CCIT Report #788 April 2011

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

Abstract

The average redundancy of the Shannon code, R_n, as a function of the block length n, is known to exhibit two very different types of behavior, depending on the rationality or irrationality of certain parameters of the source: It either converges to 1/2 as n grows without bound, or it may have a non–vanishing, oscillatory, (quasi–) periodic pattern around the value 1/2 for all large n. In this paper, we make an attempt to shed some insight into this erratic behavior of R_n, by drawing an analogy with the realm of physics of wave propagation, in particular, the elementary theory of scattering and diffraction. It turns out that there are two types of behavior of wave diffraction patterns formed by crystals, which are correspondingly analogous to the two types of patterns of R_n. When the crystal is perfect, the diffraction intensity spectrum exhibits very sharp peaks, a.k.a. Bragg peaks, at wavelengths of full constructive interference. These wavelengths correspond to the frequencies of the harmonic waves of the oscillatory mode of R_n. On the other hand, when the crystal is imperfect and there is a considerable degree of disorder in its structure, the Bragg peaks disappear, and the behavior of this mode is analogous to the one where R_n is convergent.

Index Terms: Lossless source coding, redundancy, Shannon code, scattering, diffraction, Bragg peaks, disorder.

1 Introduction

The analysis of the average redundancy of lossless codes for data compression schemes is a topic that attracted the attention of considerably many researchers throughout the history of Information Theory (cf. e.g., [1],[3],[6],[7],[8],[9],[10],[11],[12],[13] and many references therein).

In [13] Szpankowski has derived the asymptotic behavior of the average redundancy R_n, as a function of the block length n, for the Shannon code, the Huffman code, and other codes, focusing