Quantum limit for information transmission

Marcelo Schiffer
Instituto de Física Teórica, Rua Pamplona 145, São Paulo, São Paulo 01405 Brazil
(Received 29 October 1990)

In this paper, we give two independent and rigorous derivations for the quantum bound on the information transmission rate proposed independently by Bekenstein [Phys. Rev. Lett. **46**, 623 (1981)] and Bremermann [*Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability,* edited by L. M. LeCam and J. Neyman (University of California Press, Berkeley, 1967], preceded by a heuristic argument showing why such a bound must hold. In both approaches, information carriers are quanta for some field. The first method resembles the microcanonical approach to statistical mechanics, where the strategy of overestimating the real number of states by relaxing the indistinguishability of quanta was adopted. The second is based entirely upon maximum-entropy methods. Amazingly enough, the results obtained by these physically unrelated premises turn out to be identical, namely, that the single (noiseless) channel capacity is $\dot{I}_{\max}=E/2\pi\hbar$ bits s^{-1}. It is further shown that, in a finite time τ, no information can ever be conveyed unless the energy threshold $2\pi\hbar/\tau$ is reached, allowing the reinterpretation of the time-energy uncertainty in informational-theoretic language.

INTRODUCTION

Although the concept of "information" might be clear for the layman's mind, it only acquired a scientific status in 1948, when Shannon developed the mathematical theory of information.[1] Suppose that the information mediator is one possible state of a system (a symbol, a quantum state, etc.) such that the actual state the system is in is not known *a priori* but only its occurrence probability p_n. The amount of ignorance concerning a given message content was defined by him to be[1-3]

$$H=-k\sum_n p_n \ln p_n \ . \tag{1}$$

The constant k fixes the units of information; for $k=1/\ln 2$ it is measured in bits, etc. In the decodification process, the observer picks up one from all possible states, gaining an amount of information which is equal to Shannon's entropy (1).

Whenever information transmission is concerned, the important question arises, both from the theoretical and practical standpoints, whether or not the laws of nature might constrain the flux of information through a channel. A first limitation is imposed by causality, since no signal can travel faster than light. A second and less obvious one is the rate at which any information may be transferred through a given channel—the channel capacity. According to Shannon and Weaver,[2] this capacity is determined by the noise present in the channel. If this noise is statistically independent of the signal, then the channel capacity for a narrow-band channel of bandwidth $\Delta\omega/2\pi$ reads[2]

$$\dot{I}_{\max}=\frac{\Delta\omega}{2\pi}\ln\left[1+\frac{P}{N}\right] , \tag{2}$$

where P and N stand for the signal output and noise powers, respectively. This equation tells us that in the absence of noise (which can be achieved, say, by freezing a channel subjected to thermal noise), no limitation on the information flow rate exists. Since the formula was deduced on purely classical premises, it is quite legitimate to ask whether similar conclusions translate into the quantum domain or not.

Quantum channel capacity theory originated in the early 1960s. Gordon[4,5] gave two early derivations of the quantum channel capacity for a noiseless channel. Neither gave the correct coefficient in the quantum channel capacity, but both gave the correct dependence $\dot{I}_{\max}\propto(P/\hbar)^{1/2}$. Stern[6] and Marko[7] had a similar measure of success by other approaches. For a noiseless channel, there is a thermodynamic derivation of the quantum capacity due to Lebedev and Levitin,[8] which includes the effects of thermal noise. The simplest situation regarding communication or information transfer is when the steady state is obtained. Physically, the problem is somewhat analogous to equilibrium thermodynamics, and, indeed, thermodynamics has played an important role in the development of steady-state communication theory. Pendry,[9] following this assumption, and dealing specifically with a noiseless broadband channel, obtained the quantum capacity formula

$$\dot{I}_{\max}=(\pi P/3\hbar)^{1/2}\log_2 e \ \text{bits s}^{-1} \ . \tag{3}$$

This result may be criticized from two different aspects. First, it is framed by means of the signal's mean energy. Now, in Bose statistics of one level, the ratio of mean energy E to energy standard deviation ΔE is $N^{-1/2}$, where N is the total number of quanta. Thus, when the system has few quanta, the energy spread is not small compared to the mean energy itself. Hence, for signals of modest information (low-excitation configurations), the mean energy is far from representing the actual energy employed to encode the message. A second aspect is that this formula relies upon the steady-state assumption, an idealization seldom met in practice. The conclusion must be that since (3) rests on the characterization of a signal by its mean energy (canonical theory) and on the steady state (thermodynamical equilibrium), this formula can only be trusted on specific regimes and cannot be regarded as a general limit for the information transmission rate.

An alternative formula on the quantum limitations on the information flow rate was proposed by Bremermann,[10,11] using an obscure argument, whose crux is to equate Shannon's noise with the energy uncertainty $\delta E \geq \hbar/\tau$ required by the time-energy uncertainty relation for a signal of duration τ. His final result is

$$\dot{I}_{\max} = \frac{E}{2\pi\hbar}\log_2(1+4\pi) \text{ bits s}^{-1} . \tag{4}$$

Bremermann's argument has been criticized for relying on the classical Shannon formula to get an ostensibly quantum result, and for the obscurity surrounding the connection of noise power with the time-energy uncertainty relation, itself a principle that invites confusion.

An alternative road to a bound like (4) was proposed by Bekenstein[12] and relies on causality considerations combined with the bound on the entropy H that may be contained by a physical system of definite linear size R and proper energy E:[13]

$$H \leq \frac{2\pi ER}{\hbar c} . \tag{5}$$

This bound was originally inferred from black-hole thermodynamics, but has since been established by detailed numerical experiments[14] and analytical arguments.[15,16] Now, by transporting a system with information inscribed in it, one has a form of communication, albeit not the generic one. According to Shannon's information theory, the peak entropy $H_{\max}$ that *could* be in a system limits the maximum information $I_{\max}$ that can be stored in it. Because the system cannot travel faster than light, it sweeps by a given point in time $\tau \geq R/c$. Thus an appropriate "receiver" can acquire from it information at a rate not exceeding $H_{\max}\log_2 e/\tau$ (as usual, $\log_2 e$ converts to bits). Substituting from Eq. (5), we have

$$\dot{I}_{\max} \leq \frac{2\pi E}{\hbar}\log_2 e \text{ bits s}^{-1} , \tag{6}$$

which is quite similar to (4).

A few years later, Bekenstein[17] reproduced a similar result based on the canonical approach for signals of finite duration and mean energy E. In this approach, the quantity H/E is maximized, and the vacuum is excluded as a legal state for information transmission. At some stage, some numerical calculations had to be performed, and his final result was

$$\dot{I}_{\max} \leq 0.2279 E/\hbar \text{ bits s}^{-1} . \tag{7}$$

In this paper we give a heuristic argument based on very general grounds supporting a linear bound like (5), which is then followed by two independent and rigorous proofs.

THE HEURISTIC DERIVATION

Let us now give a *heuristic* argument for a communication bound which does not rely on the entropy bound Eq. (5). Suppose the information we wish to transmit is inscribed in a bosonic carrying field by populating its energy levels with quanta; each quantum configuration represents a different message. Let τ and E be the signal's duration and energy, respectively, ε the lowest non-zero one-quantum energy level, and $\Delta\varepsilon$ the smallest energy separation between levels beneath E. Evidently the total number of occupied levels is $N \leq E/\Delta\varepsilon$, while the total number of quanta is $M \leq E/\varepsilon$. The total number of configurations is bounded from above by a formula well known from Bose statistics:

$$\Omega = \frac{(N+M-1)!}{M!(N-1)!} . \tag{8}$$

All these configurations are *a priori* equally likely so that the peak entropy of the signal is bounded according to

$$H_{\max} \leq \ln[(N+M-1)!] - \ln M! - \ln[(N-1)!] . \tag{9}$$

Assuming N and M are large, the logarithms may be approximated with Stirling's formula. Substituting the bounds on N and M, equating $H_{\max}$ with the peak information, and converting to bits, we get

$$I_{\max} \leq \frac{E}{\sqrt{\varepsilon\Delta\varepsilon}}\left[\left[\frac{\varepsilon}{\Delta\varepsilon}\right]^{1/2}\log_2\left[1+\frac{\Delta\varepsilon}{\varepsilon}\right] + \left[\frac{\Delta\varepsilon}{\varepsilon}\right]^{1/2}\log_2\left[1+\frac{\varepsilon}{\Delta\varepsilon}\right]\right] \text{ bits} . \tag{10}$$

The function $f(x) = x^{-1/2}\log_2(1+x)$ inside the brackets has an upper bound ≈ 2.32. Now, in order to be able to decode the information, the receiver must be able to distinguish between the various energy levels, which calls for energy measurement with precision $\delta E \leq \Delta\varepsilon$. According to the time-energy uncertainty principle, the finiteness of the measurement interval τ imposes an uncertainty $\delta E \geq h/\tau$. Thus $\Delta\varepsilon \geq h/\tau$ in order that the useful information approaches $I_{\max}$. Furthermore, if R is the spatial extent of the signal, we can use the momentum-position uncertainty relation to set the bound $\varepsilon/c \geq \hbar/R$. In addition, on grounds of causality, the inequality $\tau \geq R/c$ must apply. Therefore, $\sqrt{\varepsilon\Delta\varepsilon} \geq \sqrt{2\pi}\hbar/\tau$. Putting all these inequalities together, it follows from (10) that

$$\hat{I}_{\max} \leq \frac{0.925E}{\hbar} \text{ bits s}^{-1} , \tag{11}$$

which is of the same form as (6) and (7), although the numerical factor is different.

This argument, appealing as it is, suffers from two drawbacks: it is only valid for large N and M, where Stirling's approximation may be trusted, and it makes use of the popular but nonrigorous version of the time-energy uncertainty relation. We now turn to two exact derivatives of the linear bound.

SIGNALS WITH SPECIFIED ENERGY BUDGET

Instead of specifying the signal by its mean energy, a misleading concept for low excitations, one can instead specify the energy "budget" or energy "ceiling" for signaling—the maximum available energy per signal. Shannon's entropy, Eq. (1), reduces in this case to

$$H_{\max}=k\ \ln\Omega(E)\ , \tag{12}$$

where $\Omega(E)$ is the number of states compatible with this energy budget, since all signal states with energies below the maximum are equally likely. The problem reduces to counting the number of signal states as a function of the energy budget. This is a difficult problem in general, as has long been known from its analog in microcanonical statistical mechanics. This counting was carried out numerically for a few examples by Gibbons,[18] and later by Bekenstein himself.[14] Recently progress has been made towards the limited goal of establishing *bounds* on the number of quantum states up to a given ceiling energy for three-dimensional systems.[16] We shall next briefly review this approach.

Let $\Omega(E)$ be the number of states with energy up to and including E that are accessible to a quantum system. Evidently $\Omega(E)$ depends on the one-quantum energy spectrum $\{\omega_a\}$. It is convenient to focus attention on configurations with a fixed number of quanta m. If the one-particle levels are ordered by energy, so that $\omega_{a_i}\le\omega_{a_j}$ when $a_i<a_j$ (degenerate levels are to be ordered arbitrarily), an m-quanta configuration is specified by the set of occupied one-particle levels $\{\omega_{a_i}\}$ (of course, some of them may be repeated, corresponding to multiple occupation of a level). The number $\Omega_m(E)$ of m-quanta states with total energy $\le E$ can be written as

$$\Omega_m(E)\equiv\sum_{a_1\le a_2\le\cdots a_m}\Theta(E-\omega_{a_1}-\omega_{a_2}-\cdots-\omega_{a_m})\ ,\qquad m\ge 1\ , \tag{13}$$

where Θ is Heavyside function. The disposition of the limits on the summation has the effect of avoiding the double counting of states that differ only by the exchange of (identical) quanta. We shall assume a nondegenerate vacuum, so that $\Omega_0(E)=1$ for $E\ge 0$. The number of one-quantum states with energy up to E will play an important role in a discussion later:

$$n(E)\equiv\Omega_1(E)=\sum_{a=0}^{\infty}\Theta(E-\omega_a)\ . \tag{14}$$

It is tacitly assumed that there is no zero mode, i.e., $\omega_a>0$. Thus $n(E)=0$ for $E\le 0$.

The problem of finding the number of accessible states $\Omega(E)$ can evidently be reduced to that of counting all possible m-quanta states:

$$\Omega(E)=\sum_{m=0}^{\infty}\Omega_m(E)\ . \tag{15}$$

An explicit calculation of $\Omega(E)$ by this means is in general hopeless. Nevertheless, for the purpose of setting an upper bound on $\Omega(E)$, we focus attention on the alternative quantity $N_m(E)$, which omits the field quanta indistinguishability, thus overcounting the actual number of m quanta states $\Omega_m(E)$. Technically speaking, this is accomplished by relaxing the energy ordering in Eq. (13). In analogy with Eq. (15), one defines

$$N(E)\equiv\sum_m N_m(E)\ , \tag{16}$$

which obviously satisfies

$$H_{\max}=\ln\Omega(E)<\ln N(E)\ . \tag{17}$$

The advantage of this procedure is that $N(E)$, which is an infinite sum of Θ functions, satisfies a very simple integral equation[16] connecting the one-particle problem to the field theory:

$$N(E)=\Theta(E)+\int_0^E N(E-E')\left[\frac{dn}{dE'}\right]dE'\ , \tag{18}$$

as may be easily checked by iterating (18), having in mind that dn/dE is a sum of δ functions [see Eq. (14)].

This equation can be converted into an algebraic equation by taking its Laplace transform:

$$\tilde{N}(s)=\frac{1}{s[1-s\tilde{n}(s)]}\ , \tag{19}$$

where $\tilde{f}(s)$ stands for the Laplace transform of a function $f(E)$.

Now, a signal of finite duration as seen from a fixed point may be represented by some function $F(t)$ that has compact support in time, i.e., it is nonvanishing only in the interval $[0,\tau]$. In fact, it is mathematically convenient to regard F as periodic with period τ. This "periodic boundary condition," well known from quantum physics, captures the essence of the finiteness of the duration, while keeping the mathematics simple. Resolve $F(t)$ into its Fourier components involving the angular frequencies $2\pi j/\tau$ for all positive integers j (negative integers are superfluous—recall that under second quantization of a Bose field negative frequencies just duplicate the modes). The $j=0$ (dc) mode is to be ignored as relating to a condensate of the field. So the spectrum is $\omega_j=2\pi\hbar j/\tau$, with $j=1,2,\ldots,$ and with no degeneracies. Thus, the Laplace transform of the one-quantum particle number function is

$$\tilde{n}(s)=\int_0^{\infty}dE\ e^{-Es}\sum_{j=1}^{\infty}\Theta(E-j\varepsilon)\ , \tag{20}$$

where $\varepsilon=2\pi\hbar/\tau$. This equation is equivalent to

$$\tilde{n}(s)=s^{-1}\sum_{j=1}^{\infty}e^{-j\varepsilon s}\ . \tag{21}$$

Performing the sum in (21), substituting in (19), and inverting the Laplace transform $\tilde{N}(s)$, we have

$$N(E)=\frac{1}{2\pi i}\int_{\gamma-i\infty}^{\gamma+i\infty}\frac{e^{s\varepsilon}-1}{s(e^{s\varepsilon}-1)}e^{Es}ds\ . \tag{22}$$

This integral is performed in the Appendix, and the result is

$$N(E)=2^{[[E/\varepsilon]]}\ , \tag{23}$$

and $[[x]]$ stands for the integral part of x. Combining (23) and (17) gives

$$H_{\max}(E)\begin{cases}=0 & \text{if } E\tau/2\pi\hbar<1\\ \le E\tau\ln 2/2\pi\hbar & \text{otherwise}\ .\end{cases} \tag{24}$$

This result teaches us that for a pulse of duration τ, no information can ever be conveyed unless the energy threshold $E>\tau^{-1}2\pi\hbar$ is reached. Furthermore, it leads to the channel capacity

$$\dot{I}_{\max}\le\frac{E}{2\pi\hbar}\ \text{bits s}^{-1}\ . \tag{25}$$

We shall now proceed to an alternative derivation of Eq. (25).

THE INFORMATIONAL-THEORETIC APPROACH

In this section we deal with an approach which is very akin to what information-theory practioners used to consider. As before, the information carriers we consider are quanta distributed among the various energy levels ω_j. If N_j is the number of quanta in a given mode j, and N is the total number of quanta consumed to encode a message, then the *total energy* spent in the message is

$$E=\sum_{j=1}^{n}N_j\omega_j=N\sum_{j=1}^{n}p_j\omega_j\ , \tag{26}$$

where $p_j=N_j/N$ stands for the probability of finding a quanta in the jth energy level. The above sum is performed for all $j\le n$ levels such that $\omega_n\le E$. Obviously,

$$\sum_{j=1}^{n}p_j=1\ . \tag{27}$$

The information conveyed per quanta is

$$h=-\sum_{j=1}^{n}p_j\ln p_j\ . \tag{28}$$

Since we wish to obtain the optimal transmission rate, the various probabilities p_i must be obtained by maximizing the total information conveyed by all N quanta, $H=Nh$, while enforcing the conditions (26) and (27). This is accomplished by varying the quantity

$$H'=-N\sum_{j=1}^{n}p_j\ln p_j-\alpha\sum_{j=1}^{n}p_j-\beta N\sum_{j=1}^{n}p_j\omega_j\ , \tag{29}$$

where α and β are Lagrange multipliers to be determined later. Variation of this equation yields

$$-\delta N\left[\sum_{j=1}^{n}p_j\ln p_j+\beta\sum_{j=1}^{n}p_j\omega_j\right]-\sum_{j=1}^{n}[\alpha+\beta N\omega_j+N(1+\ln p_j)]\delta p_j=0\ . \tag{30}$$

Here δN and δp_j are to be regarded as arbitrary and independent variations. Thus

$$\sum_{j=1}^{n}p_j(\ln p_j+\beta\omega_j)=0 \tag{31}$$

and

$$\ln p_j=-\frac{\alpha}{N}-\beta\omega_j-1\ . \tag{32}$$

Having in mind the normalization condition (27), we solve (32) and (31) for α and $\ln p_j$:

$$\alpha=-N \tag{33}$$

and

$$p_j=e^{-\beta\omega_j}\ . \tag{34}$$

The insertion of (34) into (28), with the aid of (26) and (27), yields

$$H_{\max}=\beta E\ . \tag{35}$$

Notwithstanding the kinship between this formalism and canonical theory in statistical mechanics, these formalisms are conceptually very different. Here E stands for the actual energy spent upon encoding the message, rather than being the mean energy. Furthermore, the "inverse temperature" β is not a free parameter fixed by the heat reservoir, but is determined by the normalization condition (27)

$$\sum_{j=1}^{n}g_je^{-\beta\omega_j}=1\ , \tag{36}$$

where g_j stands for the jth-level degeneracy. In order to solve this equation, first notice that since

$$\sum_{j=1}^{n}g_je^{-\beta\omega_j}<\sum_{j=1}^{\infty}g_je^{-\beta\omega_j}\ ,$$

if we define β^* through

$$\sum_{j=1}^{n}g_je^{-\beta\omega_j}=\sum_{j=1}^{\infty}g_je^{-\beta^*\omega_j}\ , \tag{37}$$

then

$$\beta<\beta^*\ , \tag{38}$$

and β^* satisfies the equation

$$\sum_{j=1}^{\infty}g_je^{-\beta^*\omega_j}=1\ . \tag{39}$$

Therefore, putting (38) and (35) together,

$$H_{\max}\le\beta^*E\ . \tag{40}$$

Owing to the fact the pulse duration τ is the unique dimensional parameter involved in the problem, on dimensional grounds Eq. (39) should yield $\beta^*\propto\tau^{-1}$. Thus, regardless of the nature of the spectrum $\{\omega_j\}$, a bound like (5) must hold [see Eq. (40)], although the numerical factor could only be determined after the signal's spectrum is specified. Adopting the periodic boundary condition already discussed in the preceding section (no degeneracies means $g_j=1$), Eq. (39) for β^* reads

$$\sum_{j=1}^{\infty} \exp\left[-\frac{2\pi\hbar\beta^*}{\tau}\right]=1 . \quad (41)$$

Summing up the above series yields the equation

$$\frac{1}{1-\exp(-2\pi\hbar\beta^*/\tau)}=2 ,$$

whose solution is

$$\beta^*=\frac{\tau \ln 2}{2\pi\hbar} . \quad (42)$$

Putting (40) and (42) together and converting information to units of bits leads to the noiseless (single) channel capacity

$$\dot{I}_{\max}=\frac{E}{2\pi\hbar} \text{ bits s}^{-1} . \quad (43)$$

A remark is in order. The sum in Eq. (36) runs over $j=1,\ldots,n$, where n is the number of levels beneath E, the energy stored in the message. Thus, if $E<\omega_1$, $\dot{I}_{\max}=0$. For the period boundary condition prescription, this means that, unless the energy threshold $E>2\pi\hbar/\tau$ is crossed, no information can be conveyed, in complete agreement with our earlier result.

It is needless to stress the technical importance of enhancing the information transmission capacity. But how could the above limit be overcome? The obvious way to do it is to endow the communication system with many channels. Fortunately, the above formalism can be easily adapted to tackle many channels. If N is the number of identical channels present in the configuration, then we should have N times the replica of the single-channel spectrum. Thus the multiple-channel configuration can be regarded as a single-channel problem, provided we replace the primitive spectrum by an N-fold-degenerate spectrum, i.e., $g_j \rightarrow g_j'=Ng_j$. Therefore, in this situation, although Eq. (40) follows unchanged, (39) must be replaced by

$$N\sum_{j=1}^{\infty} g_j e^{-\beta^*\omega_j}=1 . \quad (44)$$

The counterpart of Eq. (41) for N-channel configuration is

$$N\sum_{j=1}^{\infty} \exp\left[-\frac{2\pi\hbar\beta^* j}{\tau}\right]=1 . \quad (45)$$

Summing up the above series yields for β^*,

$$\beta^*=\frac{\tau}{2\pi\hbar}\ln(N+1) . \quad (46)$$

Inserting this solution into the capacity formula (40) and converting to units of bits yield the final result for multichannel communication systems:

$$\dot{I}_{\max}(N)=\frac{E}{2\pi\hbar}\log_2(N+1) \text{ bits s}^{-1} . \quad (47)$$

It is worthwhile to remark that the enhancement of the information transmission rate is payed at the expense of reducing the single-channel capacity. Indeed, if we define the single-channel efficiency in a multiple-channel transmission system as $r=\dot{I}_{\max}(N)/N\dot{I}_{\max}(1)$, then $r=\log_2 N/N$ drops very fast with the increasing number of channels N (for instance, for a ten-channel communication system, the single-channel efficiency drops to $\approx 33\%$).

REMARKS AND CONCLUSIONS

Focusing attention on the issue of the information carriers (photons or phonons in a crystal, sound waves in a pipe, etc.) allowed us to insert information theory within the realm of physical science and to draw important conclusions regarding the process of information transmission. Here we considered the information carriers to be field quanta, while a message consisted of a particular quantum-field configuration. We dealt with the problem of information flow from two entirely different physical standpoints.

Three ingredients went into both proofs: the periodic condition approximation, the assumption that the zero-frequency mode cannot be used in signaling, and the characterization of signals by occupation number. Let us discuss them in turn.

By viewing the signal as periodic, one obtains a simple form for the frequency spectrum. This sort of approach is quite common in physics. Arguably, it would have been more realistic to look at signals that turn on and off abruptly. In that case, there are no sharp one-quantum energies; rather, all levels are broadened. One way to proceed then is to use Gabor frequency-time cells[19] to partition the phase space occupied by the signal. To each such cell is assigned a Gaussian modulated sinusoidal wave which takes over the role of the pure sinusoidals in the Fourier representation of the periodic signal and embodies the idea that the energy levels must be broadened in inverse proportion to the duration τ. If all cells are chosen to extend a time τ, it is natural to choose the central frequencies of the Gaussian wave packets to correspond to the energies $\omega_j=2\pi j\hbar/\tau$, precisely the frequencies figuring in the periodic-boundary-condition approximation. The energy spread of a wave packet is then $\sim 2\pi\hbar/\tau$. With this choice it is easy to grasp the effect of the periodic-boundary-condition approximation.

For a given energy E, a many-quanta state with $\sum_j \omega_j > E$ was excluded in the periodic-boundary-condition approximation. However, if the energy sum exceeds E only by a quantity of order $2\pi\hbar/\tau$, the state is allowed in the present description because it is possible for the true energies of several of the quanta to be on the low side of the central energies of their Gaussian packets. Of course, the larger the excess of $\sum_j \omega_j$ over E, the less probable the state, for if the state is a one-quantum state, the quantum's energy must lie on the outskirts of the Gaussian packet to keep below E. This situation has low probability. If we deal with a several-quanta state, the individual energies can lie closer to the central energies, but there must be a trend toward the lower-energy side. Thus, although the individual quanta are not at very improbable energies, the product of several probabilities smaller than 1 will cause the overall configuration to be unlikely. Thus, in the exact treatment, extra states become available, but these states have low probability.

We must also note that some states which were permitted in the periodic-boundary-condition approximation become, in the exact treatment, low probability states. The states in question are those with $\sum_j \omega_j$ within a quantity $\lesssim 2\pi\hbar/\tau$ on the low side of E. This is because with non-negligible probability some of the quanta involved can lie on the high side of their Gaussian peaks, and cause the true total energy to exceed E. This effect partly neutralizes the gain of states discussed above. The conclusion must be that the periodic-boundary-condition approximation is likely to only somewhat underestimate the number of states. We thus venture to conclude that (25) is likely to be only a little below the true linear bound in the exact treatment.

In our derivations we excluded the mode with $j=0$. If it were included in our derivation it would have led to an infinity of states for any energy. This is because we can form arbitrarily many states by having a varying number of quanta with zero energy, since all these have no energetic cost. To understand why the zero-frequency mode must be excluded, one must distinguish between the situation where the signal is periodic, and the one where it is sharply limited in time. In the first case, the periodic boundary condition is exact; the zero-frequency mode in question sets the dc level of the signal. This dc level cannot serve to send information. It is permanent, and does not turn on when the signal is sent, so that the signal's information is not coded in it. At best, the dc level conveys some information about the channel, but is not specific to the signal. The zero-frequency mode thus has no role in signaling. When the signal is sharply bounded temporally, the spreading of frequencies precludes the existence of a mode with exactly zero energy. Even as the center of a Gaussian wave packet, $\omega_j=0$ is very far from reality, because it would have as many negative frequencies as positive. Hence, in the periodic-boundary-condition approximation, we should not include the $j=0$ mode.

In our derivation, the signal states were classified by occupation number. This means that, strictly speaking, our bound (25) is valid only for a communication system with a transmitter that prepares occupation number states, a channel in which the occupation number operator is a constant of the motion, i.e., propagation in vacuum, and a receiver which measures occupation number, e.g., a photomultiplier for an optical channel. As yet no study has been made of the influence of other choices of states (coherent, squeezed, etc.) on communication via signals with a definite energy budget. However, if the situation for signals with definite *mean* energy is any guide,[17] the capacity for these alternative signal states can only be smaller than for occupation number state signals. We thus conjecture that bound (25) is true for any type of signaling states. Progress in this direction will be reported elsewhere.

The remarkable agreement between the quantum capacity we obtained starting from two completely unrelated physical premises [see Eq. (25) and (43)], namely (i) the overestimation of the number of possible states for signaling by dismissing the indistinguishability of quanta in the microcanonical formalism, and (ii) the maximum-entropy approach, seems to point out the universal character of our result. Last, but not least, the fact that no information can ever be conveyed unless the signal duration τ and proper energy E satisfy the relation $E\tau \geq h$ allows the reinterpretation of the time-energy uncertainty relation in informational-theoretic language.

ACKNOWLEDGMENTS

I am thankful for Ben Burion University's hospitality and to Conselho Nacional de Desenvolvimento Científico e Tecnológico for partial financial support. Last, but not least, I wish to thank J. D. Bekenstein for introducing me to this fascinating subject and for many enlightening discussions.

APPENDIX

We shall evaluate the integral (21):

$$N(E)=\frac{1}{2\pi i}\int_{\gamma-i\infty}^{\gamma+i\infty}\frac{e^{\varepsilon s}-1}{s(e^{\varepsilon s}-\alpha)}e^{sE}ds \ . \tag{A1}$$

For computational reasons, it is convenient to define new variables $\sigma \equiv \varepsilon s - \ln\alpha$ and $x \equiv E/\varepsilon$. The above expression then reads

$$N(E)=\alpha^x I_\alpha(x) \ , \tag{A2}$$

where

$$I_\alpha(x)=\frac{1}{2\pi i}\int_{\gamma-i\infty}^{\gamma+i\infty}\frac{\alpha e^{\sigma}-1}{\alpha(\sigma+\ln\alpha)(e^{\sigma}-1)}e^{\sigma x}d\sigma \ . \tag{A3}$$

This integral is performed by pushing the contour leftwards to minus infinity while indenting it so as not to overrun all the infinite poles located at the imaginary axis, $\sigma=2\pi i n$, as shown in Fig. 1. By Cauchy's theorem, the integral is

$$I(x)=\frac{\alpha-1}{\alpha}\sum_{n=-\infty}^{\infty}\frac{e^{2\pi x n i}}{2\pi n i+\ln\alpha} \ , \tag{A4}$$

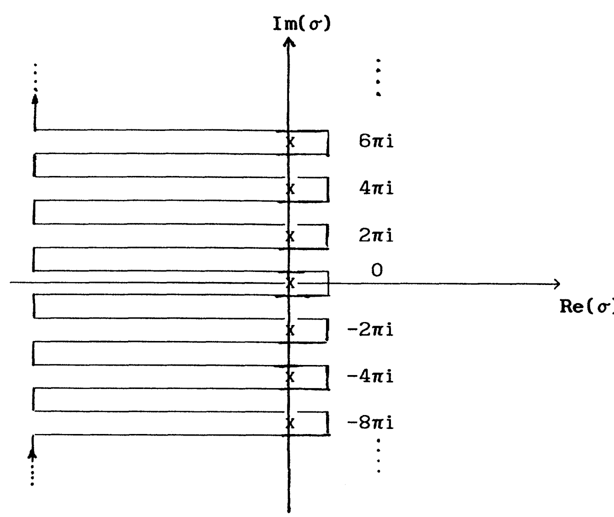

FIG. 1. The countour integral.

and can be expressed in the form

$$I(x)=\frac{\alpha-1}{\alpha}\left[\frac{1}{\ln\alpha}+\frac{\ln\alpha}{2\pi^2}\sum_{n=1}^{\infty}\frac{\cos(2\pi x n)}{n^2+\beta^2}+\frac{1}{\pi}\sum_{n=1}^{\infty}\frac{n\sin(2\pi n x)}{n^2+\beta^2}\right], \tag{A5}$$

where $\beta\equiv\ln\alpha/2\pi$. Recalling the identities[20]

$$\sum_{n=1}^{\infty}\frac{\cos(ny)}{n^2+\beta^2}=\frac{\pi}{2\beta}\frac{\cosh\beta(\pi-y)}{\sinh\beta\pi}-\frac{1}{2\beta^2}$$

and

$$\sum_{n=1}^{\infty}\frac{n\sin(ny)}{n^2+\beta^2}=\frac{\pi}{2}\frac{\sinh\beta(\pi-y)}{\sinh\beta\pi},\quad 0\le y\le 2\pi\ , \tag{A6}$$

after some tedious algebra, we obtain for $I(x)$,

$$I_\alpha(x)=\alpha^{-[x]}\ , \tag{A7}$$

where $[x]$ stands for "x *modulo integer*," i.e., the noninteger part of x.

[1]C. E. Shannon, 1948, Bell Syst. Tech. J. **27**, 379 (1948).
[2]C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication* (University of Illinois Press, Urbana, 1949).
[3]L. Brillouin, *Science and Information Theory* (Academic, London, 1965).
[4]J. P. Gordon, Proc. IRE **50**, 1898 (1962).
[5]J. P. Gordon, in *Advances in Quantum Electronics,* edited by J. R. Singer (Columbia University Press, New York, 1961).
[6]T. E. Stern, IEEE Trans. Inf. Theory **IT-6**, 435 (1960).
[7]H. Marko, Kybernetik **2**, 274 (1965).
[8]D. S. Lebedev and L. B. Levetin, Dokl. Akad. Nauk SSSR **149**, 1299 (1963) [Sov. Phys.—Dokl. **8**, 377 (1963)].
[9]J. B. Pendry, J. Phys. A **16**, 2161 (1983).
[10]H. J. Bremermann, in *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability,* edited by L. M. LeCam and J. Neyman (University of California Press, Berkeley, 1967).
[11]H. J. Bremermann, Int. J. Theor. Phys. **21**, 203 (1982).
[12]J. D. Bekenstein, Phys. Rev. Lett. **46**, 623 (1981).
[13]J. D. Bekenstein, Phys. Rev. D **23**, 287 (1981).
[14]J. D. Bekenstein, Phys. Rev. D **30**, 1669 (1984).
[15]I. Kahn and A. Qadir, Lett. Nuovo Cimento **41**, 483 (1984).
[16]M. Schiffer and J. D. Bekenstein Phys. Rev. D **39**, 1669 (1989).
[17]J. D. Bekenstein, Phys. Rev. A **37**, 3434 (1988).
[18]G. W. Gibbons (unpublished).
[19]D. Gabor, Philos. Mag. **41**, 7 (1950); **41**, 1161 (1950).
[20]I. S. Gradshteyn and I. M. Ryzhik, *Table of Series and Integrals, 1.445* (Academic, New York, 1965).