



Information degradation by the gravitational field

M.Schiffer*

CERN, CH-1211 Geneva 23, Switzerland

November 29, 1991

Abstract

In this paper we show that the gravitational field unavoidably degrades information. We quantify information by means of Shannon's entropy, and considered information carriers that are quanta of some field. Next we derive the appropriate 'channel capacity' formula, which quantifies the maximum amount of information that can be transmitted per pulse in a cosmological background. There are two main consequences of this formula. First, if *any* departure from the Friedman-Robertson-Walker (FRW) background (no matter how small) occurred close to the singularity, then all information concerning the first instants of the Universe was completely washed out by the gravitational 'noise'. In other words, no imprints of the initial conditions of the Universe are left over! A second consequence is that the degradation of information occurs in the same direction as anisotropies (clumsiness) emerge in the Universe.

*Supported by a John Stewart Bell Scholarship

I. Introduction

When Hubble established the recession of galaxies and the expansion of our cosmos, at the end of the 1920's, the typical red shift of the galaxies he was observing was of order of one part in ten. Nearly sixty years passed since then, and it is already possible, by now, to observe radio galaxies with red shift of 3.8[1]; QSO's at a red shift over 4.43 [2] were detected. Furthermore, with the advent of modern cosmology, it has been possible to predict the correct cosmic microwave background temperature, and the correct abundance of light elements, corresponding to the observation of the childhood of the Universe, when it was only 10^{-2} s old. These achievements have the effect of boosting our curiosity and lead us to a further development of theories/models which could explain the first instants of our Universe, and also of the means to observe that far.

Given the importance of the subject, one is naturally led to ask the question: Do the known laws of physics limit our ability to observe events in the Universe? Causality imposes a first constraint, in that one cannot observe events outside our particle horizon. But is that all? In this paper we show that quantum mechanics further constraints the amount of information that can be gathered by any detection device. In the framework of information theory we shall show that the degradation of information is a consequence of the cosmological expansion. Furthermore, we also show that all information concerning the Universe initial conditions was completely erased by the gravitational noise.

Physically the situation here is identical to that of a pendulum which would be harnessed as a data storage device. Shannon's information quantifies the maximum information that could be encoded in the pendulum using all its microscopic degrees of freedom [3]:

$$H = - \sum_n p_n \log_2 p_n \text{ bits} , \quad (1)$$

where p_n are the various probabilities of finding the pendulum in one of its various (quantum) states. Now, if its string is suddenly pulled, this quantity no longer represents the information that can be recovered after the string is pulled. The act of pulling induces transitions among the various states, and $p_n \rightarrow p'_n$, causing a degradation of information: we are submitting the pendulum to *quantum noise*. Similarly, during the cosmological expansion, particles are created, corresponding to transitions among the various states of a field whose quanta are used as information carriers. Accordingly, the gravitational field produces a quantum noise that degrades the information which was originally emitted by some distant object, say, the spectral composition of a gas cloud in a distant QSO.

Noise has the effect that the received signal differs in a stochastic way from the transmitted one. This limits the receiver's ability to recover the encoded information. In effect, the received signal is associated with larger entropy than the transmitted one, because the noise has introduced a further measure of uncertainty. However, as proved by Shannon [3, 4], it is in principle still possible to encode the information at the transmitter in such a way that it can all be recovered at the receiver – *provided the information is not transmitted too fast*. In effect, every communication channel is ascribed a *capacity*, which represents the maximum rate, in bits s^{-1} , at which information can be transmitted through it with negligible probability of error. Should one try to surpass this transmission rate, then the difference $I - I_{\max}$ represents information that will inevitably be degraded

by errors due to noise, and that will never be recovered. In informational jargon, we would define our communication channel as consisting of the gravitational background plus the field quanta which are the information carriers.

When one performs photometric measurements in a telescope, filters of given (narrow) frequency range are placed before the photon counter. Then one is tempted to recall Shannon's celebrated capacity formula [3, 4], which quantifies the maximum amount of information (per unit time) which could, in principle, be transmitted between the source and the receiver:

$$I_{\max} = \frac{\Delta\omega}{2\pi} \log_2 \left(\frac{N + P}{N} \right). \quad (2)$$

Here P and N stand, respectively, for the signal and noise powers and $\Delta\omega$ for the (narrow) range of frequencies of photons that reach the photon counter.

This formula was derived under the assumptions that noise and signal: i. are statistically independent; ii. both have a Gaussian distribution, i.e. they are classical. Unfortunately, none of these conditions are met in the cosmological problem. There are many important results in the literature about quantum capacity formulas, both for noisy and noiseless communication [12, 17], but none of them deals with the sort of quantum noise peculiar to the gravitational field. Thus, if we wish to address this problem in an appropriate manner, we ought to derive a quantum capacity formula for information transmission in a gravitational background. This is one of the purposes of this paper. To this end, we briefly review in Section II the appropriate tools of information theory we need. The interested reader is also referred to the standard monograph in information theory by Brillouin [6]. In Section III, we develop the relevant relations for noisy communication. Then, in Section IV we derive the (quantum) gravitational noise and, finally, the quantum channel capacity formula for the cosmological problem. Section V is devoted to the discussion of this equation, while in Section VI we summarize our results, prospects, etc. Lengthy calculations which could have disrupted the main argument were displayed in two appendices.

II. Information theory tools

How can the information that reaches the telescope photon counter be quantified? Should it be given by eq. (1) through the probabilities $p_o(m)$ of detecting m quanta in the telescope or by the probabilities $p_i(n)$ that n quanta were emitted? As we already mentioned, in the presence of (quantum) gravitational noise, the received signal is associated with an entropy larger than the one that was emitted, because the noise has introduced a further measure of uncertainty. None the less, this larger entropy does not correspond to a larger amount of useful information, meaning the one which bears information concerning the object we are observing, spectral lines, red shifts, etc. Thus, the answer is that neither is correct! The latter would have quantified the information that could be borne by the signal, but this information has since been adulterated by noise, whilst in the former the number of quanta m is partly a result of noise.

The procedure for dealing with this situation was outlined by Shannon [4]. There is a *joint* probability distribution $p_{o,i}(m, n)$ for input and output numbers of quanta which supplies a complete statistical description of the (noisy) communication system. From

it we can compute the two *marginal* probability distributions mentioned above, one, $p_i(n)$, by summing out m , and the second one, $p_o(m)$, by summing out n , as well as two *conditional* distributions. One,

$$p(m|n) \equiv \frac{p_{o,i}(m,n)}{p_i(n)}, \quad (3)$$

stands for the probability of m quanta in the detector, given that m were sent. The second,

$$q(n|m) \equiv \frac{p_{o,i}(m,n)}{p_o(m)}, \quad (4)$$

gives the probability that n quanta were sent, given that the detector registered m .

There is an entropy for each one of these distributions. The generic definitions are

$$H_{o|i} \equiv - \sum_{n,m} p_{o,i}(n,m) \ln p(m|n) \quad (5)$$

and

$$H_{i|o} \equiv - \sum_{n,m} p_{o,i}(m,n) \ln q(n|m). \quad (6)$$

The following identities [4] are easily verified:

$$H_{o,i} \equiv H_i \pm H_{o|i} = H_o \pm H_{i|o}. \quad (7)$$

Shannon noted that $H_{i|o}$, the conditional entropy of the input *when the output is known*, must represent the extra uncertainty introduced by the noise which hinders reconstruction of the initial signal even when the output is known. Thus, he interpreted

$$I \equiv H_i - H_{i|o} \quad (8)$$

to be the useful information, meaning the one which we could, in principle, recover from the output signal – even in the face of noise. We can also regard $H_{o|i}$ as the uncertainty in the output *for given input*, as the effect of the noise. Therefore, it should be subtracted from the full entropy of the output H_o in order to get the information borne by the signal. Indeed, with the help of the identity eq. (7), the above definition can be alternatively expressed as

$$I \equiv H_o - H_{o|i}. \quad (9)$$

III. Noisy communication: the optimal transmission rate

During the observation of, say, a distant galaxy in a telescope, the mean number of quanta that reach the telescope is known: it is related to the number of pulses in the photon counter. Thus, in order to find the maximum amount of information that could be transmitted, we need to find the maximum of I , i.e. we should vary the above quantity

by introducing two Lagrange multipliers: α to enforce normalization of the probabilities and β , the knowledge of $\langle m \rangle$. Namely

$$I' = - \sum_m \sum_n p(m|n)p_i(n) \ln p_o(m) \pm \sum_m \sum_n p(m|n)p_i(n) \ln p(m|n) - \sum_m \sum_n p(m|n)p_i(n) (\alpha - 1 \pm m\beta). \quad (10)$$

For convenience, we introduced the Lagrange multiplier $\alpha - 1$ instead of α and expressed $p_o(m) = \sum_n p(m|n)p_i(n)$. Variation of this quantity yields,

$$\sum_n \left\{ \sum_m [\ln p_o(m) \pm \beta m \pm \alpha - \ln p(m|n)] p(m|n) \right\} \delta p_i(n) = 0. \quad (11)$$

In order to solve this equation, we try the ansatz

$$p_o(m) = e^{-(\alpha + \beta m + B(m))}. \quad (12)$$

Inserting eq. (12) into eq. (11), we obtain an equation for $B(m)$,

$$\sum_m B(m)p(m|n) = - \sum_m p(m|n) \ln p(m|n). \quad (13)$$

What is the physical meaning of $B(m)$? In the absence of noise, it would have vanished. Thus, it should somehow quantify the amount of noise. As a matter of fact, if we multiply eq. (13) by $p_i(n)$ and sum over n , we obtain that,

$$\langle B \rangle = \sum_m B(m)p_o(m) = - \sum_{m,n} p(m,n) \ln p(m|n) = H_{o|i}. \quad (14)$$

This equation tells us that the mean value $\langle B(m) \rangle$ quantifies the amount of noise.

Substituting our ansatz (12) into eq. (9), we obtain the maximum amount of information that can be transmitted in the presence of noise,

$$I_{\max} = \sum_m p_o(m)(B(m) \pm \alpha \pm \beta m) \pm \sum_m \sum_n p_{o,i}(m,n) \ln p(m|n). \quad (15)$$

Now, by virtue of eq. (14), the first term and the double sum cancel out, and the above expression reads simply as

$$I_{\max} = \alpha \pm \beta \langle m \rangle. \quad (16)$$

In the above equation, α and β should be determined by normalization and mean number of quanta conditions. Namely,

$$\alpha = - \ln \sum_m e^{-(\beta m + B(m))} \quad (17)$$

and,

$$\langle m \rangle = - \frac{\partial \alpha}{\partial \beta}. \quad (18)$$

Therefore, once the noise is specified through $p(m|n)$, finding I_{\max} (the maximal amount of information that could be conveyed by a pulse containing $\langle m \rangle$ quanta of given frequency) has been reduced to the problem of solving eq. (13) for B .

Our intuition would say that as $\langle B(m) \rangle$ grows, less and less information can be conveyed. Indeed if we take the variation of I_{\max} [eq. (16)] with respect to $B(m)$,

$$\delta I_{\max} = \delta \alpha \pm \langle m \rangle \delta \beta \quad (19)$$

and combine this with the variation of α (eq. (17)) with respect to $B(m)$,

$$\delta \alpha = -\langle m \rangle \delta \beta - e^{-\alpha} \sum_m \delta B(m) e^{-(\beta m + B(m))}, \quad (20)$$

we obtain

$$\delta I_{\max} = -\langle \delta B(m) \rangle. \quad (21)$$

in accordance with what we suspected.

A point I would like to underline is that, contrary to (my) naive expectations, the variation of I with respect to $p_o(m)$ is not completely equivalent to the variation with respect to $p_i(n)$. Some extra information can be gathered in this process. To see this, let $\langle n \rangle$ be the mean number of quanta emitted by some distant object and let us take the variation of eq. (8) with respect to $p_i(n)$. As before, α^* and β^* enforce normalization and mean number of quanta conditions. Thus we shall now vary the quantity

$$\begin{aligned} I' = & - \delta \sum_n p_i(n) (\ln p_i(n) \pm \alpha^* - 1 + n\beta^*) \\ & \pm \delta \sum_{n,m} p(m|n) p_i(n) \ln q(n|m). \end{aligned} \quad (22)$$

Accordingly,

$$\begin{aligned} \delta I' = & - \sum_n (\ln p_i(n) \pm \alpha^* + (\beta^* n)) \delta p_i(n) \\ & \pm \sum_{n,m} p(m|n) \ln q(n|m) \delta p_i(n) \pm \sum_{n,m} \frac{p(m|n) p_i(n)}{q(n|m)} \delta q(n|m). \end{aligned} \quad (23)$$

Observe that in this expression the variations of $q(n|m)$ had to be included, since this quantity is related to $p_i(n)$ and $p_o(m)$ through Jaynes identity [compare eq. (3) with (4)]

$$p(m|n) p_i(n) = q(n|m) p_o(m). \quad (24)$$

We can eliminate $q(n|m)$ throughout eq. (23), obtaining

$$\begin{aligned} \delta I' = & - \sum_n (\ln p_i(n) \pm \alpha^* + \beta^* n) \delta p_i(n) \\ & \pm \sum_{n,m} p(m|n) (\ln p(n|m) \pm \ln p_i(n) - \ln p_o(m)) \delta p_i(n) \\ & \pm \sum_{n,m} p_o(m) \delta q(n|m). \end{aligned} \quad (25)$$

Some simplifications can now be done. First, if we sum over m the second term in the second sum, this contribution just cancels out the first term in the first one. Secondly, the last sum vanishes owing to normalization of the joint probability $p_{o,i}(m, n)$,

$$\sum_{n,m} \delta q(n|m) p_o(m) = - \sum_{n,m} q(n|m) \delta p_o(m) = \sum_m \delta p_o(m) = 0. \quad (26)$$

Next, if we introduce in this expression our earlier ansatz for $p_o(m)$ [eq. (12)] and impose that these variations vanish, we obtain after some algebra,

$$\begin{aligned}
& - (\alpha^* \pm \beta^* n) \pm \alpha \pm \beta \sum_m mp(m|n) \\
& \pm \sum_m p(m|n)(\ln p(m|n) \pm B(m)) = 0
\end{aligned} \tag{27}$$

Last, if we recall the equation we obtained for $B(m)$ [eq. (13)], the sum vanishes. Thus

$$\alpha^* \pm \beta^* n = \alpha \pm \beta \langle m \rangle_n, \tag{28}$$

where $\langle m \rangle_n \equiv \sum_m mp(m|n)$. After multiplying this equation by $p_i(n)$ and summing over n , we see that

$$\alpha^* \pm \beta^* \langle n \rangle = \alpha \pm \beta \langle m \rangle, \tag{29}$$

meaning that I_{\max} can be alternatively expressed in terms of α^* , β^* and $\langle n \rangle$. If this were everything, there would be no real gain after so much effort. However, if in the expression (28) we take $n = 0$ and $n = 1$, we obtain two equations for the unknowns α^* and β^* . Next, if we solve these equations and insert the results back into eq. (29), it becomes possible to express the mean number of quanta that reach the telescope in terms of those previously emitted, without having to solve the equation for $B(m)$. The result is

$$\langle m \rangle = (\langle m \rangle_1 - \langle m \rangle_0) \langle n \rangle \pm \langle m \rangle_0. \tag{30}$$

This expression can be put in a very suggestive form if, in the first term $[\langle m \rangle_1 \equiv \sum_m mp(m|1)]$, we substitute the dummy variable m by $m \pm 1$,

$$\langle m \rangle = \langle n \rangle \pm \langle m \rangle_0 \pm \langle n \rangle \left(\sum_m mp(m \pm 1|1) - \langle m \rangle_0 \right). \tag{31}$$

Accordingly, the first term is the number of incoming quanta, the second is the number of particles created out of the vacuum (spontaneous emission) and the last represents the stimulated emission. As a consistency check, let us consider a noise which is statistically independent of the signal. This translates mathematically into $p(m|n) = f(m - n)$, in which case the last term vanishes (no stimulated emission), as it should be. It is worth while to recall that eq. (18) was derived assuming the optimization of information transmission and, accordingly, this relation holds only in this regime.

We already have all the tools we need for the cosmological problem. So let us move to Section IV.

IV. Cosmological noise

In practice, by cooling down the photon counter, one can bring the mean kinetic energy of atoms in the detector (photoelectric device) beneath the work function energy threshold, making the thermal noise negligible. Thus, we shall henceforth assume that the communication system is not subjected to any noise other than gravitational. Furthermore, in order to render the calculations feasible, we shall consider a spatially flat cosmological

model where both the observer and the emitter are placed at asymptotically static regions or, at least, at regions which do not expand too fast. This guarantees that we have a meaningful definition of particles [5, 7, 8].

Let $|m_{\vec{k}}, r_{-\vec{k}}\rangle_{\text{out}}$ represent a state of m quanta of momentum $|\vec{k}|$ reaching the telescope from the direction of observation and r coming from the opposite direction. These states can be built up out of the vacuum state as

$$|m_{\vec{k}}, r_{-\vec{k}}\rangle_{\text{out}} = \frac{1}{\sqrt{m!r!}} b_{\vec{k}}^{+m} b_{-\vec{k}}^{+r} |\bar{0}\rangle, \quad (32)$$

where $b_{\vec{k}}^{\pm}/b_{\vec{k}}$ stand for the creation/annihilation operators of (outgoing) particles of some given species and momentum \vec{k} . Of course,

$$b_{\vec{k}}|\bar{0}\rangle = 0. \quad (33)$$

Similarly, $|n_{\vec{k}}, s_{-\vec{k}}\rangle_{\text{in}}$ represents the state of n quanta of momentum $|\vec{k}|$ which were emitted towards the observer and s in the opposite direction. Again,

$$|n_{\vec{k}}, s_{-\vec{k}}\rangle_{\text{in}} = \frac{1}{\sqrt{n!s!}} a_{\vec{k}}^{+n} a_{-\vec{k}}^{+s} |0\rangle \quad (34)$$

where $a_{\vec{k}}^{\pm}/a_{\vec{k}}$ stand for the creation/annihilation operators of (incoming) particles of the same species and momentum \vec{k} . Obviously,

$$a_{\vec{k}}|0\rangle = 0. \quad (35)$$

Owing to the spatial homogeneity of the model, the normal modes associated both to the incoming and outgoing creation/annihilation operators (u_k and v_k , respectively) consist of plane waves with some unusual time dependence. Therefore, the scalar products between incoming and outgoing normal modes have the generic form [10, 7]

$$\alpha_{\vec{k}, \vec{k}'} \equiv (u_{\vec{k}}, v_{\vec{k}'}) = A(k) \delta_{\vec{k}, \vec{k}'}, \quad (36)$$

and

$$\beta_{\vec{k}, \vec{k}'} \equiv (u_{\vec{k}}, v_{\vec{k}'}^*) = B(k) \delta_{-\vec{k}, \vec{k}'}, \quad (37)$$

the above quantities being the standard Bogolubov coefficients [10, 8]. For convenience we shall adopt A and B as a short notation for $A(k)$ and $B(k)$. These coefficients are known to satisfy the relation [10, 8, 7]

$$|A|^2 - |B|^2 = 1. \quad (38)$$

These relations among the incoming and outgoing normal modes translate, in terms of the corresponding creation/annihilation operators, as shown in refs. [7, 10] as

$$a_{\vec{k}}^{\pm} = A b_{\vec{k}}^{\pm} - B b_{-\vec{k}}^{\pm}, \quad (39)$$

and

$$b_{\vec{k}}^{\pm} = A a_{\vec{k}}^{\pm} \pm B^* a_{-\vec{k}}^{\pm}. \quad (40)$$

We shall now assume that the distant object we wish to observe emits n quanta. Of course, the most efficient way of transmitting information would be to send all quanta towards the observer and *none* in the opposite direction, i.e. the initial state is $|n_{\vec{k}}, 0_{-\vec{k}}\rangle_{\text{in}}$. Let m be the number of quanta that reach the telescope. Then the final state should be represented by $|m_{\vec{k}}, r_{-\vec{k}}\rangle_{\text{out}}$. Owing to the homogeneity of the model, r is not a free parameter but is constrained by momentum conservation, $n = m - r$. In the spirit of quantum mechanics we shall define the condition probability for m detected quanta, while n were sent, by

$$p(m|n) = \left| \langle 0_{-\vec{k}}, n_{\vec{k}} | m_{\vec{k}}, r_{-\vec{k}} \rangle_{\text{out}} \right|^2. \quad (41)$$

This transition probability was derived in Appendix A, where it is shown to correspond to a *negative binomial distribution*:

$$p(m|n) = \binom{m}{n} (1-x)^{(n+1)} x^{(m-n)}, \quad (42)$$

where $x = (|B|/|A|)^2$.

As a warming up exercise, we shall compute the mean number of particles that reach the telescope $\langle m \rangle$ in terms of $\langle n \rangle$, the number originally sent by the source. According to eq. (31), it entails the calculation of

$$\langle m \rangle_0 = \sum_m m p(m|0) = (1-x) \sum_m m x^m = (1-x) x \frac{d}{dx} \sum_m x^m = \frac{x}{1-x} \quad (43)$$

and

$$\begin{aligned} \sum_m m p(m \pm 1|1) &= \sum_m m (1-x)^2 \binom{m \pm 1}{1} x^m \\ &= (1-x)^2 \sum_m m (m \pm 1) x^m = x(1-x)^2 \frac{d^2}{dx^2} \sum_m x^{(m+1)} \\ &= x(1-x)^2 \frac{d^2}{dx^2} x \frac{x}{1-x} = \frac{2x}{1-x}. \end{aligned} \quad (44)$$

Inserting these results into eq. (31) we obtain the desired relation

$$\langle m \rangle = \langle n \rangle \pm \frac{x}{1-x} (\langle n \rangle \pm 1), \quad (45)$$

showing explicitly that there is spontaneous emission associated with this distribution. This could be foreseen since the latter does not depend exclusively on the difference $m - n$. This fact will have important consequences for us.

In the case of a negative binomial distribution [eq. (42)], the equation for $B(m)$ [eq. (13)] reads

$$\sum_{m \geq n} \binom{m}{n} x^m B(m) = - \sum_{m \geq n} \binom{m}{n} \ln p(m|n), \quad (46)$$

where we already cancelled out an overall factor $(1-x)^n x^{-n}$. Now, if we multiply both members by a fiducial quantity λ^n and sum over n we obtain

$$\sum_i (1+\lambda)^i x^i B(i) = -\mathcal{F}(\lambda), \quad (47)$$

where

$$\mathcal{F}(\lambda) = \sum_j \sum_i' \binom{i}{j} \lambda^j \ln p(i|j) \quad (48)$$

(the ' here is to remind us that the sum must be performed over $j \leq i$). The trick consists in taking the m -th derivative of eq. (47) with respect to λ and, at the end, the limit $\lambda \rightarrow -1$. Then

$$B(m) = -\frac{x^{-m}}{m!} \left. \frac{d^m \mathcal{F}(\lambda)}{d\lambda^m} \right|_{\lambda \rightarrow -1} \quad (49)$$

It is convenient to split $\mathcal{F}(\lambda)$ into two terms

$$\mathcal{G}(\lambda) = \sum_i \sum_j' \binom{i}{j} \lambda^j \ln [(1-x)^{(n+1)} x^{(m-n)}] \quad (50)$$

and

$$\mathcal{H}(\lambda) = \sum_i \sum_j' \binom{i}{j} \lambda^j \ln \left[\binom{i}{j} \right]. \quad (51)$$

The calculation of $\mathcal{G}(\lambda)$ is straightforward, and is performed in Appendix B. The result is

$$\mathcal{G}(\lambda) = \frac{(1-y) \ln(1-y) + y \ln y}{[1 - (1+\lambda)y]^2}. \quad (52)$$

Unfortunately, the logarithm of the binomial which appears in $\mathcal{H}(\lambda)$ precludes its calculation in an *exact closed form*. In order to overcome this difficulty, our strategy will be: i.) to adopt an approximation scheme which allows us to perform the sum in a closed form; ii.) to compare it with the exact value we shall obtain numerically. When comparing these results, it will turn out that the approximation is fairly good, either for modest numbers of quanta m or for small x . The idea is very simple: the function $\ln \binom{m}{n}$, as regarded as a function of n for given m , is concave; vanishes at both $n = 0$ and m and it reaches its maximum at $m/2$. This fact is very propitious for fitting this function with a concave parabola. These maxima are adjusted so as to coincide with each other. Last, this maximum is evaluated with the help of Stirling's approximation. This procedure yields

$$\ln \binom{m}{n} \approx 4 \ln 2 \left(n - \frac{n^2}{m} \right). \quad (53)$$

Inspection of fig.I, where these two functions are plotted for $m = 100$, reveals a good agreement between them, in spite of the roughness of approximation (53). In order to obtain (approximately) $\mathcal{F}(\lambda)$, we insert eq. (53) into eq. (51) and perform the double sum. This calculation was also carried out in Appendix B and the result is

$$\mathcal{H}(\lambda) \approx 4 \ln 2 \frac{\lambda x^2}{[1 - (1+\lambda)x]^2} \quad (54)$$

The quantity $B(m)$ can now be obtained in this approximation by adding these contributions and inserting $\mathcal{F}(\lambda) = \mathcal{G}(\lambda) + \mathcal{H}(\lambda)$ into eq. (49). After some trivial algebra,

$$B(m) \approx \nu + \mu m \quad (55)$$

with

$$\begin{aligned}\nu &= (4 \ln 2) x^2 - x \ln x - (1-x) \ln(1-x) \\ &\text{and} \\ \mu &= (4 \ln 2) x \pm \nu.\end{aligned}\tag{56}$$

In order to be in a position of comparing this result with the exact one, we insert the expression for $\mathcal{H}(\lambda)$ [eq. (51)] into eq. (49) and add it to the previously obtained contribution which stemmed from $\mathcal{G}(\lambda)$. The exact $B(m)$ can be cast in the form

$$\begin{aligned}B(m) &= -[x \ln x \pm (1-x) \ln(1-x)](m+1) \\ &\quad - \sum_i \sum_j \binom{i}{j} \binom{j}{m} (-1)^{(j-m)} x^{(i-m)} \ln \binom{i}{j}.\end{aligned}\tag{57}$$

The program which performs the above double sum was constructed to sum over $i \leq j$ and to halt at i 's when the difference between two consecutive iterations was less than 5%. The numerical results, as well as those obtained through the analytical approximation of eq. (56), are displayed for various values of x through tables I - IV. The very good agreement between these B 's is particularly remarkable for small values of m , where we would not expect this to happen. This should be attributed to the fact that the contribution coming from $\mathcal{G}(\lambda)$ helps to dilute the difference between the exact and the approximate values of $\mathcal{H}(\lambda)$. As we can inspect in the tables, these agreements are unfortunately spoiled for larger x . The reason is that the convergence of the double sum is very slow for larger x and, consequently, as x increases, many more terms must be considered; if the sum halts at, say, $i = k$ the number of terms added so far is given by $(k+1)(k+2)/2$. It is thus no wonder that they do not agree, after so many errors are added up.

We shall now take our analytical approximation for $B(m)$, eq.(56) and compute $p_o(m)$ [eq. 12]. In terms of the new variables $\sigma = \nu \pm \alpha$ and $\rho = \mu \pm \beta$, it reads

$$p_o(m) = e^{-(\sigma+m\rho)}.\tag{58}$$

The normalization of probabilities yields the equation

$$\sigma = -\ln(1 - e^{-\rho}),\tag{59}$$

which should be supplemented by the condition

$$\langle m \rangle = -\frac{d\sigma}{d\rho} = \frac{1}{e^\rho - 1}.\tag{60}$$

The equation for I_{\max} can be alternatively written in terms of these new variables

$$I_{\max} = \sigma \pm \rho \langle m \rangle - \nu - \mu \langle m \rangle\tag{61}$$

Thus, expressing both ρ and σ in terms of $\langle m \rangle$, and making the substitution in eq. (61), we obtain our final result

$$I_{\max} = \ln(\langle m \rangle \pm 1) \pm \ln \left(\frac{m \pm 1}{m} \right) - \nu - \mu \langle m \rangle.\tag{62}$$

After so much effort, it would be nice to plot I_{\max} against x for various values of $\langle m \rangle$ (see fig. II to IV for $m = 5, 50, 100$). In fig. II, we observe an unphysical tail at $x \approx 0.8$, where $I_{\max} \lesssim 0$. This feature is also blamed on the slow convergence of the series and superposition of errors. This tail moves towards lower values of x with increasing $\langle m \rangle$ (see figs. III and IV). Now, if we inspect table V, we see that the approximation for $B(5)$ is reasonable up to $x \approx 0.75$ (errors up to $\approx 40\%$), and that this difference blows up for larger x . Therefore, up to $x \approx 0.75$, fig. II should give a qualitatively good description of the situation for $\langle m \rangle = 5$. All these plots strongly suggest that $I_{\max} \rightarrow 1$ for any $\langle m \rangle$. Indeed it must be so, since the ratio between the number of quanta which bear useful information to the number of those which have been added to the signal (noise) [see eq. (45)]:

$$\frac{\langle n \rangle}{\langle m \rangle - \langle n \rangle} = \left(\frac{n+1}{n} \right) \left(\frac{1-x}{x} \right), \quad (63)$$

vanishes as $x \rightarrow 1$. On the other hand, eq. (62) should represent a very good approximation for small values of x (adiabatic limit), essentially because the series converges very fast.

V. Discussion

Let us now finally address the questions we posed in the introduction. The first is: Could we in principle observe the initial singularity? To this end, I will very briefly review some important results of quantum field theory in curved space-time. I shall follow the discussion contained in Birrel and Davies [5] and also in the review article by Parker [7]. For definiteness we shall consider a scalar field. In other words, our information carriers can be thought as of mesons. The situation should be similar for all other bosons (for gauge bosons, different helicities and color states should be regarded as corresponding to different channels). Let us also consider some (small) departures from isotropy, say, via the Kasner model,

$$ds^2 = -dt^2 + \sum_i a_i^2(t) dx_i^2. \quad (64)$$

The equation of motion for the scalar field is, as usual,

$$[\square - \xi R - m^2] \phi(x) = 0, \quad (65)$$

where m is the mass of the field quanta and ξ is the coupling constant ($\xi = 0$ for minimal coupling; $\xi = 1/6$ for conformal coupling). One can now expand the field $\phi(x)$ operator in terms of normal modes $u_{\vec{k}}(x)$, which can be expressed in the convenient form [5]:

$$u_{\vec{k}}(x) = \frac{1}{(2\pi)^{3/2}} e^{i\vec{k}\cdot\vec{x}} C^{-1/2}(\eta) \chi_{\vec{k}}(\eta), \quad (66)$$

where $C(t) = (a_1 a_2 a_3)^{2/3}$ and $\eta = \int^t C^{-1/2} dt'$ is the conformal time variable. Insertion of eq. (66) into the equation of motion for the scalar field yields

$$\frac{d^2 \chi_{\vec{k}}}{d\eta^2} + \omega_{\vec{k}}^2(\eta) \chi_{\vec{k}} = 0, \quad (67)$$

with

$$\omega_{\vec{k}}(\eta) = C \left[\sum_i \frac{k_i^2}{C_i} + m^2 + \left(\xi - \frac{1}{6} \right) R \right] + Q, \quad (68)$$

where $C_i(\eta) \equiv a_i^2$ and $Q \equiv \frac{1}{72} \sum_{i < j} (\dot{C}_i - \dot{C}_j)^2$.

Equation (67) corresponds to a parametric oscillator, an oscillator whose frequency is time-dependent. This shows that the analogy we draw in the introduction between the cosmological problem with the pendulum is very appropriate.

Now, in the limit where the mass vanishes, the field is conformally coupled to gravity ($\xi = 1/6$) and the space time is isotropic ($C_1 = C_2 = C_3; Q = 0$), the frequency is time-independent: no transition occurs and the Bogolubov coefficient β vanishes. Therefore, *in the conformally trivial configuration, there is no degradation of information*. However, our Universe is neither completely homogeneous nor isotropic: there are structures in it, galaxies, clusters of galaxies, superclusters, ..., which causes deviations of the FRW background. Even in the very early Universe, one expects quantum fluctuations from this FRW background to occur.

In the non-adiabatic limit (early Universe) one can mimic the effect of initial anisotropies through a minimal coupling ($\xi = 0$), which is much easier to solve. In this case, eq. (67) reads

$$\frac{d^2 \chi_{\vec{k}}}{dt^2} + \vec{k}^2 a^4 \chi_{\vec{k}} = 0. \quad (69)$$

In order to make a very general discussion, one would like to fit the cosmic scale factor by a function with three parameters, the initial and final radii of the Universe as well as its expansion rate

$$a^4(t) = a_0^4 + b e^{\tau/s}. \quad (70)$$

The solution found by Parker [7] is given in terms of gamma and Bessel functions. By taking the scalar product between in and out normal modes he found that

$$x = \left[\frac{\beta_k}{\alpha_k} \right]^2 = e^{-4\pi s k a_0^2}. \quad (71)$$

Since the energy scale at which we could in principle observe quanta is much smaller than the Planck mass, when the Universe had a radius of the order $a_0 \sim l_p$, $s k a_0 \ll 1$ (even for a moderate expansion rate s) or, equivalently, $x \rightarrow 1$. Therefore, according to an earlier discussion, all the information concerning this epoch was completely washed out, including the specification of the initial conditions of all matter and gravitational fields!!

A second point I would like to address is related to the Penrose scenario [9]. What this means is that the Universe is absolutely smooth and then, at a given moment, anisotropies start to build up. To study this scenario let me introduce the line element,

$$ds^2 = -dt^2 + a^2(t)(\delta_{ij} + h_{ij}(x, t))dx^i dx^j, \quad (72)$$

where $h_{ij}(x, t)$ represents fluctuations around the FRW metric. Now, if the field is massless and conformally coupled to gravity ($\xi = 1/6$), instead of dealing directly with $g_{\mu\nu}$ and

$\phi(x)$, it becomes convenient to deal with the conformally related quantities $\tilde{g}_{\mu\nu} = \eta_{\mu\nu} \pm h_{\mu\nu}$ and $\tilde{\phi} = a\phi$. In terms of these new quantities the equation of motion reads

$$\left[\square - \frac{1}{6} \tilde{R} \right] \tilde{\phi} = 0 \quad (73)$$

where \square and \tilde{R} are the d'Alembertian and the scalar curvature related to $\tilde{g}_{\mu\nu}$, respectively. Since we are interested in tracing the evolution of structures in the Universe through the fluctuations of the metric, it is appropriate to restrict ourselves to the trace of h_{ij} . This is because it is the unique component which couples directly to the density fluctuations: the transverse traceless part of h_{ij} represents gravitational waves propagating in the FRW background while the vectorial part couples only to matter vorticity [18, 21]. In other words, we shall assume

$$h_{\mu\nu} = \begin{cases} h_{0\mu} = 0, \\ h_{ij} = \frac{1}{3} \delta_{ij} h \end{cases} \quad (74)$$

We can now expand both the scalar curvature and the d'Alembertian up to linear terms in h ,

$$\tilde{R} \approx \square_o h \pm \frac{1}{3} \nabla^2 h \quad (75)$$

and

$$\square = \frac{1}{\sqrt{-\tilde{g}}} \partial_\mu \left(\sqrt{-\tilde{g}} \tilde{g}^{\mu\nu} \partial_\nu \right) \approx \square_o - \frac{1}{3} h \nabla^2 - \frac{1}{3} (\vec{\nabla} h) \cdot \vec{\nabla} \pm \frac{1}{2} \eta^{\mu\nu} (\partial_\mu h) \partial_\nu, \quad (76)$$

where \square_o stands for the d'Alembertian in flat space-time. Therefore, in this approximation the equation of motion (73) reads

$$\left[\square_o - \frac{1}{3} h \nabla^2 - \frac{1}{3} (\vec{\nabla} h) \cdot \vec{\nabla} \pm \frac{1}{2} \eta^{\mu\nu} (\partial_\mu h) \partial_\nu - \frac{1}{6} (\square_o h) - \frac{1}{18} (\nabla^2 h) \right] \tilde{\phi} = 0. \quad (77)$$

Now, if we assume the scale of the density fluctuations to be macroscopic, in the sense that the typical wavelength of h is much larger than the quanta Compton wavelengths ($|\vec{\nabla} \tilde{\phi}(x)/\tilde{\phi}(x)| \gg |\vec{\nabla} \tilde{h}(x)/\tilde{h}(x)|$), then we can safely ignore all spatial derivatives of h in the above equation. Accordingly,

$$\left[\square_o - \frac{1}{3} h \nabla^2 - \frac{1}{2} \frac{\partial h}{\partial \eta} \frac{\partial}{\partial \eta} \pm \frac{1}{6} \frac{\partial^2 h}{\partial \eta^2} \right] \tilde{\phi} \approx 0. \quad (78)$$

If we now seek normal modes in the form

$$u_k(\vec{x}, \eta) = \frac{1}{(2\pi)^{3/2}} \frac{e^{-h/4}}{a(\eta)} e^{i\vec{k} \cdot \vec{x}} \chi_k, \quad (79)$$

then eq. (78) can be expressed as a Schroedinger-like equation,

$$\left[\frac{\partial^2}{\partial \eta^2} \pm \omega^2 - V_k(\eta) \right] \chi_k = 0, \quad (80)$$

where

$$\omega^2 \equiv k^2 \left(1 - \frac{1}{3} h_0 \right) \quad (81)$$

and

$$V_k(\eta) = -\frac{1}{3}k^2(h_0 - h(\eta)) \pm \frac{5}{12}\ddot{h} \pm \frac{1}{16}\dot{h}^2. \quad (82)$$

Here h_0 stands for the perturbations that already existed at time $\eta = \eta_0$. Following Birrel and Davies [5, 22], we express eq. (80) as an integral equation,

$$\chi_k(\eta) = \chi_k^0(\eta) - \frac{1}{\omega} \int_{\eta_0}^{\eta} \chi_k(\eta') V(\eta') \chi(\eta') \sin \omega(\eta - \eta'), \quad (83)$$

where $\chi_k^0(\eta)$ are the ingoing modes at $\eta = \eta_0$. The outgoing modes $\chi(\eta)$ can be expanded in terms of the ingoing ones,

$$\chi_k(\eta) = \alpha_k(\eta) \chi_k^0(\eta) \pm \beta_k(\eta) \chi_k^{0*}(\eta). \quad (84)$$

In the limit where the fluctuations grow very slowly close to η_0 , then $V_k(\eta)$ would vanish [see eq. (82)]. Accordingly, the incoming modes have the standard oscillatory behaviour and we can express $\beta_k(\eta)$ in the adiabatic approximation simply as

$$\beta_k(\eta) \approx \int_{\eta_0}^{\eta} V_k(\eta) e^{-2i\omega\eta} d\eta. \quad (85)$$

Inspecting eqs. (85) and (82), we see that the Bogolubov coefficient vanishes as long as the anisotropies remain frozen, during the cosmological evolution. However, by the time the density fluctuations start to grow, so does also h , since these quantities are related through Einstein's equations [20]. According to eqs. (82) and (85) the same must be true for the modulus of the Bogolubov coefficient $|\beta_k(\eta)|$. Therefore, the degradation of the information is a monotonically increasing function of clumsiness in the Universe!

VI. Summary

Having in mind the important consequences of the quantum constraint on the information transmission discussed in this paper, we must face the question: Does our result represent a universal limitation on information transmission on a curved background? Could it be overcome by considering another quantum field representation such as coherent or squeezed states representations? It is widely known in quantum optics that the 'number of particles' representation optimizes the transmission rate [17, 23, 24, 25]. Furthermore, we recently proved a theorem that says that the transmission rate is maximized for any representation consisting of Hamiltonian eigenstates [26]. But what about trying to transmit information by means of fermions (neutrinos)? Could we do better than by bosons? The answer is no, because the exclusion principle restricts the number of available quantum states that can be obtained by distributing a fixed number of quanta. Another related issue has to do with the fact that the signal and noise are not stochastically independent. In communication problems, it often happens that the noise is statistically independent of the input signal, e. g. the thermal noise. It is then always possible to transmit any information, provided we code it with a number of quanta that is much larger than that introduced later by the noise in the signal. However, as we saw, the gravitational quantum noise precludes such a possibility, because it embodies stimulated as well as spontaneous

emission. Stated in another way, if tried the same trick to overcome the noise (code with more quanta), we would end with even more noise. In summary, our results represent a fundamental quantum mechanical restriction on the ability of conveying information on a curved background.

The main consequences of the quantum capacity formula in a gravitational background are two. First, we saw that as the cosmological singularity is approached, the gravitational (quantum) noise blows up and all information is washed out. Therefore, the issue of what the initial conditions of the Universe were cannot be decided by experiment (unless there are topological imprints). Accordingly, this question can have but a philosophical taste since no imprints of these conditions are left in our Universe. The second consequence is related to the fundamental time-arrow puzzle [9, 10]. We found out that the direction in which information is degraded (thermodynamical arrow) is aligned with the gravitational arrow, the direction in which structures evolve and clumsiness grows in our Universe. Recently a possible alternative explanation based on the inflationary scenario for the alignment of these arrows has also been proposed in the literature [27, 28].

There are a myriad of questions that remain to be addressed. For instance, how do the gravitational waves degrade information of a quantum system? It is known that for (gravitational) plane waves the Bogolubov coefficient vanishes [29]. However this is only true for perturbative solutions. There are non-perturbative solutions which even display horizons [30]. What about quantum gravity effects? It is known that gravitons also degrade information [11]. Can these effects be tackled by some a variation of the argument we considered? What about singularities which are shielded by event horizons? Naked singularities?

Acknowledgements

I am indebted to J.D. Bekenstein for his valuable comments on this paper. The author is partially supported by The World Laboratory.

References

- [1] Warren, J., Hewett, P.C., Osmer, P.S. and Irwin, M.J. (1987) *Nature*, **330**, 453.
- [2] Chambers, K. , Miley, G. and Breugel, W. van (1988) *Ap. J.* **327**, 47.
- [3] Shannon, C.E. (1948) *Bell Syst. Tech. J.* **27**, 379.
- [4] Shannon, C.E. and Weaver, W. (1949) *The Mathematical Theory of Communication* (Urbana: University of Illinois Press).
- [5] Birrell, N.D. and Davies, P.C.W. (1982) *Quantum Fields in Curved Space* (Cambridge : Cambridge University Press).
- [6] Brillouin, L. (1965) *Science and Information Theory* (London: Academic Press).

- [7] Parker, L. (1977) ‘The production of elementary particles by strong gravitational fields’ in *Asymptotic Structure of Space-Time*, eds. F.P. Esposito and L. Witten (New York: Plenum).
- [8] DeWitt, B. S. (1975) *Phys. Rep.* **19C**, 297.
- [9] Penrose, R. (1979) ‘Singularities and time-asymmetry’ in *General Relativity: An Einstein Centennial Survey*, eds. S.W. Hawking and W. Israel (Cambridge: Cambridge University Press).
- [10] Davies, P.C.W. *The Physics of Time Asymmetry* (London: Surrey University Press).
- [11] Smolin, L. (1985) *Gen. Rel. Grav.* **17**, 417.
- [12] Bekenstein, J.D. and Schiffer, M. (1990) *Int. J. Mod Phys.* **C1**, 355.
- [13] Schiffer, M. (1991) *Phys. Rev.* **A43**, 5337.
- [14] Schiffer, M. (1991) unpublished.
- [15] Lebedev, D.S. and Levitin, L.B. (1963) *Dokl. Akad. Nauk SSSR* **149**, 1299 [*Sov. Phys. Dokl.* **8**, 377].
- [16] Pendry, J.B. (1983) *J. Phys.* **A16**, 2161.
- [17] Caves, C. M. (1991) ‘Quantum Limits on Communication rates’, to appear in *Rev. Mod. Phys.*
- [18] Weinberg, S. (1972) *Gravitation and Cosmology* (New York: Wiley).
- [19] Peebles, P. J. E. (1980) *The Large-Scale Structure of the Universe* (Princeton: Princeton Univ. Press).
- [20] Kolb, W. and Turner, M. S. (1990) *The Early Universe* (Reading (MA): Addison-Wesley).
- [21] Zel’dovich Ya.B and Novikov B. (1983) *Relativistic Astrophysics*, Vol II, (Chicago: Univ. Chicago Press).
- [22] Birrel, N.D. and Davies, P.C.W. (1980) *J. Phys. A: Gen Phys.* **13**, 2109.
- [23] Yamamoto, Y. and Haus, H.A. (1986) *Rev. Mod. Phys.* **58**, 1001.
- [24] Takahashi, H. (1965) in *Advances in Communication Systems*, ed. A.V. Balakrishnan, (New York: Academic Press).
- [25] Saleh, B. and Teich, M. (1987) *Phys. Rev. Lett.* **58**, 2656.
- [26] Schiffer, M. (1991) *Phys. Lett.* **A154**, 361.
- [27] Goldwirth, D S. and Piran, T. (1991) *Class. Quantum Grav.* **8**, L155.

[28] Goldwirth D.S. (1991) *Phys. Lett* **256B**, 354.

[29] Gibbons, G. (1975) *Commun. Math. Phys.* **45**, 191.

[30] Bondi, H.; Pirani, F.A.E. and Robinson, I. (1959) *Proc. R. Soc. London* **A251**, 519.

Appendix A: The transition probabilities

Our task here is to express this $p(m|n)$ in terms of m and n . In the amplitude

$$\text{out}\langle r_{-\vec{k}}, m_{\vec{k}} | n_{\vec{k}}, 0_{-\vec{k}} \rangle_{\text{in}} = \frac{1}{\sqrt{n!}_{\text{out}}} \langle r_{-\vec{k}}, m_{\vec{k}} | a_{\vec{k}}^{+n} | 0 \rangle, \quad (86)$$

we can express $a_{\vec{k}}^{+n}$ with the aid of the Bogolubov transformation (39) and perform a binomial expansion since $[b_{\vec{k}}^+, b_{-\vec{k}}] = 0$. Thus,

$$\text{out}\langle r_{-\vec{k}}, m_{\vec{k}} | n_{\vec{k}}, 0_{-\vec{k}} \rangle_{\text{in}} = \frac{1}{\sqrt{n!}} \sum_{p=0}^n \binom{n}{p} A^p (-B)^{n-p} \text{out}\langle r_{-\vec{k}}, m_{\vec{k}} | b_{\vec{k}}^{+p} b_{-\vec{k}}^{n-p} | 0 \rangle. \quad (87)$$

One can easily show that

$$\begin{aligned} M_{m,n} &\equiv \text{out}\langle n_{-\vec{k}}, m_{\vec{k}} | b_{\vec{k}}^{+p} b_{-\vec{k}}^{(n-p)} | 0 \rangle \\ &= \sqrt{\frac{m!}{(m-p)!}} \sqrt{\frac{(r \pm n - p)!}{r!}}_{\text{out}} \langle (r \pm n - p)_{-\vec{k}}, (m-p)_{\vec{k}} | 0 \rangle. \end{aligned} \quad (88)$$

After some algebra and recalling that $m = r \pm n$, one obtains

$$\begin{aligned} M_{m,n} &= \left[\binom{m}{n} \right]^{1/2} \sum_{p=0}^n \binom{n}{p} A^p (-B)^{n-p} \\ &\text{out}\langle (m-p)_{-\vec{k}}, (m-p)_{\vec{k}} | 0 \rangle. \end{aligned} \quad (89)$$

On the other hand, the matrix element

$$\text{out}\langle s_{-\vec{k}}, s_{\vec{k}} | 0 \rangle = \frac{1}{n!} \langle \bar{0} | b_{\vec{k}}^s b_{-\vec{k}}^s | 0 \rangle \quad (90)$$

can also be calculated with the aid of the Bogolubov transformations. Indeed, if we apply eq. (40) and its adjoint to the initial vacuum $|0\rangle$,

$$b_{\vec{k}} | 0 \rangle = B^* a_{-\vec{k}}^+ | 0 \rangle \quad (91)$$

and

$$b_{-\vec{k}}^+ | 0 \rangle = A^* a_{-\vec{k}}^+ | 0 \rangle, \quad (92)$$

we see that

$$b_{-\vec{k}} | 0 \rangle = \frac{B^*}{A^*} b_{\vec{k}}^+ | 0 \rangle. \quad (93)$$

Under s repeated uses of this relation we obtain

$$b_{-\vec{k}}^s |0\rangle = \left[\frac{B^*}{A^*} \right]^s b_{\vec{k}}^{+s} |0\rangle. \quad (94)$$

Therefore,

$$\text{out} \langle s_{-\vec{k}}, s_{\vec{k}} | 0 \rangle = \frac{1}{m!} \left[\frac{B^*}{A^*} \right]^s \langle \bar{0} | b_{\vec{k}}^s b_{\vec{k}}^{+s} | 0 \rangle. \quad (95)$$

It is trivial to show that the above matrix element is just $m! \langle \bar{0} | 0 \rangle$. Thus,

$$\text{out} \langle s_{-\vec{k}}, s_{\vec{k}} | 0 \rangle = \left[\frac{B^*}{A^*} \right]^s \langle \bar{0} | 0 \rangle. \quad (96)$$

The vacuum to vacuum persistence can be estimated by enforcing the normalization of probabilities $p_{o \rightarrow s}$ (up to an arbitrary phase). Then,

$$\text{out} \langle s_{-\vec{k}}, s_{\vec{k}} | 0 \rangle = |A|^{-1} \left[\frac{B^*}{A^*} \right]^s. \quad (97)$$

Fortunately, we are now close to the end of this messy calculation. Inserting our last result [eq. (97)] into eq.(89),

$$\begin{aligned} \text{out} \langle r_{-\vec{k}}, m_{\vec{k}} | n_{\vec{k}}, 0_{-\vec{k}} \rangle_{\text{in}} &= |A|^{-1} (B^*)^{(m-n)} (A^*)^{-m} \left[\binom{m}{n} \right]^{1/2} \\ &\sum_{p=0}^n \binom{n}{p} (|A|^2)^p (-|B|^2)^{(n-p)} \end{aligned} \quad (98)$$

It can be seen easily that the sum corresponds to $(|A|^2 - |B|^2)^n = 1$, [see eq. (38)]. Further, defining $x = (|B|/|A|)^{1/2}$, we finally obtain

$$\left| \text{in} \langle 0_{-\vec{k}}, n_{\vec{k}} | m_{\vec{k}}, (n-m)_{-\vec{k}} \rangle_{\text{out}} \right|^2 = \binom{m}{n} (1-x)^{(n+1)} x^{(m-n)}. \quad (99)$$

Appendix B: Calculation of $\mathcal{G}(\lambda)$ and $\mathcal{H}(\lambda)$

Let me first recall the definition of $\mathcal{G}(\lambda)$

$$\begin{aligned} \mathcal{G}(\lambda) &= \sum_m \sum_{n \leq m} \binom{m}{n} x^m \lambda^n \ln \left[(1-x)^{(n+1)} x^{(m-n)} \right] \\ &= \sum_m \sum_{n \leq m} \binom{m}{n} x^m \lambda^n [\ln(1-x) \\ &\quad + m \ln x + m (\ln(1-x) - \ln x)] \\ &= \left[\ln(1-x) + \ln x \left(x \frac{d}{dx} \right) \right. \\ &\quad \left. + (\ln(1-x) - \ln x) \left(\lambda \frac{d}{d\lambda} \right) \right] \sum_m \sum_{n \leq m} \binom{m}{n} x^m \lambda^n. \end{aligned} \quad (100)$$

It is very easy to evaluate the series

$$\sum_m \sum_{n \leq m} \binom{m}{n} x^m \lambda^n = \sum_m [x(1 + \lambda)]^m = \frac{1}{1 - (1 + \lambda)x}. \quad (101)$$

Therefore combining these two equations,

$$\begin{aligned} \mathcal{G}(\lambda) &= \left[\ln(1 - x) \pm \ln x \left(x \frac{d}{dx} \right) \right. \\ &\quad \left. \pm \left(\ln(1 - x) - \ln x \right) \left(\lambda \frac{d}{d\lambda} \right) \right] \frac{1}{1 - (1 + \lambda)x}. \end{aligned} \quad (102)$$

Fortunately, after some algebra it becomes possible to rewrite this expression in a very simple form

$$\mathcal{G}(\lambda) = \frac{(1 - x) \ln(1 - x) \pm x \ln x}{[1 - (1 + \lambda)x]^2}. \quad (103)$$

We now turn our attention to the second contribution to $\mathcal{F}(\lambda)$. Adopting the approximation discussed,

$$\begin{aligned} \mathcal{H}(\lambda) &= \sum_m \sum_{n \leq m} \binom{m}{n} x^m \lambda^n \ln \binom{m}{n} \\ &\approx 4 \ln 2 \sum_m \sum_{n \leq m} \binom{m}{n} x^m \lambda^n \left(n - \frac{n^2}{m} \right) \end{aligned} \quad (104)$$

Because the function we are approximating, the logarithm of the binomial coefficient, does not give any contribution for $m = 0$, one should always take $n \geq 1$ in the above sums. Then,

$$\begin{aligned} \mathcal{H}(\lambda) &= 4 \ln 2 \lambda \frac{d}{d\lambda} \left[\sum_{m \geq 1} \sum_{n \leq m} \binom{m}{n} x^m \lambda^n - \sum_{m \geq 1} \frac{x^m}{m} \lambda \frac{d}{d\lambda} \sum_{n \leq m} \binom{m}{n} \lambda^n \right] \\ &= 4 \ln 2 \lambda \frac{d}{d\lambda} \left[\frac{1}{1 - (1 + \lambda)x} - \sum_{m \geq 1} \frac{x^m}{m} \lambda \frac{d}{d\lambda} (1 + \lambda)^m \right] \\ &= 4 \ln 2 \lambda \frac{d}{d\lambda} \left[\frac{1}{1 - (1 + \lambda)x} - \lambda x \sum_{m \geq 0} x^m (1 + \lambda)^m \right] \end{aligned} \quad (105)$$

where, in our last step we substituted $m \rightarrow m - 1$. The last sum is trivial:

$$\mathcal{H}(\lambda) = 4 \ln 2 \lambda \frac{d}{d\lambda} \frac{1 + \lambda x}{1 - (1 + \lambda)x} = 4 \ln 2 \frac{\lambda x^2}{[1 - (1 + \lambda)x]^2}. \quad (106)$$

Figure captions

1. **Fig. I.** Shows the superposition of the logarithm of the Binomial coefficients with the fitting function $4 \ln 2(n - n^2/m)$ as a function of n for $m = 100$.
2. **Fig. II.** Displays the behaviour of I_{\max} for $m = 5$
3. **Fig. III.** Displays the behaviour of I_{\max} for $m = 50$ Observe that the unphysical tail where $I_{\max} \approx$ moved to lower x 's.
4. **Fig. IV.** Displays the behaviour of I_{\max} for $m = 100$ Now the tail moved to even smaller values of x .

Table captions

1. **Table I.** Comparison of $B(m)$ obtained analytically through the quadratic approximation for the logarithm of the binomial coefficient for $x = 10^{-6}$, for various values of m .
2. **Table II.** The same, for $x = 10^{-3}$
3. **Table III.** Idem, for $x = 0.1$
4. **Table IV.** Idem, for $x = 0.4$
5. **Table V.** The same, now for fixed $m = 5$ and various values of x .

Table I. $x = 1.0 \cdot 10^{-6}$

m	$\frac{\Delta B(m)}{B(m)}$	$B(m)x10^{-3}$	$B^*(m)x10^{-3}$
0	9.010^{-8}	1.4810^{-2}	1.4810^{-2}
10	0.0098	0.136	0.135
20	0.034	0.247	0.255
30	0.066	0.353	0.376
40	0.091	0.455	0.496
50	0.11	0.555	0.617
60	0.13	0.653	0.737
70	0.14	0.749	0.857
80	0.16	0.844	0.978
90	0.17	0.938	1.10
100	0.18	1.03	1.22
110	0.19	1.12	1.34
120	0.20	1.21	1.46
130	0.21	1.30	1.58
140	0.22	1.39	1.82
160	0.24	1.57	1.95
170	0.24	1.65	2.06
180	0.25	1.74	2.18
190	0.26	1.83	2.30
200	0.27	1.91	2.42
210	0.27	2.00	2.54
220	0.28	2.08	2.66
230	0.29	2.16	2.79
240	0.29	2.25	2.90
250	0.29	2.33	3.02
260	0.30	2.41	3.14
270	0.31	2.50	3.26
280	0.31	2.58	3.38
290	0.32	2.66	3.51
300	0.32	2.74	3.63

Table II. $x = 1.0 \cdot 10^{-3}$

m	$\frac{\Delta B(m)}{B(m)}$	$B(m)$	$B^*(m)$
0	$1.75 \cdot 10^{-4}$	0.00791	0.00791
5	0.084	0.0367	0.0336
10	0.020	0.0606	0.0593
15	0.034	0.0821	0.0850
20	0.083	0.102	0.111
25	0.13	0.121	0.136
30	0.17	0.139	0.162
35	0.20	0.155	0.188
40	0.24	0.172	0.213
45	0.27	0.187	0.239
50	0.30	0.203	0.265
55	0.33	0.217	0.290
60	0.37	0.231	0.316
65	0.39	0.245	0.342

Table III. $x = 0.1$

m	$\frac{\Delta B(m)}{B(m)}$	$B(m)$	$B^*(m)$
0	0.041	0.339	0.353
2	0.28	0.704	0.504
4	0.30	0.943	0.655
6	0.28	1.12	0.806
8	0.24	1.26	0.957
10	0.19	1.37	1.11
12	0.15	1.49	1.26
14	0.11	1.58	1.41
16	0.062	1.66	1.56
18	0.017	1.74	1.71
20	0.062	1.75	1.86
22	0.12	1.80	2.01
24	0.18	1.83	2.16
26	0.20	1.93	2.32
28	0.24	1.99	2.47
30	0.28	2.04	2.62
32	0.39	1.99	2.77

Table IV. $x = 0.4$

m	$\frac{\Delta B(m)}{B(m)}$	$B(m)$	$B^*(m)$
0	0.25	0.895	1.11
1	0.14	1.30	1.12
2	0.27	1.56	1.13
3	0.34	1.74	1.14

Table V. $m = 5$

x	$\frac{\Delta B(m)}{B(m)}$	$B(m)$	$B^*(m)$
0.0001	0.06	0.00505	0.00474
0.0501	0.23	0.699	0.54
0.1001	0.29	1.04	0.73
0.1501	0.35	1.27	0.83
0.2001	0.39	1.48	0.89
0.2501	0.41	1.62	0.95
0.3001	0.43	1.77	1.00
0.3501	0.43	1.88	1.07
0.4001	0.42	1.99	1.15
0.4501	0.40	2.10	1.26
0.5001	0.37	2.20	1.38
0.5501	0.33	2.31	1.54
0.6001	0.29	2.41	1.71
0.6501	0.25	2.53	1.90
0.7001	0.20	2.63	2.11
0.7501	0.15	2.76	2.33
0.8001	17.6	-42.6	2.56

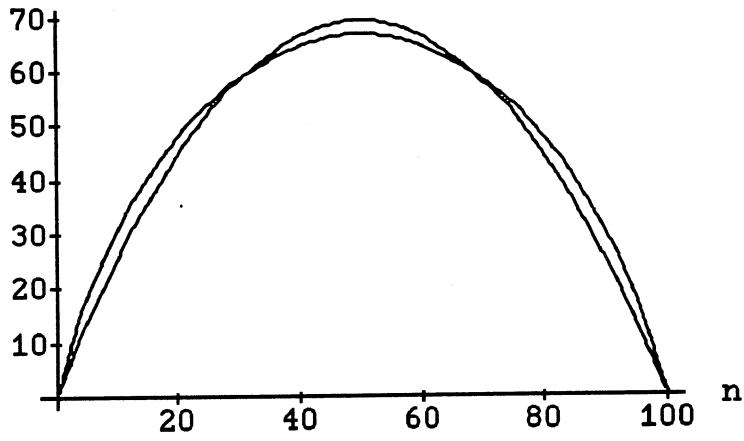


Figure I: Fit of $\ln\left(\frac{m}{n}\right)$ by a parabola, $m = 100$

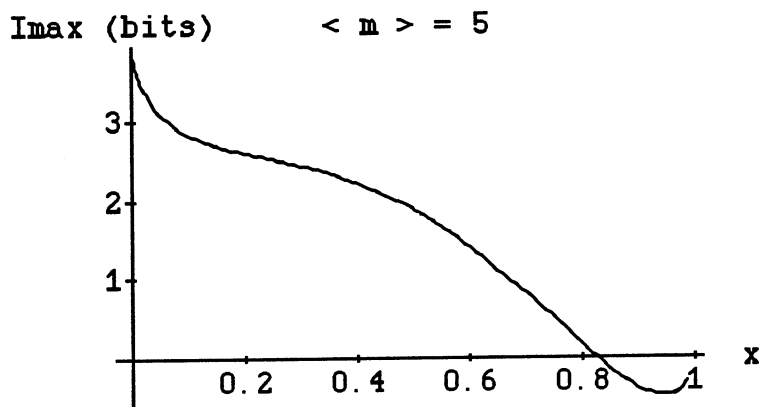


Figure II: I_{\max} for $\langle m \rangle = 5$

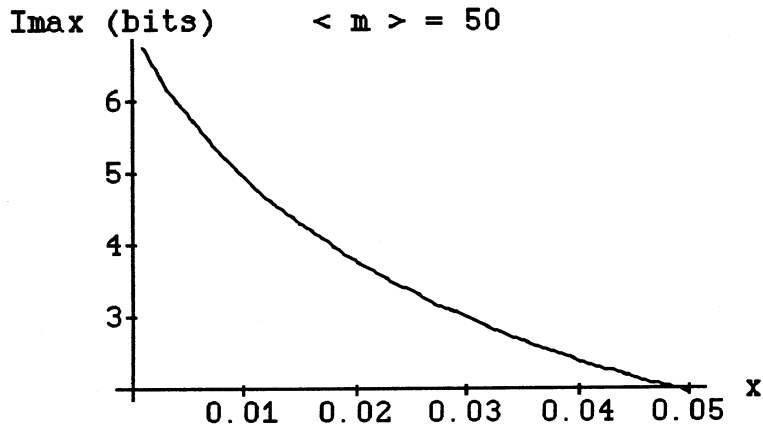


Figure III: I_{\max} for $\langle m \rangle = 50$

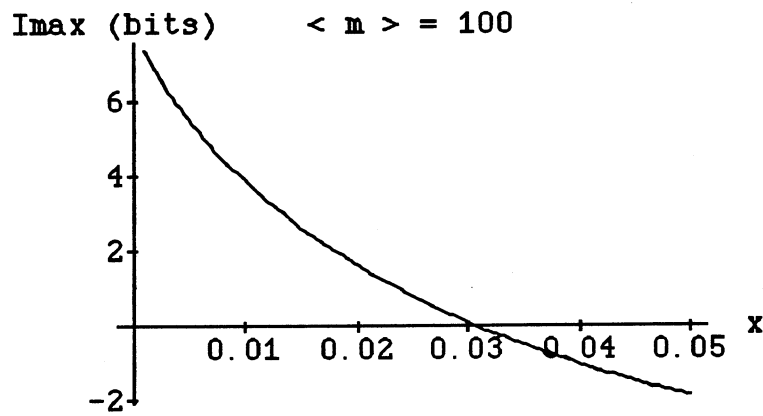


Figure IV: I_{\max} for $\langle m \rangle = 100$