Measures of spread for periodic distributions and the associated uncertainty relations

G. W. Forbes[a] and M. A. Alonso[b]
Physics Department, Macquarie University, Sydney, NSW 2109, Australia

(Received 22 March 2000; accepted 8 August 2000)

A physicist's intuition in Fourier theory is generally established from the parallels between Fourier series and transforms. Remarkably, one element of this theory that is especially significant in physics, namely the uncertainty principle, is never treated for Fourier series. We resolve this by first showing that a natural measure of spread for a periodic distribution follows simply upon regarding the distribution as a mass density on a ring. Even though the centroid of this ring is expressed in terms of just a first moment, its distance from the geometric center gives a close analog of variance. We then derive direct analogs of the uncertainty principle for both the Fourier series of a continuous periodic function as well as the fast Fourier transform of discrete data. The results have similar applications to those of the standard uncertainty principle.
[DOI: 10.1119/1.1317562]

I. INTRODUCTION

Fourier theory plays a central role in any problem that is both linear and shift invariant[1]—or even only approximately so—and this includes many wave-related phenomena. It is therefore no surprise that Fourier methods are at the foundations of the modeling of sound and light, for example, as well as the related technological devices. As a result, the notion of frequency decomposition is widespread. One of the standard properties of this decomposition is that, as a signal becomes localized, its spectrum broadens. The *uncertainty principle* extends this by stating that there is, in fact, a minimum width to the spectrum that is inversely proportional to the localization of the signal.[2] There are many variations on this result,[3] but the standard form is stated in terms of the product of the second moments in coordinate and frequency space.

Although the uncertainty principle acquired its name within the context of quantum mechanics, it has wide-ranging significance. For example, it gives a resolution limit for the determination of a signal's spectrum in terms of the duration of the measurement. It also gives resolution limits for imaging systems in terms of their aperture size or, equivalently, fixes the minimum possible spread of a beam from the size of its source. In fact, it serves as the basis of one of the standard measures of laser beam quality.[4] In quantum mechanics, of course, its significance is more profound and, like Schrödinger's wave function itself, its interpretation and consequences are still widely studied and disputed. However, issues such as quantum measurement can obscure the mathematical essence of the limits to simultaneous localization in coordinate and frequency space. Similarly, the baggage of quantized fields makes the related work on phase in quantum optics[5] effectively inaccessible for general applications.

A. The standard uncertainty principle

The numerical factors in the familiar uncertainty principle depend on the form that is adopted for the Fourier transform. When the transform is taken to be defined by

$$F(p) := \frac{1}{\sqrt{2\pi}} \int f(x) e^{-ipx}\, dx, \tag{1.1}$$

its inverse is given by

$$f(x) = \frac{1}{\sqrt{2\pi}} \int F(p) e^{ipx}\, dp, \tag{1.2}$$

where integrals are to be taken over all space unless otherwise specified. The Fourier transform of a product, say $(2\pi)^{-1/2}\int f(x)h(x)\exp(-ip'x)dx$, then follows simply upon replacing $h(x)$ in this expression by using Eq. (1.2) and reversing the order of integration:

$$\frac{1}{\sqrt{2\pi}} \int f(x)h(x)e^{-ip'x}\, dx = \frac{1}{\sqrt{2\pi}} \int F(p'-p)H(p)\,dp. \tag{1.3}$$

In particular, with $p'=0$ and introducing $g := f^*$ [so $F(p) = G^*(-p)$], Eq. (1.3) becomes

$$\int g^*(x)h(x)\,dx = \int G^*(p)H(p)\,dp. \tag{1.4}$$

Equations (1.3) and (1.4) give relations between operations in coordinate and frequency space that are central to matters like the uncertainty principle.

Centered second moments give the standard measures of spread, i.e.,

$$\Delta_x^2 := \frac{1}{N} \int (x-\overline{x})^2 |f(x)|^2\, dx, \tag{1.5a}$$

$$\Delta_p^2 := \frac{1}{N} \int (p-\overline{p})^2 |F(p)|^2\, dp, \tag{1.5b}$$

where

$$\overline{x} := \frac{1}{N} \int x|f(x)|^2\, dx, \tag{1.6a}$$

$$\overline{p} := \frac{1}{N} \int p|F(p)|^2\, dp, \tag{1.6b}$$

with [according to Eq. (1.4) with $g=h=f$]

$$N:=\int |f(x)|^2\,dx=\int |F(p)|^2\,dp. \tag{1.7}$$

The uncertainty principle is usually derived from the Cauchy–Schwarz inequality,[6] i.e.,

$$\int |g(x)|^2\,dx\int |h(x)|^2\,dx\geqslant\left|\int g^*(x)h(x)dx\right|^2, \tag{1.8}$$

by choosing

$$g(x)=(a-x)f(x)/\sqrt{N}, \tag{1.9a}$$

$$h(x)=\left(ib-\frac{d}{dx}\right)f(x)/\sqrt{N}, \tag{1.9b}$$

where a and b are constants. There are a number of steps in getting to the standard form, however, and it is helpful to review them.

Since the Fourier transform of df/dx is readily found via integration by parts to be $ipF(p)$, it follows from Eq. (1.4) that the second factor on the left-hand side of Eq. (1.8) can be written as

$$\frac{1}{N}\int\left|\left(ib-\frac{d}{dx}\right)f(x)\right|^2 dx=\frac{1}{N}\int |p-b|^2|F(p)|^2\,dp. \tag{1.10}$$

It is now natural to choose $a=\overline{x}$ and $b=\overline{p}$ since this not only gives the connection to Δ_x and Δ_p, but it also individually minimizes each of the two factors on the left-hand side of Eq. (1.8)—thereby generally strengthening the result. With this choice, the integral on the right-hand side can also be rewritten by using Eq. (1.4) as[7]

$$\begin{aligned}\frac{1}{N}&\int (x-\overline{x})(f'-i\overline{p}f)f^*\,dx\\&=\frac{1}{N}\int xf^*f'\,dx-i\overline{xp}\\&=-\frac{1}{2}+i\left[\frac{1}{N}\int x\,\mathrm{Re}(-if^*f')dx-\overline{xp}\right].\end{aligned} \tag{1.11}$$

It now follows from Eqs. (1.8) to (1.11) that

$$\Delta_x^2\Delta_p^2-\left[\frac{1}{N}\int x\,\mathrm{Re}(-if^*f')dx-\overline{xp}\right]^2\geqslant\frac{1}{4}. \tag{1.12}$$

A weaker relation follows when the second term on the left-hand side of Eq. (1.12) is dropped, and this gives the standard uncertainty principle:

$$\Delta_x\Delta_p\geqslant\tfrac{1}{2}. \tag{1.13}$$

Outside the context of quantum mechanics, this is no longer a physical principle but simply a mathematical relation. Equation (1.13) simply states that there is a lower bound on the spectral width that is inversely proportional to the localization of the signal. Because translation of f (or F) just puts a linear phase on F (or f, respectively), it is clear from the definitions that the value of the uncertainty product, i.e., $\Delta_x\Delta_p$, is unchanged by translation of, or by the incorporation of a linear phase on, either f or F. Such notions of invariance are fundamental to the ideas developed below. Also notice that Eq. (1.8) is an equality only when $g(x)\equiv\mu h(x)$ for some constant μ. From Eq. (1.9), it can then be shown that the only minimum uncertainty distributions have a Gaussian form.

B. Variants of the standard form

The extra term on the left-hand side of Eq. (1.12) gives invariance under the incorporation of a quadratic phase on either f or F. (Such a phase is acquired by F under propagation in time of a Schrödinger wave function in free space, or under propagation in space of a classical wave field in the Fresnel approximation.) This added invariance has a clearer significance when the relation is considered in terms of phase space densities such as the Wigner distribution. The uncertainty relation can then be taken as a statement of maximal localization in phase space. A quadratic phase simply shears phase space and can thereby change Δ_x or Δ_p, but it does not modify the value of the expression in Eq. (1.12) because of the extra covariance term.[8] Although he does not consider such shearing, de Bruijn[9] takes an entirely different approach to the derivation given above and goes on to develop remarkable variants and generalizations of Eq. (1.13). [He shows, for example, that the lower bound given in Eq. (1.13) can be increased by a factor of 3 if f is odd, and he also derives analogous relations for higher order moments.] Phase space is central to his discussion.

As documented in Ref. 3, there are other generalizations and cousins of Eq. (1.13). There are also analogous extensions for the fractional Fourier transform.[10] Slepian[11] developed a framework that is based on the idea of measuring concentration within subintervals. Another important relative[12,13] is based on an entropic measure of ''bunching'' rather than a measure of spread like those in Eq. (1.5). (For a double-lobed distribution, for example, the entropic measure is unchanged if the lobes are moved apart—or even subdivided and rearranged arbitrarily—so it does not measure localization in the usual sense.) Nevertheless, the application to discretely sampled signals is also considered in both of these references. Since the discrete Fourier transform[14] (DFT) is effectively periodic, this touches upon the obvious idea of finding an analog of Eq. (1.12) for periodic functions. Such a result would not only be applicable for DFTs, but would also have clear significance for processes such as classical phase measurement. In fact, as documented in Ref. 5, the search for an understanding of phase at a quantum level has led to a range of relevant ideas that followed upon the trouble-ridden early quantum treatment of angle variables.[15] However, these two areas of work are weighed down by the extra challenges that are associated with multidimensional quantum mechanics and quantized fields.

The challenge in developing an analog for Eq. (1.12) for a periodic function lies in finding an appropriate measure of spread. For $f(\theta)$, where $f(\theta+2\pi)\equiv f(\theta)$, the simplest analog is based on considering $\int(\theta-\theta_0)^2|f(\theta)|^2d\theta$ where the integral is taken over an interval of length 2π. While such an extension has been considered in some detail,[16] the limits of integration are problematic: This measure of spread depends artificially on the end points because $(\theta-\theta_0)^2$ is not periodic. Even the natural choice of $(\theta_0-\pi,\theta_0+\pi)$ gives an awkward measure where the contribution from two lobes in the distribution can depend strongly on the choice of θ_0. What is more, even if a prescription is developed to choose θ_0 so that the results are invariant under translation, this approach is not amenable to an analytic treatment. More important, the final measure of spread does not treat all points

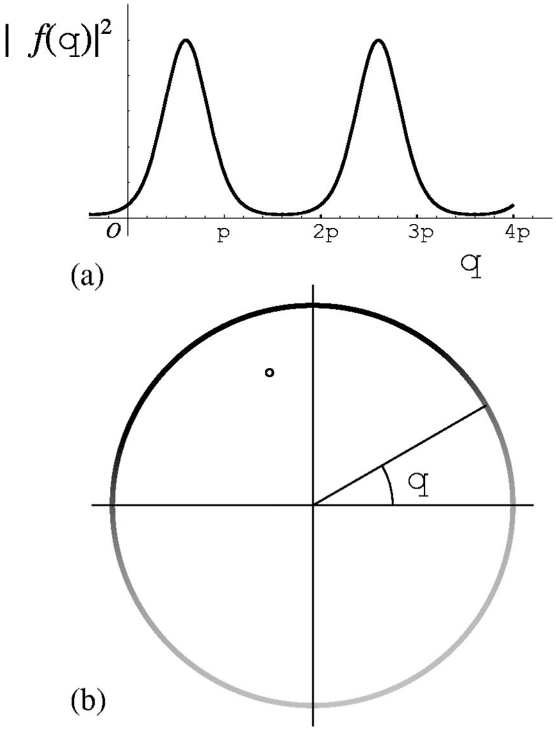

Fig. 1. Two representations of the same periodic distribution. The distribution is shown as a conventional plot in (a) while grayscale is used in (b) to denote the density on a ring as a function of angle. The ring's centroid is shown as a hollow circle in the second quadrant in (b). The angle to the centroid is evidently a natural measure of the mean of the distribution. More importantly, the centroid's distance from the origin is found to give a simple measure of the degree of localization in the distribution. Neither of these entities has natural counterparts when, as has often been the case, (a) is used to guide the developments.

equally: The circle is effectively cut open and regarded as an interval. A more even-handed variation has also been considered recently,[17] but the connection to Eq. (1.13) is not at all straightforward.

A convenient and intuitive analog of variance is discussed in Sec. II, and considered for both continuous and discrete periodic functions. While this alone is enough to develop the analog of Eq. (1.12) for the DFT, the case of a Fourier series is treated first in Sec. III. Since the Fourier coefficients are not themselves periodic, the more conventional measure of spread is used in frequency space for this case. The manner in which the resulting uncertainty relation approaches Eq. (1.12) when $f(\theta)$ becomes localized is also treated. The case of the DFT is considered in Sec. IV where the process of sampling a continuous distribution is used to establish a clear connection to the familiar form given in Eq. (1.13).

II. MEASURING THE SPREAD OF A PERIODIC FUNCTION

One option for measuring the degree of localization of $|f(\theta)|^2$ is to consider it as a distribution around the rim of the unit disk, and then to calculate the centroid of this ring. See Fig. 1. Although this is just a first moment—and it evidently gives the mean position of the distribution—it simultaneously determines the degree of localization: Whenever the distribution is localized, the centroid is near the edge.[18] This can be appreciated by considering the mean of $[\mathbf{u}(\theta)-\overline{\mathbf{u}}]^2$ where

$$\mathbf{u}(\theta):=(\cos\theta,\sin\theta), \tag{2.1}$$

$$\overline{\mathbf{u}}:=\frac{1}{N}\int \mathbf{u}(\theta)|f(\theta)|^2\,d\theta, \tag{2.2a}$$

$$N:=\int |f(\theta)|^2\,d\theta. \tag{2.2b}$$

Here, and throughout this work, any integral over a periodic function is taken over a full period. The measure of spread proposed here can now be written as

$$\Delta_\theta^2:=\frac{1}{N}\int [\mathbf{u}(\theta)-\overline{\mathbf{u}}]^2|f(\theta)|^2\,d\theta=1-\overline{\mathbf{u}}^2. \tag{2.3}$$

That is, in this case, $\overline{\mathbf{u}^2}$ is unity and it follows that $\overline{\mathbf{u}}$ gives both the mean of the distribution as well as a measure of its localization. Notice also that $0<\Delta_\theta<1$. While it may appear to be natural to consider using the complex plane so that these ideas can be presented in terms of scalars, this vectorial approach is more convenient for the derivation in Sec. III. Some general properties of Δ_θ are treated first.

A. Example and limit of high localization for continuous case

The character of Δ_θ can be appreciated in part through examples. For instance, consider the rectangular distribution defined for $\theta\in(-\pi,\pi)$ by

$$f(\theta)=\begin{cases}1, & |\theta|<\alpha\\ 0, & |\theta|>\alpha\end{cases}, \tag{2.4}$$

where $0<\alpha<\pi$. In this case, $N=2\alpha$ and $\overline{\mathbf{u}}=(\sin\alpha/\alpha,0)$, so

$$\Delta_\theta^2=1-(\sin\alpha/\alpha)^2. \tag{2.5}$$

Notice that, for $\alpha=\pi$, $\Delta_\theta=1$. The limit of a localized distribution is considered in the following sections to establish a connection with the standard results presented in Sec. I. For this particular example, as α becomes small, it is found from Eq. (2.5) that

$$\Delta_\theta=\alpha/\sqrt{3}+O(\alpha^3), \tag{2.6}$$

and this matches the usual measure of spread for the corresponding rectangular nonperiodic distribution on a line.

More generally, the limit of a localized distribution can be considered in terms of $f(\theta)=g(\gamma\theta)$ where $\theta\in(-\pi,\pi)$ and g is a square-summable function on the real line. With $N_g:=\int|g(x)|^2\,dx$ and $\overline{x^n}:=N_g^{-1}\int x^n|g(x)|^2\,dx$, the usual measure of the spread of g is just $\Delta_x^2=(\overline{x^2}-\overline{x}^2)$. For large γ, $f(\theta)$ becomes localized about $\theta=0$, Eq. (2.2b) gives $N\approx N_g/|\gamma|$, and Eq. (2.2a) becomes

$$\overline{\mathbf{u}}\approx[1-\overline{x^2}/(2\gamma^2),\overline{x}/\gamma]. \tag{2.7}$$

It now follows from Eqs. (2.3) and (2.7) that

$$\Delta_\theta^2\approx(\overline{x^2}-\overline{x}^2)/\gamma^2, \tag{2.8}$$

so $\Delta_\theta\approx\Delta_x/\gamma$, and this is the general case of Eq. (2.6). That is, when its value is small compared to unity, Δ_θ evidently gives the familiar measure of spread in units of radians. When it approaches unity, however, the uncertainty in the localization is of the order of a full period.

B. Discrete or sampled functions

Oftentimes Fourier-related computations can be performed efficiently by using a fast Fourier transform (FFT) which is just an effective means to evaluate a DFT.[19] In such cases the original data are sometimes a set of samples of a function of a continuous variable at uniformly distributed values, say $a_m := f[(2\pi/M)m]$ for $m=0,1,\ldots,M-1$. On the other hand, in problems where the original density is truly discrete from the outset, it is natural to regard the distribution as an array of delta functions so that

$$\overline{\mathbf{u}} := \frac{1}{N}\sum_m \mathbf{u}\left(\frac{2\pi}{M}m\right)|a_m|^2, \tag{2.9a}$$

$$N := \sum_m |a_m|^2, \tag{2.9b}$$

where all sums are over 0 to $M-1$ unless indicated otherwise. Again, $\Delta_\theta^2 = 1-\overline{\mathbf{u}}^2$. With this, Δ_θ is zero whenever a_m vanishes at all but one value of m. When the data results from sampling, however, this may not be a realistic measure of the uncertainty. An alternative is to interpolate between the sample values and measure Δ_θ for the associated continuous function. The case of linear interpolation is outlined in the Appendix where it is shown that Δ_θ then has a more intuitive lower bound. In particular, the uncertainty is then never less than about one-third of the sample spacing.

III. UNCERTAINTY RELATION FOR FOURIER SERIES

We first seek an analog of Eq. (1.12) for the case of a Fourier series. That is, $f(\theta)$ is written as

$$f(\theta) = \frac{1}{\sqrt{2\pi}}\sum_{m=-\infty}^{\infty} F_m e^{im\theta}, \tag{3.1}$$

where

$$F_m := \frac{1}{\sqrt{2\pi}}\int f(\theta)e^{-im\theta}\,d\theta. \tag{3.2}$$

In this case, while the original signal is a periodic function of a continuous variable, the conjugate representation is not; it is an (inherently discrete and ordered) infinite set of coefficients. While the spread in f is to be measured by using Eq. (2.3), the spread in the conjugate representation is naturally measured by

$$\Delta_m^2 := \frac{1}{N}\sum_{m=-\infty}^{\infty}(m-\overline{m})^2|F_m|^2, \tag{3.3}$$

where

$$\overline{m} := \frac{1}{N}\sum_{m=-\infty}^{\infty} m|F_m|^2, \tag{3.4a}$$

$$N := \sum_{m=-\infty}^{\infty}|F_m|^2 = \int |f(\theta)|^2\,d\theta. \tag{3.4b}$$

Notice that the second relation in Eq. (3.4b) follows easily upon deriving the analog of Eq. (1.4) in just the same manner:

$$\int f^*(\theta)g(\theta)\,d\theta = \sum_{m=-\infty}^{\infty} F_m^* G_m. \tag{3.5}$$

(The factors of $\sqrt{2\pi}$ in the definition of the Fourier series given above allow these relations to avoid a factor of 2π.) The goal therefore is to find an analog of Eq. (1.12) for Δ_θ and Δ_m.

One option is to start once again with the Cauchy–Schwarz inequality, but now in the form[20]

$$\int |\mathbf{v}(\theta)|^2\,d\theta \int |\mathbf{w}(\theta)|^2\,d\theta \geqslant \left|\int \mathbf{v}^*(\theta)\cdot\mathbf{w}(\theta)\,d\theta\right|^2. \tag{3.6}$$

In this case, in keeping with Eq. (1.9), we take

$$\mathbf{v}(\theta) = [\overline{\mathbf{u}} - \mathbf{u}(\theta)]f(\theta)/\sqrt{N}, \tag{3.7a}$$

$$\mathbf{w}(\theta) = \mathbf{w}_0\left(i\overline{m} - \frac{d}{d\theta}\right)f(\theta)/\sqrt{N}, \tag{3.7b}$$

where $\mathbf{w}_0$ is a constant vector that is yet to be fixed. The first factor on the left-hand side of Eq. (3.6) is now just Δ_θ^2. By using Eq. (3.1), the second factor can be written as

$$\frac{|\mathbf{w}_0|^2}{N}\int \left|\left(-i\frac{d}{d\theta} - \overline{m}\right)f(\theta)\right|^2 d\theta = \frac{|\mathbf{w}_0|^2}{N}\frac{1}{2\pi}\int h^*(\theta)h(\theta)\,d\theta, \tag{3.8}$$

where

$$h(\theta) = \sum_{m=-\infty}^{\infty}(m-\overline{m})F_m e^{im\theta}. \tag{3.9}$$

When Eq. (3.9) is used to eliminate h from Eq. (3.8) and the integral is taken inside the sums, it is found that $\int|\mathbf{w}(\theta)|^2\,d\theta = |\mathbf{w}_0|^2\Delta_m^2$.

The integral on the right-hand side of Eq. (3.6) can be transformed much like Eq. (1.11):

$$\begin{aligned}\mathbf{w}_0\cdot\frac{1}{N}\int(\mathbf{u}-\overline{\mathbf{u}})(f' - i\overline{m}f)f^*\,d\theta &= \mathbf{w}_0\cdot\left(\frac{1}{N}\int \mathbf{u}f^*f'\,d\theta - i\overline{m}\overline{\mathbf{u}}\right)\\ &= \mathbf{w}_0\cdot\left[-\frac{1}{2}\overline{\mathbf{u}'} + i\left(\frac{1}{N}\int \mathbf{u}\,\mathrm{Re}(-if^*f')\,d\theta - \overline{m}\overline{\mathbf{u}}\right)\right],\end{aligned} \tag{3.10}$$

where

$$\overline{\mathbf{u}'} = \frac{1}{N}\int \mathbf{u}'(\theta)|f(\theta)|^2\,d\theta.$$

Notice that $\overline{\mathbf{u}'}^2 = \overline{\mathbf{u}}^2$ and $\overline{\mathbf{u}'}\cdot\overline{\mathbf{u}} = 0$. When $\overline{\mathbf{u}}$ is nonzero it is convenient to choose $\mathbf{w}_0$ to be $\overline{\mathbf{u}'}$ and, with this choice, Eq. (3.6) gives

$$\Delta_\theta^2\Delta_m^2 \geqslant \frac{1}{4}\overline{\mathbf{u}}^2 + \frac{1}{\overline{\mathbf{u}}^2}\left(\overline{\mathbf{u}'}\cdot\frac{1}{N}\int \overline{\mathbf{u}}\,\mathrm{Re}(if^*f')\,d\theta\right)^2. \tag{3.11}$$

It is the straightforward separation into real and imaginary parts in Eq. (3.10) [to arrive at Eq. (3.11)] that justifies the use of vectors here in place of the more attractive complex form, i.e., $\mathbf{u}(\theta)$ in place of $e^{i\theta}$. Equation (3.11) can be put in a more suggestive form by using $\overline{\mathbf{u}}^2=1-\Delta_\theta^2$ to replace the first term on the right-hand side to find

$$\Delta_\theta^2\left(\Delta_m^2+\frac{1}{4}\right)\geqslant\frac{1}{4}+\frac{1}{\overline{\mathbf{u}}^2}\left\{\overline{\mathbf{u}'}\cdot\frac{1}{N}\int\mathbf{u}(\theta)\right.$$
$$\left.\times\mathrm{Re}[-if^*(\theta)f'(\theta)]d\theta\right\}^2, \tag{3.12}$$

and the analog of Eq. (1.13) is now just

$$\Delta_\theta^2(\Delta_m^2+\tfrac{1}{4})\geqslant\tfrac{1}{4}. \tag{3.13}$$

Equation (3.13) establishes that if Δ_θ becomes small, Δ_m must be large. (Of course, unity sets the scale for these dimensionless quantities, and recall that Δ_θ cannot exceed unity.) Conversely, if Δ_m becomes small, Δ_θ must approach unity, hence the distribution must be widely spread in θ. This is precisely the analog of Eq. (1.13) that was sought. In fact, these two relations approach one another when $f(\theta)$ becomes highly localized: Just as in Eq. (2.8), with $f(\theta)=g(\gamma\theta)$ where $\theta\in(-\pi,\pi)$ and g is characterized by Δ_x and Δ_p, it can shown that $\Delta_m\approx\gamma\Delta_p$ for large γ. (Notice that, as required for $\overline{m^2}$ to be well defined, f effectively becomes continuous at π, as γ increases.) Taken together with Eq. (2.8), it follows that Eq. (3.13) approaches Eq. (1.13) in this limit.

The extra term in Eq. (3.12) can be shown to vanish whenever $\mathrm{Arg}[f(\theta)]$ is linear in θ. In dealing with this stronger form, it is helpful to express the entities in Eq. (3.12) in terms of the Fourier coefficients. It turns out that the results involve

$$S:=\frac{1}{N}\sum_{m=-\infty}^{\infty}F_m^*F_{m-1}, \tag{3.14a}$$

$$T:=\frac{1}{N}\sum_{m=-\infty}^{\infty}mF_m^*F_{m-1}. \tag{3.14b}$$

For example, by using Eq. (3.1) in Eq. (2.2) it is found that

$$\overline{\mathbf{u}}=[\mathrm{Re}(S),\mathrm{Im}(S)]. \tag{3.15}$$

By similarly reworking the extra term in Eq. (3.12), this equation can be put in an alternative form:

$$\Delta_\theta^2(\Delta_m^2+\tfrac{1}{4}\geqslant\tfrac{1}{4}+[\mathrm{Im}(S^*T)/|S|]^2$$
$$=\tfrac{1}{4}+|T|^2\sin^2[\mathrm{Arg}(S)-\mathrm{Arg}(T)]. \tag{3.16}$$

When $\overline{\mathbf{u}}$ is zero, $\mathrm{Arg}(S)$ is undefined and Eq. (3.16) is itself then ill-defined [although its problems are less glaring than those of Eq. (3.12)]. In this case, however, if $\mathbf{w}_0$ is instead chosen to be $(1/N)\int\mathbf{u}\,\mathrm{Re}(-if^*f')d\theta$ in Eq. (3.10), it is found that

$$\Delta_m^2\geqslant\left\{\frac{1}{N}\int\mathbf{u}(\theta)\mathrm{Re}[-if^*(\theta)f'(\theta)]d\theta\right\}^2=|T|^2. \tag{3.17}$$

That is, in this apparently troublesome limit (where $\Delta_\theta=1$), Eq. (3.16) holds regardless of the value adopted for $\mathrm{Arg}(S)$.

From the alternative forms developed in the previous paragraph, it follows that (i) the second term on the right-hand side of Eq. (3.12) does not diverge as $\overline{\mathbf{u}}$ goes to zero, and (ii) Eq. (3.17) can be used in place of Eq. (3.12) in this limit. [A uniform blend of these results could presumably be found by adopting a more general choice for $\mathbf{w}_0$, or even for $\mathbf{w}(\theta)$, but our main goal was Eq. (3.13).] Also, notice that equality is attained in Eq. (3.6) only when $\mathbf{v}(\theta)\equiv\lambda\mathbf{w}(\theta)$ for some constant λ. This condition means that $\overline{\mathbf{u}}\cdot\mathbf{v}(\theta)\equiv\lambda\overline{\mathbf{u}}\cdot\mathbf{w}(\theta)$ and, given the choices of $\mathbf{v}(\theta)$ and $\mathbf{w}(\theta)$ considered above [see Eq. (3.7) and recall $\mathbf{w}_0=\overline{\mathbf{u}'}$], this can be satisfied only when $\overline{\mathbf{u}}$ is zero. By now considering the condition $\mathbf{u}'(\theta)\cdot\mathbf{v}(\theta)\equiv\lambda\mathbf{u}'\theta)\cdot\mathbf{w}(\theta)$ (with the form adopted above for $\mathbf{w}_0$ when $\overline{\mathbf{u}}$ is zero), it can be seen that $f'(\theta)=i\overline{m}f(\theta)$. Because $f(\theta)$ is required to be continuous (otherwise Δ_m is infinite), it follows that $f_n(\theta)=ce^{in\theta}$ for $n=0,\pm1,\pm2,\ldots$ are the only minimum uncertainty distributions. Notice that Eq. (3.13) is also an equality for these distributions because $\Delta_m=0$ and $\Delta_\theta=1$.

IV. UNCERTAINTY RELATION FOR THE DFT

The DFT maps $\{a_m,m=0,1,\ldots,M-1\}$ to $\{A_m,m=0,1,\ldots,M-1\}$ by

$$A_n:=\frac{1}{\sqrt{M}}\sum_m a_m\exp\left(-i\frac{2\pi}{M}nm\right), \tag{4.1}$$

with an inverse given by

$$a_n=\frac{1}{\sqrt{M}}\sum_m A_m\exp\left(i\frac{2\pi}{M}nm\right) \tag{4.2}$$

(recall the summation convention introduced in Sec. II B). The analog of Eqs. (1.4) and (3.5) for this case is easily found by a similar path:

$$\sum_m a_m^*b_m=\sum_m A_m^*B_m, \tag{4.3}$$

so, with $b_m=a_m$, the normalization is seen to be the same in both spaces. If measures of spread like those discussed in Sec. II B for discrete data are applied, we can once again work from the Cauchy–Schwarz inequality to arrive at an uncertainty relation.

In contrast to the case in Sec. III, the same measure of spread is applicable now to both the original function and its Fourier conjugate (so the symmetry of the standard case discussed in Sec. I has returned). Because the independent variable is now an index that is given the same name in both spaces, it is convenient to distinguish the spreads of the data set and of its conjugate by writing them as Δ_a and Δ_A, respectively. That is, for example,

$$\Delta_a^2:=1-\overline{\mathbf{u}}_a^2, \tag{4.4}$$

where $\overline{\mathbf{u}}_a$ can be taken to be defined by either Eq. (2.9a) or by Eq. (A4)—or any of the analogs for higher-order interpolants. Only the first, simplest option is considered in detail here. Since the spacing between the samples when they are arrayed around the unit disk is $2\pi/M$ radians, this sets the scale for converting either Δ_a or Δ_A to sample spaces for cases of localized distributions.

It is evident that $0\leqslant\Delta_a\Delta_A\leqslant1$ and, with the simpler measure associated with Eq. (2.9), both bounds can be attained. Nevertheless, it is natural to seek the analog of Eq. (1.12) [even though the right-hand side must be zero in the analog of Eq. (1.13)]. The analysis begins with the appropriate form of Cauchy–Schwarz, namely

$$\left(\sum_m |\mathbf{v}_m|^2\right)\left(\sum_m |\mathbf{w}_m|^2\right) \geqslant \left|\sum_m \mathbf{v}_m^* \cdot \mathbf{w}_m\right|^2, \tag{4.5}$$

and is based on the simple result [derived in the same way as Eq. (4.3)] that

$$S_a := \frac{1}{N}\sum_m a_m^* a_m \exp\left(i\frac{2\pi}{M}m\right) = \frac{1}{N}\sum_m A_m^* A_{m-1}. \tag{4.6}$$

It follows that, just as in Eq. (3.15),

$$\overline{\mathbf{u}}_a = [\mathrm{Re}(S_a), \mathrm{Im}(S_a)] \tag{4.7}$$

and then $\Delta_a^2 = 1 - |S_a|^2$.

To work in parallel to the derivations given above, an analog to the derivative in Eq. (1.9b) or (3.7b) must be found. Equation (4.6) essentially gives the answer: an index shifting operator. If we define two operators, say δ_+ and δ_-, by

$$\delta_\pm a_m := (a_{m+1} \pm a_{m-1})/2, \tag{4.8}$$

then the two sequences with coefficients $b_m^{(\pm)} := \delta_\pm a_m$ have DFTs of the form

$$B_m^{(+)} = \cos\left(\frac{2\pi}{M}m\right)A_m, \quad B_m^{(-)} = i\sin\left(\frac{2\pi}{M}m\right)A_m. \tag{4.9}$$

It now follows from Eq. (4.3) that

$$\frac{1}{N}\sum_m A_m^* A_m \mathbf{u}\left(\frac{2\pi}{M}m\right) = \frac{1}{N}\sum_m a_m^*(\delta_+, -i\delta_-)a_m, \tag{4.10}$$

so it is convenient to introduce the vector operator $\boldsymbol{\delta} := (\delta_+, -i\delta_-)$.

In Eq. (4.5), it is now clear that [in place of Eq. (1.9) or (3.7)] we can choose

$$\mathbf{v}_m = \left[\overline{\mathbf{u}}_a - \mathbf{u}\left(\frac{2\pi}{M}m\right)\right]a_m/\sqrt{N}, \tag{4.11a}$$

$$\mathbf{w}_m = [\overline{\boldsymbol{\delta}}_a - \boldsymbol{\delta}]a_m/\sqrt{N}, \tag{4.11b}$$

where by using Eq. (4.10) it is found that

$$\overline{\boldsymbol{\delta}}_a := \frac{1}{N}\sum_m a_m^* \boldsymbol{\delta} a_m = \overline{\mathbf{u}}_A. \tag{4.12}$$

The left-hand side of Eq. (4.5) can now be seen to be just $(1-\overline{\mathbf{u}}_a^2)(1-\overline{\mathbf{u}}_A^2)$, i.e., $\Delta_a^2\Delta_A^2$. The sum on the right-hand side can be treated in the usual way:

$$\begin{aligned}&\frac{1}{N}\sum_m a_m^*\left[\mathbf{u}\left(\frac{2\pi}{M}m\right) - \overline{\mathbf{u}}_a\right]\cdot(\boldsymbol{\delta} - \overline{\boldsymbol{\delta}}_a)a_m \\ &\quad = \frac{1}{N}\sum_m a_m^* \mathbf{u}\left(\frac{2\pi}{M}m\right)\cdot\boldsymbol{\delta} a_m - \overline{\mathbf{u}}_a\cdot\overline{\boldsymbol{\delta}}_a = e^{i\pi/M}V - \overline{\mathbf{u}}_a\cdot\overline{\mathbf{u}}_A,\end{aligned} \tag{4.13}$$

where

$$V := \frac{1}{N}\mathrm{Re}\left\{\sum_m a_m^* a_{m-1}\exp\left[i\frac{2\pi}{M}\left(m-\frac{1}{2}\right)\right]\right\}. \tag{4.14}$$

With this, Eq. (4.5) can be written as

$$\Delta_a^2\Delta_A^2 \geqslant V^2\sin^2(\pi/M) + [V\cos(\pi/M) - \overline{\mathbf{u}}_a\cdot\overline{\mathbf{u}}_A]^2, \tag{4.15}$$

and this is precisely the type of analog of Eq. (1.12) that was sought. Again, the vector approach used here is convenient to achieve the separation into explicitly real and imaginary parts in Eq. (4.13). [A number of variations of Eq. (4.15) can be derived by inserting a unitary matrix into the definition of $\mathbf{w}_m$ in Eq. (4.11b).]

One option for establishing a link to the familiar uncertainty relation of Sec. I is to consider sampling a continuous square-summable function. For example, consider $g(x)$ where the coordinate origin is chosen to be at the maximum value of g, and take $a_m = g(ms)$ where, for an even value of M, $m = -M/2, -M/2+1, \ldots, M/2-1$. In the limit of small sample spacing, i.e., $s \to 0$, it is clear that $\Delta_a \to 1$ and $\Delta_A \to 0$ because the sample values become uniform. As a result, the uncertainty product, i.e., $\Delta_a\Delta_A$, approaches zero. Conversely, as $s \to \infty$, $\Delta_a \to 0$ hence $\Delta_A \to 1$, so $\Delta_a\Delta_A$ also vanishes in the limit of large sample spacing. For intermediate values of s, $\Delta_a\Delta_A$ does not vanish. In numerical work where g and its Fourier transform are of interest, a natural option for choosing the sample spacing is to maximize this uncertainty product. This could be expected to balance the trade-off between the gain from finer sampling in one space and the loss due to coarser sampling in the other. While there are better ways to get around the issue when M is sufficiently large,[21] this idea allows an interesting observation for the current purposes. It turns out that $\Delta_a\Delta_A$ is not only roughly constant over a wide range of sample spacing—well clear of the two extremes just mentioned—but it is also simply linked to the traditional uncertainty product for $g(x)$.

First, notice that it follows from Eqs. (1.1) and (4.1) that

$$\begin{aligned}A_n &= \frac{1}{\sqrt{M}}\sum_{m=-M/2}^{M/2-1} g(ms)\exp\left(-i\frac{2\pi}{M}nm\right) \\ &\approx \frac{1}{s}\sqrt{\frac{2\pi}{M}}\,G\left(n\frac{2\pi}{Ms}\right).\end{aligned} \tag{4.16}$$

That is, the conjugate data set can be approximated by sampling $G(p)$ at a spacing of $v = 2\pi/Ms$—hence the trade-off mentioned above. This approximation is most valid for $n = -M/2, -M/2+1, \ldots, M/2-1$, and it is better for larger M. For optimal sampling, it is natural to have $\Delta_x/s \approx \Delta_p/v$, and this condition gives

$$s \approx \sqrt{\frac{2\pi}{M}\frac{\Delta_x}{\Delta_p}} = s_0. \tag{4.17}$$

If we therefore choose $s = \gamma s_0$ for some constant γ, the number of samples across the interval Δ_x is

$$\Delta_x/s = \frac{1}{\gamma}\sqrt{\frac{M}{2\pi}\Delta_x\Delta_p}.$$

This is a small fraction of M provided γ is much larger than $\sqrt{\Delta_x\Delta_p/M}$, and it then follows (from the relation established in Sec. II A) that

$$\Delta_a \approx \frac{2\pi}{M}\frac{\Delta_x}{s} = \frac{1}{\gamma}\sqrt{\frac{2\pi}{M}\Delta_x\Delta_p}. \tag{4.18}$$

Similarly, the number of samples across Δ_p is

$$\Delta_p/v = \gamma\sqrt{\frac{M}{2\pi}\Delta_x\Delta_p}.$$

This is a small fraction of M provided γ is much less than $\sqrt{M/\Delta_x\Delta_p}$, and then

$$\Delta_A \approx \frac{2\pi}{M}\frac{\Delta_p}{v} = \gamma\sqrt{\frac{2\pi}{M}\Delta_x\Delta_p}. \tag{4.19}$$

When $\sqrt{\Delta_x\Delta_p/M} \ll \gamma \ll \sqrt{M/\Delta_x\Delta_p}$, it now follows from Eqs. (4.18) and (4.19) that the two uncertainty products are related simply by

$$\Delta_a\Delta_A \approx \frac{2\pi}{M}\Delta_x\Delta_p\,. \tag{4.20}$$

Not only is the new uncertainty product effectively constant over this (typically wide) range of values for the sample spacing, but it is evidently simply related to the traditional uncertainty product of the function being sampled. However, when $\Delta_x\Delta_p > M$, the stable region of γ becomes empty. As an extreme example, notice that when g is discontinuous [e.g., $g(x)=\mathrm{rect}(x)$] it follows that $\Delta_x\Delta_p$ is necessarily infinite. In such cases, $\Delta_a\Delta_A$ no longer has a broad flattop when plotted as a function of γ; instead, it attains a peak value that increases with M.

V. CONCLUDING REMARKS

The uncertainty relations for Fourier series and FFTs that have been derived here can have significance and applications much like those described in Sec. I for the standard uncertainty principle. In other words, the uncertainty relations of Eqs. (3.12) and (4.15) can become uncertainty principles once they are associated with physical processes. For example, we have proposed an application in the context of beam quality measures.[22] It is also important to appreciate that the measure of spread discussed in Sec. II has significance on its own: Whenever experimental measurements are made, an expected error must accompany the nominal value. This simply quantifies the width of the distribution that reflects the uncertainty in the result. Just as the conventional measure is the variance (or centered second moment) of the distribution, we claim that the natural analog for a periodic entity is precisely Δ_θ of Eq. (2.3). Further, this measure of spread generalizes easily for distributions on a sphere (as we show in Ref. 22).

A number of natural extensions to these results could also be considered. For example, it would be interesting to derive the analogs of Eq. (4.15) that use the interpolated measure of spread given in Eq. (A4)—or even the higher-order analogs of Eq. (A4). As pointed out in the Appendix, the minimum uncertainty is then nonzero and so, therefore, is the lower bound on the new uncertainty product. It may be possible, even without this generalization, to derive a more useful alternative to Eq. (4.15) that makes it evident that if Δ_a becomes small, Δ_A must approach unity (and vice versa). In all of these options, the associated minimum uncertainty states may also be of interest. Just as the derivations given here can add to the understanding of limits to simultaneous localization in coordinate and frequency space, the extra challenges just outlined are also ideal for consideration in advanced courses that cover Fourier methods in physics.

ACKNOWLEDGMENTS

We are grateful to Dr. Barry Sanders and Dr. Sam Braunstein for drawing relevant publications to our attention. This project is funded by the Australian Research Council and by Macquarie University.

APPENDIX: INTERPOLATION-BASED MEASURES OF SPREAD

When the original data are a set of samples of a function of a continuous variable, the results given in Eq. (2.9) are not always appropriate. One natural alternative follows upon interpolating between the sample values and measuring Δ_θ for the associated continuous function. The simplest case of linear interpolation is considered explicitly here. The periodic interpolating function can therefore be written, for $\theta \in (0,2\pi)$, as

$$\tilde{f}(\theta) := \sum_{m=0}^{M} a_m \Lambda\left(\frac{M}{2\pi}\theta - m\right), \tag{A1}$$

where, here and in what follows, $a_{m+M}=a_m$ for all m, and

$$\Lambda(x) := \begin{cases} 1-|x|, & |x|<1 \\ 0, & |x|>1 \end{cases}. \tag{A2}$$

In contrast to Eq. (2.9b), the normalization factor is now found from Eq. (A1) to be

$$\begin{aligned}\tilde{N} := \int |\tilde{f}(\theta)|^2\, d\theta &= \frac{2\pi}{M}\frac{1}{6}\sum_m (4a_m + a_{m-1} + a_{m+1})a_m^* \\ &= \frac{2\pi}{M}\frac{1}{3}\sum_m [2|a_m|^2 + \mathrm{Re}(a_m^* a_{m-1})],\end{aligned} \tag{A3}$$

and Eq. (2.2a) with $f=\tilde{f}$ and $N=\tilde{N}$ leads similarly to

$$\begin{aligned}\overline{\mathbf{u}} = \frac{2\pi}{M\tilde{N}}\sum_m \Bigg\{ & D\left(\frac{2\pi}{M}\right)|a_m|^2 \mathbf{u}\left(\frac{2\pi}{M}m\right) \\ & + B\left(\frac{2\pi}{M}\right)\mathrm{Re}(a_m^* a_{m-1})\mathbf{u}\left[\frac{2\pi}{M}\left(m-\frac{1}{2}\right)\right]\Bigg\},\end{aligned} \tag{A4}$$

where

$$D(v) := \frac{4}{v^2}\left(1-\frac{\sin v}{v}\right) = \frac{2}{3} - v^2/30 + O(v^4), \tag{A5}$$

$$B(v) := \frac{4}{v^2}\left[\cos(v/2) - \frac{\sin(v/2)}{v/2}\right] = \frac{1}{3} - v^2/120 + O(v^4). \tag{A6}$$

Notice that the centroid given in Eq. (A4) approaches the one given in Eq. (2.9) as M becomes large. It is also satisfying to see from Eq. (A4) that if only one of the sample values is nonzero, the uncertainty is then given by

$$\Delta_\theta^2 = 1 - \frac{9}{4}D^2\left(\frac{2\pi}{M}\right) = \frac{1}{10}\left(\frac{2\pi}{M}\right)^2 + O\left[\left(\frac{2\pi}{M}\right)^4\right]. \tag{A7}$$

For large M therefore,

$$\Delta_\theta \approx \frac{1}{\sqrt{10}}\left(\frac{2\pi}{M}\right).$$

As it must be [given Eq. (2.8)], this is consistent with the variance of $\Lambda[(2\pi/M)x]$ and means that the uncertainty is now never less than about one-third of the sample spacing.

[a]Electronic mail: forbes@physics.mq.edu.au
[b]Electronic mail: alonso@physics.mq.edu.au

[1]If a system's response to an input signal is both linear and shift invariant, it is easy to see that the response to any input that satisfies a linear differential equation with constant coefficients is itself a solution of this same equation. As a result, a sinusoidal input (satisfying $y''+k^2y=0$) gives a sinusoidal output of the same frequency. Fourier analysis is then a natural tool.

[2]A general treatment along with a brief history is presented in L. Cohen, *Time-frequency Analysis* (Prentice–Hall, Englewood Cliffs, NJ, 1995), Chap. 3. Notice that the uncertainty principle is described in this book as ''one of the greatest achievements of the century.''

[3]G. Folland and A. Sitaram, ''The uncertainty principle: A mathematical survey,'' J. Fourier Anal. Appl. **3**, 207–238 (1997).

[4]The conceptual foundation is discussed in T. F. Johnston Jr., ''M^2 concept characterizes beam quality,'' Laser Focus World **26** (5), 173–183 (1990). Interestingly, the associated ISO standard has been the subject of significant criticisms relating largely to the noise sensitivity of the moments that are so fundamental to the standard uncertainty principle. See, for example, G. N. Lawrence, ''Proposed international standard for laser-beam quality falls short,'' Laser Focus World **30** (7), 109–114 (1994).

[5]R. Lynch, ''The quantum phase problem: A critical review,'' Phys. Rep. **256**, 367–436 (1995).

[6]This follows simply from the fact that $\int\int|g(x)h(y)-g(y)h(x)|^2dx\,dy \geqslant 0$.

[7]The second step follows by using integration by parts on $\int x f^* f'\,dx$ and then taking the average of the two expressions. This gives a convenient separation into real and imaginary parts.

[8]See, for example, Ref. 2.

[9]N. G. de Bruijn, ''Uncertainty principles in Fourier analysis,'' in *Inequalities*, edited by O. Shisha (Academic, New York, 1967), pp. 57–71.

[10]D. Mustard, ''Uncertainty principles invariant under the fractional Fourier transform,'' J. Aust. Math. Soc. B, Appl. Math. **33**, 180–191 (1991).

[11]D. Slepian, ''Prolate spheroidal wavefunctions, Fourier analysis and uncertainty. V. The discrete case,'' Bell Syst. Tech. J. **57**, 1371–1430 (1978). See also the earlier works in this five-part series of papers (spanning two hundred pages).

[12]D. C. Brody and B. K. Meister, ''Discrete uncertainty relations,'' J. Phys. A **32**, 4921–4930 (1999). This reference includes a brief history of other entropic approaches.

[13]V. DeBrunner *et al.*, ''Resolution in time-frequency,'' IEEE Trans. Signal Process. **47**, 783–788 (1999).

[14]See, for example, W. H. Press *et al.*, *Numerical Recipes* (Cambridge U.P., Cambridge, 1992), Chap. 12.

[15]See P. Carruthers and M. M. Nieto, ''Phase and angle variables in quantum mechanics,'' Rev. Mod. Phys. **40**, 411–440 (1968). A treatment for small angular speads is presented in C. L. Roy and A. B. Sannigrahi, ''Uncertainty relation between angular momentum and angle variable,'' Am. J. Phys. **47**, 965–967 (1979).

[16]See R. Ishii and K. Furukawa, ''The uncertainty principle in discrete signals,'' IEEE Trans. Circuits Syst. **33**, 1032–1034 (1986); L. Calvez and P. Vilbé, ''On the uncertainty principle in discrete signals,'' *ibid.* **39**, 394–395 (1992). Also see P. Carruthers and M. M. Nieto in Ref. 15.

[17]M. I. Doroslovački, ''Product of second moments in time and frequency for discrete-time signals and the uncertainty limit,'' Signal Process. **67**, 59–76 (1998).

[18]This basic observation is made in Sec. 3a. 7 of C. R. Rao, *Linear Statistical Inference and its Applications* (Wiley, New York, 1973), 2nd ed. However, the connection given in Eq. (2.3) is not clear there and, perhaps as a result, the idea is not seriously pursued. (In particular, it is remarkable that Rao does not go on to consider the natural analog of the Cramer–Rao lower bound on this measure of uncertainty in an estimator for a translation parameter of a periodic distribution.)

[19]See Ref. 14, Chap. 12 and Sec. 13.9.

[20]This result can be derived much like Eq. (1.8), but now by starting with $\Sigma_{ij}\int\int|v_i(\theta)w_j(\phi)-w_i(\theta)v_j(\phi)|^2 d\theta d\phi \geqslant 0$, where $v_i(\theta)$ are the components of the vector function $\mathbf{v}(\theta)$.

[21]D. H. Bailey and P. N. Swarztrauber, ''A fast method for the numerical evaluation of continuous Fourier and Laplace transforms,'' SIAM J. Sci. Comput. (USA) **55**, 1105–1110 (1994).

[22]M. A. Alonso and G. W. Forbes, ''Uncertainty products for nonparaxial fields,'' J. Opt. Soc. Am. A (to be published).

THE TANTALUS PRINCIPLE

A cautionary note is in order, however. The most distant supernovae are fainter than would be expected in the Einstein–de Sitter case. How do we know that is not because the more distant supernovae are less luminous? The authors present careful checks, but the case has to be indirect: no one is going to examine any of these supernovae up close, let alone make the trip back in time to compare distant supernovae to nearer examples. In short, the supernovae measurement is a great advance, beautifully and carefully done, but it does not come with a guarantee. The point is obvious to astronomers but not always to their colleagues in physics, and so might well be encoded in the Tantalus Principle: in astronomy you can look but never touch (with a few exceptions, such as objects in the Solar System, that are quite irrelevant for our purpose).

P. J. E. Peebles, ''Is Cosmology Solved,'' Publ. Astron. Soc. Pacific **111**, 274 (1999).