2007 Paper 7 Question 8

Information Theory and Coding

(a) Suppose that the following sequence of Yes/No questions was an optimal strategy for playing the "Game of seven questions" to learn which of the letters $\{A, B, C, D, E, F, G\}$ someone had chosen, given that their *a priori* probabilities were known:

"Is it A?"	"No."
"Is it a member of the set $\{B, C\}$?"	"No."
"Is it a member of the set $\{D, E\}$?"	"No."
"Is it F?"	"No."

   (*i*) Write down a probability distribution for the seven letters, $p(A), \ldots, p(G)$, for which this sequence of questions was an optimal strategy. [3 marks]

   (*ii*) With this probability distribution, what was the uncertainty, in bits, associated with each question? [1 mark]

   (*iii*) What is the entropy of this alphabet? [1 mark]

   (*iv*) Now specify a variable length, uniquely decodable, prefix code for this alphabet that would minimise the average code word length. [3 marks]

   (*v*) What is your average coding rate R for letters of this alphabet? [1 mark]

   (*vi*) How do you know that a more efficient code could not be developed?
   [1 mark]

(b) The signal-to-noise ratio SNR of a continuous communication channel might be different in different parts of its frequency range. Explain how the information capacity C of a noisy continuous communication channel, whose available bandwidth spans from frequency ω_1 to ω_2, may be defined in terms of its signal-to-noise ratio as a function of frequency, $SNR(\omega)$. Define the bit rate for such a channel's information capacity, C, in bits/second, in terms of the $SNR(\omega)$ function of frequency. [5 marks]

(c) An invertible transform generates projection coefficients by integrating the product of a signal with each of a family of functions. In a reverse process, expansion coefficients can be used on those same functions to reproduce the signal. If the functions in question happen to form an orthonormal set, what is the consequence for the projection coefficients and the expansion coefficients?
   [2 marks]

(d) In the Information Diagram (a plane whose axes are time and frequency), why does the Gabor–Heisenberg–Weyl *Uncertainty Principle* imply that information is *quantised* – i.e. that it exists in only a limited number of independent quanta? [3 marks]