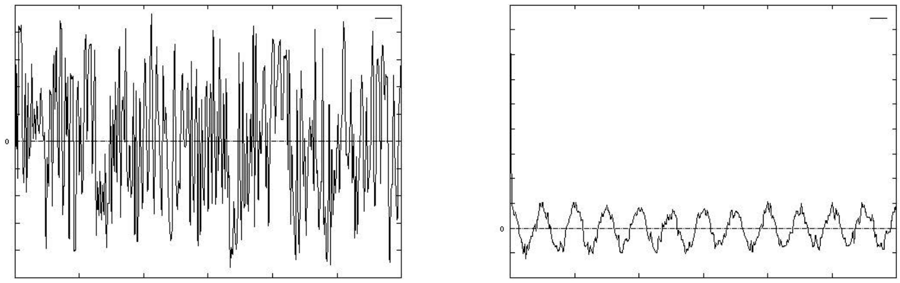
7 Information Theory (JGD)

(*a*) The left panel shows a very noisy cosmic signal $f(t)$, within which is buried a coherent quasi-periodic signal emitted by a pulsar (a collapsed neutron star). Write an auto-correlation integral which, when applied to the noisy signal $f(t)$, allows the clean quasi-periodic signal in the right panel to be detected in it. Explain why this works, and how computing the Fourier transform $F(\omega)$ of $f(t)$ can make the auto-correlation operation efficient. [5 marks]



(*b*) X and Y are discrete random variables described by entropies $H(X)$ and $H(Y)$, mutual information $I(X;Y)$, conditional entropies $H(X|Y)$ and $H(Y|X)$, and joint entropy $H(X,Y)$. Using these quantities as the labels for sets and subsets, with $\cup$ and $\cap$, evaluate the following into simpler quantities: [5 marks]

(*i*) $H(X|Y) \cup I(X;Y)$

(*ii*) $(H(X) \cup H(Y)) \cap I(X;Y)$

(*iii*) $(H(X|Y) \cup H(Y|X)) \cap I(X;Y)$

(*iv*) $H(X,Y) - H(X|Y)$

(*v*) $H(X|Y) \cap H(Y|X)$

(*c*) Explain how dictionary coding as a data compression strategy exploits the sparseness of strings, that is, the fact that the space of possible combinations is only very sparsely populated with those that actually occur. In the case of encoding English text with a coding budget of just 15 bits, contrast the richness of encoding letter-by-letter versus defining 15-bit pointers to a lexicon. How much of English vocabulary can be captured by such a vector quantisation strategy? What added cost does this strategy incur? [5 marks]

(*d*) Continuous random variables X and Y both have uniform probability density distributions on some interval. For X, $p(x) = \frac{1}{2}$ if $x \in [0,2]$ else 0, while for Y, $p(y) = \frac{1}{8}$ if $y \in [0,8]$ else 0. Calculate the differential entropies $h(X)$ and $h(Y)$. Provide an upper bound on the joint entropy $h(X,Y)$, and state the condition for reaching it. [5 marks]