Chapter 1

Superoscillations

Paulo J. S. G. Ferreira

Dedicated to Paul L. Butzer, in friendship and high esteem.

Abstract

A band-limited function may oscillate faster than its maximum Fourier component, and it may do so over arbitrarily long intervals. The goal of this chapter is discuss this phenomenon, which has been called "superoscillation". Although the theoretical interest in superoscillating functions is relatively recent, a number of applications are already known (in quantum physics, super-resolution, subwavelength imaging and antenna theory). This chapter gives a brief account of how superoscillations appeared and developed, discusses their cost and some of their implications.

1.1 Introduction

Classical sampling theory, to which Paul Butzer has contributed so extensively[1], has band-limited functions at the core. The corresponding sampling theorems determine the value of a band-limited function in terms of its values at countable, discrete sets of sufficiently high density. Typically, if f belongs to an appropriate space of band-limited functions, its value at any point t of its domain can be determined by the values $\{f(t_i)\}_{i\in\mathbb{Z}}$, where $\Lambda = \{t_i\}$ is a fixed set.

The density of Λ must be higher than a certain critical value, which depends on the bandwidth of f. This critical value is usually called the Nyquist rate, or, in somewhat different contexts, the Beurling-Landau density[2].

Naturally, the higher the bandwidth, the denser Λ needs to be. In the simplest case, in which the points in Λ form an arithmetic progression, one can define a sampling period $T = t_{i+1} - t_i$ and a sampling frequency $1/T$. Classical sampling theory

[1]Paul Butzer has contributed significantly to sampling itself and to the study of its history. His work along the latter line includes [10, 14] and the more recent articles [11–13]. The history of the sampling principle is particularly challenging because the main ideas appeared in several independent works across the world: USA, Europe, Japan [13] and Russia (see the translation of Kotel'nikov's work in [3]). The multiple discovery of the sampling principle is the subject of [20].

[2]See [31, 32], in which the so-called Nyquist rate is shown to be the minimum rate at which stable reconstruction can be performed.

then asserts that the sampling frequency must be at least twice of the maximum frequency of f. In this setting, the Nyquist rate is precisely "twice the maximum frequency" or, better, "twice the bandwidth of f".

Sampling theorems are sometimes stated along with heuristic explanations of the following type: a band-limited function contains no frequencies above a limit $\mu/2$, and so it cannot change to substantially new values in a time less than one-half cycle of its highest frequency, that is, $1/\mu$. Another commonly found statement is the following: since a band-limited function has a maximal frequency component, it cannot oscillate faster than that.

These statements are false. Counterexamples do exist, and they involve a phenomenon that has been called "superoscillation". Briefly, superoscillating functions are band-limited functions that oscillate faster than their maximal Fourier component. They provide direct (maybe striking) refutations of statements of the kind quoted above. To better understand what is at stake, consider the following facts:

1. There exist finite-energy signals band-limited to (say) 1 Hz, the values of which change from (say) -1 to 1 in an interval of duration δ, where δ is as small as desired.

2. There exist finite-energy signals f, band-limited to (say) 1 Hz, that 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 $i = 1, 2, ..., N$, where N is as large as desired.

In engineering terms, the first assertion means that a signal band-limited to a very low frequency can still have an arbitrarily high slew rate.

The second assertion is more interesting. It raises a few questions in the context of approximation theory (another area to which Paul Butzer has contributed so much). It implies that over a finite interval, a signal band-limited to (say) 1 Hz may look very much like a sinusoid of frequency f_0, with f_0 as large as desired. The implications for spectrum analysis are clear: the analysis of a signal over a finite observation window may suggest a certain frequency content and bandwidth, but the signal may as well have an arbitrarily smaller bandwidth.

The goal of this chapter is to discuss in more detail superoscillations and their consequences, in terms as simple as possible. I begin with a brief account of how superoscillations appeared. Then I discuss the construction of superoscillations and examine their energy cost. The chapter closes with a few remarks about the significance and impact of superoscillations in connection with information theory, sampling and other topics.

1.2 From zeros to superoscillations: a brief review

1.2.1 Distortion and real zeros

Given a real function f and a fixed positive real α, consider the function g defined in the following way:

$$g(t) = \begin{cases} \alpha, & \text{for all } t \text{ where } |f(t)| \geq \alpha, \\ f(t), & \text{elsewhere.} \end{cases}$$

The function g is often called a clipped version of f. Clearly, g retains whatever information f carries, at amplitude levels below α. The function

$$h(t) = \begin{cases} \dfrac{f(t)}{|f(t)|}, & f(t) \neq 0, \\ 0, & f(t) = 0 \end{cases}$$

can be regarded as an extreme form of clipping because the only information it retains about f is its sign (and obviously the position of its real zeros). The signal h can be generated very easily: drive f through an operational amplifier or comparator, and set the comparison threshold to zero.

The early discovery that speech signals remain intelligible[3] even after such extreme clipping prompted interest in the possibility of transmitting them by preserving the occurrence of their zero crossings (and no other information).

The transmission of binary signals (that is, signals with values ± 1) leads to a related problem. Assume that it is possible to generate a band-limited function with a prescribed set of zero-crossings, from which the binary signal can be recovered. This would reduce the problem of representing binary signals to the problem of representing band-limited ones, which can be solved by means of classical sampling, leading to a way of describing binary data over a channel of limited bandwidth.

However, the idea raises other questions: Is the process always possible? How should the bandwidth be chosen? If, to fix ideas, one considers finite-energy functions only, what energy would be necessary to generate the signals? How would that energy depend on the bandwidth?

1.2.2 Zero manipulation and its cost: early work

Such questions suggest a study of the representation and manipulation of signals by means of their real and complex zeros, which is related to the oscillatory behaviour of band-limited functions (discussed for example in [24, §7.4]). The topic is of interest for the study of superoscillations, of course. But, conversely, the study of superoscillations can also shed light on the (local) oscillatory behaviour of band-limited functions.

[3]As an example, Voelcker and Requicha [45] mention that speech that has been SSB-clipped (SSB modulated, clipped, retranslated to baseband) was known to possess much higher quality than LP-clipped speech. The references can be traced back to an Information Theory meeting held in 1956.

One of the first works that explore the representation and manipulation of signals through their zeros is [9], by Bond and Cahn, published in 1958.

Its authors knew that successive differentiation of a signal might lead to signals with zeros at the Nyquist rate, a conclusion that they reached based on the work of Rice [38, 39], who obtained the average rate at which a random band-limited noise with a flat spectral density crosses the zero axis. They were also aware of results on the zeros of entire functions, due to Titchmarsh, who found the average zero density of a band-limited signal. The paper [9] gives a solution to the problem of synthesising a band-limited function with a given set of zeros.

However, its authors also observed that the distribution of the zeros could affect the amplitude of the generated band-limited signal, a fact that may have important practical consequences.

When physically generating a signal, or when computing signal values, one cannot deal with arbitrarily large signal amplitudes. The zero distributions that lead to extreme amplitudes, even if theoretically possible, are not practical and cannot be used for experimentation. The article [9] represents an early example of work that distinguishes between possible and practical zero distributions of band-limited functions.

1.2.3 Superoscillations and their cost

This question of the amplitudes, already mentioned in [9], was clearly treated by Berry [4], in 1994. Berry defines superoscillations and attributes the idea to Aharonov, who had told him that functions "could oscillate faster than any of their Fourier components" and had constructed such "superoscillations" guided by quantum-mechanical arguments [1]. Using integral representations, Berry constructed superoscillations and discussed their cost: the value of the function containing the superoscillations is exponentially larger in the range where the function oscillates conventionally. He gave a striking example: the reproduction of Beethoven's ninth symphony as superoscillations with a 1 Hz bandwidth would require a signal $\exp(10^{19})$ stronger than with conventional oscillations.

The cost of superoscillations in terms of energy is also discussed in [21] and especially [22]. These papers adopt a signal processing perspective. The paper [22] investigates the dynamical range and energy required by superoscillating signals as a function of the superoscillation's frequency, number, and maximum derivative. It discusses some of the implications of superoscillating signals, in the context of information theory and time-frequency analysis. It 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 period of superoscillation. It also shows that there is no contradiction between Shannon's capacity formula and superoscillating signals, and the role that the amplitude and energy of such signals play in the matter.

Also in the perspective of signal processing, a recent article [27] considers the problem of optimising superoscillatory signals. The authors maximise 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 solved superoscillations.

The constrained optimisation leads to a generalised eigenvalue problem, which can be solved numerically.

The article [2] is interesting from the viewpoint of approximation theory. It defines the set of superoscillation in terms of the uniform convergence of functions on such a set and studies the problem of the approximation of a function by superoscillating functions.

1.2.4 Superoscillations in applications

Superoscillations and Beethoven's symphony reappear in [28], which discusses superoscillations applied to the problem of transplanckian frequencies in black hole radiation. For another application to physics see [5], which discusses superoscillations in the context of a quantum billiards problem. The article [29] considers superoscillations in quantum mechanical wave functions, and unusual associated phenomena that are of measurement theoretic, thermodynamic and information theoretic interest.

Applications of superoscillations to superresolution and subwavelength imaging have also been given. The article [8] discusses optical superresolution without evanescent waves, whereas [25] 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 [52] discusses approaches capable of beating the diffraction limit, pointing out superoscillations as a possible alternative. Another related work is [26], which proposes a solution that removes any need for evanescent fields. The object being imaged does not need to be in the immediate proximity of the superlens or field concentrator. Instead, 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 method introduced in [48, 49] stems from the observation that superdirectivity and superoscillation are related phenomena. The results are subwavelength focusing schemes in free space and within a waveguide. The simulations reported by the authors demonstrate subwavelength focusing down to 0.6 times the diffraction limit, five wavelengths away from the source. The work [50] 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 localised, but the work [51] 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 [16] 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 [30] to obtain results on a quantum gravity motivated natural ultraviolet cutoff which describes an upper bound on the spatial density of information. The article [7] deals with superoscillations in monochromatic waves in several dimensions. Other applications to physics include [17, 44] and [6], the latter on backflow, a phenomenon related to superoscillation.

1.3 How to build superoscillations

The sinc function

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

(with $\operatorname{sinc}(0) = 1$, of course) illustrates the typical band-limited function of finite energy. Its maximum frequency is 1/2 Hz (or π rad/seg) and it shows "oscillations" of period 2, which match the maximum frequency perfectly. The separation between its zeros is one unit, with one exception (there is an interval of length greater than one without any zeros). If all the zeros occurred at the Nyquist rate, the function would have to be identically zero (this follows from the sampling theorem, with the zero instants as sampling points).

This suggests that the existence of a larger zero-free interval cannot be a peculiarity of the sinc function alone. It will be shown that the zero-free interval can be arbitrarily extended, but there is a lower bound to its size.

Functions with "very large" zero-free intervals may appear to "oscillate less" than the sinc function. The more interesting functions with "too many" real zeros appear (locally) to oscillate faster than the sinc (yet have the same bandwidth). These matters will now be discussed in more detail.

1.3.1 "Too few" zeros

It is possible to construct band-limited functions with arbitrarily long zero-free intervals. In fact, it is possible to construct band-limited functions with no real zeros at all. For example, it is possible to pick α so that

$$f(t) = \operatorname{sinc}^2(t - \alpha) + \operatorname{sinc}^2(t + \alpha)$$

has no real zeros. Another example is the function

$$g(t) = \frac{1 - \operatorname{sinc} t}{t^2}$$

which is discussed in [24, §7.4].

Let $n(r)$ denote the number of zeros of a signal in the region $|z| \leq r$ of the complex z-plane. Titchmarsh's theorem asserts that, for finite-energy signals band-limited to 1/2 Hz (that is, for PW_π functions),

$$\frac{n(r)}{r} \to 2, \quad r \to \infty,$$

which of course matches the zero density of $\operatorname{sinc}(t)$.

PW_π functions with larger zero-free gaps that the sinc function are easy to build, but there must always be one gap greater than 1. This is the essence of a theorem due to Walker [46, 47], also discussed in [24, §7.4]. The theorem asserts that if $f \in PW_{\pi w}$, then $|f| > 0$ on at least one open interval of the real axis whose length exceeds $1/w$. Its elegant proof is by contradiction: if all the zeros were separated by less than $1/w$, an application of Wirtinger's inequality to a pair of consecutive zeros, followed by a summation over all such pairs and an appeal to Bernstein's inequality would lead to a contradiction.

Successive differentiation of a signal usually increases the number of its real zeros. Higgins posed the question whether some derivative of any real PW function has infinitely many real zeros. The negative answer was given in [36].

1.3.2 "Too many" zeros

The preceding constructions yield PW functions with a sparser zero-crossings than those of the sinc. They have "fewer zeros" than one could have anticipated, given the behaviour of the archetypal PW function, $\mathrm{sinc}(t)$. There are functions with arbitrarily large zero-free gaps, and even functions with no real zeros at all. In a sense, some of them seem to "sub-oscillate".

However, PW functions may also have "too many" zeros, or "superoscillate". The construction in [37] is maybe the simplest: the superoscillations are built by replacing zeros in the sinc function. One starts with the standard expansion

$$\mathrm{sinc}(t) = \prod_{n=1}^{\infty} \left(1 - \frac{t^2}{n^2}\right),$$

and the result after shifting N zeros is

$$g(t) = \prod_{n=1}^{N} \left(1 - \frac{k^2 t^2}{n^2}\right) \prod_{n=N+1}^{\infty} \left(1 - \frac{t^2}{n^2}\right)$$

where $k > 1$. The function $g(t)$ is in PW but locally behaves as $\mathrm{sinc}(kt)$.

1.3.3 Prolate spheroidals

The well-known prolate spheroidal functions $\psi_n(t)$ [33, 43] also provide examples of superoscillations. Recall that they are band-limited to 1/2 Hz and have finite energy. They also possess a double orthogonality property: they are orthonormal on the real line *and* on the interval $(-1, 1)$. Moreover, $\psi_n(t)$ has n zeros inside $(-1, 1)$. This means that the set of prolate spheroidal functions contains functions that oscillate faster than any given prescribed number, in the interval $(-1, 1)$. This implies nothing about their behaviour (amplitude, derivative, etc.) outside this interval.

However, each ψ_n is an eigenfunction of a certain operator, associated with an eigenvalue λ_n, which also measures the energy concentration of ψ_n inside $(-1, 1)$. Since $\lambda_n \to 0$ as $n \to \infty$, the energy of the ψ_n inside $(-1, 1)$ tends to zero as $n \to \infty$.

Thus, as the number of superoscillations inside $(-1,1)$ increases, the energy (and amplitude) of the superoscillating function inside $(-1,1)$ decreases. It is of course 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 and amplitude of the rescaled ψ_n must increase with n, a fact consistent with the remarks that were made before.

1.3.4 Controlling the superoscillating segment

The methods discussed so far do not allow any control over the shape of the superoscillating segment. Interpolation renders this possible. Consider, for example, that f is required to satisfy

$$f(k\delta) = a_k, \qquad 0 \le k < N.$$

Then, for fixed N, the energy of f increases polynomially with $\mu\delta$:

$$E = \mathcal{O}\left[\frac{1}{(\mu\delta)^{2N-1}}\right]. \tag{1.1}$$

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

$$E = \mathcal{O}\left[N^{-1/2}\left(\frac{4}{\pi\mu\delta}\right)^{2N-1}\right]. \tag{1.2}$$

Moreover, there is exactly one data set a_k for which the equality sign is valid. Alternating sign data such as $a_k = (-1)^k$ lead to energies very close to this upper bound. The rest of this section is dedicated to these matters (see also [22]).

Imagine that it is necessary to interpolate N points $f(t_k) = (-1)^k$, on a grid $t_k = k\delta$, using a signal band-limited to $\mu/2$. Consider the expression

$$\sum_{k=0}^{N-1} f(t_k)a^k = \int_{-\mu/2}^{+\mu/2} \widehat{f}(\omega)\left(\sum_{k=0}^{N-1} a^k\, e^{i2\pi\omega t_k}\right) d\omega.$$

Setting $t_k = k\delta$ and summing the geometric series leads to

$$\sum_{k=0}^{N-1} f(k\delta)a^k = \int_{-\mu/2}^{+\mu/2} \widehat{f}(\omega)\left(\frac{1 - a^N\, e^{i2\pi\omega N\delta}}{1 - a\, e^{i2\pi\omega\delta}}\right) d\omega.$$

The special case $a = e^{i\alpha}$ yields

$$
\begin{aligned}
\left|\sum_{k=0}^{N-1} f(k\delta)a^k\right| &\le \int_{-\mu/2}^{+\mu/2} |\widehat{f}(\omega)| \left|\frac{1-a^N\, e^{i2\pi\omega N\delta}}{1-a\, e^{i2\pi\omega\delta}}\right| d\omega \\
&= \int_{-\mu/2}^{+\mu/2} |\widehat{f}(\omega)| \left|\frac{\sin(\pi\omega N\delta + \alpha N/2)}{\sin(\pi\omega\delta + \alpha/2)}\right| d\omega \\
&\le \int_{-\mu/2}^{+\mu/2} |\widehat{f}(\omega)| \frac{1}{|\sin(\pi\omega\delta + \alpha/2)|}\, d\omega \\
&\le \|f\|_2 \left(\int_{-\mu/2}^{+\mu/2} \frac{1}{\sin^2(\pi\omega\delta + \alpha/2)}\, d\omega\right)^{1/2} \\
&= \|f\|_2 \left(\frac{2\sin\pi\mu\delta}{\pi\delta\cos\pi\mu\delta - \pi\delta\cos\alpha}\right)^{1/2}.
\end{aligned}
$$

For $a = -1$, that is, $\alpha = \pi$, this leads to

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

Thus, the energy required is bounded below by

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

which behaves as

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

for small $\mu\delta$, and fixed δ. This can be compared with the best possible bound, which follows from the following variational problem: Among all finite energy signals of bandwidth at most μ, find one that satisfies the constraints

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

and has minimum energy. The solution can be obtained by standard methods [15,35], and it is:

$$
f(t) = \mu \sum_{k=1}^{N} x_k \operatorname{sinc}[\mu(t - t_k)], \tag{1.3}
$$

where a is the vector of amplitudes $\{a_k\}_{k=1}^N$. 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 (1.3) is

$$
\widehat{f}(\omega) = \sum_{k=1}^{N} x_k\, e^{-i2\pi\omega t_k}\, \chi(\omega)
$$

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

$$
\begin{aligned}
E &= \int_{-\mu/2}^{+\mu/2} \left| \sum_{k=1}^{N} x_k \, e^{-i2\pi\omega t_k} \right|^2 d\omega \\
&= \sum_{k=1}^{N} x_k \sum_{\ell=1}^{N} x_\ell \int_{-\mu/2}^{+\mu/2} e^{i2\pi\omega(t_k - t_\ell)} d\omega \\
&= x^T S x \;=\; a^T S^{-1} a,
\end{aligned}
$$

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

$$S_{k\ell} = \mu \; \mathrm{sinc}[\mu\delta(k-\ell)]. \tag{1.4}$$

In this case, the notation $S(\mu, \delta)$ will be used. It is a convenient way of stressing the dependence of S on μ and δ.

It is easy to generate examples of superoscillating signals using this method: pick δ, select the amplitudes a_k, form the matrix S, find the coefficients x and use (1.3). Examples are given in [22].

The matrix α sinc$[\alpha(i-j)]$, with $0 < \alpha < 1$, is positive definite with eigenvalues in $(0,1)$ (see [18] for a more general discussion). However, it can be extremely ill-conditioned. As its smallest eigenvalue approaches zero, the energy $E = a^T S^{-1} a$ tends to infinity.

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

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

and its (real and distinct) eigenvalues are denoted by

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

The matrix is positive definite. It follows from [19, Theorem 3] that $\lambda_0(N, W)$ satisfies

$$\lambda_0(N, W) \leq \min(2WN, 1). \tag{1.5}$$

Using the results in [42, Section 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

$$\lambda_{N-1}(N,W) \sim \sqrt{\pi}\,(2\pi W)^{2N-1}\frac{(N-1)^{3/2}}{2^{4N-4}(2N-1)}. \tag{1.6}$$

Let us consider $S(\mu,\delta)$ again. Its elements are given by (1.4). The matrix is clearly a multiple of the prolate matrix:

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

It follows from (1.5) and (1.6) that its extreme eigenvalues can be approximated by

$$\lambda_0 \leq \min(\mu\delta N, 1),$$

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

We may now compute the energy $E = \|f\|_2^2$ required to force a signal f band-limited to $\mu/2$ Hz to interpolate N given data,

$$f(k\delta) = a_k, \qquad k = 0, 1, \ldots, N-1.$$

The energy 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 summarised in the following theorem.

Theorem 1 Consider an integer N, N arbitrary reals $a = \{a_k\}_{k=1}^N$ and N distinct reals $\{t_k\}_{k=1}^N$. Among all signals of finite energy band-limited to $\mu/2$ Hz, there is one that satisfies

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

and that has minimum energy. 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 $t_k = k\delta$, $k = 0, 1, \ldots, N-1$, and for small $\mu\delta$,

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

The equations (1.1) and (1.2) given at the beginning of this section are corollaries (to obtain the first fix N, to reach the second fix μ and δ).

1.4 The behaviour at different scales

This section gives an alternative construction of superoscillatory functions, which explores the connection between the fine and coarse structure of a band-limited function [23]. As in the previous section, the behaviour of the superoscillating segment is kept under control, but the method used is very different.

The band-limited functions that will be considered are determined by N coefficients c_k:

$$f(t) = \sum_{k=0}^{N-1} c_k \,\mathrm{sinc}(t-k).$$

Thus, they form a linear N-th dimensional subspace of PW. Plainly, $c_k = f(k)$, $0 \leq k < N$, whereas $f(k) = 0$ for any other integer k. The functions considered are therefore concentrated, as far as the values $f(k)$ are concerned, in the interval $[0, N-1]$.

The samples $f(k)$ can be associated with "scale 1 behavior". The samples $f(kT)$, on the other hand, describe "scale T behavior". Because $f(k) = c_k$, the coefficients c_k determine the behavior of $f(t)$ at scale 1. The surprising fact is that they can also control some of the behavior of $f(t)$ at scale T, and T can be as small as desired. In fact, the c_k can be chosen so that N relations of the form $f(kT) = a_k$ are satisfied. This remains true even if the kT lie at an arbitrary distance from the interval $[0, N-1]$.

1.4.1 The mappings between scale 1 and scale T

Any square-integrable function f, band-limited to 1/2 Hz, satisfies the classical sampling formula [3, 24]

$$f(t) = \sum_{k=-\infty}^{+\infty} f(k) \,\mathrm{sinc}(t-k). \tag{1.7}$$

A simple variant of a well-known proof of (1.7) leads to the "oversampled expansion" [18]

$$f(t) = T \sum_{k=-\infty}^{+\infty} f(kT) \,\mathrm{sinc}(t-kT), \quad (T<1). \tag{1.8}$$

Both equations are needed to obtain the mappings. First, set $t = \ell T$ in (1.7). This leads to

$$f(\ell T) = \sum_{k=-\infty}^{+\infty} f(k)\ \mathrm{sinc}(\ell T - k). \tag{1.9}$$

Then, set $t = \ell$ in (1.8). The result is

$$f(\ell) = T \sum_{k=-\infty}^{+\infty} f(kT)\ \mathrm{sinc}(\ell - kT). \tag{1.10}$$

Equation (1.9) specifies the mapping between the samples $f(k)$ at the Nyquist rate and the samples $f(kT)$ taken at the higher rate $1/T$. Equation (1.10) specifies the mapping in the opposite direction. These linear mappings between the values of f at grids of size 1 and T, which were first described in [23], will be needed later on.

Before going on, note that there is no nonzero function that satisfies the oversampled expansion (1.8) with finitely many nonzero coefficients. In contrast with the coefficients $f(k)$ in equation (1.7), which are independent of each other, those of the oversampled expansion (1.8) cannot be arbitrarily prescribed. As a result, if all but a finite number of the coefficients in (1.8) are zero, then the function itself must be zero. This can be shown easily, in more than one way [23].

It is therefore impossible to construct superoscillating band-limited functions by prescribing finitely many amplitudes as in

$$f(kT) = a_k, \qquad k \in J, \qquad \mathrm{card}(J) < \infty, \quad 0 < T < 1, \tag{1.11}$$

and then setting all remaining samples to zero: $f(kT) = 0, k \notin J$. If this were possible, the behaviour of the superoscillating segment could be controlled simply by controlling the a_k, say, by setting $a_k = (-1)^k$. But, as seen, the only band-limited function that satisfies $f(kT) = 0$ for $k \notin J$ is $f = 0$.

The impossibility of building a superoscillating function based on a finite oversampled series of the type (1.8) raises a question. Can a superoscillating function be built by prescribing its values on a grid of size T as in (1.11), and simultaneously enforcing a finite expansion of type (1.7)? In other words, are there functions satisfying (1.11) and also

$$f(t) = \sum_{k \in I} f(k)\ \mathrm{sinc}(t - k) \tag{1.12}$$

for some *finite* index set I, say, $I = \{0, 1, \ldots, N-1\}$?

The obvious necessary condition is that the number of degrees of freedom in this sum (given by the cardinal of I) must be at least equal to the number of interpolatory constraints (given by the cardinal of J). For simplicity, I will assume $\mathrm{card}(I) = \mathrm{card}(J)$ (although $\mathrm{card}(I) \geq \mathrm{card}(J)$ could be considered too.)

To show that the necessary condition

$$\mathrm{card}(I) = \mathrm{card}(J)$$

is also sufficient, it must be shown that there are functions given by the finite expansion (1.12) that interpolate a number of prescribed points separated by T, as in (1.11). Any such function, by construction, is a PW function.

To construct such functions, it will be shown that their coefficients satisfy a set of linear equations with a nonsingular matrix. To obtain the equations, rewrite (1.10) as

$$f(\ell) = T \sum_{k \notin J} f(kT) \ \mathrm{sinc}(\ell - kT) + g(\ell)$$

where

$$g(\ell) = T \sum_{k \in J} a_k \ \mathrm{sinc}(\ell - kT).$$

The quantities $g(\ell)$ are known for any ℓ, since the a_k and J are given. Substitution of (1.12), which is the finite version of (1.9), leads to

$$\begin{aligned} f(\ell) &= T \sum_{k \notin J} \sum_{m \in I} f(m) \ \mathrm{sinc}(kT - m) \ \mathrm{sinc}(\ell - kT) + g(\ell) \\ &= T \sum_{m \in I} f(m) \sum_{k \notin J} \mathrm{sinc}(kT - m) \ \mathrm{sinc}(\ell - kT) + g(\ell). \end{aligned}$$

Restriction to $\ell \in I$ leads to a set of linear equations of the form $\mathbf{f} = M\mathbf{f} + \mathbf{g}$, where $\mathbf{f}$ is the vector of all $f(\ell)$, for $\ell \in I$, and similarly for $\mathbf{g}$. The square matrix M has elements

$$M_{m\ell} = T \sum_{k \notin J} \mathrm{sinc}(kT - m) \ \mathrm{sinc}(\ell - kT), \qquad m, \ell \in I.$$

It is convenient to convert this equation to a more convenient form. The replacement of $f(t)$ with $\mathrm{sinc}(t - m)$ in (1.8) leads to

$$T \sum_{k=-\infty}^{+\infty} \mathrm{sinc}(kT - m) \ \mathrm{sinc}(\ell - kT) = \mathrm{sinc}(\ell - m) = \delta_{\ell m},$$

and consequently

$$\begin{aligned} M_{m\ell} = T \sum_{k \notin J} \mathrm{sinc}(kT - m) \ \mathrm{sinc}(\ell - kT) = \\ = \delta_{\ell m} - T \sum_{k \in J} \mathrm{sinc}(kT - m) \, \mathrm{sinc}(\ell - kT). \end{aligned}$$

The advantage of this form is that it allows the computation of M using a finite sum (J has finitely many elements). The equations $\mathbf{f} = M\mathbf{f} + \mathbf{g}$ become

$$\begin{aligned} \sum_{m \in I} f(m) \, T \sum_{k \in J} \mathrm{sinc}(kT - m) \ \mathrm{sinc}(\ell - kT) = \\ = \sum_{k \in J} a_k \ \mathrm{sinc}(\ell - kT), \end{aligned}$$

or, equivalently,

$$A\mathbf{f} = \mathbf{g}, \tag{1.13}$$

with

$$A_{m\ell} = T \sum_{k \in J} \mathrm{sinc}(kT - m) \ \mathrm{sinc}(\ell - kT), \qquad m, \ell \in I.$$

Note that A and $\mathbf{g}$ depend only on finitely many known quantities. Unlike $\mathbf{g}$, the matrix A is independent of the prescribed amplitudes $f(kT) = a_k$, and depends only on T and the index sets I and J.

The nonsingularity of the matrix A still needs to be investigated. The elements of the matrix are

$$A_{m\ell} = T \sum_{k \in J} \int_{-1/2}^{1/2} e^{i2\pi(kT-m)x}\, dx \int_{-1/2}^{1/2} e^{i2\pi(\ell-kT)y}\, dy$$

and therefore the quadratic form $\mathbf{v}^H A \mathbf{v}$ can be written

$$\begin{aligned} T \sum_m \sum_\ell v_m^* v_\ell \sum_{k \in J} \int_{-1/2}^{1/2} e^{i2\pi(kT-m)x}\, dx \int_{-1/2}^{1/2} e^{i2\pi(\ell-kT)y}\, dy = \\ = \; T \sum_{k \in J} \int \int_{-1/2}^{1/2} \left(\sum_m v_m^* e^{-i2\pi m x} \right) \left(\sum_\ell v_\ell e^{i2\pi \ell y} \right) e^{i2\pi kT(x-y)}\, dx\, dy. \end{aligned}$$

Upon setting

$$V(z) = \sum_\ell v_\ell e^{i2\pi \ell z}$$

this becomes

$$\begin{aligned} \mathbf{v}^H A \mathbf{v} &= T \sum_{k \in J} \int_{-1/2}^{1/2} V^*(x) e^{i2\pi kTx}\, dx \int_{-1/2}^{1/2} V(y) e^{-i2\pi kTy}\, dy \\ &= T \sum_{k \in J} \left| \int_{-1/2}^{1/2} V(x) e^{-i2\pi kTx}\, dx \right|^2 . \end{aligned}$$

This expression is obviously nonnegative, but this is not enough to establish the nonsingularity. A detailed study of the rank of the matrix in a more general context can be found in [34].

The preceding construction shows that a function can be made to oscillate at rates arbitrarily higher than any of its Fourier components, simply by enforcing finitely many interpolatory constraints at scale T, with T as small as desired. The method requires only sampling expansions. The nature of the superoscillating stretch can be controlled by varying the number N of constraints, their separation and amplitudes.

The superoscillating behaviour, at the finer scale T, is induced by the samples $f(k)$, $k = 0, 1, \ldots N-1$, which determine the behaviour of the function at scale 1. Surprisingly, the superoscillating stretch can be moved an arbitrary distance away from the interval $[0, N-1]$, where the scale 1 behaviour is determined.

1.5 Final remarks and conclusion

A number of methods to generate superoscillating signals were discussed. One of the simplest methods, in theory, is based on the manipulation of the zeros of a band-limited signal. It is indeed simple, but the superoscillating segment cannot be constrained to approximate or interpolate a given target waveform. Then methods based on linear equations were discussed. They seem to require straightforward computations, and allow any given prescribed set of data to be interpolated. However, the matter is not always that straightforward. The following remarks address a few remaining issues, among others which could also be listed.

1.5.1 Conditioning

The obvious problem with the approaches based on linear operators is the numerical conditioning of the problem. The matrix S can be very ill-conditioned, and thus the determination of the superoscillating signal by solving $Sx = a$ cannot be guaranteed in practice.

The condition number of S, denoted by $\kappa(S)$, 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. Let the data a be replaced with $a + \Delta a$. Denote the corresponding solution by $x + \Delta x$. For a given matrix norm, the condition number sets a bound on the relative error in x, with respect to that norm:

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

It can be shown that the condition number in the spectral norm satisfies [22]

$$\kappa[S(\mu,\delta)] \sim \min(\mu\delta N, 1)\, N^{-1/2} \left(\frac{4}{\pi\mu\delta}\right)^{2N-1}, \tag{1.14}$$

and thus the problem $Sx = a$ can become numerically very difficult. Nevertheless, the approach remains practical within certain bounds (and has been used in several papers given in the references to generate diagrams of superoscillating signals). As a rule of thumb, a condition number of 10^d generally means a loss of d significant figures. For a given precision and condition number, the numerical difficulties and even the computation limits can be determined ahead. Beethoven over 1 Hz is hopeless, but "more modest bandwidth compression" (to quote from [4]) can indeed be achieved.

1.5.2 Channel capacity

Superoscillations seem to allow the encoding of an arbitrarily large amount of information into an arbitrarily short segment of a low-bandwidth signal. The signal would then be able to pass through any channel of correspondingly narrow bandwidth. This is the essence of Berry's "Beethoven example" [4]. But, as Berry also pointed out, the amplitude of the conventional oscillations would have to be too large for practical purposes. Numerical conditioning, for constructions based on interpolation, is responsible for other difficulties, as discussed before. Information theory yields yet another way of looking at the problem.

Superoscillations are not incompatible with Shannon's capacity formula for a band-limited channel [40,41], which shows that a Gaussian channel of bandwidth B can convey information on average at most at a rate $B\log_2(1+S/N)$, where S/N is the signal-to-noise ratio. Thus, if the noise is very weak, the capacity can be large (even if the bandwidth is small).

In intuitive terms, if the number of samples per second allowed is very small but there is almost no noise, one can still encode many symbols in each sample. For example: one may associate a symbol with a sample amplitude a, another with $a+\delta$, another with $a+2\delta$ and so forth. If the noise is sufficiently weak (much smaller than δ) the amplitudes can always be distinguished.

In more rigorous terms, Shannon's formula allows data transmission at rates higher than the bandwidth if the signal power grows exponentially with the amount of compressed information. This agrees with the fact that the norm of maximally superoscillatory wave functions grows exponentially with the number of prescribed superoscillations.

1.5.3 Spectrum analysis

Consider a sampled segment of length T of a signal, band-limited to $\mu/2$ and containing N superoscillations of period δ. The samples of this superoscillating segment would certainly lead to estimation errors in the bandwidth of the signal, regardless of the spectrum analysis technique used. Theoretically, superoscillations show that it is impossible to infer the bandwidth of a finite energy signal f from a sampled segment of length T, regardless of how large T is. This is true because there are signals of arbitrarily small bandwidth that oscillate throughout an interval of length T with arbitrarily small period δ.

The error will be detected only if the observer is allowed to increase T at will. The amplitude of the signal will sooner or later increase rapidly as the limits of the superoscillating segment are reached. In a sense, superoscillations are created by cancellation, and the amplitude inside the superoscillating segment is kept small due to the cancellation. Outside of it conventional oscillations of much higher amplitude appear.

◊◊◊

From the signal processing point of view, superoscillations raise a number of interesting questions about the behaviour of band-limited signals. The claim originally made by Aharonov still seems, as Berry put it, "unbelievable, even paradoxical". Nevertheless, a function *can* be made to oscillate at rates arbitrarily higher than any of its Fourier components.

We mentioned some of the consequences and applications of superoscillations, concentrating on their energy cost, as function of the number of superoscillations, N, the period of the superoscillations, δ, and the bandwidth of the signal, μ. The energy increases exponentially with N, and polynomially with $1/\mu$ or $1/\delta$. Some of the implications of superoscillations were also discussed. The topic is relatively recent but has already given rise to a number of interesting applications and so it seems appropriate to give it some attention.

Bibliography

[1] 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.*, 64(25):2965–2968, June 1990.

[2] Y. Aharonov, F. Colombo, I. Sabadini, D. C. Struppa, and J. Tollaksen. Some mathematical properties of superoscillations. *J. Phys. A: Math. Theor.*, 44:365304 (16pp), 2011.

[3] J. J. Benedetto and P. J. S. G. Ferreira, editors. *Modern Sampling Theory: Mathematics and Applications.* Birkhäuser, Boston, 2001.

[4] M. Berry. Faster than Fourier. In J. S. Anandan and J. L. Safko, editors, *Quantum Coherence and Reality; in celebration of the 60th Birthday of Yakir Aharonov*, pages 55–65. World Scientific, Singapore, 1994.

[5] M. V. Berry. Evanescent and real waves in quantum billiards and Gaussian beams. *J. Phys. A: Math. Gen.*, 27(11):L391–L398, 1994.

[6] M. V. Berry. Quantum backflow, negative kinetic energy, and optical retro-propagation. *J. Phys. A: Math. Theor.*, 43:415302 (15pp), 2010.

[7] M. V. Berry and M. R. Dennis. Natural superoscillations in monochromatic waves in d dimensions. *J. Phys. A: Math. Theor.*, 42:022003 (8pp), 2009.

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

[9] F. E. Bond and C. R. Cahn. On sampling the zeros of bandwidth limited signals. *IRE Trans. Infor. Theory*, 4(3):110–113, Sept. 1958.

[10] P. L. Butzer. A survey of the Whittaker-Shannon sampling theorem and some of its extensions. *J. Math. Res. Expos.*, 3(1):185–212, Jan. 1983.

[11] P. L. Butzer, M. M. Dodson, P. J. S. G. Ferreira, J. R. Higgins, O. Lange, and P. Seidler. Herbert Raabe's work in multiplex signal transmission and his development of sampling methods. *Sig. Proc.*, 90(5):1436–1455, May 2010.

[12] P. L. Butzer, M. M. Dodson, P. J. S. G. Ferreira, J. R. Higgins, O. Lange, P. Seidler, and R. L. Stens. Multiplex signal transmission and the development of sampling techniques: the work of Herbert Raabe in contrast to that of Claude Shannon. *Appl. Analysis*, 90(3):643–688, Mar. 2011.

[13] P. L. Butzer, P. J. S. G. Ferreira, J. R. Higgins, S. Saitoh, G. Schmeisser, and R. L. Stens. Interpolation and sampling: E. T. Whittaker, K. Ogura and their followers. *J. Fourier Anal. Appl.*, 17(2):320–354, Apr. 2011.

[14] P. L. Butzer and R. L. Stens. Sampling theory for not necessarily band-limited functions: a historical overview. *SIAM Rev.*, 34(1):40–53, Mar. 1992.

[15] R. Courant and D. Hilbert. *Methods of Mathematical Physics*, volume 1. Interscience Publishers, Inc., New York, 1953. Reprinted: John Wiley & Sons, Wiley Classics Library Edition, 1989.

[16] M. R. Dennis and J. Lindberg. Natural superoscillation of random functions in one and more dimensions. In *Proceedings of SPIE — The International Society for Optical Engineering*, volume 7394, page 73940A (9pp), 2009.

[17] V. Eisler and I. Peschel. Free-fermion entanglement and spheroidal functions. *Journal of Statistical Mechanics: Theory and Experiment*, 2013(4):P04028, Apr. 2013.

[18] P. J. S. G. Ferreira. Incomplete sampling series and the recovery of missing samples from oversampled band-limited signals. *IEEE Trans. Signal Processing*, 40(1):225–227, Jan. 1992.

[19] P. J. S. G. Ferreira. The stability of a procedure for the recovery of lost samples in band-limited signals. *Sig. Proc.*, 40(3):195–205, Dec. 1994.

[20] P. J. S. G. Ferreira and R. Higgins. The establishment of sampling as a scientific principle — a striking case of multiple discovery. *Notices of the AMS*, 58(10):1446–1450, Nov. 2011.

[21] P. J. S. G. Ferreira and A. Kempf. The energy expense of superoscillations. In *Signal Processing XI — Theories and Applications. Proceedings of EUSIPCO-2002, XI European Signal Processing Conference*, volume II, pages 347–350, Toulouse, France, Sept. 2002.

[22] P. J. S. G. Ferreira and A. Kempf. Superoscillations: Faster than the Nyquist rate. *IEEE Trans. Signal Processing*, 54(10):3732–3740, Oct. 2006.

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

[24] J. R. Higgins. *Sampling Theory in Fourier and Signal Analysis: Foundations.* Oxford University Press, Oxford, 1996.

[25] 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.*, 9:S285–S288, 2007.

[26] F. M. Huang and N. I. Zheludev. Super-resolution without evanescent waves. *Nano Letters*, 9(3):1249–1254, 2009.

[27] E. Katzav and M. Schwartz. Yield–optimized superoscillations. *IEEE Trans. Signal Processing*, 61(12):3113–3118, 2013.

[28] A. Kempf. Black holes, bandwidths and Beethoven. *J. Math. Physics*, 41(4):2360–2374, Apr. 2000.

[29] A. Kempf and P. J. S. G. Ferreira. Unusual properties of superoscillating particles. *J. Phys. A: Math. Gen.*, 37:12067–12076, 2004.

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

[31] H. J. Landau. Necessary density conditions for sampling and interpolation of certain entire functions. *Acta Math.*, 117:37–52, 1967.

[32] H. J. Landau. Sampling, data transmission, and the Nyquist rate. *Proc. IEEE*, 55(10):1701–1706, Oct. 1967.

[33] H. J. Landau and H. O. Pollak. Prolate spheroidal wave functions, Fourier analysis and uncertainty – II. *Bell Syst. Tech. J.*, 40:65–84, Jan. 1961.

[34] D. G. Lee and P. J. S. G. Ferreira. Direct construction of superoscillations. Preprint, 2014.

[35] L. Levi. Fitting a bandlimited signal to given points. *IEEE Trans. Inform. Theory*, pages 372–376, July 1965.

[36] I. Ostrovskii and A. Ulanovskii. Non-oscillating Paley-Wiener functions. *C. R. Acad. Sci., Paris, Sér. I*, 333:735–740, 2001.

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

[38] S. O. Rice. Mathematical analysis of random noise. *Bell Syst. Tech. J.*, 23:282–332, July 1944.

[39] S. O. Rice. Mathematical analysis of random noise. *Bell Syst. Tech. J.*, 24:46–156, Jan. 1945.

[40] C. E. Shannon. A mathematical theory of communication. *Bell Syst. Tech. J.*, 27:379–423, 623–656, July, October 1948.

[41] C. E. Shannon. Communication in the presence of noise. *Proc. IRE*, 37:10–21, Jan. 1949.

[42] D. Slepian. Prolate spheroidal wave functions, Fourier analysis and uncertainty — V: The discrete case. *Bell Syst. Tech. J.*, 57(5):1371–1430, May 1978.

[43] D. Slepian and H. O. Pollak. Prolate spheroidal wave functions, Fourier analysis and uncertainty — I. *Bell Syst. Tech. J.*, 40(1):43–63, Jan. 1961.

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

[45] H. B. Voelcker and A. A. Requicha. Clipping and signal determinism: Two algorithms requiring validation. *IEEE Trans. Commun.*, 21:738–744, June 1973.

[46] W. J. Walker. Almost periodic functions with bounded Fourier exponents. *J. Math. Anal. Appl.*, 162:537–541, 1991.

[47] W. J. Walker. Zeros of the Fourier transform of a distribution. *J. Math. Anal. Appl.*, 154:77–79, 1991.

[48] 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. Letters*, 9:315–318, 2010.

[49] 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 *2010 IEEE Antennas and Propagation Society International Workshop (APSURSI)*, pages 1–4, Toronto, Canada, July 2010.

[50] 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.*, 59(12):4766–4776, Dec. 2011.

[51] A. M. H. Wong and G. V. Eleftheriades. Temporal pulse compression beyond the Fourier transform limit. *IEEE Trans. Microwave Theory Tech.*, 59(9):2173–2179, Sept. 2011.

[52] N. I. Zheludev. What diffraction limit? *Nature Mater.*, 7:420–422, June 2008.