#366

The Empirical Distribution of The Fourier Coefficients of a Sequence Of Independent, Identically Distributed Long-Tailed Random Variables

by

David Freedman, University of California*

and

David Lane, University of Minnesota

revised March 1981

*Research partially supported by National Science Foundation Grant MCS 80-02535

Abstract

Suppose $X_1, X_2, \cdots$ are independent, identically distributed random variables, and suppose $n^{-1/\alpha}(X_1 + \cdots + X_n)$ converges in distribution to a symmetric stable law of index $\alpha < 2$. For $s = 1, \cdots, n$, set

$$Y_{ns} = n^{-1/\alpha} \Sigma_{j=1}^{n} X_j \cos(2\pi js/n).$$

Let μ_n be the empirical distribution of $\{Y_{ns}: s = 1, \cdots, n\}$. Then μ_n converges in distribution, but not in probability.

1. Introduction

In an unpublished Bell Labs memorandum [5], Colin Mallows noted an interesting empirical phenomenon: the normality-inducing behavior of orthogonal transformations. If X is a random vector with independent coordinates and H an orthogonal matrix, then the coordinates of HX "behave in some ways like members of a random normal sample." This idea was taken up by others, and some empirical work suggests that the phenomenon might occur even when the distribution of the coordinates of X were far from normal. In particular, investigators have reported that, for standard Cauchy coordinates and Hadamard H, normal probability plots of the coordinates of Y appeared linear.

In [4], we began an investigation of one aspect of this phenomenon and showed that if the coordinates of X were identically distributed L^2 random variables, then the empirical distribution of the coordinates of Y tended with high probability to be close to the normal distribution. The proof depended strongly on the L^2-ness of the coordinates of X, but we wondered whether the result might still hold even if the coordinates of X had long tails. The mathematics of the problem became much more complicated in this setting, and so we restricted attention to Fourier coefficients. This paper presents our findings on this problem. In brief, the reported empirical results seem to be in conflict with the asymptotic theory. The empirical distribution of the coefficients does not converge in probability: there is a weak limit, but the limit does not concentrate on the normal distributions; and the scaling for the long-tailed X is quite different. Similar conclusions apply to the Hadamard case.

To describe the asymptotic theory, we first give a formal statement of the theorem in [4]. Let R denote the real line and $\mathbb{C}$ the complex plane. Let n be a positive integer, and $i = \sqrt{-1}$. Suppose $x = (x_1, \cdots, x_n)$ is a vector in R^n. The <u>discrete Fourier transform</u> $\hat{x}$ is the vector in $\mathbb{C}^n$ whose coordinates are given by

$$\hat{x}_s = \Sigma_{j=1}^n \exp(2\pi ijs/n)\, x_j \quad \text{for} \quad s = 1, \cdots, n.$$

Here, $\exp(x) = e^x$. The coordinates of $\hat{x}$ are the <u>Fourier coefficients</u> of x.

<u>Theorem</u>: Suppose $X_1, X_2, \cdots$ are independent, identically distributed random variables with mean 0 and variance 1. Let μ_n be the empirical distribution of $(Y_{n1}, \cdots, Y_{nn})$, where $\sqrt{n}\, Y_{ns}$ is the s^{th} Fourier coefficient of $(X_1, \cdots, X_n)$. Then μ_n converges in probability to a complex normal measure.

Now suppose the common distribution of the X_i's is in the domain of attraction of a symmetric stable law with parameter less than 2. We found that the transforms of these long-tailed variables behave very differently from the transforms of the L^2 - variables. Theorem (47) shows that the law of the empirical distributions of the Fourier coefficients (when properly normalized) does converge, but the limiting distribution is a nondegenerate measure on the set of probability measures on the complex plane. This limit law depends upon the index of the stable law attracting the X_i's. Proposition (50) shows that the empirical distributions themselves do not converge, even in probability.

These results hold for a class of transforms which include the Fourier transform as a special case. This class is quite different from the orthogonal transforms considered in [4]. To state the main results of this paper for the more general transform, let $X_1, X_2, \cdots$ be independent and identically distributed random variables on $(\Omega, \mathcal{F}, P)$, such that

$$n^{-1/\alpha} (X_1 + \cdots + X_n)$$

converges in law to the symmetric stable law of index α. Let h be a nonzero continuous, real-valued function on R of period 1;

$$h(x) = \cos(2\pi x)$$

is the leading special case. For $s = 1, \cdots, n$, let

$$Y_{ns} = n^{-1/\alpha} \Sigma_{j=1}^{n} h(sj/n) X_j.$$

In particular, if $h(x) = \cos(2\pi x)$, then $n^{1/\alpha} Y_{ns}$ is the real part of the s^{th} Fourier coefficient of $X_1, \cdots, X_n$.

Now let μ_n be the empirical measure of $Y_{n1}, \cdots, Y_{nn}$: that is, μ_n assigns mass $1/n$ to each Y_{ns}. Thus, μ_n is a random measure on the real line and has a law λ_n. This λ_n is a measure on the space of measures.

To go at this a bit more slowly, we introduce $M(R)$, the space of probability measures on the Borel real line. Endowed with the weak-star topology, $M(R)$ is a complete separable metric space. And μ_n is a Borel measurable mapping from Ω to $M(R)$. Now $\lambda_n = P\mu_n^{-1}$ is a probability on $M(R)$, that is, an element of $M[M(R)]$. Our main result can be stated as follows:

<u>Theorem</u>: λ_n converges weak-star to a limit λ in M[M(R)].

Notice that μ_n is a random element of M(R), and the theorem says that μ_n converges in law. Does μ_n converge in probability? The answer is no, unless $\alpha = 2$ and $\int_0^1 h(t)dt = 0$. This is the content of proposition (50).

The main results of this paper are proved in Section 3. Readers may wish to begin with this section, and refer to Section 2, which sets out some preliminary lemmas, only when needed. Several of the lemmas of Section 2 may be interesting in themselves. Lemma (1) is a version for rationals of Weyl's theorem on the equidistribution of multiplicative sequences generated by irrationals in R mod 1. Lemmas (4) and (15) establish inequalities for the sums of independent L^α random variables, for $0 < \alpha < 2$. If X and Y have unsymmetric distributions, lemma (4) shows that $E|X - Y|^\alpha < E|X + Y|^\alpha$. If $X_1, \ldots, X_n$ all have symmetric distributions, then according to lemma (15) $E|\Sigma_{k=1}^n X_k|^\alpha < \Sigma_{k=1}^n E|X_k|^\alpha$. Finally, in Section 4 appear some facts about the limit laws of μ_n. In particular, a class of interesting stochastic processes are discussed there, all of whose finite dimensional distributions are stable.

We would like to thank M. L. Eaton for many helpful conversations around the ideas of this paper, and W. Pruitt for suggesting the method of proof used in (4).

2. Some Preliminary Results

The first result of this section is a variation on a famous theorem of Weyl's. For x in R, let $\{x\}$ denote the fractional part of x, that is, $x \bmod 1$. Fix real numbers $\alpha_1, \cdots, \alpha_k$. Let y_j be the k-tuple $\{\alpha_1 j\}, \cdots, \{\alpha_k j\}$. Weyl's theorem states that $y_1, y_2, \cdots$ is equidistributed over the unit cube in R^k, unless the α's are rationally related. But suppose $\alpha_1 = a_1/n, \cdots, \alpha_k = a_k/n$, with integer a_i's: so $\alpha_1, \cdots, \alpha_k$ is a k-tuple of rationals of order n. Lemma (1) shows that, for most such k-tuples, the corresponding sequence $y_1, \cdots, y_n$ is close to being equidistributed over the unit cube in R^k.

For integers n and k, let $N(k)$ denote the set of k-tuples $\underline{a} = (a_1, \cdots, a_k)$ with integer coordinates between 1 and n inclusive (here, the dependence on n is suggested by the N). For $\underline{a}$ in $N(k)$, define the probability ν_{na} by the requirement that it assign weight $1/n$ to each of the k-vectors $(\{a_1 j/n\}, \cdots, \{a_k j/n\})$, for $j = 1, \cdots, n$. Let I denote the unit interval. For f continuous on I^k and $\varepsilon > 0$, define

$$A(n, \varepsilon, f) = \{\underline{a}\colon \ \underline{a} \in N(k) \text{ and } \left| \int f(x)\, \nu_{na}(dx) - \int f(x)dx \right| < \varepsilon\}.$$

Denote the cardinality of a finite set A by $|A|$. In particular, $|N(k)| = n^k$.

(1) Lemma: For every continuous function f on I^k, for every $\varepsilon > 0$,

$$|A(n, \varepsilon, f)|/n^k \to 1 \text{ as } n \to \infty.$$

Proof: Consider the class of complex-valued continuous functions on I^k which satisfy the displayed relation. This class is linear and closed under uniform limits. Thus, it contains every continuous function on I^k if it contains functions of the form

$$f(x) = \exp[2\pi i(v \cdot x)]$$

where "$\cdot$" denotes inner product in R^k and v is a vector in R^k with integer coordinates. The case where v vanishes identically is trivial. So fix v with at least one coordinate nonzero. In this case,

$$\int f(x)dx = 0 .$$

For any real number y and integer p, we have

$$py = p\{y\} + \text{an integer}$$

so

$$\exp(2\pi ipy) = \exp(2\pi ip\{y\}).$$

Thus

$$\int \exp[2\pi i(v \cdot x)]\nu_{na}(dx) = 1/n \, \Sigma_{j=1}^{n} \exp[2\pi i(v \cdot \underline{a})j/n]. \tag{2}$$

Suppose that n does not divide $v \cdot \underline{a}$. We will show that $\underline{a}$ is in $A(n, 0, f)$. Indeed, the right-hand side of (2) is a finite geometric series whose sum is zero. On the other hand, $\int f(x)dx = 0$ too.

Now consider the set S of $\underline{a}$ in $N(k)$ such that $v \cdot \underline{a}$ is divisible by n. We claim that $|S| = 0(n^{k-1})$, which would complete the argument for (1). Let $K = \max_s |v_s|$. Clearly, $|v \cdot \underline{a}|$ is bounded by Kkn for $\underline{a}$ in $N(k)$. For $j = -Kk, \cdots, Kk$ set

$$S_j = \{\underline{a}:\underline{a} \in N(k) \text{ and } v \cdot \underline{a} = jn\}.$$

Then $S = \cup_{j=-Kk}^{Kk} S_j$. But S_j consists of all the integer lattice points in the intersection of the $(k-1)$- dimensional hyperplane $\{x: x \in R^k$ and $v \cdot x = jn\}$ with the hypercube $[1,n]^k$. As such, $|S_j| \leq n^{k-1}$, by induction on k. And so, $|S| \leq (2Kk+1)n^{k-1}$. □

(3) Corollary: Let $\mathcal{F}$ be a family of complex-valued continuous functions f on I^k which is precompact in the sup norm. Let Φ be a bounded continuous function on a closed disk in the complex plane which contains $f(x)$, for all f in $\mathcal{F}$ and x in I^k. Then as $n \to \infty$,

$$n^{-k} \sum_{\underline{a} \in N(k)} \Phi\left(\frac{1}{n} \sum_{j=1}^{n} f(\{a_1 j/n\},\ldots,\{a_k j/n\})\right) \to \Phi\left(\int f(x)dx\right)$$

uniformly in $f \in \mathcal{F}$.

The next three lemmas represent small, but for our purposes critical, improvements on results of Clarkson [3] and von Bahr and Essen [7]. In particular, the strict inequality (11) improves upon the corresponding weak inequality (i.e., with "$\leq$" in place of "$<$") proved in [3]. The representation of $|x|^\alpha$ used in the proof of (4) appears in [7].

(4) Lemma: Suppose X and Y are independent, identically distributed random variables. Let $0 < \alpha < 2$. If $E\{|X|^\alpha\} < \infty$, then

(5) $$E\{|X-Y|^\alpha\} < 2E\{|X|^\alpha\}$$

unless X is degenerate. Also

(6) $$E\{|X-Y|^\alpha\} < E\{|X+Y|^\alpha\} ,$$

unless the distribution of X is symmetric.

Proof: For x in R,

$$|x|^{\alpha} = C_{\alpha} \int_{-\infty}^{\infty} [1 - \cos(ux)]\, |u|^{-\alpha-1}\, du.$$

Here, C_{α} is a real constant whose exact value is immaterial. So, for any random variable X with characteristic function ϕ_X,

(7) $$E\{|X|^{\alpha}\} = C_{\alpha} \int_{-\infty}^{\infty} [1 - \operatorname{Re} \phi_X(u)]\, |u|^{-\alpha-1}\, du.$$

In particular,

(8) $$E\{|X-Y|^{\alpha}\} = C_{\alpha} \int_{-\infty}^{\infty} [1 - |\phi_X(u)|^2]\, |u|^{-\alpha-1}\, du,$$

and

(9) $$E\{|X+Y|^{\alpha}\} = C_{\alpha} \int_{-\infty}^{\infty} [1 - \operatorname{Re} \phi_X^2(u)]\, |u|^{-\alpha-1}\, du.$$

Now (5) follows from (7) and (8), because for any complex number z with $|z| \leq 1$,

$$1 - |z|^2 < 2[1 - \operatorname{Re} z],$$

unless $z = 1$, when equality obtains. But $\phi_X(u) = 1$ for almost all u's if and only if X is degenerate.

Likewise, (6) follows from (8) and (9), because

$$\operatorname{Re}(z^2) < |z|^2,$$

unless z is real, in which case equality obtains. But $\phi_X(u)$ is real for almost all u's if and only if X is symmetric.□

Note: (6) does not extend to $\alpha > 2$. In particular, if EX and EX^3 have opposite signs, the inequality is reversed for $\alpha = 4$. For general $\alpha > 2$, let X and Y be independent, identically distributed random variables, with $P[X = -L] = 1 - P[X = 1] = L^{-\frac{1}{2}\alpha}$. For L sufficiently large, $E\{|X+Y|^\alpha\} < E\{|X-Y|^\alpha\}$.

(10) Corollary: Suppose h is a measurable function on the unit interval, $0 < \alpha < 2$, and $0 < \int_0^1 |h(x)|^\alpha \, dx < \infty$. Then

$$\int_0^1 \int_0^1 |h(x_1) - h(x_2)|^\alpha \, dx_1 \, dx_2 < 2 \int_0^1 |h(x)|^\alpha \, dx.$$

This inequality holds also for $\alpha = 2$, unless $\int_0^1 h(x) = 0$.

Proof: For $\alpha = 2$, the calculation is immediate. Otherwise, let U and V be independent random variables, uniform on the unit interval. Set $X = h(U)$ and $Y = h(V)$. Then apply (5). □

(11) Lemma: Suppose x and y are nonzero real numbers and $1 < \alpha < 2$. Then

$$|x-y|^\alpha + |x+y|^\alpha < 2(|x|^\alpha + |y|^\alpha).$$

Proof: Divide both sides of the inequality by the larger in absolute value of x and y. This reduces (11) to the claim that, for $0 < x \leq 1$,

(12) $$\phi(x) < 1+x^\alpha \ , \quad \text{where} \quad \phi(x) = \tfrac{1}{2}[(1-x)^\alpha + (1+x)^\alpha]$$

Expand in a Taylor series:

$$\phi(y) = 1 + \Sigma_{n=1}^{\infty} c_n \, y^{2n} \ ,$$

where

(13) $$c_n = \frac{\alpha(\alpha-1) \cdots (\alpha-2n+1)}{(2n)!} > 0$$

because $1 < \alpha < 2$. In particular, $\phi(y)$ increases with y for $0 \le y \le 1$. But $\phi(1) = 2^{\alpha-1} < 2$, so

$$\Sigma_{n=1}^{\infty} c_n y^{2n} < 1 \quad \text{for} \quad 0 \le y \le 1.$$

Substitute $y = x^{1-\frac{1}{2}\alpha}$:

(14) $$\Sigma_{n=1}^{\infty} c_n x^{(2-\alpha)n} < 1 \quad \text{for} \quad 0 \le x \le 1.$$

Now $x^{-\alpha} \ge 1$, so $x^{-n\alpha} \ge x^{-\alpha}$. By (13) and (14)

$$\Sigma_{n=1}^{\infty} c_n x^{2n-\alpha} < 1 \quad \text{for} \quad 0 \le x \le 1.$$

Multiply by x^{α}:

$$\Sigma_{n=1}^{\infty} c_n x^{2n} < x^{\alpha} \quad \text{for} \quad 0 < x \le 1.$$

Adding 1 to both sides gives (12).□

(15) Lemma: Suppose $0 < \alpha < 2$. Let $X_1, \cdots, X_k$ be nondegenerate, symmetric, independent random variables, with $E\{|X_i|^{\alpha}\} < \infty$. Then

$$E\{|\Sigma_{i=1}^{k} X_i|^{\alpha}\} < \Sigma_{i=1}^{k} E\{|X_i|^{\alpha}\}.$$

Proof: First, suppose $0 < \alpha < 1$. For x a nonnegative real number, let $\phi(x) = 1 + x^{\alpha}$ and $\psi(x) = (1 + x)^{\alpha}$. Then ϕ and ψ are equal at zero, while for all positive x, the derivative of ψ is strictly less than the derivative of ϕ. Thus ϕ is strictly greater than ψ, for positive x. Now let x_1 and x_2 be any nonzero real numbers. Then

$$\begin{aligned} |x_1 + x_2|^{\alpha} &\le (|x_1| + |x_2|)^{\alpha} \\ &\le |x_1|^{\alpha}\, \psi(|x_2| \,/\, |x_1|) \\ &< |x_1|^{\alpha}\, \phi(|x_2| \,/\, |x_1|) \\ &= |x_1|^{\alpha} + |x_2|^{\alpha}. \end{aligned}$$

By induction, if $x_1, \cdots, x_k$ are real numbers at least two of which are nonzero,

$$|\Sigma_{i=1}^k x_i|^\alpha < \Sigma_{i=1}^k |x_i|^\alpha.$$

The conclusion of (15) in this case follows by integration.

Next, suppose $\alpha = 1$. Certainly,

$$|X_1 + \cdots + X_n| \leq |X_1| + \cdots + |X_n|.$$

Let A be the event that at least two of the X_i's have different signs. Since the X_i's are independent and have symmetric distributions, A has positive probability. On A,

$$|X_1 + \cdots + X_n| < |X_1| + \cdots + |X_n| .$$

The conclusion of (15) follows for this case.

Finally, suppose $1 < \alpha < 2$. Consider the case $k = 2$. From (11),

$$E\{|X_1 + X_2|^\alpha\} + E\{|X_1 - X_2|^\alpha\} < 2\ (E\{|X_1|^\alpha\} + E\{|X_2|^\alpha\}).$$

Since X_1 and X_2 are independent and X_2 is symmetric,

$$E\{|X_1 + X_2|^\alpha\} = E\{|X_1 - X_2|^\alpha\},$$

and so the result follows. The inequality is obtained for $k > 2$ by induction.□

(16) <u>Corollary</u>: Suppose $0 < \alpha < 2$. Let $X_1, \cdots, X_k$ be nondegenerate, symmetric, independent random variables, all with the same distribution, and with $E|X_1|^\alpha < \infty$. Let $t_1, \cdots, t_k$ be real numbers, with $\Sigma_{i=1}^k |t_i|^\alpha \leq 1$. Then

$$E\{|\Sigma_{i=1}^k t_i X_i|^\alpha\} < E\{|X_1|^\alpha\}.$$

The next main result is lemma (19), a characterization of the domain of attraction to the symmetric stable laws. The preliminaries in lemmas (17) and (18) give a careful treatment of the logarithm of the characteristic function. Proofs are omitted, being routine applications of the method of analytic continuation. The material is well known, but we cannot supply references.

(17) Lemma: Let $0 < T \leq \infty$. Let t be a real variable, with $0 \leq |t| < T$. Let ψ be a continuous, complex-valued non-vanishing function of t, with $\psi(0) = 1$.

(a) There is a unique continuous, complex-valued function λ of $t \in (-T, T)$ such that $\lambda(0) = 0$ and $\lambda(t)$ is a value of $\log[\psi(t)]$. Write $\lambda(t) = (\log, \psi)(t)$.

(b) $(\log, \psi^n)(t) = n(\log, \psi)(t)$.

(c) Let $0 < T_0 < T$. Suppose $|1 - \psi(t)| < 1$ for $|t| \leq T_0$. Then for $|t| \leq T_0$,

$$(\log, \psi)(t) = -\Sigma_{k=1}^{\infty} \frac{1}{k} (1 - \psi(t))^k.$$

(d) Let $0 < T_1 \leq T$. Suppose $\psi(t)$ is real-valued for $0 \leq |t| < T_1$. Then $(\log, \psi)(t)$ is the ordinary real logarithm of $\psi(t)$, for $|t| < T_1$.

(18) Lemma: Let $0 < T \leq \infty$. Let θ_n and θ be continuous, complex-valued, non-vanishing functions of the real variable t for $0 \leq |t| < T$, with $\theta_n(0) = \theta(0) = 1$. Suppose $\theta_n \to \theta$, uniformly for $|t| \leq T_0 < T$. Then $(\log, \theta_n) \to (\log, \theta)$ uniformly for $|t| \leq T_0$. This can fail for pointwise convergence.

(19) Lemma: Let $0 < \alpha \leq 2$ and $0 < c < \infty$. Let be a characteristic function. Then

$$\phi(t/n^{1/\alpha})^n \to \exp(-c|t|^\alpha)$$

uniformly on bounded intervals if and only if

$$\phi(t) = 1 - c|t|^\alpha + o(|t|^\alpha) \quad \text{as} \quad t \to 0.$$

<u>Proof</u>: The "if" part is easy. For "only if", set

(20) $$\delta(t) = 1 - \phi(t)$$

and

$$\sigma(t) = \Sigma_{k=1}^{\infty} \frac{1}{k}\,\delta(t)^k.$$

Choose $T_0 > 0$ so small that

(21) $$|\delta(t/n^{1/\alpha})| \leqq \tfrac{1}{2} \quad \text{for all } n \text{ and all } t \text{ with } |t| \leqq T_0.$$

Use lemma (17), with $\psi_n(t) = \phi(t/n^{1/\alpha})$ in place of ψ; this function does not vanish by (21): the conclusion is

$$(\log, \psi_n^{\,n})(t) = n(\log, \psi_n)(t) = -n\sigma(t/n^{1/\alpha})$$

for $|t| \leqq T_0$. In view of (18), there is a sequence $\varepsilon_n \to 0$ such that

(22) $$\left|c|t|^\alpha - n\sigma(t/n^{1/\alpha})\right| \leqq \varepsilon_n, \quad \text{for all } n \text{ and all } t \text{ with } |t| \leqq T_0.$$

Divide by n and put $u = t/n^{1/\alpha}$: for all n,

(23) $$\left|c|u|^\alpha - \sigma(u)\right| \leqq \varepsilon_n/n \quad \text{for all } u \text{ with } |u| \leqq T_0/n^{1/\alpha}.$$

Given u with $0 < u \leqq T_0$ choose n so that

(24) $$T_0/(n+1)^{1/\alpha} < |u| \leqq T_0/n^{1/\alpha}.$$

But then $1/n \leqq 2/(n+1) \leqq 2T_0^{-\alpha}\,|u|^{\alpha}$. So

(25) $$\left|c|u|^{\alpha} - \sigma(u)\right| \leqq 2\,\varepsilon_n\, T_o^{-\alpha}\,|u|^{\alpha} \quad \text{for } 0 < |u| \leqq T_o, \text{ with } n \text{ defined by (24).}$$

As $u \to 0$, clearly $n \to \infty$ and $\varepsilon_n \to 0$. Hence

(26) $$\sigma(u) = c\,|u|^{\alpha} + o(|u|^{\alpha}) \quad \text{as} \quad u \to 0.$$

Recall (20). Clearly, $\sigma(u) = \delta(u) + \rho(u)$, where

$$\rho(u) = \Sigma_{k=2}^{\infty} \frac{1}{k}\,\delta(u)^k.$$

Recall (21). For $|u| \leqq T_0$,

(27) $$|\rho(u)| \leqq |\delta(u)|^2 \;\; \Sigma_{k=2}^{\infty} \frac{1}{k}\,(\frac{1}{2})^k \leqq \frac{1}{4}\,|\delta(u)|^2$$

and

(28) $$|\rho(u)| \leqq \frac{1}{8}\,|\delta(u)|.$$

In particular, $\delta(u) = \sigma(u) - \rho(u) = O(|u|^{\alpha})$ by (26) and (28). Then $\rho(u) = o(|u|^{\alpha})$ by (27), so in fact $\delta(u) = c|u|^{\alpha} + o(|u|^{\alpha})$. □

The next three results are well known.

(29) <u>Lemma</u>: Let z_j and z_j' be complex numbers, with absolute values bounded by A. Then

$$|\Pi_{j=1}^{n} z_j - \Pi_{j=1}^{n} z_j'| \leqq A^{n-1}\Sigma_{j=1}^{n} |z_j - z_j'|.$$

(30) <u>Lemma</u>: Let z be a complex number. Then

$$|e^z - 1 - z| \leqq \tfrac{1}{2}\,|z|^2\, e^{|z|}$$

(31) Lemma: Let X be a random variable with characteristic function ϕ, and $\varepsilon > 0$. Then

$$P\{|X| \geq 2/\varepsilon\} \leq \frac{1}{\varepsilon} \int_{-\varepsilon}^{\varepsilon} \mathrm{Re}[1 - \phi(t)]dt.$$

We will be considering the law of an empirical distribution, that is, a measure on measures. Some technical machinery is developed in the remainder of this section to handle this complication. Suppose X is a complete separable metric space. Let $M(X)$ denote the space of probability measures on the Borel σ-field of X; equip $M(X)$ with the weak star topology. By definition, a subset S of $M(X)$ is tight if for each $\varepsilon > 0$, there is a compact subset K_ε of X such that $m(K_\varepsilon) > 1 - \varepsilon$ for all m in S. By Prohorov's Theorem [1, p. 37], a subset of $M(X)$ is tight if and only if it is relatively compact. $M(X)$ is itself a complete separable metric space [6, Theorem 6.2, p. 43]. Thus, $M[M(X)]$ is well defined.

(32) Lemma: A subset T of $M[M(X)]$ is tight if and only if for each $\varepsilon > 0$, there is a compact subset K_ε of X such that for all λ in T,

$$\lambda\{m: m(K_\varepsilon) \geq 1 - \varepsilon\} \geq 1 - \varepsilon.$$

Proof: "If". Suppose the condition holds. Fix $\delta > 0$. We need to find a compact subset C_δ of $M(X)$ such that $\lambda(C_\delta) \geq 1 - \delta$ for all λ in T. Pick a sequence of positive numbers ε_n such that $\Sigma_{n=1}^{\infty} \varepsilon_n < \delta$. For each n, choose a compact subset K_n of X according to the condition with ε_n in place of ε. Let A_n be the compact set of m in $M(X)$ with $m(K_n) \geq 1 - \varepsilon_n$. Then, let $C_\delta = \cap_n A_n$.

"Only if". Suppose T is tight. For each $\varepsilon > 0$, there is a compact subset C_ε

of $M(X)$ such that $\lambda(C_\varepsilon) \geq 1 - \varepsilon$ for all λ in T. By Prohorov's Theorem, each C_ε is tight: so there exists a compact subset K_ε of X such that $m(K_\varepsilon) \geq 1 - \varepsilon$ for all m in C_ε. □

The proof of the next result is omitted as routine.

(33) <u>Lemma</u>: Suppose $f, f_1, f_2, \cdots$ are uniformly bounded continuous real-valued functions on X, and f_n converges to f uniformly on compacts. Suppose λ and γ are in $M[M(X)]$ and for each integer n, the random variable $m \to \int f_n dm$ on $M(X)$ has the same distribution under λ as it has under γ. Then the λ-distribution of $m \to \int fdm$ coincides with the γ - distribution.

Let B denote the space of bounded continuous functions from R into $\mathbb{C}$. With the topology of uniform convergence on compacts, B is a complete separable metric space. For m in $M(R)$, denote the characteristic function of m by $\hat{m}$; that is, $\hat{m}(t) = \int \exp(itx)\, m(dx)$.

(34) <u>Lemma</u>: Suppose λ and γ are in $M[M(R)]$, and for each integer k and k-tuple of real numbers $(t_1, \cdots, t_k)$, the λ - distribution of $m \to [\hat{m}(t_1), \cdots, \hat{m}(t_k)]$ coincides with the γ - distribution. Then $\lambda = \gamma$.

<u>Proof</u>: Let $f_1, \cdots, f_n$ be bounded continuous real-valued functions on R. It is enough to show that the λ- and γ- distributions of vectors of the form $(\int f_1 dm, \ldots, \int f_n dm)$ coincide. We begin with the case in which $n = 1$. Suppose f is a complex trigonometric polynomial, namely $f(x) = \Sigma_{j=1}^k a_j \exp(it_j x)$ for some integer k, complex k-tuple $\underline{a}$ and real k-tuple t. If m is in $M(R)$, then $\int fdm = \Sigma_{j=1}^k a_j \hat{m}(t_j)$. By assumption, for any Borel subset A of $\mathbb{C}$

$$\lambda\{m:\ m \in M(R) \text{ and } \Sigma_{j=1}^{k} a_j\ \hat{m}(t_j) \in A\} = \gamma\{m:\ m \in M(R) \text{ and } \Sigma_{j=1}^{k} a_j\ \hat{m}(t_j) \in A\}.$$

This settles the case of one trigonometric polynomial.

Next, let f be any bounded, continuous real-valued function on the line. There is a sequence of real trigonometric polynomials which are uniformly bounded and converge to f uniformly on compacts. By (33), the λ - and γ - distributions of $\int f\,dm$ coincide. This settles the case $n = 1$.

Finally, let $f_1, \cdots, f_n$ be bounded continuous real-valued functions on R, and $c_1, \cdots, c_n$ arbitrary real numbers. Then

$$\Sigma_{j=1}^{n} c_j \int f_j\ dm = \int (\Sigma_{j=1}^{n} c_j f_j)\ dm \quad \text{for } m \text{ in } M(R),$$

and the right-hand side has the same distribution under λ as it has under γ. By Radon's Theorem, the λ - and γ - distributions of the n-vector $m \to (\int f_1\ dm, \cdots, \int f_n\ dm)$ must coincide also.□

(35) <u>Lemma</u>: Let k be a positive integer and $Z_1, \cdots, Z_k$ complex-valued random variables, with $|Z_j| \leq 1$ for all j. Then the joint distribution of $Z_1, \cdots, Z_k$ is determined by the moments

$$E\left(Z_1^{a_1}\ \bar{Z}_1^{b_1} \cdots Z_k^{a_k}\ \bar{Z}_k^{b_k}\right),$$

where the a_j and b_j range over all nonnegative integers.

<u>Proof</u>: Immediate from the Stone-Weierstrass Theorem.□

<u>Note</u>: The conjugate moments really are needed. For example, suppose Z is uniform over the circle with radius $r \leq 1$. Then $E(Z^a) = 0$ unless $a = 0$, in which case the expectation is 1. This is so whatever r may be.

(36) <u>Proposition:</u> For each n, let λ_n be an element of $M[M(R)]$. Suppose that for each $\varepsilon > 0$, there is a compact subset K_ε of R such that

(37) $$\lambda_n \{m: m \in M(R) \text{ and } m(K_\varepsilon) \geq 1 - \varepsilon\} \geq 1 - \varepsilon, \text{ for all } n.$$

Suppose too that for every integer k and k-tuple of real numbers $(t_1, \cdots, t_k)$,

(38) $$\int \hat{m}(t_1) \cdots \hat{m}(t_k)\, \lambda_n(dm) \text{ converges as } n \text{ goes to infinity.}$$

Then λ_n converges weak-star to some element λ of $M[M(R)]$.

The limit λ is point-mass at some point in $M(R)$ if and only if for all t, the λ_n - variance of $m \to \hat{m}(t)$ goes to zero as n goes to infinity.

<u>Note:</u> The λ_n- variance of $\hat{m}(t)$ is

$$\int |\hat{m}(t) - \int \hat{m}(t)\, \lambda_n(dm)|^2\, \lambda_n(dm) = \int |\hat{m}(t)|^2\, \lambda_n(dm) - |\int \hat{m}(t)\, \lambda_n(dm)|^2.$$

<u>Proof:</u> By condition (37) and lemma (32), the sequence $\{\lambda_n\}$ is tight. By Prohorov's Theorem, then, $\{\lambda_n\}$ is relatively compact. Suppose λ is a subsequential limit of $\{\lambda_n\}$. Let k be a positive integer and $(t_1, \cdots, t_k)$ a k-tuple of real numbers. By condition (38) and lemma (35), the λ- distribution of

$$m \to [\hat{m}(t_1), \cdots, \hat{m}(t_k)]$$

is determined: the complex conjugate of $\hat{m}(t)$ is just $\hat{m}(-t)$. By lemma (34), then, λ is unique. Therefore, λ_n converges weak-star to λ.

When is λ a point-mass? Clearly, if and only if for all real t, the λ- variance of $m \to \hat{m}(t)$ is zero. But the λ- variance is the limit of the λ_n-variances, because $|\hat{m}(t)| \leq 1$. □

3. The Convergence Theorem

Let h be a continuous function on R with period 1. Let α be a real number with $0 < \alpha \leq 2$. Let $X_1, X_2, \cdots$ be independent, identically distributed random variables on (Ω, F, P), such that $n^{-1/\alpha}(X_1 + \cdots + X_n)$ converges in distribution to the symmetric stable law of order α. Let ψ denote the characteristic function of X_1.

For $n = 1, 2, \cdots$ and $s = 1, \cdots, n$, define Y_{ns} by

(39) $$Y_{ns} = n^{-1/\alpha} \Sigma_{j=1}^{n} h(sj/n)\, X_j.$$

In particular, if $h(x) = \cos 2\pi x$, then $n^{1/\alpha} Y_{ns}$ is the real part of the s^{th} Fourier coefficient of $X_1, \cdots, X_n$.

Recall from (1) that $N(k)$ is the set of k-tuples $\underline{a} = (a_1, \ldots, a_k)$ of integers between 1 and n inclusive. Let $t = (t_1, \cdots, t_k)$ be a k-tuple of real numbers. Define functions ϕ_n and θ_k by

(40) $$\phi_n(\underline{a}, t) = E \exp[i(t_1 Y_{na_1} + \cdots + t_k Y_{na_k})]$$

and

(41) $$\theta_k(t) = \exp[-\textstyle\int_{I^k} |t_1 h(x_1) + \cdots + t_k h(x_k)|^\alpha \, dx],$$

where dx is Lebesgue measure on the k- dimensional unit cube I^k.

(42) Proposition: $n^{-k} \Sigma_{\underline{a} \in N(k)} \phi_n(a,t)$ converges uniformly on compact t-sets to $\theta_k(t)$.

Proof: Clearly

$$\phi_n(\underline{a}, t) = E[\exp(iY)]$$

where

$$Y = t_1 Y_{na_1} + \cdots + t_k Y_{na_k}.$$

Next, by collecting terms,

$$Y = n^{-1/\alpha} \Sigma_{j=1}^{n} h(\underline{a},j,t)X_j$$

where

$$h(\underline{a},j,t) = t_1 h(a_1 j/n) + \cdots + t_k h(a_k j/n).$$

Thus

$$(43) \qquad \phi_n(\underline{a},t) = E[\exp(iY)] = \Pi_{j=1}^{n} \psi[h(\underline{a},j,t)/n^{1/\alpha}].$$

Now h is bounded, so in view of (19),

$$(44) \qquad \psi[h(\underline{a},j,t)/n^{1/\alpha}] = 1 - \frac{1}{n}|h(\underline{a},j,t)|^{\alpha} + o(1/n),$$

the error being uniform in $\underline{a}$, j, and compact t.

Of course,

$$(45) \qquad \exp[-\frac{1}{n}\Sigma_{j=1}^{n}|h(\underline{a},j,t)|^{\alpha}] = \Pi_{j=1}^{n} \exp[-\frac{1}{n}|h(\underline{a},j,t)|^{\alpha}].$$

By (29), applied to (43) and (45), the difference between $E[\exp(iY)]$ and $\exp[-\frac{1}{n}\Sigma_{j=1}^{n}|h(\underline{a},j,t)|^{\alpha}]$ is at most

$$\Sigma_{j=1}^{n} |\psi[h(\underline{a},j,t)/n^{1/\alpha}] - \exp[-\frac{1}{n}|h(\underline{a},j,t)|^{\alpha}]|,$$

which by the triangle inequality is at most $T_1 + T_2$, where

$$T_1 = \Sigma_{j=1}^{n}|\psi[h(\underline{a},j,t)/n^{1/\alpha}] - [1 - \frac{1}{n}|h(\underline{a},j,t)|^{\alpha}]|$$

and

$$T_2 = \Sigma_{j=1}^{n}|1 - \frac{1}{n}|h(\underline{a},j,t)|^{\alpha} - \exp[-\frac{1}{n}|h(\underline{a},j,t)|^{\alpha}]|.$$

Restrict t to a compact set in R^k. Then (44) implies that T_1 is $o(1)$, uniformly in $\underline{a}$, j, and t. Next, h is uniformly bounded and so

$\frac{1}{n}\Sigma_{j=1}^{n}|h(\underline{a},j,t)|^{2\alpha} = O(1)$ uniformly in $\underline{a}$ and j. By (30), then, T_2 is $o(1)$. Thus

$$\phi_n(\underline{a},t) = \exp[-\frac{1}{n}\Sigma_{j=1}^{n}|h(\underline{a},j,t)|^{\alpha}] + o(1),$$

uniformly in $\underline{a}$, j, and t restricted to a compact.

We may now apply (3) to estimate the exponent in (45). For F, take the functions on I^k of the form

$$|t_1h(x_1) + \cdots + t_kh(x_k)|^{\alpha},$$

as t ranges over a compact set in R^k. Also, take $\Phi: x \to \exp(-x)$ on $0 \leq x \leq (\|h\|_\infty \sup_t \Sigma|t_k|^{\alpha})$. This completes the proof.□

Remark: The same argument shows that $n^{-k}\Sigma_{\underline{a}\in N(k)}|\phi_n(\underline{a},t)|^2$ converges to $[\theta_k(t)]^2$. So $\phi_n(\underline{a},t)$ is nearly $\theta_k(t)$ for most k-tuples $\underline{a}$.

(46) Corollary: $\frac{1}{n}\Sigma_{s=1}^{n} P\{|Y_{ns}| > L\}$ converges to zero as L goes to infinity, uniformly in n.

Proof: Use the case $k = 1$ of (42), and then (31).□

We are now ready to state and prove the main theorem of this paper. Let μ_n be the empirical measure of $\{Y_{ns}\}$; that is, μ_n assigns mass $1/n$ to each Y_{ns}. Thus, μ_n is a Borel measurable mapping from Ω into $M(R)$. Let λ_n be the distribution of μ_n, so λ_n is in $M[M(R)]$.

(47) Theorem: λ_n converges weak-star to a limit λ in $M[M(R)]$.

Proof: We will use (36). We first verify condition (37). Fix $L < \infty$. Then

$$\int_{M(R)} m[-L, L]\ \lambda_n(dm) = \frac{1}{n} \Sigma_{s=1}^n\ P\{|Y_{ns}| \leq L\}$$

is uniformly close to 1, by (46). So (37) follows by Chebyschev's inequality. Next, we verify condition (38). But

$$\int \hat{m}(t_1) \cdots \hat{m}(t_k)\ \lambda_n(dm)$$

can be evaluated as

$$n^{-k} \quad \Sigma_{\underline{a} \in N(k)}\ E\{\exp[i(t_1 Y_{na_1} + \cdots + t_k Y_{na_k})]\}$$

whose limit was computed in (42).□

(48) Corollary: The limit λ in (47) depends only on α and h, but not on the distribution of the X_i's. Indeed, λ is characterized by the fact that

$$\int \hat{m}(t_1) \cdots \hat{m}(t_k)\ \lambda(dm) = \theta_k(t_1, \ldots, t_k)$$

where θ_k is defined in (41).

Proof: Use (34) and (35).□

(49) Corollary: The limit λ in (47) is a point mass if and only if $\alpha = 2$ and $\int_0^1 h(t)\ dt = 0$.

Proof: To apply (36), we need to compute the λ_n - variance of $\hat{m}(t)$. This variance is $T_1 - T_2$, where

$$T_1 = E\{|\frac{1}{n}\ \Sigma_{s=1}^n\ \exp(itY_{ns})|^2\}$$

and

$$T_2 = |E\{\frac{1}{n}\ \Sigma_{s=1}^n\ \exp(itY_{ns})\}|^2.$$

From (42) with $k = 1$,

$$T_2 \to \exp[-2|t|^\alpha \int_0^1 |h(x)|^\alpha dx].$$

Similarly, using (42) with $k = 2$,

$$T_1 = \frac{1}{n^2} \Sigma_{\underline{a} \in N(2)} E\{\exp[i(tY_{na_1} - tY_{na_2})]\}.$$

$$\to \exp[-|t|^\alpha \int |h(x_1) - h(x_2)|^\alpha dx_1\, dx_2].$$

Now use (10).□

<u>Note</u>: We have proved that the empirical measure μ_n converges in distribution. But μ_n is a random element of $M(R)$, which is endowed with the weak-star topology. The next result shows μ_n does not converge in probability, except for a special case.

(50) <u>Proposition</u>: μ_n converges in probability if and only if $\alpha = 2$ and $\int_0^1 h(t)\, dt = 0$.

<u>Proof</u>: "If." This follows from (49): if λ is a point mass and μ_n converges to λ in distribution, it converges in probability also.

"Only if." Fix $t > 0$. Let

$$\phi_n = \int_{-\infty}^{\infty} \exp(itx)\, \mu_n(dx),$$

a complex-valued random variable bounded in absolute value by 1. If μ_n converges in probability, then $\{\phi_n\}$ is a Cauchy sequence in L^2:

$$E\{|\phi_n - \phi_{n'}|^2\} \to 0$$

as $n, n' \to \infty$. We will derive a contradiction. First,

$$E\{|\phi_n - \phi_{n'}|^2\} = V_1 - V_2,$$

where

$$V_1 = E\{|\phi_n|^2\} + E\{|\phi_{n'}|^2\}$$

and

$$V_2 = 2 \text{ Re } E\{\phi_n \overline{\phi_{n'}}\}.$$

By (42), as $n, n' \to \infty$,

$$V_1 \to 2 \exp[-|t|^\alpha \iint_{I^2} |h(x_1) - h(x_2)|^\alpha dx_1 dx_2]. \tag{51}$$

Now send n' to infinity before n. By (42),

$$V_2 \to 2 \exp[-|t|^\alpha \, 2\int_0^1 |h(x)|^\alpha dx].$$

Finally, apply (10) to (51) and (52) to obtain the contradiction.□

4. The Limiting Measures

In this section, we suppose $0 < \alpha < 2$ and study the measure λ, the weak limit of λ_n in (47). As stated in (48), λ is determined by the quantities

(53) $$\int_{M(R)} \hat{m}(t_1) \cdots \hat{m}(t_k)\, \lambda(dm) = \theta_k(t)$$

over all integers k and k-tuples of real numbers $t = (t_1,\cdots,t_k)$, with $\theta_k(t)$ defined in (41).

Equation (53) can be interpreted as follows. Choose m at random according to λ. This m is a probability on the line: given m, construct a sequence ξ_1, ξ_2, $\cdots$ of independent random variables with common distribution m. Unconditionally, the members of this sequence form an exchangeable process, and (53) gives their joint characteristic function:

(54) $$E\{\exp[i\, \Sigma_{j=1}^{k} t_j\xi_j] = \exp[-\int_{I^k} |\Sigma_{j=1}^{k} t_j h(x_j)|^\alpha dx_1\cdots dx_k].$$

From this point of view, proposition (42) states that the generalized Fourier coefficients $Y_{n1},\cdots,Y_{nn}$ are "nearly" distributed like $\{\xi_1,\xi_2,\cdots\}$ in the following sense: as $n \to \infty$, most k-tuples of these generalized Fourier coefficients are distributed like $\xi_1,\cdots,\xi_k$. According to (47), then, the empirical distribution of $Y_{n1},\cdots,Y_{nn}$ behaves like the empirical distribution of $\xi_1,\cdots,\xi_n$, namely, its law goes to λ.

The next result shows that $(\xi_1,\cdots,\xi_k)$ has a multivariate distribution which is symmetric stable of index α.

(55) Proposition: a) Fix real numbers $c_1,\cdots,c_k$. Then $\Sigma_{j=1}^{k} c_j\xi_j$ is symmetric stable of order α.

b) Let $\zeta_1,\cdots,\zeta_n$ be independent k-vectors, each distributed like

$(\xi_1,\cdots,\xi_k)$. Then $n^{-1/\alpha}(\zeta_1 + \cdots + \zeta_n)$ is also distributed like $(\xi_1,\cdots,\xi_k)$.

Proof: This is immediate from (54).□

Call an element of $M[M(R)]$ symmetric stable of order β if it assigns measure 1 to the set of measures in $M(R)$ which are symmetric stable of order β, having arbitrary scale parameters. If μ is a symmetric stable measure in $M[M(R)]$ of order β, then there is a measure γ in $M(R)$ concentrated in $(0,\infty)$ such that, for all integers k and k-tuples of real numbers $t_1,\cdots t_k$,

(56) $$\int_{M(R)} \hat{m}(t_1)\cdots\hat{m}(t_k)\mu(dm) = \int_0^\infty \exp(-c\,\Sigma_{j=1}^k |t_j|^\beta)\gamma(dc).$$

Here, c is the arbitrary scale parameter.

The reported results that the empirical distribution of $(Y_{n1},\cdots,Y_{nn})$ is nearly normal suggest that λ should by symmetric stable of order 2. The following proposition shows that this is not so.

(57) Proposition: Suppose k is not constant, $0 < \alpha < 2$, and $0 < \beta \leq 2$. Then λ is not symmetric stable of order β.

Proof: Suppose the contrary. Let γ be the measure corresponding to λ as in (56). We will obtain a contradiction between the representations (56) and (53).

Case 1: $\alpha < \beta$. Let $t = [\Sigma_{j=1}^k |t_j|^\beta]^{1/\beta}$. Then (56) entails

(58) $$\int_{M(R)} \hat{m}(t_1)\cdots\hat{m}(t_k)\lambda(dm) = \int_{M(R)} \hat{m}(t)\lambda(dm).$$

Evaluate both sides of (58) by (53) and (41):

(59) $$\int_{I^k} |t_1 h(x_1) + \cdots + t_k h(x_k)|^\alpha dx = t^\alpha \int_I |h(x)|^\alpha dx.$$

Let $U_1, U_2, \cdots$ be independent random variables, each uniform on $[0,1]$. Let $V_j = h(U_j)$. With $k = 2$, the right hand side of (59) gives the same evaluation for two cases: $t_1 = t_2 = 1$ and $t_1 = -t_2 = 1$. Thus, the left hand sides for the two cases also coincide, which shows that

(60) $$E[|V_1 - V_2|^\alpha] = E[|V_1 + V_2|^\alpha].$$

Thus, by (4), the V's are symmetric. They are nondegenerate because h is not constant. Now (16) implies that

$$E[|\Sigma_{j=1}^k t_j V_j|^\alpha] < [\Sigma_{j=1}^k |t_j|^\alpha] E|V_1|^\alpha.$$

Since $\alpha < \beta$,

$$[\Sigma_{j=1}^k |t_j|^\alpha]^{1/\alpha} \le [\Sigma_{j=1}^k |t_j|^\beta]^{1/\beta} = t,$$

so the left side of (59) is strictly smaller than the right side, a contradiction.

Case 2: $\beta < \alpha$. Use (56) with $k = 1$ to see that

$$\exp[-|t|^\alpha \int_0^1 |h(x)|^\alpha dx] = \int_0^\infty \exp(-c|t|^\beta)\gamma(dc).$$

Put x for c, λ for $|t|^\beta$, and k for $\int_0^1 |h(x)|^\alpha dx$:

$$\int_0^\infty \exp(-\lambda x)\gamma(dx) = \exp(-k\lambda^{\alpha/\beta}).$$

However, since $\alpha > \beta$, the right side is not a Laplace transform.

Case 3: $\beta = \alpha$. As in Case 2, we get

$$\int_0^\infty \exp(-\lambda x)\gamma(dx) = \exp(-k\lambda),$$

so $\gamma\{k\} = 1$. But this contradicts (49).□

Here is a probabilistic construction of ξ. First, it is convenient to embed ξ in a continuous time process: for each instant t, $\xi^t = (\xi_1^t, \xi_2^t, \cdots)$ will be an infinite random vector, with ξ^1 equal to the original ξ. To construct ξ^t, proceed as follows. For simplicity, suppose $h(U)$ and $-h(U)$ have the same distribution, where U is a uniform random variable on $[0,1]$.

<u>Step 1</u>: Construct an infinite supply of random vectors $V_1, V_2, \cdots$, each $V_i = (v_{i1}, v_{i2}, \cdots)$ where the v_{ij}'s are independent random variables distributed as $h(U)$.

<u>Step 2</u>: Construct a one-dimensional symmetric stable process of index α. Call this process $\{\eta_t\}$.

<u>Step 3</u>: Take each jump of η_t; say the height of the jump is u; replace this by the vector $C_\alpha \cdot u \cdot V$, where V is one of the vectors constructed in Step 1, and C_α is the constant in (7).

<u>Step 4</u>: Sum the vector "jumps": the sum of these "jumps" to time t is ξ^t. (Note: if $\alpha \leq 1$, caution is in order. Take the sum over the jumps corresponding to u such that $|u| \geq t$, and let $t \to 0$.)

The log characteristic function of $(\xi_1, \cdots, \xi_k)$ is

$$-\int |\Sigma_{j=1}^k t_j h(x_j)|^\alpha dx_1 \cdots dx_k.$$

By (7), this is

$$C_\alpha \int_{-\infty}^{\infty} [\Pi_{j=1}^k \phi(t_j u) - 1] \frac{du}{|u|^{1+\alpha}},$$

where ϕ is the characteristic function of the random variable $h(U)$. Now $C_\alpha |u|^{-(1 + \alpha)}$ is the canonical Lévy measure for the process η.

The relevant infinite dimensional canonical measure is then as follows: select u from the canonical measure for η; make $v_1, v_2, \cdots$ independent and distributed like $h(U)$; then, take the distribution of $uv_1, uv_2, \cdots$.

Given η, the processes $\xi_1^t, \xi_2^t, \cdots$ are independent and identically distributed. Now λ can be described as follows: a "typical" m selected from λ is the distribution of $\Sigma u_j v_j$, where the v_j's are independent and distributed like $h(U)$. So the u's are parameters, which are randomly selected by λ, as the jumps of the process η between 0 and 1.

References

[1] Billingsley, P. (1968), Convergence of Probability Measures, J. Wiley, New York.

[2] Chung, K. L. (1974), A Course in Probability Theory (second edition), Academic Press, New York.

[3] Clarkson, J. A. (1936), "Uniformly Convex Spaces, " Trans. Amer. Math. Soc., Vol. 40, pp. 396-414.

[4] Freedman, D. and D. Lane (1978), "The Empirical Distribution of Fourier Coefficients," to appear Annal of Statistics, Nov. 1980.

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

[6] Parthasarathy, K. R. (1967), Probability Measures on Metric Spaces, Academic Press, New York.

[7] von Bahr, B. and C. - G. Esseen (1965), "Inequalities for the r^{th} Absolute Moment of a Sum of Random Variables, $1 \leq r \leq 2$," Annals of Math. Stat., Vol. 36, pp. 299-303.