No-Broadcasting Theorem and Its Classical Counterpart

Amir Kalev*
Department of Physics, Technion-Israel Institute of Technology, Haifa 32000, Israel

Itay Hen†
Raymond and Beverly Sackler School of Physics and Astronomy, Tel-Aviv University, Tel-Aviv 69978, Israel
(Received 20 November 2007; published 29 May 2008)

Although it is widely accepted that ''no-broadcasting''—the nonclonability of quantum information—is a fundamental principle of quantum mechanics, an impossibility theorem for the broadcasting of general density matrices has not yet been formulated. In this Letter, we present a general proof for the no-broadcasting theorem, which applies to arbitrary density matrices. The proof relies on entropic considerations, and as such can also be directly linked to its classical counterpart, which applies to probabilistic distributions of statistical ensembles.

DOI: 10.1103/PhysRevLett.100.210502 PACS numbers: 03.67.Hk, 89.70.Hj

Concepts from quantum information theory have been shown to provide new insights into profound topics relating to fundamental features of quantum mechanics, such as the uncertainty principle [1], interference [2], entanglement [3], and the connection to the second law of thermodynamics [4]. A hallmark of quantum mechanics is that quantum information cannot be cloned [5–7]. The enormous impact of this theorem, which was called the ''no-broadcasting'' theorem, is reflected by several studies that focus on various aspects of the nonclonability of quantum information [8–12].

Although it is widely accepted that no-broadcasting is a fundamental principle of quantum mechanics, an impossibility theorem for the broadcasting of general arbitrary (i.e., finite- as well as infinite-dimensional) density matrices has not yet been formulated. In the literature two separate proofs for no-broadcasting are to be found, one applies only to pure states (the no-cloning theorem) [5] while the other applies only to invertible density matrices [6]. These two classes of states exclude each other, and hence, trivially, none of the two proofs is derivable from one another. (Although for the finite-dimensional case, a generalization for noninvertible density matrices exists [7].)

In this Letter we present a general proof for the no-broadcasting theorem which applies to arbitrary density matrices. Our proof relies on fundamental principles from information theory, mainly on entropic considerations. As such, it also enables us to directly link the theorem to its classical analogue which applies to probabilistic distributions of statistical ensembles [13].

A general broadcasting machine consists of a source system whose unknown state σ is to be broadcast, a target system onto which the source state should be copied, and an auxiliary system, or a ''machine,'' which interacts unitarily with the source and target systems. Labeling the three subsystems by subscripts s, t, and m, respectively, the broadcasting process then reads

$$\rho^{\text{in}} = \sigma_s \otimes \tau_t \otimes \Sigma_m \longrightarrow \rho^{\text{out}}, \tag{1}$$

where the final state ρ^{out} obeys

$$\text{Tr}_{t,m}[\rho^{\text{out}}] = \text{Tr}_{s,m}[\rho^{\text{out}}] = \sigma, \tag{2}$$

where $\text{Tr}_{t(s),m}$ denote partial traces over the target (source) and auxiliary systems. In what follows we show that no unitary (quantum mechanical) transformation which performs process (1) exists for arbitrary source states.

Our proof is based on the concept of relative entropy. The relative entropy of a state ρ_1 with respect to another ρ_2 [14]

$$S(\rho_1|\rho_2) = \text{Tr}[\rho_1(\log\rho_1 - \log\rho_2)] \tag{3}$$

is a measure of the ''closeness'' between the two. For some pairs of states (''perfectly distinguishable'' ones) the relative entropy is ill defined. This happens if (and only if) $\ker(\rho_2) \subseteq \ker(\rho_1)$, yielding $S(\rho_1|\rho_2) = \infty$ [15]. For what follows we consider only the case $S(\rho_1|\rho_2) < \infty$, and address in detail the problematic issues which may arise from this ill-definiteness, later on.

One important property of relative entropy is that it is invariant under dynamical changes. The evolution of a general quantum system represented by a density operator ρ is given by $\rho(t) = U(t)\rho(0)U^\dagger(t)$ where $U(t)$ may be any unitary operator. Since the relative entropy is defined by a trace operation, it is easy to check that it is conserved under time evolution, namely

$$S(\rho_1(t)|\rho_2(t)) = S(\rho_1(0)|\rho_2(0)). \tag{4}$$

Let us now consider two general broadcasting processes (1), whose initial states are given by $\rho_i^{\text{in}} = \sigma_{i,s} \otimes \tau_t \otimes \Sigma_m$, where $i = 1, 2$, σ_i are arbitrary density matrices, and the initial states of the target and auxiliary systems, τ and Σ, are the same for both processes. The relative entropy of the two states is

$$\begin{aligned} S(\rho_1^{\rm in}|\rho_2^{\rm in}) &= {\rm Tr}[\sigma_{1,s}\otimes\tau_t\otimes\Sigma_m(\log\sigma_{1,s}\oplus\log\tau_t\oplus\log\Sigma_m \\ &\quad -\log\sigma_{2,s}\oplus\log\tau_t\oplus\log\Sigma_m)] \\ &= {\rm Tr}_s[\sigma_{1,s}(\log\sigma_{1,s}-\log\sigma_{2,s})]{\rm Tr}_{m,t}[\tau_t\otimes\Sigma_m] \\ &= {\rm Tr}_s[\sigma_{1,s}(\log\sigma_{1,s}-\log\sigma_{2,s})] = S(\sigma_1|\sigma_2). \end{aligned} \quad (5)$$

That is, the relative entropy of the two initial states is exclusively given by the relative entropy of the two source systems. Using this and the conservation of relative entropy in time (4), it is clear that

$$S(\sigma_1|\sigma_2) = S(\rho_1^{\rm in}|\rho_2^{\rm in}) = S(\rho_1^{\rm out}|\rho_2^{\rm out}). \quad (6)$$

The relative entropy of the final states of any two broadcasting processes is equal to the relative entropy of the sources prior to copying.

We now proceed to show that Eq. (6) is violated for broadcasting processes. To do this, we invoke the theorem of monotonicity of relative entropy [16] which reads

$$S(\rho_{1,AB}|\rho_{2,AB}) \geq S(\rho_{1,B}|\rho_{2,B}), \quad (7)$$

where $\rho_{1,AB}$ and $\rho_{2,AB}$ are two density operators of a composite system AB, whereas $\rho_{1,B}$ and $\rho_{2,B}$ denote the corresponding density operators of a subsystem B. The equality holds if and only if the condition

$$\log\rho_{1,AB} - \log\rho_{2,AB} = I_A\otimes(\log\rho_{1,B}-\log\rho_{2,B}), \quad (8)$$

evaluated after a restriction to the support of $\rho_{2,AB}$, is satisfied, and I_A denotes the identity operator of subsystem A [17]. Intuitively, Eq. (8) means that ignoring part of two physical systems reduces the "distance" between them, unless the ignored part contains no information at all. Using (7), we can establish a lower bound for the relative entropy of the two final states $\rho_i^{\rm out}$. The monotonicity inequality (7) implies that the final states fulfill

$$S(\rho_1^{\rm out}|\rho_2^{\rm out}) \geq S(\rho_{1,k}^{\rm out}|\rho_{2,k}^{\rm out}), \quad (9)$$

for $k = s, t$ where $\rho_{i,s(t)}^{\rm out}$ denotes ${\rm Tr}_{t(s),m}[\rho_i^{\rm out}]$. According to Eq. (8), the equality in (9) holds if and only if the equalities

$$\begin{aligned} \log\rho_1^{\rm out} - \log\rho_2^{\rm out} &= (\log\rho_{1,s}^{\rm out} - \log\rho_{2,s}^{\rm out})\otimes I_t\otimes I_m \\ &= I_s\otimes(\log\rho_{1,t}^{\rm out}-\log\rho_{2,t}^{\rm out})\otimes I_m, \end{aligned} \quad (10)$$

evaluated on the support of $\rho_2^{\rm out}$, are satisfied. Under broadcasting, Eq. (10) thus reads

$$\begin{aligned} \log\rho_1^{\rm out} - \log\rho_2^{\rm out} &= (\log\sigma_{1,s}-\log\sigma_{2,s})\otimes I_t\otimes I_m \\ &= I_s\otimes(\log\sigma_{1,t}-\log\sigma_{2,t})\otimes I_m. \end{aligned} \quad (11)$$

The above condition, however, is satisfied only if σ_1 and σ_2 are diagonal, reflecting the fact that a realization of a broadcasting machine may be possible only provided that all its input states are mutually commuting and the basis in which they are diagonal is known [18]. For any two noncommuting arbitrary states the inequalities in (9) are strict, that is,

$$S(\rho_1^{\rm out}|\rho_2^{\rm out}) > S(\sigma_1|\sigma_2), \quad (12)$$

in contradiction with equality (6). We have therefore shown that under broadcasting, the monotonicity of relative entropy is in conflict with quantum dynamics, rendering universal broadcasting impossible.

To complete our proof, let us consider the case of $S(\rho_1|\rho_2) = \infty$, and show that the no-broadcasting theorem may be extended to this case as well [19]. This is done using a proof by contradiction. Let us first assume the existence of a machine capable of broadcasting states with infinite relative entropy, and consider two such noncommuting states σ_1 and σ_2 for which $S(\sigma_1|\sigma_2) = \infty$. Because of the linearity of the broadcasting procedure (containing only unitary operations and partial traces), it immediately follows that the machine is also capable of broadcasting the mixture $\sigma_{\rm mix} = \lambda\sigma_1 + (1-\lambda)\sigma_2$ for any $0 < \lambda < 1$. However, our main proof rules out the existence of a machine which broadcasts both σ_1 and $\sigma_{\rm mix}$, since $S(\sigma_1|\sigma_{\rm mix}) < \infty$ [20]. Therefore, the existence of a machine which broadcasts both σ_1 and σ_2 is also ruled out, contradictory to our initial assumption, and this completes the proof.

At this point, we turn to show that the proof given above enables a direct link between the quantum theorem and its classical analogue [13]. The classical no-broadcasting theorem states that it is impossible to broadcast classical probability distributions with unit fidelity in a deterministic manner once infinitely-narrow distributions (i.e., delta-function distributions) are excluded; assuming Liouville evolution for the broadcasting process, the monotonicity of the classical relative entropy (the Kullback-Leibler information distance) between two classical probability distributions $P_1(x, p, t)$ and $P_2(x, p, t)$, defined by [21]

$$\mathcal{K}(P_1|P_2) = \int dxdpP_1(\log P_1 - \log P_2), \quad (13)$$

is in conflict with broadcasting. (We note here, however, that approximate classical broadcasting machines may in principle be realized with any desired degree of accuracy [22].)

As we shall now show, the quantum no-broadcasting theorem translates in the $\hbar\to 0$ limit to its classical analogue if Hamiltonian dynamics, which is a subclass of Liouville dynamics, is concerned. This will be accomplished in two steps. First, we show that for every classical probability distribution one can construct a corresponding density matrix such that in the classical limit, quantum relative entropy reduces to classical relative entropy. Second, we show that quantum (unitary) dynamics reduces in the classical limit to Hamiltonian dynamics under this correspondence. Even though these two statements seem reasonable, even expected, to the best of our knowledge they have not yet been shown explicitly.

As a preliminary step, we make a classical-quantum correspondence by assigning to each classical probability distribution over phase space $P(x, p)$ (we drop the time index t), a quantum state according to

$$\rho = \int dx dp P(x, p) |\alpha\rangle\langle\alpha|, \tag{14}$$

where $|\alpha\rangle$ is a coherent state with $\alpha = (1/\sqrt{2\hbar\omega})(\omega x + ip)$ (we shall fix $\omega = 1$ in the following).

This correspondence is of course the identification of classical probability distributions with the P representations [23] of density matrices. Although the P representation is known to be problematic, being highly singular, negative, or even undefined, we stress here that these types of states are not of our concern here, since in our correspondence, the P distributions we consider are *bona fide* classical distributions.

First, we show that in the classical limit, the relative entropy of two density matrices constructed from two classical statistical distributions by (14) reduces to the relative entropy of the two distributions, namely

$$\lim_{\hbar\to 0} S(\rho_1|\rho_2) = \mathcal{K}(P_1|P_2). \tag{15}$$

Expanding the logarithms appearing in the expression for the quantum relative entropy in a Taylor series and then tracing term by term, it becomes sufficient to show that

$$\lim_{\hbar\to 0} \mathrm{Tr}[\rho_1(\rho_2/\hbar)^{n-1}] = \int dx dp P_1(x, p) P_2^{n-1}(x, p), \tag{16}$$

where $n \in \mathbb{Z}^+$, and the extra $\hbar$ factors appearing in (16), are introduced into the relative entropy by rewriting $(\log\rho_1 - \log\rho_2)$ as $[\log(\rho_1/\hbar) - \log(\rho_2/\hbar)]$. Inserting (14) into the left-hand-side of Eq. (16), we have

$$\lim_{\hbar\to 0} \mathrm{Tr}[\rho_1(\rho_2/\hbar)^{n-1}] = \int dx_0 dp_0 P_1(x_0, p_0) \int \left(\prod_{i=1}^{n-1} dx_i dp_i P_2(x_i, p_i) \right) \lim_{\hbar\to 0} \frac{\exp[-(2\hbar)^{-1} \mathbf{u}^\dagger V \mathbf{u}]}{(2\pi\hbar)^{n-1}}. \tag{17}$$

where $\mathbf{u}^\dagger = (x_0, p_0, x_1, p_1, \cdots, x_{n-1}, p_{n-1})$ and V, presented in a $(2 \times 2) \otimes (n \times n)$ block form is

$$V_{(2n\times 2n)} = \begin{pmatrix} 1 & B & 0 & \cdots & 0 & B^T \\ B^T & 1 & B & 0 & \cdots & 0 \\ 0 & B^T & 1 & B & 0 & \vdots \\ \vdots & 0 & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & B^T & 1 & B \\ B & 0 & \cdots & 0 & B^T & 1 \end{pmatrix}_{(n\times n)}, \tag{18}$$

1 and 0 being the (2×2) unit and zero matrices, respectively, and B^T is the transpose of $B = -\frac{1}{2}(1 - \sigma_y)$. To evaluate the limit, we note that V is a normal matrix and as such it can be written in the form $V = UDU^\dagger$ where D is its diagonal form and U is unitary with orthonormal eigenvector basis as its columns. Computation of these eigenvectors yields

$$\mathbf{e}_{kj} = \frac{1}{\sqrt{2n}} \begin{pmatrix} (-1)^k \\ i \end{pmatrix} \otimes \begin{pmatrix} 1 \\ \omega_j \\ \omega_j^2 \\ \vdots \\ \omega_j^{n-1} \end{pmatrix}, \tag{19}$$

with corresponding eigenvalues $\mu_{kj} = 1 - \omega_j^{(-1)^k}$ where $\omega_j = e^{2\pi i j/n}$, $k = 1, 2$ and $j = 0, \ldots, n-1$. Noting that $\mu_{1,0} = \mu_{2,0} = 0$, the term $\mathbf{u}^\dagger V \mathbf{u}$ in the exponent of (17) can thus be simplified to $\mathbf{u}^\dagger V \mathbf{u} = \mathbf{v}^\dagger D \mathbf{v} = \sum_{k=1}^{2} \sum_{j=1}^{n-1} \mu_{kj} v_{kj}^2$, with $\mathbf{v}^\dagger \equiv \mathbf{u}^\dagger U$. The limit thus becomes

$$\lim_{\hbar\to 0} \frac{\exp[-(2\hbar)^{-1} \mathbf{v}^\dagger D \mathbf{v}]}{(2\pi\hbar)^{n-1}} = \prod_{i=1}^{n-1} \delta(x_i - x_0)\delta(p_i - p_0). \tag{20}$$

This completes the derivation of Eq. (16), and thus proves (15).

Next, we turn to prove that the limit given in (15) holds under time evolution. This is achieved by showing that quantum dynamics is reduced to Hamiltonian dynamics in the $\hbar \to 0$ limit, provided an appropriate correspondence between classical and quantum systems is made. The proof is as follows.

In classical mechanics, a statistical distribution $P_C(x, p)$ evolving in time (the time index t is suppressed) under some Hamiltonian $H(x, p)$ obeys the well-known Liouville equation [24]. In terms of the characteristic (Fourier transformed) function defined by $P_C(x, p) = \int d\lambda d\mu \tilde{P}_C(\lambda, \mu) e^{i(\lambda x + \mu p)}$, and an analogous definition for $H(x, p)$, the equation translates to

$$\partial_t \tilde{P}_C(\lambda, \mu) = \int d\lambda' d\mu' \tilde{P}_C(\lambda', \mu') \tilde{H}(\lambda - \lambda', \mu - \mu') \times K_C(\lambda, \mu, \lambda', \mu'), \tag{21}$$

with a "classical kernel" $K_C = \lambda'\mu - \lambda\mu'$. Accordingly, a general quantum state (also written in characteristic form)

$$\rho = \int dx dp \int d\lambda d\mu e^{i(\lambda x + \mu p)} \tilde{P}_Q(\lambda, \mu) |\alpha\rangle\langle\alpha|, \tag{22}$$

whose evolution is governed by the Hamiltonian [25]

$$\hat{H} = \frac{1}{2\pi\hbar}\int dxdp\int d\lambda d\mu e^{i(\lambda x+\mu p)}\tilde{H}(\lambda,\mu)|\alpha\rangle\langle\alpha| \tag{23}$$

obeys the von Neumann equation $\partial_t\rho = i\hbar^{-1}[\rho, \hat{H}]$. Expressing the equation in terms of $\tilde{P}_Q(\lambda,\mu)$ and $\tilde{H}(\lambda,\mu)$, the equation takes the form (21) but with a "quantum kernel" $K_Q = \frac{2}{\hbar}e^{\frac{\hbar}{2}(\lambda'(\lambda-\lambda')+\mu'(\mu-\mu'))}\sin\frac{\hbar}{2}(\lambda'\mu - \mu'\lambda)$. It is easy to check that $\lim_{\hbar\to 0}K_Q = K_C$; thus, we have shown that *every* classical system may be viewed as a limiting case of an appropriately constructed quantum system. Together with the result of the classical limit of the relative entropy (15), the no-broadcasting theorem which states that the monotonicity of relative entropy of two density operators is in conflict with quantum dynamics under broadcasting, translates in the classical limit to its classical version [13], stating that the monotonicity of (classical) relative entropy is in conflict with Hamiltonian dynamics.

As with other results from quantum mechanics that have their analogies and parallels in classical probabilistic theories [22,26–29], the classical no-broadcasting theorem can also be recovered from its quantum version. This reduction is attributed to the fact that both quantum and classical information theories are based on common grounds and are described by analogous measures.

We have thus shown that no-broadcasting is indeed a general principle, originating from fundamental concepts of information theory, in particular, the monotonicity of relative entropy.

We believe that this may help to gain a better understanding of the relations between nonclonability and reversibility properties both in quantum and in classical physics. This proof may also provide a further clarification on "quantumness" versus "classicality" in that context, in particular, in connection with a recent result by Walker and Braunstein [22], who proved the realizability of approximate classical broadcasting of statistical distributions with any desired degree of accuracy.

We thank Venketeswara Pai, Sam Braunstein, Gilad Gour, Ady Mann, and two anonymous referees for useful comments. This work was supported in part by a grant from the Israel Science Foundation administered by the Israel Academy of Sciences and Humanities.

*amirk@techunix.technion.ac.il
†itayhe@post.tau.ac.il

[1] C. A. Fuchs and A. Peres, Phys. Rev. A **53**, 2038 (1996).
[2] D. Kaszlikowski *et al.*, Phys. Rev. Lett. **91**, 037901 (2003).
[3] C. H. Bennett *et al.*, Phys. Rev. Lett. **70**, 1895 (1993).
[4] J. Gemmer, A. Otte, and G. Mahler, Phys. Rev. Lett. **86**, 1927 (2001); M. Horodecki *et al.*, Found. Phys. **35**, 2041 (2005).
[5] W. K. Wootters and W. H. Zurek, Nature (London) **299**, 802 (1982); D. Dieks, Phys. Lett. A **92**, 271 (1982).
[6] H. Barnum *et al.*, Phys. Rev. Lett. **76**, 2818 (1996).
[7] G. Lindblad, Lett. Math. Phys. **47**, 189 (1999); H. Barnum *et al.*, Phys. Rev. Lett. **99**, 240501 (2007).
[8] C. H. Bennett, G. Brassard, and N. D. Mermin, Phys. Rev. Lett. **68**, 557 (1992).
[9] V. Buzek and M. Hillery, Phys. Rev. A **54**, 1844 (1996).
[10] N. Gisin and S. Massar, Phys. Rev. Lett. **79**, 2153 (1997).
[11] N. Gisin and B. Huttner, Phys. Lett. A **228**, 13 (1997).
[12] T. Mor, Phys. Rev. Lett. **80**, 3137 (1998).
[13] A. Daffertshofer, A. R. Plastino, and A. Plastino, Phys. Rev. Lett. **88**, 210601 (2002).
[14] H. Umegaki, Ködai Math. Sem. Rep. **14**, 59 (1962).
[15] A. Wehrl, Rev. Mod. Phys. **50**, 221 (1978).
[16] G. Lindblad, Commun. Math. Phys. **40**, 147 (1975).
[17] M. B. Ruskai, J. Math. Phys. (N.Y.) **43**, 4358 (2002); D. Petz, *Quantum Information Theory and Quantum Statistics* (Springer, Berlin, 2008), Chap. 9. This condition is symmetric with respect to A and B.
[18] That mutually commuting states are broadcastable provided the basis in which they are diagonal is known, was also established in the formulation of the original no-cloning and no-broadcasting theorems [5,6].
[19] We thank M. J. Donald for pointing this out to us.
[20] First we note that for every vector $|x\rangle$ and a density matrix σ, $\langle x|\sigma|x\rangle \geq 0$ with equality iff $|x\rangle$ belongs to the kernel of σ. Considering $\langle x|\sigma_{\text{mix}}|x\rangle = \lambda\langle x|\sigma_1|x\rangle + (1-\lambda)\times\langle x|\sigma_2|x\rangle$, we notice that the left-hand side is zero iff both bracketed terms on the right-hand side are zero as well, meaning that $\ker(\sigma_{\text{mix}}) = \ker(\sigma_1)\bigcap\ker(\sigma_2)$ and in particular $\ker(\sigma_{\text{mix}}) \subseteq \ker(\sigma_1)$.
[21] S. Kullback and R. A. Leibler, Ann. Math. Stat. **22**, 79 (1951).
[22] T. A. Walker and S. L. Braunstein, Phys. Rev. Lett. **98**, 080501 (2007).
[23] R. J. Glauber, Phys. Rev. Lett. **10**, 84 (1963); E. C. G. Sudarshan, *ibid.* **10**, 277 (1963).
[24] N. G. van Kampen, *Stochastic Processes in Physics and Chemistry* (North-Holland, Amsterdam, 1992).
[25] Unlike the P-representation of ρ, the definition of the Hamiltonian operator $\hat{H}$ requires an extra factor proportional to $\hbar^{-1}$.
[26] A. S. Holevo, Probl. Inf. Transm. **9**, 110 (1973) [Prob. Peredachi Inf. **9**, 3 (1973)].
[27] B. Schumacher, Phys. Rev. A **51**, 2738 (1995).
[28] M. Horodecki, J. Oppenheim, and A. Winter, Nature (London) **436**, 673 (2005).
[29] A. Kalev, A. Mann, and M. Revzen, Found. Phys. **37**, 125 (2007).