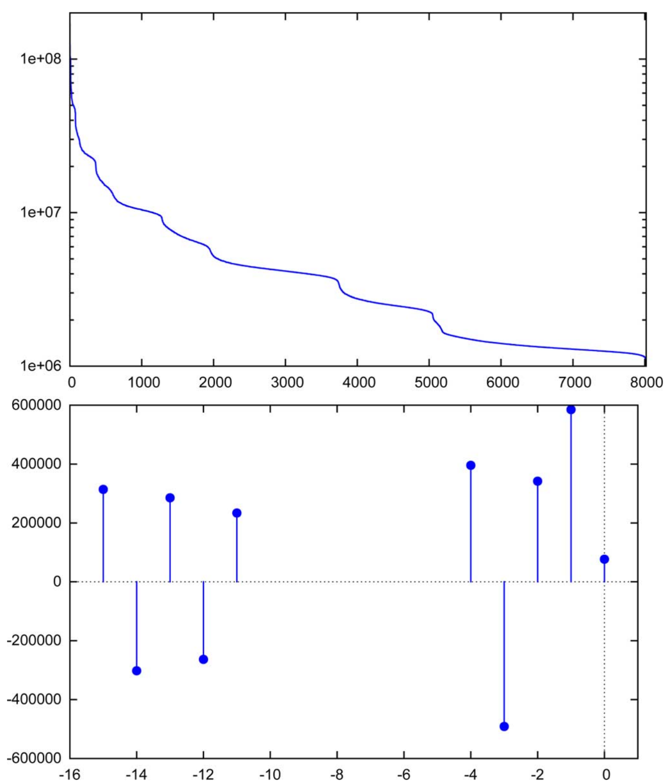

Fig. 6. The set J (belonging to $[1, 2]$) and the vector $\mathbf{c}$ were held fixed to create superoscillations at 10 times the Nyquist rate. Then we considered all possible sets I of cardinal 10 within the range $-15 : 0$. Top: the ordered norms of the solutions $A\mathbf{f} = \mathbf{c}$. Bottom: the solution that had the least norm, which does not coincide with the set I closest to J.

by means of the equations $A\,\mathbf{f} = \mathbf{c}$. We now fix the set J and the vector $\mathbf{c}$, and ask how I alone impacts the solution. We consider $|I| \geq |J|$ to make sure that the constraints are met. To complement the previous examples, we move I away from J. Assume that A has full rank at the start. By Theorem 1, A will remain of full rank. How will the changes in I impact the norm of the solution?

Fig. 6 shows the ordered set of norms for all possible sets I of a fixed cardinal within a fixed range.

The frequency of the superoscillations was 10 times greater than the Nyquist rate, the set J, located inside $[1, 2]$, was held fixed across all experiments, and the vector $\mathbf{c}$ contained alternating values ± 1. The set I was restricted to 10 elements and all possible sets of cardinal 10 in the range $-15 : 0$ were considered. The set I furthest from J is $I_1 = \{-15, -14, \ldots, -6\}$, and the closest is $I_2 = \{-9, -8, \ldots, 0\}$. We found that I_1 lead to the solution with the largest norm, as expected (cf. Example 3). However, as already hinted in Example 4, the contiguous set I_2 did not lead to the superoscillations of least norm. The solution of least norm is shown in Fig. 6 (bottom) and contains the integers $\{-15, -14, -13, -12, -11\}$ at the furthest end of the range. The set I_2 is within the 30% best and the solution of least norm has about 0.75 of its norm.

III. Conclusion

We have built superoscillations by considering a function of the form

$$f(t) = \sum_{k \in I} f(k)\,\mathrm{sinc}(t - k)$$

and directly imposing constraints on its values at a fine scale:

$$f(kT) = c_k, \quad k \in J.$$

We also considered the more general situation

$$f(t_k) = c_k, \quad k = 1, 2, \ldots, K,$$

in which the values are prescribed on irregularly spaced points.

The conditions constrain the behavior of f on a grid that is arbitrarily finer than the reference grid. Thus, they may force f to oscillate at a much higher frequency than its bandlimit $1/2$ Hz. Furthermore, the sets I and J can be contained in intervals very far apart, reducing the apparent influence of the constraints on the coefficients of f.

Nevertheless, we found that the constraints are in general compatible with the nature of the signals. We considered the matrix

$$A = [\mathrm{sinc}(t_j - k)]_{t_j \in J, k \in I},$$

which determines the constraint equation $A\,\mathbf{f} = \mathbf{c}$, and gave necessary and sufficient conditions for it to have full rank. We found that the number of elements in J that also belong to I plays the crucial role in the matter.

The conditions on the rank yield necessary and sufficient conditions for the nonsingularity of A when the equations are exactly determined. However, we also studied the underdetermined and overdetermined cases (including the case $|I| = \infty$) and gave generalized solutions to $Af = c$, which were given a unified treatment using the pseudo-inverse $A^\dagger$.

The problem of constructing superoscillations is in itself difficult from the numerical point of view, and the computation of the pseudo-inverse is an ill-posed problem when the matrix A does not have full rank. This is another reason why understanding the rank of A is important.

We derived expressions for the energy of the signal and related it with the results obtained using the variational approach followed in [4], [42].

We found that superoscillations can be made to occur at a large distance from the nonzero samples of the signal (cf. Examples 3 and 4). A reviewer pointed out that there are two ways in which this may occur. One that was recently explored [43] is related to pointer shifts corresponding to the weak measurement of an operator in Aharonov's scheme. The other occurs e.g., in sub-wavelength optical microscopy, when superoscillations are propagated in monochromatic waves. The superoscillations can be fitted so as to propagate without distortion [44], but the weaker requirement of distortion-free reproduction of superoscillations at a series of prescribed distances leads to a matrix inversion problem involving a nonuniform grid and a real symmetric matrix ([45], (3.7)).

The results obtained establish a connection between two seemingly different approaches to the construction of superoscillations: the variational approach, which seeks to minimize the energy of the interpolant in the space $PW_{1/2}$, and the direct approach discussed in the present paper, which is based on the mere specification of signal values at the finest scale. The variational approach is powerful and flexible (e.g., it can be used to construct superoscillations with controlled amplitude and derivative [42]). The direct approach discussed in the present paper is more elementary (there is no need to solve the Euler-Lagrange equation). However, it can be directly applied to any set I, finite or infinite, and yields insight on the energy required by the superoscillations as a function of the signal parameters and constraints, as well as their structure.

Acknowledgment

Dae Gwan Lee gratefully acknowledges the generous hospitality of IEETA, Universidade de Aveiro, during the period in which this paper was written. He also acknowledges the support provided by BK21 PLUS of the Department of Mathematical Sciences, KAIST.

References

[1] E. T. F. Rogers, J. Lindberg, T. Roy, S. Savo, J. E. Chad, M. R. Dennis, and N. I. Zheludev, "A super-oscillatory lens optical microscope for subwavelength imaging," *Nature Mater.*, vol. 11, pp. 432–435, 2012.

[2] P. J. S. G. Ferreira and A. J. Pinho, "The natural scale of signals: Pulse duration and superoscillations," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Florence, Italy, 2014, pp. 4209–4212.

[3] P. J. S. G. Ferreira, A. Kempf, and M. J. C. S. Reis, "Construction of Aharonov-Berry's superoscillations," *J. Phys. A, Math. Gen.*, vol. 40, pp. 5141–5147, 2007.

[4] P. J. S. G. Ferreira and A. Kempf, "Superoscillations: Faster than the Nyquist rate," *IEEE Trans. Signal Process.*, vol. 54, no. 10, pp. 3732–3740, Oct. 2006.

[5] Y. Aharonov, D. Z. Albert, and L. Vaidman, "How the result of a measurement of a component of the spin of a spin-1/2 particle can turn out to be 100," *Phys. Rev. Lett.*, vol. 60, no. 14, pp. 1351–1354, Apr. 1988.

[6] Y. Aharonov and D. Rohrlich, *Quantum Paradoxes: Quantum Theory for the Perplexed*. Weinheim, Germany: Wiley-VCH, 2005.

[7] F. E. Bond and C. R. Cahn, "On sampling the zeros of bandwidth limited signals," *IRE Trans. Inf. Theory*, vol. 4, no. 3, pp. 110–113, Sep. 1958.

[8] P. J. S. G. Ferreira, "Superoscillations," in *New Perspectives on Approximation and Sampling Theory: Festschrift in Honor of Paul Butzer's 85th Birthday*, G. Schmeisser and A. Zayed, Eds. New York, NY, USA: Springer, 2014.

[9] H. J. Landau, "Extrapolating a band-limited function from its samples taken in a finite interval," *IEEE Trans. Inf. Theory*, vol. 32, no. 4, pp. 464–470, Jul. 1986.

[10] L. Levi, "Fitting a bandlimited signal to given points," *IEEE Trans. Inf. Theory*, pp. 372–376, Jul. 1965.

[11] A. A. G. Requicha, "The zeros of entire functions: Theory and engineering applications," *Proc. IEEE*, vol. 68, no. 3, pp. 308–328, Mar. 1980.

[12] W. Qiao, "A simple model of Aharonov-Berry's superoscillations," *J. Phys. A, Math. Gen.*, vol. 29, pp. 2257–2258, 1996.

[13] M. Berry, "Faster than Fourier," in *Quantum Coherence and Reality; in Celebration of the 60th Birthday of Yakir Aharonov*, J. S. Anandan and J. L. Safko, Eds. Singapore: World Scientific, 1994, pp. 55–65.

[14] P. J. S. G. Ferreira and A. Kempf, "The energy expense of superoscillations," in *Proc. XI Eur. Signal Process. Conf. Signal Process. XI—Theories Appl. (EUSIPCO)*, Toulouse, France, Sep. 2002, vol. II, pp. 347–350.

[15] Y. Aharonov, F. Colombo, I. Sabadini, D. C. Struppa, and J. Tollaksen, "Some mathematical properties of superoscillations," *J. Phys. A, Math. Theory*, vol. 44, 2011, 365304, 16 pp.

[16] E. Katzav and M. Schwartz, "Yield-optimized superoscillations," *IEEE Trans. Signal Process.*, vol. 61, no. 12, pp. 3113–3118, 2013.

[17] N. Schwartz and M. Schwartz, "On a periodicity measure and superoscillations," 2014 [Online]. Available: http://arxiv.org/abs/1402.4092, arXiv:1402.4092 [physics.data-an]

[18] M. Schwartz and E. Perlsman, "Sensitivity of interpolated yield optimized superoscillations," 2014 [Online]. Available: http://arxiv.org/abs/1402.3089, arXiv:1402.3089 [physics.data-an]

[19] Y. Aharonov, J. Anandan, S. Popescu, and L. Vaidman, "Superpositions of time evolutions of a quantum system and a quantum time-translation machine," *Phys. Rev. Lett.*, vol. 64, no. 25, pp. 2965–2968, Jun. 1990.

[20] A. Kempf, "Black holes, bandwidths and Beethoven," *J. Math. Phys.*, vol. 41, no. 4, pp. 2360–2374, Apr. 2000.

[21] M. V. Berry, "Evanescent and real waves in quantum billiards and Gaussian beams," *J. Phys. A, Math. Gen.*, vol. 27, no. 11, pp. L391–L398, 1994.

[22] A. Kempf and P. J. S. G. Ferreira, "Unusual properties of superoscillating particles," *J. Phys. A, Math. Gen.*, vol. 37, pp. 12 067–12 076, 2004.

[23] M. V. Berry and S. Popescu, "Evolution of quantum superoscillations and optical superresolution without evanescent waves," *J. Phys. A, Math. Gen.*, vol. 39, pp. 6965–6977, 2006.

[24] F. M. Huang, Y. Chen, F. J. Garcia de Abajo, and N. I. Zheludev, "Optical super-resolution through super-oscillations," *J. Opt. A, Pure Appl. Opt.*, vol. 9, pp. S285–S288, 2007.

[25] N. I. Zheludev, "What diffraction limit?," *Nature Mater.*, vol. 7, pp. 420–422, Jun. 2008.

[26] F. M. Huang and N. I. Zheludev, "Super-resolution without evanescent waves," *Nano Lett.*, vol. 9, no. 3, pp. 1249–1254, 2009.

[27] A. M. H. Wong and G. V. Eleftheriades, "Adaptation of Schelkunoff's superdirective antenna theory for the realization of superoscilattory antenna arrays," *IEEE Antennas Wireless Propag. Lett.*, vol. 9, pp. 315–318, 2010.

[28] A. M. H. Wong and G. V. Eleftheriades, "Superoscillatory antenna arrays for sub-diffraction focusing at the multi-wavelength range in a waveguide environment," in *Proc. IEEE Antennas Propag. Soc. Int. Workshop (APSURSI)*, Toronto, ON, Canada, Jul. 2010, pp. 1–4.

[29] A. M. H. Wong and G. V. Eleftheriades, "Sub-wavelength focusing at the multi-wavelength range using superoscillations: An experimental demonstration," *IEEE Trans. Antennas Propag.*, vol. 59, no. 12, pp. 4766–4776, Dec. 2011.

[30] A. M. H. Wong and G. V. Eleftheriades, "Temporal pulse compression beyond the Fourier transform limit," *IEEE Trans. Microw. Theory Tech.*, vol. 59, no. 9, pp. 2173–2179, Sep. 2011.

[31] M. R. Dennis and J. Lindberg, "Natural superoscillation of random functions in one and more dimensions," in *Proc. SPIE—Int. Soc. Opt. Eng.*, 2009, vol. 7394, p. 73940A, 9 pp.

[32] A. Kempf and R. Martin, "Information theory, spectral geometry, and quantum gravity," *Phys. Rev. Lett.*, vol. 100, Jan. 2008, 021304.

[33] M. V. Berry and M. R. Dennis, "Natural superoscillations in monochromatic waves in d dimensions," *J. Phys. A, Math. Theory*, vol. 42, 2009, 022003, 8 pp.

[34] M. V. Berry, "A note on superoscillations associated with Bessel beams," *J. Opt.*, vol. 15, 2013, 044006, 5 pp.

[35] D. Sokolovski and R. S. Mayato, "Superluminal transmission via entanglement, superoscillations, and quasi-Dirac distributions," *Phys. Rev. A*, vol. 81, 2010, 022105.

[36] V. Eisler and I. Peschel, "Free-fermion entanglement and spheroidal functions," *J. Statist. Mechan., Theory Experiment*, vol. 2013, no. 4, Apr. 2013, P04028.

[37] M. V. Berry, "Quantum backflow, negative kinetic energy, and optical retro-propagation," *J. Phys. A, Math. Theory*, vol. 43, 2010, 415302, 15 pp.

[38] S. L. Campbell and C. D. Meyer, Jr., *Generalized Inverses of Linear Transformations*. New York, NY, USA: Dover, 1991.

[39] A. Ben-Israel and T. N. E. Greville, *Generalized Inverses: Theory and Applications*. New York, NY, USA: Springer, 2003.

[40] G. Pólya and G. Szegö, *Problems and Theorems in Analysis II.*. Berlin, Germany: Springer, 1976, vol. 2.

[41] E. K. P. Chong and S. H. Żak, *An Introduction to Optimization*. Hoboken, NJ, USA: Wiley-Interscience, 2008.

[42] D. G. Lee and P. J. S. G. Ferreira, "Superoscillations of prescribed amplitude and derivative," *IEEE Trans. Signal Process.*, 2014 [Online]. Available: http://www.ua.pt/~pjf/preprints/deriv-superosc.pdf, submitted for publication

[43] M. V. Berry and P. Shukla, "Pointer supershifts and superoscillations in weak measurements," *J. Phys. A, Math. Theory*, vol. 45, 2012, 015301, 14 pp.

[44] K. G. Makris and D. Psaltis, "Superoscillatory diffraction-free beams," *Opt. Lett.*, vol. 36, no. 22, pp. 4335–4337, Nov. 2011.

[45] M. V. Berry, "Exact nonparaxial transmission of subwavelength detail using superoscillations," *J. Phys. A, Math. Theory*, vol. 46, 2013, 205203, 15 pp.

Dae Gwan Lee was born in Seoul, S. Korea. He received the B.S. degree in mathematical sciences in 2008 from KAIST, S. Korea. Currently, he is a Ph.D. candidate in Mathematical Sciences, KAIST, S. Korea. His research interests include signal processing, sampling theory, and Fourier and wavelet analysis.

Paulo Jorge S. G. Ferreira is Professor in the Departamento de Electrónica, Telecomunicações e Informática at Universidade de Aveiro, Portugal. He is an Associate Editor of the IEEE SIGNAL PROCESSING LETTERS and other journals and a past Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING. His research interests include bioinformatics, sensor networks and signal processing, particularly sampling, approximation and superoscillations.