Relations Between Redundancy Patterns of the Shannon Code and Wave Diffraction Patterns of Partially Disordered Media

Neri Merhav

Department of Electrical Engineering
Technion—Israel Institute of Technology
Haifa 32000, Israel

The analysis of redundancy of lossless coding has been studied extensively: Capocelli & De Santis (‘92), Gallager (‘78), Jacquet & Szpankowski (‘95), Krichevsky (‘68), Louchard & Szpankowski (‘97), Savari & Gallager (‘97),... Szpankowski (‘00) analyzed the asymptotic (unnormalized) redundancy of the Shannon code, the Huffman code and others for a DMS, and discovered a weird behavior:

For a binary DMS, let $\alpha \triangleq \log_2 \frac{1-p}{p}$. Then, for the Shannon code

$$R_n \triangleq \mathbb{E}\{L(X_1, \ldots, X_n)\} - nH = \begin{cases} \frac{1}{2} + o(n) & \alpha \text{ irrational} \\ \text{oscillatory} & \alpha \text{ rational} \end{cases}$$

The frequency of the oscillations in the 2nd line – dictated by the integer denominator of $\alpha$. 
Q: What is the explanation for this erratic behavior?

Our purpose in this work is to try to provide some insight by drawing an analogy with the theory of wave diffraction:

- Perfect crystal (periodic lattice) $\Rightarrow$ Bragg peaks.
- Disordered medium $\Rightarrow$ continuous diffraction pattern (no Bragg peaks).

The first case is parallel to rational $\alpha$, and the second – to irrational $\alpha$. 
Bragg Diffraction

\[ 2d \sin \theta = n\lambda, \quad n = 0, 1, 2, \ldots \]
Bragg Diffraction (Cont’d)
The Hendricks–Teller (HT) Model (1942)

Assume that the distances between consecutive layers are selected independently at random from a finite set.

For example, distance $d_0$ w.p. $p$ and distance $d_1$ w.p. $1 - p$.

$d_0$ and $d_1$ are **commensurate** ($d_1/d_0$ – rational) $\Rightarrow$ $\exists$ wavelength $\lambda_1$ s.t. both $2d_0 \sin \theta$ and $2d_1 \sin \theta$ are multiples of $\lambda_1$: Bragg peaks (constructive interference) appear at all wavelengths $\lambda_n = \lambda_1/n$.

$d_0$ and $d_1$ incommensurable $\Rightarrow$ no such wavelengths exist – no Bragg peaks.

$p$ and $1 - p$ of the source are the same as those of the HT model.

$\alpha = \log_2 \frac{1-p}{p}$ is analogous to $d_1/d_0$.

In the oscillatory case, the fundamental frequency of the oscillations of $R_n$ is related to the fundamental wavenumber of the Bragg peaks.
The Main Common Mathematical Facts

Let $p_0, p_1, \ldots, p_{M-1}$ be probabilities, and define

$$C_m = p_0 + \sum_{j=1}^{M-1} p_j \exp\{2\pi i m \alpha_j\}, \quad \alpha_j \in \mathbb{R}. $$

- $|C_m| \leq 1$ for all $m$.
- $C_m = 1$ for some $m \neq 0$ if $\{\alpha_j\}$ are all rational. Otherwise, $|C_m| < 1 \ \forall \ m$.
- In the commensurate case, the smallest $m \neq 0$ for which $C_m = 1$ is the common denominator $m_0$ of all $\{\alpha_j\}$.
- $C_m = 1$ for multiples of $m_0$ and only for these integers.
Szpankowski derived the following asymptotic formula for the binary source whose extension to the $M$–ary source is as follows:

$$R_n = \begin{cases} \frac{1}{2} + \frac{1}{m_0} \left(\frac{1}{2} - \langle \beta m_0 n \rangle\right) + o(1) & \text{all } \{\alpha_j\} \text{ are rational} \\ \frac{1}{2} + o(1) & \text{otherwise} \end{cases}$$

where $\langle x \rangle = x - \lfloor x \rfloor = \text{fractional part of } x$,

$$\alpha_j = \log_2 \left(\frac{p_0}{p_j}\right),$$

$$\beta = -\log_2 p_0.$$  

and

$$m_0 = \text{common denominator of all } \{\alpha_j\}$$
Sketch of Analysis (Cont’d)

Consider the Fourier series expansion of the periodic function $\langle x \rangle$:

$$\langle x \rangle = \frac{1}{2} - \sum_{m \neq 0} \frac{\exp\{2\pi imx\}}{2\pi im}$$

Applying to $R_n = E[-\log_2 P(X_1, \ldots, X_n)] - nH$:

$$R_n = \frac{1}{2} + \sum_{m \neq 0} \frac{e^{-2\pi imn \log p_0}}{2\pi im} \left[ p_0 + \sum_{j=1}^{M-1} p_j \exp \left\{ 2\pi im \log \left( \frac{p_0}{p_j} \right) \right\} \right]^n$$

where

$$C_m = p_0 + \sum_{j=1}^{M-1} p_j \exp \left\{ 2\pi im \log \left( \frac{p_0}{p_j} \right) \right\}.$$
Sketch of Analysis (Cont’d)

\[ R_n = \frac{1}{2} + \sum_{m \neq 0} \frac{e^{2\pi imn\beta}}{2\pi im} \cdot (C_m)^n \]

- Not all \( \alpha_j = \log \frac{p_0}{p_j} \) rational:
  \[ \Rightarrow |C_m| < 1 \forall m \Rightarrow C_m^n \rightarrow 0 \Rightarrow R_n \rightarrow 1/2. \]

- All \( \alpha_j = \log \frac{p_0}{p_j} \) rational \( \Rightarrow C_{km_0} = 1: \)

\[ R_n \approx \frac{1}{2} + \sum_{k \neq 0} \frac{e^{2\pi ikm_0\beta n}}{2\pi ikm_0} \]

\[ = \frac{1}{2} + \frac{1}{m_0} \sum_{k \neq 0} \frac{e^{2\pi ik\beta m_0n}}{2\pi ik} \]

\[ = \frac{1}{2} + \frac{1}{m_0} \left( \frac{1}{2} - \langle \beta m_0n \rangle \right), \quad (1) \]

Fundamental frequency of oscillations: \( \omega_0 = 2\pi m_0\beta. \)
The Hendrick–Teller Model (1D)

Let the atoms be in locations $Z_0, Z_1, Z_2, \ldots, Z_{n-1}$ with spacings $\Delta_j = Z_j - Z_{j-1}$, $j = 1, 2, \ldots, n - 1$ being i.i.d. RV's taking values in \{ $d_0, \ldots, d_{M-1}$ \} with probabilities \{ $p_0, \ldots, p_{M-1}$ \}.

Each point (atom) at $Z_i$ contributes a scattered wave designated by the phasor $e^{-iqZ_j}$ with $q = 2\pi/\lambda$ being understood as a wavenumber. Thus, the superposition

$$U(q) = \sum_j e^{-iqZ_j}$$

gives rise to the structure function (intensity):

$$I(q) = \mathbb{E}\{|U(q)|^2\} = \sum_{k,\ell} \mathbb{E}\{e^{iq(Z_k-Z_\ell)}\} = n + I_0(q) + I_0^*(q)$$

with

$$I_0(q) = \sum_{k>\ell} \mathbb{E}\{e^{iq(Z_k-Z_\ell)}\} = \sum_{k>\ell} \mathbb{E}\{e^{iq\Delta_1}\}^{k-\ell}.$$
The intensity is then given by

\[
I(q) \approx n \cdot \frac{1 - |C(q)|^2}{|1 - C(q)|^2}.
\]
The Hendrick–Teller Model (Cont’d)

There are singularities (= Bragg peaks) if there are values of \( q \) with \( C(q) = 1 \). Let \( q_m = 2\pi m/d_0 \):

\[
C_m = C(q_m) = p_0 + \sum_{j=1}^{M-1} p_j e^{2\pi i m d_j / d_0},
\]

and we have the same \( C_m \) but now

\[
\alpha_j = \frac{d_j}{d_0},
\]

which are all rational iff \( \{d_j\} \) are commensurate.

Bragg peaks at all multiples of \( q_{m_0} = 2\pi m_0 / d_0 \).
Summarizing the Analogy

- Letter probabilities $\{p_j\} \Leftrightarrow$ Distance probabilities $\{p_j\}$.
- Log-probability ratios $\alpha_j = \log p_0/p_j \Leftrightarrow$ distance ratios $\alpha_j = d_j/d_0$.
- Oscillatory behavior or $R_n \Leftrightarrow$ Bragg peaks.
- Fund. frequency $\omega_0 = 2\pi m_0 \beta \Leftrightarrow$ fund. wavenumber $q_0 = 2\pi m_0/d_0$.

Ongoing work: Extension from i.i.d. to the Markov case (with W. Szpankowski).
Thank You!