Approximations to the quantum phase operator

1995 J. Phys. A: Math. Gen. 28 6809

(http://iopscience.iop.org/0305-4470/28/23/026)

J. Phys. A: Math. Gen. 28 (1995) 6809–6856. Printed in the UK

Approximations to the quantum phase operator

M A Hennings†, T B Smith‡ and D A Dubin§

† Sidney Sussex College, Cambridge University, Cambridge, UK
‡ Physics Department, The Open University, Milton Keynes, UK
§ Pure Mathematics Department, The Open University, Milton Keynes, UK

Received 8 February 1995, in final form 1 September 1995

Abstract. In a series of papers, Barnett, Pegg and various co-authors have proposed a description of quantum phase by means of a collection of s-dimensional states and operators, $s \geqslant 1$. We analyse the limiting procedure they employ for large s, which is known not to be compatible with quantum mechanics in the usual sense. Further, we supply a rigorous demonstration of the asymptotic limits of the 'mean' and 'variance' of their system of operators in coherent states. These values had previously been given but not justified mathematically. Our analysis, based on the asymptotic analysis of certain random variables, shows that the physical deductions that can be drawn from these limits are limited.

We also prove that the $(s+1)$-dimensional 'pure phase' LHW-states they consider form a sequence of approximate eigenvectors for the Weyl-quantized angle operator $\Delta(\varphi)$ and the Toeplitz phase operator X proposed by Garrison and Wong, Popov and Yarunin, and others.

These states $\chi_s(\theta)$, to our knowledge first introduced by Lerner, Huang and Walters, can be used to construct a system of measurement as in the usual quantum theory, sensitive to certain qualities of phase, but not all. Indeed, a feature of the Barnett–Pegg method, when it gives finite answers, is the construction of associated measurement systems for different observables. We give examples of sequences of $(s+1)$-dimensional devices which represent measurements significantly closer to ideal for X. This serves as a model for corresponding devices for $\Delta(\varphi)$, or indeed, any observable with a continuous spectrum, contingent on its spectral decomposition being obtained explicitly.

1. Introduction

It is generally accepted that electromagnetic waves in an optical cavity can exist in states characterized as pure photon states, or very nearly so. To a good approximation this situation can be modelled as an assemblage of Boson oscillators, representing the discrete wavenumber modes of the field. The pure photon states are then Hermite functions, the eigenstates of the number operators for the modes.

In other circumstances, there are states of the field that seem to exhibit definite phase relations between the modes. A naive application of the principle of complementarity suggests that such states are eigenstates of an operator canonically conjugate to the number operator for each mode. That such a canonical phase operator does not exist in quantum mechanics is well known. Hence any scheme to describe these states of the electromagnetic field must do so through an observable which is in some sense a non-canonical phase operator, and various candidates for such an operator have been proposed. As a matter of terminology, we shall refer to the states in question as phase states, but we do not suppose them to be eigenstates of an observable, since we believe them to be associated with an observable having a continuous spectrum.

0305-4470/95/236809+48$19.50

Barnett and Pegg and others have proposed a description of phase states through what is claimed to be a phase operator, and have investigated the consequences of this claim in a number of papers, of which [1–10] provides a representative sample.

It will be argued in this paper that there are a number of important objections to their theory. In particular, their calculational procedures lead outside quantum theory. We do not accept that the phenomenon being described requires such a drastic revision of physics, particularly as there do exist candidates for a phase operator within quantum theory. For example there is the Toeplitz operator X of Garrison and Wong [12] introduced as an operator on the Hardy space H^2 over the unit circle, also considered by Popov and Yarunin [13], and the operator $\Delta(\varphi)$ introduced by ourselves [14, 15] and Royer [21] as the Weyl quantization of the angle function on phase space. For two recent review papers concerning the problem of phase, see Lynch [27] and Dubin *et al* [28].

In section 2 of this paper, we describe the formalism of Barnett and Pegg as we understand it, and offer some mathematical criticisms of their theory: their 'phase states' do not converge strongly in Hilbert space and their limiting procedure does not give well defined and finite expectations for all reasonable quantum mechanical observables. Nor—in the limit—do the 'moments' of their 'phase operator' with respect to the 'phase state' correspond to the moments with respect to a state of any Hilbert space operator.

In section 2 we also contrast the distinct predictions for the mean and variance with respect to oscillator states of the Toeplitz phase operator X, the 'phase operator' of Barnett and Pegg, and the Weyl quantization of the phase angle, $\Delta(\varphi)$.

In section 3 we discuss the variance of the 'phase' with respect to the coherent state Φ_α for these three theories, for large values of the parameter $|\alpha|$. The asymptotic dependence on α claimed by Barnett and Pegg [3] is verified rigorously. This verification entails considerable analysis, and we show that the simple arguments offered by Barnett and Pegg are mathematically insufficient.

For large $|\alpha|$, the variance of X in the state Φ_α is of the same order as for the Barnett and Pegg theory. In contrast, the variance of $\Delta(\varphi)$ is of a different order. We suspect that measurements to date are of quantization of periodic functions of the phase. As the experiments involve only small values of $|\alpha|$ in addition, it seems premature to choose a theory on the basis of the experimental evidence to date. Analysis of the variance for $\Delta(\varphi)$ will be found in the companion paper [18].

In section 4 we discuss the measurement process in quantum mechanics. After considering some problems inherent in any system of measurement we turn to what is our particular interest here: the measurement of bounded observables which have a continuous spectral component. After pointing out why measuring such an observable requires a measuring apparatus responding to a discrete spectrum only, we construct 'ideal' systems of measurement for such observables, based on their spectral decompositions. Such systems have optimal properties as regards both spectral response and output states.

As a preparation for considering the formalism of Barnett and Pegg as a system of measurement for phase, we use our knowledge of spectral systems of measurement to construct systems of a loosely analogous type. These will give good spectral response, but, in contrast to the 'ideal' measurement systems, the output states are not constrained to approach the generalized eigenvectors of the observable to be measured.

We then note additional conditions proposed by Barnett and Pegg to compensate for this lack of convergence, and set them in the measurement context. This allows their conditions to be examined in a full quantum mechanical setting.

Other than 'ideal' systems of measurements, in section 4 we consider systems determined by sequences of approximate eigenvectors, which are defined there. These systems have less desirable properties than the 'ideal' systems, but have their uses nonetheless.

In particular, we show that the states $\chi_s(\theta)$ introduced by Lerner, Huang and Walters [11] constitute approximate eigenvectors for a number of operators of interest. In section 5 we prove that they constitute such a sequence for the Toeplitz phase and shift operators, and in section 6 we do the same for the Weyl quantizations, $\Delta(\varphi)$, of angle, and $\Delta(e^{\pm i\varphi})$ of complex exponentials of angle. With this analysis in hand, we return to measurement theory in section 7, and show that much of the theory of Barnett and Pegg may be recast as a system of measurement for these operators based on the $\chi_s(\theta)$. From this point of view, the undesirable properties of the Barnett and Pegg variance and higher moments considered in section 2 are seen to be a feature of the measurement system. We also point out in section 7 that the characteristic 'ideal' system is available for X in more or less explicit form as a result of the work of Garrison and Wong [12].

As a bonus, the information we obtain about the spectrum of $\Delta(\varphi)$ from our analyis enables us to say that its spectrum includes the interval $[-\pi, \pi]$, and that its norm lies between π and $3\pi/2$. This reinforces our belief that the spectrum of $\Delta(\varphi)$ is $[-\pi, \pi]$, which we conjectured in [14] partly as a result of computer work reported there.

2. Three theories of phase: oscillator states

In order to make our objections definite, let us describe what we understand Barnett, Pegg and their collaborators to have done. For brevity, we shall refer to their theory as the BP theory. As usual in this subject, we shall restrict attention to one degree of freedom, so that the system Hilbert space is $L^2(\mathbb{R})$.

To begin with, we note that the mathematical analysis in this field often suffers from a lack of rigour. As an example, consider the 'vector' defined by the limit

$$\psi_\theta = \lim_{s\to\infty} \chi_s(\theta) \tag{2.1a}$$

where, for each $s \geqslant 0$, $\chi_s(\theta)$ is the normalized Hilbert space vector

$$\chi_s(\theta) = \frac{1}{\sqrt{s+1}} \sum_{n=0}^{s} e^{in\theta} h_n \tag{2.1b}$$

and $\{h_n : n \geqslant 0\}$ are the familiar Hermite functions, comprising an orthonormal basis for the system Hilbert space $L^2(\mathbb{R})$ of eigenvectors of the harmonic oscillator Hamiltonian.

The $\chi_s(\theta)$ are the 'pure phase' states of Barnett and Pegg [3]. As far as we can ascertain, they were first introduced by Lerner, Huang and Walters [11], who considered them in their analysis of their various shift operators. For this reason we shall refer to the $\chi_s(\theta)$ as LHW states.

The limit (2.1*a*) is also referred to in the first edition (only) of the book of Loudon [16], for example, where it is used to calculate a number of expectation values, and is supposed to determine a pure state of well defined phase. However, this limit does not define a state. To see this, we need only note that the sequence $\{\chi_s(\theta) : s \geqslant 0\}$ of vectors converges weakly but not strongly to the zero vector.

In more operational terms, each vector $\chi_s(\theta)$ defines a state on the algebra of observables by the rule

$$\omega_{s,\theta}(A) = \langle \chi_s(\theta), A\chi_s(\theta)\rangle = \frac{1}{s+1} \sum_{m,n=0}^{s} e^{i(n-m)\theta} \langle h_m, Ah_n\rangle \tag{2.2}$$

for every observable A. It is easy to show that the limit

$$\omega_\theta(A) = \lim_{s\to\infty} \omega_{s,\theta}(A) \tag{2.3}$$

does not exist for all reasonable A. For instance, taking the number operator N to be an observable [7], one has

$$\omega_{s,\theta}(N) = \frac{s}{2}$$

so that the limit as $s \to \infty$ diverges. Thus, there is no such state as ω_θ.

This conclusion does not change if we form mixed states using the $\omega_{s,\theta}$. To do this, multiply $\omega_{s,\theta}$ with an amplitude function $F(\theta)$, integrate over θ, and then take the limit of large s.

Since at this stage we do not know which class of functions F should be used in order that a well defined state results, we must proceed formally. We define

$$\begin{aligned}\omega_F(A) &= \lim_{s\to\infty} \int_{-\pi}^{\pi} F(\theta)\omega_{s\theta}(A)\,\mathrm{d}\theta \\ &= \lim_{s\to\infty} \frac{2\pi}{s+1} \sum_{m,n=0}^{s} \hat{F}_{m-n}\langle h_m, Ah_n\rangle\end{aligned}$$

for all observables A, where $\hat{F}_k$ is the kth Fourier coefficient of F. The probability interpretation for states requires that they be normalized, and so

$$1 = \omega_F(I) = 2\pi\hat{F}_0.$$

Proceeding as before, we consider the observable N, obtaining

$$\omega_F(N) = \lim_{s\to\infty} 2\pi\hat{F}_0\left(\frac{s}{2}\right) = \lim_{s\to\infty}\left(\frac{s}{2}\right) = \infty.$$

We conclude that no such mixed state exists, no matter what choice is made for F.

The BP theory is an attempt to refine the large s limit of the $\chi_s(\theta)$, but in doing so they are forced outside the realm of quantum mechanics, as they acknowledge. They end up not with vectors and operators on Hilbert space, but with collections of these: 'The key feature in the development of the Hermitian optical phase operator was the abandonment of the conventional infinite Hilbert space for the description of the states of a single field mode. In its place, a state space (Ψ) of formally finite dimensions is employed together with a prescription for taking the infinite dimensional limit only after c-number expectation values and moments have been calculated.' [7].

Proceeding with our description of the BP theory, for any integer $s \geqslant 0$, attention is restricted to the $(s+1)$-dimensional subspace $\mathcal{H}_s$ of $L^2(\mathbb{R})$ spanned by the first $s+1$ Hermite functions. Defining the $s+1$ equally spaced angles

$$\theta_{s,j} = \theta_0 + \frac{j}{s+1}2\pi \qquad 0 \leqslant j \leqslant s \tag{2.4}$$

where θ_0 is some fixed angle, the set $\{\chi_s(\theta_{s,j}) : 0 \leqslant j \leqslant s\}$ is an orthonormal basis for $\mathcal{H}_s$.

For given s, the vectors $\chi_s(\theta_{s,j})$ and the numbers $\theta_{s,j}$ are used to construct a finite-rank self-adjoint operator X_s on $L^2(\mathbb{R})$ in standard fashion, by setting

$$X_s = \sum_{j=0}^{s} \theta_{s,j} P_{s,j} \tag{2.5a}$$

where the $P_{s,j}$ are the projections onto the $\chi_s(\theta_{s,j})$:

$$P_{s,j} f = \langle \chi_s(\theta_{s,j}), f \rangle \chi_s(\theta_{s,j}). \tag{2.5b}$$

The operator X_s has been constructed in diagonal form, with

$$X_s \chi_s(\theta_{s,j}) = \theta_{s,j} \chi_s(\theta_{s,j}) \qquad 0 \leqslant j \leqslant s. \tag{2.5c}$$

Thus, we can write down any function of X_s immediately. In particular, it generates the one-parameter unitary group

$$\mathrm{e}^{\mathrm{i}tX_s} = \sum_{j=0}^{s} \mathrm{e}^{\mathrm{i}t\theta_{s,j}} P_{s,j} \qquad t \in \mathbb{R} \tag{2.5d}$$

and the spectrum of $\mathrm{e}^{\mathrm{i}tX_s}$ can be read off as $\{\mathrm{e}^{\mathrm{i}t\theta_{s,j}} : 0 \leqslant j \leqslant s\}$.

Moments of the operator X_s can be calculated in any given state. Now in quantum theory, one might attempt to take the limit of these moments as s tends to infinity, so as to seek to define

$$\mathrm{BP}_r(\psi) = \lim_{s\to\infty} \langle \psi, (X_s)^r \psi \rangle \qquad r \in \mathbb{N} \bigcup \{0\},\ \psi \in L^2(\mathbb{R}). \tag{2.6}$$

In particular, we would be interested in trying to define the quantities

$$\mathrm{E}_{\mathrm{BP}}(\psi) = \lim_{s\to\infty} \langle \psi, X_s \psi \rangle = \mathrm{BP}_1(\psi) \tag{2.7a}$$

and

$$\mathrm{V}_{\mathrm{BP}}(\psi) = \lim_{s\to\infty} [\langle \psi, (X_s)^2 \psi \rangle - \langle \psi, X_s \psi \rangle^2] = \mathrm{BP}_2(\psi) - \mathrm{BP}_1(\psi)^2. \tag{2.7b}$$

These latter quantities have been interpreted in BP theory as the mean and variance of their 'phase operator' in the state ψ. Indeed, BP theory would interpret $\mathrm{BP}_r(\psi)$ as the rth moment of their 'phase operator' in the state ψ. However, we shall now show that this interpretation is not valid.

For any function $a \in L^\infty(\mathbb{T})$ we can define a bounded operator $T(a)$ on the Hardy space $H^2(\mathbb{T})$ by

$$T(a)f = P_+(af) \qquad f \in H^2(\mathbb{T}) \tag{2.8a}$$

where P_+ is the orthogonal projection from $L^2(\mathbb{T})$ onto $H^2(\mathbb{T})$ and af indicates the pointwise product of a and f. The norm of $T(a)$ satisfies the bound

$$\|T(a)\| \leqslant \|a\|_\infty. \tag{2.8b}$$

The map $T : L^\infty(\mathbb{T}) \to \mathcal{L}[H^2(\mathbb{T})]$ is linear and preserves adjoints in the sense that

$$T(a)^* = T(\overline{a}) \qquad a \in L^\infty(\mathbb{T}) \tag{2.9}$$

and we see that

$$\langle f, T(a)g\rangle = \frac{1}{2\pi}\int_{-\pi}^{+\pi} \overline{f(\mathrm{e}^{\mathrm{i}\theta})}a(\mathrm{e}^{\mathrm{i}\theta})g(\mathrm{e}^{\mathrm{i}\theta})\,\mathrm{d}\theta \qquad a \in L^\infty(\mathbb{T})\ f, g \in H^2(\mathbb{T}) \tag{2.10}$$

but M is not an algebra homomorphism. However,

$$\begin{aligned}\langle f, T(|a|^2)f\rangle &= \frac{1}{2\pi}\int_{-\pi}^{+\pi} |f(\mathrm{e}^{\mathrm{i}\theta})|^2|a(\mathrm{e}^{\mathrm{i}\theta})|^2\,\mathrm{d}\theta \\ &= \|af\|^2 \\ &\geqslant \|P_+(af)\|^2 \\ &= \|T(a)f\|^2\end{aligned}$$

for all $a \in L^\infty(\mathbb{T})$ and $f \in H^2(\mathbb{T})$. It follows that

$$T(|a|^2) \geqslant T(a)^*T(a) \tag{2.11}$$

for all $a \in L^\infty(\mathbb{T})$.

With respect to the standard orthonormal basis $\{e_n : n \geqslant 0\}$ of $H^2(\mathbb{T})$, where

$$e_n(\mathrm{e}^{\mathrm{i}\theta}) = \mathrm{e}^{\mathrm{i}n\theta} \qquad n \geqslant 0 \tag{2.12}$$

we see that

$$\langle e_m, T(a)e_n\rangle = \hat{a}_{m-n} \qquad a \in L^\infty(\mathbb{T}),\ m, n \geqslant 0 \tag{2.13}$$

proving that $T(a)$ is a Toeplitz operator.

If we introduce the unitary map U from $L^2(\mathbb{R})$ onto $H^2(\mathbb{T})$, given by

$$Uh_n = \mathrm{i}^{-n}e_n \qquad n \geqslant 0 \tag{2.14}$$

we can combine U and T to construct a linear map $\hat{T}(a) : L^\infty(\mathbb{T}) \to \mathcal{L}[L^2(\mathbb{R})]$ by

$$\hat{T}(a) = U^*T(a)U \qquad a \in L^\infty(\mathbb{T}) \tag{2.15}$$

with the properties

$$\|\hat{T}(a)\| \leqslant \|a\|_\infty \tag{2.16a}$$

$$\hat{T}(a)^* = \hat{T}(\overline{a}) \tag{2.16b}$$

$$\langle h_m, \hat{T}(a)h_n\rangle = \mathrm{i}^{m-n}\hat{a}_{m-n} \tag{2.16c}$$

where $a \in L^\infty(\mathbb{T})$ and $m, n \geqslant 0$; just as for T, $\hat{T}$ is not an algebra homomorphism.

In particular,

$$X = \hat{T}(\Theta) \tag{2.17}$$

is the Toeplitz phase operator of Garrison and Wong [12], Popov and Yarunin [13], and others. (Taking into account that in this and our previous work we have adopted a definition of angle complementary to that usually chosen, we must make minor modifications to the definitions of the Toeplitz phase operator and the states $\chi_s(\theta)$ ordinarily used. This will become clearer in the course of this work.) We have introduced the coordinate function Θ on the unit circle $\mathbb{T}$, and will use

$$\Theta^r(\mathrm{e}^{\mathrm{i}\theta}) = \theta^r \qquad -\pi < \theta < \pi,\ r \in \mathbb{N} \tag{2.18}$$

for its powers.

Here, as elsewhere, we shall work with the BP operators for which the fiducial angle θ_0 has been chosen to be equal to $-\pi$. In this case, for any integers r, m, $n \geqslant 0$, direct calculation shows that

$$\langle h_m, (X_s)^r h_n\rangle = \frac{\pi^r \mathrm{i}^{n-m}}{(s+1)^{r+1}} \sum_{k=0}^{s} (2k-s-1)^r \mathrm{e}^{2\pi\mathrm{i}(n-m)k/(s+1)} \tag{2.19}$$

for all $s \geqslant \max(m, n)$. Hence we can prove that

$$\lim_{s\to\infty} \langle h_m, (X_s)^r h_n\rangle = \mathrm{i}^{m-n}\widehat{\Theta^r}_{m-n} = \langle h_m, \hat{T}(\Theta^r) h_n\rangle. \tag{2.20}$$

Since $\{(X_s)^r : s \in \mathbb{N}\}$ *is a uniformly bounded sequence of operators on* $L^2(\mathbb{R})$, *with* $\|(X_s)^r\| \leqslant \pi^r$ for all $s \in \mathbb{N}$, we deduce that

$$\lim_{s\to\infty} (X_s)^r = \hat{T}(\Theta^r) \tag{2.21}$$

weakly in $\mathcal{L}[L^2(\mathbb{R})]$ for any integer $r \geqslant 0$. However, the sequence does not converge strongly, as can be seen, for example, from the fact that

$$\hat{T}(\Theta^2) \neq \hat{T}(\Theta)^2. \tag{2.22}$$

The first consequence of this calculation is that the quantities

$$\mathrm{BP}_r(\psi) = \langle \psi, \hat{T}(\Theta^r)\psi\rangle = \frac{1}{2\pi}\int_{-\pi}^{+\pi} \theta^r |(U\psi)(\mathrm{e}^{\mathrm{i}\theta})|^2 \,\mathrm{d}\theta \qquad r \geqslant 0,\ \psi \in L^2(\mathbb{R}) \tag{2.23}$$

all exist, and so if ψ is a unit vector in $L^2(\mathbb{R})$, $\mathrm{BP}_r(\psi)$ can be interpreted as the rth moment of a $[-\pi, \pi]$-valued random variable whose density function is

$$P_\psi(\theta) = \frac{1}{2\pi}|(U\psi)(\mathrm{e}^{\mathrm{i}\theta})|^2 \qquad -\pi \leqslant \theta \leqslant \pi. \tag{2.24}$$

It follows that $\mathrm{E}_{\mathrm{BP}}(\psi)$ and $\mathrm{V}_{\mathrm{BP}}(\psi)$ can be interpreted as the mean and variance of this random variable.

Alternatively, it can be seen that

$$\mathrm{BP}_r(\psi) = \langle U\psi, Z^r U\psi\rangle \qquad r \in \mathbb{N}\bigcup\{0\} \tag{2.25}$$

where $U\psi \in H^2(\mathbb{T}) \subset L^2(\mathbb{T})$ and Z is the bounded operator on $L^2(\mathbb{T})$—and not $H^2(\mathbb{T})$—given by multiplication by the function Θ:

$$Zf = \Theta f \qquad f \in L^2(\mathbb{T}). \tag{2.26}$$

But Z does not preserve Hardy space $H^2(\mathbb{T})$, which is the subspace of $L^2(\mathbb{T})$ which corresponds to the quantum mechanical Hilbert space $L^2(\mathbb{R})$. So while, for example, $\mathrm{E}_{\mathrm{BP}}(\psi)$ and $\mathrm{V}_{\mathrm{BP}}(\psi)$ are the mean and variance of the self-adjoint operator Z on $l^2(\mathbb{T})$ in the state $U\psi$, in general *there does not exist a single bounded operator* X_∞ on $L^2(\mathbb{R})$ such that

$$\mathrm{BP}_r(\psi) = \langle \psi, X_\infty^r \psi\rangle \tag{2.27}$$

for all $\psi \in L^2(\mathbb{R})$ and all $r \geqslant 0$. In consequence, there does not seem to be any consistent quantum mechanical interpretation of the quantities $\mathrm{BP}_r(\psi)$.

An interesting relationship can be noted between the BP 'moments' and the Toeplitz phase operator X. Certainly

$$\mathrm{E}_{\mathrm{BP}}(\psi) = \mathrm{BP}_1(\psi) = \langle \psi, \hat{T}(\Theta)\psi\rangle = \mathrm{Exp}[X;\psi]$$

the expectation of X for any state $\psi \in L^2(\mathbb{R})$. We also have

$$\mathrm{BP}_2(\psi) = \langle \psi, \hat{T}(\Theta^2)\psi\rangle \geqslant \langle \psi, \hat{T}(\Theta)^2\psi\rangle = \langle \psi, X^2\psi\rangle$$

so that we get the general bound

$$\mathrm{V}_{\mathrm{BP}}(\psi) = \mathrm{BP}_2(\psi) - \mathrm{BP}_1(\psi) \geqslant \langle \psi, X^2\psi\rangle - \langle \psi, X\psi\rangle^2 = \mathrm{var}[X;\psi]$$

for any $\psi \in L^2(\mathbb{R})$.

It is easy to derive the BP result

$$\mathrm{V}_{\mathrm{BP}}(h_n) = \frac{\pi^2}{3} \qquad n \geqslant 0 \tag{2.28}$$

which has been interpreted as realizing a minimal uncertainty between the number and phase operators (which is suggested as an expected feature of a phase operator [6]). This result should be contrasted with the corresponding result [20] for X:

$$\mathrm{var}[X; h_n] = \frac{\pi^2}{3} - \sum_{k=n+1}^{\infty} \frac{1}{k^2} \qquad n \geqslant 0 \tag{2.29}$$

(note that $\mathrm{var}[X; h_n] \leqslant \mathrm{V}_{\mathrm{BP}}(h_n)$ in accordance with our general bound), and with the result

$$\mathrm{var}[\Delta(\varphi); h_n] = \frac{3\pi^2}{8} + \sum_{\substack{0\leqslant l,m\leqslant[\frac{n-2}{2}]\\ l+m\leqslant[\frac{n-2}{2}]}} \frac{1}{(2l+1)(2m+1)} - \left(\sum_{m=0}^{[\frac{n-1}{2}]} \frac{1}{2m+1}\right)^2 \tag{2.30a}$$

$$= \frac{\pi^2}{3} + O\left(\frac{\log n}{n}\right) \qquad n \to \infty \tag{2.30b}$$

for the phase operator $\Delta(\varphi)$ obtained by quantizing the angle function φ on phase space, using the Wigner–Weyl correspondence [18].

We believe that there is no reason to assume *a priori* that a phase operator ought to have the variance $\pi^2/3$ in the Hermite states. Alternatively, one might suggest that the deviations from $\pi^2/3$ represents a quantum effect, which disappears as n tends to infinity.

3. Three theories of phase: coherent states

By a coherent state with parameter $\alpha \in \mathbb{C}$ we mean the Hilbert space vector

$$\begin{aligned}\Phi_\alpha(q) &= e^{-\frac{1}{4}|\alpha|^2}\sum_{n\geqslant 0}\frac{i^n\overline{\alpha}^n}{\sqrt{2^n n!}}h_n(q)\\ &= \frac{1}{\pi^{1/4}}\exp[\tfrac{1}{4}(\overline{\alpha}^2-|\alpha|^2)]\exp[-\tfrac{1}{2}q^2+i\overline{\alpha}q].\end{aligned}$$

Note that the usual parametrization would be in terms of $a = i\overline{\alpha}/\sqrt{2}$. For example, the expectation of the number operator N in the state Φ_α is $|a|^2 = |\alpha|^2/2$. We have chosen this slightly different parameter for convenience, and to accommodate our choice of phase angle, which is complementary to that used by other authors.

Coherent states have the property of being states of minimal uncertainty for position and momentum. As such, they are viewed as possessing essentially classical attributes. It is widely believed that the number and phase operators will be most nearly canonically conjugate in these states, increasingly so for large values of $R = |\alpha|$. A heuristic argument can be given (see [19] for example) that might lead one to expect that the variance of a phase operator in the state Φ_α should exhibit behaviour like $1/(2R^2)$ as R tends to infinity.

In [18], the authors proved that

$$\mathrm{var}[\Delta(\varphi);\Phi_R] = \|\Delta(\varphi)\Phi_R\|^2 \sim \frac{\pi^{1/2}}{R} \tag{3.1}$$

as R tends to infinity, which is not of the form expected by the above reasoning. However, if the infinite series used to calculate the variance in [18] were truncated, or if the phase space angle function were restricted to a region which excluded the branch line (both of which processes seem to us to be unjustified), the expected $1/2R^2$ behaviour would result.

It is claimed in [3] that $\mathrm{BP}_2(\Phi_R) \sim 1/(2R^2)$, and in [11] that $\mathrm{var}[X;\Phi_R]$ is of order $o(1)$ as R tends to infinity. Regarding this latter quantity, what we can say with certainty is that

$$\mathrm{var}[X;\Phi_R] = O[\mathrm{BP}_2(\Phi_R)] \qquad R\to\infty. \tag{3.2}$$

However, to our knowledge, no explicit claim as to the exact asymptotic behaviour of X has been made, although computations do imply that $\mathrm{var}[X;\Phi_R] \sim 1/2R^2$ [20].

Regarding the claim about the first quantity, it deserves further investigation. Note that (equations (2.23) and (2.24))

$$\mathrm{BP}_2(\Phi_R) = \frac{1}{2\pi}\int_{-\pi}^{+\pi}\theta^2|(U\Phi_R)(e^{i\theta})|^2\,d\theta \tag{3.3}$$

where

$$\begin{aligned}(U\Phi_R)(e^{i\theta}) &= e^{-\frac{1}{4}R^2}\sum_{n=0}^{\infty}\frac{R^n}{\sqrt{2^n n!}}e^{in\theta}\\ &= e^{-\frac{1}{4}R^2}\left(\sum_{n=0}^{\infty}\frac{R^n}{\sqrt{2^n n!}}\right)\chi_R(\theta)\end{aligned}$$

where χ_R is the characteristic function (in the sense of probability theory) of an $\mathbb{N}\bigcup\{0\}$-valued random variable Y_R with distribution

$$\mathcal{P}[Y_R = n] = \left(\sum_{m=0}^{\infty} \frac{R^m}{\sqrt{2^m m!}}\right)^{-1} \frac{R^n}{\sqrt{2^n n!}} \qquad n \geqslant 0. \tag{3.4}$$

In the calculation for $\mathrm{BP}_2(\Phi_R)$ in [3], and in analogous calculations for X in [12], $\chi_R(\theta)$ is replaced by the characteristic function of a normal distribution with the same mean and variance as those of Y_R. This approximation is said to be justified by the analogous procedure for the Poisson distribution.

If this approximation is made, we obtain

$$(U\Phi_R)(\mathrm{e}^{\mathrm{i}\theta}) = (2R)^{1/2}\pi^{1/4}\mathrm{e}^{-\frac{1}{2}R^2\theta^2+\frac{1}{2}\mathrm{i}R^2\theta} \tag{3.5a}$$

so that

$$|(U\Phi_R)(\mathrm{e}^{\mathrm{i}\theta})|^2 = 2R\pi^{1/2}\mathrm{e}^{-R^2\theta^2}. \tag{3.5b}$$

Substituting this into the expression for BP_2 yields

$$\mathrm{BP}_2(\Phi_R) \cong \frac{R}{\pi^{1/2}} \int_{-\pi}^{+\pi} \theta^2 \mathrm{e}^{-R^2\theta^2}\,\mathrm{d}\theta \cong \frac{1}{2R^2}. \tag{3.5c}$$

However, there are several objections to this line of argument. Firstly, no error analysis is presented concerning the effect of this approximation, which is only described as 'good', but no statement as to how good. Nor is there any indication that, whatever measure of 'goodness' holds, it is maintained *uniformly* in θ, so that $\mathrm{BP}_2(\Phi_R)$ can be usefully approximated by this process.

To illustrate these concerns, let us consider the analogous problem of an $\mathbb{N}\bigcup\{0\}$-valued random variable Z_R which has a Poisson distribution,

$$\mathcal{P}[Z_R = n] = \mathrm{e}^{-R^2}\frac{R^{2n}}{n!} \qquad n \geqslant 0 \tag{3.6a}$$

and characteristic function

$$\begin{aligned}\xi_R(\theta) &= \sum_{n=0}^{\infty} \mathcal{P}[Z_R = n]\mathrm{e}^{\mathrm{i}n\theta}\\ &= \exp[R^2(\mathrm{e}^{\mathrm{i}\theta} - 1)]\\ &= \exp[-2R^2\sin^2(\tfrac{1}{2}\theta) + \mathrm{i}R^2\sin\theta].\end{aligned}$$

Calculating the expectation, we find

$$\begin{aligned}\mathrm{Exp}\left(\exp\left[\mathrm{i}\theta\left(\frac{Z_R - R^2}{R}\right)\right]\right) &= \mathrm{e}^{-\mathrm{i}R\theta}\xi_R\left(\frac{\theta}{R}\right)\\ &= \exp[R^2(\mathrm{e}^{\mathrm{i}\theta/R} - 1) - \mathrm{i}R\theta].\end{aligned}$$

Taking the limit in this expression as R tends to infinity,

$$\lim_{R\to\infty} \operatorname{Exp}\left(\exp\left[\mathrm{i}\theta\left(\frac{Z_R - R^2}{R}\right)\right]\right) = \mathrm{e}^{-\frac{1}{2}\theta^2} \tag{3.7}$$

uniformly for θ lying in compact subsets of $\mathbb{R}$. Thus, in the sense of *convergence in distribution*,

$$\frac{Z_R - R^2}{R} \to N(0,1) \qquad \text{as } R \to \infty. \tag{3.8}$$

The notation $N(m, v)$ indicates the normal distribution with mean m and variance v.

It is tempting to believe that this result justifies the approximation (in distribution)

$$Z_R \sim N(R^2, R^2) \tag{3.9}$$

but care needs to be taken here. To see this, let ζ_R be the characteristic function for the process $N(R^2, R^2)$, so that

$$\zeta_R(\theta) = \exp[-\tfrac{1}{2}R^2\theta^2 + \mathrm{i}R^2\theta]. \tag{3.10}$$

Certainly it is not true that

$$\xi_R(\theta) \sim \zeta_R(\theta) \qquad \text{as } R \to \infty$$

except for the value $\theta = 0$. Indeed,

$$\zeta_R(\theta) = o[\xi_R(\theta)] \qquad 0 < |\theta| < \pi.$$

Were it possible to write

$$|\xi_R(\theta)|^2 = |\zeta_R(\theta)|^2 + \epsilon_R(\theta)$$

with

$$\epsilon_R(\theta) = O\left(\frac{1}{R^4}\right)$$

as $R \to \infty$ uniformly in $\theta \in [-\pi, \pi]$, then we would have

$$\begin{aligned}
\frac{1}{2\pi}\int_{-\pi}^{+\pi} \theta^2|\xi_R(\theta)|^2\,\mathrm{d}\theta &= \frac{1}{2\pi}\int_{-\pi}^{+\pi} \theta^2|\zeta_R(\theta)|^2\,\mathrm{d}\theta + \frac{1}{2\pi}\int_{-\pi}^{+\pi} \theta^2\epsilon_R(\theta)\,\mathrm{d}\theta \\
&= \frac{1}{2\pi}\int_{-\pi}^{+\pi} \theta^2\mathrm{e}^{-R^2\theta^2}\,\mathrm{d}\theta + \frac{1}{2\pi}\int_{-\pi}^{+\pi} \theta^2\epsilon_R(\theta)\,\mathrm{d}\theta \\
&= \frac{1}{2\pi R^2}\int_{-\pi R}^{\pi R} \theta^2\mathrm{e}^{-\theta^2}\,\mathrm{d}\theta + \frac{1}{2\pi}\int_{-\pi}^{+\pi} \theta^2\epsilon_R(\theta)\,\mathrm{d}\theta \\
&= \frac{1}{4\pi^{1/2}R^3} + O\left(\frac{1}{R^4}\right).
\end{aligned}$$

This would justify the substitution, since the term $1/4\pi^{1/2}R^3$ is the contribution from ζ_R, the error term ϵ_R contributing only effects of smaller order.

However, *this is not true*, since

$$\epsilon_R\left(\frac{1}{R}\right) = \exp\left[-4R^2\sin^2\left(\frac{1}{2R}\right)\right] - \mathrm{e}^{-1}$$
$$= \mathrm{e}^{-1}\left(\exp\left[1 - 4R^2\sin^2\left(\frac{1}{2R}\right)\right] - 1\right)$$
$$= O\left(\frac{1}{R^2}\right).$$

This is the method that would have to be used to justify such a substitution in a general case, and it does not work here. But special circumstances prevail for this example, and it is possible to justify the substitution, which we show as follows.

In this case, we know the exact form of $\xi_R(\theta)$, and by using Laplace's method of asymptotic expansion we obtain

$$\frac{1}{2\pi}\int_{-\pi}^{+\pi}\theta^2|\xi_R(\theta)|^2\,\mathrm{d}\theta = \frac{1}{2\pi}\int_{-\pi}^{+\pi}\theta^2\exp\left[-4R^2\sin^2\left(\frac{1}{2}\theta\right)\right]\mathrm{d}\theta \sim \frac{1}{4\pi^{1/2}R^3}$$

and

$$\frac{1}{2\pi}\int_{-\pi}^{+\pi}\theta^2|\zeta_R(\theta)|^2\,\mathrm{d}\theta = \frac{1}{2\pi}\int_{-\pi}^{+\pi}\theta^2\exp(-R^2\theta^2)\,\mathrm{d}\theta \sim \frac{1}{4\pi^{1/2}R^3}$$

as R tends to infinity.

Thus, in spite of being quite different, these two expressions have the same asymptotic form, which is essentially determined by the behaviour of $\xi_R(\theta)$ and $\zeta_R(\theta)$ at $\theta = 0$, at which point they happen to agree.

Thus, the region around $\theta = 0$ is the only area of influence in these calculations, which shows that the branch line in the definition of the angle function is unimportant for these calculations.

Having shown that replacing the Poisson distribution by a normal distribution does give the correct results in the above calculation, but that justifying the substitution is only possible because we can calculate all integrals directly, it is clear that, while replacing the 'root Poisson' random variable Y_R by a normal distribution may indeed give valid results (for certain asymptotic expressions), justifying the procedure may not be straightforward, since we cannot explicitly calculate the characteristic function $\chi_R(\theta)$ of Y_R.

After this background, we shall now investigate the behaviour of

$$\mathrm{BP}_2(\Phi_R) = \frac{1}{2\pi}\int_{-\pi}^{+\pi}\theta^2|(U\Phi_R)(\mathrm{e}^{\mathrm{i}\theta})|^2\,\mathrm{d}\theta \tag{3.11}$$

as $R \to \infty$. The method we use hinges upon turning the sum for Φ_R into an integral.

Given a series

$$\sum_{n\geqslant 0} a_n$$

which may be finite, one may consider an analytic function $a(z)$, defined so that $a(n)$ is just a_n. Depending on the properties of $a(z)$, it might be possible to choose a contour so that the integral

$$\frac{1}{2\mathrm{i}}\oint_C a(z)\cot(\pi z)\,\mathrm{d}z$$

is equal to the series via the residue theorem, together with other terms that can be managed. If so, an integral representation for the series may be obtained.

By deforming the contour, various other useful representations might be possible, perhaps leading to asymptotic expressions. This technique is well known in Regge-pole theory, where it is used to derive the Sommerfeld–Watson formula.

Of course great care has to be exercised in doing this. We begin with a finite sum and a correspondingly bounded contour: by $C(n)$ we mean the rectangular contour with vertices

$$-\tfrac{1}{2}-\mathrm{i}(n+\tfrac{1}{2}) \qquad n-\tfrac{1}{2}-\mathrm{i}(n+\tfrac{1}{2}) \qquad n-\tfrac{1}{2}+\mathrm{i}(n+\tfrac{1}{2}) \qquad -\tfrac{1}{2}+\mathrm{i}(n+\tfrac{1}{2}).$$

Then

$$\sum_{k=0}^{n-1}\frac{R^k}{\sqrt{k!}}\mathrm{e}^{\mathrm{i}k\theta} = \frac{1}{2\mathrm{i}}\oint_{C(n)}\frac{R^z}{\Gamma(z+1)^{1/2}}\mathrm{e}^{\mathrm{i}z\theta}\cot(\pi z)\,\mathrm{d}z. \tag{3.12}$$

By considering the form of the integral along each of the four sides, we are able to take the limit as $n\to\infty$, and obtain an integral representation:

$$\sum_{n=0}^{\infty}\frac{R^n}{\sqrt{n!}}\mathrm{e}^{\mathrm{i}n\theta} = \int_{-\frac{1}{2}}^{\infty}\frac{R^t\mathrm{e}^{\mathrm{i}t\theta}}{\Gamma(t+1)^{1/2}}\,\mathrm{d}t + \frac{1}{2R^{1/2}\mathrm{e}^{\mathrm{i}\theta/2}}\int_0^{\infty}\frac{R^{\mathrm{i}s}\mathrm{e}^{-\pi s}}{\cosh(\pi s)}\left[\frac{\mathrm{e}^{-s\theta}}{\Gamma(\frac{1}{2}+\mathrm{i}s)^{1/2}} - \frac{\mathrm{e}^{s\theta}}{\Gamma(\frac{1}{2}-\mathrm{i}s)^{1/2}}\right]\mathrm{d}s. \tag{3.13}$$

By applying Stirling's formula to this representation, we find that

$$(U\Phi_R)(\mathrm{e}^{\mathrm{i}\theta}) = \mathrm{e}^{-\frac{1}{4}R^2}\int_0^{\infty}\left(\frac{R}{\sqrt{2}}\right)^t\frac{\mathrm{e}^{\mathrm{i}t\theta}}{\Gamma(t+1)^{1/2}}\,\mathrm{d}t + o(\mathrm{e}^{-\frac{1}{4}R^2}) \tag{3.14}$$

as $R\to\infty$, uniformly for $\theta\in[-\pi,\pi]$.

It is from this integral representation that we shall extract the asymptotic expression of interest to us. As we do not know of any standard expression for $1/\Gamma(z)^{1/2}$, we must proceed directly. To that end we consider the function

$$F(t) = t\log\left(\frac{R}{\sqrt{2}}\right) - \frac{1}{2}\log[\Gamma(t+1)] \qquad t>0,\ R>0. \tag{3.15}$$

Then

$$F'(t) = \log\left(\frac{R}{\sqrt{2}}\right) - \frac{1}{2}\psi(t+1) \tag{3.16}$$

where ψ is the digamma function, the logarithmic derivative of the gamma function.

Let $T = T(R)$ be the value of t where F' vanishes, $F'(T) = 0$; since

$$F'(t) \text{ is } \begin{cases} > 0 & \text{if } 0 < t < T \\ < 0 & \text{if } t > T \end{cases}$$

there is one and only one such point. Implicitly,

$$\psi(T+1) = 2\log\left(\frac{R}{\sqrt{2}}\right). \tag{3.17a}$$

We can solve this for T to order $O(R^{-2})$ by working from

$$\psi(R^2 + \tfrac{1}{2}) = 2\log R + O(R^{-4})$$

from which we find, after applying the intermediate value theorem, that

$$T(R) = \tfrac{1}{2}(R^2 - 1) + O(R^{-2}). \tag{3.17b}$$

In order to obtain uniform bounds, we are going to have to change variables in the integral, in a way which depends on F near T. We define the constant ($F(0) = 0$ and $F(T) > 0$)

$$A = \sqrt{F(T)} \tag{3.18a}$$

from which we find that

$$e^{A^2} = \frac{e^{R^2/4}}{\pi^{1/4} R^{1/2}}\left[1 + O\left(\frac{\log R}{R^2}\right)\right]. \tag{3.18b}$$

The change of variables we want is from t to w, where

$$w = A + \mathrm{Sign}(t - T)\sqrt{A^2 - F(t)}. \tag{3.19a}$$

Then

$$\theta\int_0^\infty \left(\frac{R}{\sqrt{2}}\right)^t \frac{e^{it\theta}}{\Gamma(t+1)^{1/2}}\,dt = \theta\int_0^\infty e^{F(t)+it\theta}\,dt$$
$$= \theta e^{A^2}\int_0^\infty e^{-(w-A)^2 + it(w)\theta}\,t'(w)\,dw$$

integrating by parts, with

$$t'(w)e^{it(w)\theta} = \frac{1}{i\theta}\frac{d}{dw}e^{it(w)\theta}$$

gives

$$\theta\int_0^\infty \left(\frac{R}{\sqrt{2}}\right)^t \frac{e^{it\theta}}{\Gamma(t+1)^{1/2}}\,dt = i + \frac{2e^{A^2}}{i}\int_0^\infty (w - A)e^{-(w-A)^2 + it(w)\theta}\,dw$$

so that

$$\theta(U\Phi_R)(e^{i\theta}) = \frac{2e^{A^2-\frac{1}{4}R^2}}{i}\int_0^\infty (w-A)e^{-(w-A)^2+it(w)\theta}\,dw + O(e^{-\frac{1}{4}R^2})$$

as $R \to \infty$, uniformly in $\theta \in (-\pi, \pi)$.

We can find $0 < \eta < 1$ such that $T(w)$ is well approximated in the interval $[\eta A, 2A]$ by

$$\begin{aligned}\hat{t}(w) &= t(A) + t'(A)(w-A) + \tfrac{1}{2}t''(A)(w-A)^2 \\ &= T + \frac{2}{\sqrt{\psi'(T+1)}}(w-A) - \frac{2}{3}\frac{\psi''(T+1)}{\psi'(T+1)^2}(w-a)^2 \qquad \eta A \leqslant w \leqslant 2A\end{aligned}$$

in the sense that there exists a constant $K > 0$ such that

$$|\hat{t}(w) - t(w)| \leqslant \frac{K}{R}(1+|w-A|) \qquad \eta A \leqslant w \leqslant 2A$$

from which it follows that

$$\theta(U\Phi_R)(e^{i\theta}) = \frac{2e^{A^2-\frac{1}{4}R^2}}{i}\int_{-\infty}^\infty (w-A)e^{-(w-A)^2+i\hat{t}(w)\theta}\,dw + O(R^{-3/2})$$

as $R \to \infty$, uniformly in $\theta \in (-\pi, \pi)$.

This last integral can be evaluated explicitly, giving

$$\theta(U\Phi_R)(e^{i\theta}) = \frac{\pi^{1/2}t'(A)e^{A^2-\frac{1}{4}R^2}\theta}{(1-\frac{i}{2}t''(A)\theta)^{3/2}}\exp\left[iT\theta - \frac{t'(A)^2\theta^2}{4(1-\frac{i}{2}t''(A)\theta)}\right] + O(R^{-3/2})$$

as $R \to \infty$, uniformly in $\theta \in (-\pi, \pi)$.

The expression on the right can now be squared and integrated over θ:

$$\begin{aligned}\frac{1}{2\pi}&\int_{-\pi}^{+\pi}\left|\frac{\pi^{1/2}t'(A)e^{A^2-\frac{1}{4}R^2}\theta}{(1-\frac{i}{2}t''(A)\theta)^{3/2}}\exp\left[iT\theta - \frac{t'(A)^2\theta^2}{4(1-\frac{i}{2}t''(A)\theta)}\right]\right| d\theta \\ &= \frac{t'(A)^2e^{2A^2-\frac{1}{2}R^2}}{2}\int_{-\pi}^{+\pi}\frac{\theta^2}{(1+\frac{1}{4}t''(A)^2\theta^2)^{3/2}}\exp\left[-\frac{t'(A)^2\theta^2}{2[1+\frac{1}{4}t''(A)^2\theta^2]}\right] d\theta \\ &\sim \frac{t'(A)^2e^{2A^2-\frac{1}{2}R^2}}{2}\int_{-\infty}^{\infty}\theta^2\exp[-\tfrac{1}{2}t'(A)^2\theta^2]\,d\theta \\ &= \frac{e^{2A^2-\frac{1}{2}R^2}}{2t'(A)}\int_{-\infty}^{\infty}\theta^2e^{-\frac{1}{2}\theta^2}\,d\theta \\ &= \sqrt{\frac{\pi}{2}}\frac{e^{2A^2-\frac{1}{2}R^2}}{t'(A)} \\ &= \sqrt{\frac{\pi}{8}}\psi'(T+1)^{1/2}e^{2A^2-\frac{1}{2}R^2} \\ &\sim \frac{1}{2R^2}.\end{aligned}$$

Therefore,

$$\mathrm{V_{BP}}(\Phi_R) = \frac{1}{2\pi}\int_{-\pi}^{+\pi} \theta^2 |(U\Phi_R)(\mathrm{e}^{\mathrm{i}\theta})|^2 \,\mathrm{d}\theta \doteq \frac{1}{2R^2} + O\left(\frac{1}{R^3}\right) \tag{3.20}$$

as $R \to \infty$. This justifies the statements found in the literature concerning the asymptotic form of the BP 'variance' in the state Φ_R for large R.

Thus while it is certainly true that as R tends to infinity we have

$$\mathrm{V_{BP}}(\Phi_R) \sim \frac{1}{2R^2} \tag{3.21a}$$

$$\mathrm{var}[X;\, \Phi_R] = O\left(\frac{1}{R^2}\right) \tag{3.21b}$$

$$\mathrm{var}[\Delta(\varphi);\, \Phi_R] \sim \frac{\pi^{1/2}}{R} \tag{3.21c}$$

and so the behaviour of $\Delta(\varphi)$ is substantially different from that of X and the BP formulation, we note that it seems likely that the BP result comes from restricting attention to a wedge around $\varphi = 0$ in phase space, namely that region which is 'furthest away' from the cut in the plane.

It has been noted elsewhere [14, 15] that the fact that N and $\Delta(\varphi)$ are not canonically conjugate, indeed, the fact that their classical counterparts, the phase space functions $\frac{1}{2}(r^2 - 1)$ and φ are not canonically conjugate with respect to the Poisson bracket, comes from the effect of the cut in the phase plane. Consequently, N and $\Delta(\varphi)$, correspondingly $\frac{1}{2}(r^2 - 1)$ and φ, are in some sense canonically conjugate in the region of the plane near $\varphi = 0$, and so it is only to be expected that modifying the functions and operators to eliminate the influence of the region of the cut will result in $1/(2R^2)$ for the asymptotic behaviour of the variance for any of the modified operators. However, we have seen [18] that the large asymptotic behaviour of the variance of $\Delta(\varphi)$ in the state Φ_R comes precisely from the behaviour of φ near the cut.

Noting the identites

$$|\langle \chi_s(\theta_{s,j}), h_n\rangle|^2 = \frac{1}{s+1} \qquad 0 \leqslant j \leqslant s,\ s \geqslant n \tag{3.22}$$

Barnett, Pegg and their collaborators argue that, together with the $1/(2R^2)$ variance result, the collection of operators (X_s) represents an approximation to a probability distribution for the phase angle which is uniform on the interval $[\theta_0, \theta_0 + 2\pi)$, and this, it is claimed, is supported by experiment.

Regarding experiments, it is not clear to us, nor to the experimenters as far as we can tell, whether they are measuring a phase operator or some function of it. As far as we can tell, the experiments in question seem to measure observables representing quantized cosine and sine operators rather than the angle operator directly. Therefore, the experimental results should be compared with the variances of $\Delta(\cos\varphi)$ and $\Delta(\sin\varphi)$ in various states. If this is done, then the Weyl quantization produces the expected results [18], and there is no longer any contradiction. Moreover, it is not clear that the experiments have been refined enough to distinguish between exact variances and their asymptotic limits. In addition, all currently available experimental data comes in the region $R < 10$, and nothing is known for really large values of R. We believe that direct comparison with experiment to select one phase operator proposal from the others is not yet possible.

We must also say that it seems to us that these semiclassical heuristics may not do justice to the rich range of phenomena being considered. In particular, they may omit the possibility of new and interesting quantum effects, thereby reducing the predictive value of the theory.

We have stated our objections to the procedures applied to the sequences $[\chi_s(\theta_{s,j})]$ and (X_s) in the BP theory: were they within the purview of quantum theory, they would obtain no more than the theory of the Toeplitz phase operator, which we feel to be of less obvious physical significance than the quantization of the angle function in any event. Hence, the manner in which they do proceed constitutes an *ad hoc* scheme which is supposed to generalize quantum theory. Yet the fact that it does not subsume quantum theory calls into question its acceptability as a correct physical theory.

4. Measuring operators with a continuous spectrum

In this section we shall formulate the phase theory of Barnett and Pegg within the usual framework of quantum theory. Barnett and Pegg themselves do not do so. Paraphrasing their reasons, as given in the paper by Barnett and Dalton [7], they believe that phase is an attribute of a sort different from those ordinarily considered in quantum theory, and cannot be described by a single self-adjoint operator as can, say, position. They even state that 'the nature of phase cannot be determined via experimental test!' We feel that even if phase does require a number of observables to fully describe it, this is in no way incompatible with quantum theory. All quantum theory demands is that we provide a means of measuring quantities we call observables. What seems to be the case is that a multiplicity of associated, but different, quantum qualities coalesce into a single classical property we call phase. It may even be more complicated than this, since relative and absolute phase may be considered as different in a certain sense. Whatever the results will turn out to be, as determined by experiments, we must emphasize that none of this requires modifying quantum theory. Nor, indeed, do we know of any experimental results that do.

The theory proposed by Barnett and Pegg makes use of the operators X_s. Whereas X_s is viewed as an operator on the space $\mathcal{H}_s$ spanned by the first $s+1$ Hermite functions in BP theory, we shall consider every X_s as an operator on $\mathcal{L}^2(\mathbb{R})$. As such they are self-adjoint operators whose spectrum consists of non-degenerate eigenvalues, and so they are observable. It is precisely this consideration which will enable us to recast their work as part of quantum mechanics. There are some unusual features which result, and these require a particular treatment of quantum measurement theory. Our plan is to consider the necessary form of measurement theory as a general proposition in this section, and then to apply it to the BP theory in section 7, after we have obtained certain additional mathematical results in sections 5 and 6.

We will assume the reader to be familiar with standard quantum measurement theory for self-adjoint and bounded operators whose spectrum consists wholly of eigenvalues. Very briefly, if A is such an observable, it can be written as

$$A = \sum_{j\geqslant 1} \lambda_j P_j \tag{4.1}$$

where $(\lambda_j)_{j\geqslant 1}$ are its eigenvalues and $(P_j)_{j\geqslant 1}$ are the corresponding eigenprojections. A measurement of A can result in a registration of one of its eigenvalues only, and the value

λ_j results in the collapse of the input state $\psi_{\rm in}$ to the output state

$$\psi_{\rm out} = \frac{P_j \psi_{\rm in}}{\|P_j \psi_{\rm in}\|}. \tag{4.2}$$

In the special case where λ_j is non-degenerate, the eigenspace is spanned by a normalized vector, say φ_j, unique up to a phase factor.

For definiteness, we shall consider this theory as realizing an arrangement where a beam is incident on a measuring device—we use the terms apparatus and instrument as synonymous with device. The beam is in the state $\psi_{\rm in}$, and the device is completely defined by an observable such as A. To distinguish this example from the others we need, we shall call operators such as A ideal observables and the instruments they represent as ideal instruments.

All of the operators associated with phase that we have encountered are different in character from this, having a continuous component to their spectra. For definiteness, we shall consider the simplest such case, where A is a bounded self-adjoint operator whose spectrum is absolutely continuous and multiplicity free: the Toeplitz operator X is a good example. We shall write $\sigma(A)$ for the spectrum of A, which we assume to be equal to the real closed and bounded interval $[a, b]$.

No measuring device can be constructed which will precisely measure A. The two sources of error are (1) the continuity of the spectrum and (2) the absence of eigenvectors to use as output states. The consequence is that A must be measured by an ideal device which is as close to it as possible. Closeness must be defined here, and we take it to mean accuracy with respect to the two sources of error just noted.

These two points form the basis of our understanding of the theory of Barnett and Pegg, and so it is essential to what follows to understand what they imply. Elaborating point (1), we observe that there is an ineradicable inaccuracy inherent in the response of any instrument: small though it may be, it is always present. The result is, as Kemble [25] states categorically in his 1937 text on Quantum Mechanics: 'exact predictive measurements of continuous spectrum eigenvalues are fundamentally impossible. Experience shows that exact restrospective measurement of such eigenvalues are equally impossible ... we assume that inexact measurements which conform (*to the exact measurements possible for observables with a discrete spectrum*) to an arbitrarily high degree of approximation are to take their place.'

Let us quantify the inaccuracy of an instrument by a (strictly) positive number, call it the tolerance, measured in the same units as the observable we are using it to measure. The instrument cannot distinguish between spectral values which differ by less than the tolerance. Now were we measuring an observable with a discrete spectrum, this would not be an important concern, as we should only need an instrument whose tolerance was small compared to the distance between neighbouring eigenvalues. A registered spectral value would then be closest to one, and only one, eigenvalue, and the error in assigning the registration value to that eigenvalue would be insignificant.

But no matter how small the tolerance, an observable with a continuous spectrum has a continuum of spectral values within the tolerance interval around any registration value, and so the device cannot distinguish between those values. (This is also true of limit points of discrete spectra, such as the ionization energy of a hydrogen atom; such points are treated in the same way as the continuous spectrum.)

It follows that any instrument effectively partitions the continuous spectrum into intervals determined by the tolerance. The fact that all spectral points within such an interval are indistinguishable by the instrument is physically equivalent to associating the instrument with an observable with a discrete spectrum: call it the instrument observable. Each eigenvalue of the instrument observable corresponds to that continuum of spectral points of the observable we wish to measure lying in the appropriate interval. This effective replacement of an observable with a continuous spectrum by one with a discrete spectrum is inescapable.

Summing up, the existence of a non-zero tolerance in any measurement device means that it is represented by a device observable with a discrete spectrum. Hence, an operator with a continuous spectrum necessarily differs from any device observable, and it is actually the latter which is measured. It follows that we must have means of constructing instrument operators whose differences from the operators to be measured is small in some appropriate sense.

Fortunately, this first source of error, spectral readings, can be dealt with in a satisfactory manner by using the construction of the Riemann integral as a model. By a partition π_s of the spectrum of A we shall mean a finite ordered set

$$\pi_s = \{a = x_0 < x_1 < \cdots < x_{s+1} = b\}. \tag{4.3a}$$

The closed intervals

$$I_{s,j} = [x_{s,j}, x_{s,j+1}] \qquad j = 0, 1, \ldots, s \tag{4.3b}$$

so defined we call the partition intervals, and they are of length

$$|I_{s,j}| = x_{s,j+1} - x_{s,j}. \tag{4.3c}$$

These lengths define the partition norm, as it is called in integration theory,

$$\|\pi_s\| = \max_{0\leqslant j\leqslant s} |I_{s,j}|. \tag{4.3d}$$

Along with the partition π_s we are free to choose a distinguished point $\lambda_{s,j} \in I_{s,j}$ in each interval; this is analogous to choosing a value of the function to be integrated in each interval and using it in the Riemann sum. These values, $\lambda_{s,j}$, will be the eigenvalues of the ideal measurement device under construction.

For each spectral value λ of A, there is a nearest eigenvalue $\lambda_{s,j}$ to λ (possibly there are two such), and the spectral value λ will be registered as $\lambda_{s,j}$. This is an error, but is not greater than the corresponding length $|\lambda_{s,j}|$. If we supplement the partition π_s with a corresponding set of output states—which we shall do below—we obtain an ideal device to measure A that we shall denote by A_s. From what we have just said, it is obvious that whatever the spectral value that would have registered were perfect A-devices possible, the error obtained by using A_s will be no greater than $\|\pi_s\|$.

We now see that it is appropriate to consider a sequence (A_s) of devices, constructed through a sequence of partitions (π_s) with partition norms converging to zero:

$$\lim_{s\to\infty} \|\pi_s\| = 0. \tag{4.4}$$

The physical significance of this is that if you require a device to measure A with a spectral accuracy no worse than some strictly positive tolerance ε, you need only choose a large enough value of s, depending on ε, and all the A_t with $t > s$ will have this property. For this reason we refer to a sequence of partitions whose norms converge to 0 as satisfying the spectral accuracy condition. Note that we are free to choose the $\lambda_{s,j}$ as we will, just as in integration theory. Our conclusion is that the sensitivity to spectral readings for observables such as A can be dealt with by an effective approximation process. As a matter of terminology we refer to the sequence (A_s) of ideal observables as a measurement system, and will assume the spectral accuracy condition as a matter of course.

The second source of inaccuracy in constructing devices to measure A is the choice of output states. For a system of measurement (A_s), this is equivalent to a choice of projection operators $P_{s,j}$ for each s, so that

$$A_s = \sum_{j=0}^{s} \lambda_{s,j} P_{s,j}. \tag{4.5}$$

The only general condition required of these projections is that

$$P_{s,j} P_{s,k} = 0 \qquad \text{if } j \neq k \tag{4.6}$$

so that A_s is self-adjoint. Of course without further conditions, such a system is of only marginal interest, since there is only a very weak relation between the A_s and A. Spectral accuracy is not a very strong recommendation for a measurement system. For example, the spectra of position and momentum are the same, and one would not think much of measuring the momentum of a particle by registering only its position.

The sorts of conditions that will provide a more acceptable system are those that cause the A_s to converge to A in some appropriate fashion. We take this to mean in the weak, strong or uniform operator topologies. However, weak convergence is hardly acceptable, owing to the fact that if we have sequences of bounded self-adjoint operators (K_n), (L_n) converging weakly to K, L, respectively, then $(K_n L_n)$ need not converge even weakly; and if it does, its limit need not be KL. This cannot happen for strong (hence uniform convergence *a fortiori*), and so without strong convergence we cannot make sense of variance and higher moment calculations.

Our treatment of the theory of Barnett and Pegg is based on taking the LHW states $\chi_s(\theta)$ as output states for a system of measurement. The instrument observables then turn out to be the operators (X_s). It is the following property of these states that provides some justification for choosing them as output states.

Proposition 4.1. For a bounded self-adjoint operator A on a Hilbert space $\mathcal{H}$, a point λ is in its spectrum, $\sigma(A)$, if and only if there is a sequence $(\psi_s^{(\lambda)})$ of normalized vectors in $\mathcal{H}$ such that

$$\lim_{s\to\infty} \|(A-\lambda)\psi_s^{(\lambda)}\| = 0. \tag{4.7}$$

Then $(\psi_s^{(\lambda)})_{s\geqslant 0}$ is said to be a sequence of approximating eigenvectors for λ (SAE for short) [26].

To make a connection with the more familiar notion of an eigenvector, note that if λ should happen to be an eigenvalue of A and ψ a corresponding eigenvector, the sequence $\psi_s^{(\lambda)} = \psi$ is an SAE for the point λ.

An SAE is anything but unique. For one thing, there are many different SAE for a given spectral point λ of A. For another, a given sequence of normalized vectors can be an SAE for many different operators. In particular this is true for the LHW states: in sections 5 and 6 we shall prove that they are SAE for many of the operators considered in phase theory. Granting this, we may expect that instrument systems based on them will not have good convergence properties, as a system can converge to at most one observable.

An SAE is defined as a sequence of vectors associated with a single spectral point, and there need not be any connection between an SAE for one spectral value and an SAE for another. The LHW states $\chi_s(\theta)$, however, are related for different values θ by virtue of the relation

$$\chi_s(\theta) = \mathrm{e}^{\mathrm{i}\theta N}\chi_s(0). \tag{4.8}$$

Examination of the consequences of this leads us to consider families of SAE related across the spectrum of a given operator in the following way.

The spectral uniformity condition (4.2). Let A be a bounded normal operator on $L^2(\mathbb{R})$, and let

$$\sigma_0(A) = \sigma(A) \setminus \{\lambda_1, \ldots, \lambda_N\} \tag{4.9}$$

be the indicated subset of its spectrum; the points omitted are said to be exceptional. Given a family $(\psi_s^{(\lambda)})_{s,\lambda}$ of SAEs for A, define

$$\varepsilon_{s,K} = \sup_{\lambda\in K} \|(A-\lambda)\psi_s^{(\lambda)}\| \tag{4.10}$$

for any subset K of the spectrum $\sigma(A)$.

We shall say that the family $(\psi_s^{(\lambda)})_{s,\lambda}$ satisfies the spectral uniformity condition if

$$\lim_{s\to\infty} \varepsilon_{s,K} = 0 \tag{4.11}$$

for one of the increasingly general options for the possible range of choices for the subset K:

(i) $K = \sigma(A)$;
(ii) K can be any compact subset of $\sigma_0(A)$;
(iii) K can be any compact subset of any open subset of $\sigma_0(A)$ which is dense in $\sigma(A)$.

Anticipating the results of the remaining sections, we shall see that option (i) will hold for Y and $\Delta(\mathrm{e}^{\mathrm{i}\varphi})$, (ii) will hold for X, and (iii) will hold for $\Delta(\varphi)$. The exceptional points for X are $\pm\pi$, and for $\Delta(\varphi)$ they are $\pm\pi$ and 0. The need to exclude $\pm\pi$ for X and $\Delta(\varphi)$ is not at all surprising, nor is the fact that we do not need to make a similar exclusion for Y and $\Delta(\mathrm{e}^{\mathrm{i}\varphi})$. The point is that the LHW states $\chi_s(\theta)$ are periodic functions of θ, and the spectrum of Y and $\Delta(\mathrm{e}^{\mathrm{i}\varphi})$ is $\mathbb{T}$. In contrast, the spectrum of X is the interval $[-\pi, \pi]$, whose endpoints correspond to one and the same point on $\mathbb{T}$. Since there is a discontinuity in the spectrum which is not matched by a discontinuity in the output states, it is not surprising that the mathematics requires us to avoid this part of the spectrum. This is another curious effect of the cut plane.

We have *conjectured that the spectrum of $\Delta(\varphi)$ is the same as that of X. For this* reason, the fact that we need to exclude the spectral points $\pm\pi$ when considering $\Delta(\varphi)$ follows from the same arguments used above for X. The fact that we must also exclude the spectral point 0 in considering $\Delta(\varphi)$ is curious, and we can find no reason for it.

We use these conditions as follows. Returning to our self-adjoint operator A with spectrum $[a, b]$, choose a family (π_s) of partitions satisfying the spectral accuracy condition. Choose next a family of SAE satisfying the spectral uniformity condition. Here we have a choice of excluding exceptional points or not. If we include a point we should have excluded, the result will be that the instrument system we are constructing will work less well in the neighbourhood of the point.

Now write

$$\psi_{s,j} = \psi_s^{(\lambda_{s,j})}. \tag{4.12a}$$

If—and only if—the orthonormality conditions

$$\langle \psi_{s,j}, \psi_{s,k} \rangle = \delta_{j,k} \qquad 0 \leqslant j, k \leqslant s \tag{4.12b}$$

hold, we may use the projections $P_{s,j}$ defined by these vectors,

$$P_{s,j} f = \langle \psi_{s,j}, f \rangle \psi_{s,j} \qquad 0 \leqslant j \leqslant s, \ s \geqslant 0 \tag{4.12c}$$

to define the A_s by the usual formula:

$$A_s = \sum_{j=0}^{s} \lambda_{s,j} P_{s,j}.$$

Thus, if quantity to be measured 'is' within the interval $I_{s,j}$, the value $\lambda_{s,j}$ will be registered and the output state will be $\psi_{s,j}$. We shall refer to the sequence (A_s) constructed in this way as an SAE system.

In general, while appearing fairly restrictive, the condition of spectral uniformity is not strong enough for the resulting SAE system (A_s) to converge to A in any sense. It is clearly of interest to know whether there might be further conditions which do guarantee convergence. The following conditions, while still not strong enough, are found in the BP theory and are felt by some workers to be motivated by physical considerations in some circumstances.

The ascending subspace condition (4.3). Let $(\psi_{s,j})_{0 \leqslant j \leqslant s; s \geqslant 0}$ be a spectrally uniform family for A. Write $\mathcal{H}_s$ for the $(s+1)$-dimensional Hilbert space $\mathcal{H}_s$ spanned by the $(\psi_{s,j})_{0 \leqslant j \leqslant s}$. The family is said to satisfy the ascending subspace condition if $\mathcal{H}_s$ is a subspace of $\mathcal{H}_{s+1}$ for all s, and the inner product space spanned by the $\psi_{s,j}$ for all j and s, $\mathcal{H}$, is dense in $L^2(\mathbb{R})$.

For the applications we have in mind, it would be a spurious generality to assume that $\mathcal{H}$ was not dense.

In this description, the vectors $\psi_{s,j}$ are given and the subspaces $\mathcal{H}_s$ are formed from them. Alternatively, one might specify an ascending family of Hilbert subspaces and then seek a basis for each of them so as to constitute a spectrally uniform SAE family for A. Clearly it might be difficult to find such vectors; it might even be impossible. The difficulty is increased if it is further required that the vectors sought have additional specified properties. We shall consider this again in section 7.

We have found a condition which does ensure strong convergence for an SAE system. While theoretically interesting, it does not apply to any of the operators we are considering for the SAE system constructed from LHW states.

Proposition 4.4. Suppose there is given an SAE system (A_s) for A, where we may take K *to be* $\sigma(A)$, *and satisfying the ascending subspace condition.* Then *a sufficient condition* that system (A_s) converges strongly to A is that

$$\lim_{s\to\infty}(s+1)^{1/2}\varepsilon_s = 0. \tag{4.13}$$

Proof. For any $f \in \mathcal{H}$ there is a t for which $f \in \mathcal{H}_t$. For any $s \geqslant t$ we can write

$$f = \sum_{j=0}^{s} f_{s,j}\psi_{s,j}. \tag{4.14}$$

Then

$$(A - A_s)f = \sum_{j=0}^{s} f_{s,j}(A - \lambda_{s,j})\psi_{s,j}. \tag{4.15}$$

Using the Cauchy–Schwarz inequality,

$$\begin{aligned}
\|(A - A_s)f\| &\leqslant \left(\sum_{j=0}^{s}|f_{s,j}|^2\right)^{1/2}\left(\sum_{j=0}^{s}\|(A-\lambda_{s,j})\psi_{s,j}\|^2\right)^{1/2} \\
&\leqslant \|f\|\left(\sum_{j=0}^{s}\varepsilon_s^2\right)^{1/2} \\
&= (s+1)^{1/2}\varepsilon_s\|f\|.
\end{aligned}$$

With the given condition on the (ε_s), the result is seen to be true for any $f \in \mathcal{H}$.

The family $(\|A_s\|)_s$ is uniformly bounded. This most accessible way to see this is by writing

$$A_s = \sum_{j=0}^{s}\lambda_{s,j}P_{s,j}$$

so that for any $g \in L^2(\mathbb{R})$,

$$\|A_s g\|^2 = \sum_{j=0}^{s}|\lambda_{s,j}|^2|\langle\psi_{s,j}, g\rangle|^2.$$

As the spectrum of A is $[a, b]$,

$$\|A_s g\|^2 \leqslant (\max\{|a|, |b|\})^2\sum_{j=0}^{s}|\langle\psi_{s,j}, g\rangle|^2 \leqslant (\max\{|a|, |b|\})^2\|g\|^2.$$

Thus

$$\sup_s \|A_s\| \leqslant \max\{|a|, |b|\}.$$

Alternatively, as $\mathcal{H}$ is a norm determining subspace for bounded operators on $L^2(\mathbb{R})$, this result follows from the uniform boundedness theorem. It is a standard argument that under these circumstances, the convergence extends from $\mathcal{H}$ to $L^2(\mathbb{R})$ and we omit it. □

With a little more work of an essentially book-keeping nature, we could extend this to the more general choices for K, but as there still would be no applications to BP theory, we shall not do so.

Given any ascending sequence of subspaces $\mathcal{H}_s$ whose union is dense in $L^2(\mathbb{R})$, we can construct a measurement system which converges strongly to A, but not by using the notion of an SAE. The mathematical construction is quite simple.

The restriction model (4.5). Let $(\mathcal{K}_s)$ be a family of Hilbert subspaces of $L^2(\mathbb{R})$, nested in the sense that $\mathcal{K}_s \subset \mathcal{K}_{s+1}$ for all s, and such that their union is dense in $L^2(\mathbb{R})$. Let $P^{(s)}$ be the associated family of projection operators. Define

$$A^{(s)} = P^{(s)} A P^{(s)}. \tag{4.16}$$

As $P^{(s)}$ converges strongly—but not uniformly—to the identity, and as strong convergence is respected by products of bounded operators,

$$\lim_{s\to\infty} A^{(s)} = \lim_{s\to\infty} P^{(s)} A \lim_{s\to\infty} P^{(s)} = A \qquad \text{(strongly).} \tag{4.17}$$

Thus, the system $(A^{(s)})$ constructed in this way converges strongly to A.

In particular, were we to be given an SAE system (A_s) for A, say with $\mathcal{K} = \sigma(A)$, and satisfying the ascending subspace condition, we could choose $\mathcal{K}_s = \mathcal{H}_s$, enabling us to compare the convergence properties of (A_s) and $(A^{(s)})$.

Naturally it is of interest to know whether or not a slightly weaker condition on the ε_s would lead to weak convergence. The examples found in BP theory suggest that no such theorem exists, since each LHW-based SAE system may be associated with more than one operator. There is no conflict with our convergence theorem since we shall prove that the fall-off rates are not as fast as $o(s^{-1/2})$. The most we can say is that since the sequence (A_s) is uniformly bounded, it has weak limit points. And since $L^2(\mathbb{R})$ is separable, any weak limit point is the limit of a subsequence, as opposed to a subnet. This seems to be as far as this line of analysis will go.

The following construction shows that essential properties of an SAE system do not depend on the choice of the $\lambda_{s,j}$. To make things general, we suppose that the spectral uniformity condition of type (iii) holds. Then let U be an open subset of $\sigma_0(A)$ which is dense in $\sigma(A)$. If K is any compact subset of U, we can find a compact subset L of U such that K is contained in the interior of L. For any $\lambda \in \sigma(A)$, let $J(s,\lambda)$ be such that $\lambda_{s,J(s,\lambda)}$ is the eigenvalue of A_s nearest to λ. We can find an $s(K)$ such that $\lambda_{s,J(s,\lambda)} \in L$ for all $\lambda \in K$ and all $s \geqslant s(K)$. Since

$$\|(A-\lambda)\psi_{s,J(s,\lambda)}\| \leqslant \|(A-\lambda_{s,J(s,\lambda)})\psi_{s,J(s,\lambda)}\| + \|(\lambda-\lambda_{s,J(s,\lambda)})\psi_{s,J(s,\lambda)}\|$$

we deduce that

$$\|(A-\lambda)\psi_{s,J(s,\lambda)}\| \leqslant \varepsilon(s,L) + \|\pi_s\| \tag{4.18}$$

for all $\lambda \in K$ and $s \geqslant s(K)$. We interpret this statement as extending the concept of spectral accuracy beyond the specified registration values $\lambda_{s,j}$ to general spectral values $\lambda \in \sigma(A)$.

Our analysis has shown that SAE systems are of only limited value as measuring systems. We have constructed a strongly convergent system, the restriction model, without very much difficulty. To underpin our discussion we note that it is even possible to find a system (A_s) which converges uniformly to A, provided we know the spectral decomposition of A.

The spectral model (4.6). Suppose we know the spectral decomposition of the observable we wish to measure

$$A = \int_I \lambda \, \mathrm{d}P_\lambda. \tag{4.19}$$

Now simply choose the output eigenprojections to be

$$P_{s,j} = P(x_{s,j+1}) - P(x_{s,j}) = P(I_{s,j}) \tag{4.20}$$

and define

$$\hat{A}_s = \sum_{j=0}^{s} \lambda_{s,j} P_{s,j}.$$

The fact that

$$\lim_{s\to\infty} \|A - \hat{A}_s\| \leqslant \|\pi_s\| \tag{4.21a}$$

hence that we have the uniform convergence limit

$$\lim_{s\to\infty} \|A - \hat{A}_s\| = 0 \tag{4.21b}$$

is now a special case of a standard theorem in operator theory.

In the next two sections we shall prove that the LHW states $\chi_s(\theta)$ form an SAE for the various operators of interest. We shall do this in such a way that we can determine the $\varepsilon_{s,K}$ in the different cases. It will be seen that none falls off sufficiently rapidly to apply the above theorem, as we said. In section 7 we shall combine these results with the SAE system construction given here to analyse certain aspects of BP theory.

5. The Toeplitz phase operator

The operator obtained by Weyl quantization of the phase function, $\Delta(\varphi)$, was constructed using a definition of the angle complementary to that generally found in the literature. For consistency with this convention we shall define the Toeplitz phase operator to be the bounded self-adjoint operator $X = X^*$ on $L^2(\mathbb{R})$ given by the rule obtained by linear *continuous extension from*

$$Xh_n = \sum_{m\geqslant 0, m\neq n} \frac{\mathrm{i}^{n-m+1}}{m-n} h_m \qquad n \in \mathbb{N} \bigcup \{0\} \tag{5.1}$$

where h_n is the nth Hermite function.

The Hermite matrix elements of X are obtained from those of $\Delta(\varphi)$ by replacing the factors $g_{m,n}$ by unity [14, 15]. It should be noted that our definition of X assumes values for the angle function lying in $[-\pi, \pi)$, and in what follows we shall restrict our attention to values of θ lying in that range unless otherwise noted.

The Toeplitz shift operators may be obtained from the operators $\Delta(e^{i\varphi})$ and $\Delta(e^{-i\varphi}) = \Delta(e^{i\varphi})^*$ by using the same conventions, i.e. taking the $g_{m,n} = 1$ and the fixed angle θ_0 equal to $-\pi$, giving

$$Yh_n = ih_{n+1} \qquad n \in \mathbb{N} \bigcup \{0\} \tag{5.2a}$$

and

$$Y^* h_n = \begin{cases} -ih_{n-1} & \text{if } n \in \mathbb{N} \\ 0 & \text{if } n = 0. \end{cases} \tag{5.2b}$$

For consistency, we must also redefine the LHW vectors by setting

$$\chi_s(\theta) = \frac{1}{\sqrt{s+1}} \sum_{n=0}^{s} i^n e^{-in\theta} h_n \qquad s \in \mathbb{N},\ \theta \in \mathbb{R}. \tag{5.3}$$

The 'maximally random' property of the $\chi_s(\theta)$ with respect to the number operator is represented by the fact that

$$|\langle \chi_s(\theta), h_n \rangle|^2 = \frac{1}{s+1} \qquad 0 \leqslant n \leqslant s,\ \theta \in \mathbb{R}. \tag{5.4}$$

We begin our analysis proper by determining the relation between the $\chi_s(\theta)$ and the shift operators Y and Y^*. (The shift operator Y differs from the customary unilateral shift operator by a factor of i.)

The spectral decompositions of Y and Y^* are well known. By mapping $L^2(\mathbb{R})$ onto the sequence space l^2, they are favourite examples in operator theory [22]. For example, the spectrum of Y is the closed unit disc $\overline{\mathbb{D}}$ in the complex plane and its eigenvalues are the points of the open disc. The boundary of the disc, the unit circle $\mathbb{T}$, is then the continuous spectrum. What is of interest to us is the relation of the points of the unit circle with the vectors $\chi_s(\theta)$.

Proposition 5.1. The set

$$\{\chi_s(\theta) : s \in \mathbb{N}\}$$

of LHW states satisfies the s-equalities

$$\|(Y - e^{i\theta})\chi_s(\theta)\|^2 = \frac{2}{s+1} \tag{5.5a}$$

and

$$\|(Y^* - e^{-i\theta})\chi_s(\theta)\|^2 = \frac{1}{s+1}. \tag{5.5b}$$

Hence they constitute an approximating sequence of unit vectors for the spectral value $e^{i\theta}$ of Y, and for the spectral value $e^{-i\theta}$ of Y^*.

Proof. Using the definitions

$$Y\chi_s(\theta) = \frac{1}{\sqrt{s+1}}\sum_{n=0}^{s} \mathrm{i}^{n+1}\mathrm{e}^{-\mathrm{i}n\theta}h_{n+1}$$
$$= \frac{\mathrm{e}^{\mathrm{i}\theta}}{\sqrt{s+1}}\sum_{n=1}^{s+1} \mathrm{i}^{n}\mathrm{e}^{-\mathrm{i}n\theta}h_{n}$$

and so

$$(Y - \mathrm{e}^{\mathrm{i}\theta})\chi_s(\theta) = \frac{\mathrm{e}^{\mathrm{i}\theta}}{\sqrt{s+1}}[\mathrm{i}^{s+1}\mathrm{e}^{-\mathrm{i}(s+1)\theta}h_{s+1} - h_0]$$

so that

$$\|(Y - \mathrm{e}^{\mathrm{i}\theta})\chi_s(\theta)\|^2 = \frac{2}{s+1}.$$

Taking limits, we see that for any $\theta \in \mathbb{R}$,

$$\lim_{s\to\infty} \|(Y - \mathrm{e}^{\mathrm{i}\theta})\chi_s(\theta)\| = 0.$$

Similarly,

$$Y^*\chi_s(\theta) = \frac{1}{\sqrt{s+1}}\sum_{n=1}^{s} \mathrm{i}^{n-1}\mathrm{e}^{-\mathrm{i}n\theta}h_{n-1}$$
$$= \frac{\mathrm{e}^{-\mathrm{i}\theta}}{\sqrt{s+1}}\sum_{n=0}^{s-1} \mathrm{i}^{n}\mathrm{e}^{-\mathrm{i}n\theta}h_{n}$$

and so

$$(Y^* - \mathrm{e}^{-\mathrm{i}\theta})\chi_s(\theta) = -\mathrm{i}^s\frac{\mathrm{e}^{-\mathrm{i}(s+1)\theta}}{\sqrt{s+1}}h_s$$

so that

$$\|(Y^* - \mathrm{e}^{-\mathrm{i}\theta})\chi_s(\theta)\|^2 = \frac{1}{s+1}.$$

Taking limits, we see that for any $\theta \in \mathbb{R}$,

$$\lim_{s\to\infty} \|(Y^* - \mathrm{e}^{-\mathrm{i}\theta})\chi_s(\theta)\| = 0.$$

□

This, coupled with the fact that

$$|\langle\chi_s(\theta), h_n\rangle| = \frac{1}{s+1} \qquad 0 \leqslant n \leqslant s$$

gives reasonable motivation for choosing $\chi_s(\theta)$ as an 'eigenstate' of Y (or Y^*) in $\mathcal{H}_s$ with 'eigenvalue' $\mathrm{e}^{\mathrm{i}\theta}$ (or $\mathrm{e}^{-\mathrm{i}\theta}$). Moreover, the same sort of results are true for the operator X itself.

We next consider the expectation of the Toeplitz phase operator X in the state $\chi_s(\theta)$. This will give us a bound enabling us to relate the $\chi_s(\theta)$ to the spectrum of X.

Lemma 5.2. For $|\theta| < \pi$, we have the limit result

$$\lim_{s\to\infty} \langle \chi_s(\theta), X\chi_s(\theta)\rangle = \theta. \tag{5.6}$$

Proof. If we expand $\chi_s(\theta)$ in terms of Hermite functions, and then substitute the expression for the Hermite matrix elements of X we find

$$\langle \chi_s(\theta), X\chi_s(\theta)\rangle = -\frac{2}{s+1}\sum_{0\leqslant n<m\leqslant s}\frac{(-1)^{m-n}}{m-n}\sin[(m-n)\theta].$$

This is essentially a one-variable sum, and introducing the new variable $N = m-n$ enables us to reduce the expression to the form

$$\langle \chi_s(\theta), X\chi_s(\theta)\rangle = -\frac{2}{s+1}\sum_{N=1}^{s}\frac{(-1)^N}{N}(s+1-N)\sin(N\theta).$$

Replacing the sine by a difference of exponentials and recalling the function $\Theta \in L^\infty(\mathbb{T})$ given by

$$\Theta(\theta) = \theta \qquad -\pi < \theta < \pi \tag{5.7}$$

and its Nth Fourier coefficient $\hat{\Theta}_N$, we may rewrite our expression as

$$\langle \chi_s(\theta), X\chi_s(\theta)\rangle = \sum_{N=-s}^{s}\hat{\Theta}_N \mathrm{e}^{\mathrm{i}N\theta} + \frac{2}{s+1}\sum_{N=1}^{s}(-1)^N\sin(N\theta)$$

for any $s \in \mathbb{N}$ and $\theta \in \mathbb{R}$.

Noting the sum

$$\sum_{N=1}^{s}(-1)^N\sin(N\theta) = \sin\left[\frac{s}{2}(\theta+\pi)\right]\sin\left[\frac{s+1}{2}(\theta+\pi)\right]\sec\frac{\theta}{2}$$

it follows that

$$\left|\langle \chi_s(\theta), X\chi_s(\theta)\rangle - \sum_{N=-s}^{s}\hat{\Theta}_N \mathrm{e}^{\mathrm{i}N\theta}\right| \leqslant \frac{2}{s+1}\sec\frac{\theta}{2} \tag{5.8}$$

for $s \in \mathbb{N}$ and $|\theta| < \pi$. The result now follows from standard Fourier series theory. □

We also require a useful expression for the quantity $\|X\chi_s(\theta)\|^2$, the second moment of X with respect to the state $\chi_s(\theta)$. To obtain it we deduce the following expression for the cross terms, the Hermite matrix elements of X^2.

Lemma 5.3. The matrix elements of X^2 with respect to the Hermite functions are given by

$$\langle Xh_m, Xh_n\rangle = \begin{cases} \dfrac{\mathrm{i}^{n-m}}{m-n}\left(\dfrac{2}{m-n} + \displaystyle\sum_{k=1}^{n}\frac{1}{k} - \sum_{k=1}^{m}\frac{1}{k}\right) & \text{if } m \neq n \\ \dfrac{\pi^2}{3} - \displaystyle\sum_{k=n+1}^{\infty}\frac{1}{k^2} & \text{if } m = n. \end{cases} \tag{5.9}$$

Proof. Using the definition of X in the case $m \neq n$ we get

$$\langle Xh_m, Xh_n\rangle = \mathrm{i}^{n-m} \sum_{k\geqslant 0, k\neq n,m} \frac{1}{(k-m)(k-n)}$$
$$= \frac{\mathrm{i}^{n-m}}{m-n} \sum_{k\geqslant 0, k\neq n,m} \left(\frac{1}{k-m} - \frac{1}{k-m}\right).$$

This expression simplifies to give the required result, while for $n = m$ we have

$$\langle Xh_n, h_n\rangle = \|Xh_n\|^2 = \sum_{m\geqslant 0, m\neq n} \frac{1}{(m-n)^2}.$$

Changing variables from m to $k = m - n$ and remembering that

$$\sum_{j\geqslant 1} \frac{1}{j^2} = \frac{\pi^2}{6}$$

the result is evident. □

With these preliminary calculations done, we can prove one of the principal results of this paper, namely that the states $\chi_s(\theta)$ are approximate eigenvectors for X, and the spectrum of X *includes the interval* $(-\pi, \pi)$.

The proof of this result contains the germ of the methods needed for analogous proofs concerning the other operators we shall consider. First of all, we must know the answer; the proposition is verification, not discovery. We must then work with the various matrix elements so as to extract the limiting s behaviour. This requires doing estimates of the sort just above, as we do not have closed form expressions for the various quantities we need, and suspect that they may not exist, in any event.

Proposition 5.4. For any $|\theta| < \pi$,

$$\lim_{s\to\infty} \|X\chi_s(\theta)\|^2 = \theta^2 \tag{5.10a}$$

so that for any $|\theta| < \pi$,

$$\lim_{s\to\infty} \|(X-\theta)\chi_s(\theta)\| = 0. \tag{5.10b}$$

Thus θ is a spectral value for X, and $\{\chi_s(\theta) : s \in \mathbb{N}\}$ is an approximating sequence of unit vectors for any $|\theta| < \pi$.

Proof. We begin the proof by expressing $\|X\chi_s(\theta)\|^2$ in terms of Hermite functions. This will involve the Hermite matrix elements of X^2, for which we substitute the expression obtained in the lemma above, yielding

$$\|X\chi_s(\theta)\|^2 = \frac{1}{s+1} \sum_{m,n=0}^{s} \mathrm{i}^{m-n} \mathrm{e}^{\mathrm{i}(m-n)\theta} \langle Xh_m, Xh_n\rangle$$
$$= \frac{1}{s+1} \sum_{n=0}^{s} \left(\frac{\pi^2}{3} - \sum_{k=n+1}^{\infty} \frac{1}{k^2}\right)$$
$$+ \frac{2}{s+1} \sum_{0\leqslant n<m\leqslant s} \frac{(-1)^{m-n}}{m-n} \cos[(m-n)\theta] \left(\frac{2}{m-n} - \sum_{k=n+1}^{m} \frac{1}{k}\right).$$

Writing out the sums in a less condensed form results in the unwieldy expression

$$\|X\chi_s(\theta)\|^2 = \frac{\pi^2}{3} - \frac{1}{s+1}\sum_{k=1}^{\infty}\frac{1+\min(s,k-1)}{k^2} + \frac{4}{s+1}\sum_{N=1}^{s}\frac{(-1)^N\cos N\theta}{N^2}(s+1-N)$$
$$-\frac{2}{s+1}\sum_{N=1}^{s}\sum_{n=0}^{s-N}\sum_{k=n+1}^{n+N}\frac{(-1)^N\cos N\theta}{Nk}$$

which we rearrange into the slightly more useful form

$$\|X\chi_s(\theta)\|^2 = \sum_{N=-s}^{s}\widehat{(\Theta^2)}_N e^{iN\theta} - \frac{1}{s+1}\sum_{k=1}^{\infty}\frac{1+\min(s,k-1)}{k^2} - \frac{4}{s+1}\sum_{N=1}^{s}\frac{(-1)^N\cos N\theta}{N}$$
$$-\frac{2}{s+1}\sum_{k=1}^{s}\sum_{n=0}^{k-1}\sum_{N=k-n}^{s-n}\frac{(-1)^N\cos N\theta}{Nk}$$

for any $\theta \in \mathbb{R}$ and $s \in \mathbb{N}$.

We estimate the last three terms in turn. The first of these estimates is

$$\begin{aligned}\left|\frac{1}{s+1}\sum_{k=1}^{\infty}\frac{1+\min(s,k-1)}{k^2}\right| &\leqslant \frac{1}{s+1}\sum_{k=1}^{s}\frac{1}{k} + \sum_{k=s+1}^{\infty}\frac{1}{k^2}\\ &\leqslant \frac{1+\log s}{s+1} + \int_s^{\infty} x^{-2}\,dx\\ &\leqslant \frac{1+\log s}{s+1} + \frac{1}{s}.\end{aligned}$$

The next estimate is

$$\left|\frac{4}{s+1}\sum_{N=1}^{s}\frac{(-1)^N\cos N\theta}{N}\right| \leqslant \frac{8}{(s+1)\cos(\frac{1}{2}\theta)}.$$

The third estimate bounds each of its three sums in turn:

$$\begin{aligned}\left|\frac{2}{s+1}\sum_{k=1}^{s}\sum_{n=0}^{k-1}\sum_{N=k-n}^{s-n}\frac{(-1)^N\cos N\theta}{Nk}\right| &\leqslant \frac{4}{(s+1)\cos(\frac{1}{2}\theta)}\sum_{k=1}^{s}\sum_{n=0}^{k-1}\frac{1}{k(k-n)}\\ &\leqslant \frac{4}{(s+1)\cos(\frac{1}{2}\theta)}\sum_{k=1}^{s}\frac{1+\log k}{k}\\ &\leqslant \frac{8}{(s+1)\cos(\frac{1}{2}\theta)}\sum_{k=1}^{s}\frac{1}{k^{1/2}}\\ &\leqslant \frac{32}{(s+1)^{1/2}\cos(\frac{1}{2}\theta)}\end{aligned}$$

for $|\theta| < \pi$.

It is seen that all three terms will vanish in the limit $s \to \infty$. Thus we deduce that

$$\lim_{s\to\infty}\left(\|X\chi_s(\theta)\|^2 - \sum_{N=-s}^{s}\widehat{(\Theta^2)}_N e^{iN\theta}\right) = 0$$

for $|\theta| < \pi$.

We recognize the sum as a partial sum of the Fourier series for the function Θ^2. Standard theory assures us of the pointwise convergence of these partial sums. Remembering that $\Theta^2(\theta) = \theta^2$, we let $s \to \infty$, obtaining

$$\lim_{s\to\infty} \|X\chi_s(\theta)\|^2 = \theta^2 \qquad |\theta| < \pi$$

as required. The remainder of the proposition is immediate. □

We have shown, therefore, that the states $\chi_s(\theta)$ for $|\theta| < \pi$ are of some importance for the Toeplitz phase operator X and to the shift operators Y and Y^*. Their importance stems from the fact that they are sequences of approximate eigenvectors for these operators, and satisfy the maximal randomness criterion given above, equation (5.4), in addition. The arguments of section 4 now apply, leading us to consider the measurement systems (X_s), (Y_s) and (Y_s^*) for X, Y and Y^*, respectively.

We believe that the operators X, Y and Y^* are not as immediately physically relevant as the operator $\Delta(\varphi)$ and $\Delta(e^{\pm i\varphi})$, since these latter are Weyl quantizations of phase space functions of the angle alone. It is the relation of the $\chi_s(\theta)$ to these operators to which we now turn.

6. The phase operators $\Delta(\varphi)$, $\Delta(e^{i\varphi})$ and $\Delta(e^{-i\varphi})$

We now present similar considerations for the phase operator and the quantized exponentials. we recall [15] that the phase operator is the bounded operator on $L^2(\mathbb{R})$ given by the formula

$$\Delta(\varphi)h_n = \sum_{n,m\geqslant 0, m\neq n} \frac{i^{n-m+1}}{m-n} g_{m,n} h_m \qquad n \geqslant 0 \tag{6.1}$$

where the g-factor, which distinguishes $\Delta(\varphi)$ from the Toeplitz phase operator, is

$$g_{m,n} = \frac{\xi_{\min(m,n),s(m,n)}}{\xi_{\max(m,n),s(m,n)}} \qquad m, n \geqslant 0. \tag{6.2}$$

The quantities ξ are

$$\xi_{n,j} = 2^{n/2}\Gamma\left(\frac{n}{2} + j\right)(n!)^{-1/2} \tag{6.3a}$$

with

$$s(m,n) = \begin{cases} 1 & \text{if } \min(m,n) \text{ is odd} \\ \frac{1}{2} & \text{if } \min(m,n) \text{ is even.} \end{cases} \tag{6.3b}$$

The quantization of the phase space exponentials is given by the formulae

$$\Delta(e^{i\varphi})h_n = \lambda_{n+1}h_{n+1} \qquad n \geqslant 0 \tag{6.4a}$$

and

$$\Delta(e^{-i\varphi})h_n = \begin{cases} \overline{\lambda_n} h_{n-1} & \text{if } n \geqslant 1 \\ 0 & \text{if } n = 0. \end{cases} \tag{6.4b}$$

The factors λ_n are given by

$$\lambda_{n+1} = \mathrm{i}\left(\frac{n+1}{2}\right)^{1/2} \frac{\Gamma(\frac{n}{2}+s_n)}{\Gamma(\frac{n}{2}+s_n+\frac{1}{2})} \qquad n \geqslant 0 \tag{6.5a}$$

where

$$s_n = \begin{cases} \frac{1}{2} & \text{if } n \text{ is even} \\ 1 & \text{if } n \text{ is odd.} \end{cases} \tag{6.5b}$$

The λ_n were obtained from the $g_{m,n}$, and when $g_{m,n} = 1$ the operators $\Delta(\mathrm{e}^{\mathrm{i}\varphi})$ and $\Delta(\mathrm{e}^{-\mathrm{i}\varphi})$ go over to Y and Y^*, respectively.

As before, let us start with the operators $\Delta(\mathrm{e}^{\mathrm{i}\varphi})$ and $\Delta(\mathrm{e}^{-\mathrm{i}\varphi})$, for which the calculations are simpler. From the definition of λ_n we can prove that there exists a strictly positive constant A such that

$$||\lambda_n| - 1| \leqslant \frac{A}{n} \qquad n \in \mathbb{N}. \tag{6.6}$$

This enables us to prove the precise counterpart to the fact that $\mathrm{e}^{\mathrm{i}\theta}$ and $\mathrm{e}^{-\mathrm{i}\theta}$ are in the spectrum of the Toeplitz shift operators Y and Y^*, respectively. The result is an illustration of the fact that the spectrum and an approximating sequence of unit vectors for each point of it is not sufficient to completely characterize an operator. In this it differs from a full spectral decomposition, which completely defines an operator.

Proposition 6.1. For any $\theta \in \mathbb{R}$, the limits

$$\lim_{s\to\infty} \|(\Delta(\mathrm{e}^{\mathrm{i}\varphi}) - \mathrm{e}^{\mathrm{i}\theta})\chi_s(\theta)\| = 0 \qquad \text{and} \qquad \lim_{s\to\infty} \|(\Delta(\mathrm{e}^{-\mathrm{i}\varphi}) - \mathrm{e}^{-\mathrm{i}\theta})\chi_s(\theta)\| = 0 \tag{6.7}$$

hold, so that $\mathrm{e}^{\mathrm{i}\theta}$ is an spectral value for $\Delta(\mathrm{e}^{\mathrm{i}\varphi})$, $\mathrm{e}^{-\mathrm{i}\theta}$ is an spectral value for $\Delta(\mathrm{e}^{-\mathrm{i}\varphi})$, and $\{\chi_s(\theta) : s \in \mathbb{N}\}$ is an approximating sequence of unit vectors in both cases.

Proof.

$$\Delta(\mathrm{e}^{\mathrm{i}\varphi})\chi_s(\theta) = \frac{1}{\sqrt{s+1}} \sum_{n=0}^{s} \mathrm{i}^n \mathrm{e}^{-\mathrm{i}n\theta} \lambda_{n+1} h_{n+1} = \frac{\mathrm{e}^{\mathrm{i}\theta}}{\sqrt{s+1}} \sum_{n=1}^{s+1} \mathrm{i}^n \mathrm{e}^{-\mathrm{i}n\theta} |\lambda_n| h_n$$

so that

$$[\Delta(\mathrm{e}^{\mathrm{i}\varphi}) - \mathrm{e}^{\mathrm{i}\theta}]\chi_s(\theta) = \frac{\mathrm{e}^{\mathrm{i}\theta}}{\sqrt{s+1}} \left(\mathrm{i}^{s+1} \mathrm{e}^{-\mathrm{i}(s+1)\theta} |\lambda_{s+1}| h_{s+1} + \sum_{n=1}^{s+1} \mathrm{i}^n \mathrm{e}^{-\mathrm{i}n\theta} (|\lambda_n| - 1) h_n - h_0 \right).$$

Taking the inner product of this with itself, we can apply the above bound on the λ_n to obtain the leading term for large s:

$$\begin{aligned} \|[\Delta(\mathrm{e}^{\mathrm{i}\varphi}) - \mathrm{e}^{\mathrm{i}\theta}]\chi_s(\theta)\|^2 &= \frac{1}{s+1} \left(|\lambda_{s+1}|^2 + \sum_{n=1}^{s} (|\lambda_n| - 1)^2 + 1 \right) \\ &\leqslant \frac{1}{s+1} \left[(1+A)^2 + \frac{\pi^2}{6} A^2 + 1 \right] \end{aligned} \tag{6.8}$$

for all $s \in \mathbb{N}$ and $\theta \in \mathbb{R}$, so certainly

$$\lim_{s\to\infty} \|[\Delta(e^{i\varphi}) - e^{i\theta}]\chi_s(\theta)\| = 0.$$

For the operator $\Delta(e^{-i\varphi})$ we proceed as follows.

$$\Delta(e^{-i\varphi})\chi_s(\theta) = \frac{1}{\sqrt{s+1}}\sum_{n=1}^{s} i^n e^{-in\theta}\overline{\lambda_n}h_{n-1} = \frac{e^{-i\theta}}{\sqrt{s-1}}\sum_{n=0}^{s-1} i^n e^{-in\theta}|\lambda_{n+1}|h_n$$

so that

$$[\Delta(e^{-i\varphi}) - e^{-i\theta}]\chi_s(\theta) = \frac{e^{-i\theta}}{\sqrt{s+1}}\left(-i^s e^{-is\theta}h_s + \sum_{n=0}^{s-1} i^n e^{-in\theta}(|\lambda_{n+1}| - 1)h_n\right).$$

Taking the inner product of this with itself, we now use the result for $\Delta(e^{i\varphi})$ *to get the* estimate we need.

$$\|[\Delta(e^{-i\varphi}) - e^{-i\theta}]\chi_s(\theta)\|^2 = \frac{1}{s+1}\left(\sum_{n=1}^{s}(|\lambda_n| - 1)^2 + 1\right)$$
$$\leqslant \|[\Delta(e^{i\varphi}) - e^{i\theta}]\chi_s(\theta)\|^2 \tag{6.9}$$

so for all $s \in \mathbb{N}$ and $\theta \in \mathbb{R}$ we have

$$\lim_{s\to\infty} \|[\Delta(e^{-i\varphi}) - e^{-i\theta}]\chi_s(\theta)\| = 0$$

as well. □

Moving on to consider the phase operator $\Delta(\varphi)$, we need to obtain estimates on the behaviour of expressions analogous to those dealt with when studying the operator of Popov and Yarunin, but with extra terms coming from the presence of the g-factor. In the analysis below, we intend to show that for the quantities we are calculating, we may replace g by unity, so that the results for X apply, without affecting the end results.

The factor g is quite subtle to work with, and it is useful to attack the problem by proceeding in stages. The first stage we have done: the operator X is just the phase operator $\Delta(\varphi)$ with the g-factor replaced by unity. In the next stage we express the $g_{m,n}$ as

$$g_{m,n} = Y_{m,n}\left[1 + \frac{C(m,n)}{\min(m,n)+1}\right] \qquad m, n \geqslant 0$$

where

$$Y_{m,n} = \left(\frac{\min(m,n)+1}{\max(m,n)+1}\right)^{s(m,n)-\frac{3}{4}}$$
$$= \left(\frac{\min(m,n)+1}{\max(m,n)+1}\right)^{\frac{1}{4}(-1)^{1+\min(m,n)}}$$

A precise expression for the $C(m,n)$ may be obtained from the expression for $g_{m,n}$. However, the only property of the $C(m,n)$ that we shall need to use is its boundedness: there exists a strictly positive constant C such that

$$|C(m,n)| \leqslant C \qquad m, n \geqslant 0.$$

Effectively, then, we replace $\Delta(\varphi)$ by the simpler operator involving $Y_{m,n}$ in place of $g_{m,n}$.

In the final stage, we compare this simpler operator with X in the large s limit.

The matrix $(g_{m,n})$ has an alternating pattern of increase and decrease, which stems from the quantities $s(m,n)$. The result is the appearance of the minimum and maximum functions. The matrix $(Y_{m,n})$ has analogous properties. The precise property is the subject of this next lemma.

Lemma 6.2. (i) If n is even, then the sequence whose Nth term is $N^{-1}(Y_{n,n+N}-1)$ is positive and monotonically decreasing.

(ii) If n is odd, then the sequence whose Nth term is $N^{-1}(1-Y_{n,n+N})$ is positive and monotonically decreasing.

Proof. For the first sequence, with n even, we observe that

$$\begin{aligned}\frac{Y_{n,n+N}-1}{N} &= \frac{1}{N}\left[\left(\frac{n+N+1}{n+1}\right)^{1/4}-1\right]\\ &= (n+1)^{-1/4}[(n+N+1)^{1/4}+(n+1)^{1/4}]^{-1}[(n+N+1)^{1/2}+(n+1)^{1/2}]^{-1}\end{aligned}$$

and for the second sequence, with n odd,

$$\begin{aligned}\frac{1-Y_{n,n+N}}{N} &= \frac{1}{N}\left[1-\left(\frac{n+1}{n+N+1}\right)^{1/4}\right]\\ &= (n+N+1)^{-1/4}[(n+N+1)^{1/4}+(n+1)^{1/4}]^{-1}[(n+N+1)^{1/2}+(n+1)^{1/2}]^{-1}\end{aligned}$$

and the assertions are now immediate. □

We can now establish the first result concerning $\Delta(\varphi)$ by using the method outlined above.

Proposition 6.3. For all $|\theta| < \pi$, the limit

$$\lim_{s\to\infty}\langle\chi_s(\theta), \Delta(\varphi)\chi_s(\theta)\rangle = \theta \tag{6.10}$$

holds.

Proof. From the definitions,

$$\langle\chi_s(\theta), \Delta(\varphi)\chi_s(\theta)\rangle = -\frac{2}{s+1}\sum_{N=1}^{s}\sum_{n=0}^{s-N}\frac{(-1)^N}{N}\sin(N\theta)g_{n,n+N}. \tag{6.11a}$$

For comparison, we know that

$$\langle\chi_s(\theta), X\chi_s(\theta)\rangle = -\frac{2}{s+1}\sum_{N=1}^{s}\sum_{n=0}^{s-N}\frac{(-1)^N}{N}\sin(N\theta). \tag{6.11b}$$

As mentioned above, it will be useful to use an intermediate stage between X and $\Delta(\varphi)$ through the quantities $Y_{m,n}$. Hence we define

$$\alpha(s) = -\frac{2}{s+1}\sum_{N=1}^{s}\sum_{n=0}^{s-N}\frac{(-1)^N}{N}\sin(N\theta)Y_{n,n+N}. \tag{6.11c}$$

Expressing $g_{m,n}$ in terms of $Y_{m,n}$ and using the boundedness of the $C(m,n)$, we find

$$\begin{aligned}|\langle\chi_s(\theta), \Delta(\varphi)\chi_s(\theta)\rangle-\alpha(s)| &= \frac{2}{s+1}\left|\sum_{N=1}^{s}\sum_{n=0}^{s-N}\frac{(-1)^N}{N}\sin(N\theta)\frac{1}{n+1}Y_{n,n+N}C(n,n+N)\right|\\ &\leqslant \frac{2C}{(s+1)^{3/4}}\left[\sum_{N=1}^{s}\frac{1}{N}\right]\left[\sum_{n=0}^{s-1}\frac{1}{(n+1)^{5/4}}\right]\end{aligned}$$

from which it follows that

$$\langle \chi_s(\theta), \Delta(\varphi)\chi_s(\theta)\rangle - \alpha(s) = O\left(\frac{\log s}{(s+1)^{3/4}}\right) \tag{6.12}$$

as s tends to infinity. It now follows that

$$\lim_{s\to\infty}[\langle \chi_s(\theta), \Delta(\varphi)\chi_s(\theta)\rangle - \alpha(s)] = 0. \tag{6.13}$$

This reduces the asymptotic form of the first moment of $\Delta(\varphi)$ in the state $\chi_s(\theta)$ (the expectation) to the limit of the $Y_{m,n}$ problem. We obtain this limit by doing the sum over N first, and comparing the result with the properties of the operator X:

$$\begin{aligned} |\alpha(s) - \langle \chi_s(\theta), X\chi_s(\theta)\rangle| &= \frac{2}{s+1}\left|\sum_{N=1}^{s}\sum_{n=0}^{s-N}\frac{(-1)^N}{N}\sin(N\theta)(Y_{n,n+N}-1)\right| \\ &\leqslant \frac{4}{(s+1)\cos(\frac{1}{2}\theta)}\sum_{n=0}^{s-1}|Y_{n,n+1}-1| \\ &\leqslant \frac{1}{(s+1)\cos(\frac{1}{2}\theta)}\sum_{n=1}^{s}\frac{1}{n} \end{aligned}$$

for any $s \in \mathbb{N}$ and $|\theta| < \pi$. Thus

$$\alpha(s) - \langle \chi_s(\theta), X\chi_s(\theta)\rangle = O\left(\frac{\log s}{s+1}\right) \tag{6.14}$$

as s tends to infinity, $|\theta| < \pi$, so we have

$$\lim_{s\to\infty}[\alpha(s) - \langle \chi_s(\theta), X\chi_s(\theta)\rangle] = 0 \tag{6.15}$$

for any $|\theta| < \pi$.

Thus,

$$\lim_{s\to\infty}\langle \chi_s(\theta), \Delta(\varphi)\chi_s(\theta)\rangle = \lim_{s\to\infty}\langle \chi_s(\theta), X\chi_s(\theta)\rangle \tag{6.16}$$

and the results of the previous section tell us that

$$\lim_{s\to\infty}\langle \chi_s(\theta), \Delta(\varphi)\chi_s(\theta)\rangle = \theta$$

as required. □

As an aid to the next stage of the calculation of the asymptotic form of the second moment of $\Delta(\varphi)$, we note the following intermediate result.

Lemma 6.4. Let $|\theta| < \pi$. Then

$$\lim_{s\to\infty}[\|(\Delta(\varphi) - X)\chi_s(\theta)\|^2 - \|(\Delta(\varphi) - \theta)\chi_s(\theta)\|^2] = 0 \tag{6.17}$$

and so we can say that

$$\lim_{s\to\infty}\|(\Delta(\varphi) - \theta)\chi_s(\theta)\|^2 = 0 \qquad \text{if and only if } \lim_{s\to\infty}\|(\Delta(\varphi) - X)\chi_s(\theta)\|^2 = 0. \tag{6.18}$$

Proof. Consider

$$\begin{aligned}\|(\Delta(\varphi) - X)\chi_s(\theta)\|^2 &= \|(\Delta(\varphi) - \theta)\chi_s(\theta) - (X - \theta)\chi_s(\theta)\|^2 \\ &= \|(\Delta(\varphi) - \theta)\chi_s(\theta)\|^2 + \|(X - \theta)\chi_s(\theta)\|^2 \\ &\quad - 2\,\mathrm{Re}\langle(\Delta(\varphi) - \theta)\chi_s(\theta), (X - \theta)\chi_s(\theta)\rangle.\end{aligned}$$

Using the Schwarz inequality for the last term,

$$\begin{aligned}|\,\mathrm{Re}\langle(\Delta(\varphi) - \theta)\chi_s(\theta), (X - \theta)\chi_s(\theta)\rangle| &\leqslant \|(\Delta(\varphi) - \theta)\chi_s(\theta)\|\,\|(X - \theta)\chi_s(\theta)\| \\ &\leqslant (\tfrac{3}{2}\pi + |\theta|)\|(X - \theta)\chi_s(\theta)\|\end{aligned}$$

for any $s \in \mathbb{N}$ and $\theta \in \mathbb{R}$. Thus

$$\begin{aligned}&|\|(\Delta(\varphi) - X)\chi_s(\theta)\|^2 - \|(\Delta(\varphi) - \theta)\chi_s(\theta)\|^2| \\ &\qquad \leqslant (\tfrac{3}{2}\pi + |\theta|)\|(X - \theta)\chi_s(\theta)\| + \|(X - \theta)\chi_s(\theta)\|^2.\end{aligned}$$

From our previous work on the spectrum of the Toeplitz phase operator, we know that $\|(X - \theta)\chi_s(\theta)\|^2$ converges to zero as s tends to infinity if $|\theta| < \pi$. The result is now immediate. □

In this way we are led to study the second moment of the difference operator $\Delta(\varphi) - X$, which is in accordance with the general scheme of analysis we have been pursuing.

This second moment involves two factors of g:

$$\|(\Delta(\varphi) - X)\chi_s(\theta)\|^2 = \frac{1}{s+1}\sum_{m,n=0}^{s}(-1)^{m-n}\mathrm{e}^{\mathrm{i}(m-n)\theta}\sum_{k\geqslant 0, k\neq m,n}\frac{g_{k,m}-1}{k-m}\frac{g_{k,n}-1}{k-n}.$$

Remembering that substituting $Y_{m,n}$ for $g_{m,n}$ was useful in considering one g-factor, we suppose that it will prove equally useful for two. The two-factor analogue of $\alpha(s)$ is the function

$$\beta_s(\theta) = \frac{1}{s+1}\sum_{m,n=0}^{s}(-1)^{m-n}\mathrm{e}^{\mathrm{i}(m-n)\theta}\sum_{k\geqslant 0, k\neq m,n}\frac{Y_{k,m}-1}{k-m}\frac{Y_{k,n}-1}{k-n}. \tag{6.19}$$

Following the pattern employed before, we take the difference between the $g_{m,n}$ and $Y_{m,n}$ expressions:

$$\begin{aligned}\|(\Delta(\varphi) - X)\chi_s(\theta)\|^2 - \beta_s(\theta) &= \frac{1}{s+1}\sum_{n=0}^{s}\sum_{\substack{k\geqslant 0\\ k\neq n}}\frac{Z_{k,n,n}}{(k-n)^2} \\ &\quad + \frac{2}{s+1}\sum_{0\leqslant n<m\leqslant s}(-1)^{m-n}\cos[(m-n)\theta]\sum_{\substack{k\geqslant 0\\ k\neq m,n}}\frac{Z_{k,m,n}}{(k-m)(k-n)}\end{aligned}$$

where the quantity $Z_{k,m,n}$ is given by

$$
\begin{aligned}
Z_{k,m,n} &= (g_{k,m}-1)(g_{k,n}-1)-(Y_{k,m}-1)(Y_{k,n}-1)\\
&= Y_{k,m}Y_{k,n}\left(\left[1+\frac{C(k,m)}{\min(k,m)+1}\right]\left[1+\frac{C(k,n)}{\min(k,n)+1}\right]-1\right)\\
&\quad - Y_{k,m}\frac{C(k,m)}{\min(k,m)+1} - Y_{k,n}\frac{C(k,n)}{\min(k,n)+1}.
\end{aligned}
$$

Applying the bound, we obtain

$$
|Z_{k,m,n}| \leqslant \frac{C(C+4)}{\min(k,m,n)+1}\left(\frac{\max(k,m)+1}{\min(k,m)+1}\,\frac{\max(k,n)+1}{\min(k,n)+1}\right)^{1/4}
$$

for all $k, m, n \geqslant 0$. There now follows a long calculation which reduces part of the problem to the convergence of $\beta_s(\theta)$ as s tends to infinity. In order to reduce the length of the proof, we shall consider only one part of it in any detail, as being characteristic of the rest, which we shall only indicate.

Proposition 6.5. For all $|\theta| < \pi$, the limit

$$
\lim_{s\to\infty} \|(\Delta(\varphi)-X)\chi_s(\theta)\| = 0 \qquad \text{holds if and only if } \lim_{s\to\infty}\beta_s(\theta)=0. \tag{6.20}
$$

Proof. The first of the sums we must consider is

$$
\begin{aligned}
\left|\frac{1}{s+1}\sum_{n=0}^{s}\left(\sum_{k\geqslant 0,k\neq n}\frac{Z_{k,n,n}}{(k-n)^2}\right)\right| &\leqslant \frac{1}{s+1}\sum_{n=0}^{s}\left(\sum_{k=0}^{n-1}\frac{C(C+4)}{(k+1)(k-n)^2}\left(\frac{n+1}{k+1}\right)^{1/2}\right.\\
&\quad\left. + \sum_{k=n+1}^{\infty}\frac{C(C+4)}{(n+1)(k-n)^2}\left(\frac{k+1}{n+1}\right)^{1/2}\right)\\
&\leqslant \frac{C(C+4)}{(s+1)^{1/2}}\sum_{n=1}^{s}\sum_{k=0}^{n-1}\frac{1}{(k+1)(k-n)^2} + \frac{C(C+4)}{(s+1)}\sum_{n=0}^{s}\sum_{k=1}^{\infty}\frac{(k+n+1)^{1/2}}{k^2(n+1)^{3/2}}\\
&\leqslant 2\frac{C(C+4)}{(s+1)^{1/2}}\left[\sum_{n=1}^{s}\frac{1}{n+1}\right]\left[\sum_{k=1}^{n}\frac{1}{k}\right] + 2\frac{C(C+4)}{(s+1)}\left[\sum_{n=0}^{s}\frac{1}{(n+1)}\right]\left[\sum_{k=1}^{\infty}\frac{1}{k^{3/2}}\right].
\end{aligned}
$$

This is a type of expression we have encountered before, and we deduce that

$$
\frac{1}{s+1}\sum_{n=0}^{s}\sum_{\substack{k\geqslant 0\\ k\neq n}}\frac{Z_{k,n,n}}{(k-n)^2} = O\left(\frac{(\log s)^2}{(s+1)^{1/2}}\right)
$$

as s tends to infinity, and so

$$
\lim_{s\to\infty}\frac{1}{s+1}\sum_{n=0}^{s}\left(\sum_{k\geqslant 0,k\neq n}\frac{Z_{k,n,n}}{(k-n)^2}\right) = 0.
$$

The second bound is obtained in the same way, and we find

$$\left|\frac{2}{s+1}\sum_{0\leqslant n<m\leqslant s}(-1)^{m-n}\cos[(m-n)\theta]\sum_{k<n}\frac{Z_{k,m,n}}{(k-m)(k-n)}\right|\leqslant 64\frac{C(C+4)}{(s+1)^{1/4}}\sum_{N=1}^{s-1}\frac{1}{N}.$$

In the usual way,

$$\frac{2}{s+1}\sum_{0\leqslant n<m\leqslant s}(-1)^{m-n}\cos[(m-n)\theta]\sum_{k<n}\frac{Z_{k,m,n}}{(k-m)(k-n)}=O\left(\frac{\log s}{(s+1)^{1/4}}\right)$$

as s tends to infinity, and so

$$\lim_{s\to\infty}\frac{2}{s+1}\sum_{0\leqslant n<m\leqslant s}(-1)^{m-n}\cos[(m-n)\theta]\sum_{k<n}\frac{Z_{k,m,n}}{(k-m)(k-n)}=0.$$

The next bound is

$$\frac{2}{s+1}\sum_{0\leqslant n<m\leqslant s}(-1)^{m-n}\cos[(m-n)\theta]\sum_{n<k<m}\frac{Z_{k,m,n}}{(k-m)(k-n)}=O\left(\frac{(\log s)^2}{(s+1)^{3/4}}\right)$$

as s tends to infinity, and so

$$\lim_{s\to\infty}\frac{2}{s+1}\sum_{0\leqslant n<m\leqslant s}(-1)^{m-n}\cos[(m-n)\theta]\sum_{n<k<m}\frac{Z_{k,m,n}}{(k-m)(k-n)}=0.$$

For the final sum, we note without proof that

$$\sum_{k\geqslant 1}\frac{(k+n+N+1)^{1/2}}{k(k+N)}\leqslant 5\frac{(n+N+1)^{1/2}}{N^{1/2}}.$$

If we substitute this into the estimate for the final sum we find that

$$\left|\frac{2}{s+1}\sum_{0\leqslant n<m\leqslant s}(-1)^{m-n}\cos[(m-n)\theta]\sum_{k>m}\frac{Z_{k,m,n}}{(k-m)(k-n)}\right|$$
$$\leqslant 10\frac{C(C+4)}{(s+1)^{3/4}}\left[\sum_{n\geqslant 1}\frac{1}{n^{5/4}}\right]\left[\sum_{N=1}^{s}\frac{1}{N^{1/2}}\right].$$

Thus

$$\frac{2}{s+1}\sum_{0\leqslant n<m\leqslant s}(-1)^{m-n}\cos[(m-n)\theta]\sum_{k>m}\frac{Z_{k,m,n}}{(k-m)(k-n)}=O\left(\frac{1}{(s+1)^{1/4}}\right)$$

for large s, so

$$\lim_{s\to\infty}\frac{2}{s+1}\sum_{0\leqslant n<m\leqslant s}(-1)^{m-n}\cos[(m-n)\theta]\sum_{k>m}\frac{Z_{k,m,n}}{(k-m)(k-n)}=0.$$

Putting these estimates together,

$$\lim_{s\to\infty}[\|(\Delta(\varphi)-X)\chi_s(\theta)\|^2-\beta_s(\theta)]=0$$

for all $|\theta|<\pi$, from which the proposition is immediate. □

To study the behaviour of $\beta_s(\theta)$ as s tends to infinity, we must split the expression for $\beta_s(\theta)$ into four parts. The resulting limit will be seen to depend on the large s limit of the function

$$\gamma_s(\theta) = \frac{2}{s+1} \sum_{0 \leqslant n < m \leqslant s} (-1)^{m-n} \cos[(m-n)\theta] \sum_{k>m} \frac{Y_{k,m}-1}{k-m} \frac{Y_{k,n}-1}{k-n}.$$

The relation between $\beta_s(\theta)$ and $\gamma_s(\theta)$ is

$$\begin{aligned}
\beta_s(\theta) - \gamma_s(\theta) &= \frac{1}{s+1} \sum_{n=0}^{s} \sum_{k \geqslant 0, k \neq n} \left(\frac{Y_{k,n}-1}{k-n} \right)^2 \\
&\quad + \frac{2}{s+1} \sum_{N=1}^{s-1} \sum_{n=1}^{s-N} \sum_{k=0}^{n-1} \frac{Y_{k,n+N}-1}{k-n-N} \frac{Y_{k,n}-1}{k-n} (-1)^N \cos(N\theta) \\
&\quad + \frac{2}{s+1} \sum_{N=2}^{s} \sum_{n=0}^{s-N} \sum_{k=n+1}^{n+N-1} \frac{Y_{k,n+N}-1}{k-n-N} \frac{Y_{k,n}-1}{k-n} (-1)^N \cos(N\theta).
\end{aligned}$$

Recalling that the sequence $((-1)^n N^{-1}[Y_{n,n+N} - 1])$ is positive and monotonic decreasing for any $n \geqslant 0$, we can establish the next result.

Proposition 6.6. For all $|\theta| < \pi$, the limit

$$\lim_{s \to \infty} \beta_s(\theta) = 0 \qquad \text{if and only if } \lim_{s \to \infty} \gamma_s(\theta) = 0. \tag{6.21}$$

Proof.

$$\begin{aligned}
0 &\leqslant \frac{1}{s+1} \sum_{n=0}^{s} \sum_{\substack{k \geqslant 0 \\ k \neq n}} \left(\frac{Y_{k,n}-1}{k-n} \right)^2 \\
&\leqslant \frac{1}{s+1} \sum_{n=0}^{s} \sum_{\substack{k \geqslant 0 \\ k \neq n}} [\min(k,n)+1]^{-1/2} [(k+1)^{1/4} + (n+1)^{1/4}]^{-2} \\
&\quad \times [(k+1)^{1/2} + (n+1)^{1/2}]^{-2} \\
&\leqslant \frac{1}{s+1} \left(\sum_{n=1}^{s} \sum_{k \geqslant 0}^{n-1} (k+1)^{-1/2} (n+1)^{-3/2} + \sum_{n=0}^{s} \sum_{k=n+1}^{\infty} (n+1)^{-1/2} (k+1)^{-3/2} \right)
\end{aligned}$$

from which we deduce that

$$\frac{1}{s+1} \sum_{n=0}^{s} \sum_{k \geqslant 0, k \neq n} \left(\frac{Y_{k,n}-1}{k-n} \right)^2 = O\left(\frac{\log s}{s+1} \right)$$

as s tends to infinity, and so

$$\lim_{s \to \infty} \frac{1}{s+1} \sum_{n=0}^{s} \sum_{k \geqslant 0, k \neq n} \left(\frac{Y_{k,n}-1}{k-n} \right)^2 = 0.$$

Similarly,

$$\left|\frac{2}{s+1}\sum_{N=1}^{s-1}\sum_{n=1}^{s-N}\sum_{k=0}^{n-1}\frac{Y_{k,n+N}-1}{k-n-N}\frac{Y_{k,n}-1}{k-n}(-1)^N\cos(N\theta)\right| \leqslant \frac{8}{(s+1)\cos(\frac{1}{2}\theta)}\sum_{k=0}^{s-2}\frac{1}{k+1}$$

and so is of the same order as the previous expression and, hence, converges to zero as s tends to infinity.

The last expression is of different order:

$$\frac{2}{s+1}\sum_{N=2}^{s}\sum_{n=0}^{s-N}\sum_{k=n+1}^{n+N-1}\frac{Y_{k,n+N}-1}{k-n-N}\frac{Y_{k,n}-1}{k-n}(-1)^N\cos(N\theta) = O\left(\frac{(\log s)^2}{s+1}\right)$$

and so converges to zero as s tends to infinity, from which the proposition now follows. □

Putting everything together, we have reduced the asymptotic behaviour of $\Delta(\varphi)$ to that of $\gamma_s(\theta)$:

Corollary 6.7. For any $|\theta| < \pi$,

$$\lim_{s\to\infty}\|(\Delta(\varphi)-\theta)\chi_s(\theta)\| = 0 \qquad \text{if and only if } \lim_{s\to\infty}\gamma_s(\theta) = 0. \tag{6.22}$$

Up until now we have made substantial use of the fact that if (A_N) is a monotonic decreasing sequence of positive terms, and if $|\theta| < \pi$, then

$$\left|\sum_{N=K}^{M}(-1)^N A_N\cos(N\theta)\right| \leqslant \frac{2A_K}{\cos(\frac{1}{2}\theta)}$$

for any $K < M$. However, in those cases where we consider the function $\gamma_s(\theta)$, it turns out that we must consider a sequence (A_N) which is not monotonic decreasing. Instead—and this case will occur here because of the interlacing property of the g-factor—the even and odd subsequences (A_{2N}) and (A_{2N+1}) are monotonic. The first is positive and decreasing, the second negative and increasing. In the next—and last—calculation, we take advantage of this by utilizing the following lemma, which we present without proof.

Lemma 6.8. If $(A_{2N})_{N\geqslant 1}$ and $(A_{2N+1})_{N\geqslant 0}$ are each positive monotonic decreasing sequences, then

$$\left|\sum_{N=1}^{M}A_N\cos N\theta\right| \leqslant \frac{2}{|\sin\theta|}(A_1+A_2) \tag{6.23}$$

for any $M \in \mathbb{N}$ and $0 < |\theta| < \pi$.

Finally, the culmination of our previous work identifies part of the spectrum of the phase operator $\Delta(\varphi)$ and the states $\chi_s(\theta)$ as approximate eigenvectors.

Theorem 6.9. For any $0 < |\theta| < \pi$, the limit

$$\lim_{s\to\infty}\|(\Delta(\varphi)-\theta)\chi_s(\theta)\| = 0 \tag{6.24}$$

holds. Thus θ is a spectral value for $\Delta(\varphi)$ and $\{\chi_s(\theta) : s \in \mathbb{N}\}$ is an approximating sequence of unit vectors.

Proof. Writing $\gamma_s(\theta)$ in terms of the $Y_{m,n}$, and reordering the sums, we find

$$|\gamma_s(\theta)| = \left|\frac{2}{s+1}\sum_{k=1}^{\infty}\sum_{n=0}^{s-1}\sum_{N=1}^{s-n}\frac{|Y_{k+n+N,n}-1|}{k}\frac{|Y_{k+n+N,n+N}-1|}{k+N}\cos(N\theta)\right|.$$

Applying the previous lemma,

$$|\gamma_s(\theta)| \leqslant \frac{4}{(s+1)\sin\theta}\sum_{k=1}^{\infty}\sum_{n=0}^{s-1}\frac{|Y_{k+n+1,n+1}-1|}{k}\frac{|Y_{k+n+1,n}-1|}{k+1} + \frac{|Y_{k+n+2,n+2}-1|}{k}\frac{|Y_{k+n+2,n}-1|}{k+2}.$$

With estimates based on the $Y_{m,n}$ of the sort obtained previously, we find that

$$|\gamma_s(\theta)| \leqslant \frac{16}{(s+1)|\sin\theta|}\sum_{n=0}^{s-1}\frac{1}{n+1}$$

which gives us the asymptotic behaviour

$$\gamma_s(\theta) = O\left(\frac{\log s}{s+1}\right) \tag{6.25}$$

as $s \to \infty$. Then

$$\lim_{s\to\infty}\gamma_s(\theta) = 0 \qquad |\theta| < \pi \tag{6.26}$$

from which the result follows. □

Taking stock, we have shown that the spectrum of $\Delta(\varphi)$ satisfies

$$\mathrm{spec}[\Delta(\varphi)] \supseteq (-\pi, 0)\bigcup(0, \pi) \tag{6.27}$$

and so contains the interval $[-\pi, \pi]$. We cannot conclude that this is the entire spectrum of $\Delta(\varphi)$, since we have only been able to prove that

$$\pi \leqslant \|\Delta(\varphi)\| \leqslant \frac{3\pi}{2} \tag{6.28}$$

a result obtained in [18] by considering coherent states. Were we able to prove that the *norm of $\Delta(\varphi)$ is equal to π, which we conjectured previously, this would complete the* determination of the spectral values of $\Delta(\varphi)$. Unfortunately, we have as yet made no progress in this direction.

It is rather intriguing, not to say irritating, that the value $\theta = 0$ is not dealt with by the above theory. Of course, we have been forced to make a number of approximations to obtain the above results, and it might have been the case that these approximations were too simplistic to deal with the value $\theta = 0$. This is not the case, however, since we certainly know that

$$\lim_{s\to\infty}\|\Delta(\varphi)\chi_s(0)\| = 0 \qquad \text{if and only if } \lim_{s\to\infty}\gamma_s(0) = 0. \tag{6.29}$$

But does $\gamma_s(0)$ converge to zero as $s \to \infty$?

To answer this we observe that

$$\sum_{n=0}^{M-1} \frac{1}{(k+M+1)^{1/4}+(n+1)^{1/4}} \frac{1}{(k+M+1)^{1/2}+(n+1)^{1/2}}$$
$$\geqslant \int_1^{M+1} \frac{1}{(k+M+1)^{1/4}+x^{1/4}} \frac{1}{(k+M+1)^{1/2}+x^{1/2}}\,\mathrm{d}x$$
$$\geqslant \frac{M}{4(k+M+1)^{3/4}}$$

for any k, $M \in \mathbb{N}$, while

$$\sum_{k=1}^{\infty} \frac{1}{(k+M+1)^{5/4}} \frac{1}{(k+M+1)^{1/4}+(M+1)^{1/4}} \frac{1}{(k+M+1)^{1/2}+(M+1)^{1/2}}$$
$$\geqslant \int_{M+2}^{\infty} \frac{1}{x^{5/4}} \frac{1}{(M+1)^{1/4}+x^{1/4}} \frac{1}{(M+1)^{1/2}+x^{1/2}}\,\mathrm{d}x$$
$$\geqslant \frac{1}{4(M+2)}$$

for any $m \in \mathbb{N}$.

With these bounds in hand, we calculate that

$$\gamma_s(0) \geqslant \frac{8}{s(s+1)} \sum_{M=1}^{s} \frac{M}{M+2} \geqslant \frac{s}{24(s+1)}$$

for all $s \in \mathbb{N}$, so that $\gamma_s(0)$ certainly does not converge to zero as $s \to \infty$. We conclude that:

Proposition 6.10. The sequence $(\|\Delta(\varphi)\chi_s(0)\|)$ does not tend to 0 as $s \to \infty$, and so $[\chi_s(0)]$ is not an approximating sequence of unit vectors for the spectral value $\theta = 0$ of $\Delta(\varphi)$.

7. Discussion

We now wish to resume our discussion of the theory of Barnett and Pegg from the point of view of the measurement systems of the SAE type we discussed in section 4. The new material for the discussion is the results of our asymptotic analysis in the previous two sections, namely

$$\|(Y - \mathrm{e}^{\mathrm{i}\theta})\chi_s(\theta)\| = \left(\frac{2}{s+1}\right)^{1/2} \qquad \|(Y^* - \mathrm{e}^{-\mathrm{i}\theta})\chi_s(\theta)\| = \left(\frac{1}{s+1}\right)^{1/2} \tag{7.1a}$$

$$\|(\Delta(\mathrm{e}^{-\mathrm{i}\varphi}) - \mathrm{e}^{-\mathrm{i}\theta})\chi_s(\theta)\| \leqslant \left(\frac{C}{s+1}\right)^{1/2} \qquad \|(\Delta(\mathrm{e}^{\mathrm{i}\varphi}) - \mathrm{e}^{\mathrm{i}\theta})\chi_s(\theta)\| \leqslant \left(\frac{C}{s+1}\right)^{1/2} \tag{7.1b}$$

for the exponential operators, where C is a strictly positive constant. For the Toeplitz phase operator we have

$$\|(X-\theta)\chi_s(\theta)\|^2 \leqslant \frac{1}{s} + \frac{1+\log s}{s+1} + \frac{32}{(s+1)^{1/2}}\frac{1}{|\cos(\theta/2)|} + \frac{8}{s+1}\frac{1}{|\cos(\theta/2)|}. \tag{7.2}$$

The Weyl quantization phase operator is estimated in terms of the Toeplitz phase operator by

$$\|[\Delta(\varphi)-\theta]\chi_s(\theta)\| \leqslant \|(X-\theta)\chi_s(\theta)\| + \|(\Delta(\varphi)-X)\chi_s(\theta)\| \tag{7.3a}$$

together with

$$\|[\Delta(\varphi)-X]\chi_s(\theta)\|^2 \leqslant O\left[\frac{1}{(s+1)^{1/4}}\right] + O\left[\frac{(\log s)^2}{(s+1)^{1/2}}\right] + O\left[\frac{\log s}{s+1}\right]\frac{1}{|\sin\theta|} + O\left[\frac{(\log s)^2}{s+1}\right]. \tag{7.3b}$$

The points $\theta = \pm\pi$ are omitted for X and $\Delta(\varphi)$ and, in addition, the point $\theta = 0$ is omitted from the $\Delta(\varphi)$ estimate. These are the exceptional points in the sense discussed in section 4.

Exceptional spectral values aside, this shows that the sequence $\{\chi_s(\theta) : s \geqslant 0\}$ of LHW states is an SAE for each of the above operators. For the exponential operators we have a globally spectrally uniform family, since we can take

$$\varepsilon_s[Y] = \left(\frac{2}{s+1}\right)^{1/2} \qquad \varepsilon_s[Y^*] = \left(\frac{1}{s+1}\right)^{1/2} \tag{7.4a}$$

$$\varepsilon_s[\Delta(\mathrm{e}^{-\mathrm{i}\varphi})] \leqslant \left(\frac{C}{s+1}\right)^{1/2} \qquad \varepsilon_s[\Delta(\mathrm{e}^{\mathrm{i}\varphi})] \leqslant \left(\frac{C}{s+1}\right)^{1/2}. \tag{7.4b}$$

For the phase operators we have local uniformity. For the Toeplitz operator X, let K be any compact subset of $(-\pi,\pi)$. From the above estimate we see that we may choose

$$\varepsilon_{s,K}[X] = \left(\frac{1}{s}\right)^{1/2} + \left(\frac{1+\log s}{s+1}\right)^{1/2} + \eta_K[X]\frac{(32)^{1/2}}{(s+1)^{1/4}} + \eta_K[X]\left(\frac{8}{s+1}\right)^{1/2} \tag{7.5a}$$

where

$$\eta_K[X] = \sup_{\theta\in K}|\cos(\theta/2)|^{-1/2}. \tag{7.5b}$$

For the operator $\Delta(\varphi)$, let K be any compact subset of $(-\pi,\pi)$ not meeting the point $\theta = 0$. Then we may choose

$$\varepsilon_{s,K}[\Delta(\varphi)] = \varepsilon_{s,K}[X] + C_K\left[\frac{1}{(s+1)^{1/4}} + \frac{(\log s)^2}{(s+1)^{1/2}} + \frac{(\log s)^2}{(s+1)} + \zeta_K[\Delta(\varphi)]\frac{\log s}{s+1}\right]^{1/2} \tag{7.6a}$$

where C_K is a suitable positive constant which could be made definite by chasing carefully through the bounds, and where

$$\zeta_K[\Delta(\varphi)] = \sup_{\theta\in K}|\sin\theta|^{-1}. \tag{7.6b}$$

We have chosen not to simplify these choices for two reasons. The first is that we do not calculate with them, and so their complexity does not matter, and second, in the form given it is easy to determine where each term comes from, by consulting sections 5 and 6.

The BP theory is based in large part on these LHW states, and we now recognize them to provide a uniform spectral SAE family for these operators, global or local as the case may be. But they have other properties which are of interest, namely:

The BP *subspace conditions (7.1).*

(a) The family $(\chi_{s,j})$ is an orthonormal basis for the $(s+1)$-dimensional Hilbert space $\mathcal{H}_s$ spanned by the first $s+1$ Hermite functions, and the collection $(\mathcal{H}_s)_s$ satisfies the ascending subspace condition.

(b) The $(\chi_{s,j})$ are uniformly distributed over the eigenvalues of the number operator.

Let us now examine the implications of using the LHW family to construct an SAE system for Y and Y^*. In principle, of course, we should consider C and S and not Y and Y^*, but results for the latter pair transfer to the former pair easily enough. With this latitude, we consider the family

$$Y_s = \sum_{j=0}^{s} \mathrm{e}^{\mathrm{i}\theta_{s,j}} P_{s,j} \qquad s \geqslant 0 \tag{7.7}$$

where $P_{s,j}$ projects onto the subspace spanned by $\chi_{s,j}$. As with all SAE systems, we have spectral accuracy.

Let us consider whether or not Y_s converges to Y, and if so, in what sense; and similarly for Y_s^* and Y^*.

Proposition 7.2. The family (Y_s) converges strongly, but not uniformly to Y. The adjoint family (Y_s^*) converges weakly, but not strongly, to Y.

Proof. A simple calculation shows that

$$Y_s h_n = \begin{cases} Y h_n & \text{if } n \leqslant s-1 \\ \mathrm{i}^{-s}\mathrm{e}^{\mathrm{i}(s+1)\theta_0} h_0 & \text{if } n = s \\ 0 & \text{if } n > s. \end{cases} \tag{7.8}$$

Then if f belongs to the subspace $\mathcal{H}$ spanned by the Hermite functions, we have

$$Y_s f = Y f \qquad f \in \mathcal{H} \tag{7.9}$$

for s large enough.

There are several ways to extend this result to all of $L^2(\mathbb{R})$. The most elementary is to note that

$$\|Y_s f\|^2 = \sum_{j=0}^{s} |(\chi_{s,j}, f)|^2 \leqslant \|f\|^2.$$

Hence the norms of the Y_s are uniformly bounded by 1. Now if g is any vector in $L^2(\mathbb{R})$ and ε is any strictly positive number, we may choose an $f \in \mathcal{H}$ such that $\|f-g\|$ is less than $(\|Y\|+1)^{-1}\varepsilon/2$. Then for s large enough so that $\|(Y-Y_s)f\| < \varepsilon/2$ we have

$$\begin{aligned} \|(Y-Y_s)g\| &\leqslant \|(Y-Y_s)f\| + \|Y-Y_s\|\|g-f\| \\ &\leqslant \|(Y-Y_s)f\| + (\|Y\|+\|Y_s\|)\|g-f\| \\ &\leqslant \|(Y-Y_s)f\| + (\|Y\|+1)\|g-f\| \\ &\leqslant \varepsilon. \end{aligned}$$

Thus Y_s converges strongly to Y.

Taking the adjoint of Y_s we find that

$$Y_s^* h_n = \begin{cases} i^s e^{-i(s+1)\theta_0} h_s & \text{if } n = 0 \\ Y^* h_n & \text{if } 1 \leqslant n \leqslant s \\ 0 & \text{if } n > s. \end{cases}$$

Then

$$\begin{aligned} Y_s^* \sum_{n=0}^{N} f_n h_n &= f_0 Y_s^* h_0 + Y_s^* \sum_{n=1}^{N} f_n h_n \\ &= i^s e^{-i(s+1)\theta_0} f_0 h_s + Y^* \sum_{n=1}^{N} f_n h_n \\ &= i^s e^{-i(s+1)\theta_0} f_0 h_s + Y^* \sum_{n=0}^{N} f_n h_n. \end{aligned}$$

Thus, if $f \in \mathcal{H}$,

$$(Y_s^* - Y) f = i^s e^{-i(s+1)\theta_0} f_0 h_s$$

for large enough s. Taking norms,

$$\|(Y_s^* - Y) f\| = |f_0|$$

which does not converge to 0 as $s \to \infty$. For weak convergence we calculate

$$|\langle g, (Y_s^* - Y) f \rangle| = |f_0 g_s|$$

which does converge to 0 for all $g \in L^2(\mathbb{R})$. This completes the proof. □

Thus the system (Y_s) is (just) reasonable as a measurement system for Y, but its adjoint is not a reasonable way to measure Y^*. We can deduce from this that we have no better than weak convergence for the Hermitian components (C_s) and (S_s) as measurement systems for C and S, respectively. For the operators $\Delta(e^{\pm i\varphi})$ there is necessarily no convergence at all, and so this is a very poor system for those measurements. The strong convergence of the Y_s to Y is the best convergence result we have found for SAE systems constructed from LHW states.

If the reader is surprised that the systems (Y_s) and (Y_s^*) behave differently as regards convergence, it should be remembered that the adjoint is not continuous in the strong operator topology.

More than for the exponential or trigonometric operators, the BP theory is meant to provide a substitute for a phase operator. For us, that means considering the system (X_s) as measuring X or $\Delta(\varphi)$. We now know that this is a locally, but not globally, uniform SAE system for these operators. We know its convergence properties from section 2, which we may summarize as follows:

Proposition 7.3. The sequence (X_s) converges weakly, but not strongly, to X. The power sequences (X_s^n) converge weakly, but not strongly, to operators $\hat{T}(\Theta^n)$.

This result is the reason that in using the BP theory, the limit $s \to \infty$ must be taken after all the matrix elements have been calculated. For if the BP theory were using a system that converged strongly, the limiting results would be just the same as calculations using the limit operator. The fact that a phase operator is replaced by a multiplicity of operators is known to those that use BP theory. Here we see what some of that multiplicity is. Each power sequence determines a new operator on $L^2(\mathbb{R})$. The SAE system determined by (X_s^n) does not converge to X^n, which further indicates the poor response of this system. This is also the case for many functions F—but not all—in considering $(F(X_s))$ for measuring $F(X)$.

Something which seems to have been overlooked is that as the $\hat{T}(\Theta^n)$ are themselves observables, they have a distribution in every state, and one should consider the interpretation of the powers $(\hat{T}(\Theta^n)^m)_{m\geqslant 0}$ for each n.

Further insight into the meaning of the BP theory comes from considering how close to being unique the LHW states are. If we demand the BP subspace conditions for some unknown family $(\psi_{s,j})$, we have the ansatz

$$\psi_{s,j} = \sum_{n=0}^{s} a(n; s, j) h_n$$

from (a). From (b) it follows that

$$|a(n; s, j)| = |a(m; s, j)| = (s+1)^{-1/2} \qquad 0 \leqslant n, m \leqslant s$$

the last equality following from normalization. The result may be written as

$$\psi_{s,j} = \frac{1}{\sqrt{s+1}} \sum_{n=0}^{s} \mathrm{i}^n \mathrm{e}^{-\mathrm{i}n\theta_{s,j}} u(n; s, j) h_n \tag{7.10}$$

where the $u(n; s, j)$ are of unit modulus. The only latitude in satisfying these conditions, then, is the appearance of phase factors as indicated. The SAE systems determined by such $(\psi_{s,j})$ will have the essentially same properties as the system obtained from the LHW states, certainly as regards convergence.

In examining this problem, we considered the possibility of keeping only the ascending subspace part of the BP subspace condition, and replacing the uniform distribution condition by a least-squares condition. Doing this for the operator Y, this means that we must find the normalized vectors $\psi_s(\theta)$ which minimize the quantity $\|(Y - \mathrm{e}^{\mathrm{i}\theta}) f\|$ over all (normalized) states $f \in \mathcal{H}_s$.

The solution to the minimum problem is obtained by noting that

$$\min\{\|(Y - \mathrm{e}^{\mathrm{i}\theta}) f\| : f \in \mathcal{H}_s, \|f\| = 1\} = 2 \sin\left(\frac{\pi}{2(s+2)}\right)$$

and that this minimum value is achieved for

$$\psi_s(\theta) = \left(\frac{2}{s+2}\right)^{1/2} \sum_{n=0}^{s} \sin\left[\frac{n+1}{s+2}\pi\right] \mathrm{i}^n \mathrm{e}^{-\mathrm{i}n\theta} h_n. \tag{7.11a}$$

Then

$$\psi_{s,j} = \left(\frac{2}{s+2}\right)^{1/2} \sum_{n=0}^{s} \sin\left[\frac{n+1}{s+2}\pi\right] \mathrm{i}^n \mathrm{e}^{-\mathrm{i}n\theta_{s,j}} h_n. \tag{7.11b}$$

However, the $\psi_{s,j}$ are not mutually orthogonal for fixed s as j varies, so this is not an acceptable SAE family.

This shows the stringency of the BP subspace conditions. As they were proposed in that theory as realizing crucial features of phase, any critique of the BP theory must consider the legitimacy of requiring these conditions. Given an operator A to be measured, as in section 4, a measurement system based on its spectral decomposition was constructed, which converged uniformly to A. After the construction, not before, one can determine what sort of subspaces are associated with the index s, and how they relate to other operators. From the no-go theorem, one would not expect a phase operator to be uniformly distributed over the number operator eigenstates. In fact, it is part of the useful results obtained from such a model to determine this distribution. Our judgment is that it is the deviation from the BP subspace conditions that give quantum phase theory its non-classical character.

One of the lessons this study teaches is not to confuse a measurement system with the observable to be measured in the non-ideal case. Even in the case of a system based on the spectral decomposition of A, there are infinitely many such systems obtained by rounding off the staircase function in various ways. Moreover, spectral accuracy by itself is certainly not a sufficient characterization of A.

In conclusion, we have argued in this paper that the only way that phase operators, which have continuous spectra, can be measured is by apparatus observables with discrete spectra. In particular, we have recast certain aspects of the BP theory into this framework, through the interpretation of the operators X_s as apparatus observables. In this setting, many of the BP calculations are perfectly understandable within standard quantum theory. The fact that the X_s do not converge strongly to the Toeplitz phase operator X accounts for certain other results which seem difficult to understand when not viewed in this way.

The essential distinction between different measuring systems arises from the different choices of output eigenvectors. We constructed models different from that obtained from the BP theory. Utilizing the spectral representation of the phase operator, we could construct systems of measurement which seem natural and give significantly more quantum information than does the BP system. Evidently, the experimental distinctions between different candidates for a phase operator are not a simple matter. They are further complicated by the fact that measuring, say $\Delta(\mathrm{e}^{\mathrm{i}\varphi})$ is not at all the same thing as measuring $\Delta(\varphi)$, or even $\mathrm{e}^{\mathrm{i}\Delta(\varphi)}$. These problems provide a nice challenge.

References

[1] Barnett S M and Pegg D T 1986 *J. Phys. A: Math. Gen.* **19** 3845
[2] Pegg D T and Barnett S M 1988 *Europhys. Lett.* **6** 483
[3] Barnett S M and Pegg D T 1989 *J. Mod. Opt.* **36** 7
[4] Pegg D T and Barnett S M 1989 *Phys. Rev.* A **39** 1665
[5] Pegg D T, Vaccaro J A and Barnett S M 1990 *J. Mod. Opt.* **37** 1703
[6] Barnett S M and Pegg D T 1992 *J. Mod. Opt.* **39** 2121
[7] Barnett S M and Dalton B J 1993 *Phys. Scr.* T **48** 13
[8] Vaccaro J A and Pegg D T 1993 *Phys. Scr.* T **48** 22
[9] Freyberger M and Schleich W P 1993 *Phys. Rev.* A **47** R30
[10] Schleich W P and Barnett S M (ed) 1993 *Phys. Scr.* T **48**
[11] Lerner E C, Huang H W and Walters G E 1970 *J. Math. Phys.* **11** 1679
[12] Garrison J C and Wong J 1970 *J. Math. Phys.* **11** 2242
[13] Popov V N and Yarunin V S 1992 *J. Mod. Opt.* **39** 1525
[14] Smith T B, Dubin D A and Hennings M A 1992 *J. Mod. Opt.* **39** 1603
[15] Dubin D A, Hennings, M A and Smith T B 1994 *Publ. RIMS Kyoto Univ.* **30** in press
[16] Loudon R 1973 *The Quantum Theory of Light* 1st edn (Oxford: Clarendon)

[17] Dubin D A and Hennings M A 1990 *Quantum Mechanics, Algebras and Distributions* (Harlow: Longman)
[18] Hennings M A, Smith T B and Dubin D A 1994 *Preprint* Open University
[19] Loudon R and Knight P L 1987 *J. Mod. Opt.* 34 709
[20] Gantsog Ts, Miranowicz A and Tanás R 1992 *Phys. Rev.* A **46** 2870
[21] Royer A 1993 Hermitian phase operators for the quantum harmonic oscillator *Preprint* Ecole Polytechnique, Montréal
[22] Taylor A E 1957 *Introduction to Functional Analysis* (New York: Wiley)
[23] Gel'fand I M and Vilenkin N Ya 1964 *Generalized Functions* vol 4 (New York: Academic)
[24] Haag R 1992 *Local Quantum Physics* (Berlin: Springer)
[25] Kemble E C 1937 *Fundamental Principles of Quantum Mechanics* (New York: Dover)
[26] Halmos P R 1982 *A Hilbert Space Problem Book* 2nd edn (New York: Springer)
[27] Lynch R 1995 *Phys. Rep.* **256** 367
[28] Dubin D A, Hennings M A and Smith T B 1995 *Int. J. Mod. Phys.* B to appear