2001 Paper 9 Question 10

Information Theory and Coding

(*a*) (*i*) Construct an efficient, uniquely decodable binary code, having the prefix property and having the shortest possible average code length per symbol, for an alphabet whose five letters appear with these probabilities:

Letter	*Probability*
A	1/2
B	1/4
C	1/8
D	1/16
E	1/16

[4 marks]

(*ii*) How do you know that your code has the shortest possible average code length per symbol? [2 marks]

(*b*) (*i*) For a string of data of length N bits, what is the upper bound for its Minimal Description Length, and why? [2 marks]

(*ii*) Comment on how, or whether, you can know that you have truly determined the Minimal Description Length for a set of data. [2 marks]

(*c*) Suppose you have sampled a strictly bandlimited signal at regular intervals more frequent than the Nyquist rate; or suppose you have identified all of the zero-crossings of a bandpass signal whose total bandwidth is less than one octave. In either of these situations, provide some intuition for why you now also have knowledge about exactly what the signal must be doing at all points between these observed points. [3 marks]

(*d*) (*i*) Explain how autocorrelation can remove noise from a signal that is buried in noise, producing a clean version of the signal. For what kinds of signals, and for what kinds of noise, will this work best, and why? What class of signals will be completely unaffected by this operation except that the added noise has been removed? Begin your answer by writing down the autocorrelation integral that defines the autocorrelation of a signal $f(x)$. [5 marks]

(*ii*) Some sources of noise are additive (the noise is just superimposed onto the signal), but other sources of noise are multiplicative in their effect on the signal. For which type would the autocorrelation clean-up strategy be more effective, and why? [2 marks]