Quantum bounds on the information capacity of narrow-band free-space links without extraneous noise

1981 J. Phys. A: Math. Gen. 14 133

(http://iopscience.iop.org/0305-4470/14/1/013)

J. Phys. A: Math. Gen. **14** (1981) 133–143. Printed in Great Britain

Quantum bounds on the information capacity of narrow-band free-space links without extraneous noise

W G Chambers

Department of Mathematics, Westfield College (University of London), Kidderpore Avenue, London, NW3 7ST, UK

Received 10 June 1980

Abstract. An upper bound for the information rate in a communication link is set by the finite number of orthogonal states available to the electromagnetic field, if the link is subject to constraints on the average power and bandwidth. This bound may not be attainable in a free-space link since the transmitter does not have complete control over what is received. Some hypothetical systems are examined to see how close they get. At high photon rates it is possible with difficulty to improve slightly on an 'x–p' system, and at low photon rates on a photon-detecting system, but it seems that these systems are nearly the best, and certainly they are not far below the upper bound, in the sense that the ratios of the information rates to the maximum are close to unity. In all these systems the normal modes (Planck oscillators) are treated as independent channels. It is suggested that a beam splitter can provide a simple model for a free-space link.

1. Introduction

One would expect an upper bound on the information capacity of an electromagnetic link (subject to an average-power constraint) because of quantum effects, even if there were no extraneous noise (Gordon 1962, Yu 1976). We might fancifully assign this limitation to 'photon shot noise', although more precisely it is due to the finite number of orthogonal quantum states available to the electromagnetic field when subjected to limitations of space, time and energy. It seems that this bound can be approached arbitrarily closely for guided-wave links. However, in free-space links only a small fraction of the transmitted energy is received, and one might expect that there is added a kind of 'partition noise', because some photons hit the receiving antenna while others miss, or because the transmitter does not have complete control over what is received. This presumably reduces the capacity of the channel below the bound just mentioned, and the question is by how much. As a partial answer we discuss some lower bounds for the capacity, in the limiting cases when the photon rate is very low and when it is very high. These lower bounds are obtained by considering hypothetical transmitter–receiver systems. In each case we consider synchronous links when there is available to both transmitter and receiver a common reference clock–oscillator, and for comparison 'photon-counting' links. The basic quantum theory may be found in Helstrom (1976).

2. The upper bound

We first imagine that the transmitter sends the receiver a succession of well-separated particles of spin S. Each particle thus can be put into one of $2S+1$ orthogonal spin states by the transmitter, and the receiver can determine with certainty the state of each particle. Thus each spin state can represent a letter in an alphabet of $2S+1$ letters, and the maximal rate of transmission of information is $\ln(2S+1)$ natural units per particle (if the time of arrival of the particles is not used).

Let us next imagine a very long waveguide carrying a unidirectional stream of electromagnetic waves of mean frequency ν, bandwidth B† ($B \ll \nu$), and average power P. Suppose this system is chopped up into lengths with duration T ($T \gg B^{-1}$). Then each length can be regarded as a 'particle', although since it is a macroscopic system it has an enormously large number N of orthogonal quantum states available to it. However, the logarithm of this number can be calculated by the same mathematics as is used to calculate the entropy of a microcanonical ensemble, with a given total energy PT (Kubo 1971, Takahasi 1965). It is found that

$$\ln N \approx BT[(\bar{n}+1)\ln(\bar{n}+1)-\bar{n}\ln\bar{n}],$$

where

$$\bar{n} = P/h\nu B \tag{1}$$

is the photon rate divided by B, or the average number of photons in each normal mode or Planck oscillator‡. The usual postulates of quantum theory permit us to imagine that the link using spinning particles could be set up, so there seems no basic objection to the setting up of the electromagnetic link just described. The result is equivalent to the zero-temperature value of the capacity given by Gordon (1962). The rate $C_{\max}$ is given by

$$C_{\max}/B = (\bar{n}+1)\ln(\bar{n}+1)-\bar{n}\ln\bar{n}. \tag{2}$$

We would also expect this formula to provide an upper bound for the capacity of a free-space link, since the waveguide could be placed between the receiving antenna and the receiver proper. (Naturally the power P is the received power!) Presumably the receiver cannot receive information at a rate higher than that conveyed by the waveguide.

Let us next imagine an arrangement where each of the BT normal modes or Planck oscillators acts as a separate channel, used once every T seconds, and with an average photon number given by (1). We may code the channel by putting the oscillator into various eigenstates of the photon number operator, which can in principle be determined precisely by measuring this dynamical variable. If we then choose the probabilities to maximise the information rate subject to the average number being $\bar{n}$, we find that the rate is given by (2) (Takahasi 1965). The use of such a particular arrangement provides us with a way of finding a lower bound for the capacity, and in this instance this happens also to be the upper bound, which therefore seems attainable in

† It is assumed that there is only one polarisation in use. If both polarisations are used, then we have to redefine B to be twice the bandwidth.

‡ To specify the state of such an oscillator in classical theory we need a complex number or two real numbers to give the amplitude and phase. In some treatments each oscillator is made to correspond to two normal modes with real amplitudes. We only consider normal modes with positive wavenumbers, because of the unidirectional propagation.

principle. In all the arrangements to be described we shall again regard each normal mode as a separate channel, so that we need only consider a single normal mode in isolation from the others. In a free-space link the receiving oscillator may be put into any chosen 'coherent state' (Glauber 1963), but the idea just described fails because these states are not orthogonal.

Approximations to (2) for large $\bar{n}$ and small $\bar{n}$ will be needed. These are, for large $\bar{n}$

$$C_{\max}/B = \ln \bar{n} + 1 + O(\bar{n}^{-1}), \tag{3}$$

and for small $\bar{n}$

$$C_{\max}/B = \bar{n} \ln(1/\bar{n}) + \bar{n} + O(\bar{n}^2). \tag{4}$$

3. Beam splitters and coherent states

We may model the tremendous attenuation of a free-space signal on its journey from the transmitter to the receiver by means of a beam splitter. We consider only one particular normal mode and its corresponding wave-packet. A beam splitter divides an incident wave-packet into two outgoing wave-packets on separate outputs. The first output goes to the receiver and the other is lost. (There is another input for the beam splitter, but this is not used here.) Let U denote the unitary evolution operator which transforms initial states into final states. The vacuum state $|\psi_{\text{vac}}\rangle$ is supposed not to change, so that $U|\psi_{\text{vac}}\rangle = |\psi_{\text{vac}}\rangle$. A single-photon input $a^{\dagger}|\psi_{\text{vac}}\rangle$ changes into a linear combination of single-photon output states $\lambda_1 b_1^{\dagger}|\psi_{\text{vac}}\rangle + \lambda_2 b_2^{\dagger}|\psi_{\text{vac}}\rangle$. Here λ_1 and λ_2 are the complex amplitudes for the outputs with unit input, $a^{\dagger}$ is the creation operator for a photon in the input wave-packet, and $b_1^{\dagger}$ and $b_2^{\dagger}$ are creation operators for photons in the output wave-packets. We assume $[b_i, b_j^{\dagger}] = \delta_{ij}$ and $[b_i, b_j] = 0$, and because of normalisation we find $|\lambda_1^2| + |\lambda_2^2| = 1$. (Naturally we expect $|\lambda_1|^2$ to be very small.) More generally we set $Ua^{\dagger}U^{\dagger} = \lambda_1 b_1^{\dagger} + \lambda_2 b_2^{\dagger}$, so that an incident l-photon wave-packet $|\psi_l\rangle = (l!)^{-1/2}(a^{\dagger})^l|\psi_{\text{vac}}\rangle$ will evolve into

$$U|\psi_l\rangle = (l!)^{-1/2}(Ua^{\dagger}U^{\dagger})^l U|\psi_{\text{vac}}\rangle = (l!)^{-1/2}(\lambda_1 b_1^{\dagger} + \lambda_2 b_2^{\dagger})^l|\psi_{\text{vac}}\rangle.$$

The transmitted power is usually high enough to let us regard the input signal classically, with a well-defined complex amplitude. This is modelled by allowing the input state to be a 'coherent' state (Glauber 1963)

$$|\alpha\rangle = \exp(-\tfrac{1}{2}|\alpha|^2) \sum_l (l!)^{-1/2}\alpha^l|\psi_l\rangle = K(a^{\dagger}, \alpha)|\psi_{\text{vac}}\rangle, \tag{5}$$

where $K(a^{\dagger}, \alpha)$ stands for the operator $\exp(\alpha a^{\dagger} - \frac{1}{2}|\alpha|^2)$. (Note that $|0\rangle = |\psi_0\rangle = |\psi_{\text{vac}}\rangle$.) Here α is the complex amplitude, which will be referred to as the 'α-value' of the coherent state. The initial density matrix is then $\rho_{\text{i}} = |\alpha\rangle\langle\alpha|$, which evolves into the final density matrix $\rho_{\text{f}} = U|\alpha\rangle\langle\alpha|U^{\dagger}$. The state $U|\alpha\rangle$ is obtained by replacing $a^{\dagger}$ by $Ua^{\dagger}U^{\dagger}$, and so by (5) we find

$$U|\alpha\rangle = K(Ua^{\dagger}U^{\dagger}, \alpha)|\psi_{\text{vac}}\rangle = K(b_1^{\dagger}, \lambda_1\alpha)K(b_2^{\dagger}, \lambda_2\alpha)|\psi_{\text{vac}}\rangle$$

by factorising the exponential $K(Ua^{\dagger}U^{\dagger}, \alpha)$. Thus $U|\alpha\rangle$ can be regarded as a direct product of coherent states on the outputs, and so ρ_{f} factorises into a direct product of density matrices. This factorisation means that measurements made on the two output channels will give independent results; this fact will be used in the next section. In the present context the second output is not used, and so we take a partial trace over the states of the second output. Hence its density matrix is replaced by unity, and we are left with a reduced density matrix corresponding to the first output being in a coherent state with α-value $\lambda_1\alpha$. Thus the receiver may be put into any specified coherent state.

In a photon-counting system we are interested in the probability distribution of the number of photons. For a coherent state with α-value α the probability of finding m photons is easily shown to be the Poisson distribution

$$p_m = (m!)^{-1}\bar{N}^m \exp(-\bar{N}), \tag{6}$$

where $\bar{N} = |\alpha|^2$ is the mean number of photons.

In most cases the absolute phase of the received carrier has no significance, as the distance of the transmitter is not determined to within a fraction of the wavelength of the carrier. However, the relative phases of the various normal modes need not be lost. The overall receiver density matrix is an outer product of BT density matrices, one for each normal mode, if the normal modes are used in a statistically independent manner, and we should then average this product over all phases of λ_1. This is very different from averaging each density matrix over the phase of λ_1, which then diagonalises it in the number representation, so that the system counts the photons in each normal mode. Probably only a few of the BT normal modes need be given over to providing a phase reference for the carrier, so that in principle the information rate hardly falls at all. However, although of interest, this problem is not strictly relevant to the present discussion.

This is a convenient place to summarise some required properties of the coherent states (Helstrom 1976). The expectation value of the photon number is given by

$$\langle\alpha|a^\dagger a|\alpha\rangle = |\alpha|^2. \tag{7}$$

The overlap integrals are given by

$$\langle\alpha|\beta\rangle = \exp[\mathrm{i}\,\mathrm{Im}(\alpha^*\beta)]\exp(-\tfrac{1}{2}|\alpha-\beta|^2). \tag{8}$$

A unitary displacement operator $D(\gamma)$ with γ complex can be defined by

$$D(\gamma) = \exp(\gamma a^\dagger - \gamma^* a)$$

so that $|\gamma\rangle = D(\gamma)|0\rangle$. This operator has the composition rule

$$D(\gamma)D(\delta) = D(\gamma+\delta)\exp[-\mathrm{i}\,\mathrm{Im}(\gamma^*\delta)]. \tag{9}$$

4. The '*x–p*' detector

A simple synchronous receiver was described by Takahasi (1965). (See also Arthurs and Kelly (1965).) The signal is evenly split by a beam splitter and used to excite two oscillators. In one oscillator we measure the 'x coordinate' (the hermitean operator $2^{-1/2}(a+a^\dagger)$), and in the other the 'p coordinate' ($2^{-1/2}(a-a^\dagger)/\mathrm{i}$). With a received signal represented by the coherent state $|\lambda\rangle$, with λ equal to $\lambda'+\mathrm{i}\lambda''$, the oscillators are each put into a coherent state with an α-value $2^{-1/2}\lambda$. Moreover, for reasons given in § 3 the results of the measurements are independent. (This is true even if we measure the photon numbers of the two outputs; these results would not be independent if the input was an eigenstate of the photon number, with a definite number of photons.) In the Schrödinger representation (with $p = -\mathrm{i}\,\mathrm{d}/\mathrm{d}x$) the wavefunction satisfies $a\psi =$

$2^{-1/2}(x+\mathrm{d}/\mathrm{d}x)\psi=(2^{-1/2}\lambda)\psi$, so that $|\psi|^2$ is proportional to $\exp[-(x-\lambda')^2]$. Since $\langle x\rangle=\lambda'$, we may imagine that a measurement of x is directly a measurement of λ', so that a detected value λ'_r comes up with a probability density proportional to $\exp[-(\lambda'_r-\lambda')^2]$. Similarly the p measurement determines λ'', and overall in the complex α plane the probability density of detecting α_r when α comes in is just $(\pi R^2)^{-1}\exp(-|\alpha_r-\alpha|^2/R^2)$, with $R=R_N=1$. It is as though there is Gaussian noise. It is well known (Yu 1976) that, subject to the restriction that the average number of photons is $\bar{n}$, the maximum information is obtained by choosing a similar Gaussian probability density $(\pi R^2)^{-1}\exp(-|\alpha|^2/R^2)$ for the values of α coming in, with $R=R_S=\bar{n}^{1/2}$. The information is then $\ln[(R_N^2+R_S^2)/R_N^2]=\ln(1+\bar{n})$. This is the information sent every T seconds by each of BT normal modes, and hence it is the value of C/B, where C is the rate and B the bandwidth.

For large values of $\bar{n}$ this gives $C/B=\ln(\bar{n})+O(\bar{n}^{-1})$, which compares favourably with the upper bound (3). For very small values of $\bar{n}$ it is not so good because $C/B\approx\bar{n}$, much less than the right-hand side of (4). It is advantageous at low levels to use the null signal $\alpha=0$ with high probability, and this condition is not detected faithfully, that is, there are 'false alarms'.

5. 'Lattice' system for large $\bar{n}$

The coherent states $|\alpha\rangle$ may be represented by points (x, y) in the xy plane, with $\alpha=x+\mathrm{i}y$. Let us suppose (for a given normal mode) that we use a system where the possible receiver states are coherent states represented by points r_i forming a square lattice with squares of area Δ. Suppose also (for the moment) that the corresponding coherent states are orthogonal, so that in principle at least there is a quantum observable whose measurement can distinguish these states with certainty (Dirac 1958). Then if the ith point is used with probability p_i, we have to maximise the entropy $-\Sigma p_i \ln p_i$ subject to the conditions $\Sigma p_i=1$, $\Sigma p_i n_i=\bar{n}$, where $n_i(=r_i^2)$ is the expectation value of the number of photons for this point. We find that $p_i\propto\exp(-\lambda-\mu n_i)=\exp(-\lambda-\mu r_i^2)$, where λ and μ are Lagrange multipliers chosen to satisfy the constraints. In the case when $\bar{n}\gg\Delta$ the lattice points are in effect closely spaced, so that the sums can be replaced by integrals. If we put $p_i=(\pi R^2)^{-1}\exp(-r_i^2/R^2)\Delta$ with R to be determined, we find that Σp_i becomes

$$(\pi R^2)^{-1}\iint \exp(-r^2/R^2)\,\mathrm{d}^2r=1.$$

Similarly $\Sigma n_i p_i$ becomes

$$(\pi R^2)^{-1}\iint r^2\exp(-r^2/R^2)\,\mathrm{d}^2r=R^2,$$

so that $R=\bar{n}^{1/2}$. A similar evaluation gives for the information (Shannon 1948)

$$-\sum p_i \ln p_i \approx \ln\bar{n}+1-\ln(\Delta/\pi), \qquad (10)$$

which is then also the value of C/B, where C is the rate. At least the moderately dominant $\ln\bar{n}$ behaviour in (3) is also found here. The problem is choosing a value for Δ. The coherent states are not orthogonal, but because of the Gaussian behaviour of the overlap integrals they should be effectively orthogonal for Δ large enough.

However, with this periodic arrangement of the points representing the coherent states we can set up a genuinely orthogonal set of states, by the use of techniques from solid state band theory (Ziman 1964). Any lattice point in the square lattice may be written as $(a, b)\Delta^{1/2}$, where a and b are integers, and so if we define the unitary displacement operators

$$\tau(a, b) = D[(a + \mathrm{i}b)\Delta^{1/2}],$$

we find from (9) that

$$\tau(a, b)\tau(a', b') = \tau(a + a', b + b') \exp[-\mathrm{i}(ab' - ba')\Delta]. \tag{11}$$

Because of the phase factors, these operators do not in general form an ordinary Abelian translation group, but instead a 'magnetic translation group' of the sort studied in the 1960s in connection with the motion of charged particles in a magnetic field and in a periodic lattice (Brown 1964, Zak 1964). If Δ is given a value $\pi\lambda/\mu$, where λ and μ are integers with no common factors, then it is straightforward to show that $\tau(1, 0)$ and $\tau(0, \mu)$ commute, and thus can be used as a basis for a superlattice. We have a choice for the value of Δ, and in this paper we shall choose the simplest cases, $\mu = 1$ with $\lambda = 1$ or 2. For λ even, the phase factors in (11) are always unity and we have the usual situation. For λ odd, the phase factor can be either 1 or -1.

The basic method is as follows. We use the translation group to produce from the coherent states a set of automatically orthogonal Bloch functions. These are then normalised, and used to produce a set of automatically orthonormal localised states (Wannier functions). It is convenient to define $e(x) = \exp(2\pi\mathrm{i}x)$, and it is worth noting that $e(\frac{1}{2}\lambda) = e(-\frac{1}{2}\lambda)$ for integral λ. We also impose periodic boundary conditions over an enormous square of $N \times N$ lattice points.

We define the Bloch states

$$|\psi_{pq}\rangle = N^{-1} \sum_{ab} e(pa + qb)e(-\tfrac{1}{2}\lambda ab)\tau(a, b)|0\rangle,$$

where p and q are integers divided by N, and $|0\rangle$ is the coherent state with α-value zero. The vector $|\psi_{pq}\rangle$ is periodic in p and q with period 1, and so we confine p and q to a 1×1 Brillouin zone. It is not hard to show that

$$\tau(a', 0)|\psi_{pq}\rangle = e(-pa')|\psi_{pq}\rangle, \qquad \tau(0, b')|\psi_{pq}\rangle = e(-qb')|\psi_{pq}\rangle,$$

and thus for $p \neq p'$ or $q \neq q'$ we find that $|\psi_{pq}\rangle$ and $|\psi_{p'q'}\rangle$ are orthogonal as they are eigenvectors of unitary operators with different eigenvalues. It is also worth noting that $|\psi_{pq}\rangle$ is simply a linear combination of rephased coherent states $|\phi_{ab}\rangle = e(-\frac{1}{2}\lambda ab)\tau(a, b)|0\rangle$ on the lattice. We define the norm $\gamma_{pq} = \langle\psi_{pq}|\psi_{pq}\rangle$ which must be real and positive. Thus we find

$$\gamma_{pq} = \sum_{AB} e(pA + qB)e(-\tfrac{1}{2}\lambda AB)M_{AB},$$

where $M_{AB} = \langle 0|\tau(A, B)|0\rangle = \exp[-\frac{1}{2}\Delta(A^2 + B^2)]$ by (8). We may therefore divide $|\psi_{pq}\rangle$ by $\gamma_{pq}^{1/2}$ to produce an orthonormal set of Bloch waves, and from these the Wannier functions

$$|w_{ab}\rangle = N^{-1} \sum_{pq} e(-ap - bq)\gamma_{pq}^{-1/2}|\psi_{pq}\rangle$$

which are then orthonormal. Finally we need the overlap integrals with the coherent states

$$\langle w_{ab}|\phi_{a'b'}\rangle = N^{-2}\sum_{pq} e[p(a-a')+q(b-b')]\gamma_{pq}^{1/2}$$

which are thus translationally invariant. We set

$$J(\boldsymbol{r}_j-\boldsymbol{r}_i)=|\langle w_{ab}|\phi_{a'b'}\rangle|^2,$$

where $\boldsymbol{r}_j=(a,b)\Delta^{1/2}$ and $\boldsymbol{r}_i=(a',b')\Delta^{1/2}$. It is worth mentioning that the Bloch states can be rephased, so that we can include an arbitrary function $\exp(\mathrm{i}\theta_{pq})$ in the above summation. However, it can be shown that the choice $\theta_{pq}=0$ minimises the sum

$$\sum_{ab}(a^2+b^2)|\langle w_{ab}|0\rangle|^2 \qquad \text{or} \sum_j r_j^2 J(\boldsymbol{r}_j).$$

The operation of the link is as follows. The transmitter puts the receiver into a coherent state represented by $\boldsymbol{r}_i$. The detector then makes a measurement for which the Wannier functions form an orthonormal basis, and so returns a result $\boldsymbol{r}_j$ with probability $J(\boldsymbol{r}_j-\boldsymbol{r}_i)$ (Dirac 1958). The Shannon (1948) formula then gives

$$C/B=-\sum_j q_j \ln q_j+\sum_i p_i\sum_j q(j|i)\ln q(j|i), \tag{12a}$$

with

$$q_j=\sum_i q(j|i)p_i, \tag{12b}$$

where p_i is the probability of the point $\boldsymbol{r}_i$ being used, $q(j|i)=J(\boldsymbol{r}_j-\boldsymbol{r}_i)$ is the probability of the point $\boldsymbol{r}_j$ being received given that $\boldsymbol{r}_i$ was actually used, and q_j is the probability of the point $\boldsymbol{r}_j$ being received. We take p_i as before with a Gaussian distribution of radius $\bar{n}^{1/2}$. If the range of the overlap $J(\boldsymbol{r}_j-\boldsymbol{r}_i)$ is very much less than this, we may approximate q_j by p_j, which in fact reduces the first term and so lowers the result. We obtain by (10)

$$C/B\approx \ln\bar{n}+1-\ln(\Delta/\pi)-E,$$

where E is an 'equivocation' given by the second term in (12a), that is,

$$E=-\sum_j J(\boldsymbol{r}_j)\ln J(\boldsymbol{r}_j).$$

The obvious value of λ to choose is $\lambda=1$, which gives $\Delta=\pi$. Unfortunately this leads to trouble as it seems that γ_{pq} vanishes for $p=q=\frac{1}{2}$, or very nearly. (This is equivalent to the result

$$\sum_{ab}(-1)^{a+b+ab}\exp[-\tfrac{1}{2}\pi(a^2+b^2)]\approx 0$$

which is easily verified numerically. However, the author has no proof that it is zero!) If γ_{pq} does vanish, then the coherent states must be linearly dependent. A similar result is found in the hexagonal lattice, and it seems that the coherent states have been placed too close.

The next value to try is the rather tame case $\lambda=2$, so that the phase factors in (11) can be quietly forgotten. An even greater simplification takes place as the two-dimensional Fourier transforms factorise into products of one-dimensional transforms, and so the computing is trivial. This time $\Delta = 2\pi$ and the result is

$$C/B \approx \ln \bar{n} + 1 - \ln 2 - 0{\cdot}01621 = \ln \bar{n} + 0{\cdot}2906.$$

This is a lower bound which could probably be improved. Thus a hexagonal lattice may well be better. We might also try values of λ/μ between 1 and 2, say $\frac{3}{2}$. (A very rough calculation suggests that it may be possible to increase C/B by about 0·2.) Unfortunately the theory is more complicated, involving eigenvalues and eigenvectors of $\mu \times \mu$ matrices, and the phasing problem briefly mentioned above becomes more serious.

6. Photon-counting link for large $\bar{n}$

This brief section is put in mostly for comparison with the last. As we might expect, the rate is about half that of the last section, that is, $C/B \approx \frac{1}{2} \ln \bar{n}$, since the information about the phase of the carrier for each normal mode is lost, leaving only the information about the magnitude. An approximate answer for $\bar{n} \gg 1$ can be derived as follows.

If the ith transmitted symbol is arranged to give an average reception of n_i photons, then by (6) the probability of receiving exactly j photons is $Q_j = (j!)^{-1} n_i^j \exp(-n_i)$, and the probability δQ of receiving a count in a range δj about j is $\delta Q \approx Q_j \delta j$ (with $1 \ll \delta j \ll n_i^{1/2}$). We set $x_r = j^{1/2}$, $x = n_i^{1/2}$. By applying Stirling's formula and a second-order Taylor expansion to $\ln Q_j$ and by putting $\delta j \approx 2x_r \delta x_r$, we find

$$\delta Q \approx (2\pi)^{-1/2} 2 \exp[-2(x_r - x)^2] \delta x_r.$$

Thus we have a situation similar to that in § 4, with Gaussian noise, except that (a) x_r and x are real, not complex, (b) x_r and x are restricted to positive values, and (c) the value of R_N^2 is $\frac{1}{2}$. The first problem can be circumvented by artificially pairing the normal modes (so that $R_S^2 = 2\bar{n}$), and the second by allowing x and x_r to have either sign and then throwing away the information conveyed by the sign. In consequence the information per pair is about $\ln[(R_N^2 + R_S^2)/R_N^2] \approx \ln(4\bar{n})$, so that the final answer is roughly $\frac{1}{2}\ln(4\bar{n}) - \ln 2 = \frac{1}{2} \ln \bar{n}$.

7. Links with small $\bar{n}$

For $\bar{n} \ll 1$ we first consider a 'photon-detecting' link against which to compare a synchronous link. We use two transmitted symbols $i = 0$ and 1 say, with $n_0 = 0$ (a null) and $n_1 = \bar{n}/\beta$, where the parameter β is the probability that $i = 1$, so that $p_0 = 1 - \beta$. (Naturally n_1 is greater than or equal to $\bar{n}$.) The receiver gives two outputs, $j = 0$ for no photon received, and $j = 1$ for one or more photons received. A photon detector is not quite as good as a photon counter, but for small values of $\bar{n}$ it is hard enough to receive one photon, let alone two or more, so there is scarcely any loss of performance. The conditional probabilities $q(j|i)$ are by the Poisson distribution (6) $q(0|0) = 1$, $q(1|0) = 0$, $q(0|1) = \exp(-n_1)$, $q(1|1) = 1 - \exp(-n_1)$. We then maximise (12) with respect to β, keeping $\bar{n}$ fixed.

For a synchronous link we imagine that we can put the receiver into one of $N+1$ coherent states $|\alpha_c\rangle$, $|\alpha_0\rangle, \ldots, |\alpha_{N-1}\rangle$ with $\alpha_c = 0$, $\alpha_k = n_1^{1/2} \exp(2\pi i k/N)$. We assign probabilities $\bar{n}/Nn_1$ to the outer points and $1-\bar{n}/n_1$ to the central point $\alpha_c = 0$. As in § 5 we take orthogonal linear combinations of these states and assume that there is a measurement process corresponding to this basis. The new centre state $|b_c\rangle$ may be chosen as $|\alpha_c\rangle$ or the linear combination $|\alpha_c\rangle + N^{-1}\lambda\ \Sigma_j\ |\alpha_j\rangle$, where λ is an arbitrary complex number. (This mixing does not gain much over $\lambda = 0$ and is hardly worth the bother.) Then we orthogonalise the states $|\alpha_j\rangle$ of the outer ring to $|b_c\rangle$ by setting $|b_j\rangle = |\alpha_j\rangle - \mu|b_c\rangle$ and choosing μ so that $\langle b_j|b_c\rangle = 0$ for $j = 0, \ldots, N-1$. Then we may use the $|b_j\rangle$ to set up automatically orthogonal Bloch functions round the ring, which we normalise and use to form localised Wannier functions which are also automatically orthogonal. The results are as follows (with the numerical subscripts evaluated mod N). The overlap integrals are obtained from (8) and are given by

$$I_j = \langle\alpha_{k+j}|\alpha_k\rangle = \exp[-2n_1 \sin^2(\pi j/N)] \exp[-\mathrm{i}n_1 \sin(2\pi j/N)],$$

$$I_{\mathrm{oc}} = \langle\alpha_k|\alpha_c\rangle = \exp(-\tfrac{1}{2}n_1), \qquad I_{\mathrm{cc}} = \langle\alpha_c|\alpha_c\rangle = 1.$$

The norm $|\eta|^2$ of $|b_c\rangle$ is given by

$$|\eta|^2 = I_{\mathrm{cc}} + \lambda^* I_{\mathrm{oc}} + \lambda I^*_{\mathrm{oc}} + |\lambda|^2 \sum_j I_j/N.$$

The quantity μ is given by $\mu|\eta|^2 = I^*_{\mathrm{oc}} + \lambda^*\ \Sigma_j\ I_j/N$. The quantities K_k and γ_j are defined by

$$K_k = \Big(\sum_j e(-kj/N)I_j\Big) - N|\mu|^2|\eta|^2\delta_{k0}$$

and

$$\gamma_j = N^{-1} \sum_k e(kj/N) K_k^{1/2}.$$

We obtain $q(j|i) = |\gamma_{j-i}|^2$, $q(\mathrm{c}|i) = |\mu|^2|\eta|^2$, $q(j|\mathrm{c}) = |\lambda|^2 K_0/N^2$, $q(\mathrm{c}|\mathrm{c}) = |\eta|^2|1-\lambda\mu|^2$ for the conditional probabilities for j or c given i or c. These results are then used in the Shannon formula (12) with $p_i = \bar{n}/Nn_1$ and $p_c = 1 - \bar{n}/n_1$.

The upper bound for $\bar{n} \ll 1$ is given approximately by (4), so that $C_{\max}$ tends to zero like $-\bar{n}\ \ln(\bar{n})$ as $\bar{n}$ tends to zero, and so do the other values for the rates. Therefore to illustrate the results we have plotted the ratios $C/C_{\max}$ against $\bar{n}$ on a log–linear plot (figure 1). For very small values of $\bar{n}$ ($\approx 10^{-6}$) the rate for the photon detector is up to 75% of $C_{\max}$, and in fact it can be shown that the ratio $C/C_{\max}$ tends to unity very slowly as $\bar{n} \to 0$, an interesting result perhaps, but not of much practical interest. The other curve shows the corresponding result for the synchronous link with $N = 4$, and evidently it is not much better, especially for $\bar{n}$ small. (Here the maximisation was carried out with respect to n_1 and λ.) Other choices for N, the number of outer points on the 'star', give very similar results, and in fact it seems that for very small values of $\bar{n}$ lower values of N are slightly better, although the difference would not show on the graph.

8. Concluding remarks

We have discussed some links which have rates quite close to the theoretical upper bound, at least in the limits of low and high photon rates. At high rates the 'x–p' detector works well, and at low rates the 'photon detector'. It is also evident how in

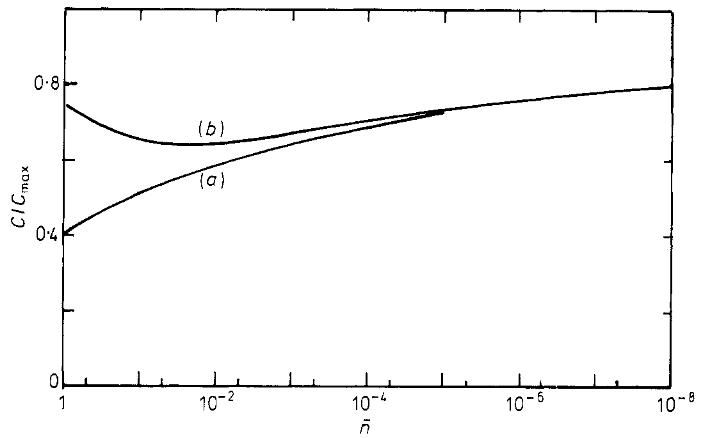

Figure 1. Information rates divided by the upper bound C_{max} (equation (2)) against $\bar{n}$ (equation (1)) at small values of $\bar{n}$, (*a*) for a 'photon-detecting' system, (*b*) for the 'star' coherent system with $N = 4$. Both these systems are described in § 7.

principle such receivers might be constructed. (A description of a complete 'x–p' receiver is given by Takahasi (1965).) We have postulated systems that can do slightly better, but there is no obvious way in which they might be constructed. How close are they to the optimal for a free-space link? The very difficulty of making any improvement seems to suggest that they are fairly close. It might be worth emphasising that the ratios of the rates to the maximum approach unity for both large and small $\bar{n}$.

The results for low photon rates ($\bar{n} \ll 1$) seem to be the most interesting. The author at least is surprised by how good the photon detector is, and by how little better the coherent 'star' system is. This is in contrast with what happens for large $\bar{n}$. The fact that in a coherent system the transmitter can put the receiver into any chosen 'coherent' state suggests that there should be complete control over what is received. An analogy based on the 'spinning-particle' system of § 2 shows that this need not be so. For simplicity let us choose spin-$\frac{1}{2}$ particles. If the transmitter can send a particle with any given direction of spin, then, by choosing the orthogonal spin-up and spin-down states, a noise-free link is established. But if the transmitter is restricted to sending particles with the spin direction restricted to lie within a certain small angle from the z axis, then, although the transmitter may have perfect control over the spin state of each particle sent, the receiver will suffer from quantum uncertainties. (Such a restriction would arise if there was an energy constraint, and if the particles possessed a spin magnetic moment and were in a magnetic field parallel to the z axis.)

Acknowledgments

The author would like to thank the University of London and Westfield College Computing Centres for the use of their facilities, and the Numerical Algorithms Group for the use of the minimising routines.

Note added in proof. Since the submission of this manuscript Professor J Brown (Imperial College, University of London) has proved that the Gaussian sum in § 5 which was conjectured to vanish does indeed do so (private communication).

References

Arthurs E and Kelly J L 1965 *Bell Syst. Tech. J.* **44** 725–9
Brown E 1964 *Phys. Rev.* **133** A1038–44
Dirac P A M 1958 *Principles of Quantum Mechanics* 4th edn (Oxford: Clarendon)
Glauber R J 1963 *Phys. Rev.* **131** 2766–88
Gordon J P 1962 *Proc. IRE* **50** 1898–908
Helstrom C W 1976 *Quantum Estimation and Detection Theory* (New York: Academic)
Kubo R 1971 *Statistical Mechanics* (Amsterdam: North-Holland)
Shannon C 1948 *Bell Syst. Tech. J.* **27** 379–423, 623–55
Takahasi H 1965 in *Advances in Communication Systems* vol 1, ed A Balakrishnan (New York: Academic) pp 227–310
Yu F T S 1976 *Optics and Information Theory* (New York: Wiley)
Zak J 1964 *Phys. Rev.* **134** A1602–6, A1607–11
Ziman J M 1964 *Principles of the Theory of Solids* (Cambridge: UP) p 148