

Superoscillations: Faster Than the Nyquist Rate

Paulo J. S. G. Ferreira, *Member, IEEE*, and Achim Kempf

Abstract—It is commonly assumed that a signal bandlimited to $\mu/2$ Hz cannot oscillate at frequencies higher than μ Hz. In fact, however, for any fixed bandwidth, there exist finite energy signals that oscillate arbitrarily fast over arbitrarily long time intervals. These localized fast transients, called superoscillations, can only occur in signals that possess amplitudes of widely different scales. This paper investigates the required dynamical range and energy (squared L^2 norm) as a function of the superoscillation's frequency, number, and maximum derivative. It briefly discusses some of the implications of superoscillating signals, in reference to information theory and time-frequency analysis, for example. It also shows, among other things, that the required energy grows exponentially with the number of superoscillations, and polynomially with the reciprocal of the bandwidth or the reciprocal of the superoscillations' period.

Index Terms—Bandlimited signals, information rates, quantum theory, signal sampling, superoscillations, time-frequency analysis.

I. INTRODUCTION

It is commonly assumed that “a signal bandlimited to $\mu/2$ Hz cannot oscillate at a frequency higher than μ Hz” or that “a signal bandlimited to $\mu/2$ Hz cannot change substantially in an interval shorter than $1/\mu$ seconds.”

Strictly speaking, these statements are false. Locally, bandlimited signals may vary in an arbitrarily rapid way. For example, there exist finite-energy signals bandlimited to (say) 1 Hz, the values of which change from $f(0) = -1$ to $f(\delta) = 1$, no matter how small δ is.

Even more surprisingly, it is theoretically possible to construct finite-energy signals bandlimited to 1 Hz (or to any other fixed and arbitrarily small frequency) which interpolate N samples of a sinusoid or a square wave or any other waveform of arbitrarily high frequency. More precisely, given arbitrarily chosen times $\{t_i\}_{i=1}^N$ and amplitudes $\{a_i\}_{i=1}^N$, it is possible to find f such that $f(t_i) = a_i$ for $i = 1, 2, \dots, N$, where N is arbitrarily large.

The purpose of this paper is to discuss this phenomenon, which has been called “superoscillation” [1]–[3], and in particular to give bounds for the energy (squared L^2 norm) of superoscillating signals. We generalize and complement the preliminary results reported in [4] and relate energy, maximum deriva-

Manuscript received October 21, 2004; revised September 17, 2005. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Carlos H. Muravchik. The work of P. J. S. G. Ferreira was supported in part by FCT. The work of A. Kempf was supported by PREA, CFI, OIT, NSERC, and the Canada Research Chairs Program.

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

A. Kempf is with the Department of Applied Mathematics, University of Waterloo, Waterloo, ON N2L 3G1, Canada (e-mail: akempf@uwaterloo.ca).

Digital Object Identifier 10.1109/TSP.2006.877642

Minimum energy solutions

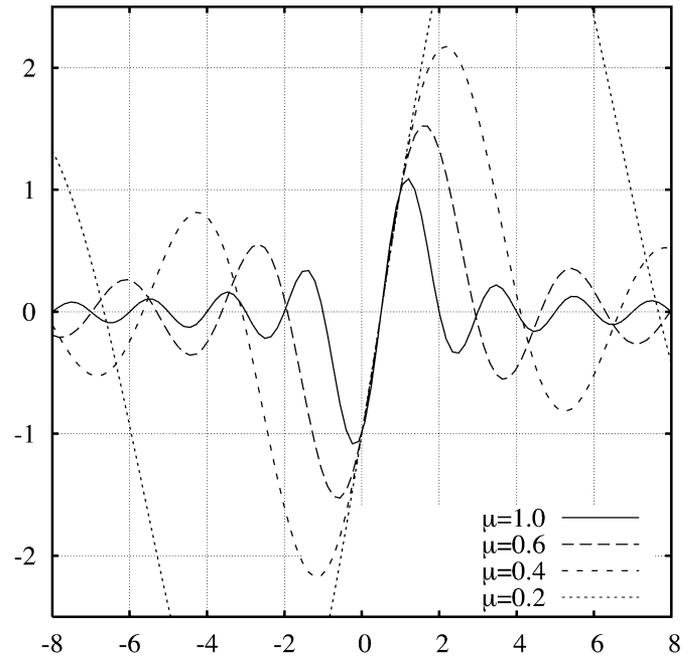


Fig. 1. The signal of minimum energy bandlimited to $\mu/2$ Hz that satisfies $f(0) = -1$ and $f(1) = 1$, for a few values of μ .

ive, and bandwidth, examining the role that the prolate matrix plays in the matter.

We find that a finite-energy signal f , bandlimited to $\mu/2$ Hz, may change its value from -1 to $+1$ in an interval of length δ provided that its energy satisfies

$$E = \|f\|_2^2 \geq \frac{2}{\mu - \mu \operatorname{sinc}(\mu\delta)}. \quad (1)$$

The constants are the best possible, and the signal for which the bounds are met is described (see Fig. 1). When f is required to take several amplitudes alternating in sign, say, $f(k\delta) = a_k$, for $k = 0, 1, \dots, N-1$, we find that for fixed N the energy required increases polynomially¹ with $\mu\delta$

$$E = \mathcal{O} \left[\frac{1}{(\mu\delta)^{2N-1}} \right]. \quad (2)$$

On the other hand, for fixed μ and δ , the energy increases exponentially with N the number of superoscillations

$$E = \mathcal{O} \left[N^{-1/2} \left(\frac{4}{\pi\mu\delta} \right)^{2N-1} \right]. \quad (3)$$

¹Recall that the notation $f(x) = \mathcal{O}(g(x))$ means that there exists a positive constant C such that $|f(x)| \leq C|g(x)|$ as x tends to some limit (usually $x \rightarrow \infty$).

Our analysis shows that there is exactly one data set a_k , leading to the constraints $f(k\delta) = a_k$, for which the equality sign is valid. Alternating sign data such as $a_k = (-1)^k$ imply energies very close to this upper bound.

At first glance these findings appear to contradict principles that are generally accepted. The following remarks are intended to clarify some of the issues, and to motivate the mathematical analysis that follows.

A. Conditioning

Our first remark has to do with the numerical conditioning of the problem that springs from the problem $f(k\delta) = a_k$. We will see that it leads to a set of linear equations with a certain matrix S , and thus the determination of the superoscillating signal is not complete in practice until an equation of the form $Sx = a$ is solved. The size of the matrix is determined by the number N of prescribed amplitudes $\{a_i\}_{i=1}^N$, and the matrix is nonsingular. How difficult can the problem $Sx = a$ be?

This question suggests the study of the condition number of S , denoted by $\kappa(S)$. Recall that the condition number of a nonsingular matrix S , with respect to the matrix norm $\|\cdot\|$, is $\kappa(S) := \|S\| \|S^{-1}\|$. Its meaning, in connection with the linear problem $Sx = a$, is the following. If the data a are replaced with $a + \Delta a$, the solution x will change to $x + \Delta x$. For a given matrix norm, the condition number sets a bound on the relative error in x , with respect to that norm:

$$\frac{\|\Delta x\|}{\|x\|} \leq \kappa(S) \frac{\|\Delta a\|}{\|a\|}.$$

Our results lead to the following approximation for the condition number in the spectral norm:

$$\kappa[S(\mu, \delta)] \sim \min(\mu\delta N, 1) N^{-1/2} \left(\frac{4}{\pi\mu\delta} \right)^{2N-1} \quad (4)$$

which shows how difficult numerically the problem $Sx = a$ can be.

B. Prolate Spheroidals

The signal f of least energy that satisfies $f(t_i) = a_i$ for $i = 1, 2, \dots, N$, is determined by linear equations $Sx = a$, and S turns out to be a multiple of the prolate matrix. We will explore the connection between the coefficients of a superoscillating signal and the size of the eigenvalues of the prolate matrix, which naturally appears in the study of the *discrete* prolate spheroidal wave functions (PSWFs) [5].

One of the reviewers pointed out the connection with the *continuous-time* PSWF [6], denoted by ψ_n . These functions are bandlimited, orthonormal on the real line, orthogonal on $(-1, 1)$, and odd or even with n . Furthermore, ψ_n has exactly n zeros in $(-1, 1)$ [7]. In other words, the PSWF ψ_n , for large n , contains superoscillations.

Each ψ_n is an eigenfunction of a certain operator, associated with the eigenvalue λ_n . The energy concentration of ψ_n inside $(-1, 1)$ is measured by λ_n . Since $\lambda_n \rightarrow 0$ as $n \rightarrow \infty$, the energy of the ψ_n inside $(-1, 1)$ tends to zero as $n \rightarrow \infty$. In other words, as the number of superoscillations inside $(-1, 1)$ increases, the energy (and amplitude) of the function inside $(-1, 1)$ decreases.

It is possible to rescale ψ_n so that the fraction of energy in the superoscillating segment $(-1, 1)$ remains constant with n . Naturally, this means that the total energy (square L^2 norm) of the rescaled ψ_n must increase with n .

Although the PSWFs provide ready-to-use examples of superoscillating functions, the superoscillations are not under control. All that is known is that the number of oscillations inside $(-1, 1)$ increases with n . By contrast, in this paper, we emphasize interpolation problems such as $f(t_k) = a_k$ to keep the behavior of the function under control. We study how the energy increases with the number of constraints $f(t_k) = a_k$ and with the bandlimit $\mu/2$ and separation $\delta = \Delta t_k$.

C. Information Theory

At first sight, superoscillations seem to open the possibility of encoding an arbitrary amount of information into an arbitrarily short segment of a low-bandwidth signal. This signal would then be able to pass through any channel of correspondingly narrow bandwidth.

But, on one hand, we show that the superoscillations' amplitude decreases exponentially with the length of the superoscillating stretch. On the other hand, every channel has some level of noise. Shannon showed [8], [9] that signals that pass through a Gaussian channel of bandwidth B can deliver information on average at most at a rate $B \log_2(1 + S/N)$, where S/N is the signal-to-noise ratio. Thus, Shannon's theorem allows superoscillatory information compression to an arbitrary extent—but demands that the signal power grows exponentially with the amount of compressed information, i.e., with the length of the part of the message that is superoscillatory. Interestingly, this corresponds to our finding that the norm of maximally superoscillatory wave functions grows exponentially with the number of prescribed superoscillations.

D. Interpolation and the Nyquist Rate

The possibility of solving a bandlimited interpolation problem depends on the density of the interpolation knots. Beurling, Landau, and others have shown that the density of a set of points $\Lambda = \{t_k\}$ determines the properties of the functions $\mathcal{E}(\Lambda) = \{e^{j2\pi t_k x}\}$, and that these properties are closely related to interpolation and sampling.

In essence, the results state that the interpolation problem $f(t_k) = a_k$, $k \in \mathbf{Z}$, can be solved in the Paley–Wiener space only when the t_k are separated and their density is small enough (below the Nyquist rate). Precise definitions of “separation” and “density” will be given, but roughly speaking the t_k are separated if they do not get arbitrarily close, and their density is related to the number of t_k that fall in an interval of length r , divided by r . Our results in no way contradict these because the interpolation problem considered is finite. The separation condition is not violated, since a set of N distinct numbers t_k is separated. It is true that the local density of the t_k can reach arbitrarily high values, but on finite intervals only. The average density will be zero.

E. Quantum Physics

This remark concerns superoscillations in quantum physics, where superoscillations and associated phenomena are also

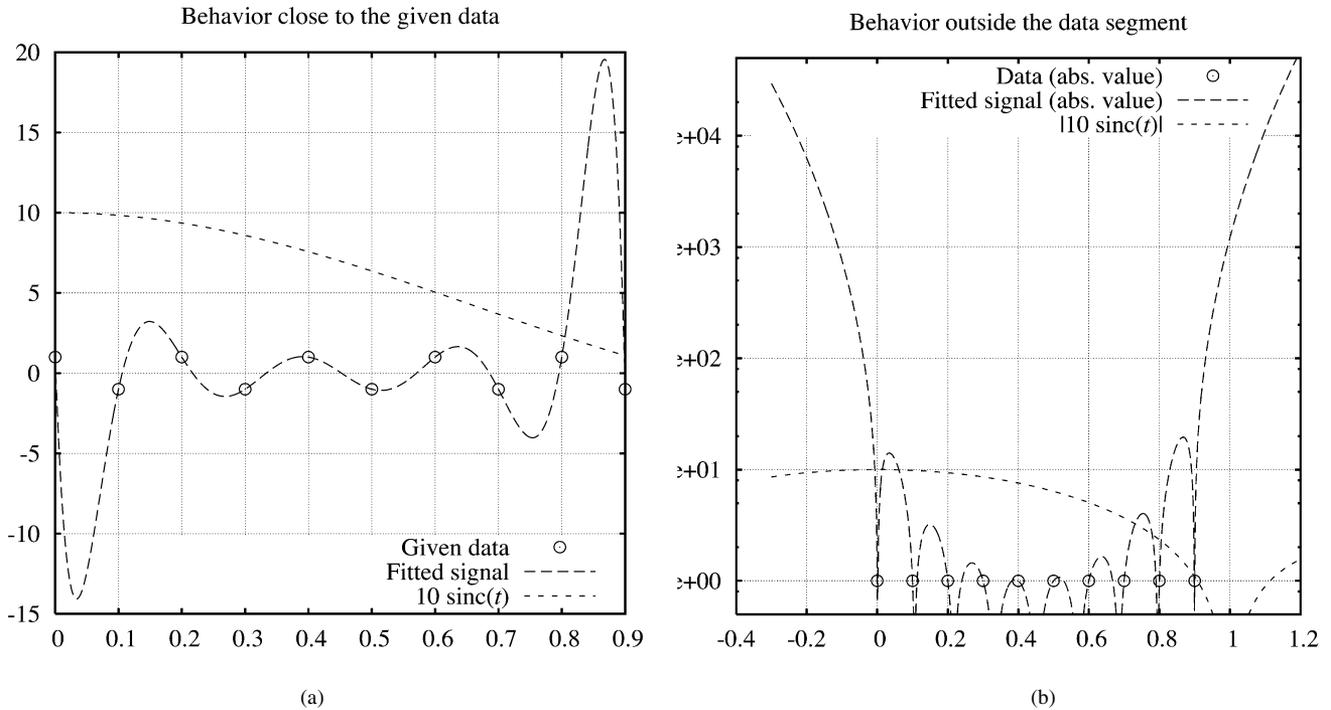


Fig. 2. (a) Example of a superoscillating signal, with $N = 10$ and $\delta = 0.1$, compared with a sinc signal of the same bandwidth. (b) Absolute value of the signal, in a logarithmic scale. Compare the behavior inside and outside the segment that contains the fitted data.

of interest; see [3]. Indeed, the wave functions of elementary particles are bandlimited whenever the particles' momentum is bounded. While superoscillating wave functions of course do not contradict the Heisenberg uncertainty relations, the properties of superoscillating wave functions do illuminate some unexpected aspects of the uncertainty relations, as was recently shown in [10].

F. Time-Frequency Analysis

Superoscillations have implications on time-frequency analysis. Consider a sampled segment of length T of a signal, bandlimited to $\mu/2$, and containing N superoscillations of period δ . The underlying signal will resemble the one depicted in Fig. 2. It is clear that the samples would certainly mislead the observer as to the bandwidth of the signal. In fact, it is impossible to infer the bandwidth of a finite energy signal f from a sampled segment of length T , even if T is large. The reason is that there exists signals of arbitrarily small bandwidth that oscillate throughout an interval of length T , with arbitrarily small period δ .

However, if the observer is able to increase T at will, the true character of f will become apparent. The amplitude of a superoscillating signal will increase rapidly outside the superoscillating segment. In a sense, superoscillations are created by cancellation, and the amplitude inside the superoscillating segment is kept small due to that cancellation.

In the next sections, we concentrate on the energy (or L^2 norm) required to achieve a certain slope or a certain number N of oscillations, separated by δ seconds, with signals of maximum frequency $\mu/2$ Hz. The asymptotic expressions for the required energy show an exponential increase of the energy with

N and a polynomial increase with $1/\mu$ or $1/\delta$, and the expressions and results mentioned in this Introduction will be mathematically justified.

II. ENERGY AND MAXIMUM SLOPE

Fast varying signals are commonly identified with wide-band signals, but this is not necessarily the case. In fact, a (finite) segment of a signal of small bandwidth can exhibit a behavior that is often associated with wide-band signals. We will try to clarify the extent to which this is true in the following sections.

The maximum value of the derivative of the signal $a \sin(bt)$ is ab . If b is fixed, the maximum derivative can be large if and only if a is large. A sinusoidal signal such as $a \sin(bt)$ does not have finite energy, but the argument suggests that fast varying finite energy signals might occur as a result of either a very large bandwidth or a combination of small bandwidth and very high energy.

To make this precise, we start by giving a lower bound for the energy required to attain a given slope. Before going on, we explain the notation: the L^2 norm of f is

$$\|f\|_2 := \left(\int_{-\infty}^{\infty} |f(t)|^2 dt \right)^{1/2}.$$

If f is bandlimited to $\mu/2$, then

$$\|\hat{f}\|_2 = \left(\int_{-\mu/2}^{\mu/2} |\hat{f}(\xi)|^2 d\xi \right)^{1/2} = \|f\|_2.$$

We need a form of Bernstein’s inequality. To obtain it, consider $f \in L^2$, bandlimited to $\mu/2$, where μ is any fixed positive real, that is

$$f(t) = \int_{-\mu/2}^{+\mu/2} \widehat{f}(\omega) e^{i2\pi\omega t} d\omega. \tag{5}$$

It follows that

$$\|f'\|_2^2 = \|2\pi\omega \widehat{f}(\omega)\|_2^2 \leq (\pi\mu)^2 \|\widehat{f}\|_2^2$$

and hence

$$\|f'\|_2 \leq \pi\mu \|f\|_2. \tag{6}$$

To relate the energy of a signal and the maximum value of its derivative, we proceed as follows. Schwarz’s inequality applied to (5) leads to

$$|f(t)|^2 \leq \mu \int_{-\mu/2}^{+\mu/2} |\widehat{f}(\omega)|^2 d\omega = \mu \|\widehat{f}\|_2 = \mu \|f\|_2^2.$$

This inequality also holds for the derivative f' , which is also bandlimited to the same value, and so

$$|f'(t)|^2 \leq \mu \|f'\|_2^2.$$

Combination of this with Bernstein’s inequality (6) leads to

$$|f'(t)|^2 \leq \pi^2 \mu^3 \|f\|_2^2 \tag{7}$$

and so

$$\|f\|_2 \geq \frac{\mu^{-3/2}}{\pi} |f'(t)|. \tag{8}$$

These expressions relate the bandwidth of the signal, its maximum derivative, and its L^2 norm. Equation (7) sets an upper bound to the magnitude of the signal derivative, controlled by the signal bandwidth and energy. The reverse inequality (8) is more interesting for our purposes. It does not forbid sharp transitions in signals of small bandwidth. It merely states that high derivatives can be achieved only at the expense of energy.

Consider the problem of fitting a bandlimited finite-energy signal $f(t)$ to points

$$(t_0, 1), (t_1, -1), (t_2, 1), \dots, (t_n, (-1)^n).$$

When $t_k = k\delta$, with $\delta = 1/n\mu$, this roughly requires an oscillation at a frequency n times greater than the Nyquist rate. We begin by considering the problem of interpolating two points of amplitude ± 1 separated by $\Delta t = 1/n\mu$. This implies a derivative of at least $2n\mu$ (locally). Inserting in (8)

$$\|f\|_2 \geq \frac{\mu^{-3/2}}{\pi} 2n\mu = \frac{2}{\pi} \frac{n}{\sqrt{\mu}}. \tag{9}$$

Assume now that we want to force a signal, bandlimited to $\mu/2$ as in (5), to jump from $f(a) = -1$, say, to $f(b) = 1$ in time

$b - a = \delta$. A bound similar to (8) must hold. But there is another way of seeing why it works

$$\begin{aligned} 2 = |f(a) - f(b)| &= \left| \int_{-\mu/2}^{+\mu/2} \widehat{f}(\omega) (e^{j2\pi\omega a} - e^{j2\pi\omega b}) d\omega \right| \\ &\leq 2 \int_{-\mu/2}^{+\mu/2} |\widehat{f}(\omega)| |\sin \pi\omega\delta| d\omega \\ &\leq 2 \|\widehat{f}\|_2 \left(\int_{-\mu/2}^{+\mu/2} |\sin \pi\omega\delta|^2 d\omega \right)^{1/2} \\ &= 2 \|\widehat{f}\|_2 \left(\frac{\pi\mu\delta - \text{sinc} \pi\mu\delta}{2\pi\delta} \right)^{1/2} \\ &= \sqrt{2\mu} \|f\|_2 \sqrt{1 - \text{sinc} \mu\delta} \\ &\leq \frac{\pi}{\sqrt{3}} \|f\|_2 \mu^{3/2} \delta \end{aligned}$$

where $\text{sinc} x = \sin(\pi x)/(\pi x)$. This leads to

$$\|f\|_2 \geq \frac{2\sqrt{3}}{\pi} \mu^{-3/2} \frac{1}{\delta}. \tag{10}$$

Taking, as before, $\delta = 1/n\mu$, we obtain

$$\|f\|_2 \geq \frac{2\sqrt{3}}{\pi} \frac{n}{\sqrt{\mu}}. \tag{11}$$

This slightly improves (9).

These approaches based on Bernstein’s inequality and Fourier analysis connect the energy, the bandwidth, the maximum derivative, and the period of the superoscillations. The bounds to which they lead hold for any signal within the limits considered. However, these methods do not yield information concerning the optimality of the bounds.

To recognize how far or how close they are from the best possible bound, we need to consider one specific signal, namely, the signal of smallest energy satisfying the stated constraints. To find it, we need to consider a simple variational problem: The minimization of

$$E = \int_{-\infty}^{+\infty} |f(t)|^2 dt = \int_{-\mu/2}^{+\mu/2} |\widehat{f}(\xi)|^2 d\xi$$

subject to the two constraints

$$f(0) = -1, \quad f(\delta) = 1.$$

Computing the Lagrangian and solving the problem in the standard way leads to

$$f(t) = -\frac{\text{sinc}(\mu t) - \text{sinc}[\mu(t - \delta)]}{1 - \text{sinc}(\mu\delta)} \tag{12}$$

which does indeed satisfy the interpolatory constraints. The energy of this signal is easily computed

$$E = \|f\|_2^2 = \int_{-\mu/2}^{+\mu/2} |\widehat{f}(\xi)|^2 d\xi = \frac{2}{\mu - \mu \text{sinc}(\mu\delta)}.$$

This establishes (1). Furthermore, asymptotically as $\mu\delta \rightarrow 0$, this becomes

$$\|f\|_2 \sim \frac{2\sqrt{3}}{\pi} \mu^{-3/2} \frac{1}{\delta}$$

which shows that (10) and (11) cannot be improved.

We have established the following theorem.

Theorem 1: The derivative of a finite energy signal f , bandlimited to $\mu/2$ Hz, may have arbitrarily high magnitude for any prescribed, fixed $\mu > 0$. For f to change from the value -1 to the value $+1$ in an interval of length δ , its energy must satisfy

$$E = \|f\|_2^2 \geq \frac{2}{\mu - \mu \operatorname{sinc}(\mu\delta)}$$

or, as $\mu\delta$ approaches zero

$$E = \|f\|_2^2 \sim \frac{12}{\pi^2} \frac{1}{\mu^3 \delta^2}.$$

The constants are the best possible, and the signal for which the bounds are met is given by (12).

The signal (12) is illustrated in Fig. 1 for several values of μ and $\delta = 1$. Note how the energy increases as the bandwidth μ decreases in order to satisfy the constraint.

III. INTERPOLATING N POINTS

We now turn to the problem of bounding the energy required to interpolate N points $f(t_k) = (-1)^k$, on a grid $t_k = k\delta$, using a signal bandlimited to $\mu/2$. We start with the expression

$$\sum_{k=0}^{N-1} f(t_k) a^k = \int_{-\mu/2}^{+\mu/2} \hat{f}(\omega) \left(\sum_{k=0}^{N-1} a^k e^{j2\pi\omega t_k} \right) d\omega.$$

If we set $t_k = k\delta$ and sum the geometric series, we get

$$\sum_{k=0}^{N-1} f(k\delta) a^k = \int_{-\mu/2}^{+\mu/2} \hat{f}(\omega) \left(\frac{1 - a^N e^{j2\pi\omega N\delta}}{1 - a e^{j2\pi\omega\delta}} \right) d\omega.$$

The special case $a = e^{j\alpha}$ leads to

$$\begin{aligned} \left| \sum_{k=0}^{N-1} f(k\delta) a^k \right| &\leq \int_{-\mu/2}^{+\mu/2} |\hat{f}(\omega)| \left| \frac{1 - a^N e^{j2\pi\omega N\delta}}{1 - a e^{j2\pi\omega\delta}} \right| d\omega \\ &= \int_{-\mu/2}^{+\mu/2} |\hat{f}(\omega)| \left| \frac{\sin(\pi\omega N\delta + \frac{\alpha N}{2})}{\sin(\pi\omega\delta + \frac{\alpha}{2})} \right| d\omega \\ &\leq \int_{-\mu/2}^{+\mu/2} |\hat{f}(\omega)| \frac{1}{|\sin(\pi\omega\delta + \frac{\alpha}{2})|} d\omega \\ &\leq \|f\|_2 \left(\int_{-\mu/2}^{+\mu/2} \frac{1}{\sin^2(\pi\omega\delta + \frac{\alpha}{2})} d\omega \right)^{1/2} \\ &= \|f\|_2 \left(\frac{2 \sin \pi\mu\delta}{\pi\delta \cos \pi\mu\delta - \pi\delta \cos \alpha} \right)^{1/2}. \end{aligned}$$

Finally, for $a = -1$, that is, $\alpha = \pi$, we obtain

$$N^2 \leq \|f\|_2^2 \frac{2 \sin \pi\mu\delta}{\pi\delta \cos \pi\mu\delta + \pi\delta}.$$

We find that the energy is therefore bounded below by

$$E = \|f\|_2^2 \geq N^2 \pi\delta \frac{1 + \cos \pi\mu\delta}{2 \sin \pi\mu\delta}$$

which behaves as

$$\frac{N^2}{\mu}$$

for small $\mu\delta$ and fixed δ . To check how far this result is from the best possible bound, we need to consider the following variational problem. Among all finite energy signals of bandwidth at most μ , find one that satisfies the constraints

$$f(t_k) = a_k, \quad k = 1, 2, \dots, N$$

and has minimum energy. The solution can be obtained in the standard way [11], [12] and is

$$f(t) = \mu \sum_{k=1}^N x_k \operatorname{sinc}[\mu(t - t_k)] \quad (13)$$

where a is the vector of amplitudes $\{a_k\}_{k=1}^N$ and the vector of coefficients x is the solution to

$$Sx = a.$$

The $N \times N$ matrix S has elements

$$S_{ij} = \mu \operatorname{sinc}[\mu(t_i - t_j)].$$

The Fourier transform of the interpolating signal (13) is

$$\hat{f}(\omega) = \sum_{k=1}^N x_k e^{-j2\pi\omega t_k} \chi(\omega)$$

where $\chi(\omega)$ is the characteristic function of the interval $[-\mu/2, \mu/2]$. Thus, the energy of the signal can be expressed as

$$\begin{aligned} E &= \int_{-\mu/2}^{+\mu/2} \left| \sum_{k=1}^N x_k e^{-j2\pi\omega t_k} \right|^2 d\omega \\ &= \sum_{k=1}^N x_k \sum_{\ell=1}^N x_\ell \int_{-\mu/2}^{+\mu/2} e^{j2\pi\omega(t_k - t_\ell)} d\omega \\ &= x^T Sx = a^T S^{-1} a \end{aligned}$$

since $Sx = a$. If the $\{t_k\}_{k=1}^N$ are equidistant, one has $t_k - t_\ell = (k - \ell)\delta$, and the elements of S are

$$S_{k\ell} = \mu \operatorname{sinc}[\mu\delta(k - \ell)]. \quad (14)$$

In this case, to emphasize the dependence of S on μ and δ , we denote it by $S(\mu, \delta)$. An example of a superoscillating signal, obtained using this method, is given in Fig. 2.

The matrix $\alpha \operatorname{sinc}[\alpha(i - j)]$, with $0 < \alpha < 1$, is a positive definite matrix, and its eigenvalues belong to $(0, 1)$ (see [13] for a more general discussion). It can, however, be extremely ill conditioned. As the smallest eigenvalue of S approaches zero, the energy $E = a^T S^{-1} a$ rapidly grows to infinity, as discussed in the next section.

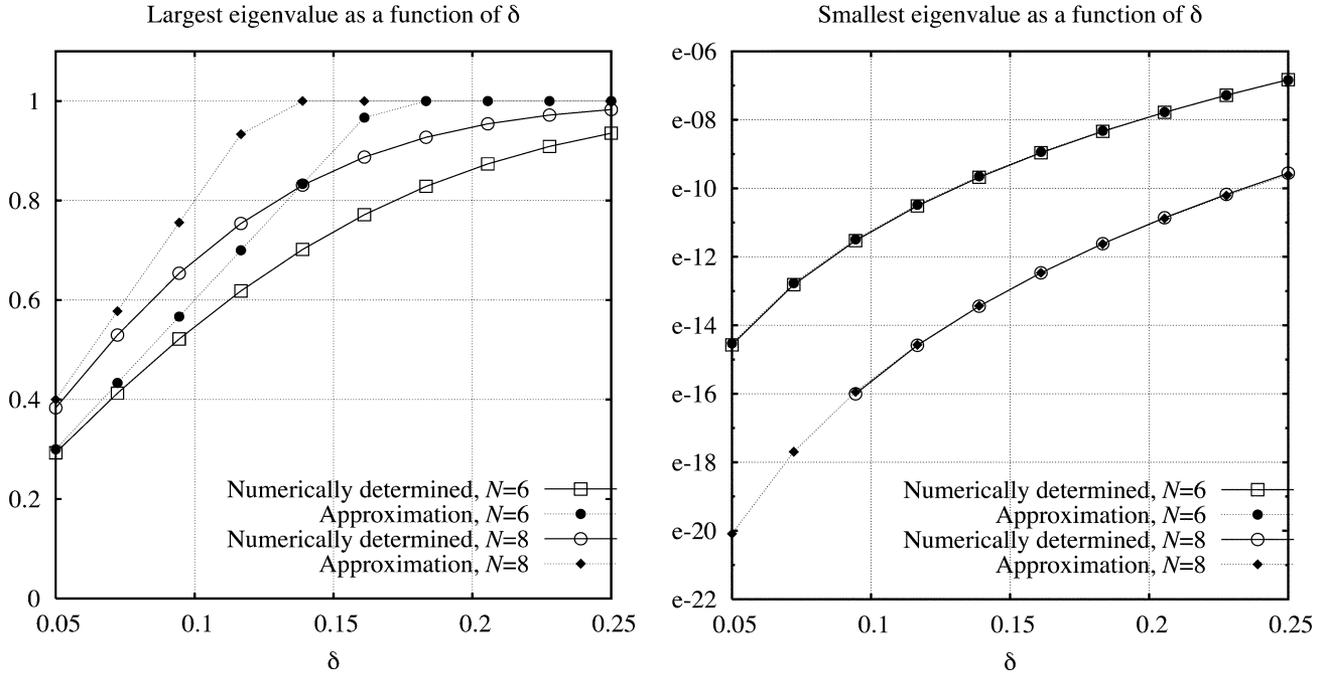


Fig. 3. The extreme eigenvalues of S as functions of δ for $N = 6$ and $N = 8$.

IV. THE REQUIRED ENERGY

Let $0 < W < 1/2$. The prolate matrix is the $N \times N$ matrix with elements

$$\rho(N, W)_{mn} = \frac{\sin 2\pi W(n - m)}{\pi(n - m)}, \quad m, n = 0, 1, \dots, N - 1$$

and its (real and distinct) eigenvalues are denoted by

$$\lambda_0(N, W) > \lambda_1(N, W) > \dots > \lambda_{N-1}(N, W).$$

The matrix is positive definite. It follows from [14, Th. 3] that $\lambda_0(N, W)$ satisfies

$$\lambda_0(N, W) \leq \min(2WN, 1). \tag{15}$$

Using the results in [5, Sec. 2.5], we derive that, for small W

$$\lambda_{N-1}(N, W) \sim \frac{G(N)}{\pi} (2\pi W)^{2N-1} [1 + O(W)]$$

where

$$G(N) = \frac{2^{2N-2}}{(2N - 1) \binom{2N-2}{N-1}^3}.$$

Expansion of the binomial coefficient using Stirling's formula yields the approximation

$$G(N) \sim \frac{\pi^{3/2}(N - 1)^{3/2}}{2^{4N-4}(2N - 1)}$$

and as a result

$$\lambda_{N-1}(N, W) \sim \sqrt{\pi} (2\pi W)^{2N-1} \frac{(N - 1)^{3/2}}{2^{4N-4}(2N - 1)}. \tag{16}$$

Let us return to $S(\mu, \delta)$, the elements of which are given by (14). The matrix is a multiple of the prolate matrix

$$S(\mu, \delta) = \frac{1}{\delta} \rho \left(N, \frac{\mu\delta}{2} \right).$$

It follows from (15) and (16) that its extreme eigenvalues can be approximated by

$$\begin{aligned} \lambda_0 &\leq \min(\mu\delta N, 1) \\ \lambda_{N-1} &\sim \sqrt{\pi} (\pi\mu\delta)^{2N-1} \frac{(N - 1)^{3/2}}{2^{4N-4}(2N - 1)}. \end{aligned}$$

See Fig. 3 for examples. We may now compute the energy $E = \|f\|_2^2$ required to force a signal f bandlimited to $\mu/2$ Hz to interpolate N given data,

$$f(k\delta) = a_k, \quad k = 0, 1, \dots, N - 1.$$

It is

$$\begin{aligned} E = \|f\|_2^2 &= a^T S^{-1} a \leq \|a\|^2 \lambda_0(S^{-1}) \\ &= \frac{\|a\|^2}{\lambda_{N-1}(S)} \\ &\sim \|a\|^2 \frac{2^{4N-4}(2N - 1)}{\sqrt{\pi} (\pi\mu\delta)^{2N-1} (N - 1)^{3/2}}. \end{aligned}$$

Here, $\|a\|$ is the Euclidean norm of the data vector. The equality holds when a is the eigenvector of S that corresponds to $\lambda_{N-1}(S)$. These conclusions are summarized in the following theorem.

Theorem 2: Let there be given an integer N , a set of N arbitrary reals $a = \{a_k\}_{k=1}^N$, and N distinct reals $\{t_k\}_{k=1}^N$. Among all signals of finite energy, bandlimited to $\mu/2$ Hz, there is one that has minimum energy and satisfies

$$f(t_k) = a_k, \quad k = 1, 2, \dots, N.$$

The signal is given by

$$f(t) = \mu \sum_{k=1}^N x_k \operatorname{sinc}[\mu(t - t_k)]$$

where $x = S^{-1}a$, and S is the (nonsingular) $N \times N$ matrix with elements

$$S_{ij} = \mu \operatorname{sinc}[\mu(t_i - t_j)].$$

The energy of f is given by

$$E = \|f\|_2^2 = a^T S^{-1} a \leq \frac{\|a\|^2}{\lambda_{N-1}(S)}$$

with equality when a is an eigenvector of S that corresponds to its smallest eigenvalue $\lambda_{N-1}(S)$. In that case, if $t_k = k\delta$, $k = 0, 1, \dots, N-1$, and for small $\mu\delta$

$$E \sim \|a\|^2 \frac{2^{4N-4}(2N-1)}{\sqrt{\pi}(\pi\mu\delta)^{2N-1}(N-1)^{3/2}}.$$

Equations (2) and (3) follow as corollaries (to obtain the first fix N and to reach the second fix μ and δ).

We have already obtained a bound for the energy in the case $N = 2$, using Bernstein's inequality. We have also verified the optimality of that bound using a Lagrangian (variational) formulation. It should be noted that the energy predicted in Theorem 2, for the case $N = 2$, is different. To understand why, note that this theorem gives an upper bound, not a lower bound, and that this upper bound is met if and only if a is an eigenvector corresponding to the smallest eigenvalue of S . This is not the case when $N = 2$ and $a = [-1, 1]$, a fact that explains the difference. One should note, however, that for larger N , alternating sign data such as $a_k = (-1)^k$ lead to energies close to the upper bound, and it is for these larger values of N that this theorem is of interest.

V. REMARKS AND CONCLUSIONS

A. The Condition Number of the Problem

The condition number of a nonsingular matrix S , with respect to the matrix norm $\|\cdot\|$, is $\kappa(S) := \|S\| \|S^{-1}\|$ and provides an important measure of the numerical difficulty of the linear problem $Sx = a$. If the exact data a are replaced with $a + \Delta a$, the solution x will change to $x + \Delta x$. The condition number sets a bound on the relative error in x

$$\frac{\|\Delta x\|}{\|x\|} \leq \kappa(S) \frac{\|\Delta a\|}{\|a\|}.$$

Condition number as a function of δ

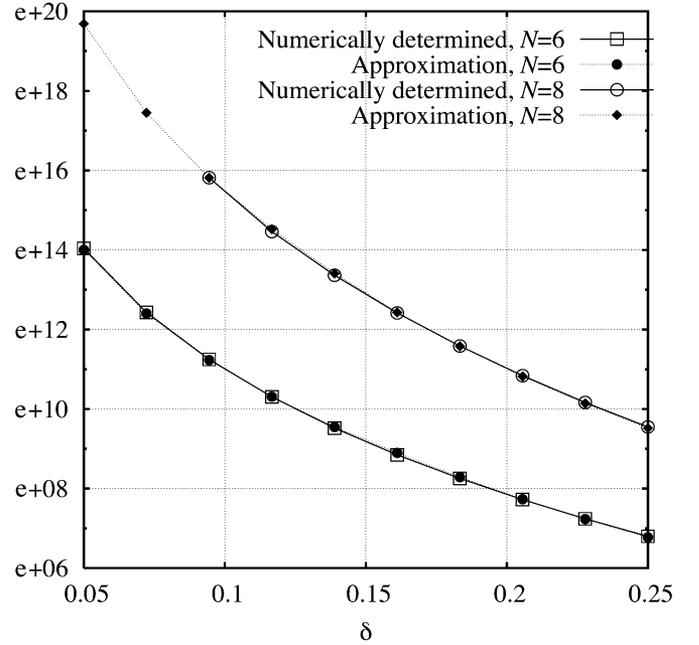


Fig. 4. The condition number of S as a function of δ for $N = 6$ and $N = 8$.

The preceding results lead to the following approximation for the condition number, with respect to the spectral norm

$$\kappa[S(\mu, \delta)] \sim \min(\mu\delta N, 1) \frac{2^{4(N-1)}(2N-1)}{\sqrt{\pi}(\pi\mu\delta)^{2N-1}(N-1)^{3/2}}.$$

The exact value of $\kappa(S)$ and this approximation are plotted in Fig. 4. The results clearly show how difficult the problem $Sx = a$ can be. Recall that, as a rule of thumb, in well-scaled problems a condition number of 10^d means a loss of roughly d significant figures. Thus, for $\delta = 0.1$ and $N = 8$, for example, we would lose about 15 significant figures. The numerical difficulties that must be faced when designing functions that have a large number of high-frequency superoscillations are obvious.

B. Superoscillations and Information Theory

We have already pointed out that the existence of superoscillations and the possibility of encoding arbitrary amounts of information into an arbitrarily short segment of a low-bandwidth signal do not contradict the basic laws of information theory. The reasons for this are twofold. First, the superoscillations' amplitude decreases exponentially with the length of the superoscillating stretch. Secondly, any channel through which the superoscillating signal is passed will add some noise to the signal. As Shannon showed [8], [9], signals that pass through a Gaussian channel of bandwidth B can deliver information on average at most at a rate $B \log_2(1 + S/N)$, where S/N is the signal-to-noise ratio. The theorem allows superoscillatory information compression to an arbitrary extent—but demands that the signal power grows exponentially with the amount of compressed information, i.e., with the length of the part of the message that is superoscillatory. Our findings are consistent

with this requirement. The L^2 norm of a maximally superoscillatory wave function grows exponentially with the number N of prescribed superoscillations.

C. Superoscillations, Interpolation, and the Nyquist Rate

We have mentioned that the well-known work of Beurling, Landau, and many others has explained how the density of a set of points $\Lambda = \{t_k\}$ determines the properties of the functions $\mathcal{E}(\Lambda) = \{e^{j2\pi t_k x}\}$ and how these properties relate to the stable sampling and interpolation problems.

Before proceeding, we recall that Λ is separated when $|t_m - t_n| \geq \delta$, for some positive δ , and that the upper density of a separated sequence is

$$D^+(\Lambda) = \lim_{r \rightarrow \infty} \frac{n^+(r)}{r}$$

where $n^+(r)$ is the largest number of elements of the sequence in any interval of length r .

We need to mention a few fundamental concepts (see [15] for a detailed exposition). The set Λ is an interpolating set for the Paley–Wiener space PW (the space of all finite energy functions bandlimited to 1/2 Hz) if and only if $\mathcal{E}(\Lambda)$ is a Riesz sequence for $L^2(-1/2, 1/2)$. On the other hand, for $\mathcal{E}(\Lambda)$ to be a Riesz sequence, it is necessary that Λ be separated and $D^+(\Lambda) \leq 1$, and it is sufficient that Λ be separated and $D^+(\Lambda) < 1$.

In essence, these results go back to the work of Beurling, and roughly speaking they state that the interpolation problem $f(t_k) = a_k$, $k \in \mathbf{Z}$, can be solved in the Paley–Wiener space only when the t_k are separated and their upper density is small enough.

Our results in no way contradict these. The interpolation problem studied in here is finite, and hence the distinct N numbers t_k are of course separated. As to their upper density, it is true that $n^+(r)$ can be large, since the number of t_k inside the superoscillating segment is large. However, the length of the superoscillating stretch is bounded, and there is only one such stretch. Therefore, for sufficiently large r , $n^+(r)/r$ will approach zero. This means that the upper density is zero, and that the interpolation problem is possible.

D. Superoscillation and Quantum Physics

The interest of superoscillations and associated phenomena is not restricted to signal processing. We have mentioned that superoscillations are of interest also in quantum physics, where the unusual properties of superoscillating wave functions shed new light on the Heisenberg uncertainty principle. A quantum mechanical wave function superoscillates if it contains local oscillations at length scales much smaller than the smallest wave length contained in its Fourier spectrum. As we have seen, however, superoscillating functions require excessively large dynamical ranges. This indicates that that it should be very difficult to produce superoscillating quantum mechanical wave functions experimentally.

E. Superoscillation and Time-Frequency Analysis

The existence of a signal bandlimited to 1 Hz and satisfying $f(k\delta) = a_k$ for $k = 1, 2, \dots, N$, for arbitrary δ , N , and data $\{a_k\}_1^N$ justifies the claim “that even the recording of an entire

Beethoven symphony can occur as part of a signal with 1 Hz bandwidth” (see [3] for a detailed discussion).

To understand the relevance of this for time-frequency analysis, consider a segment of length T of a signal bandlimited to $\mu/2$ and containing N superoscillations of period δ . The segment will be similar to the one depicted in Fig. 2.

It will without doubt mislead the observer as to the bandwidth of the signal. It is possible to commit large errors when estimating the bandwidth of a finite energy signal f from a sampled segment of length T , even if T is relatively large. The reason is that there exist signals of arbitrarily small bandwidth that possess the same set of samples.

If the observer is able to increase T at will, the true character of f will become apparent. The amplitude of a superoscillating signal will increase rapidly outside the superoscillating segment. In a sense, superoscillations are created by cancellation, and the amplitude inside the superoscillating segment is kept small only due to that cancellation (see Fig. 2). This also happens with the PSWF $\psi_n(t)$. Asymptotically, for large t , they behave like a sinc function of amplitude k_n and of the “correct” bandwidth (see [7] for details). If the observer is able to inspect $\psi_n(t)$ for large t , the bandwidth can be accurately estimated.

F. Conclusions

From the signal-processing point of view, superoscillating signals are an interesting topic and raise a number of issues concerning the behavior of bandlimited signals that at first glance appear surprising, counterintuitive, or even wrong.

We have examined some of the implications of superoscillations, concentrating on the energy (or L^2 norm) required to achieve a certain slope or a certain number N of oscillations, separated by δ seconds, with signals of maximum frequency $\mu/2$ Hz. We found asymptotic expressions for this energy that show that the energy increases exponentially with N and polynomially with $1/\mu$ or $1/\delta$. We have also identified the vectors a for which the problem $f(t_k) = a_k$, $k = 1, 2, \dots, N$, requires the smallest and the largest possible energy. Some of the implications of superoscillations were also considered and briefly discussed.

REFERENCES

- [1] Y. Aharonov, J. Anandan, S. Popescu, and L. Vaidman, “Superpositions of time evolutions of a quantum system and a quantum time-translocation machine,” *Phys. Rev. Lett.*, vol. 64, no. 25, pp. 2965–2968, Jun. 1990.
- [2] M. V. Berry, “Evanescence and real waves in quantum billiards and Gaussian beams,” *J. Phys. A: Math. Gen.*, vol. 27, no. 11, pp. L391–L398, 1994.
- [3] A. Kempf, “Black holes, bandwidths and Beethoven,” *J. Math. Phys.*, vol. 41, no. 4, pp. 2360–2374, Apr. 2000.
- [4] P. J. S. G. Ferreira and A. Kempf, “The energy expense of superoscillations,” in *Signal Process. XI—Theories Applcat.: Proc. EUSIPCO-2002 XI Eur. Signal Process. Conf.*, Toulouse, France, Sep. 2002, vol. II, pp. 347–350.
- [5] D. Slepian, “Prolate spheroidal wave functions, Fourier analysis and uncertainty — V: The discrete case,” *Bell Syst. Tech. J.*, vol. 57, no. 5, pp. 1371–1430, May 1978.
- [6] D. Slepian and H. O. Pollak, “Prolate spheroidal wave functions, Fourier analysis and uncertainty—I,” *Bell Syst. Tech. J.*, vol. 40, no. 1, pp. 43–63, Jan. 1961.
- [7] D. Slepian, “Some comments on Fourier analysis, uncertainty and modeling,” *SIAM Rev.*, vol. 25, no. 3, pp. 379–393, Jul. 1983.
- [8] C. E. Shannon, “A mathematical theory of communication,” *Bell Syst. Tech. J.*, vol. 27, pp. 379–423, Jul. 1948.

- [9] —, "Communication in the presence of noise," *Proc. IRE*, vol. 37, pp. 10–21, Jan. 1949.
- [10] A. Kempf and P. J. S. G. Ferreira, "Unusual properties of superoscillating particles," *J. Phys. A: Math. Gen.*, vol. 37, pp. 12067–12076, 2004.
- [11] R. Courant and D. Hilbert, *Methods of Mathematical Physics*. New York: Interscience, 1953, vol. 1.
- [12] L. Levi, "Fitting a bandlimited signal to given points," *IEEE Trans. Inf. Theory*, vol. IT-11, pp. 372–376, Jul. 1965.
- [13] P. J. S. G. Ferreira, "Incomplete sampling series and the recovery of missing samples from oversampled band-limited signals," *IEEE Trans. Signal Process.*, vol. 40, pp. 225–227, Jan. 1992.
- [14] —, "The stability of a procedure for the recovery of lost samples in band-limited signals," *Signal Process.*, vol. 40, no. 3, pp. 195–205, Dec. 1994.
- [15] H. J. Landau, "Sampling, data transmission, and the Nyquist rate," *Proc. IEEE*, vol. 55, pp. 1701–1706, Oct. 1967.



Paulo J. S. G. Ferreira (S'85–M'91) was born in Torres Novas, Portugal.

He is Professor Catedrático at the Departamento de Electrónica e Telecomunicações/IEETA, Universidade de Aveiro, Portugal. He is currently a Member of the Editorial Board of the *Journal of Applied Functional Analysis* and an Editor-in-Chief of *Sampling Theory in Signal and Image Processing*. He coedited (with J. J. Benedetto) *Modern Sampling Theory: Mathematics and Applications* (Boston, MA: Birkhauser, 2001). His research interests

include bioinformatics and symbolic signals, coding, sampling, and signal reconstruction.

He was an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING.



Achim Kempf was born in Heidelberg, Germany. He was educated at the Universities of Heidelberg, Karlsruhe, and Munich, Germany.

A mathematical physicist by training, he has held postdoctoral positions with the University of Florida and Cambridge University, where he has also been a Research Fellow of Corpus Christi College. He now holds the Canada Research Chair in the Physics of Information in the Department of Applied Mathematics, University of Waterloo, Waterloo, ON, Canada. His research interests range from

classical sampling theory to quantum information theory, quantum gravity, and inflationary cosmology.

Prof. Kempf is a member of the Institute for Quantum Computing and an affiliated member of the Perimeter Institute for Theoretical Physics.