The Empirical Distribution of Fourier Coefficients

by

David Freedman — University of California, Berkeley

and

David Lane — University of Minnesota

Technical Report No. 329

January, 1979

The Empirical Distribution of

Fourier Coefficients

by

David Freedman   University of California, Berkeley

and

David Lane   University of Minnesota

Technical Report No. 329

ABSTRACT

Suppose $X_1, X_2, \cdots$ are independent, identically distributed complex-valued L^2 random variables with $EX_1 = 0$ and $E(|X_1|^2) = 1$. Let Y_{nk} be the k^{th} Fourier coefficient of $X_1, \cdots, X_n$:

$$Y_{nk} = \sum_{j=1}^{n} X_j \exp\left(\frac{2\pi\sqrt{-1}\,kj}{n}\right).$$

Let μ_n be the empirical distribution of $\{n^{-\frac{1}{2}} Y_{nk}: k = 1, \cdots, n\}$. Then μ_n converges to the distribution of $U + iV$, where U and V are independent normal variables with mean 0 and variance $\frac{1}{2}$. This theorem is derived from a similar result for the Fourier coefficients of random permutations of the coordinates of z^n, where z^n is a vector with n coordinates such that $\max_k |z^n_k| = o(n^{-\frac{1}{2}})$, as $n \to \infty$.

I. Introduction

Suppose x is a vector in C^n, where C is the complex plane. That is, x has n coordinates, each a complex number. The empirical distribution of x is the probability measure on C which places mass n^{-1} on each coordinate of x; it will be denoted by μ_x. The discrete Fourier transform $\hat{x}$ is the vector in C^n whose coordinates are given by

$$\hat{x}_k = \sum_{j=1}^{n} x_j \exp\left(\frac{2\pi\sqrt{-1}\,kj}{n}\right),\ 1 \le k \le n.$$

The coordinates of $\hat{x}$ are the Fourier coefficients of x.

Now suppose $X_1,\cdots,X_n$ are independent complex-valued random variables with a common L^2-distribution and suppose x is an observation on $(X_1,\cdots,X_n)$. It seems to be a well-known fact, at least in the case that the X_i's are real valued, that normal probability plots of the real and imaginary parts of the coordinates of $\hat{x}$ tend to be close to linear. This phenomenon is discussed, for example, in Brillinger (1975, pp. 95-97) and Mallows (1969), and it is illustrated by some examples in the appendix to this paper.

If $X_1,\cdots,X_n$ have either real or complex normal distributions, there is a simple explanation for this phenomenon. (Recall that $Z = X + iY$ has a complex normal distribution if X and Y are independent real-valued normal variables with the same variance.) In the real case, the first $[\frac{n-1}{2}]$ Fourier coefficients of $(X_1,\cdots,X_n)$ are independent identically distributed complex normal variables, and $\hat{X}_i$ and $\hat{X}_{n-i}$ are conjugate for $1 \le i \le [\frac{n-1}{2}]$. In the complex case, the first n-1 Fourier coefficients have independent, identical complex normal distributions. Thus in both cases, if x is an observation on $(X_1,\cdots,X_n)$ for n reasonably large, the empirical distribution of $\hat{x}$ should be close to complex normal. Consequently, normal probability plots of the coordinates of Re $\hat{x}$ and Im $\hat{x}$ should be close to linear.

If the distribution of the X_i's is not normal, the situation is not so simple. There are theorems which establish the asymptotic joint normality of a fixed finite number of the Fourier coefficients of $(X_1,\cdots,X_n)$: see, for example, Brillinger (1975, Theorem 4.4.1). However, these theorems do not by themselves prove that the empirical distribution of all the Fourier coefficients will converge to a normal distribution. That seems to depend on the joint distribution of all n Fourier coefficients, which is hard to estimate. The purpose of this paper is to provide a rigorous mathematical proof for the asymptotic normality of the empirical distribution of the Fourier coefficients. This is the content of Theorem 2 below. We prove this theorem without characterizing the asymptotic joint distribution of all n Fourier coefficients.

Theorem 2 is a consequence of Theorem 1. Suppose z is a vector with complex coordinates, none of which has a particularly large modulus relative to the others. Consider all possible permutations of the coordinates of z. For each permutation, calculate the discrete Fourier transform. Theorem 1 shows that for most permutations, the empirical distribution of the Fourier coefficients will be close to complex normal. Theorem 1 applies in particular, of course, to the case in which z has only real coordinates.

A consequence of Theorems 1 and 2, pointed out in Theorem 3, is that the empirical distribution of the periodogram for data of the sort considered in the theorems should be close to exponential. This phenomenon has been observed empirically by Brillinger (1975, Figure 5.2.5., p. 127).

We want to thank Christopher Bingham and David Brillinger for several helpful discussions. Bingham suggested the problem to us. Sandy Weisberg helped prepare the plots appearing in the appendix.

2. Preliminaries

Let C_o be the set of continuous real-valued functions on C with compact support. Give C_o the sup norm (denoted $\| \ \|$), and let $f_1, f_2, \cdots$ be a dense countable subset of C_o. Let M denote the space of probability measures on C. Metrize M as follows: for μ, ν in M

$$d(\mu, \nu) = \sum_{i=1}^{n} \frac{|\int f_i d\mu - \int f_i d\nu|}{2^i \|f_i\|} .$$

d induces the weak topology on M, but M is not complete with respect to d. Let $\mathfrak{M}$ denote the σ-field on M generated by the weak open sets. The space of probabilities on C^k may be given the weak topology in a similar way.

A <u>random measure</u> on C is a measurable map from some probability space into $(M, \mathfrak{M})$. If μ is a random measure on $(\Omega, \mathcal{F}, P)$, and f a bounded Borel function on C, then $\int f d\mu$ is a random variable on $(\Omega, \mathcal{F}, P)$. The set function $E\mu$ is given by $E\mu(A) = \int \mu(A) dP$, for A a Borel subset of C. Thus, $E\mu$ is an element of M.

Suppose μ is a random measure satisfying $P(\mu = m) = 1$ for some m in M. Then μ is a <u>constant measure</u>, and the random measure μ will sometimes be identified with its value m. Lemma 1 provides a criterion for convergence in probability of a sequence of random measures to a constant measure; its easy proof is omitted.

Lemma 1: Suppose $\mu_1, \mu_2, \cdots$ are random measures on $(\Omega, \mathfrak{F}, P)$, and m is an element of M. Suppose for each f in C_o :

i) $E\int f d\mu_n \to \int f dm$

ii) $\mathrm{Var}(\int f d\mu_n) \to 0$.

Then μ_n converges in probability to m: that is, $P[d(\mu_n, m) > \epsilon] \to 0$ as $n \to \infty$ for each $\epsilon > 0$.

Lemma 2: Suppose m is an element of M. Let m^2 be the product of m with itself, a probability on C^2. Let $\{X_{nk}\}$ be an array of complex-valued random variables, with $n = 1, 2, \cdots$, and $k = 1, \cdots, k_n$. Let $m_{nk\ell}$ be the joint distribution of X_{nk} and $X_{n\ell}$. Suppose the following conditions are satisfied:

i) $k_n \to \infty$ as $n \to \infty$;

ii) $m_{nk\ell} \to m^2$ as $n \to \infty$, uniformly in pairs (k, ℓ), except for indices $nk\ell$ in an exceptional set E, with $\dfrac{\#\{(k,\ell): nk\ell \text{ in } E\}}{k_n^2} \to 0$ as $n \to \infty$.

For each n, let μ_n be the empirical distribution of $(X_{n1}, \cdots, X_{nk_n})$, so that μ_n is a random measure. Then μ_n converges in probability to m.

Proof: For f in C_o,

$$\int f d\mu_n = (k_n)^{-1} \sum_{k=1}^{k_n} f(X_{nk}),$$

so $$E\int f d\mu_n = (k_n)^{-1} \sum_{k=1}^{kn} \int f dm_{nk},$$

where m_{nk} is the distribution of X_{nk}. Now (ii) implies that $m_{nk} \to m$

as $n \to \infty$ uniformly in k, except for a set of k's with limiting density 0. Together with (i), this implies that

$$E\int f d\mu_n \to \int f dm.$$

Next,

$$E(\int f d\mu_n)^2 = (k_n)^{-2}[\sum_{k\neq \ell}\int f(x)f(y)\; m_{nk\ell}\,(dxdy) + \sum_{\ell}\int f^2 dm_{n\ell}]$$

so $\quad \mathrm{Var}(\int f d\mu_n) = T_1(n) + T_2(n)$, where

$$T_1(n) = (k_n)^{-2}\sum_{k\neq \ell}(\int f(x)f(y)m_{nk\ell}(dxdy) - \int f dm_{nk}\int f dm_{n\ell})$$

and

$$T_2(n) = (k_n)^{-2}\sum_{k=1}^{k_n}[\int f^2 dm_{nk} - (\int f dm_{nk})^2].$$

But $T_2(n)$ is bounded by $k_n^{-1}\|f\|^2$, which converges to 0 by (i). Furthermore, $T_1(n)$ converges to 0 by (ii). Thus, Lemma 1 implies that μ_n converges to m in probability. □

<u>Lemma 3</u>: Suppose X_{nk} and Y_{nk} are real-valued random variables, for $n = 1,2,\cdots$, and $k = 1,\cdots,k_n$, where $k_n \to \infty$ as $n \Rightarrow \infty$. Suppose that for each four-tuple of real numbers $\lambda_1,\cdots,\lambda_4$, the distribution of $\lambda_1 X_{nk} + \lambda_2 Y_{nk} + \lambda_3 X_{n\ell} + \lambda_4 Y_{n\ell}$ converges to $N(0, \frac{1}{2}(\lambda_1^2 + \cdots + \lambda_4^2))$ as $n \to \infty$, uniformly in pairs (k,ℓ), except for indices $nk\ell$ in an exceptional set E with $\frac{\#\{(k,\ell):\, nk\ell \text{ in } E\}}{k_n^2} \to 0$ as $n \to \infty$. Set $Z_{nk} = X_{nk} + \sqrt{-1}\,Y_{nk}$, and let μ_n be the empirical distribution of $(Z_{n1},\cdots,Z_{nk_n})$. Then μ_n converges in probability to the standard complex normal distribution (that is, the distribution of $Z = U + iV$ where U and V are independent real-valued normal variables with mean 0 and variance $\frac{1}{2}$).

Proof: Let γ be the distribution of four independent normal random variables, each with mean 0 and variance $\frac{1}{2}$. Let $\gamma_{nk\ell}$ denote the distribution of the random vector $(X_{nk}, Y_{nk}, X_{n\ell}, Y_{n\ell})$, for $n = 1,2,\cdots$, and $k,\ell = 1,\cdots,k_n$. Now, for all indices $nk\ell$ not in E, order the distributions $\gamma_{nk\ell}$ to form a single sequence p_r, $r = 1, 2, \cdots$, in such a way that if $\gamma_{nk\ell}$ corresponds to p_j and γ_{stv} corresponds to p_k, then $n > s$ implies $j > k$. Clearly, the characteristic function of p_r converges pointwise to the characteristic function of γ. Thus, if g is a metric on probabilities on R^4 inducing the topology of weak convergence, then $g(p_r,\gamma) \to 0$. But because of the ordering of the sequence $\{p_r\}$, this implies that $g(\gamma_{nk\ell},\gamma) \to 0$ as $n \to \infty$ uniformly in pairs (k,ℓ) with $nk\ell$ not in E.

Now let $m_{nk\ell}$ denote the distribution of $(Z_{nk},Z_{n\ell})$ for $n = 1,2,\cdots$, $k,\ell = 1,\cdots,k_n$. Let m denote the standard complex normal distribution. $m_{nk\ell}$ is of course determined by $\gamma_{nk\ell}$, and so $m_{nk\ell}$ converges to m^2 as $n \to \infty$ uniformly in pairs (k,ℓ) with $nk\ell$ not in E. By Lemma 2, then, μ_n converges in probability to m. □

Suppose x is a vector with n coordinates. Each permutation ψ on $\{1,\cdots,n\}$ yields a new vector x_ψ with coordinates $(x_\psi')_i = x_{\psi(i)}$. In the statement of the next lemma, Φ denotes the distribution function of a real-valued normal variable with mean 0 and variance 1.

Lemma 4: Suppose x, y, a, and b are vectors in R^n satisfying:

$$\text{i)}\quad \sum_{j=1}^{n} x_j = \sum_{j=1}^{n} y_j = \sum_{j=1}^{n} a_j = \sum_{j=1}^{n} b_j = 0$$

$$\text{ii)}\quad \sum_{j=1}^{n} (x_j^2 + y_j^2) = n,$$

$$\sum_{j=1}^{n} a_j^2 = \sum_{j=1}^{n} b_j^2 = n, \text{ and}$$

$$\sum_{j=1}^{n} a_j b_j = 0.$$

Set $V = \max\{|a_j|, |b_j|: 1 \le j \le n\}$ and $U = \max\{|x_j|, |y_j|: 1 \le j \le n\}$. Let ρ be a random permutation on $\{1,\cdots,n\}$ taking on any particular permutation with probability $1/n!$. Set $W = n^{-\frac{1}{2}} [\sum_{j=1}^{n} (x_{\rho(j)} a_j + y_{\rho(j)} b_j)]$. If F is the distribution function of W then

$$\sup_{x \in R} | F(x) - \Phi(x) | \le 48\, n^{-\frac{1}{2}} V\ U + h(n),$$

where h does not depend on x, y, a, or b, and $h(n) \to 0$ as $n \to \infty$.

Proof: This lemma can be derived by straightforward calculations from Corollary 3.2 of Ho and Chen (1978). □

3. Theorems

The main result of this section is Theorem 1, which asserts that for vectors z in C^n, the empirical distribution of the Fourier coefficients of z_ψ is close to complex normal, for most permutations ψ. For each integer n, let $z^n = (z_1^n, \ldots, z_n^n)$ be a vector in C^n. Let ρ be a random permutation of $\{1, \ldots, n\}$, taking on any particular permutation with probability $1/n!$. Let $Z_k^n = n^{-\frac{1}{2}} \sum_{j=1}^{n} z_{\rho(j)}^n \exp\left(\frac{2\pi\sqrt{-1}\, kj}{n}\right)$. Let μ_n be the empirical distribution of $(Z_1^n, \ldots, Z_n^n)$.

Theorem 1: Suppose the sequence $\{z^n\}$ satisfies:

i) $\sum_{k=1}^{n} z_k^n = 0$ and $\sum_{k=1}^{n} |z_k^n|^2 = n$

ii) $\max_{1\le k\le n} |z_k^n| = o(n^{-\frac{1}{2}})$ as $n \to \infty$.

Then μ_n converges in probability to the standard complex normal distribution.

Proof: The idea of the proof is to use Lemma 3 with $X_{nk} = \mathrm{Re}(Z_k^n)$, $Y_{nk} = \mathrm{Im}(Z_k^n)$ and $k_n = n$. The exceptional set E consists of all indices $nk\ell$ such that either,

a) $k = n$ or $\ell = n$

b) $k = n - \ell$ or

c) $k = \ell$.

Then $\frac{\#\{(k,\ell): nk\ell \text{ in } E\}}{n^2} \le \frac{4n}{n^2} \to 0$ as $n \to \infty$.

Fix the real numbers $\lambda_1, \cdots, \lambda_4$, and set $\lambda^2 = \lambda_1^2 + \cdots + \lambda_4^2$. Fix k and ℓ between 1 and n, with $nk\ell$ not in E. For the rest of the proof, for notational convenience, drop the superscript n from z^n. Let $z = x + iy$. Then

$$\lambda_1 X_{nk} + \lambda_2 Y_{nk} + \lambda_3 X_{n\ell} + \lambda_4 Y_{n\ell} =$$
$$n^{-\frac{1}{2}} \sum_{j=1}^{n} (x_{\rho(j)} \; [\lambda_1 \cos(\tfrac{2\pi kj}{n}) + \lambda_2 \sin \tfrac{2\pi kj}{n} + \lambda_3 \cos \tfrac{2\pi \ell j}{n} + \lambda_4 \sin \tfrac{2\pi \ell j}{n} \;])$$
$$+ n^{-\frac{1}{2}} \sum_{j=1}^{n} y_{\rho(j)} \; [-\lambda_1 \sin \tfrac{2\pi kj}{n} + \lambda_2 \cos \tfrac{2\pi kj}{n} - \lambda_3 \sin(\tfrac{2\pi \ell j}{n}) + \lambda_4 \cos(\tfrac{2\pi \ell j}{n})].$$

So,

$$\lambda_1 X_{nk} + \lambda_2 Y_{nk} + \lambda_3 X_{n\ell} + \lambda_4 Y_{n\ell} = (\tfrac{\lambda^2}{2})^{\frac{1}{2}} n^{-\frac{1}{2}} \sum_{j=1}^{n} (x_{\rho(j)}\, a_j + y_{\rho(j)}\, b_j),$$

where for $1 \le j \le n$,

$$a_j = (\tfrac{\lambda^2}{2})^{-\frac{1}{2}} [\lambda_1 \cos(\tfrac{2\pi kj}{n}) + \lambda_2 \sin(\tfrac{2\pi kj}{n}) + \lambda_3 \cos(\tfrac{2\pi \ell j}{n}) + \lambda_4 \sin(\tfrac{2\pi \ell j}{n})]$$

and

$$b_j = (\tfrac{\lambda^2}{2})^{-\frac{1}{2}} [-\lambda_1 \sin(\tfrac{2\pi kj}{n}) + \lambda_2 \cos(\tfrac{2\pi kj}{n}) - \lambda_3 \sin(\tfrac{2\pi \ell j}{n}) + \lambda_4 \cos(\tfrac{2\pi \ell j}{n})].$$

Note that x, y, a, and b satisfy the conditions of Lemma 4, and

$$\sup_{1\le j\le n}(|a_j|, |b_j|) \le (\tfrac{\lambda^2}{2})^{-\frac{1}{2}} (|\lambda_1| + |\lambda_2| + |\lambda_3| + |\lambda_4|) \le 2\sqrt{2}.$$

Thus, if F is the distribution of $(\frac{\lambda^2}{2})^{-\frac{1}{2}}(\lambda_1 X_{nk} + \lambda_2 Y_{nk} + \lambda_3 X_{n\ell} + \lambda_4 Y_{n\ell})$, Lemma 4 allows us to conclude

$$\sup_{x \in R} |F(x) - \Phi(x)| \le 96\sqrt{2}\,(n^{-\frac{1}{2}} \max_{1 \le k \le n} |z_k^n|) + h(n)$$

The right hand side tends to 0 as $n \to \infty$ by condition ii) and does not depend on k and ℓ. Thus the distribution function of $\lambda_1 X_{nk} + \lambda_2 Y_{nk} + \lambda_3 X_{n\ell} + \lambda_4 Y_{n\ell}$ converges uniformly as $n \to \infty$ to the distribution function of a real-valued normal variable with mean 0 and variance $\frac{\lambda^2}{2}$, and this convergence is uniform for pairs (k, ℓ) with $nk\ell$ not in E. This implies that the distribution of $\lambda_1 X_{nk} + \lambda_2 Y_{nk} + \lambda_3 X_{n\ell} + \lambda_4 Y_{n\ell}$ converges to $N(0, \frac{\lambda^2}{2})$ as $n \Rightarrow \infty$, uniformly in pairs (k, ℓ) with $nk\ell$ not in E. Now apply Lemma 3 to conclude that μ_n converges in probability to the standard complex normal distribution. □

The following corollary will be used to prove theorem 2.

<u>Corollary 1</u>: Suppose the sequence z^n satisfies:

i) $\sum_{k=1}^{n} z_k^n = o(n)$ and $\sum_{k=1}^{n} |z_k^n|^2 = n + o(n)$

as $n \to \infty$

ii) $\max_{1 \le k \le n} |z_k^n| = o(n^{-\frac{1}{2}})$ as $n \to \infty$.

Let μ_n correspond to z^n as in Theorem 1. Then μ_n converges in probability to the standard complex normal distribution.

Proof: Let $a_n = 1/n \sum_{h=1}^{n} z_k^n$ and $s_n^2 = 1/n \sum_{h=1}^{n} |z_k^n - a_n|^2$.

Then $a_n \to 0$ and $s_n \to 1$ as $n \to \infty$. Apply Theorem 1 to $\frac{(z_i^n - a_n)}{s_n}$. $\square$

Theorem 2: Suppose $X_1, X_2, \cdots$ are independent identically distributed complex-valued random variables with $E(X_1) = 0$ and $E|X_1|^2 = 1$. For each n, let $\hat{X}_n$ be the discrete Fourier transform of $(X_1, \cdots, X_n)$, and let μ_n be the empirical distribution of $n^{-\frac{1}{2}} \hat{X}_n$. Then μ_n converges in probability to the standard complex normal distribution.

Proof: Suppose $X_1, X_2, \cdots$ are defined on the probability triple $(\Omega, \mathfrak{F}, P)$. Let the triple $(\Omega', \mathfrak{F}', P')$ support a sequence $\rho_1, \rho_2, \cdots$ independent of $(X_1, X_2, \cdots)$, where ρ_n is a random permutation of $\{1, \cdots, n\}$ taking on each permutation with probability $1/n!$. A typical point of Ω' will be denoted ω'.

Now consider $(\Omega \times \Omega', \mathfrak{F} \times \mathfrak{F}', P \times P')$. On this product space, let $Y_n(\omega, \omega')$ be the discrete Fourier transform of

$$X_{\rho_n(\omega',1)}(\omega), \cdots, X_{\rho_n(\omega',n)}(\omega).$$

Let $\nu_n(\omega,\omega')$ be the empirical distribution of $n^{-\frac{1}{2}} Y_n(\omega,\omega')$. Fix $\varepsilon > 0$. Let m be the standard complex normal distribution. Let

$$Z_n(\omega) = P'\{\omega': d(\nu_n(\omega,\omega'), m) > \varepsilon\}.$$

Now, for almost all ω,

$$X_1(\omega) + \dots + X_n(\omega) = o(n)$$

$$X_1^2(\omega) + \dots + X_n^2(\omega) = n + o(n)$$

and

$$\max_{1 \le k \le n} |X_k(\omega)| = o(n^{-\frac{1}{2}}).$$

The corollary to theorem 1 implies that $Z_n(\omega) \to 0$ for almost all ω. By Fubini and dominated convergence,

$$(P \times P')\{d(\nu_n, m) > \epsilon\} = E(Z_n) \to 0.$$

Thus ν_n converges in probability to the standard complex normal distribution. Finally, note that the law of μ_n coincides with the law of ν_n. □

For a vector z in C^n, let $\hat{z}_k$ denote the k^{th} Fourier coefficient. The k^{th}/ <u>periodogram ordinate</u> is defined by

$$I(k) = (1/n)|\hat{z}_k|^2.$$

Under the conditions of Theorem 1 or Theorem 2, the empirical distribution of $n^{-\frac{1}{2}}\hat{z}$ is approximately standard complex normal: that is, the real and imaginary parts of the distribution are approximately independent $N(0, \frac{1}{2})$. Thus, if each mass point is squared and the moduli of the corresponding real and imaginary parts are summed, the resulting empirical distribution is approximately exponential with parameter 1. Formally:

<u>Theorem 3</u>: Under the conditions of Theorem 1 or Theorem 2, the empirical distribution of the periodogram ordinates converges in probability to an exponential distribution with parameter 1.

This theorem provides an explanation for the linearity of the χ_2^2-probability plot of 500 periodogram ordinates noted by Brillinger (1975, p. 127).

4. Notes and Questions

1) Consider the n x n matrix F with entries $F_{jk} = n^{-\frac{1}{2}} \exp(\frac{2\pi\sqrt{-1}kj}{n})$. F is a unitary matrix, and for y in C^n, $n^{-\frac{1}{2}}\hat{y} = Fy$. The questions considered in this paper about the coordinates of Fy may be raised with arbitrary unitary matrices H, and in fact the theorems of section 3 generalize immediately with the Fourier transformation replaced by arbitrary unitary transformation. That is, suppose for each n, H^n is an n x n unitary matrix, with $\max_{i,j}|H^n_{ij}| \le cn^{-\frac{1}{2}}$, where c is a constant which does not depend on n. For a triangular array X_{nk}, $n \ge 1$, $1 \le k \le n$, consider $y_n = \sqrt{n}\, H^n x_n$. Then the theorems of section 3 hold if $\hat{x}_n$ is replaced by y_n. Thus, broadly speaking, unitary transformations take arbitrary vectors into vectors with approximately normal empirical distribution. Some other aspects of this "normality-inducing" behavior of unitary transformations have been considered by Mallows (1969).

2) In the setting of Theorem 3, do the empirical distributions μ_n converge almost surely? We have not been able to settle this question yet. Another related question of some statistical interest is to determine the distribution of the largest Fourier coefficient. Gersho, Gopinath and Odlyzko (1978), building on theoretical work of Halasz(1973), have shown that if $X_1,\cdots,X_n$ are independent with the same L^6-distribution and $Var(X_i) = 1$, the maximum Fourier coefficient is with high probability close to $\sqrt{n \log n}$. Can this result be extended to more general L^2-distributions?

3) Suppose X_1, X_2, are independent complex-valued random variables with a common distribution, but $EX_1^2 = \infty$. In this case, is there any sequence c_n, such that if μ_n is the empirical distribution of $\{c_n Y_{nk}: h = 1,\cdots,n\}$, where $Y_{nk} = \sum_{j=1}^{n} X_j \exp(\frac{2\pi\sqrt{-1}\, kj}{n})$, then μ_n converges to some distribution?

If the X's have the distribution of a symmetric stable law of order p, $0 < p < 2$, then such μ_n cannot converge to a constant measure. However, if $c_n = n^{-1/p}$, we can show that the corresponding μ_n converge to a random measure, and plan to discuss this in a future paper.

References

Brillinger, David R. (1975), Time Series Data Analysis and Theory (New York, Holt, Rinehard, and Winston).

Gersho, A., Gopinath, B. and Odlyzko, A.M. (1978), "Coefficient Inaccuracy in Transversal Filtering," Bell Laboratories Technical Memorandum.

Halasz, G. (1973), "On a Result of Salem and Zygmund concerning Random Polynomials", Studia Scient. Math Hung., Vol. 8, pp. 369-377.

Ho, Soo-Thong and Chen, Louis H.Y. (1978), "An L_p bound for the remainder in a combinatorial central limit theorem," Annals of Probability, Vol. 6, pp. 231-249.

Mallows, Colin L. (1969), "Joint normality induced by orthogonal transformations," Bell Telephone Laboratories Technical Memorandum.

Appendix

This appendix presents some rankit plots to illustrate the results of Theorem 2. In each plot, the expected values of the order statistics from an appropriate-sized normal sample are plotted along the x-axis, while various data or Fourier coefficients are plotted along the y-axis.

Figure 1 - Normal data: A pseudorandom sample of 1000 observations from the standard normal distribution was generated. Figure 1a) is a rankit plot of the original data while Figures 1b) and 1c) are rankit plots of the real and imaginary parts respectively of the second through 499th Fourier coefficients of the data. Of course, all three of these plots are linear. They are included for comparison with the next three examples.

Figure 2 - Exponential Data: The arrangement is the same as Figure 1, but here the pseudorandom sample of size 1000 is from an exponential distribution with parameter 1.

Figure 3 - Uniform Data: Again, the arrangement is the same as Figure 1, but the data is a pseudorandom sample of size 1000 from a uniform distribution on [0,1].

Figure 4 - Cauchy Data: Here the data is a pseudorandom sample of size 1000 from a Cauchy distribution. The situation is quite different from the L^2 distributions illustrated in the preceding three figures. Different Cauchy samples lead to differently shaped plots of the Fourier coefficients, but the S-shape which appears in Figures 4b) and 4c) is common. It indicates a short-tailed empirical distribution, and is probably accounted for by the fact that the Cauchy sample has one dominating extreme point, so that the Fourier transform is essentially a cosine wave.

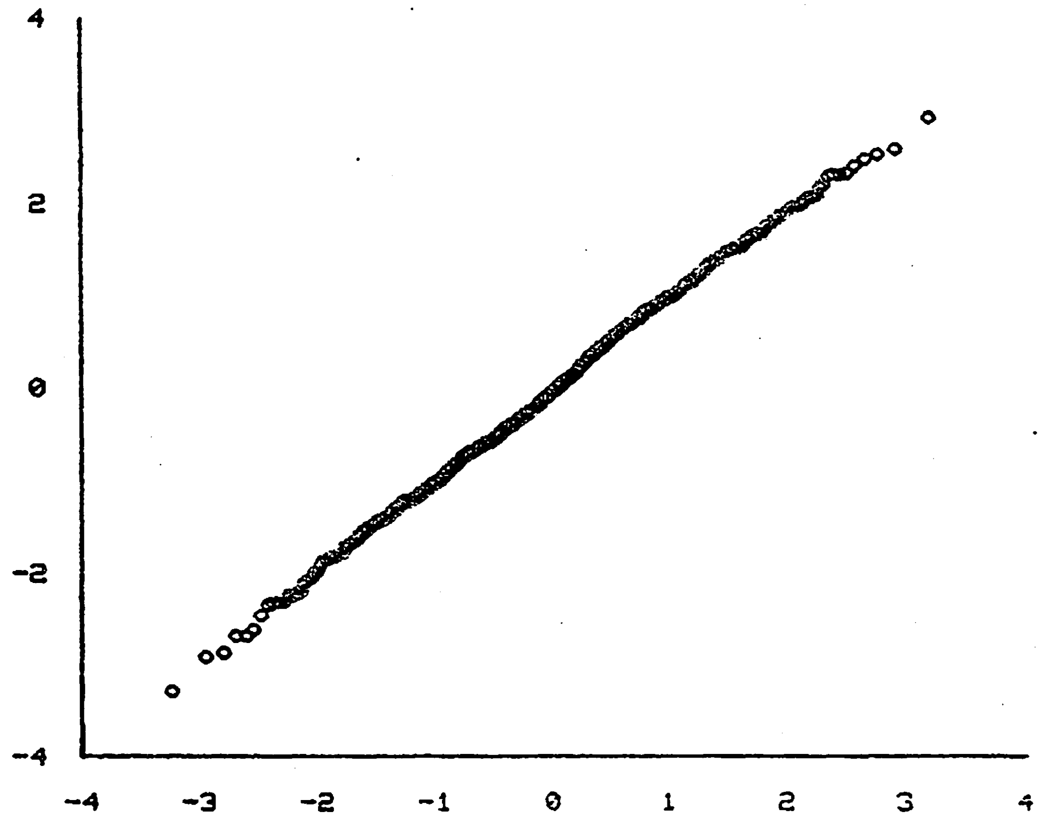

Figure 1a): Y-axis: Pseudorandom normal sample of size 1000
X-axis: Rankits

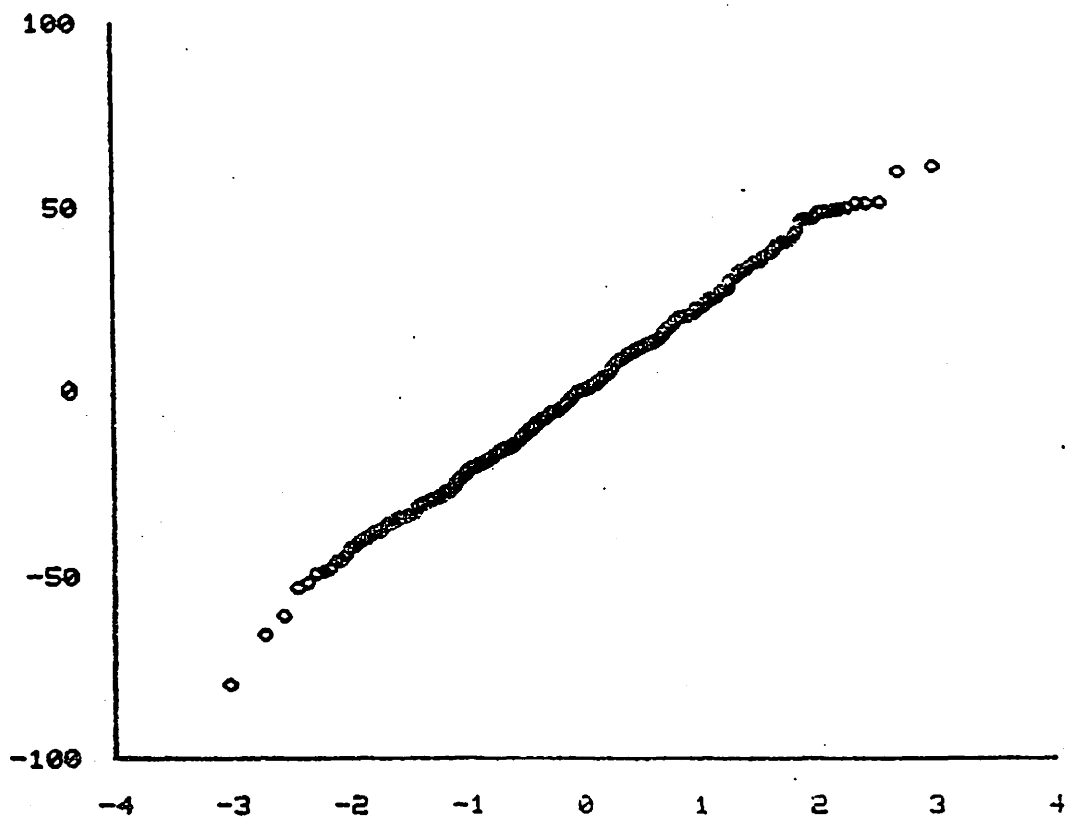

Figure 1b): Y-axis: Real parts of second through 499th Fourier coefficients of data plotted in 1a).

X-axis: Rankits

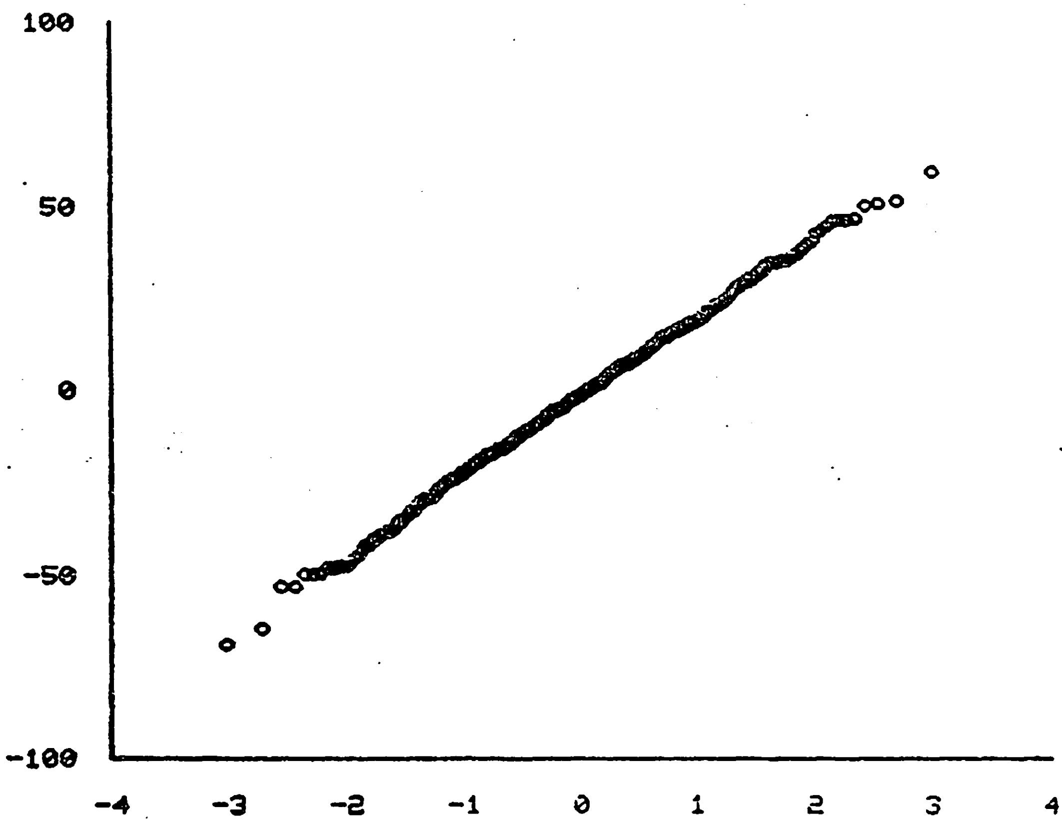

Figure 1c): Y-axis: Imaginary parts of second through 499th Fourier coefficients of data plotted in 1a).

X-axis: Rankits

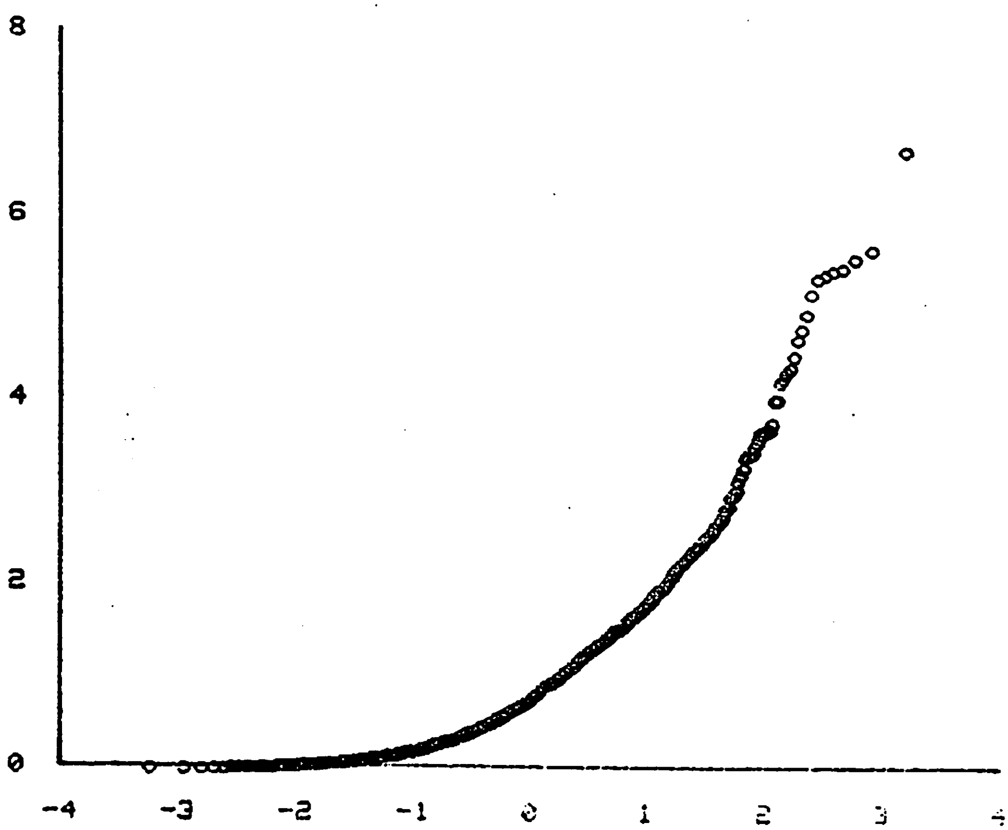

Figure 2a): Y-axis: Pseudorandom sample of size 1000 from an exponential distribution with parameter 1.

X-axis: Rankits

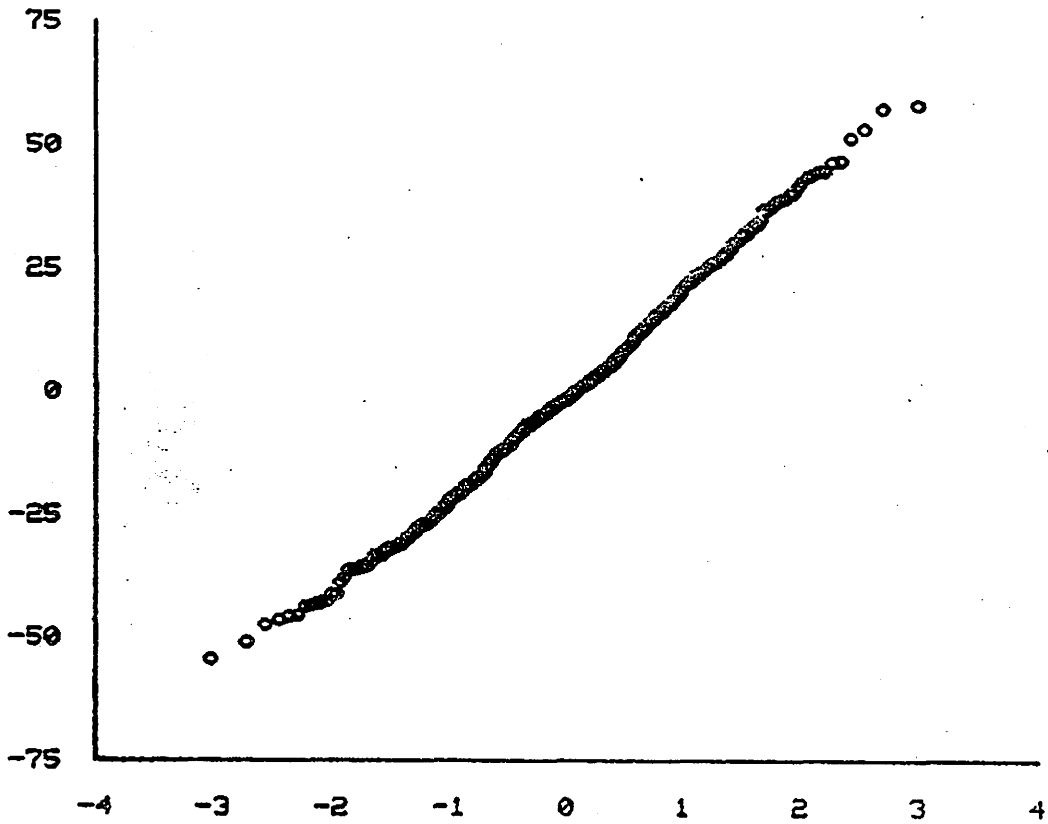

Figure 2b): Y-axis: Real parts of second through 499th Fourier coefficients of data plotted in 2a).

X-axis: Rankits

Note: Since the variance of an exponential (1) variate is 1, the normalization required by Theorem 2 for the coefficients would be to divide each of them by $\sqrt{500} = 23.6$ to obtain an approximately standard normal empirical distribution.

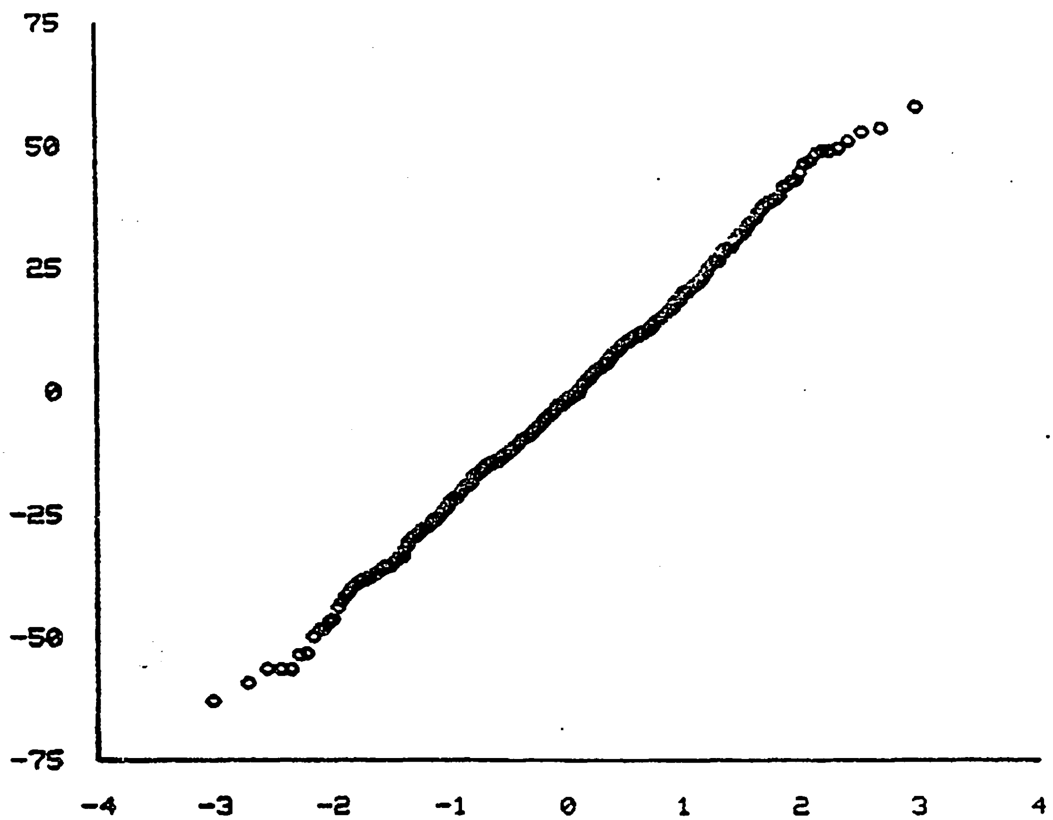

Figure 2c): Y-axis: Imaginary parts of second through 499th Fourier coefficients of data plotted in 2a).

X-axis: Rankits

(see Note on Figure 2b)).

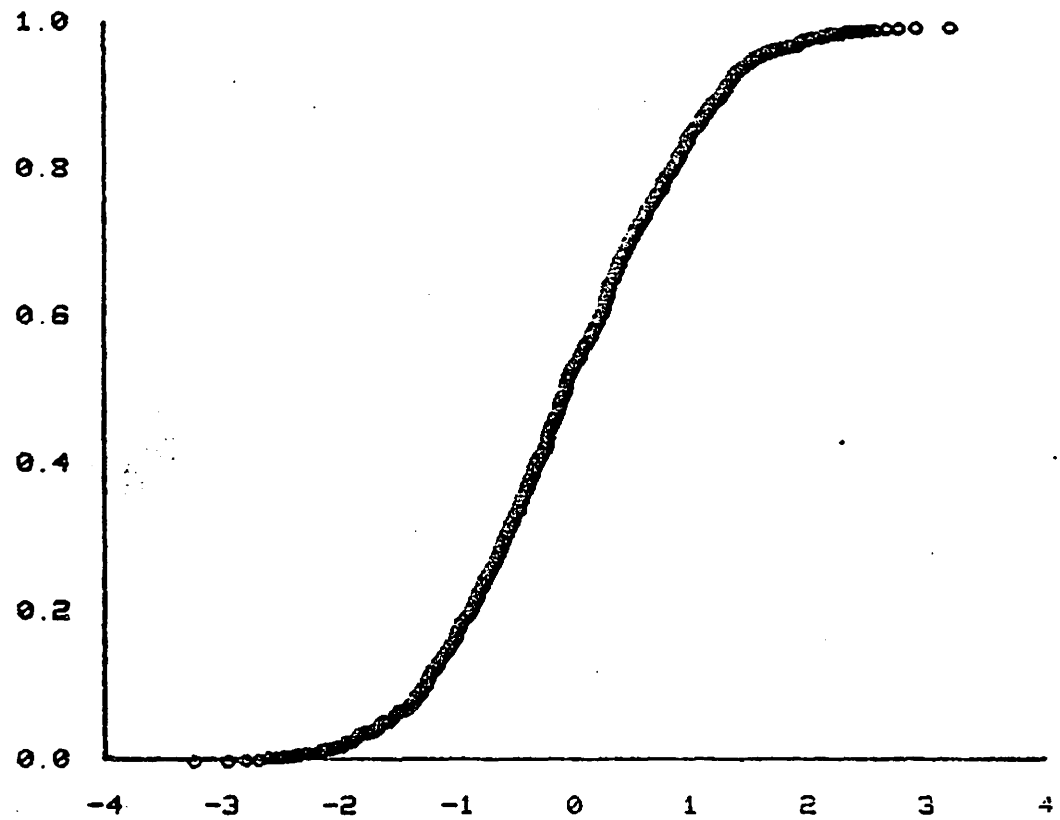

Figure 3a): Y-axis: Pseudorandom sample of size 1000 from a uniform distribution on [0,1].

X-axis: Rankits

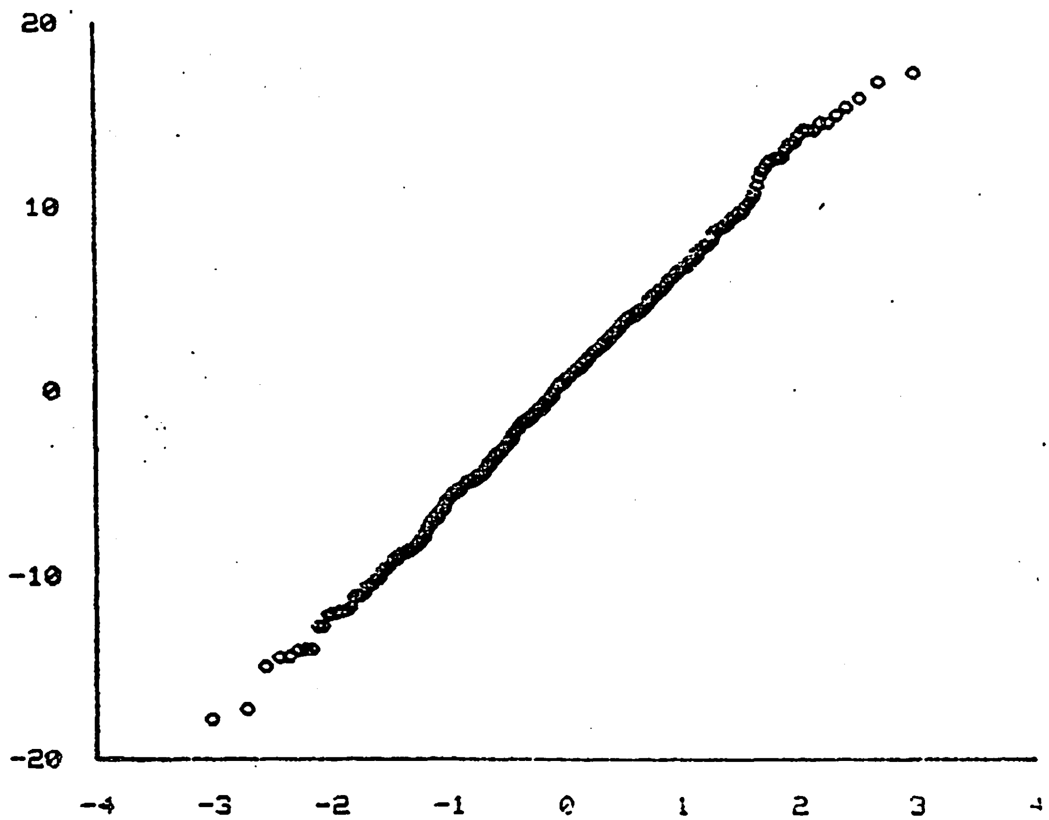

Figure 3b): Y-axis: Real parts of second through 499th Fourier coefficients of data plotted in 3a).

X-axis: Rankits

Note: Since the variance of a U[0,1] variate is 1/12, the normalization required by Theorem 2 for coefficients would be to divide each of them by $\sqrt{\frac{1500}{12}} = 6.45$.

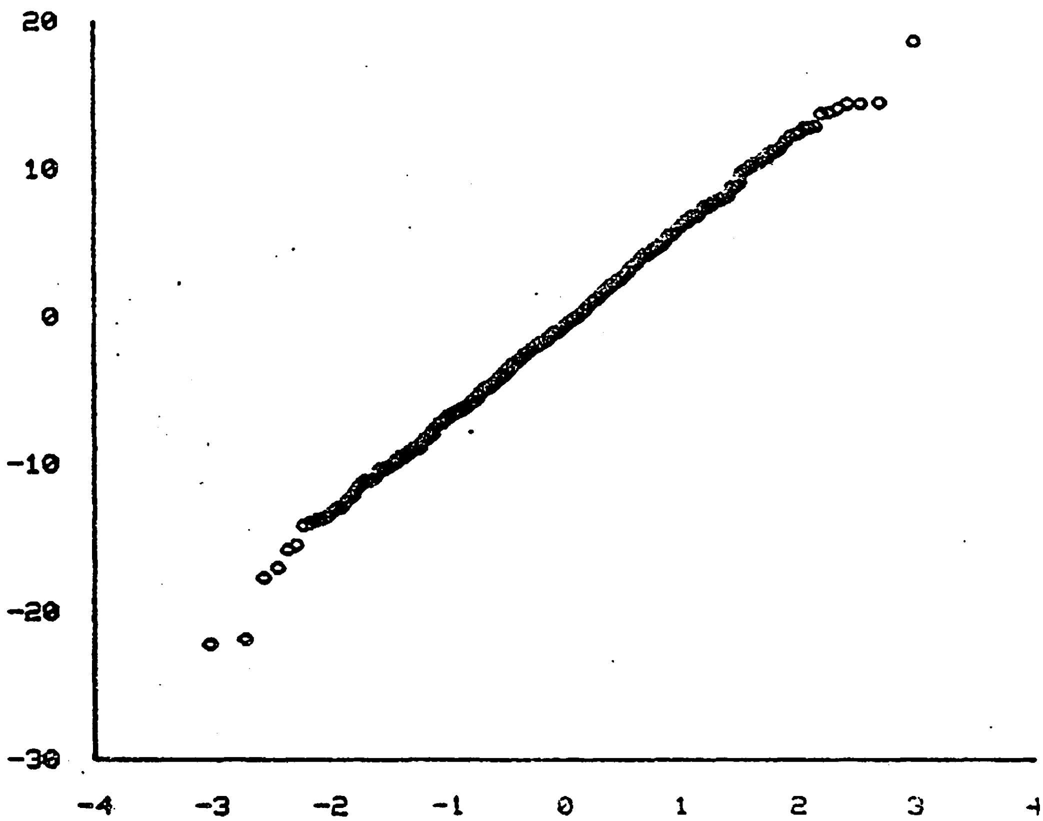

Figure 3c): Y-axis: Imaginary parts of second through 499th Fourier coefficients of data plotted in 3a).

X-axis: Rankits

(See Note on Figure 3b)).

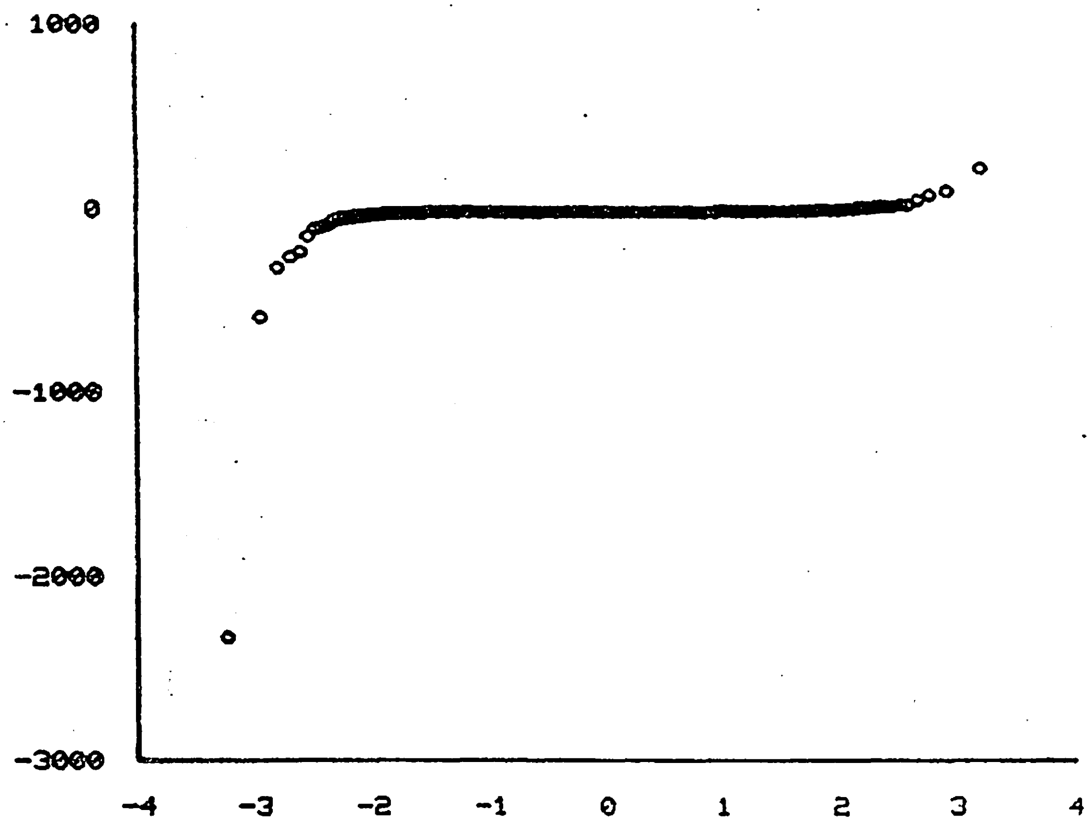

Figure 4a): Y-axis: Pseudorandom sample of size 1000 from Cauchy data.

X-axis: Rankits

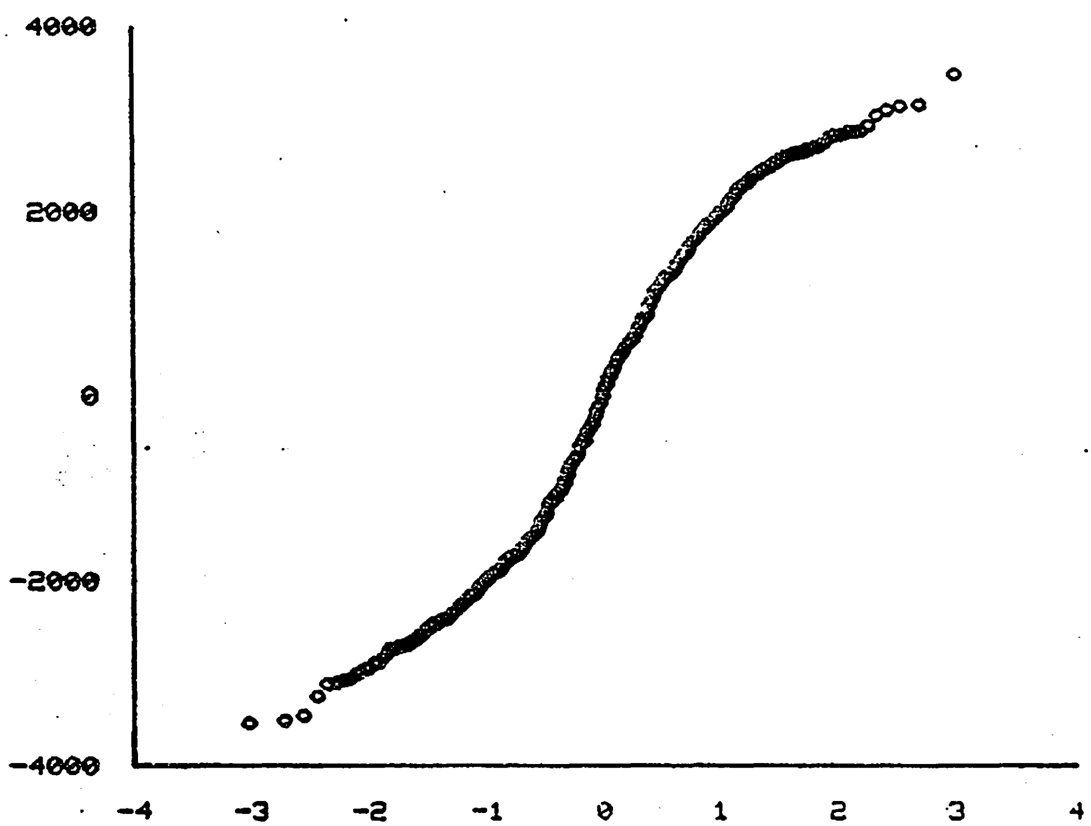

Figure 4b): Y-axis: Real parts of second through 499th Fourier coefficients of data plotted in 4a).

X-axis: Rankits

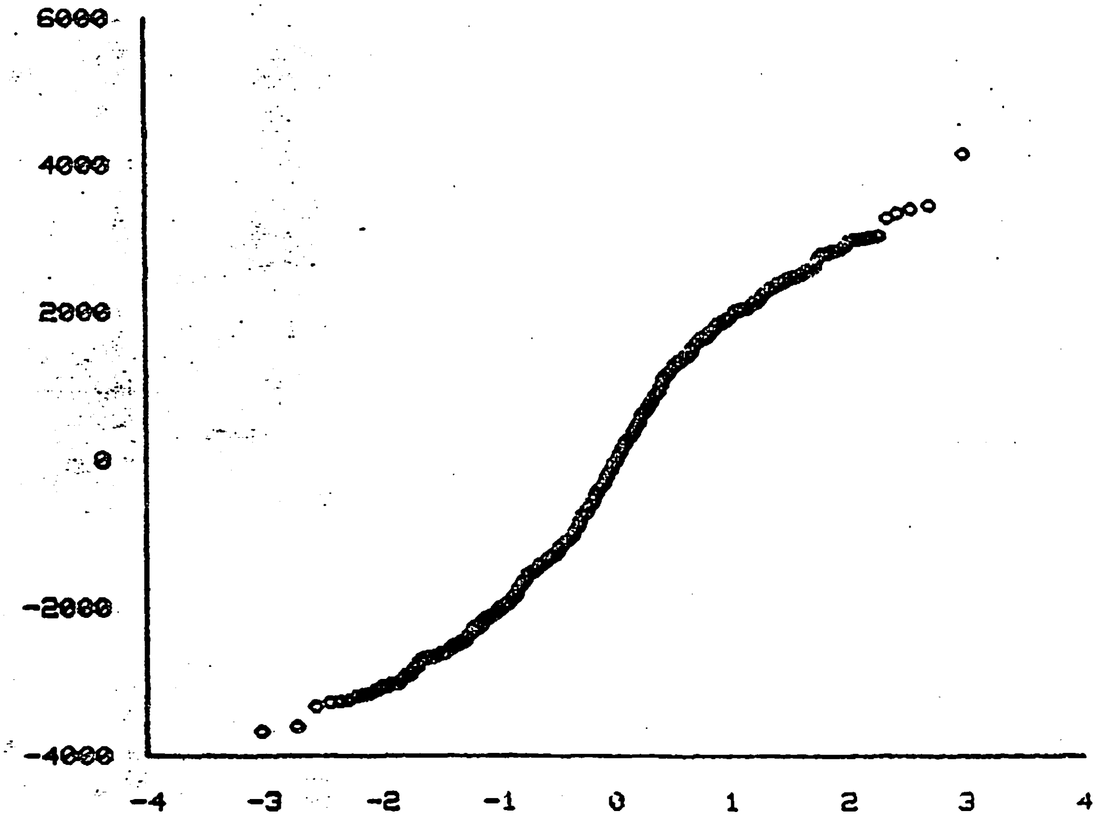

Figure 4c): Y-axis: Imaginary parts of second through 499th Fourier coefficients of data plotted in 4a).

X-axis: Rankits