The Annals of Probability
1981, Vol. 9, No. 3, 365-376

SUBADDITIVE EUCLIDEAN FUNCTIONALS AND NONLINEAR GROWTH IN GEOMETRIC PROBABILITY

BY J. MICHAEL STEELE

Stanford University

A limit theorem is established for a class of random processes (called here subadditive Euclidean functionals) which arise in problems of geometric probability. Particular examples include the length of shortest path through a random sample, the length of a rectilinear Steiner tree spanned by a sample, and the length of a minimal matching. Also, a uniform convergence theorem is proved which is needed in Karp's probabilistic algorithm for the traveling salesman problem.

1. Introduction. The main objective of the present paper is to show how the methodology of independent subadditive processes can be used to obtain strong limit laws for a wide class of problems in geometrical probability which exhibit nonlinear growth.

The problems studied here find their origin and principal motivation in a theorem of Beardwood, Halton and Hammersley (1959) of which the following is a special case.

For any bounded i.i.d. random variables $\{X_i\}$ with values in $\mathbb{R}^2$ the length of the shortest path through $\{X_1, X_2, \cdots, X_n\}$ is asymptotic to $cn^{1/2}$ with probability one.

Because of Karp's (1976) probabilistic algorithm for the traveling salesman problem, results like the above have gained accelerated practical interest. Motivated by algorithmic applications Papadimitriou and Steiglitz (1977) and Papadimitriou (1978) have taken pains to abstract the properties used in the Beardwood, Halton and Hammersley theorem. As a consequence, they have been able to treat other problems, including that of minimal matching of a random sample by Euclidean edges.

The tack taken here differs considerably from the method of Beardwood, Halton and Hammersley and is in the spirit of Kesten's lemma in the theory of independent subadditive processes (Kesten (1973), Hammersley (1974), Kingman (1976)). One benefit of the present approach is therefore a new proof of the Beardwood-Halton-Hammersley theorem, but a level of generality is maintained which permits immediate application to a number of other optimality problems in geometric probability.

The second section is devoted to developing the basic properties of subadditive Euclidean functionals which are the central object of study. The limit theorem proved there (Theorem 1) is established by a pure subadditivity argument which makes no appeal to the two-sided bounds sometimes available in specific problems.

The third and fourth sections extend Theorem 1 to nonuniform distributions and also weaken the monotonicity assumption. These sections then treat four specific examples. Section five provides a uniform convergence theorem which serves to rigorize one aspect of Karp's algorithm for the TSP. The final section makes brief comment on some unknown constants and on rates of convergence.

2. Subadditive Euclidean functionals. By L we denote a real valued function of the finite subsets of $\mathbb{R}^d$, $d \geq 2$. It will be assumed that L is a *Euclidean functional* by which it is indicated that the following properties hold:

Received August 27, 1979; revised March 1, 1980.

Research was partially supported by the National Science Foundation Grant MCS 78-07736.

AMS 1970 *subject classification*. 60D05, 60F15, 60C05, 60G55.

Key words and phrases. Subadditive process, traveling salesman problem, Steiner tree, Euclidean functional.

A1. $L(\alpha x_1, \alpha x_2, \cdots, \alpha x_n) = \alpha L(x_1, x_2, \cdots, x_n)$ for all real $\alpha > 0$.

A2. $L(x_1 + x, x_2 + x, \cdots, x_n + x) = L(x_1, x_2, \cdots, x_n)$ for all $x \in \mathbb{R}^d$.

Since L is a function on the finite subsets of $\mathbb{R}^d$, we also note that $L(x_1, x_2, \cdots, x_n)$ is the same as $L(x_{\sigma(1)}, x_{\sigma(2)}, \cdots, x_{\sigma(n)})$ for any permutation σ: [1, n] $\to$ [1, n]. The function L is also assumed to be *monotone*, i.e.,

A3. $L(x \cup A) \geq L(A)$ for any $x \in \mathbb{R}^d$ and finite subset A of $\mathbb{R}^d$.

Since L of the empty set is taken as zero, the monotonicity of L entails positivity; $L(A) \geq 0$ for all finite sets $A \subset \mathbb{R}^d$.

The required amount of boundedness of L is provided by an assumption of *finite variance*,

A4. $\mathrm{Var}(L(X_1, X_2, \cdots, X_n)) < \infty$ whenever X_i, $1 \leq i \leq n$, are independent and uniformly distributed in $[0, 1]^d$.

The preceeding assumptions are met in a huge number of contexts, and the most telling is a subadditivity restriction. Suppose that $\{Q_i: 1 \leq i \leq m^d\}$ is a partition of the d-cube $[0, 1]^d$ into cubes with edges parallel to the axle and of length $1/m$. Let $tQ_i = \{x : x = ty, y \in Q_i\}$. The *subadditivity hypothesis* is

A5. There exists a $C > 0$, such that for all positive integers m and positive reals t one has

$$L(\{x_1, x_2, \cdots, x_n\} \cap [0, t]^d) \leq \textstyle\sum_{i=1}^{m^d} L(\{x_1, x_2, \cdots, x_n\} \cap tQ_i) + Ctm^{d-1}.$$

The next result provides the key to the subsequent extentions.

THEOREM 1. *Suppose L is a monotone, Euclidean functional on $\mathbb{R}^d$ with finite variance which satisfies the subadditivity hypothesis. If $\{X_i : 1 \leq i < \infty\}$ are independent and uniformly distributed in $[0, 1]^d$, then there is a constant $\beta(L)$ such that*

$$\lim_{n\to\infty} L(X_1, X_2, \cdots, X_n)/n^{(d-1)/d} = \beta(L)$$

with probability one.

PROOF. We let Π denote a Poisson point process in $\mathbb{R}^d$ with uniform intensity parameter 1, and for any $A \subset \mathbb{R}^d$ $\Pi(A)$ denotes the random set of points in A. Next, let $\lambda(t) = L(\Pi([0, t]^d))$ and $\phi(t) = E\lambda(t)$. The first task is to prove

$$\lim_{t\to\infty} \frac{\phi(t)}{t^d} = \beta(L) \text{ exists.} \tag{2.1}$$

By the subadditivity of L,

$$\lambda(t) \leq \textstyle\sum_{i=1}^{m^d} L(\Pi(tQ_i)) + Cm^{d-1}t,$$

and since L is a Euclidean functional $EL(\Pi(tQ_i)) = \phi(t/m)$. Hence, one has

$$\phi(t) \leq m^d \phi\left(\frac{t}{m}\right) + Cm^{d-1}t.$$

Setting $t = mu$ and dividing by t^d yields

$$\frac{\phi(mu)}{(mu)^d} \leq \frac{\phi(u)}{u^d} + \frac{C}{u^{d-1}}. \tag{2.2}$$

If $\beta = \liminf \phi(u)/u^d$, we note that with $u = 1$ (2.2) implies by the monotonicity of $\phi(\cdot)$ that

$$\beta \leq \limsup_u \frac{\phi(u)}{u^d} \leq \limsup_{m\to\infty} \frac{\phi(m+1)}{m^d} \leq \phi(1) + C < \infty.$$

One now chooses u_0 so that $C/u_0^{d-1} \leq \varepsilon$ and $\phi(u_0)/u_0^d \leq \beta + \varepsilon$ to obtain for all $m = 1, 2, \cdots$

$$\frac{\phi(mu_0)}{(mu_0)^d} \le (\beta + \varepsilon) + \varepsilon.$$

Again, by the monotonicity of $\phi(\cdot)$ and the fact that $((m+1)u_0)^d/(mu_0)^d \to 1$,

$$\limsup \frac{\phi(u)}{u^d} \le \beta + 2\varepsilon,$$

so the limit, $\lim_{u\to\infty} \phi(u)/u^d$ exists and is finite. It will be denoted $\beta(L)$.

The second task is the calculation of a sum of variances by use of recursions given by subadditivity. Set $m = 2$ and let $\lambda_i(t) = L(\Pi(tQ_i))$. By A5 one has

$$\lambda(2t) \le \sum_{i=1}^{2^d} \lambda_i(2t) + C2^d t.$$

Define $\tilde{\lambda}(t) = \lambda(t) + 2Ct$ and $\tilde{\lambda}_i(t) = \lambda_i(2t) + 2Ct$, $1 \le i \le 2^d$. The nice fact is that now $\tilde{\lambda}_i(t)$ are i.i.d.,

$$\tilde{\lambda}(2t) \le \sum_{i=1}^{2^d} \tilde{\lambda}_i(t), \tag{2.3}$$

and

(2.4) the $\tilde{\lambda}_i(t)$ have the same distribution as $\tilde{\lambda}(t)$.

Next, let $\tilde{\phi}(t) = E\tilde{\lambda}(t)$ and $\psi(t) = (E(\tilde{\lambda}(t))^2)^{1/2}$. One has $V(t) = \mathrm{Var}\lambda(t) = \mathrm{Var}\,\tilde{\lambda}(t) = \psi^2(t) - \tilde{\phi}^2(t)$ and

$$\tilde{\phi}(t)/t^d = \phi(t)/t^d + 2C/t^{d-1} \to \beta(L) \quad \text{as} \quad t \to \infty. \tag{2.5}$$

Taking squares and expectations in (2.3) yields

$$\psi^2(2t) \le 2^d\psi^2(t) + (2^{2d} - 2^d)\tilde{\phi}^2(t).$$

So,

$$V(2t) = \psi^2(2t) - \tilde{\phi}^2(2t) \le 2^d V(t) + 2^{2d}\tilde{\phi}^2(t) - \tilde{\phi}^2(2t).$$

Dividing by $(2t)^{2d}$ yields

$$\frac{V(2t)}{(2t)^{2d}} - \frac{V(t)}{2^d t^{2d}} \le \frac{\tilde{\phi}^2(t)}{t^{2d}} - \frac{\tilde{\phi}^2(2t)}{(2t)^{2d}}.$$

Applying this result for $t, 2t, \cdots, 2^{M-1}t$ and summing,

$$\sum_{k=1}^{M} \frac{V(2^k t)}{(2^k t)^{2d}} - \frac{1}{2^d}\sum_{k=0}^{M-1} \frac{V(2^k t)}{(2^k t)^{2d}} \le \frac{\tilde{\phi}^2(t)}{t^{2d}} - \frac{\tilde{\phi}^2(2^M t)}{(2^M t)} \le \frac{\tilde{\phi}^2(t)}{t^{2d}}.$$

Finally, one finds for all $t > 0$,

$$\sum_{k=1}^{\infty} \frac{V(2^k t)}{(2^k t)^{2d}} \le (1 - 2^{-d})^{-1}\left(\frac{V(t)}{t^{2d}} + \frac{\tilde{\phi}^2(t)}{t^{2d}}\right) < \infty. \tag{2.6}$$

Now, let $N(t)$ be the Poisson counting process on $[0, \infty)$ and suppose $\{X_i : 1 \le i < \infty\}$ are i.i.d. uniform on $[0, 1]^d$. Since L is a Euclidean functional well-known properties of the Poisson process Π show that $\lambda(t) = L(\Pi([0, t]^d))$ has the same distribution as $tL(X_1, X_2, \cdots, X_{N(t^d)})$. Since by (2.1), $\phi(t)/t^d \to \beta(L)$ Chebyshev's inequality applied to (2.6) yields for $\varepsilon > 0$,

$$\sum_{k=0}^{\infty} P\left(\left|\frac{t2^k L(X_1, X_2, \cdots, X_{N((t2^k)^d)})}{(t2^k)^d} - \beta(L)\right| > \varepsilon\right) < \infty;$$

and consequently for each $t > 0$

$$\lim_{k\to\infty} \frac{L(X_1, X_2, \cdots, X_{N((2^k t)^d))})}{(t2^k)^{d-1}} = \beta(L) \quad \text{a.s.} \tag{2.7}$$

The monotonicity of L will now be used in a slightly more subtle way than was done with ϕ.

Let p be a fixed positive integer and note for each real $s \geq 2^P$ *there is an integer* t, $2^p \leq t < 2^p+1$ and an integer $k \geq 0$ so that $2^k t \leq s \leq 2^k(t+1)$. Since L is monotone,

$$L(X_1, X_2, \cdots, X_{N((2^k t)^d)}) \leq L(X_1, X_2, \cdots, X_{N(s^d)}) \leq L(X_1, X_2, \cdots, X_{N((2^k(t+1))^d)}).$$

Since the set of t's, $2^p \leq t \leq 2^{p+1}$ is finite (2.7) implies

$$\limsup_{s\to\infty} L(X_1, X_2, \cdots, X_{N(s^d)})/s^{d-1} \leq \beta(L)(1+2^{-p})^{d-1}$$

and

$$\liminf_{s\to\infty} L(X_1, X_2, \cdots, X_{N(s^d)})/s^{d-1} \geq \beta(L)(1+2^{-p})^{1-d}$$

Since p was arbitrary,

$$\lim_{s\to\infty} L(X_1, X_2, \cdots, X_{N(s^d)})/s^{d-1} = \beta(L) \quad \text{a.s.} \tag{2.8}$$

Next, let $\tau(n)$ be defined so that $N(\tau(n)^d) = n$, and note by the elementary renewal theorem

$$\tau(n)/n^{1/d} \to 1 \quad \text{a.s.} \tag{2.9}$$

By the definition of $\tau(n)$ one has

$$L(X_1, X_2, \cdots, X_n)/n^{(d-1)/d} = \{L(X_1, X_2, \cdots, X_{N(\tau(n)^d)})/\tau(n)^{d-1}\}\{\tau(n)^{d-1}/n^{(d-1)/d}\}.$$

So applying (2.8) and (2.9) to the first and second factors respectively, the theorem is proved. □

3. Nonuniformly distributed random variables. To extend the preceding result to nonuniformly distributed random variables some additional "localization" properties of L are needed. A Euclidean functional L will be called *scale bounded*, provided the following assumption holds:

A6. There is a constant B such that

$$L(x_1, x_2, \cdots, x_n)/tn^{(d-1)/d} \leq B \quad \text{for all} \quad n \geq 1,\ t \geq 1,$$

and

$$\{x_1, x_2, \cdots, x_n\} \subset [0, t]^d.$$

Also, L is called *simply subadditive* provided

A7. There is a constant B such that

$$L(A_1 \cup A_2) \leq L(A_1) + L(A_2) + tB$$

for any finite subsets A_1 and A_2 of $[0, t]^d$.

If $\mu(A) = p(X_i \in A)$ is the distribution function of the X_i, the Lebesgue decomposition theorem allows us to write $\mu = \mu_a + \mu_s$, where μ_a is absolutely continuous and μ_s is singular (with respect to Lebesgue measure). The support of μ_s will be called the *singular support* of the X_i, and the next lemma shows that the contribution to L of observations in the singular support is very small.

LEMMA 3.1. *Suppose that L is a scale bounded, simply subadditive Euclidean functional. Suppose also that $\{X_i: 1 \leq i < \infty\}$ are* i.i.d. *random variables with bounded singular support E, then*

$$L(\{X_1, X_2, \cdots, X_n\} \cap E) = o(n^{(d-1)/d}) \quad \text{a.s.} \tag{3.1}$$

PROOF. We can suppose that $E \subset [0, 1]^d$ and then choose disjoint cubes Q_i, $1 \leq i \leq M$, so that the Lebesgue measure of the Q_i, $m(Q_i)$, satisfies $\sum_{i=1}^M m(Q_i) < \varepsilon$ while $P(X_i \in E \setminus \cup_{i=1}^M Q_i) < \varepsilon$. Applying simple subadditivity yields

$$L(\{X_1, X_2, \cdots, X_n\} \cap E) \le MB + \textstyle\sum_{j=1}^M L\{X_i : X_i \in Q_j\} + L\{X_i : X_i \in E \setminus \cup_{j=1}^M Q_j\}. \tag{3.2}$$

By scale boundedness and Hölder's inequality,

$$\begin{aligned}\textstyle\sum_{j=1}^M L\{X_i : X_i \in Q_j\} &\le B \textstyle\sum_{j=1}^M (\sum_{i=1}^n 1_{Q_j}(X_i))^{(d-1)/d} (m(Q_j))^{1/d} \\ &\le B (\textstyle\sum_{j=1}^M \sum_{i=1}^n 1_{Q_j}(X_i))^{(d-1)/d} (\sum_{j=1}^M m(Q_j))^{1/d} \\ &\le B n^{(d-1)/d} \varepsilon^{1/d}.\end{aligned}$$

Similarly, setting $Q = E \setminus \cup_{j=1}^M Q_j$ one has by scale boundedness

$$L\{X_i : X_i \in Q\} \le B (\textstyle\sum_{i=1}^n 1_Q(X_i))^{(d-1)/d},$$

and the last term is asymptotically no larger than $Bn^{(d-1)/d}\varepsilon^{(d-1)/d}$ with probability one. Finally, the arbitrariness of $\varepsilon > 0$ completes the proof of the lemma.

The next assumption is the last one which will be needed. A Euclidean functional L will be called *upper-linear* provided

A8. For any finite collection of cubes Q_i, $1 \le i \le s$ with edges parallel to the axes and for any infinite sequence x_i, $1 \le i < \infty$, in $\mathbb{R}^d$ one has

$$\textstyle\sum_{i=1}^s L(\{x_1, x_2, \cdots, x_n\} \cap Q_i) \le L(\{x_1, x_2, \cdots, x_n\} \cap \cup_{i=1}^s Q_i) + o(n^{(d-1)/d}).$$

This condition will now be put to work.

LEMMA 3.2. *Suppose L satisfies assumptions* A1–A8. *Suppose also that Y_i, $1 \le i < \infty$, are* i.i.d. *random variables with bounded support and absolutely continuous part $\phi(x) = \sum_{i=1}^s \alpha_i 1_{Q(i)}(x)$ where $Q(i)$, $1 \le i \le s$, are disjoint cubes with edges parallel to the axes. One then has*

$$\lim_{n\to\infty} L(Y_1, Y_2, \cdots, Y_n)/n^{(d-1)/d} = \beta(L) \int_{\mathbb{R}^d} \phi(x)^{(d-1)/d}\, dx \quad \text{a.s.}$$

where $\beta(L)$ is a constant depending only on L.

PROOF. Write E for the singular support of the $\{Y_i\}$ and assume without loss of generality that the whole support of the $\{Y_i\}$ is contained in $[0, 1]^d$. By simple subadditivity

$$L(Y_1, Y_2, \cdots, Y_n) \le L(\{Y_1, Y_2, \cdots, Y_n\} \cap E) + \textstyle\sum_{i=1}^s L(\{Y_1, Y_2, \cdots, Y_n\} \cap Q(i)) + sB. \tag{3.3}$$

Since $\{Y_1, Y_2, \cdots, Y_n\} \cap Q_i$ is a uniform sample in Q_i, Theorem 1 and A1 imply

$$\lim_{n\to\infty} L(\{Y_1, Y_2, \cdots, Y_n\} \cap Q_i)/(\textstyle\sum_{j=1}^n 1_{Q(i)}(Y_j))^{(d-1)/d} = \beta(L)(m(Q(i)))^{1/d}$$

with probability one. Since $\sum_{j=1}^n 1_{Q(i)}(Y_j) \sim \alpha_i m(Q_i) n$ a.s. one concludes

$$\lim_{n\to\infty} L(\{Y_1, Y_2, \cdots, Y_n\} \cap Q(i))/n^{(d-1)/d} = \beta(L) m(Q_i) \alpha_i^{(d-1)/d},$$

so returning to (3.3) and applying Lemma 3.1 gives

$$\limsup_{n\to\infty} L(Y_1, Y_2, \cdots, Y_n)/n^{(d-1)/d} \le \beta(L) \int_{\mathbb{R}^d} \phi(x)^{(d-1)/d}\, dx.$$

To obtain a comparable bound on the lim inf, one procedes as before after noticing that monotonicity and upper-linearity imply

$$L(Y_1, Y_2, \cdots, Y_n) \ge \textstyle\sum_{i=1}^s L(\{Y_1, Y_2, \cdots, Y_n\} \cap Q(i)) - o(n^{(d-1)/d}).$$

The main result of this section can now be obtained from Lemma 3.2 and a "thinning argument."

THEOREM 2. *Suppose L is a Euclidean functional which satisfies assumptions* A1–A8. *There is a constant $\beta(L)$ such that*

$$\lim_{n\to\infty} L(X_1, X_2, \cdots, X_n)/n^{(d-1)/d} = \beta(L) \int_{\mathbb{R}^d} f(x)^{(d-1)/d}\, dx \quad \text{a.s.}$$

for any independent identically distributed random variables $\{X_i\}$ with bounded support and asolutely continuous part $f(x)\,dx$.

PROOF. As before, suppose the $\{X_i\}$ have support contained in $[0, 1]^d$ and that E is the singular support. Next choose a $\phi(x) = \sum_{i=1}^{s} \alpha_i 1_{Q(i)}(x)$ where $Q(i)$, $1 \le i \le s$, are disjoint cubes with edges parallel to the axes. The "thinning domain" A is defined by

$$A = \{x : f(x) \le \phi(x)\}. \tag{3.4}$$

Now a sequence of random variables Y_i, $1 \le i < \infty$, with density $\phi(x)$ can be generated from the X_i as follows. If $X_i \in A \cup E$, then Y_i is set equal to some fixed $a_0 \in A$; and if $X_i \notin A \cup E$, then Y_i is taken to be X_i or a_0 according to an independent randomization with probabilities $p = \phi(X_i)/f(X_i)$ and $1 - p$ respectively.

By Y_i' we denote a third sequence of i.i.d. random variables. These are chosen to have bounded support and absolutely continuous part $\phi(x)$.

The main point of the previous construction is that the two sets of random variables $\{Y_1, Y_2, \cdots, Y_n\} \cap (A \cup E)^c$ and $\{Y_1', Y_2', \cdots, Y_n'\} \cap (A \cup E)^c$ now have the same distributions. One then has that the two processes

$$\{L(\{Y_1, Y_2, \cdots, Y_n\} \cap (A \cup E)^c), n \ge 1\} = \{L_n, n \ge 1\}$$

and

$$\{L(\{Y_1', Y_2', \cdots, Y_n'\} \cap (A \cup E)^c), n \ge 1\} = \{L_n', n \ge 1\}$$

also have the same joint distributions. From this and the Hewitt-Savage zero-one law, we get

$$\liminf_{n\to\infty} L_n/n^{(d-1)/d} = \liminf_{n\to\infty} L_n'/n^{(d-1)/d} \quad \text{a.s.} \tag{3.5}$$

By simple subadditivity

$$L_n' \ge L(Y_1', Y_2', \cdots, Y_n') - L(\{Y_1', Y_2', \cdots, Y_n'\} \cap (A \cup E)) - B,$$

so by applying Lemma 3.2 to the first term of the right-hand side, then applying scale boundedness and the law of large numbers to the second one, we get

$$\liminf_{n\to\infty} L_n'/n^{(d-1)/d} \ge \beta(L) \int_{\mathbb{R}^d} \phi(x)^{(d-1)/d}\, dx - B\left(\int_{A\cup E} \phi(x)\, dx\right)^{(d-1)/d}. \tag{3.6}$$

Here one should note that $\beta(L)$ does not depend upon the auxiliary process $\{L_n' : n \ge 1\}$; Lemma 3.2 shows that $\beta(L)$ depends only on the functional L.

We can now apply the monotonicity of L to obtain

$$L(X_1, X_2, \cdots, X_n) \ge L(\{X_1, X_2, \cdots, X_n\} \cap (A \cup E)^c) \ge L_n,$$

so (3.5), (3.6), and the arbitrariness of ϕ in (3.4) imply

$$\liminf_{n\to\infty} L(X_1, X_2, \cdots, X_n)/n^{(d-1)/d} \ge \beta(L) \int_{\mathbb{R}^d} f(x)^{(d-1)/d}\, dx.$$

A slightly more elaborate thinning argument will be used to obtain the opposite inequality. We will take $\phi(x)$ as before but set $A = \{x : \phi(x) \le f(x)\}$. Let $\{Y_i\}$ be an i.i.d.

sequence with absolutely continuous part $\phi(x)$ and an atom at $a_0 \in A$ of size $1 - \int_{\mathbb{R}^d} \phi(x) dx$. For each i define $X_i' = Y_i$ if $Y_i \in A$, otherwise choose X_i' as Y_i or a_0 according to an independent randomization with probabilities $f(Y_i)/\phi(Y_i)$ and $1 - f(Y_i)/\phi(Y_i)$ respectively.

Let E denote the singular support of the $\{X_i\}$. Also, let $\{\tau_1 < \tau_2 < \cdots\} = \{i : X_i' \notin A\}$ and $\{\sigma_1 < \sigma_2 < \cdots\} = \{i : X_i \notin A \cup E\}$. The key observations are that the two-dimensional processes $\{(X_{\sigma_i}, \sigma_i)\}$ and $\{(X'_{\tau_i}, \tau_i)\}$ have the same distribution and that $\{X'_{\tau_k} : k \geq 1\}$ is just a subsequence of $\{Y_i : Y_i \notin A\}$. One now calculates,

$$
\begin{aligned}
&\limsup_{n\to\infty} L(X_1, X_2, \cdots, X_n)/n^{(d-1)/d} \\
&\qquad \leq \limsup_{n\to\infty} L(\{X_1, X_2, \cdots, X_n\} \cap E^c)/n^{(d-1)/d} \\
&\qquad \leq \limsup_{n\to\infty} L(\{X_1, X_2, \cdots, X_n\} \cap E^c \cap A^c)/n^{(d-1)/d} \\
&\qquad\quad + BP(X_1 \in E^c \cap A)^{(d-1)/d}.
\end{aligned} \tag{3.7}
$$

Further one has

$$
\begin{aligned}
&\limsup_{n\to\infty} L(\{X_1, X_2, \cdots, X_n\} \cap E^c \cap A^c)/n^{(d-1)/d} \\
&\qquad = \limsup_{n\to\infty} L(\{X_{\sigma_k} : \sigma_k \leq n\})/n^{(d-1)/d} \\
&\qquad = \limsup_{n\to\infty} L(\{X'_{\tau_k} : \tau_k \leq n\})/n^{(d-1)/d} \\
&\qquad \leq \limsup_{n\to\infty} L(Y_1, Y_2, \cdots, Y_n)/n^{(d-1)/d}
\end{aligned} \tag{3.8}
$$

Now Lemma 3.2 implies that the last limit superior actually equals $\beta(L) \int_{\mathbb{R}^d} \phi(x)^{(d-1)/d} dx$. Finally, one can choose ϕ and A so that $P(X_1 \in E^c \cap A)$ is nearly zero, and $\int_{\mathbb{R}^d} \phi(x)^{(d-1)/d} dx$ is nearly $\int_{\mathbb{R}^d} f(x)^{(d-1)/d}$. Used in (3.7) and (3.8), this implies

$$
\limsup_{n\to\infty} L(X_1, X_2, \cdots, X_n)/n^{(d-1)/d} \leq \beta(L) \int_{\mathbb{R}^d} f(x)^{(d-1)/d} dx
$$

which completes the proof.

4. Selected applications.

A. *Beardwood-Halton-Hammersley theorem.* To apply the preceding results to the Euclidean functional $L_0(x_1, x_2, \cdots, x_n)$ which equals the *length of the shortest path* through the points $\{x_1, x_2, \cdots, x_n\}$, one much check several assumptions. It is trivial that A1–A4 hold. Assumption A5 is not too hard to check, but it is perhaps most easily obtained as a consequence of the following well-known lemma.

LEMMA 4.1. *There is a constant $c = c_d$ such that for any $\{x_1, x_2, \cdots, x_n\} \subset [0, t]^d$,*

$$
L_0(x_1, x_2, \cdots, x_n) \leq c n^{(d-1)/d} t.
$$

PROOF. This has been treated by Fejes Toth (1940), S. Verblunsky (1951), and L. Few (1955), who devoted considerable effort to determining the best value of c. For a crude value of c the lemma is easily proved by partitioning the d-cube.

Now to justify A5 we construct a path through $\{x_1, x_2, \cdots, x_n\}$. First take the set of m^d path segments of length $L_0(tQ_i \cap \{x_1, x_2, \cdots, x_n\})$ through $tQ_i \cap \{x_1, x_2, \cdots, x_n\}$, and then consider the set of $2m^d$ points which are end points of these segments. By Lemma 4.1 (and a change of scale) there is a path through this set of end points of length not greater than $c2^{(d-1)/d} m^{d-1} t$. We therefore set a path through $\{x_1, x_2, \cdots, x_n\} \cap [0, t]^d$ with length not more than

$$
\sum_{i=1}^{m^d} L(\{x_1, x_2, \cdots, x_n\} \cap tQ_i) + Ctm^{d-1} \qquad \text{with} \quad C = c2^{(d-1)/d}.
$$

This completes the justification of A1–A5 and thus gives a proof of the Beardwood-Halton-Hammersley theorem. To push the result to cover the case of nonuniformly distributed random variables, we need to also verify Assumptions A6 and A7 of Theorem 2. This requires another lemma.

LEMMA 4.2. *For any finite collection Q_i, $1 \le i \le s$, of disjoint cubes and any infinite sequence $\{x_1, x_2, \cdots\} \subset \mathbb{R}^d$ one has*

$$\sum_{i=1}^{s} L_0(\{x_1, x_2, \cdots, x_n\} \cap Q_i) \le L_0(\{x_1, x_2, \cdots x_n\} \cap (\cup_{i=1}^{s} Q_i)) + O(n^{(d-2)/(d-1)}). \tag{4.1}$$

PROOF. Let P be a path which attains $L_0(\{x_1, x_2, \cdots, x_n\} \cap \cup_{i=1}^{s} Q_i) = l$ and note that by a preliminary perturbation which changes l only slightly one can suppose no segment of P is contained in any face of Q_i. Let $P_i = P \cap Q_i$ and let P_{ij}, $j = 1, 2, \cdots$ be the connected components of P_i which contain an element of $\{x_1, x_2, \cdots, x_n\}$. Let a_{ij} and b_{ij} be the points of P_{ij} which intersect ∂Q_i. (One gets at most two points since ∂Q_i cannot contain a segment and the P_{ij} are connected). Let s_i be the length of the edges of Q_i and let $F_1, F_2, \cdots, F_{2^d}$ be the faces. The set $F_k \cap \{a_{ij}: j = 1, 2, \cdots, n\} = A_{ik}$ is contained in a $d-1$ cube so by Lemma 4.1 there is a path through the elements of A_{ik} of length $c_{d-1} s_i |A_{ik}|^{(d-2)/(d-1)}$. Hence, there is a path through $\cup_{i=1}^{2^d} A_{ik}$ of length no greater than $c_{d-1} s_i \sum_{k=1}^{2^d} |A_{ik}|^{(d-2)/(d-1)} + 2c_d s_i 2^{d/2}$. Since the left side of (4.1) is not greater than $L_0(\{x_1, x_2, \cdots, x_n) \cap \cup_{i=1}^{s} Q_i)$ plus the lengths of the paths through the a_{ij} and b_{ij}, one has

$$\begin{aligned} l \le L_0(\{x_1, x_2, \cdots, x_n\} \cap \cup_{i=1}^{s} Q_i) + \gamma_1 \sum_{i=1}^{s} \sum_{k=1}^{2^d} |A_{ik}|^{(d-2)/(d-1)} \\ + \gamma_2 \sum_{i=1}^{s} \sum_{k=1}^{2^d} |B_{ik}|^{(d-2)/(d-1)} + \gamma_3, \end{aligned} \tag{4.2}$$

where the B_{ik} are defined analogously to the A_{ik} and $\gamma_1, \gamma_2, \gamma_3$ are constants not depending on n.

Now by Hölder's inequality

$$\begin{aligned} \sum_{i=1}^{s} \sum_{k=1}^{2^d} |A_{ik}|^{(d-2)/(d-1)} &\le (\sum_{i=1}^{s} \sum_{k=1}^{2^d} |A_{ik}|)^{(d-2)/(d-1)} (\sum_{i=1}^{s} \sum_{k=1}^{2^d} 1^{(d-1)})^{1/(d-1)} \\ &\le n^{(d-2)/(d-1)} (s2^d)^{1/(d-1)}. \end{aligned}$$

Returning to (4.2) one has, as claimed, that

$$l \le L_0(\{x_1, x_2, \cdots, x_n\} \cap \cup_{i=1}^{s} Q_i) + O(n^{(d-2)/(d-1)}). \ \square$$

By Lemma 4.1 one sees that L_0 is scale bounded (A6), and it is also trivial to see L_0 is simply subadditive (A7). The last assumption of upper-linearity (A8) is a consequence of Lemma 4.2 since $(n^{(d-2)/(d-1)}) = O(n^{(d-1)/d})$.

This completes the proof of the Beardwood-Halton-Hammersley theorem, but comment on the nature of this proof will be postponed to the last section. First consideration will be given to additional applications of Theorems 1 and 2.

B. *Papadimitriou's matching problem.* Let $L_1(x_1, x_2, \cdots, x_n)$ denote the length of the least Euclidean matching of the points $\{x_1, x_2, \cdots, x_n\} \subset \mathbb{R}^d$, i.e.,

$$L_1(x_1, x_2, \cdots, x_n) = \min_\sigma \sum_{i=1}^{[n/2]} \| x_{\sigma(2i-1)} - x_{\sigma(2i)} \| \tag{4.3}$$

where $\|x - y\|$ is the distance from x to y and the minimum is over all permutations $\sigma: [1, n] \to [1, n]$.

To treat limit theory of L_1, it is useful to generalize Theorems 1 and 2 slightly to accommodate functionals which do not quite satisfy the monotonicity assumption A3. One can call a Euclidean functional L *sufficiently monotone* provided

(A3)′. There exist a positive sequence $r_n = o(n^{(d-1)/d})$ such that for any infinite sequence $\{x_1, x_2, \cdots\} \subset \mathbb{R}^d$ and any $m \ge n$ one has

$$L(x_1, x_2, \cdots, x_m) \ge L(x_1, x_2, \cdots, x_n) - r_n.$$

With this assumption the proof of Theorems 1 and 2 can be repeated virtually without change to give the following:

THEOREM 3. (a). *If L satisfies* A1, A2, (A3)′, *and* (A4), *then*

$$\lim_{n\to\infty} L(X_1, X_2, \cdots, X_n)/n^{(d-1)/d} = \beta(L) \quad \text{a.s.}$$

provided $\{X_i\}$ *are independent and uniformly distributed in* $[0, 1]^d$.

(b) *If, in addition, L satisfies* (A5–A8), *then*

$$\lim L(X_1, X_2, \cdots, X_n)/n^{(d-1)/d} = \beta(L) \int_{\mathbb{R}^d} f(x)^{(d-1)/d}\, dx \quad \text{a.s.}$$

provided $\{X_i\}$ *are* i.i.d. *with bounded support and absolutely continuous part* $f(x)$.

To apply Theorem 3 to the least matching functional L_1, one first notes that all of the assumptions except A5 and A8 are easily checked. (To show A6, one applies Lemma 4.1 and the bound $L_1 \le L_0$.) To verify A5, one gets a matching of $\{x_1, x_2, \cdots, x_n\} \cap [0, t]^d$ by taking the matchings of $\{x_1, x_2, \cdots, x_n\} \cap Q_i$ together with a matching of the collection S of elements which are not used in any matching of a $\{x_1, x_2, \cdots, x_n\} \cap Q_i$. Since $|S| \le m^d$, A6 (or Lemma 4.1) now shows

$$L_1(\{x_1, x_2, \cdots, x_n\} \cap [0, t]^d) \le \textstyle\sum_{i=1}^{m^d} L_1(\{x_1, x_2, \cdots, x_n\} \cap Q_i) + Bt(m^d)^{(d-1)/d}$$

which simplifies precisely to A5.

This lemma completes the considerations which are needed to make the first part of Theorem 1 applicable to the matching functional L_1. To be able to use the second part, one further bound is needed.

LEMMA 4.3. *For any cubes Q_i, $1 \le i < s$,*

$$\textstyle\sum_{i=1}^{s} L_1(\{x_1, x_2, \cdots, x_n\} \cap Q_i) \le L_1(\{x_1, x_2, \cdots, x_n\} \cap \cup_{i=1}^{s} Q_i) + O(n^{(d-2)/(d-1)})$$

for any infinite sequence x_i, $1 \le i < \infty$, in $\mathbb{R}^d$.

PROOF. Let A be a set of arcs which attain $l = L_1(\{x_1, x_2, \cdots, x_n\} \cap \cup_{i=1}^{s} Q_i)$. By the usual perturbation argument there will be no loss in assuming A has no segment in any face of any Q_i. Hence, one can let C_i be the set of points in $A \cap \partial Q_i$ which are endpoints of segments which contain an element of $Q_i \cap \{x_1, x_2, \cdots, x_n\}$. By this definition we note $\sum_{i=1}^{s} |C_i| \le n$. Now $L_1(\{x_1, x_2, \cdots, x_n\} \cap Q_i)$ is certainly no larger than the sum of the length of the segments of A in Q_i plus the length of a patch through all of C_i. One then decomposes C_i into the subsets on faces and applies Lemma 4.1 and Hölder's inequality as in Lemma 4.2. □

Remarking again that $n^{(d-2)/(d-1)} = o(n^{(d-1)/d})$, one has upper-linearity (A8) as a consequence of the preceding lemma. Hence, both parts of Theorem 3 apply to the matching functional L_1.

C. *Steiner trees and rectilinear Steiner trees.* A Steiner tree on $\{x_1, x_2, \cdots, x_n\} = S \subset \mathbb{R}^d$ is a connected graph which contains $\{x_1, x_2, \cdots, x_n\}$ which has the least total sum of edge lengths among all such graphs. A rectilinear Steiner tree is defined similarly except that the edges are required to be parallel to the axes. One can naturally define two corresponding functionals, and these will be denoted by L_2 and L_3.

While both of these functionals were mentioned in Beardwood, Halton, and Hammersley (1958), their limit theory was not explicitly developed in that paper. Since A1–A5 are trivial to verify for L_2 and L_3, one sees that Theorem 1 applies immediately. To check the conditions of Theorem 2, first note that $L_2 \le L_0$ and $L_3 \le \sqrt{d} L_2$ so Lemma 4.1 gives scale boundedness (A6). Since simple subadditivity (A7) is trivial, only the last assumption (A8) needs individual attention; in this case, the proof of Lemma 4.2 can be applied almost without change.

5. Uniform convergence and Karp's algorithm. In Karp (1976) an algorithm for the probabilistic solution to the traveling salesman's problem is given which hinges on the Beardwood-Halton-Hammersley theorem and which actually assumes a uniform version of that theorem. The main objective of this section is to rigorize one part of Karp's procedure by establishing a uniform version of Theorem 2. For a further application of the present methods to the "independent model" of Karp's problem (B. Weide (1978)) see Steele (1979).

THEOREM 4. *Let $\mathscr{C}$ denote the class of convex Borel subsets of $\mathbb{R}^d$. If A Euclidean functional L satisfies assumptions* A1–A8, *and $\{X_i\}$ is a sequence of* i.i.d. *random variables with bounded support and absolutely continuous part given by $f(x)$, then*

$$\sup_{C\in\mathscr{C}} | L(\{X_1, X_2, \cdots, X_n\} \cap C)/n^{(d-1)/d} - \beta(L) \int_C f(x)^{(d-1)/d}\, dx |$$

converges to 0 a.s. *as $n \to \infty$.*

PROOF. Let E denote the singular support of the $\{X_i\}$ and assume without loss that the whole support is contained in $[0, 1]^d$. Let $m(\cdot)$ denote Lebesgue measure and recall that $\{C \cap [0, 1]^d : C \in \mathscr{C}\} = \mathscr{C}'$ is compact in the Hausdorf metric $d(A, B) = m(A \triangle B)$. Defining $d_f(A, B) = \int_{A\triangle B} f(x)^{(d-1)/d}\, dx$, one can use the compactness $\mathscr{C}'$ under d to prove compactness under d_f (e.g., by writing $f(x)^{(d-1)/d} = f_1(x) + f_2(x)$ where $f_1 \in L^2([0, 1]^d)$ and $\int_{\mathbb{R}^d} f_2(x)\, dx < \varepsilon$ and using Schwartz' inequality). Hence, we can choose C_i, $1 \le i \le k$, such that for all $C \in \mathscr{C}'$ there is a C_i such that

$$\int_{C_i \triangle C} f(x)^{(d-1)/d}\, dx \le \varepsilon.$$

Next, by simple subadditivity applied twice,

$$| L(\{X_1, X_2, \cdots, X_n\} \cap C) - L(\{X_1, X_2, \cdots, X_n\} \cap C_k) | \le B + L(\{X_1, X_2, \cdots, X_n\} \cap C \triangle C_k).$$

Now, consider the class of sets

$$S = \{A : A = C \triangle C', \int_A f(x)\, dx \le \varepsilon^{d/(d-1)}, C, C' \in \mathscr{C}\},$$

and note for $A \in S$, $\int_A f(x)^{(d-1)/d}\, dx \le \varepsilon$. The basic bounds are the following:

$$\sup_{C\in\mathscr{C}} \left| L(\{X_1, X_2, \cdots, X_n\} \cap C)/n^{(d-1)/d} - \int_C f(x)^{(d-1)/d}\, dx \right|$$

$$\le \max_{1\le i\le k} \left| L(\{X_1, X_2, \cdots, X_n\} \cap C_i)/n^{(d-1)/d} - \int_{C_i} f(x)^{(d-1)/d}\, dx \right| + \sup_{A\in S} (L(\{X_1, X_2, \cdots, X_n\} \cap A) + B)/n^{(d-1)d} + \varepsilon. \tag{5.1}$$

The first term on the right in (5.1) goes to zero a.s. by Theorem 2. For the second term note for all $A \in S$

$$L(\{X_1, X_2, \cdots, X_n\} \cap A) \le L(\{X_1, X_2, \cdots, X_n\} \cap E)$$

$$+ L(\{X_1, X_2, \cdots, X_n\} \cap E^c \cap A) + B$$
$$\leq L(\{X_1, X_2, \cdots, X_n\} \cap E)$$
$$+ B(\textstyle\sum_{i=1}^n 1_{A\cap E^c}(X_i))^{(d-1)/d} + B. \quad (5.2)$$

Since $L(\{X_1, X_2, \cdots, X_n\} \cap E)/n^{(d-1)/d}$ goes to zero a.s. by Lemma 3.1, it will suffice to verify that

$$\lim_{n\to\infty} \frac{1}{n} \sup_{A\in S} \textstyle\sum_{i=1}^n 1_{A\cap E^c}(X_i) \leq \varepsilon^{d/(d-1)}. \quad (5.3)$$

Since $P(X_i \in A \cap E^c) = \int_A f(x)\,dx \leq \varepsilon^{d/(d-1)}$, (5.3) is an easy consequence of the fact that the class of convex sets is a uniformity class, i.e., the law of large numbers applies uniformly to the sums $(1/n) \sum_{i=1}^n 1_C(Y_i)$ over all $C \in \mathscr{C}$ provided $\sup_C P(Y_1 \in \partial C) = 0$. (See, e.g., Ranga Rao (1962) or Steele (1978)).

6. Remarks on constants and rates. One of the persistently interesting aspects of subadditive methods is that one proves convergence to a constant which is unknown and sometimes seems unknowable. For the shortest path functional the best known bounds are to be found in Beardwood-Halton-Hammersley (1959) which improves upon Mahalanobis (1940), Marks (1948), and Ghosh (1949). Papadimitriou (1978) gives the best known bounds for the matching functional. While no independent bounds are known on the constant for the Steiner functional, the rectilinear Steiner problem has recently been studied by F. R. K. Chung and R. L. Graham (1980) and F. R. K. Chung and F. K. Hwang (1979).

A second problem of interest is that of rates of convergence. At the level of generality of Theorem 1 it is unlikely that one can say anything about rates of convergence, but by considering processes which have two-sided bounds, as in Theorem 2, it is much more likely that a rate result can be obtained. Some preliminary results in this direction have been obtained jointly with T. L. Lai and may be reported in a subsequent paper.

Acknowledgment. I would like to thank Patrick Billingsley and Steve Lalley for their detailed comments on an earlier draft of this paper.

REFERENCES

BEARDWOOD, J., HALTON, J. H. and HAMMERSLEY, J. M. (1959). The shortest path through many points. *Proc. Cambridge Philos. Soc* **55** 299–327.

CHUNG, F. R. K. and GRAHAM, R. L. (1980). On Steiner trees for bounded point sets. Unpublished manuscript.

CHUNG, F. R. K. and HWANG, F. K. (1979). The largest minimal rectilinear Steiner trees for a set of n points enclosed in a rectangle with given perimeter. *Networks* **9** 19–36.

FEJES, L. T. (1940). Über einen geometrischen Satz. *Math. Z.* **46** 83–85.

FEW, L. (1955). The shortest path and the shortest road through n points. *Mathematika* **2** 141–144.

GHOSH, H. W. (1949). Expected travel among random points. *Bull. Calcutta Statist. Assoc.* **2** 82.

HAMMERSLEY, J. M. (1974). Postulates for subadditive processes. *Ann. Probability* **2** 652–680.

KARP, R. M. (1976). The probabilistic analysis of some combinatorial search algorithms. In *Proc. Symp. New Directions and Recent Results in Algorithms and Complexity.* (F. Taub, ed.). Academic, San Francisco.

KARP, R. M. (1977). Probabilistic analysis of partitioning algorithms for the traveling-salesman problem in the plane. *Math. Op. Res.* **2** 209–224.

KESTEN, H. (1973). Contribution to J. F. C. Kingman, "Subadditive ergodic theory." *Ann. Probability* **1** 883–909.

KINGMAN, J. F. C. (1976). Subadditive processes, Ecole d'Eté de probabilités de Saint-Flour. *Lecture Notes in Math.* **539**. Springer, New York.

MAHALANOBIS, P. C. (1940). A sample survey of the acreage under jute in Bengal. *Sankhyā* **4** 511–531.

MARKS, E. S. (1948). A lower bound for the expected travel among m random points. *Ann. Math. Statist.* **19** 419–422.

PAPADIMITRIOU, C. H. (1978). The probabilistic analysis of matching heuristics. *Proc. 15th Annual*

Conference Comm. Contr. Computing. Univ. Illinois.

PAPADIMITRIOU, C. H. and STEIGLITZ, K. (1979). *Combinatorial Optimization Algorithms.* To appear.

RAO, R. RANGA. (1962). Relations between weak and uniform convergence of measures with applications. *Ann. Math. Statist.* **33** 659–680.

STEELE, J. M. (1981). Complete convergence of short paths and Karp's algorithm for the TSP. *Math. Op. Res.* (To appear).

STEELE, J. M. (1978). Empirical discrepancies and subadditive processes. *Ann. Probability* **6** 118–127.

VERBLUNSKY, S. (1951). On the shortest path through a number of points. *Proc. Amer. Math. Soc.* **2** 904–913.

WEIDE, B. (1978). Statistical methods in algorithm design and analysis. Thesis, Depart. Computer Science, Carnegie-Mellon Univ., Pittsburgh.

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA 94305