INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND GENERAL

J. Phys. A: Math. Gen. **39** (2006) 1717–1728 doi:10.1088/0305-4470/39/7/015

Convergence of matrices under random conjugation: wave packet scattering without kinematic entanglement

Lawrence S Schulman[1] **and Leonard J Schulman**[2]

[1] Physics Department, Clarkson University, Potsdam, New York 13699-5820, USA
[2] Caltech, 1200 E California Blvd. MC 256-80, Pasadena, CA 91125, USA

Received 14 October 2005
Published 1 February 2006
Online at stacks.iop.org/JPhysA/39/1717

Abstract

In previous work, it was shown numerically that under successive scattering events, a collection of particles with Gaussian wavefunctions retains the Gaussian property, with the spread of the Gaussian ('Δx') tending to a value inversely proportional to the square root of each particle's mass. We prove this convergence in all dimensions $\geqslant 3$.

PACS numbers: 02.50.Ey, 03.65.Nk, 03.65.Ud

1. Introduction

Under reasonable assumptions concerning mutual interactions and other properties, a pair of particles described by Gaussian wave packets will emerge from a scattering with Gaussian wave packets [1]. The 'Gaussian' property means that the logarithm of the wavefunction is a polynomial of degree 2 in either position or momentum coordinates. We focus on the quadratic form and its associated matrix. There is a definite relation between the pre- and post-scattering matrices. Under a succession of collisions between unequal mass particles it was found, mostly numerically, that these matrices converge to specific multiples of the identity. The limiting values have a surprising property: although momentum conservation (kinematics) suggests that particle wavefunctions should be entangled even after unique-outcome scatterings, for Gaussians with the indicated limiting values, there is no entanglement.

This particular mapping of pairs of matrices into other pairs of matrices does not seem to have appeared previously in other physical applications. It has an attractive mathematical structure and will be studied in the present paper. For further details of the physical motivation and demonstration of the results listed in the previous paragraph, we refer to [1]. In [2], the *expectation* value under Haar measure is shown to converge; this is weaker than convergence of the actual matrices, demonstrated in the present paper. Since the physical background for this problem has already been presented twice [1, 2] we do not elaborate further.

In section 2, we introduce the subject by reviewing the one-space-dimension version of the mapping. This is followed, in section 3, by the full matrix iteration in three space dimensions. Section 4 contains the principal results of the paper: a statement of the general result and its proof. Showing convergence of the mapping in N dimensions is no more difficult than showing it in three. We mention that the *two*-dimensional case has exceptional properties. It is not discussed in any depth in the present paper; the results are given in [2]. Finally, in section 5, we present a number of extensions of the present results.

2. One dimension

A two-particle Gaussian wave packet in one dimension, as a function of its momenta, has the form $\Psi_I(p_1, p_2) = \mathcal{N}\exp\left[-\sigma_1^2(p_1 - k_1)^2 + \mathrm{i}p_1a_1 - \sigma_2^2(p_2 - k_2)^2 + \mathrm{i}p_2a_2\right]$, with $\mathcal{N}$ being normalization. We focus on the quadratic form $\sigma_1^2 p_1^2 + \sigma_2^2 p_2^2$ and take σ_k to be real (if unspecified, $k = 1, 2$). The position uncertainty of the particles is $(\Delta x)_k = \sigma_k$. Let the masses of the particles be m_k and define $r_k \equiv m_k/(m_1 + m_2)$. Further define $\xi_k \equiv r_k\sigma_k^2$. In [1], it was shown that after the particles have undergone a nontrivial scattering, the particles are represented by a density matrix which can again be expressed as a sum of Gaussian pure states with a new quadratic form and as a consequence new values of ξ_k. The mapping between the old and new values is

$$\begin{aligned} \xi_1' &= (r_1 - r_2)^2\xi_1 + 4r_1r_2\xi_2 = \xi_1\cos^2 2\phi + \xi_2\sin^2 2\phi, \\ \xi_2' &= (r_2 - r_1)^2\xi_2 + 4r_2r_1\xi_1 = \xi_2\cos^2 2\phi + \xi_1\sin^2 2\phi, \end{aligned} \tag{1}$$

where $\cos^2\phi \equiv r_1$ and use has been made of $r_1 + r_2 = 1$. If one considers a gas with many type-1 and type-2 particles, repeated scatterings between different particles take place and the mapping equation (1) is applied repeatedly[3]. The relative proportions of type-1 and type-2 particles determine the rate, but not the nature of this process. Equation (1) is obtained by taking the post-scattering wavefunction and for particle k tracing over the coordinates of particle k' (where $(k, k') = (1, 2)$ or $(2, 1)$). One thereby obtains new density matrices for each particle. For each of these, a particular value of the spread, σ, is obtained. These new spreads are what are given in equation (1) (through $\xi = r\sigma^2$).

We generalize and rewrite equation (1) in anticipation of the higher dimensional case. Consider

$$\begin{aligned} \xi_1' &= (r_1 + r_2\,\mathrm{e}^{\mathrm{i}\theta})\xi_1(r_1 + r_2\,\mathrm{e}^{-\mathrm{i}\theta}) + r_1r_2(1 - \mathrm{e}^{\mathrm{i}\theta})\xi_2(1 - \mathrm{e}^{-\mathrm{i}\theta}), \\ \xi_2' &= (r_2 + r_1\,\mathrm{e}^{\mathrm{i}\theta})\xi_2(r_2 + r_1\,\mathrm{e}^{-\mathrm{i}\theta}) + r_2r_1(1 - \mathrm{e}^{\mathrm{i}\theta})\xi_1(1 - \mathrm{e}^{-\mathrm{i}\theta}), \end{aligned} \tag{2}$$

which reduces to equation (1) for $\theta = \pi$. The joint evolution of $\{\xi_1, \xi_2\}$ is simple: adding the equations of equation (2), one finds that $\xi_1 + \xi_2$ is constant, while subtraction shows that $\xi_1 - \xi_2$ is multiplied on successive iterations by

$$\mu \equiv 1 - 4r_1r_2(1 - \cos\theta) = \cos^2\frac{\theta}{2} + \sin^2\frac{\theta}{2}\cos 4\phi = \delta^2 + (1 - \delta^2)\cos\theta, \tag{3}$$

where the notation $\delta \equiv r_1 - r_2$ has been introduced. $|\mu|$ is manifestly 1 or less. $\mu = 1$ only if $\theta = 0$ (nothing happens, i.e., no scattering) or if $\cos 4\phi = 1$, implying that one of the masses is zero. The case $\mu = -1$ can occur only if $\theta = \pi$ and $\cos 4\phi = -1$, which corresponds to equal masses. In the latter situation, ξ_1 and ξ_2 exchange values on successive scatterings.

[3] In a real gas, focusing on say, particle 1, the other particle will vary and the 'ξ_2' used on a subsequent iteration need not be the 'ξ_2'' of equation (1). In section 5.4, it will be shown that single-pair convergence implies the convergence of a collection. Pure transmission collisions ($\theta = 0$ in equation (2)) can also be included, allowing loss of particle–particle correlations.

The physical interpretation of the joint convergence is that a collection of particles with Gaussian wavefunctions will, after many collisions, have spread inversely proportional to the square root of each particle's mass. This relation of spread to mass eliminates the entanglement that in general would arise from purely kinematic considerations of momentum conservation. That is, when $r_k\sigma_k^2 = \text{const}$, apart from the effects of having multiple outcomes (in the one-dimensional case, transmission and reflection), the post-scattering wavefunction is a product of wavefunctions of particle 1 and particle 2, with no intertwining of coordinates that could in principle result from momentum conservation.

3. The physical case

For a Gaussian wavefunction in three dimensions, the spreads, $(\sigma_k)^2$, are symmetric positive definite 3×3 matrices. Let $R \in SO(3)$. As before, ξ_k is r_k times the corresponding spread matrix $\left(\sigma_k^2\right)$. As shown in [1], a pair of particles with Gaussian wave packets will, after scattering, have the following new values of ξ_k:

$$\begin{aligned} \xi_1' &= (r_1 + r_2 R)\xi_1(r_1 + r_2 R^{-1}) + r_1 r_2 (1 - R)\xi_2 (I - R^{-1}), \\ \xi_2' &= (r_2 + r_1 R)\xi_2(r_2 + r_1 R^{-1}) + r_2 r_1 (1 - R)\xi_1 (I - R^{-1}), \end{aligned} \tag{4}$$

(where $I = I_3$ is the 3×3 identity matrix). It is clear that $\mathrm{Tr}(\xi_1 + \xi_2)$ is constant. The numerically determined result of [1] is that each ξ_k converges to a multiple of the 3×3 identity, the multiplier being one-sixth the trace of the initial $\xi_1 + \xi_2$. It is the purpose of the present paper to prove and extend that result. We remark that, as in one dimension, the equilibrated spreads eliminate kinematic entanglement.

For later use, we rewrite equation (4) in terms of the sum and difference of the ξs. Define $P \equiv \xi_1 + \xi_2$ and $M \equiv \xi_1 - \xi_2$. Equation (4) is equivalent to

$$P' = \frac{1}{2}(P + RPR^{-1}) + \frac{\delta}{2}(M - RMR^{-1}) \tag{5}$$

$$M' = \frac{1-\delta^2}{2}(RM + MR^{-1}) + \frac{\delta^2}{2}(M + RMR^{-1}) + \frac{\delta}{2}(P - RPR^{-1}), \tag{6}$$

where, as defined earlier, $\delta = r_1 - r_2$.

In all considerations below, we exclude $\delta = \pm 1$.

Until now we have not formally stated the 'random' selection process for the matrices, R. In this paper, the probability distribution is Haar measure (on $SO(3)$, and when we generalize, on $SO(N)$). Numerical evidence shows that the convergence in fact holds under weaker randomness assumptions; this will be discussed in section 5.1.

4. Linear transformations of matrices

Convergence under the iteration of equation (4) is far more general than the case of 3×3 positive definite symmetric matrices. In particular, as we will now prove, it is true irrespective of the symmetry of the 'ξ' matrices and for all dimensions $N \geqslant 3$ (with $R \in SO(N)$). It holds in modified form in two dimensions. The considerations of the present section apply to all dimensions greater than 2. Yet further extensions will be discussed in section 5.

The iteration (4) is a linear operation on the pair of matrices (ξ_1, ξ_2). As such it is useful to define 'super-operators' (now on $\mathbb{R}^{N^2}$ rather than $\mathbb{R}^N$). For $R \in SO(N)$ and an $N \times N$ matrix A, define

$$C_R A = RAR^{-1} \tag{7}$$

$$B_R A = RA + AR^{-1}. \tag{8}$$

In terms of these, equation (4) becomes (dropping the explicit R dependence)

$$\begin{aligned}\xi_1' &= \left(r_1^2 + r_2^2 C + r_1 r_2 B\right)\xi_1 + r_1 r_2 (I + C - B)\xi_2 \\ \xi_2' &= \left(r_2^2 + r_1^2 C + r_2 r_1 B\right)\xi_2 + r_2 r_1 (I + C - B)\xi_1\end{aligned} \tag{9}$$

(where $I = I_N$ is the $N \times N$ identity matrix). These equations are combined by considering a column vector, $\zeta \equiv \binom{\xi_1}{\xi_2}$, and an operator S acting on this. We write

$$\zeta' = S\zeta \tag{10}$$

with

$$S \equiv \begin{pmatrix} r_1^2 + r_2^2 C + r_1 r_2 B & r_1 r_2 (I + C - B) \\ r_2 r_1 (I + C - B) & r_2^2 + r_1^2 C + r_2 r_1 B \end{pmatrix}. \tag{11}$$

This operator on $\mathbb{R}^{2N^2}$ has remarkably simple properties by virtue of the (easily verified) fact that C and B commute. As such, the eigenvalue problem for S can be solved by first finding the eigenvalues of the 2×2 structure in equation (11), treating B and C as if they were scalars. Then, using the eigenvalues of the simultaneously diagonalizable B and C, one has the spectrum of S or any function of it.

The same procedure can be applied to equations (5) and (6). For the vector $Z = \binom{P}{M}$, the value Z' on the next time step is given by applying the operator

$$T \equiv \begin{pmatrix} \frac{1}{2}(I + C) & \frac{\delta}{2}(I - C) \\ \frac{\delta}{2}(I - C) & \frac{1}{2}[B + \delta^2(I + C - B)] \end{pmatrix} = \begin{pmatrix} I & I \\ I & -I \end{pmatrix} S \begin{pmatrix} I & I \\ I & -I \end{pmatrix}^{-1}. \tag{12}$$

Our method of proof will be to show convergence in the Frobenius norm of Z,

$$\|Z\|^2 \equiv \sum_{i,j} (|P_{ij}|^2 + |M_{ij}|^2). \tag{13}$$

We will show that for $N > 2$, $\|Z\|$ is strictly decreasing under the iteration until a certain subset of the matrix elements has vanished. This will imply that the limiting value of P is

$$P_\infty \equiv \frac{I}{N}\,\mathrm{Tr}\left(\xi_1^{(0)} + \xi_1^{(0)}\right) \tag{14}$$

(where (0) indicates initial values) and that the limiting value of M is zero. For $N = 2$, an additional term can survive in P. And finally, for $N = 2$ and $\delta = 0$ a yet more complicated situation prevails (and is described in [2]).

In proving norm convergence it will not be the eigenvalues of S or T that are essential, but their singular values, the square roots of the eigenvalues of $S^\dagger S$ or $T^\dagger T$. We remark that the maximal singular value is always equal to or greater than the maximum of the absolute values of the eigenvalues.

4.1. *Convergence theorem*

For any $R \in SO(N)$, there are N orthonormal real column vectors $v_1, v_1', v_2, v_2', \ldots, v_{\lfloor N/2\rfloor}, v_{\lfloor N/2\rfloor}'$ (and another vector v_0 if N is odd), such that, letting V be the orthogonal matrix of which these are (in this order) the columns, $\Gamma = VRV^{-1}$ is block diagonal with $\lfloor N/2\rfloor 2 \times 2$ blocks descending the diagonal (and possibly a 1×1 block at the bottom). Each block is an $SO(2)$ matrix (the 1×1 block contains a 1). Corresponding to each pair v_j, v_j', there is a pair of eigenvectors $(v_j + \mathrm{i}v_j')/\sqrt{2}$ and $(v_j - \mathrm{i}v_j')/\sqrt{2}$, which are complex conjugates of each other and which we have scaled to have norm 1. We use the following notation: if f_μ is such an eigenvector then $f_{\tilde{\mu}}$ is its complex conjugate. If the $SO(2)$ rotation in the jth subspace is by angle θ_j, then the corresponding eigenvalues are $\exp(\pm \mathrm{i}\theta_j)$. For N odd, $f_0 = v_0$. The identity matrix can be written as $I = \sum_\mu f_\mu f_\mu^\dagger$ and a general matrix A can be written as

$$A = \sum_{\mu,\nu} f_\mu a_{\mu\nu} f_\nu^\dagger \qquad \text{with} \qquad a_{\mu\nu} \equiv f_\mu^\dagger A f_\nu. \tag{15}$$

Note that

$$\sum_{ij} |A_{ij}|^2 = \sum_{\mu\nu} |a_{\mu\nu}|^2. \tag{16}$$

The N^2 eigenvectors of both C_R and B_R are the $N \times N$ matrices $f_\mu f_\nu^\dagger$, specifically

$$C_R f_\mu f_\nu^\dagger = \exp[\mathrm{i}(\theta_\mu - \theta_\nu)] f_\mu f_\nu^\dagger \equiv c_{\mu\nu} f_\mu f_\nu^\dagger \tag{17}$$

and

$$B_R f_\mu f_\nu^\dagger = [\exp(\mathrm{i}\theta_\mu) + \exp(-\mathrm{i}\theta_\nu)] f_\mu f_\nu^\dagger \equiv b_{\mu\nu} f_\mu f_\nu^\dagger. \tag{18}$$

Note that $c_{\mu\nu} = c_{\tilde{\nu}\tilde{\mu}}$ and $b_{\mu\nu} = b_{\tilde{\nu}\tilde{\mu}}$.

The identity matrix is an eigenvector of C_R with eigenvalue 1. (It lies in the eigenspace spanned by $\{f_\mu f_\mu^\dagger\}$.) In general, the identity is not an eigenvector of B_R. An eigenvector A of C_R with eigenvalue 1 commutes with R. For $N > 2$, only if A is a multiple of the identity can it be an eigenvector with eigenvalue 1 for all R; that is to say, only the identity is in the centre of $SO(N)$. In dimension 2, this condition also allows A to be a multiple of the single generator of $SO(2)$ (since $SO(2)$ is Abelian), which is the reason dimension 2 is special with respect to the conjugation iteration.

The matrix T can be seen (because of the fact that B and C commute) to decompose into 2×2 blocks, each acting on the two-dimensional subspace spanned by $\left(f_\mu f_\nu^\dagger, 0\right)$ and $\left(0, f_\mu f_\nu^\dagger\right)$ for some pair μ, ν. Specifically,

$$T = \oplus_{\mu,\nu} \exp(\mathrm{i}\beta) Q_{\mu,\nu} \tag{19}$$

with $Q_{\mu,\nu}$ the 2×2 matrix

$$Q_{\mu,\nu} \equiv \begin{pmatrix} \cos\beta & -\mathrm{i}\delta\sin\beta \\ -\mathrm{i}\delta\sin\beta & (1-\delta^2)\cos\alpha + \delta^2\cos\beta \end{pmatrix}, \tag{20}$$

and $\alpha = (\theta_\mu + \theta_\nu)/2$, $\beta = (\theta_\mu - \theta_\nu)/2$. To determine the asymptotic behaviour of our iteration, we analyse the singular values of Q (we drop the μ, ν subscripts when considering a particular two-dimensional subspace).

Lemma 1. *The maximal singular value of Q is $\leqslant 1$ and equal to 1 only if* $1 \in \{\cos^2\alpha, \cos^2\beta, \delta^2\}$.

The proof is deferred to section 4.2.

Corollary 2. *If $\nu \notin \{\mu, \tilde{\mu}\}$, the operator norm of Q is a.s. strictly less than 1.*

This is because, if $\nu \notin \{\mu, \tilde{\mu}\}$, α and β are both a.s. nonzero.

(The other cases are that $\nu = \mu$, in which case β is zero while α is generically nonzero, or that $\nu = \tilde{\mu}$, in which case α is zero while β is generically nonzero.)

Theorem 3. *For $N > 2$, T a.s. strictly contracts Z unless P is a multiple of the identity and M is zero. From initial conditions $(P^{(0)}, M^{(0)})$, the process tends to the limit $(I/\operatorname{Tr}(P^{(0)}), 0)$.*

Proof. We return to the near-diagonalization $\Gamma = VRV^{-1}$. The distribution on R is the Haar measure on $SO(N)$. Note that this measure is absolutely continuous with respect to the measure obtained by selecting V from the Haar measure on $SO(N)$, Γ from the product over its blocks of the Haar measure on $SO(2)$ and setting $R = V^{-1}\Gamma V$. Therefore, it is enough to show the theorem in this latter measure.

Observe that

$$C_R(A) = RAR^{-1} = V^{-1}\Gamma V A V^{-1}\Gamma^{-1} V = C_{V^{-1}}(C_\Gamma(C_V(A))). \tag{21}$$

Similarly,

$$\begin{aligned} B_R(A) &= RA + AR^{-1} = V^{-1}\Gamma V A + A V^{-1}\Gamma^{-1} V \\ &= C_{V^{-1}}(\Gamma V A V^{-1} + V A V^{-1}\Gamma^{-1}) = C_{V^{-1}}(B_\Gamma(C_V(A))). \end{aligned} \tag{22}$$

Note the similarity of the operators: in both cases there is a change of basis by V, then the B or C action (associated with the near-diagonal matrix Γ), then a change of basis back. For the identity operator, there is a similar factoring: $I(A) = C_{V^{-1}}(I(C_V(A)))$. Therefore, T_R can be written as

$$T_R(Z) = \begin{pmatrix} C_{V^{-1}}\left(\frac{1}{2}[I + C_\Gamma](C_V(P)) + \frac{\delta}{2}[I - C_\Gamma](C_V(M))\right) \\ C_{V^{-1}}\left(\frac{\delta}{2}[I - C_\Gamma](C_V(P)) + \frac{1}{2}[B_\Gamma + \delta^2(I + C_\Gamma - B_\Gamma)](C_V(M))\right) \end{pmatrix}. \tag{23}$$

Since we are only interested in controlling the Frobenius norm, we can drop the external $C_{V^{-1}}$ operator and concentrate on showing that with probability 1, if M is nonzero or P is not a multiple of the identity, the Frobenius norm decreases. Let $C^{\oplus 2}$ denote the direct sum of the conjugation operator with itself: in other words,

$$C_V^{\oplus 2}(Z) = \begin{pmatrix} C_V(P) \\ C_V(M) \end{pmatrix}.$$

Then, we wish to control the Frobenius norm of

$$C_V^{\oplus 2}(T_R(Z)) = \begin{pmatrix} \frac{1}{2}[I + C_\Gamma](C_V(P)) + \frac{\delta}{2}[I - C_\Gamma](C_V(M)) \\ \frac{\delta}{2}[I - C_\Gamma](C_V(P)) + \frac{1}{2}[B_\Gamma + \delta^2(I + C_\Gamma - B_\Gamma)](C_V(M)) \end{pmatrix} \tag{24}$$

$$= \begin{pmatrix} \frac{1}{2}[I + C_\Gamma] & \frac{\delta}{2}[I - C_\Gamma] \\ \frac{\delta}{2}[I - C_\Gamma] & \frac{1}{2}[B_\Gamma + \delta^2(I + C_\Gamma - B_\Gamma)] \end{pmatrix} C_V^{\oplus 2}(Z). \tag{25}$$

Observe that the first operator acting on Z depends only on V, while the second depends only on Γ. We already know that the second operator decomposes as a direct sum over the subspaces described by the eigenvectors of Γ (see equation (19)).

It will be helpful to make one last change of basis. Let

$$h = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & \mathrm{i} \\ 1 & -\mathrm{i} \end{pmatrix}$$

and let H be the block-diagonal matrix that is the direct sum of $\lfloor N/2 \rfloor$ copies of h, followed (if N is odd) by a single 1. Observe that $H\Gamma H^{-1}$ is diagonal and H is unitary; let $\bar\Gamma = H\Gamma H^{-1}$. Now by a derivation similar to that which led us to equation (25), the Frobenius norm of $C_V^{\oplus 2}(T_R(Z))$ is the same as that of

$$\begin{pmatrix} \frac{1}{2}[I + C_{\bar\Gamma}] & \frac{\delta}{2}[I - C_{\bar\Gamma}] \\ \frac{\delta}{2}[I - C_{\bar\Gamma}] & \frac{1}{2}[B_{\bar\Gamma} + \delta^2(I + C_{\bar\Gamma} - B_{\bar\Gamma})] \end{pmatrix} C_H^{\oplus 2}\left(C_V^{\oplus 2}(Z)\right).$$

Observe that $C_H^{\oplus 2} \circ C_V^{\oplus 2}$ leaves the identity matrix invariant within each of the 'P' and 'M' regions. We consider any matrix Z to be the sum of three terms that are orthogonal w.r.t. the Frobenius inner product: first, a multiple of the 'P' identity; second, a multiple of the 'M' identity; and third, a term which has trace 0 in both regions. To show the theorem, we show that if the second or third term is nonzero, the iteration is a.s. strictly contractive. Observe that since every $Q_{\mu,\nu}$ is weakly contractive, it suffices to show that *some* $Q_{\mu,\nu}$ strictly contracts its argument.

To begin with, if the second term is nonzero, then it passes unchanged through $C_H^{\oplus 2} \circ C_V^{\oplus 2}$ and is acted on by the matrices $Q_{\mu\mu}$. An examination of equation (20) shows that these Qs are diagonal and that with probability 1 the multiplier of each 'M' coordinate has norm strictly less than 1.

It remains to show that with probability 1, T is strictly contractive if the second term of Z is zero and the third is nonzero. Thanks to lemma 1 and corollary 2, this will follow from showing that with probability 1, $C_H^{\oplus 2}\big(C_V^{\oplus 2}(Z)\big)$ has a nonzero entry in some subspace $\mathrm{Span}\big(\big(e_j e_k^\dagger, 0\big), \big(0, e_j e_k^\dagger\big)\big)$ for which j, k do not satisfy '$j = k$' or '$|j - k| = 1$ and $\max\{j, k\}$ is even'. (Here, e_j is a standard-basis column vector.) Since $C_H^{\oplus 2} \circ C_V^{\oplus 2}$ acts separately on P and M, this claim is equivalent to showing that if A is nonzero and traceless then $C_H(C_V(A))$ a.s. has nonzero projection on such a vector $e_j e_k^\dagger$. Observe that C_H is a unitary operator which decomposes into a direct sum over the 'blocks' of A; in other words, it has a nonzero coefficient in its $((j, k), (j', k'))$ entry only if (j, k) and (j', k') are within the same 2×2 (or 2×1 or 1×1) block.

Therefore, in order to show that for some j, k of the desired type, $C_H(C_V(A))$ has a nonzero entry, it is enough to show that for some 'off-diagonal block', $C_V(A)$ has a nonzero entry in that block. In fact we will show that with probability 1, *every* off-diagonal entry of $C_V(A)$ is nonzero. This will rely only on A not being a multiple of the identity.

We consider the j, k entry ($j \neq k$). Selecting V from the Haar measure on $SO(N)$, we must show that with probability 1, $\mathrm{Tr}\left(VAV^{-1}e_j e_k^\dagger\right) \neq 0$. We can rewrite the last quantity as $e_k^\dagger VAV^{-1}e_j$. Since A is not a multiple of the identity, the probability that $V^{-1}e_j$ is an eigenvector of A is 0, and therefore a.s. $AV^{-1}e_j$ has nonzero projection on the subspace $(V^{-1}e_j)^\perp$. The vector $V^{-1}e_k$ is, by assumption, chosen uniformly from the unit sphere in this subspace. Hence, with probability 1 it has nonzero inner product with $AV^{-1}e_j$. □

4.2. The singular values of Q: proof of lemma 1

Q can be written as

$$Q = q_0 I + \mathrm{i} q_1 \sigma_1 + q_3 \sigma_3, \tag{26}$$

where

$$q_0 = \tfrac{1}{2}[(1 + \delta^2)\cos\beta + (1 - \delta^2)\cos\alpha], \tag{27}$$

$$q_1 = -\delta \sin\beta, \tag{28}$$

$$q_3 = \tfrac{1}{2}[(1 - \delta^2)\cos\beta - (1 - \delta^2)\cos\alpha], \tag{29}$$

and the Pauli matrices are

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_2 = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{30}$$

The singular values of Q are the square roots of the eigenvalues of $Q^\dagger Q$. A short calculation shows that

$$Q^\dagger Q = \left(q_0^2 + q_1^2 + q_3^2\right) I + 2q_3(q_0\sigma_3 - q_1\sigma_2). \tag{31}$$

The eigenvalues of a matrix of the form $\vec{v} \cdot \vec{\sigma}$ are $\pm\sqrt{\vec{v} \cdot \vec{v}}$, so that the eigenvalues of $Q^\dagger Q$ are

$$\lambda_\pm = \left(q_0^2 + q_1^2 + q_3^2\right) \pm 2q_3\sqrt{q_0^2 + q_1^2}. \tag{32}$$

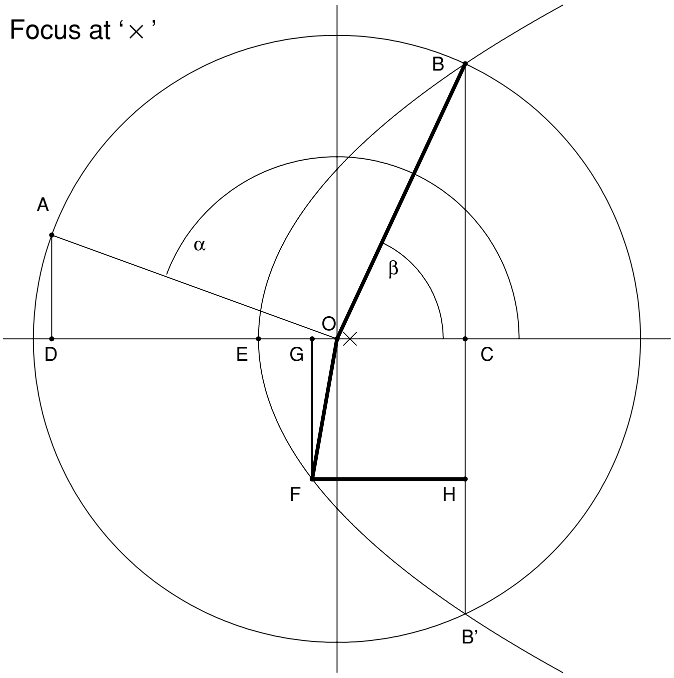

Figure 1. The unit circle and the parabola intersect at B and B'. F is at the point (q_0, q_1), so $u = |OF|$. A brief calculation shows that $q_3 = |FH|$. If the focus of the parabola is at the origin then $|OF| + |FH| = |OB| = 1$ for any F. The origin is always between the focus and base of the parabola; if it does not coincide with the focus, $|OF| + |FH|$ is strictly increasing as a function of $|FG|$, and so is uniquely maximized at $F = H = B'$, with $|OF| + |FH| = |OB| = 1$.

Letting $u \equiv \sqrt{q_0^2 + q_1^2}$, the eigenvalues can be written as

$$\lambda_\pm = (u \pm q_3)^2. \tag{33}$$

We now show that *the eigenvalues* $\lambda_\pm$ *are nonnegative reals less than or equal to 1; the larger eigenvalue equals 1 if and only if* $1 \in \{\cos^2\alpha, \cos^2\beta, \delta^2\}$.

Observe that for $\delta^2 \neq 1$ (which we assume throughout), the special cases occur with probability 0 with the following exceptions: (1) $\mu = \nu$, so $\theta_\mu = \theta_\nu$ and $\beta = 0$ and (2) for conjugate pairs, $\theta_\mu = -\theta_{\bar{\mu}}$, so $\alpha = 0$.

Proof. For $\beta = 0$ or π, Q is Hermitian and diagonal. The eigenvalues of $Q^\dagger Q$ are thus the squares of those of Q, which in turn are $(1, (1-\delta^2)\cos\alpha + \delta^2)$ for $\beta = 0$ and $(-1, (1-\delta^2)\cos\alpha - \delta^2)$ for $\beta = \pi$. We now exclude these β values and proceed to the geometric argument. □

In figure 1, we have drawn a unit circle around the origin *(O)* with the points $A = (\cos\alpha, \sin\alpha)$ and $B = (\cos\beta, \sin\beta)$ indicated. Drop a perpendicular from B to the horizontal, striking the x-axis at C; note that $|OC| = \cos\beta$. Similarly, the projection of A on the x-axis will be designated as D and $|OD| = \cos\alpha$.

Let E be midway between C and D. Draw the parabola that passes through the points B and E and is symmetric about the x-axis. The equation of this parabola is

$$y^2 = \frac{2\sin^2\beta}{\cos\beta - \cos\alpha}\left(x - \frac{\cos\alpha + \cos\beta}{2}\right). \tag{34}$$

If $|\alpha| > |\beta|$ (both are taken in $[0, \pi]$) then the parabola is open to the right, otherwise it is open to the left[4]. The focus of the parabola is at[5]

$$x_{\text{focus}} = \frac{\sin^2\alpha}{2(\cos\beta - \cos\alpha)}. \tag{35}$$

Note that whichever way the parabola is pointing, the origin lies between the point E and the focus.

Henceforth, for specificity, we assume the parabola is open to the right, although the arguments can be carried through in a nearly identical fashion for the other instance.

Let F be the point (q_0, q_1). It is easy to verify that it lies on the parabola and figure 1 reflects this fact. By excluding the case $\delta^2 = 1$, we conclude that $F \notin \{B, B'\}$. The projection from F to the x-axis is designated as G. Thus $|OG| = q_0$ and $|GF| = q_1$. It follows that $|OF| = u$. Furthermore, $|GC| = \cos\beta - q_0 = q_3$.

Let H be the projection of F on the line BC, $|FH| = |GC|$. By definition $|OB| = 1$ so the assertion that the eigenvalues of $Q^\dagger Q$ are equal to or less than 1 is equivalent to showing that $|OF| + |FH| \leqslant |OB|$.

If O and the focal point were to coincide, then by the constructive method of drawing parabolas[6], we would have the equality $|OF| + |FH| = |OB|$. The assertion finally rests on the following extension[7] of the previous statement: when the origin lies between the focus and the base of the parabola, $|OF| + |FH|$ is an increasing function of the distance of F from the x-axis.

Hence, the maximal singular value of Q is $\leqslant 1$ and equal to 1 only if $1 \in \{\cos^2\alpha, \cos^2\beta, \delta^2\}$.

The eigenvectors of Q are not needed in our convergence arguments, but may play a role in convergence rate estimates or in proofs involving measures other than Haar measure. They are given in the appendix.

5. Extensions

We outline a few conjectured extensions of our convergence theorem as well as a result on the relation of the iteration studied here to the physical process in a gas.

[4] The case $\alpha = |\beta|$ is not generic, except for special cases covered elsewhere. Aside from those cases, it can also be seen to yield eigenvalues smaller than 1. For $|\alpha| = |\beta|$, $q_0 = \cos\alpha = \cos\beta$ and $q_3 = 0$. This implies that the eigenvalue is degenerate and $\lambda_\pm = \cos^2\beta + \delta^2\sin^2\beta$, which is less than unity for $\delta^2 \neq 1$:

[5] For $y^2 \neq \mu(x - x_0)$, the x-coordinate of the focus is $x_0 + \mu/4$.

[6] The familiar construction of an ellipse, using a string of fixed length between two points (which become the foci), is extended to parabolas by holding one end of the string on a point (the focal point) and attaching the other to a bead that can move on a line (a line that will lie perpendicular to the axis of the parabola). This line need not be the 'directrix' of the parabola.

[7] For convenience in proving this assertion we simplify (and shift) the parabola to have the form $x = -1 + y^2/4$, with focus at (0, 0). Consider a point between the base of the parabola (at $x = -1$) and the focus, call it Z = $(z, 0)$ for $z < 0$. Since letting F coincide with B$'$ leads to the sum of the lengths being unity and since $\delta\sin\beta < 1$, it is sufficient to show that $u - x$ is an increasing function of y^2, where

$$u = \sqrt{y^2 + (x - z)^2} = \sqrt{y^2 + (-1 + y^2/4 - z)^2}.$$

Differentiate

$$\frac{\partial(u - x)}{\partial y^2} = \frac{2 + x - z - \sqrt{(2 + x - z)^2 + 4z}}{4\sqrt{y^2 + (x - z)^2}}. \tag{36}$$

Looking at the numerator we see that $\operatorname{sgn}\frac{\partial(u-x)}{\partial y^2} = -\operatorname{sgn}(z) > 0$. This shows that |ZP| + |PR| < |ZP| so long as Z lies to the left of the focus, O, an assertion equivalent to the required eigenvalue inequality (see figure 2).

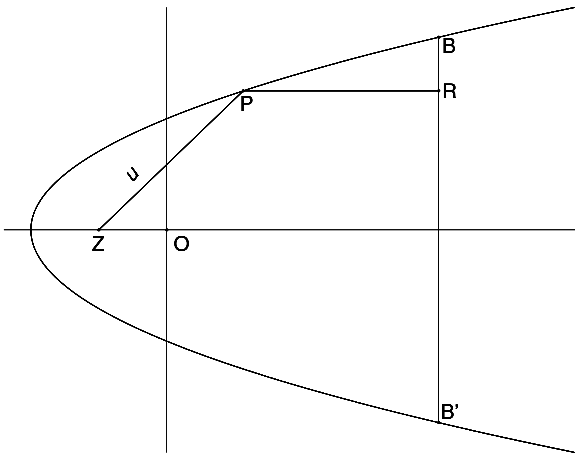

Figure 2. The line BB' is a fixed reference line and the point R is defined by its y-coordinate being the same as that of P and by its being on the line BB'. The distances of interest are $|PR|$, which is a constant minus x_P, and $|ZP|$, which corresponds to 'u'. When BB' is taken to be the y-axis, $r - x - \mu/2$ defines the parabola, $r = |OP|$. (Note that the vertical lines we use are not (necessarily) the 'directrix'.)

5.1. The measure for 'R'

Numerical evidence suggests that the convergence theorem holds if, in place of the Haar measure on $SO(N)$, a measure is used which generates a dense subset of $SO(N)$.

Even more conjecturally, it may be possible to obtain a convergence theorem only under the assumption that the group generated by the measure acts unitarily and irreducibly on $\mathbb{C}^N$.

5.2. Separate rotations

Numerical evidence suggests that if the rotation applied to ξ_1 is different from that applied to ξ_2, but both are selected randomly (as before, from some measure generating $SO(N)$), convergence still holds. In other words, in equation (4) one uses a rotation R in the first line and a different rotation, $\widetilde{R}$, in the second line. (For expectation values, the limiting behaviour can be established by the methods of [2], although in this case the trace of the sum is not preserved step by step.)

5.3. Different coefficients

Numerical evidence suggests that the basic framework still goes through when the coefficients r_1 and r_2 are replaced by any other positive numbers adding to one in the formula for ξ_2. That is, instead of equation (4), one uses

$$\xi_1' = (r_1 + r_2 R)\xi_1(r_1 + r_2 R^{-1}) + r_1 r_2 (1 - R)\xi_2(1 - R^{-1}), \tag{37}$$

$$\xi_2' = (s_2 + s_1 R)\xi_2(s_2 + s_1 R^{-1}) + s_2 s_1 (1 - R)\xi_1(1 - R^{-1}). \tag{38}$$

Like the 'r's the numbers s_1 and s_2 add to one. Again, the sum of the traces is not exactly preserved. Once again, as in [2], the expectation values converge.

5.4. Multi-particle scattering

The physical problem considered in [1] involves a gas of particles, many of each kind (i.e., each mass), scattering off one another. Thus, one should consider two collections of particles: one distribution of spreads for one kind another for the second kind. In the gas these would encounter (most likely) different particles on each scattering event, thereby justifying the tracing out of coordinates needed to establish the repeated conjugation formula studied here. Thus, in the gas, after particle 1 scatters with particle 2, it goes on to encounter particle 3. The density matrix that particle 1 next sees is not that given by the 1–2 scattering, but rather that given by whatever history particle 3 has had. Therefore, a model closer to the physical events should be constructed as follows: from each of the two distributions, randomly draw one particle, let them interact under equation (4) and modify the distributions appropriately.

We next establish that this yields the result that we have already obtained. We model the random process as follows. Take some large number, K, of each kind, and pair them off—one from each collection—with one another. Each collection has statistical properties, an average spread matrix and moments of the matrix. Each pair satisfies the usual relation, equation (4), with a superscript to indicate which member of the distribution is being considered:

$$\begin{aligned} \xi_1^{(m_1)'} &= (r_1 + r_2 R)\xi_1^{(m_1)}(r_1 + r_2 R^{-1}) + r_1 r_2 (1 - R)\xi_2^{(n_1)}(1 - R^{-1}), \\ \xi_2^{(n_1)'} &= (r_2 + r_1 R)\xi_2^{(n_1)}(r_2 + r_1 R^{-1}) + r_2 r_1 (1 - R)\xi_1^{(m_1)}(1 - R^{-1}). \end{aligned} \tag{39}$$

The indices n_k and m_k $(k = 1, \ldots, K)$ are permutations of $\{1, \ldots, K\}$. We emphasize that equation (39) is a linear equation, so that thinking of the pairs (ξ_1, ξ_2) as vectors of length $2N^2$ (where N is the dimension of the underlying space), we can write

$$X' = AX, \tag{40}$$

where $X = \binom{\xi_1}{\xi_2}$ and A is built appropriately from equation (4) or equation (39). It immediately follows that the expectation of X (over the ensemble of the $2K$ particles) obeys exactly the same law as X itself *and has the same limit*. Let the limiting value of X be designated as X_∞ and let $Y \equiv X - X_\infty$. Then Y also obeys $Y' = AY$.

We next consider the standard deviation of the spread distribution, generically denoted as 'σ^2' (not to be confused with the wavefunction spread—it is, aside from factors of r_k, the spread of these spreads). To calculate the evolution of the standard deviation of the spreads we consider $Y^\dagger Y$, which is proportional to σ^2. But then on the next time step (each step consists of a single scattering for each member of the collection) $\sigma'^2 \propto Y'^\dagger Y' = (AY)^\dagger AY$. But AY goes to zero, just as Y does, so not only is there convergence of the mean of the distribution, but also that distribution becomes ever narrower.

5.5. Many different masses

Numerical evidence suggests that convergence holds also when there are more than two kinds of particle. In that case, the spreads again tend to multiples of the identity, with the multiplier again being such that the product $m_k \sigma_k^2$ is the same for all interacting particles.

Acknowledgments

This work was supported by the grants NSF PHY 00 99471, NSF CCF 0524828 and the NSF Institute for Quantum Information.

Appendix. Eigenvectors of $Q_{\mu\tilde{\mu}}$

As apparent from section 4.2, the most subtle case of the analysis of the $Q_{\mu\nu}$ operators is when $\nu = \tilde{\mu}$, that is to say, $\alpha = 0$. In this case, the operator is not, in fact, strictly contractive. It is of course weakly contractive, but it always has an eigenvalue of norm 1. For completeness, we describe here the structure of $Q_{\mu\tilde{\mu}}$.

Let $p \equiv p_{\mu\tilde{\mu}}$, $m \equiv m_{\mu\tilde{\mu}}$ and $\eta \equiv \binom{p}{m}$. It is immediate that $|p'|^2 + |m'|^2 = \eta^\dagger Q^\dagger Q\eta$. If $v_\pm$ is the (column) eigenvector of $Q^\dagger Q$ with eigenvalue $\lambda_\pm$, then

$$|p'|^2 + |m'|^2 = \lambda_- |v_-^\dagger \eta|^2 + \lambda_+ \left|v_+^\dagger \eta\right|^2 = |v_-^\dagger \eta|^2 + (u + q_3)^2 \left|v_+^\dagger \eta\right|^2. \tag{A.1}$$

(Note that for $\alpha = 0$, $q_3 < 0$ and u is positive by definition. The eigenvalue associated with v_-, which is $(u - q_3)^2$, is in this case unity.)

To obtain the vectors $v_\pm$ it is sufficient to look at the non-identity portion of $Q^\dagger Q$. From equation (31) this is proportional to $q_0\sigma_3 - q_1\sigma_2$. Define the matrix S to be a normalized version of this,

$$S \equiv \frac{1}{q_0^2 + q_1^2}(q_0\sigma_3 - q_1\sigma_2) \equiv \begin{pmatrix} \cos\psi & \mathrm{i}\sin\psi \\ -\mathrm{i}\sin\psi & \cos\psi \end{pmatrix}. \tag{A.2}$$

Equation (A.2) defines the angle ψ and a bit of calculation shows that

$$\tan\psi = \frac{q_1}{q_0} = \frac{1}{\tan(\beta/2)}\frac{\delta}{1-\delta^2} = \frac{2\sinh(\log\delta)}{\tan(\beta/2)}. \tag{A.3}$$

The eigenvectors of S are

$$v_- = \begin{pmatrix} -\mathrm{i}\sin(\psi/2) \\ \cos(\psi/2) \end{pmatrix}, \qquad v_+ = \begin{pmatrix} \cos(\psi/2) \\ -\mathrm{i}\sin(\psi/2) \end{pmatrix} \tag{A.4}$$

which are respectively associated with the minus and plus eigenvalues of $Q^\dagger Q$, namely, 1 and $(u + q_3)^2$.

References

[1] Schulman L S 2004 Evolution of wave-packet spread under sequential scattering of particles of unequal mass *Phys. Rev. Lett.* **92** 210404

[2] Schulman L S and Schulman L J 2005 Wave packet scattering without kinematic entanglement: convergence of expectation values *IEEE Trans. Nanotechnol.* **5** 8–13