To: *C. Fuchs, K. Manne, J. Renes, and R. Schack*
From: *C. M. Caves*
Subject: **Linear dynamics that preserves maximal information is Hamiltonian**
2001 January 7; modified 2001 June 23 to include the relation to Kadison's theorem

The objective of this brief report is to show that any dynamics that preserves maximal information is unitary and thus is generated by a Hamiltonian. We start from the fact that any dynamics on density operators should be a convex linear map $\mathcal{R}$. We can extend $\mathcal{R}$ linearly to the positive cone by defining $\mathcal{L}(a\rho) = a\mathcal{R}(\rho)$, where $a \geq 0$. We can then extend it linearly to any Hermitian operator A by writing $A = E_1 - E_2$, where E_1 and E_2 are positive operators and defining $\mathcal{L}(A) = \mathcal{L}(E_1) - \mathcal{L}(E_2)$; it is easy to show that this extension is well defined even though the operators E_1 and E_2 are not unique. The resulting superoperator $\mathcal{L}$ is positive. We show that if $\mathcal{L}$ preserves maximal information—by taking pure states to pure states in a one-to-one way—then $\mathcal{L}$ corresponds to unitary or antiunitary evolution. Since antiunitary evolution cannot be reached from the identity, only unitary evolution is allowed.

This result says that when we start with maximal information about a system and we neither acquire nor lose any information, we are not free to use just any sort of positive evolution that comes into our heads. Instead we must use Hamiltonian evolution to update the subjective quantum state. Where does the Hamiltonian come from? It must be an objective property of the system, which we must know in order the preserve maximal information and which we can determine by observing the system's behavior in a wide variety of circumstances. Thus this seemingly innocuous result confirms the not surprising notion that a Hamiltonian is *the* objective property of a quantum system, embodying all the physical properties of the system. Other kinds of completely positive dynamics are seen (in the standard way) as coming from situations where we either acquire new information about a system or lose information by allowing it to become diffused in classical correlations or entanglement with other systems.

Now let's get started. All we need to assume is that the dynamics is described by a convex linear map $\mathcal{R}$ on density operators, which maps one-dimensional projectors one-to-one to one-dimensional projectors:

$$\mathcal{R}(|\psi\rangle\langle\psi|) = |\tilde{\psi}\rangle\langle\tilde{\psi}| \ . \tag{1}$$

The map $\mathcal{R}$ can be extended to a positive superoperator $\mathcal{L}$ as described above, and it is easy to show from the assumption about its action on pure states that $\mathcal{L}$ is trace preserving. The one-to-one assumption is motivated by the notion that one is losing information if many pure states map to the same pure state.

One small technical point. All we really know is that $\mathcal{L}$ is linear on the real vector space of Hermitian operators; when we extend the action of $\mathcal{L}$ to all operators, we are free to extend it in any way we wish—we could make it antilinear on operators, for example—but since it doesn't matter, we choose the extension to be linear. For example, if $\mathcal{L}$ corresponds to a unitary operator U, then it is the superoperator $\mathcal{L} = U \odot U^\dagger$. In contrast, if $\mathcal{L}$ corresponds to an antiunitary operator $U\mathcal{C}$, where $\mathcal{C}$ is complex conjugation in a standard basis, then $\mathcal{L}$ can be the antilinear superoperator $\mathcal{L}_1 = U\mathcal{C} \odot \mathcal{C}U^\dagger$ or it can be the linear superoperator $\mathcal{L}_2 = U \odot U^\dagger \circ \mathcal{T}$, where $\mathcal{T}$ is transposition in the standard basis. $\mathcal{L}_1$ and $\mathcal{L}_2$ have the same action on Hermitian operators.

The superoperator $\mathcal{L}$ induces a map on normalized vectors, the tilde map, defined by

$$T(|\psi\rangle) = |\tilde{\psi}\rangle \ . \tag{2}$$

The tilde map can be rephased as we please; i.e.,

$$|\tilde{\psi}'\rangle = T'(|\psi\rangle) = e^{i\nu(|\psi\rangle)}T(|\psi\rangle) = e^{i\nu(|\psi\rangle)}|\tilde{\psi}\rangle \tag{3}$$

works just as well as the original map. Our strategy is to show that given two normalized state vectors, $|\psi\rangle$ and $|\phi\rangle$, the tilde map preserves the absolute value of the inner product, i.e.,

$$|\langle\phi|\psi\rangle| = |\langle\tilde{\phi}|\tilde{\psi}\rangle| \ . \tag{4}$$

We can then invoke Wigner's theorem to assert that the tilde map can be rephased to be either unitary or antiunitary. Since the superoperator is determined by its action on pure states, it must be derived either from a unitary or an antinunitary. Since antiunitary maps are not connected to the identity, they are unsuitable for describing quantum dynamics, and thus we are left with unitary dynamics as the only possibility.

Now return to those two vectors, $|\psi\rangle$ and $|\phi\rangle$. They span a two-dimensional subspace. Let $|0\rangle$ and $|1\rangle$ be orthonormal vectors in this subspace. The tilde map takes these two orthonormal vectors to linearly independent (one-to-one assumption), but not necessarily orthogonal vectors $|\tilde{0}\rangle$ and $|\tilde{1}\rangle$, which also span a two-dimensional subspace. Set up Pauli operators in the input subspace, defined relative to the basis $|0\rangle$ and $|1\rangle$, and set up Pauli operators in the output subspace, defined relative to $|\tilde{0}\rangle$ and $|\tilde{0}_\perp\rangle$.

We don't know at this point that the tilde map takes all the vectors in the input subspace to vectors in the output subspace. To allow for the possibility that it doesn't, we introduce the projector $\tilde{\Pi}$ onto the output subspace. The superoperator $\mathcal{L}' \equiv \tilde{\Pi} \odot \tilde{\Pi} \circ \mathcal{L}$ ($\mathcal{L}$ followed by projection onto the output subspace) maps any operator to an operator that lives in the two-dimensional output subspace. Moreover, it takes all one-dimensional projectors in the input subspace to subnormalized one-dimensional projectors in the output subspace. Notice that the output is normalized to unity if and only if $\tilde{\Pi}$ has no effect, i.e., the projector produced by $\mathcal{L}$ already lies in the output subspace.

Now we're set. The linearity of $\mathcal{L}'$ means that

$$\begin{aligned} \mathcal{L}'(1) &= \alpha 1 + \vec{t}\cdot\vec{\sigma} \ , \\ \mathcal{L}'(\sigma_j) &= s_j 1 + M_{jk}\sigma_k \ . \end{aligned} \tag{5}$$

The 3×3 matrix M can be written as

$$M = \sqrt{G} O \;, \quad G = M M^T \;, \tag{6}$$

where O is an orthogonal matrix and G is a positive matrix. A one-dimensional input projector is mapped to

$$\begin{aligned} \mathcal{L}'\bigg(\frac{1}{2}(1 + \vec{n}\cdot\vec{\sigma})\bigg) &= \frac{1}{2}(\alpha 1 + \vec{t}\cdot\vec{\sigma} + \vec{n}\cdot\vec{s}\,1 + n_j M_{jk}\sigma_k) \\ &= \frac{1}{2}[(\alpha + \vec{s}\cdot\vec{n})1 + (\vec{t} + M^T\vec{n})\cdot\vec{\sigma}] \;. \end{aligned} \tag{7}$$

The first restriction we have on this map is that $|0\rangle\langle 0|$ ($\vec{n} = \vec{e}_z$) and $|1\rangle\langle 1|$ ($\vec{n} = -\vec{e}_z$) map to normalized one-dimensional projectors, $|\tilde{0}\rangle\langle\tilde{0}|$) and $|\tilde{1}\rangle\langle\tilde{1}|$). This means that

$$\alpha \pm s_z = 1 \quad \Longrightarrow \quad \alpha = 1 \text{ (important) and } s_z = 0 \;. \tag{8}$$

But now we see that $\vec{s}$ must be zero, for were it not, then choosing $\vec{n}$ along $\vec{s}$ would give a supernormalized output.

With these results, we are back to using $\mathcal{L}$, since the projector $\tilde{\Pi}$ has no effect (i.e., all input projectors are mapped to normalized output projectors), and we can use the qubit proof that I constructed last summer. Equation (7) simplifies to

$$\mathcal{L}\bigg(\frac{1}{2}(1 + \vec{n}\cdot\vec{\sigma})\bigg) = \frac{1}{2}(1 + \vec{m}\cdot\vec{\sigma}) \;, \tag{9}$$

where

$$\vec{m} \equiv \vec{t} + M^T\vec{n} \tag{10}$$

must be a unit vector. It's pretty obvious that this affine map takes unit vectors to unit vectors only if (i) $\vec{t} = 0$ and $M = O$ or (ii) $\vec{t}$ is a unit vector and $M = 0$, but the next two paragraphs give a formal proof.

Writing out the requirement that $\vec{m}$ be a unit vector, we get

$$1 = \vec{m}\cdot\vec{m} = t^2 + 2\,\vec{t}\cdot M^T\vec{n} + M^T\vec{n}\cdot M^T\vec{n} = t^2 + 2M\vec{t}\cdot\vec{n} + \vec{n}\cdot G\vec{n} \;. \tag{11}$$

Defining the trace-free part of G by

$$F \equiv G - \frac{1}{3}\mathrm{tr}(G) I \;, \tag{12}$$

we can write Eq. (11) as

$$1 = t^2 + \frac{1}{3}\mathrm{tr}(G) + 2\sqrt{G} O\vec{t}\cdot\vec{n} + \vec{n}\cdot F\vec{n} \;. \tag{13}$$

The uniqueness of expansions in trace-free tensors (or, equivalently, in spherical harmonics) implies that

$$\begin{aligned} 1 &= t^2 + \frac{1}{3}\mathrm{tr}(G) \\ 0 &= \sqrt{G}O\vec{t} \\ 0 &= F\ . \end{aligned} \tag{14}$$

The first and third of these requirements imply that

$$G = \frac{1}{3}\mathrm{tr}(G)I = (1 - t^2)I\ . \tag{15}$$

The second requirement implies one of two cases: (i) $\vec{t} = 0$, which implies that $G = I$ and $M = O$ or (ii) $O\vec{t}$ is a null eigenvector of G, which in view of Eq. (15), means that $G = 0 = M$ and $\vec{t}$ is a unit vector. The second case, i.e., $\vec{m} = \vec{t}$ regardless of $\vec{n}$, is the one where all pure states in the input subspace map to the same output pure state $\frac{1}{2}(1 + \vec{t}\cdot\vec{n})$. We can rule out this second case using our one-to-one requirement. The first case corresponds to a unitary mapping of the input subspace to the output subspace if $\det O = +1$ and to an antiunitary mapping if $\det O = -1$. Since unitary and antiunitary mappings both preserve absolute values of inner products, we have the desired result that absolute values of inner products are preserved by $\mathcal{L}$ and by the tilde mapping.

The result shown here is called *Kadison's theorem* [R. Kadison, *Topology* **3** (Suppl. 2), 177–198 (1965); I learned about Kadison's theorem by Howard Barnum's referring me to B. Simon, "Quantum dynamics: From automorphism to Hamiltonian," in *Studies in Mathematical Physics*, edited by E. H. Lieb, B. Simon, and A. S. Wightman (Princeton University Press, Princeton, NJ, 1976), pp. 327–349]. The theorem is usually stated in the following way

> A *Kadison automorphism* $\mathcal{Q}$ is a one-to-one, onto map from density operators to density operators that is convex linear. Any Kadison automorphism is of the form $\mathcal{Q}(\rho) = U\rho U^\dagger$, where U is unitary or antiunitary.

It would be nice to show that the premises in the two versions are equivalent (without going through the conclusion, which clearly makes them equivalent), but I'm only able to do this in one direction.

Suppose that $\mathcal{Q}$ is a Kadison automorphism. Its inverse is also a Kadison automorphism. It maps pure states to pure states for the following reason. Let

$$\mathcal{Q}(|\psi\rangle\langle\psi|) = \rho = \sum_j \lambda_j |e_j\rangle\langle e_j|\ , \tag{16}$$

where the right side is the eigendecomposition of ρ. Taking the inverse, we get

$$|\psi\rangle\langle\psi| = \sum_j \lambda_j \rho_j\ , \tag{17}$$

where $\rho_j = \mathcal{Q}^{-1}(|e_j\rangle\langle e_j|)$. But now the decomposition theorem for density operators guarantees that $\rho_j = |\psi\rangle\langle\psi|$ for all j. This implies that $\mathcal{Q}(|\psi\rangle\langle\psi|) = |e_j\rangle\langle e_j|$ for all j, which means that $\rho = \mathcal{Q}(|\psi\rangle\langle\psi|)$ is a pure state. Thus the Kadison automorphism is a convex linear map on density operators that maps pure states to pure states one-to-one.

I have not been able to show (without going through the conclusion) that the premises for my theorem imply directly that $\mathcal{R}$ is one-to-one on all density operators, which suggests that my premises are *a priori* weaker than the assumption of a Kadison automorphism.

I can think of three further projects along this line that might prove useful.

1. Show that my premises imply that $\mathcal{R}$ is a Kadison automorphism without going through the conclusion.
2. Try to push through the proof directly in D dimensions, without going through Wigner's theorem. Since the above proof gets one all the way to unitarity or antiunitarity in each two-dimensional subspace, it seems that one ought to be able to derive the unitarity or antiunitarity directly in all dimensions. One approach would be to use the complete set of trace-free, Hermitian operators λ_j. These operators give a generalized-Bloch-sphere description of quantum states, the differences from the two-dimensional Bloch sphere being, first, that although all pure states are on the surface of the generalized Bloch sphere, not all points on the surface correspond to states and, second (and closely related), that not all orthogonal transformations represent unitaries. I have, however, an efficient way to describe the pure states and the orthogonal transformations that correspond to unitaries, so it might be possible to generalize the above proof. The proof might be made easier by the fact that one knows the assertion is true.
3. See what other possibilities arise if one drops the insistence on one-to-one maps. The resulting maps in higher dimensions might be what is necessary to characterize completely the set of positive maps. It might be useful to have completed the first project before tackling this one.