APPROXIMATION WITH A FIDELITY CRITERION

J. WOLFOWITZ
CORNELL UNIVERSITY

1. Introduction

The present paper, as so many others in information theory, was stimulated by a paper of Shannon's [1]. The interesting theorem 1 below is due to him; the new result is theorem 2. We give a different proof of theorem 1. Actually this proof is not very new and is essentially the one used to prove theorem 1 of [3] (reproduced in [2] as theorem 3.2.1). The relation between the notion of "distortion" and that of "being generated" will be clear from this proof.

In the present paper we keep separate the ideas of approximating and coding. Then theorem 1 says essentially that, by embedding a certain number of sequences one can achieve a prescribed bound on the distortion, and theorem 2 says essentially that this cannot be done with fewer sequences. Shannon's results on coding are described in section 4. Some of his generalizations and additional suggestions for further generalizations are described in section 5.

It may perhaps be of interest to mention that theorem 4.9 of [4] is a special case of (4.3) below (the latter is theorem 1 of [1]). In fact, the probability of error defined in (4-65) of [4] is a special case of Shannon's distortion function ((2.1) below).

In [1], and in the present paper, the "source" digits (components of u below) are chance variables with a given (fixed) distribution. This is also true in the situation treated in theorem 4.9 of [4]. In the strong converse proved in [3], and in the others proved in [2], the messages are *not* stochastic and are chosen arbitrarily by the sender. If they should be chosen by a chance process their distribution can be arbitrary. The claims made in ([4], p. 219) on behalf of theorem 4.9 of [4] are therefore without the least basis in fact.

2. The approximating theorem

Consider the alphabets $M = \{m_1, \cdots, m_a\}$ and $Z = \{z_1, \cdots, z_b\}$. Let M^* (resp. Z^*) be the space of n-sequences (sequences of length n) in the M-alphabet (resp. the Z-alphabet). Let $\pi = (\pi_1, \cdots, \pi_a)$ be a probability a-vector which will be fixed in all that follows. When we speak of the probability distribution on M^*, we shall always mean the distribution implied by n independent chance variables with the common distribution π.

Research under contract with the Office of Naval Research.