Communication and energy

Jacob D. Bekenstein*
Canadian Institute for Theoretical Astrophysics, University of Toronto, Toronto, Ontario, Canada M5S 1A1
and Center for Theoretical Physics, University of Texas, Austin, Texas 78712
(Received 20 January 1987; revised manuscript received 21 August 1987)

The quantum limitations on the maximum communication rate (capacity) possible through a single noiseless channel with signals of finite duration are investigated. They may be summarized in the characteristic information function defined here. In the absence of dispersion this description is a Lorentz-invariant one, and applies also in the presence of exterior gravitational fields. For long duration the characteristic information function corresponds to the standard quantum-channel-capacity formula of Gordon, of Lebedev and Levitin, and of Pendry. For finite signal duration it proves useful to distinguish between heralded and self-heralding signals, according to whether their arrival is anticipated or not. The two types have different characteristic functions. These are calculated here for occupation-number signal states in channels where the carrier quantum field may be represented by independent modes. The Gordon-Lebedev-Levitin-Pendry formula provides an upper bound to the exact results for both types of signals. The linear bound on communication rate of Bremermann and the present author bounds the capacity only for self-heralding signals. However, it is the better estimator of capacity for self-heralding signals with modest information content.

I. INTRODUCTION

Information in transit is always encoded in physical entities, such as an acoustic field, or the electromagnetic field. This makes the subject of communication—the transfer of information—an issue of physics just as much as one of mathematics or technology. The study of limitations on communication, standard communication theory,[1] has long emphasized the mathematical aspects. However, over the years many workers have discussed physical limitations on communication. A general review is given by Landauer;[2] some early highlights are given in Refs. 3–9. The review by Yamamoto and Haus[10] describes the implementation in quantum optics of several once purely theoretical communications paradigms.

Not only is the stress on physics useful for a deeper understanding of communication, but it is essential when one seeks to employ information and communication concepts to scrutinize *physics* itself. The value of this last endeavor has been underlined by the use of cellular automata to model physical dynamics,[11] by information theoretic schemes that illuminate the meaning of quantum mechanics,[12] and by views of physical law as a computation process.[13] Nowhere is the need for an informational point of view in physics more visible than in the loophole-proof statement of one of the principles of relativity theory: *information* cannot be propagated at a speed exceeding that of light.

It is plausible that progress in the endeavor to see "physics as information"[14] will result from a clearer understanding of the issue of the energy cost of information. The physical nature of all information carriers suggests that each quantity of information being transmitted should be accompanied by some energy. The present wisdom is that dissipation of this energy is not mandatory.[2] But there remains the question of whether the energy in transit can be reduced arbitrarily, or whether there exists a minimum *energy cost per bit* in communication via physical channels. This issue is bound up with that of the maximum possible communication rate a channel is capable of. In the past, various studies of these questions have offered sometimes contradictory answers.[5–9,15–17] Issues have been whether the maximum communication rate is set by signal power or by signal energy, and whether the energy cost per bit can vanish.

The present paper attempts to elucidate these questions for a noiseless channel when the signals are of finite duration. The influence of the finite duration on the communication rate and the energy cost per bit are delineated with the help of elementary ideas from information theory and quantum mechanics. The contradictions mentioned above are resolved by distinguishing between self-heralding versus heralded signals, a distinction widely ignored. In common with Refs. 6 and 9, and in contrast with the model-dependent approach common in communication theory, the present attempt focuses on a quantum field description of the *carriers* of information, and de-emphasizes questions of signal coding, detection, etc.

It is well to first delineate the background and issues. In standard communication theory information is quantified by Shannon's formula,[1,3]

$$I = -\sum_a p_a \log_2 p_a \ . \tag{1}$$

With the choice of base 2 for the logarithm, I is expressed in bits. The p_a stand for the *a priori* probabilities of the various states of a system being employed to store or carry information. The formula applies to the system before it is scrutinized for its actual state. It gives the information that can be gathered when that state is actually established by measurement or reception.

Communication theory investigates the maximum rate at which information may be transferred with negligible errors (channel capacity). Its classic result in Shannon's celebrated channel-capacity formula,

$$\dot{I}_{\max}=(\Delta\omega/2\pi)\log_2(1+P/N)\ , \tag{2}$$

which gives the maximum error-free information transfer rate (bits s^{-1}) achievable through a single channel of bandwidth $\Delta\omega$ (rad s^{-1}) operating at a signal-to-noise ratio P/N (P is the signal power and N is the external noise power). In the theory signals are represented by frequency-limited continuous functions of time, i.e., the theory is wholly classical. Formula (2) successfully describes myriad systems (telephone, fiber-optics links, space telemetry, etc.).

Shannon's formula predicts that the communication rate can be increased arbitrarily by suppressing the noise power, i.e., by cooling the channel for purely thermal noise. The Shannon energy cost per bit $P/\dot{I}_{\max}$ can be written as

$$\epsilon_{\min}=N[\exp(2\pi\dot{I}_{\max}\ln 2/\Delta\omega)-1](\dot{I}_{\max})^{-1}\ . \tag{3}$$

For a given communication rate, $\epsilon_{\min}$ can be reduced arbitrarily by making the noise lower. For thermal noise and low $\dot{I}_{\max}$, $\epsilon_{\min}\approx kT\ln 2$ [see Eq. (5) below], where k is Boltzmann's constant and T is the absolute temperature of the channel. This is a well-known result for the energy associated with a bit in various contexts.[2,3,18] It too vanishes together with the noise. The divergence of $\dot{I}_{\max}$ is removed by quantum fluctuations, as demonstrated by the quantum analog of Shannon's formula for a narrow-band noiseless channel.[6,10,19]

The restriction to narrow band is primarily motivated by technological considerations. Hence we shall focus on the broadband noiseless channel. The corresponding capacity formula,

$$\dot{I}_{\max}=(\pi P/3\hbar)^{1/2}\log_2 e\ , \tag{4}$$

has been developed by Gordon,[5] by Lebedev and Levitin,[6] and Pendry[9] (GLLP) (a related result was obtained by Marko[20]). It, likewise, does not depend on external noise to give finite $\dot{I}_{\max}$. However, steady-state communication as governed by Eqs. (2) or (4) is but an idealization. Particularly, in trying to understand physics from the information point of view, communication in bursts, i.e., via signals of finite duration, must be given attention. The subject has been studied sparsely. A highlight is Bremermann's[7] early suggestion that for a signal of finite duration the communication rate is subject to the bound

$$\dot{I}<\alpha E/\hbar\ , \tag{5}$$

where E denotes the energy carried by the signal and α is a constant for which Bremermann[15] proposed the value $\log_2(1+4\pi)$. Support for a bound of the form (5) has been inferred[16] from the universal bound on the specific entropy of finite systems.[21-24] Note that the results (4) and (5) seem to be contradictory. Formula (4) restricts $\dot{I}_{\max}$ in terms of signal power, while bound (5) does so in terms of signal energy. A hint to the resolution of the paradox is provided by Landauer and Woo's remark[8] that $\dot{I}$ is maximized for a burst of energy with the overall information of order one bit. This suggests that bound (5) might be achievable for just such low-information signals while overestimating $\dot{I}_{\max}$ otherwise. With an important proviso, this turns out to be the case (see below), but the road to this insight has been paved with much confusion.[16,17,25]

The task of reaching general conclusions about communication is made arduous by the difficulty in delineating a generic communication system. We take it for granted that the electromagnetic field can convey information. But transporting a box of preprints is also communication (assuming that they contain information), albeit in a form that would be hard to explore by standard channel-capacity theory. The method of Ref. 16 is well tailored for constraining the communication rate in this case, but would not be appropriate for electromagnetic channels. Paradigms of the communication process have been considered[2] which are close to the transport method, but involve little energy accompanying the information. Which of all these is the most useful paradigm?

It can be argued that the transport and related paradigms are unlikely to disclose fundamental limitations on communication because the information they consider is attached to only a small fraction of the degrees of freedom available to the system [an ink mark, or a SQUID (superconducting quantum-interference device) are made up of millions of atoms]. To really penetrate the issue one should consider paradigms that involve as little incidental "machinery" as possible apart from that strictly required (by physics) to hold the information. Reasoning of this sort makes it compelling to focus on quantum fields, presumably the most elementary representations of matter and its excitations, as the carriers of information. This view, akin to that of Lebedev and Levitin[6] and of Pendry,[9] and contrasting the classical or purely wave-mechanical paradigms of communication that dominate the literature on physical aspects of communication (see review in Ref. 2), is the view we adopt here.

In Sec. II we review the GLLP formula[5,6,9] noting that its form is suggested by purely semiclassical considerations involving Shannon's capacity formula (2). Then we sketch Pendry's[9] thermodynamic derivation of it which brings to the fore the issues of dispersion in the channel and quantum statistics, issues which recur in our study.

In Sec. III we introduce the characteristic information function (CIF) for a channel; it relates the maximum information a signal of finite duration may bear to its energy. We point out that this description is Lorentz invariant and unaffected by gravitational fields through which the signal propagates, so long as its own self-gravitation is negligible. Finally, we describe the general procedure for computing the CIF when signal energy is interpreted as mean energy. The distinction between heralded and self-heralding signals is crucial in this respect, and is related to the issue of detectability of the signal vacuum state. Some general properties of the CIF are mentioned.

The form of the CIF depends on the kind of quantum states used in signaling. In Sec. IV we compute the CIF for occupation-number states for both heralded and self-heralding signals, and for boson or fermion quanta under the assumption that the signal field can be described by independent modes. We show that for signals for which the product of mean energy and duration is very large in units of $\hbar$, the CIF corresponds to the GLLP formula (the answer for steady-state signaling). Departures from this ideal are investigated in the framework of two different methods: periodic boundary condition, or wave-packet description of signals. The results are very similar except for signals bearing only fractions of a bit of information. The GLLP formula is found to provide an upper bound on the communication rate even for finite-duration signals. We also show that the linear bound (5) is rigorously obeyed for self-heralding signals; it gives a good rule-of-thumb estimate of the information bearing capabilities of signals in the range around a few bits. Section V summarizes our results and maps out several avenues for further inquiry closely related to the issues discussed here.

II. STEADY-STATE COMMUNICATION

A. Semiclassical justification of the GLLP formula

Formula (2) for a noisy channel, and its quantum counterpart for a noiseless one (4) are special cases of Lebedev and Levitin's capacity formulas.[6] What is seldom realized is that the *form* of the GLLP formula is directly implied by Shannon's classical capacity formula. Assume the noise is thermal. Then, in the classical regime, the noise is white (no frequency dependence) and is described by Nyquist's formula,[3]

$$N = kT(\Delta\omega/2\pi)\ . \tag{6}$$

The formula is accurate for $kT \gg \hbar\omega_0$, where ω_0 is a typical frequency of the channel. Evidently $\omega_0 > \Delta\omega/2$ (the equality is reached for a bandwidth extending from zero up to some cutoff if we take ω_0 as half the cutoff frequency). The classical regime is obtained for $\hbar\omega_0 \lesssim kT$. Putting the two inequalities in Eq. (6) we get

$$\Delta\omega \lesssim (4\pi N/\hbar)^{1/2}\ . \tag{7}$$

This inequality is merely a guarantee that the classical regime applies with fixed parameters T, $\Delta\omega$, and ω_0, not a physical restriction on N. If the inferred $\Delta\omega$ is not necessarily small, the calculation may nevertheless be justified provided the signal power P is frequency independent also. Then Shannon's capacity formula is valid for a wide band. Substituting Eq. (7) into Eq. (2) we get

$$\dot I_{\max} \lesssim (P/\pi\hbar)^{1/2} f(P/N)\ , \tag{8}$$

where $f(x) \equiv x^{-1/2}\log_2(1+x)$. Now $f(x)$ has a maximum of ≈ 1.16 at $x \approx 3.92$; therefore, we find

$$\dot I_{\max} \lesssim 0.65(P/\hbar)^{1/2}\ , \tag{9}$$

which is of the same form as (4) with a coefficient half as large.

Of course, inequality (9) only restricts the peak-communication rate in the classical regime. It does not preclude a higher rate well into the quantum regime. But it would be incorrect to try to improve on our "derivation" by replacing Eq. (6) by the quantum version of Nyquist's formula, as is sometimes done,[5] since the description of signal and noise in Shannon's theory is classical. It is thus better to rely on a purely quantum derivation.

B. Pendry's thermodynamic derivation

Two early derivations of the GLLP formula (4) were given by Gordon,[5] one based on the time-energy uncertainty relation, and one on the classical Shannon capacity formula. These are merely suggestive, and neither gives the precise coefficient. An early thermodynamic approach is due to Lebedev and Levitin.[6] It not only treats both narrow-band and broadband channels, but includes the effects of thermal noise. Particular limits of their general formula are (2) and (4). Lebedev and Levitin propose their results for electromagnetic channels, but in fact they are good for a range of bosonic channels. A more recent thermodynamic derivation of the GLLP formula for a broadband noiseless channel is due to Pendry.[9] We focus on it since it illustrates two issues we shall treat: the difference between boson and fermion channels, and the insensitivity of the communication limit to dispersion.

Pendry's description of signals is a quantum one: a particular quantum state of the excitation in the channel stands for a particular signal. For concreteness one may think of each set of occupation numbers for the various propagating modes in the channel as a distinct signal. Pendry assumes a channel uniform in the direction of propagation, which allows him to label modes by momentum p. He allows dispersion so that a quantum of momentum p has some general energy $\varepsilon(p)$. Then the propagation velocity of the quanta is the group velocity $c_*(\varepsilon) = d\varepsilon(p)/dp$.

The basic assumption is that $\dot I_{\max}$ can be identified (apart from units) with the unidirectional entropy current that the channel carries when in a thermal state. This hails back to the idea that in a thermal state the entropy in each mode is maximized. Of course, in this state there is no net flow of entropy along the channel, but if we look only at modes propagating, say, to the right, then they do carry an entropy current, and it is assumed that it is the maximal entropy current.

Now the entropy of a particular boson mode in thermal equilibrium at temperature T is[26]

$$s(p) = \frac{\varepsilon(p)/kT}{e^{\varepsilon(p)/kT}-1} - \ln(1-e^{-\varepsilon(p)/kT})\ . \tag{10}$$

The entropy current in one direction is thus

$$\dot S = \int_0^\infty s(p)c_*(\varepsilon)dp/2\pi\hbar\ , \tag{11}$$

where the factor $dp/2\pi\hbar$ is the number of modes per unit length in the interval dp which go by in one direction. This factor, when multiplied by the group velocity, gives the unidirectional current of modes. After an integration by parts on the second term, we can cast the last result into the form

$$\dot{S}=2\int_0^\infty \frac{\varepsilon(p)/kT}{e^{\varepsilon(p)/kT}-1}\frac{d\varepsilon(p)}{dp}\frac{dp}{2\pi\hbar} \ . \tag{12}$$

Now Pendry notes that the first factor in the integrand is actually the mean energy per mode divided by kT, so that the integral represents the unidirectional power in the channel,

$$\dot{S}=2P/kT \ . \tag{13}$$

Pendry then evaluates the integral in Eq. (12) by canceling the two differentials dp and assuming the energy spectrum is single valued and extends from 0 to ∞. Then the form of the dispersion relation $\varepsilon(p)$ does not enter, and Pendry obtains

$$P=\pi(kT)^2/12\hbar \ . \tag{14}$$

The last and crucial step is to eliminate kT between the expressions for $\dot{S}$ and P. Since entropy is here measured in natural units, while information is measured in bits, one identifies $\dot{S}\log_2 e$ with $\dot{I}_{\max}$. The final result is the GLLP formula (4). The calculation may easily be repeated for Fermi statistics with the result that the GLLP communication rate is reduced by a factor $\sqrt{2}$. Instead of Eq. (3) of Shannon's theory, we have here

$$\epsilon_{\min}=3\hbar\pi^{-1}(\ln 2)^2\dot{I}_{\max} \tag{15}$$

for bosons, and a factor $\sqrt{2}$ larger for a fermion channel. Lebedev and Levitin[6] generalize (15) to channels with thermal noise.

III. FINITE DURATION SIGNALS

A. The characteristic information function (CIF)

The thermodynamic derivations make it clear that the GLLP formula is, in the first instance, applicable only to steady-state communication, which implies both that the signal is of infinite duration, and that its statistical properties are stationary. Of course, such a situation is an idealization. Particularly when exploring communication theory as an avenue to a deeper understanding of physics, we should also consider signals of finite duration; of course these can also appear in more mundane contexts. Is the GLLP formula exactly valid for finite duration signals? Consider the analogous issue in thermodynamics. Purely thermodynamic results like Eqs. (10)–(14) are strictly valid in a perfectly stationary situation. Temporal variations bring about a departure from thermodynamic equilibrium and a modification of the formulas. Thus we suspect that the GLLP formula needs to be generalized for signals of finite duration. The new formula would be expected to reduce to the GLLP case as the signals grow long in a sense yet to be determined. In what follows we use very general arguments to determine the expected relation between the peak information a signal may bear and its energy. Detailed calculations are reserved for Sec. IV.

We shall often refer to the concept of channel, one widespread in communication theory. It proves particularly illuminating to think of a channel as standing for a collection of parameters and quantum numbers characterizing the *class* of signals that may be transmitted, as opposed to the particular signals. For example, a certain channel might be defined by stipulating that the signals are electromagnetic waves with right-handed circular polarization, with a definite wave-vector direction, and subject to the dispersion relation appropriate to a dielectric with a certain index of refraction. This would describe an optical fiber. Another channel might be defined by calling for signals which are longitudinal sound waves with fixed wave-vector direction and a fixed ratio between frequency and wave-vector magnitude. This would represent a pipe for voice communication.

In these examples, the parameters describing the channel, e.g., electromagnetic field, polarization, direction, etc., are distinguished from properties of the signal itself, e.g., energy, duration, etc. We take the view that the only specific and independent signal parameters are duration τ and energy E; the rest are to be descriptive of the channel. It is consistent with this point of view to regard different polarizations, quanta species, etc., as associated with separate channels. Thus unpolarized light, even if monochromatic and perfectly collimated, is regarded as propagating through two channels, say, one left- and one right-circularly polarized. A communication system involving monochromatic collimated beams of neutrinos will entail one channel for each neutrino species (flavor). This precaution is designed to remove energy degeneracies in the subsequent treatment.

The question then arises: For a particular channel, how is the maximum information $I_{\max}$ a signal may bear related to E and τ? Since reception is a necessary ingredient for the concept of signal to make sense, we suppose at the outset that all the quantities involved are measured in the receiver's Lorentz frame. In Sec. III B we show how to break free from this restriction. At any rate, the form of the relation between $I_{\max}$ and E and τ can be obtained by the following argument.[27]

Since information is dimensionless, $I_{\max}$ must be a function of dimensionless combinations of E, τ, channel parameters, and fundamental constants. We shall exclude channels which transmit massive quanta, e.g., electrons, because rest masses contribute a lot to the energy cost per bit, so that the strictest limits on the energy cost per bit and communication rate as a function of power or energy are expected for massless signal carriers.[28] Hence Compton lengths do not enter into the argument. For the reason already stated we focus on broadband channels, and exclude any frequency cutoff and its associated length. Thus no lengths describe the channel. In the absence of such lengths, there are two independent dimensionless combinations that can enter: $\xi = E\tau/\hbar$ and $\sigma\equiv\tau/T_{\mathrm{PW}}$, where $T_{\mathrm{PW}}\equiv(G\hbar/c^5)^{1/2}\approx 5.4\times 10^{-44}$ s is the Planck-Wheeler time.

In this paper we do not attempt to study signals with significant self-gravitation, such as would be of interest in black-hole physics. (However, we shall have occasion to discuss signals propagating through exterior gravitation fields.) Evidently, if self-gravitation of the signal is negligible, the parameter σ, the only parameter which contains G, should not enter into the formula determining $I_{\max}$. In everyday situations the condition for this is easily satisfied. The ratio of gravitational self-energy of the signal $G(E/c^2)^2/c\tau$ to E is of order ξ/σ^2. For usual signals σ^2 is very large (even for picosecond signal duration, $\sigma \sim 10^{64}$), so excluding signals with extremely large ξ, self-gravitation effects are negligible. Obviously non-self-gravitating signals form an important class. For such, the exclusion of σ means that

$$I_{\max} = \mathcal{I}(E\tau/\hbar) , \tag{16}$$

where $\mathcal{I}(\xi)$ is some non-negative valued function characteristic of the channel which we call the *characteristic information function* or CIF.

The reader may find it surprising that the ratio c_*/c, where c_* is the propagation speed of signals, e.g., the speed of sound, was not considered in our argument. Obviously the ratio, if different from unity, is a property of the channel, not of individual signals. Therefore, it is regarded as determining the *form* of the one-argument function $\mathcal{I}(\xi)$. We shall see in Sec. IV that in many cases c_*/c drops out entirely from the CIF.

Let us check our result. Consider steady-state communication. Because of the statistically stationary character of the signal, it should be possible to infer the peak-communication rate by considering only a finite section of the signal-bearing information $I_{\max}$ and energy E. It should matter little how long a stretch in τ is used. This can only be true if $\dot I_{\max} \equiv I_{\max}\tau^{-1}$ is determined by the power $P \equiv E\tau^{-1}$. This is consistent with $I_{\max} = \mathcal{I}(\xi)$ only if $\mathcal{I}(\xi) = \beta\sqrt{\xi}$, where β is a constant; only then does τ cancel out. It follows that $\dot I_{\max} = \beta(P/\hbar)^{1/2}$, which is precisely the GLLP formula. The argument is, however, too general to say anything about the value of β which depends sensitively on the channel's parameters.

The dividing line between steady-state communication and communication by means of very long signals is not sharp. This suggests that long signals must also obey a GLLP-type formula, albeit approximately. Indeed, long ago Marko[20] proposed that $I_{\max} \propto (E\tau/\hbar)^{1/2}$ for long-duration signals. As we shall see in Sec. IV, for $\xi = E\tau/\hbar \gtrsim 100$, $\mathcal{I}(\xi) \to \beta\sqrt{\xi}$. In this connection we may mention Joos's claim[25] that $I_{\max} \propto (E\tau)^{1/2}$ for *arbitrary* duration. He argues from extensivity of the information as a signal is partitioned into smaller pieces. Now, in thermodynamics[26] a similar argument concerning entropy gives reasonable results if the pieces of the system are themselves macroscopic, and boundary effects negligible. As will become clear in Sec. IV, for $E\tau \lesssim 100\hbar$, end effects in a signal become significant so that the extensivity assumption breaks down. As a consequence $\mathcal{I}(\xi)$ departs from the form $\sqrt{\xi}$ at low ξ.

B. Lorentz invariance and gravitational fields

It is easy to show that Eq. (16) is a Lorentz-invariant statement under wide circumstances, and valid in the presence of exterior gravitational fields. Let us demonstrate the Lorentz invariance by some examples. First, consider a "medium" such as a fluid or dielectric solid in which signals propagate with fixed speed c_* and no dispersion. The carrier quanta could be phonons propagating in the fluid, or "dressed" photons propagating in a dielectric channel, etc. We assume there are no currents (flows) in the medium so that all of it is at rest in a given Lorentz frame A. Consider another Lorentz frame B moving to the right relative to A with speed V. Without loss of generality we may assume that their origins coincide at time $t_A = 0$. Let a right-moving signal's front pass the origin of A at that same time. We assume $V < c_*$; the opposite case can be studied with appropriate changes. At some time $t_A = t_1$ the signal's rear end will pass the origin of A, at which time the origin of B has reached position $x_A = Vt_1$. At some later time $t_A = t_2$ the signal's rear has caught up with the origin of B which is then at $x_A = Vt_2$. Calculating entirely in A, we find $(c_* - V)t_2 = c_* t_1$ so that

$$t_2/t_1 = (1 - V/c_*)^{-1} . \tag{17}$$

Evidently, the duration of the signal in A is just $\tau_A = t_1$. Because of time dilation, the duration in B is just $\tau_B = t_2\gamma^{-1}$, where $\gamma \equiv (1 - v^2/c^2)^{-1/2}$ is the Lorentz factor between the frames. Then by virtue of (17) we have

$$\tau_B = \tau_A(1 - V/c_*)^{-1}\gamma^{-1} . \tag{18}$$

Let us now look at the energy. If in A the energy and momentum of a quantum are ε and p, respectively, then, by virtue of the constancy of the propagation velocity, $\epsilon = c_* p$. The total energy E_A and momentum of the signal must stand in the same ratio if there are no interactions. Therefore, by the Lorentz transformation of energy and momentum, the signal energy in frame B is

$$E_B = \gamma E_A(1 - V/c_*) . \tag{19}$$

We now see from (18) and (19) that $E_A\tau_A = E_B\tau_B$ which shows that the quantity ξ is the same in the propagating medium's frame and in some other frame which might be that of the receiver in motion with respect to the medium. It is possible to demonstrate the invariance when B is the *transmitter's* frame by having frame B move to the left, and the signal to the right, with respect to A. Of course, the information I is itself a Lorentz invariant. The end result is that the formula $I_{\max} = \mathcal{I}(E\tau/\hbar)$ is Lorentz invariant. In particular, it has the same form in the frames of the medium, the transmitter, and the receiver.

When the signal moves precisely with the speed of light, e.g., photons in empty space, the above argument may be rephrased by taking frame A as the transmitter frame, while B is some other frame, like the receiver's. The calculations go through formally as before, and demonstrate the Lorentz invariance of $I_{\max} = \mathcal{I}(E\tau/\hbar)$ in this case also.

Up to now we have implicitly assumed that the signal propagates in flat space-time (no gravitational field). Consider now its propagation through an external gravitational field, or in the expanding universe, between stationary transmitter and receiver. Redshift effects will make the E and τ at reception differ from those at transmission. However, $E\tau$ will be the same. To verify this, we focus on a single monochromatic wave component of the wave packet representing a particular signal state. Evidently, the variation of the phase from the front to the rear of the packet must be conserved in transit. At a fixed point in the frame A of the transmitter the overall phase variation is just $\omega_A \tau_A$, where ω_A is the angular frequency or time derivative of the phase in frame A. Analogously, at a fixed point in the receiver's frame B the change of phase amounts to $\omega_B \tau_B$. Now for a single quantum $\varepsilon = \hbar\omega$. Therefore, if field self-interaction may be neglected (E is the sum of ε's), if the signal transit is adiabatic (no quantum transitions between various states), and if dispersion is absent (signal does not spread), then $E\tau$ will be conserved in transit. The same adiabaticity assumption guarantees that information is not lost. Thus Eq. (16) is equally valid as applied to transmitter or receiver (or in any motionless frame in between). By combining this with our result on Lorentz invariance interpreted locally, we conclude that formula (16) must be valid in all Lorentz frames, and in the presence of external gravitational fields.

C. Self-heralding versus heralded signals

In the remainder of this paper we interpret E, the signal energy, as the mean value of the signal state's energies,

$$E \equiv \sum_a p_a E_a \ , \tag{20}$$

where the p_a's are the *a priori* probabilities that enter into the calculation of I. What set of *a priori* probabilities p_a maximizes the formal information of a signal I given the mean value of its energy? Maximizing Eq. (1) subject to this constraint, and to the normalization of probabilities by the method of Lagrange multipliers gives

$$p_a = C 2^{-\mu E_a} \ , \tag{21}$$

where μ is the Lagrange multiplier related to the energy constraint; that related to probability normalization is contained in the necessarily positive normalization constant C. So far the result is like that for thermal equilibrium.

Equation (21) is the formal answer to our question, but it may overestimate the peak information that the signal can deliver to the receiver. For one thing, noise in the channel can change the answer, as it does in classical Shannon theory.[1] Since we exclude external noise, a more immediate concern are limitations placed on the communication by the nature of the reception. This is a well-studied area, e.g., Ref. 10. Here we only draw attention to one aspect. The distribution (21) assumes that the available signal states can all be detected by the receiver and distinguished from one another. Obviously, if due to peculiarities of the receiver, several states are confused, a more representative distribution is one which makes all these *a priori* equally likely. A subtler problem is confusion between the vacuum-signal state and the absence of a signal. How would the signal receiver know that a signal has arrived with the field in the vacuum (or ground) state, rather than no signal having been received? At first sight there is no distinction between the two events. Before delving deeper into the matter it is appropriate to recall that even if the vacuum should prove impossible to detect directly, this undetectability need not always prevent its use in signaling.

For example, in a man-made channel transmitting a train of signals at equally spaced intervals, the absence of any energy in a particular time interval (not the first or last) implies that *that* signal is in the vacuum state. The embedding of the signal in a series is not even necessary for the inference. If two friends P and Q agree that if P passes his exam, he will phone Q between 2 and 3 p.m., then if Q's phone fails to ring in that period, then Q has acquired a bit of information (P has failed) by getting the vacuum state of the signal. If in a scattering experiment at an accelerator no relevant events are detected, information is obtained (upper bound on a cross section) by the vacuum state of the signal. What is common to these examples is that the signal is anticipated by virtue of being part of a structure (series), by prior agreement (phone if you pass), or by causality considerations (no scattering expected unless accelerator beam is on). A signal of this sort is aptly termed a *heralded signal. For heralded signals the vacuum-signal state, even if not* directly detectable, can be put to use in signaling jut as any other state. This means that the prescription (21) extends to the $E_a = 0$ state.

An analog to the inference that a signal has arrived in the vacuum state would be the assertion that the electromagnetic field in a cavity is in the vacuum state because no photons are detected (elimination of other states). However, quantum physics offers a method for the direct detection of the vacuum state of a field: Measure the influence on its energy of varying boundary conditions (Casimir effect). *If* this approach could be brought to bear on signals, then the vacuum state could be included in the prescription (21) even if the signal is not heralded.

What analog of the Casimir effect boundaries is required for finite duration signals? It would seem that enclosure on all sides is essential. Now a signal may be confined laterally, as in an optical fiber, but along the axis of the channel its extremities are not sharp, let alone confined. This is not conducive to the appearance of a Casimir effect.[29] Clearly, equipping the signal's front and back with "walls" in an effort to make the analogy complete transforms the situation into one of "communication by transport" of a boxed in field. But we agreed that this is not a very illuminating paradigm of communication. And insertion of the walls by the receiver will produce a Casimir effect whether a signal arrives, or not. This is so because, on account of its Lorentz invariance, the vacuum is the same for traveling signal or for receiver at rest. Particularly, this last observation strongly suggests that the signal vacuum cannot be distinguished from the lack of signal by any technique. In the absence of contrary evidence, we shall accept the necessity of this ambiguity tentatively.

In some quarters the notion that the signal-vacuum state is undetectable has met with disbelief.[2] This should disappear when the distinction between direct detection (Casimir-type measurement) and elimination of other states is kept in mind.

Consider now a signal whose arrival time is unanticipated. For reasons that will become clear forthwith, we call such a signal *self-heralding*. Its vacuum state cannot be inferred by elimination since the receiver does not know when to expect it, and so cannot carry out the necessary measurements, e.g., counting photons in a prescribed time interval. Neither can the vacuum state be detected directly. Hence such a signal, if received, is always received in a nonvacuum state: the signal heralds itself. An example might be the reception of the first burst of energy and information from a heretofore unsuspected supernova (the recent one in the large Magellanic Cloud providing a dramatic illustration). This is the first signal through a newly opened channel. Its being first insures the absence of information that could have heralded its coming: it is a self-heralding signal.

Evidently, the vacuum signal state, not being distinguishable from the "no signal" situation for a self-heralding signal, must be excluded from the list of signal states. Formally this means $p_{\rm vac} \equiv Pr(E_a=0)=0$. The derivation leading to Eq. (21) can still be carried out with the remaining states. The result can be written in a form applicable to both types of signals,

$$p_\alpha = C \times \begin{cases} 2^{-\mu E_\alpha}, & E_\alpha \neq 0 \\ 1-\zeta, & E_\alpha = 0, \end{cases} \tag{22}$$

where $\zeta=0$ for heralded signals and $\zeta=1$ for self-heralding ones. Although other values of ζ seem to have no physical relevance, all the calculations to follow are unified if we keep ζ general.

D. Generic properties of the CIF

The normalization constant C is easily calculated,

$$C=(Q-\zeta)^{-1}, \quad Q \equiv \sum_\alpha 2^{-\mu E_\alpha}, \tag{23}$$

where the sum over states Q (analogous to the thermodynamic partition function) *includes* the vacuum state. By the usual trick of statistical mechanics the expression for the mean energy can be cast into the form

$$E = \partial \log_2 C / \partial \mu . \tag{24}$$

This is to be viewed as determining μ (analogous to reciprocal temperature) in terms of the prescribed E, and is valid for both heralded and self-heralding cases.

The calculation of $I_{\max}$ from (1) with the distribution (22) is a bit more subtle, but with due caution the correct result emerges,

$$I_{\max} = \mu E - \log_2 C - C(1-\zeta)\log_2(1-\zeta) . \tag{25}$$

Formally the last term vanishes for both $\zeta=1$ and $\zeta=0$. Of course, this does not mean that self-heralding and heralded signals bear identical information because, for given E, the two will have different μ's [see (23)].

Equations (24) and (25) give, in parametric form, $I_{\max}$ as a function of E. Given Eq. (16) they thus determine the form of the CIF. Several properties of the CIF follow immediately. For example, differentiating (25) with respect to E and using (24), we get for $\zeta=0,1$ that $\partial I_{\max}/\partial E=\mu$. Since μ must be positive (otherwise Q would diverge and C would vanish), we find that $\mathcal{J}(\xi)$ is always an increasing function (τ is a fixed parameter in the present exercise). The above result is also formally valid for $0<\zeta<1$.

A look at (23) shows that in the limit of small μ (large E or ξ), the sum over states overwhelms ζ. Thus at large argument the CIF's for heralded and self-heralding signals must merge. As we have already hinted, they go over into the CIF associated with the GLLP formula, $\mathcal{J}(\xi) \propto \sqrt{\xi}$ (see Sec. IV B).

Taking the second derivative of Eq. (25), and observing that necessarily $\partial E/\partial \mu < 0$ by the analogy between μ and inverse temperature, we discover that the CIF is always a convex function of its argument. Again, this conclusion is formally valid for $0 \le \zeta \le 1$. Note that the CIF for infinitely long signals, $\mathcal{J}(\xi) \propto \sqrt{\xi}$, agrees with this. An immediate consequence of convexity is that a signal of energy NE and duration $N'\tau$ cannot carry as much information as NN' signals of energy E and duration τ.

IV. THE CIF FOR OCCUPATION-NUMBER STATES

A. Noninteracting modes model

The determination of the CIF for a given channel hinges upon the calculation of the sum over states Q [see (23)]. In accordance with earlier discussion the states referred to in Sec. III will be interpreted as (pure) quantum field states, and denoted by $|a\rangle, |b\rangle, \ldots$. These states have *a priori* probabilities $p_a, p_b, \ldots$. Notice that we do not use here the density-operator description. That operator includes, in its off-diagonal terms, correlations which are foreign to the business at hand. Were we to use the density operator, and for consistency the von-Neumann quantum formula for entropy,[26] we would introduce apparent contributions to the information of signals which could not be ferreted out by a receiver whose job is to distinguish one pure state from another. Thus the statistical description of the signal involves only the diagonal part of the density operator, $\{p_1, p_2, \ldots\}$.

A simplification we invoke is that the field in question can be described as a free field. If the field is subject to interactions (arguably it must be for communication through it to be possible), we assume the choice of *propagating* normal modes made manages to eliminate any cross-interaction terms, e.g., normal modes in an elastic solid. The field Hamiltonian will thus be equivalent to a collection of noninteracting harmonic oscillators. Depending on what it takes to do this, the quanta will be free particles, e.g., photons or quasiparticles, e.g., phonons.

What do the field states look like? Consider a single mode j. To it corresponds a harmonic-oscillator Hamiltonian H_j with a certain frequency ω_j. One type of state of mode j is the occupation-number state $|j\alpha\rangle$ defined by $H_j|j\alpha\rangle = n_\alpha \hbar\omega_j |j\alpha\rangle$, where n_α is a non-negative integer. Other choices like coherent and squeezed states[10] are not eigenstates of the mode Hamiltonian. However, any state $|j\alpha\rangle$ does have a well-defined mean energy $\varepsilon_{j\alpha}$,

$$\varepsilon_{j\alpha}=\langle j\alpha\,|\,H_j\,|\,j\alpha\rangle\ . \quad (26)$$

We can now build the signal states $|\,a\,\rangle$ by exploiting the independence of the H_j, namely

$$|\,a\,\rangle=|\,j\alpha\rangle\otimes|\,k\beta\rangle\otimes\cdots\ , \quad (27)$$

where $j,k,\ldots,$ label modes while $\alpha,\beta,\ldots,$ label one-mode states, and $a,b,\ldots,$ label signal (many-mode) states.

The probabilities p_a of the signal states are assumed to be normalized to unity. It is not necessary for the signal states to form a complete set in the sense of quantum theory. But completeness obviously favors higher communication rates by making a maximum number of states available, and will be assumed henceforth. We start by defining the *mean* energy of the signal,

$$E=\sum_a p_a(\varepsilon_{1\alpha}+\varepsilon_{2\beta}+\cdots)\ . \quad (28)$$

Two averages are involved here: a quantum expectation value over the one-mode states which yields $\varepsilon_{1\alpha}+\varepsilon_{2\beta}+\cdots$, and a statistical average over the *a priori* probabilites. Clearly, only the latter one was involved in the calculations of $I_{\max}$ in Sec. III D. Thus from our point of view the expression $\varepsilon_{1\alpha}+\varepsilon_{2\beta}+\cdots$, though formally a quantum expectation value, can be treated as a definite energy E_a.

Turn now to the sum over states, Q. The sum over a is equivalent to one over all combinations of j and α. Thus in a manner analogous to well-known thermodynamic calculations, Q can be written as $\prod_j Z_j$, where

$$Z_j(\mu)\equiv\sum_\alpha 2^{-\mu\varepsilon_{j\alpha}}=\sum_\alpha \exp(-\mu\langle j\alpha\,|\,H_j\,|\,j\alpha\rangle\ln 2)\ . \quad (29)$$

Further progress necessitates a specific choice of one-mode states since the sum in Eq. (29) is not invariant under a unitary transformation of the $|\,j\alpha\rangle$.

For the rest of this section we consider occupation-number states. To actually approach the channel capacity we are after, the receiver should be able to *count* quanta. An example for optical channels would be a photoelectric tube equipped with photon-counting electronics. However, we shall ignore aspects of the detection process and concentrate on the propagation of signals.

If the states $|\,j\alpha\rangle$ are chosen as occupation-number states, $\langle j\alpha\,|\,H_j\,|\,j\alpha\rangle=n_\alpha\hbar\omega_j$. For a bosonic field, n_α can be any non-negative integer; for a fermionic one, $n_\alpha=0,1$. Then for bosons, Z_j reduces to the partition function of a harmonic oscillator at temperature $(\mu\ln 2)^{-1}$:

$$Z_j=\sum_{n=0}^{\infty}2^{-\mu n\hbar\omega_j}=(1-2^{-\mu\hbar\omega_j})^{-1}\ . \quad (30\text{a})$$

For fermions

$$Z_j=1+2^{-\mu\hbar\omega_j}\ . \quad (30\text{b})$$

To calculate Q we first sum $\log_2 Z_j$ over modes, and then exponentiate the result to the base 2.

B. The continuum limit

Consider the case of small μ analogous to the thermodynamic high-temperature limit. This corresponds to large E, i.e., large ξ. The exponent in Eqs. (30) changes gradually with ω_j so that we may use the continuum approximation. Assuming uniformity of the channel along the propagation direction, we label modes by momentum (index j is just p). For maximum generality we shall allow for the possibility of dispersion, a variation of the mode propagation (group) velocity $c_*(\varepsilon)$ with the one-quantum energy $\varepsilon=\hbar\omega$. As Pendry's argument reviewed in Sec. II shows, for steady-state signaling, dispersion does not affect $\dot I_{\max}$. Of course, for signals of finite-duration dispersion will tend to spread the signal, and may thus be detrimental to the goal of high $\dot I$. The optimum case is obtained when the quanta in a slow mode are transmitted earlier than those in a fast mode in such a way that the state arriving at the receiver corresponds to the nearly simultaneous arrival of quanta in all modes within a fixed time interval τ. This may not be a practical arrangement, but should give an upper bound on the $\dot I_{\max}$ of a realistic channel.

The number of modes at energy ε within the duration τ and momentum interval dp is evidently $c_*(\varepsilon)\tau\,dp/2\pi\hbar$. Since $c_*(\varepsilon)=d\varepsilon/dp$ we have

$$\log_2\left[\prod_j Z_j(\mu)\right]=\mp(\tau/2\pi\hbar)\int_0^\infty \log_2(1\mp 2^{-\mu\varepsilon})d\varepsilon\ , \quad (31)$$

where upper (lower) signs are for boson (fermion) quanta. The range of integration corresponds to the momentum range $[0,\infty]$ since we are only considering modes traveling in one direction. Note that p no longer appears explicitly; this assures us that dispersion is irrelevant in any subsequent results (provided sequencing of the quanta is carried out, as explained earlier). Even the typical magnitude of c_* has disappeared. In both of these features we see reemerging the phenomenon pointed out by Pendry for steady-state communication: the maximum communication rate is independent of signal velocity. The range of integration in Eq. (31) should actually have a lower cutoff since a finite-duration signal cannot comprise modes of arbitrarily low energies (frequencies). However, the approximation entailed in setting the cutoff to zero is of the same order as that involved in the continuum approximation; therefore, our expression is consistent.

The integration in (31) is performed by first switching to natural exponential and logarithm and then integrating by parts. The final result is

$$\prod_j Z_j(\mu)=\exp\left\{\frac{\pi\tau\log_2 e}{12\mu\hbar}\right\} \quad (32)$$

for bosons; for fermions the factor 12 is replaced by 24. It is now clear that for small μ (more precisely small $y\equiv\mu\hbar/\tau$), $Q\gg\xi$ so that $C\approx Q^{-1}$ for self-heralding signals. For heralded signals this is, of course, an exact result. Equation (24) now gives

$$E=\pi\tau(\mu^{-1}\log_2 e)^2/12\hbar \ .$$

Calculating $I_{\max}$ from Eq. (25), and eliminating μ between the results gives

$$I_{\max}\rightarrow(\pi E\tau/3\hbar)^{1/2}\log_2 e\times\begin{cases}1 & \text{(boson)}\\ 2^{-1/2} & \text{(fermion)},\end{cases} \tag{33}$$

Apart from the numerical constant, this is just Marko's[20] expression for $I_{\max}$. It reduces to the GLLP formula (4) under the substitutions $E/\tau\rightarrow P$ and $I_{\max}/\tau\rightarrow\dot{I}_{\max}$.

Thus as anticipated in Sec. III A, for large ξ, $\mathcal{I}(\xi)\rightarrow\text{const}\sqrt{\xi}$ and the difference between heralded and self-heralding signals disappears. According to Eqs. (23) and (32), the merging of the results for heralded and self-heralding cases occurs when $\pi/12y$ is at least a few times unity. Our expression for E can be written as

$$\xi=\pi(y^{-1}\log_2 e)^2/12 \ .$$

Hence, the merging should be apparent for $\xi\gtrsim 10^2$. This can also be taken as the criterion for approach of the limit (33). Thus the long signals for which the GLLP formula may already be applied are those with $E\tau/\hbar\gtrsim 100$.

C. The periodic boundary condition

When $\xi\lesssim 100$ ($y\gtrsim 0.1$) the continuum approximation becomes poor, and we must resort to numerical summation of $\log_2 Z_j$. The required detailed knowledge of the spectrum of ω_j can be obtained by the known trick of imposing periodic boundary conditions. Think of the signal as seen at a fixed point of the receiver. Now, a function $F(t)$ defined only in $[0,\tau]$ can be represented over that interval by a complex Fourier series involving the circular frequencies $2\pi j\tau^{-1}$ for all integers j. In the process of field quantization, the negative frequencies enter automatically alongside the positive ones, and do not represent extra modes. The dc mode ($j=0$) may be ignored in our contexts as it relates to a condensate of the field. So the spectrum is $\omega_j=2\pi j\tau^{-1}$ with $j=1,2,\ldots$ and with no degeneracies. The usual degeneracy corresponding to oppositely directed momenta with equal magnitude is irrelevant here since one is only interested in propagation in one direction. Note that since we need not bring up the relation between the energy and momentum of a quanta, our results will again be insensitive to dispersion or the magnitude of the signal velocity. (The proviso of sequencing is, however, still in force.)

As mentioned, $\log_2 Q=\sum_j\log_2 Z_j$, a sum which can be performed directly. The mean energy E is most conveniently calculated by analytically performing the differentiation in Eq. (24): one

$$\frac{E\tau}{\hbar}=2\pi\left[\frac{Q}{Q-\xi}\right]\sum_{j=1}^{\infty}\frac{j}{2^{2\pi\mu j\hbar/\tau}\mp 1} \ , \tag{34}$$

We have numerically summed the above expressions for $\log_2 Q$ and E over the spectrum ω_j described above for a range of values of the parameter y, and for both the boson and fermion cases. $I_{\max}$ was computed from (25).

The results are plotted in a log-log scale in Fig. 1 for heralded signals and in Fig. 2 for self-heralding ones.

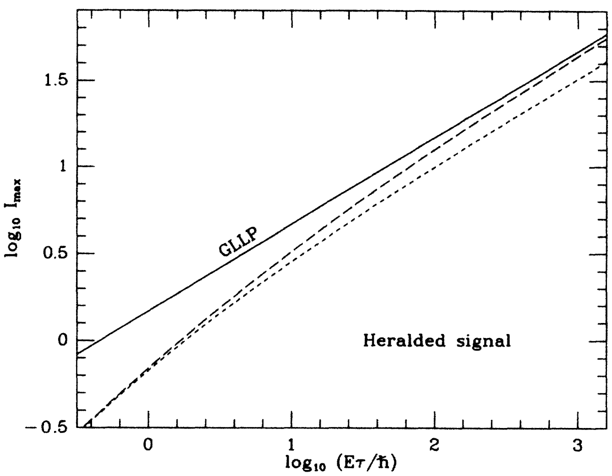

FIG. 1. Log-log plot of the maximum information a heralded signal can carry as a function of $E\tau$. The long-dashed (short-dashed) curves are for boson (fermion) carrier fields. The solid line corresponds to the limiting formula (33), itself a form of the Gordon-Lebedev-Levitin-Pendry channel-capacity formula (4).

The long-dashed lines are for bosons, while short-dashed lines are for fermions. Note that fermion and boson graphs merge at low ξ but begin to separate for $\xi\gtrsim 10$. Note also that for self-heralding signals $I_{\max}\rightarrow 0$ for $\xi=2\pi$. This is because at least one quantum must be present and the finite duration keeps it from having arbitrarily low energy.

The solid line labeled GLLP, drawn for comparison, is the limiting relation (33) which gives the same relation between $\dot{I}_{\max}$ and P as the GLLP formula (4). For clarity only the boson-limiting relation has been drawn; the fermion version lies parallel to it through $I_{\max}$'s a factor 0.707 as large. It may be seen that the GLLP-like rela-

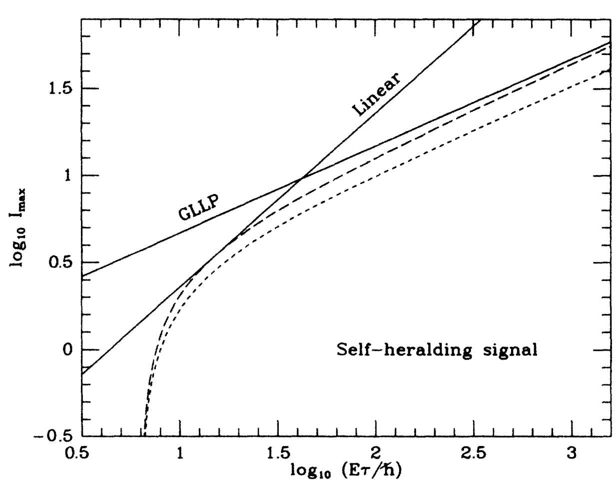

FIG. 2. Log-log plot of the maximum information a self-heralding signal may bear as a function of $E\tau$. The long-dashed (short-dashed) curves are for boson (fermion) carrier fields. The solid line labeled GLLP is the limiting relation (33), while that labeled Linear corresponds to Bremermann's bound (5), as made precise by (38).

tion is a strict upper bound on I_{max} (both for heralded and self-heralding cases), and an excellent approximation to it for signals with $\xi \gtrsim 10^3$ corresponding to $I_{max} \gtrsim 44$ bits. This roughly agrees with the estimates of Sec. IV B. As ξ decreases, true I_{max} of finite-duration signals falls below the naive prediction of (33) by factors which at $\xi = 10$ reach 2.5 and 4 for heralded and self-heralding signals, respectively. The corresponding true I_{max}'s are ≈ 3 bits for heralded signals, and ≈ 2 bits for self-heralding ones. Thus signals carrying modest information have to be treated as finite-duration signals, rather than by the GLLP formula. The line labeled Linear will be discussed in Sec. IV E.

Our results may also be displayed in the form of a plot of the energy cost per bit ε_{min} as a function of I_{max}, as is done in Fig. 3 for both heralded and self-heralding signals. For clarity only the boson case is plotted. The line labeled GLLP corresponds to Eq. (15); clearly for finite-duration signals the energy cost per bit exceeds that implied by the theory of steady-state communication. It may be seen that for self-heralding signals there exists a lower bound of $\approx 4.4\hbar/\tau$ on ϵ_{min} which is attained for $I_{max} \approx 3.5$ bits (more precise numbers are given in Sec. IV E). No such bound exists for *heralded* signals: the energy cost per bit can be low for signals with only fractions of a bit. Such low-information signals *are* meaningful. For example, if a question has three alternative answers with the first being 98% probable, then 0.3 bits suffice to single out the answer [see Eq. (1)].

D. The wave-packet approach

In the method used above the spectrum of mode frequencies is discrete. In actuality a signal confined to an interval τ should contain a continuum of frequencies. Gabor devised the method of frequency-time cells[4] to deal with just such a situation. In this method the class

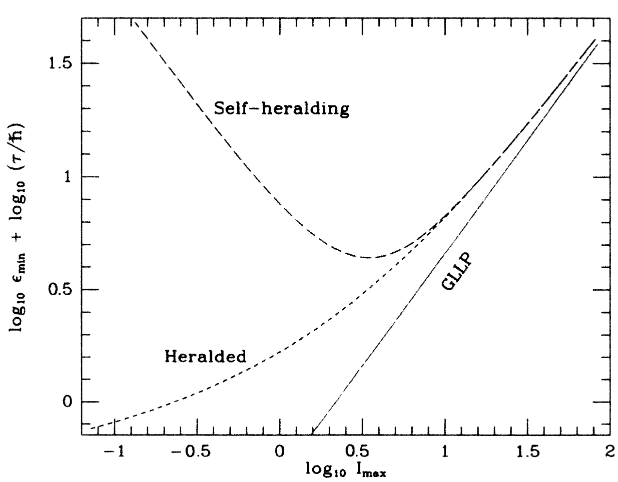

FIG. 3. Log-log plot of the energy cost per bit as a function of maximum information for heralded or self-heralding signal. Only the case of boson carriers is shown. The line labeled GLLP corresponds to Eq. (15). In order to show graphs for all possible durations on the same scale, $\log_{10}(\tau/\hbar)$ has been added to the value displayed in the ordinate.

of signals confined to an interval of time τ and containing frequencies in the range $[0, \Omega]$ is represented by subdividing the occupied part of the $\omega - t$ plane into rectangular cells with equal areas $\Delta\omega\Delta t = 2\pi$. To cell j containing the point $t = t_j$, $\omega = \omega_j$ is associated the Gaussian modulated wave (equivalent to a modulated sinusoid with general phase)

$$r_j(t) \equiv \exp[-(t-t_j)^2/2(\Delta t)^2]\exp[-i\omega_j(t-t_j)] \ . \quad (35)$$

A specific signal is represented as a superposition of $\Omega\tau/2\pi$ such functions.

The Fourier transform of $r_j(t)$,

$$R_j(\omega, \Delta t) \propto \exp[-(\Delta t)^2(\omega-\omega_j)^2/2 + i\omega t_j] \ , \quad (36)$$

represents a Gaussian distribution of pure exponentials centered at $\omega = \omega_j$ with rms $(\Delta t)^{-1}$. The cell partition with spacing $\Delta\omega = 2\pi/\Delta t$ essentially prevents overlap of the Fourier components of adjacent cells. The broadband channel we are examining requires that $\Omega \to \infty$ so that regardless of the choice of partition Δt, an infinite series of ω_j's is involved. To avoid degeneracies, we choose $\Delta t = \tau$, so that only one cell straddles each particular ω_j. Evidently, what has been done is to go over from a description of modes which are complex exponentials in time to one of Gaussian wave packets with temporal rms corresponding to the proposed signal duration.

It should be evident that the spacing of the ω_j's here is $2\pi/\tau$, just as in the periodic-boundary-condition method. The centers of the Gaussians may thus be taken at the discrete spectrum frequencies of Sec. IV C. In that approach the $\log_2 Z_j$'s were evaluated at just these ω's. In the wave-packet method they are calculated as

$$\log_2 Z_j = \mp \int_0^\infty |R_j(\omega, \tau)|^2 \log_2(1 \mp 2^{-\mu\hbar\omega}) d\omega \ , \quad (37)$$

where the R_j are assumed normalized over positive frequencies. The above convolution "smears" each mode's $\log_2 Z_j$ over a range $\sim \tau^{-1}$ of positive frequencies, thus taking account of the finite signal duration. The resulting effective $\log_2 Z_j$'s are then used as in Sec. IV C to calculate Q, E, and I_{max}.

Numerical integration confirms that for small y the results of the wave-packet method accurately reproduce those of Sec. IV C. Thus for $y = 0.25$, 0.5, and 0.7, a Z_j obtained by convolution differs from one obtained by directly substituting $2\pi j/\tau$ for ω by 1%, 3%, and 6%, respectively. Thus the accuracy of the graphs in Figs. 1 and 2 should be a few percent or better down to $I_{max} \sim 0.6$ bit for self-heralding signals and $I_{max} \sim 0.3$ bit for heralded ones, points which correspond to $y = 0.7$. The graphs in Figs. 1 and 2 can thus be trusted except for the extreme low end.

E. The linear bound on communication

The line in Figs. 1 and 2 labeled Linear corresponds to the upper bound (5) on $\dot{I}_{max}$ conjectured by Bremermann,[7,15] and by the present author,[16] for finite duration signals. It is well to recollect these early arguments for the existence of such bound before going into how it is formally derived.

Bremermann's argument[7] is that the energy alloted to a signal limits the bandwidth available to the communication system in question, which limitation results in a bound on $\dot{I}_{\max}$ via the Shannon capacity formula (2). His main point is that a signal, when looked at in quantum terms, must contain at least one quantum of some sort. Thus for alloted energy E, the angular frequencies that can appear are bounded from above by $E/\hbar$. This is interpreted as the bandwidth $\Delta\omega$ of the system. Bremermann goes on to estimate a maximum value of P/N by interpreting N in terms of the uncertainty of energy which enters into the time-energy uncertainty relation. The result is inequality (5). Bremermann's argument can be criticized for relying on the classical Shannon formula to get an ostensibly quantum result, and for the obscurity surrounding the connection of noise power and the time-energy uncertainty relation, itself a principle that invites confusion.[30]

The second road to (5) relies[16] on causality considerations combined with the bound on the entropy S that may physically be confined to a system with definite linear size R and total rest energy E.[21] This bound, namely, $S \le 2\pi ER/\hbar$, was originally inferred from black-hole thermodynamics, but has since been established by detailed numerical experiments[22] and analytic arguments.[23,24] The argument is that the peak entropy that *could* be in a system limits the total information $I_{\max}$ that can be stored in it. Communication is envisaged as resulting from transport of the system. Thus the *rate* at which information is acquired by the receiver is evidently limited by the linear dimension of the system and by the fact that it travels no faster than the speed of light, as well as by the entropic bound on $I_{\max}$. Bound (5) results from combining these limitations with geometric and Lorentz factors. A limitation of the argument is that transport is not the only way for communication to take place. Further, the argument cannot deal with signals traveling at the speed of light. For these the notion of rest energy is absent, and the bound on specific entropy is ambiguous.

Bremermann regarded bound (5) as holding for a number of channels in parallel.[15] When the communication system is looked at in a deep way, there may be some validity to such a view,[22] but the straightforward conclusion must be that a bound of the form of (5) can be relied upon only for one channel. This point was made by Landauer and Woo[8] and reiterated by Levitin.[17] The related difficulty in formulating bounds on communication for three-dimensional multichannel systems has been discussed by Pendry.[9]

Even for a single channel, our preceding results show that a bound of the form (5) cannot be everywhere valid for heralded signals, whatever the α. For finite-duration signals we can rewrite (5) as $I_{\max} < \alpha E\tau/\hbar$. In a log-log plot such as that in Fig. 1, the boundary of such an inequality is a straight line with unit slope. It can be seen that such a line must cut the graphs $\mathcal{I}(\xi)$ in the heralded case (unless the corrections to the periodic-boundary-condition method bend the graph very strongly at low ξ). Thus any linear bound will eventually be violated. However, as Fig. 2 shows, the graph of $\mathcal{I}(\xi)$ in the self-heralding case can be bounded by a line of unit slope. The tightest such bound is the line labeled Linear in Fig. 2. The dichotomy between heralded and self-heralding signals is related to Bremermann's point that a signal must involve at least one quantum. From the discussion in Sec. III C it follows that this is indeed true for self-heralding signals; for these the vacuum state is illegal, and in the context of occupation-number states, all such signals must have at least one quantum. This conclusion is not relevant for heralded signals, which thus sidestep the linear bound.

The constant α of the best linear bound has been accurately determined by seeking the maximum value of $I/E=(\epsilon_{\min})^{-1}$ subject only to normalization of probabilities.[27] As is clear from Fig. 3, such a maximum exists only for self-heralding signals. With the periodic boundary condition the results are[27]

$$I_{\max} \le \frac{E\tau}{\hbar} \times \begin{cases} 0.2279 & \text{(bosons)} \\ 0.1878 & \text{(fermions).} \end{cases} \tag{38}$$

For clarity, only the boson version of the bounding line corresponding to (38) has been plotted in Fig. 2. It also bounds the fermion $\mathcal{I}(\xi)$.

We observe that the linear bound so calibrated can actually be attained, but only for $\xi = E\tau/\hbar \approx 15.17$ and $I_{\max} \approx 3.458$ bits (boson case; for fermions the corresponding numbers are $\xi \approx 14.93$ and $I_{\max} \approx 2.803$ bits). These numbers correspond to the minimum of the graph in Fig. 3 since the equality sign in Eq. (38) corresponds to the least possible $\epsilon_{\min}$. Thus $\epsilon_{\min} \ge 4.388\hbar/\tau$. The linear bound is close to saturation in the vicinity of $I_{\max} \sim 3$ bits. (This agrees with Landauer and Woo's remark[8] mentioned in the Introduction.) Indeed, it provides a 20% or better estimate of the true $I_{\max}$ in the range $9 \lesssim \xi \lesssim 40$ or $1.6 \lesssim I_{\max} \lesssim 7$. The bound is thus a good rule of thumb for estimating the parameters of a signal bearing modest information. For comparison it should be noted that at the point of saturation of the linear bound, the limiting GLLP bosonic (fermionic) formula (33) already overestimates the exact bosonic (fermionic) $I_{\max}$ by 66% (44%).

How do the linear bound and (33) compare for very large or very small $I_{\max}$? Already for signals with $\xi \lesssim 42$ ($I_{\max} \lesssim 7.5$) the bosonic linear bound sets a tighter bound on the exact $I_{\max}$ than the bosonic limiting GLLP formula. For the fermionic case the corresponding ranges are $\xi \lesssim 21$ and $I_{\max} \lesssim 4$. However, for very large ξ the linear bound becomes overly "generous," and the limiting formula (33) is a better estimate of the true $I_{\max}$.

V. CONCLUSIONS AND FUTURE DIRECTIONS

We have deduced the characteristic information functions for a single quantum channel for finite-duration heralded and self-heralding signals carried by either bosons or fermions. The information that can be carried by such signals is bounded from above by the formula corresponding to the GLLP channel capacity. Hence, in energy terms, information transfer by finite-duration signals is more expensive than continuous signaling (though it must be stressed that dissipation of this energy seems not to be required by the physics[2]). The linear bound of Bremermann and the present author has been shown to apply only to self-heralding signals. It provides a good estimator of their maximum information content when that is around a few bits. It was shown that dispersion in the channel is not necessarily deleterious to the information performance, provided the quanta associated with the involved modes are suitable sequenced upon transmission.

All these results have been demonstrated only for occupation-number signal states. It remains to be explored whether the limitations just mentioned can be evaded by the use of other suitable signal quantum states. Unlike the situation in thermodynamics, a transformation between states does make a difference in communication. Indeed, studies of the question for the narrow-band case have made it clear that capacity depends on the type of state used.[10,31] It would be useful to extend these considerations to the broadband case along the lines drawn in Sec. IV. The issue is bound up with that of quantum-measurement theory, since only states which can be distinguished at the receiver need be accounted for separately in the formalism, and the distinction between states has to be made by quantum measurements.

In this paper we have interpreted the signal energy E as the mean of signal states' energies over the probability distribution which defines the information content. An equally interesting definition of E is the maximum alloted signal energy. After an early study by Gibbons[32] of the channel capacity within this framework, detailed numerical simulations by the present author[22] showed that the linear bound tends to be obeyed regardless of whether the vacuum state is allowed or not, i.e., there is no great gap between heralded and self-heralding signals within the new definition of signal energy. Presumably this means that the CIF are quite similar for both types. Unfortunately, there is little prospect that great analytic insight may be forthcoming in this approach to communication. Hence, more detailed numerical simulations of the information borne by signals whose energy is limited from above are needed.

Finally, the question of multichannel communication needs to be addressed for finite-duration signals. Multiple channels allow the GLLP capacity or the linear bound to be sidestepped.[8,9,17] Many natural communication channels can be envisaged which rely on parallel channels. For example, a beam of light with small opening angle entails many channels. But the number of channels is related to the geometric parameters of the beam, so a natural generalization of the narrow-band channel capacity is possible.[17] It remains to be seen whether similar multichannel analogs of the GLLP formula and the linear bound exist.

ACKNOWLEDGMENTS

I thank J. Pendry for incisive criticism that led to a better paper. D. Kondepudi, J. A. Wheeler, and W. Wootters are thanked for many comments and suggestions, and E. Joos, R. Landauer, and J. Pierce for correspondence and preprints. The hospitality of J. A. Wheeler at the University of Texas at Austin and of S. Tremaine at the Canadian Institute for Theoretical Astrophysics allowed me to concentrate on communication. Financial support by the Basic Research Foundation administered by the Israel Academy of Sciences and Humanities, and by the Arnow Chair in Astrophysics at Ben Gurion University is acknowledged.

---

*Permanent address: Physics Department, Ben Gurion University, 84105 Beer-Sheva, Israel.

[1]C. Shannon and W. Weaver, *The Mathematical Theory of Communication* (University of Illinois Press, Urbana, 1949); F. Haber, *An Introduction to Information and Communication Theory* (Addison-Wesley, Reading, MA, 1974).

[2]R. Landauer, in *Signal Processing,* edited by S. Haykin (Prentice-Hall, New York, 1988).

[3]L. Brillouin, *Science and Information Theory* (Academic, London, 1965).

[4]D. Gabor, Philos. Mag. **41**, 7 (1950); **41**, 1161 (1950).

[5]J. P. Gordon, in *Advances in Quantum Electronics,* edited by J. R. Singer (Columbia University Press, New York, 1961).

[6]D. S. Lebedev and L. B. Levitin, Dokl. Akad. Nauk SSSR **149**, 1299 (1963) [Sov. Phys.—Dokl. **8**, 377 (1963)].

[7]H. J. Bremermann, in *Self-Organizing Systems,* edited by M. C. Yovitz, T. C. Jacobi, and G. D. Goldstein (Spartan Books, Washington, D.C., 1962); 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).

[8]R. Landauer and J. W. F. Woo, in *Synergetics,* edited by H. Haken (Teubner, Stuttgart, 1973).

[9]J. B. Pendry, J. Phys. A **16**, 2161 (1983).

[10]Y. Yamamoto and H. A. Haus, Rev. Mod. Phys. **58**, 1001 (1986).

[11]S. Wolfram, Rev. Mod. Phys. **55**, 601 (1983).

[12]W. Wootters, Phys. Rev. D **23**, 357 (1981); D. Deutsch, Phys. Rev. Lett. **50**, 631 (1983).

[13]J. A. Wheeler, Int. J. Theor. Phys. **21**, 557 (1982); R. P. Feynman, *ibid.* **21**, 467 (1982).

[14]This motto is attributed by J. A. Wheeler (Ref. 13) to E. Fredkin.

[15]H. J. Bremermann, Int. J. Theor. Phys. **21**, 203 (1982).

[16]J. D. Bekenstein, Phys. Rev. Lett. **46**, 623 (1981).

[17]L. B. Levitin, Int. J. Theor. Phys. **21**, 299 (1982).

[18]R. Landauer, IBM J. Res. Dev. **5**, 183 (1961).

[19]T. E. Stern, IEEE Trans. Inf. Theory **IT-6**, 435 (1960); J. P. Gordon, Proc. IRE **50**, 1898 (1962).

[20]H. Marko, Kybernetik **2**, 274 (1965).

[21]J. D. Bekenstein, Phys. Rev. D **23**, 287 (1981).

[22]J. D. Bekenstein, Phys. Rev. D **30**, 1669 (1984).

[23]I. Kahn and A. Qadir, Lett. Nuovo Cimento **41**, 493 (1984).

[24]M. Schiffer and J. D. Bekenstein (unpublished).

[25]E. Joos (unpublished).

[26]L. D. Landau and E. M. Lifshitz, *Statistical Physics,* Pt. I (Pergamon, Oxford, 1980).

[27]J. D. Bekenstein (unpublished).

[28]The argument sometimes made, e.g., Ref. 2, that for signals the energy locked in mass should be ignored would seem to compromise any hope of profound information theoretic insights into physics (for which the mass-energy equivalence is a crucial pillar).

[29]This makes doubtful the possible relevance of negative vacuum energy in a communication argued by D. Unwin, Phys. Rev. D **26**, 944 (1982).

[30]L. D. Landau and E. M. Lifshitz, *Quantum Mechanics*, 3rd ed. (Pergamon, Oxford, 1977); E. P. Wigner, in *Aspects of Quantum Theory*, edited by A. Salam and E. P. Wigner (Cambridge University Press, Cambridge, 1972).

[31]B. Saleh and M. Teich, Phys. Rev. Lett. **58**, 2656 (1987).

[32]G. Gibbons (unpublished).