THE SHORTEST PATH THROUGH MANY POINTS

By JILLIAN BEARDWOOD, J. H. HALTON and J. M. HAMMERSLEY

Received 11 March 1958

ABSTRACT. We prove that the length of the shortest closed path through n points in a bounded plane region of area v is 'almost always' asymptotically proportional to $\sqrt{(nv)}$ for large n; and we extend this result to bounded Lebesgue sets in k-dimensional Euclidean space. The constants of proportionality depend only upon the dimensionality of the space, and are independent of the shape of the region. We give numerical bounds for these constants for various values of k; and we estimate the constant in the particular case $k = 2$. The results are relevant to the travelling-salesman problem, Steiner's street network problem, and the Loberman–Weinberger wiring problem. They have possible generalizations in the direction of Plateau's problem and Douglas' problem.

The problem of determining the shortest closed path through a given set of n points is often called the 'travelling-salesman problem', particularly in the literature on operational research and linear programming: see for example, (7), (11), (12), (20). A travelling salesman, starting at and finally returning to his depot in a certain town, has to visit $(n-1)$ other towns by the shortest possible route. If n is at all large, it is prohibitively laborious to calculate the total mileages for each of the $(n-1)!$ orders in which the towns may be visited, and to choose the smallest total. There are, however, systematic numerical methods ((21), (23)) of computing the optimum order with a practical amount of labour; though even with these the labour is not light and will usually demand the services of an electronic computer. On the other hand, given a map of the towns, it is not too hard to draw a path which looks pretty much the shortest; and, for practical purposes, such a path could be used with confidence if we knew that its length was nearly that of the actual shortest path. But no explicit formula for the latter exists; and indeed any exact formula, depending as it must on the detailed structure of the distances between pairs of towns, would presumably be too cumbersome to be of any value. As a workable substitute for an exact formula, we shall derive a simple asymptotic formula for the shortest length when n is large. Despite its simplicity, this formula holds under surprisingly general conditions.

When handling the problem numerically, one may insert as data the actual road mileages between pairs of towns. However, in an analytical approach, where we are deliberately avoiding the detailed distances between pairs of towns, it is convenient to take as the distance between two towns the length of the straight line joining them. This will not affect the result provided that the road mileage is effectively a known constant multiple of the straight-line distance. This proviso is probably met in the large, though in the small it may well be violated. For instance, superficially it might seem that one of the applications of the problem was to the delivery of telegrams in a city. Speed of delivery and operating costs are conflicting factors in a telegraph service. To have to deliver each telegram (not surcharged for special delivery) immediately upon receipt at the delivery office would call for an uneconomically large force of messengers. Accordingly, such telegrams are allowed to accumulate at the

https://doi.org/10.1017/S0305004100034095 Published online by Cambridge University Press

delivery office until a small number n have arrived; whereupon a messenger delivers all n in a single tour of their destinations. Hence one seeks that value of n which provides the best compromise between the delay caused when n is too large and the delivery expenses when n is very small. The rate of arrival of telegrams at the post office and the delivery time for a batch of n telegrams both affect the answer. This delivery time depends upon the length of the shortest route through the n destinations; but it also depends upon other factors, such as the congestion of certain city streets or restrictions in the number of bridges over rivers and railways (13). Local complications of this kind will probably vitiate application of the theory to small regions, especially if n is not very large.

The example just cited raises another complexion of the problem. The points to be visited may be distributed at random over a certain region; and the policy decision on the optimum value of n for a single delivery must rest not on the mileage through n *given* points but rather on the distribution of mileage through n *random* points. This random version of the problem also arises in connexion with agricultural surveys ((15), (18)). Previous work on the shortest path through n random points has concentrated on the expected value of the shortest length. We shall show that a much stronger result holds, namely, that this random shortest length is asymptotically equal to a non-random function of n with probability one. Thus for large n the distinction between the random and the non-random versions of the problem effectively vanishes: indeed it is this fact which ensures the existence of a simple result in the non-random case.

We shall also mention Steiner's street network problem (2), and an interesting related problem of connecting terminals in electronic equipment using the minimum amount of wire (17).

Hereafter for brevity we shall call a shortest closed path a *tour*, and speak of the *tour-length* through a set of points. The points will be distributed in k-dimensional Euclidean space $\mathfrak{R}$; and, as already mentioned, we take the distance between two points to be the ordinary Euclidean distance. Except when the contrary is explicitly stated, we assume $k \geqslant 2$; the reason being that the case $k = 1$ is slightly anomalous and, anyway, rather trivial. We use German capitals for point sets in $\mathfrak{R}$; and, without further mention, we always assume that *these point sets* (*with the sole exception of* $\mathfrak{R}$ *itself*) *are bounded*. In certain cases (explicitly mentioned) we assume that these sets are Lebesgue sets and sometimes one of the conventions (explained later) automatically implies that they are. We use boldface capitals for sequences of points in $\mathfrak{R}$: thus $\mathbf{P} = P_1, P_2, \ldots$. We write $\mathbf{P}^n$ for the first n points of $\mathbf{P}$: thus $\mathbf{P}^n = P_1, P_2, \ldots, P_n$. We write $l(\mathbf{P}^n)$ for the tour-length through $\mathbf{P}^n$. The tour through points $P_1, P_2, \ldots, P_n$ in a set $\mathfrak{E}$ will not necessarily be confined to $\mathfrak{E}$ if $\mathfrak{E}$ is not convex. The closure of a set $\mathfrak{E}$ is written $\bar{\mathfrak{E}}$. If $\mathfrak{E}$ is a Lebesgue set, $v(\mathfrak{E})$ denotes its Lebesgue measure. $\int \ldots dv$ is integration with respect to Lebesgue measure. The expected value of a random variable x is written $\mathscr{E}x$.

To state a preliminary result first, we shall prove that there exists an absolute constant α_k such that

$$\limsup_{n\to\infty} n^{-(k-1)/k}\, l(\mathbf{P}^n) \leqslant \alpha_k k^{\frac{1}{2}} [v(\bar{\mathfrak{E}})]^{1/k}, \tag{1}$$

https://doi.org/10.1017/S0305004100034095 Published online by Cambridge University Press

whenever $\mathbf{P}$ lies within $\mathfrak{E}$. This bound is best-possible in the sense that it can be attained by suitable choice of $\mathbf{P}$ in $\mathfrak{E}$, provided that $\overline{\mathfrak{E}}$ has a boundary of zero measure. We conjecture that this proviso is unnecessary. Some further statements concerning the uniformity of $\mathbf{P}$ and the best-possible character of (1) appear in Lemma 4 below. The inequality (1) was previously proved by Verblunsky (**24**) in the particular case $k = 2$ under the added assumption that $\mathfrak{E}$ was an open set with a boundary of zero measure; but Verblunsky did not investigate the conditions under which (1) was best-possible. Note that, in (1), $v(\overline{\mathfrak{E}})$ exists even when $\mathfrak{E}$ is not a Lebesgue set.

Whilst (1) is best-possible in the sense aforementioned, our main results (2) and (3) below show that much more can be said if we make the additional assumption that $\mathfrak{E}$ is a Lebesgue set of strictly positive measure. In this case there exists an absolute constant β_k (independent of $\mathbf{P}$ and of $\mathfrak{E}$) such that

$$\lim_{n\to\infty} n^{-(k-1)/k} l(\mathbf{P}^n) = \beta_k k^{\frac{1}{2}} [v(\mathfrak{E})]^{1/k} \tag{2}$$

for almost all $\mathbf{P}$ in $\mathfrak{E}$. The existence of this limit is part of the assertion (2). The phrase 'almost all' can be explained in either of two equivalent senses. In set-theoretic language we may represent the sequence $\mathbf{P}$ as a point $P_1 \times P_2 \times \dots$ in the infinite product space $\mathfrak{E} \times \mathfrak{E} \times \dots$, on which we set up the product measure $[v/v(\mathfrak{E})] \times [v/v(\mathfrak{E})] \times \dots$ in the usual way (**9**). Then 'almost all $\mathbf{P}$' means that (2) fails only for a set of $\mathbf{P}$ of zero product measure. Alternatively, in the language of probability theory, we may regard $P_1, P_2, \dots$ as a sample of random points, each independently and uniformly distributed over $\mathfrak{E}$. Then $l(\mathbf{P}^n)$ is a random variable; and the phrase 'almost all $\mathbf{P}$' means that (2) holds with probability one. More generally, if $P_1, P_2, \dots$ are independently distributed over $\mathfrak{E}$ with a common probability density function p, the density being taken with respect to Lebesgue measure, then with probability one

$$\lim_{n\to\infty} n^{-(k-1)/k} l(\mathbf{P}^n) = \beta_k k^{\frac{1}{2}} \int_{\mathfrak{E}} p^{(k-1)/k}\, dv. \tag{3}$$

Equation (3) reduces to (2) when $p = [v(\mathfrak{E})]^{-1}$. Hölder's inequality shows that the right-hand side of (3) does not exceed that of (2). Hence non-uniformity in the distribution of $\mathbf{P}$ tends to decrease $l(\mathbf{P}^n)$. The extreme case of this arises when $P_1, P_2, \dots$ are independently distributed over a Lebesgue set $\mathfrak{E}$ each with same same *arbitrary* probability law. An arbitrary probability law can be expressed as the sum of a discrete, a singular, and a continuous component; and the continuous component can be represented almost everywhere in $\mathfrak{E}$ by a density function, whose integral over $\mathfrak{E}$ will be less than unity when a discrete or a singular component exists. We shall show that (3) remains true provided p is taken to be this density function. Thus singular or discrete components of a probability law contribute nothing to the limit on the left-hand side of (3): they are, as it were, the extreme cases of non-uniformity in distribution.

It is natural to ask if (3) remains valid for probability distributions which are not confined to a bounded set $\mathfrak{E}$, that is to say, can we replace $\mathfrak{E}$ by $\mathfrak{R}$ in (3)? We have given some thought to this question; but have not been able to resolve it. It is quite easy to show that if $\int_{\mathfrak{R}} p^{(k-1)/k}\, dv$ is infinite, then so is the limit on the left-hand side of (3).

https://doi.org/10.1017/S0305004100034095 Published online by Cambridge University Press

When $\int_{\mathcal{R}} p^{(k-1)/k}\,dv$ is finite, it appears that a necessary condition for the validity of the extended version of (3) is that all the absolute $k/(k-1)$th moments of the probability distribution should exist; and we conjecture that this is also a sufficient condition.

Unfortunately, we do not possess explicit formulae for the constants α_k and β_k, though we have established the following bounds:

$$\beta_k \leqslant \alpha_k \leqslant k^{\frac{1}{2}}(k-1)^{-(k-1)/2k}, \tag{4}$$

$$\gamma_k = \frac{(2+k^{-1})!\,[(\frac{1}{2}k)!]^{1/k}}{2\pi^{\frac{1}{2}}(k^{\frac{1}{2}}+k^{-\frac{1}{2}})} \leqslant \beta_k \leqslant 6^{-\frac{1}{2}}12^{1/2k}, \tag{5}$$

where γ_k is defined by (5). As $k \to \infty$, $\gamma_k \to \gamma_\infty = (2\pi e)^{-\frac{1}{2}}$. For the special case $k = 2$, there is a stronger result

$$3^{-1/4} \leqslant \alpha_2 \leqslant (\tfrac{7}{5})^{\frac{1}{2}}; \tag{6}$$

and for $k = 3$, the stronger result

$$(\tfrac{2}{3})^{\frac{1}{2}} \leqslant \alpha_3. \tag{7}$$

The bounds in (6) are due to Verblunsky (**24**). The lower bounds in (6) and (7) arise from taking $P_1, P_2, \ldots, P_n$ as the centres of equal circles or equal spheres close-packed in $\mathfrak{E}$. We should not be surprised if α_2 and α_3 equalled these lower bounds. Numerically,†

$$\left.\begin{aligned}
&\gamma_2 = 0{\cdot}44194 \leqslant \beta_2 \leqslant 0{\cdot}6508, \quad 0{\cdot}75984 \leqslant \alpha_2 \leqslant 1{\cdot}1832;\\
&\gamma_3 = 0{\cdot}37314 \leqslant \beta_3 \leqslant 0{\cdot}61771, \quad 0{\cdot}81650 \leqslant \alpha_3 \leqslant 1{\cdot}3747;\\
&\gamma_4 = 0{\cdot}34208 \leqslant \beta_4 \leqslant 0{\cdot}55696, \qquad \beta_4 \leqslant \alpha_4 \leqslant 1{\cdot}3247;\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&\gamma_\infty = 0{\cdot}24197 \leqslant \beta_\infty \leqslant 0{\cdot}40825, \qquad \beta_\infty \leqslant \alpha_\infty \leqslant 1.
\end{aligned}\right\} \tag{8}$$

We have not proved that α_k and β_k tend to limits when $k \to \infty$; so that the symbols α_∞ and β_∞ in (8) must be interpreted as $\liminf_{k\to\infty} \alpha_k$, $\limsup_{k\to\infty} \alpha_k$, etc., as appropriate. The upper bound for β_2 in (8) is obtained by a special argument and is stronger than in (5). [For comparative purposes $6^{-\frac{1}{2}}12^{\frac{1}{4}} = 0{\cdot}7598$.] All other results in (8) come from (4), (5), (6) and (7).

We have also obtained a very rough Monte Carlo estimate of β_2 by a pair of simple sampling experiments. In the first experiment we used random number tables to distribute 202 points uniformly over the interior of a circle. We then drew with pencil and ruler what appeared (partly by eye and partly by measurement and trial and error) to be the tour of these points. We measured the tour-length and estimated β_2 from (2). The second experiment was similar, except that we distributed 400 points uniformly over the interior of a square. As a precaution against subjective errors in choosing the tour, the two experiments were done independently by different people. Any failure

† Since this paper was submitted we have seen a paper by L. Few, 'The shortest path and the shortest road through n points' *Mathematika*, 2 (1955) 141–4. Few proves there that $\alpha_2 \leqslant 1$, and that $\alpha_k \leqslant k^{\frac{1}{2}}/\{2(k-1)\}^{(k-1)/2k}$ for $k \geqslant 3$. The latter inequality gives $\alpha_3 \leqslant 1{\cdot}0911$, $\alpha_4 \leqslant 1{\cdot}0215, \ldots$, $\alpha_\infty \leqslant 0{\cdot}70709$. Few also gives the result that $\alpha_k \geqslant \{(\frac{1}{2}k)!\}^{1/k}/(\pi k)^{\frac{1}{2}}$; but this gives the following lower bounds $\alpha_2 \geqslant 0{\cdot}39895$, $\alpha_3 \geqslant 0{\cdot}35815$, $\alpha_4 \geqslant 0{\cdot}33546, \ldots, \alpha_\infty \geqslant 0{\cdot}24197$, and all these (except for the limiting case as $k \to \infty$) are weaker then our results. Our result that $0{\cdot}75984 \leqslant \alpha_2$ was given previously by L. Fejes Toth, *Math. Zeit.* 46 (1940) 83–5.

https://doi.org/10.1017/S0305004100034095 Published online by Cambridge University Press

to pick the actual tour will lead to an overestimate of β_2. Experience of carrying through this procedure leads us to believe that it is not nearly so untrustworthy and imprecise as it may seem at first sight. The first and second experiments respectively gave estimates $\beta_2 \doteqdot 0{\cdot}526$ and $\beta_2 \doteqdot 0{\cdot}536$. This agreement is gratifying, though possibly coincidental. In the absence of further evidence, we think it probable† that

$$\beta_2 \doteqdot 0{\cdot}53. \tag{9}$$

To see how (2), (3), and (9) fared in practice we compared them with some exact results for the tour of Washington, D.C., and forty-eight other towns, one in each State of the U.S.A. (4). The tour-length given in (4) is 12,345 miles. From (2) and (9) we estimate

$$l = 0{\cdot}53(2nV)^{\frac{1}{2}}, \tag{10}$$

where $n = 49$ and V is the area of the U.S.A. [Although we only prove (2) for Euclidean space, it could equally well be proved for the surface of a spheroid.] Since (10) gives 9100 miles, the agreement seems indifferent. However, in (4), the distances between towns are road distances. Accordingly, we measured the corresponding straight-line distances, using dividers and conical orthomorphic projections of the U.S.A. ((22), Maps 88/89, 90/91, 92/93) on scales of 78·9 and 126·3 miles to the inch. On the likely hypothesis that the straight-segment tour will visit the forty-nine towns in the same order as the road-mileage tour‡, we found a straight-segment tour-length of 10,070 miles (subject to rather less than 1% error for errors of measurement with the dividers). What is responsible for the disagreement between 9100 and 10,070 miles? Presumably the fact that the different States are of different sizes is not to blame; for any allowance for non-uniformity of distribution will decrease the estimated value of l. For instance, if we use a probability density function which is constant over each State and allocates a total probability $\frac{1}{49}$ to each State (regarding the District of Columbia as a State), we find from (3)

$$l = 0{\cdot}53 \sum_{i=1}^{49} (2V_i)^{\frac{1}{2}}, \tag{11}$$

where V_i is the area of the ith State. Equation (11) gives $l = 8400$ miles. A more likely cause of the discrepancy is that, although (3) allows for non-uniformity of distribution, it pays no attention to where the areas of high population density are within $\mathfrak{E}$. If n is sufficiently large this is justified; but the consideration may become important when we impose the condition that there shall only be one point in each area of high density. In the case of the U.S.A. the areas of high population density and the small States are near the seaboard or the Great Lakes; and the forty-nine towns, chosen by the authors of (4) because they were sufficiently large centres of traffic to merit appearance in a road-map table of mileages between towns, lie noticeably near the outer boundary of the U.S.A. No fewer than eleven of the forty-nine towns are actually on the coast or lakeside, or

† Of course, the Monte Carlo experiment is very slight and gives very little information. With only two experiments, the 95% confidence limits for β_2 are 0·467 and 0·595. We hope someone will be encouraged to make a more precise estimate.

‡ There is another way of looking at the matter not involving this hypothesis: namely, to the estimate 9100 miles we add 23% to allow for the empirically determined excess of road over straight-line distance, and then compare the resulting figure of 11,200 miles with the correct 12,345 miles.

https://doi.org/10.1017/S0305004100034095 Published online by Cambridge University Press

wide estuaries thereof; and a further 5 lie within 10 miles of the coast or Lakes (cf. the approximate diameter of the U.S.A., about 3000 miles). This congregation on the outer boundaries will tend to inflate the tour-length above the value estimated by (2) or (3). The example serves as a useful warning of the limitations of our asymptotic formulae.

We have proved one result which holds for finite values of n: namely, when $\mathfrak{E}$ is a Lebesgue set of positive measure and $P_1, P_2, \ldots$ are independently and uniformly distributed over $\mathfrak{E}$, then for all $n \geqslant 3$

$$\mathscr{E}l(\mathbf{P}^n) \geqslant \frac{(2+k^{-1})!\,[(\frac{1}{2}k)!]^{1/k}\,k^{\frac{1}{2}}n![v(\mathfrak{E})]^{1/k}}{2\pi^{\frac{1}{2}}(k^{\frac{1}{2}}+k^{-\frac{1}{2}})\,[(n-1+k^{-1})!]} \geqslant \gamma_k\, k^{\frac{1}{2}}n^{(k-1)/k}[v(\mathfrak{E})]^{1/k}. \tag{12}$$

This result is stronger than a previous inequality due to Marks (19) in the case $k = 2$. (There is a misprint in Marks' paper, which erroneously suggests that $\mathscr{E}l(\mathbf{P}^n) \geqslant [\frac{1}{2}nv(\mathfrak{E})]^{\frac{1}{2}}$; and this misprint is copied by Ghosh (8).) Ghosh also proves for $k = 2$

$$\limsup_{n\to\infty} \mathscr{E}l(\mathbf{P}^n) \leqslant 0{\cdot}8950[2nv(\mathfrak{E})]^{\frac{1}{2}}, \tag{13}$$

which is weaker than our result (8). Mahalanobis (18) states without proof that $\mathscr{E}l(\mathbf{P}^n) = n^{\frac{1}{2}} - n^{-\frac{1}{2}}$ when $v(\mathfrak{E}) = 1$ and $k = 2$; but his assertion is false in view of (8).

The analysis in this paper specifically concentrates on the tour-length $l(\mathbf{P}^n)$. But our arguments apply, with occasional minor modifications, to $l^*(\mathbf{P}^n)$, the shortest length of wire needed to connect together terminals situated at $P_1, P_2, \ldots, P_n$. Consequently, the same results apply to l^* as to l, although the numerical values of the relevant constants α_k^*, β_k^* may differ from those of α_k, β_k. There are two versions of the terminal problem; and an example will serve to distinguish them. Suppose we have four terminals P_1, P_2, P_3, P_4 at the vertices $(\pm\frac{1}{2}, \pm\frac{1}{2})$ of a unit square in the Euclidean plane. It is quite simple to show that $l^*(\mathbf{P}^4) = 1+\sqrt{3}$. For this an appropriate set of straight wires connects the points $(+\frac{1}{2}, \pm\frac{1}{2})$ to the point $(\frac{1}{2}-\frac{1}{6}\sqrt{3}, 0)$, the point $(\frac{1}{2}-\frac{1}{6}\sqrt{3}, 0)$ to the point $(-\frac{1}{2}+\frac{1}{6}\sqrt{3}, 0)$, and the point $(-\frac{1}{2}+\frac{1}{6}\sqrt{3}, 0)$ to the points $(-\frac{1}{2}, \pm\frac{1}{2})$. This version of the problem is known as Steiner's street network problem ((2), Chapter VII). In the other version of the problem, considered in (1), (16), (17) one is not allowed to join wires together at points which are not terminals; and the resulting shortest length of wire will then be $l^{**}(\mathbf{P}^4) = 3$. In general, in k dimensions one has

$$\tfrac{1}{2}l(\mathbf{P}^n) \leqslant l^*(\mathbf{P}^n) \leqslant l^{**}(\mathbf{P}^n) \leqslant l(\mathbf{P}^n). \tag{14}$$

To see this, one may regard each segment of wire in the $l^*(\mathbf{P}^n)$ arrangement as a closed path (going once in each direction along it) and remark that the sum of a number of closed paths with points in common is a closed path, so that $2l^*(\mathbf{P}^n) \geqslant l(\mathbf{P}^n)$. The argument used for $l(\mathbf{P}^n)$ in Lemma 3 breaks down for $l^*(\mathbf{P}^n)$; but one may use (12) and (14) instead. Results similar to (1), (2) and (3) hold for $l^{**}(\mathbf{P}^n)$ with appropriate α_k^{**}, β_k^{**}. Some slight simplifications in our argument would result if we used the easily verified fact that $l(\mathbf{P}^n)$, and for that matter $l^*(\mathbf{P}^n)$, are non-decreasing functions of n for every given $\mathbf{P}$. But we eschew these simplifications because $l^{**}(\mathbf{P}^n)$ may not have this property: for instance, if we take $P_5 = (\frac{1}{2}-\frac{1}{6}\sqrt{3}, 0)$, $P_6 = (-\frac{1}{2}+\frac{1}{6}\sqrt{3}, 0)$ in the above example, we find $1+\sqrt{3} = l^{**}(\mathbf{P}^6) < l^{**}(\mathbf{P}^4) = 3$.

The material of this paper raises the following problem, in which we write v_k for k-dimensional Lebesgue measure: *Given a region $\mathfrak{E}$ in $\mathfrak{R}$ with known measure $v = v_k(\mathfrak{E})$,*

https://doi.org/10.1017/S0305004100034095 Published online by Cambridge University Press

choose in $\mathfrak{E}$ *a random* k'*-dimensional subset* $\mathfrak{E}'$ *with prescribed finite measure* $v' = v_{k'}(\mathfrak{E}')$, *What is the distribution of the measure* $v'' = v_{k'+1}(\mathfrak{E}'')$ *of* $(k'+1)$*-dimensional sets* $\mathfrak{E}''$ (*not necessarily subsets of* $\mathfrak{E}$), *such that* v'' *is minimal subject to* $\mathfrak{E}''$ *being bounded*† *by* $\mathfrak{E}'$? The present paper considers the special case $k' = 0$, with the usual convention that the zero-dimensional Lebesgue measure of a point set is the number of points it contains ($v' = n$, $v'' = l$). The case $k' = 1$, $k = 3$ would be a randomized version of Plateau's problem ((2), Chapter VII); while, for other values of k', k, we get randomized versions of Douglas' problem ((3), (5)). The main difficulty presented by this generalized problem lies, we believe, in the precise definition of the random nature of $\mathfrak{E}'$ and in the restrictions to be imposed on the connectivity, orientability, and other topological characteristics of the spanning set $\mathfrak{E}''$. We do not consider the generalized problem in this paper; but we feel intuitively that the basic appeal to subadditivity, which is the key of our present analysis, would still remain appropriate in the general case. That is to say, roughly, that if we take v' proportional to v and regard v'' as a function of $\mathfrak{E}$, then v'' is a subadditive set-function of $\mathfrak{E}$. The proof of Lemma 5 below should afford a clearer idea of what we mean by this very condensed, and consequently rather cryptic, remark and of how it leads to a solution of the problem.

We shall now specify the detailed notation to be used in the proof of our results. After that, we shall give a short outline of the method of proof, together with a few conjectures on results we have not succeeded in proving. The lemmas assembled in the Appendix contain the relevant details of rigorous proof and exact statement; and hence in the short outline we are free to ignore these precise details and to concentrate instead on the broad spirit of the underlying ideas.

We have already stated that German capitals denote bounded point sets in $\mathfrak{R}$, that $\overline{\mathfrak{E}}$ denotes the closure of $\mathfrak{E}$, and that boldface capitals denote sequences of points. These symbols may carry suffices to denote different sets or sequences: e.g. $\mathbf{P}_j = P_{1j}, P_{2j}, \ldots$ and $\mathbf{P}_j^n = P_{1j}, P_{2j}, \ldots, P_{nj}$. We write $\mathfrak{E}_1 + \mathfrak{E}_2$ for the union of $\mathfrak{E}_1$ and $\mathfrak{E}_2$; $\Sigma_j \mathfrak{E}_j$ for the union of several sets $\mathfrak{E}_j$; $\mathfrak{E}_1\mathfrak{E}_2$ for the intersection of $\mathfrak{E}_1$ and $\mathfrak{E}_2$; $\mathfrak{E}_1 - \mathfrak{E}_2$ for the (possibly empty) largest subset of $\mathfrak{E}_1$ not in $\mathfrak{E}_2$. Sequences of points can be manipulated in the same way: e.g. $\mathbf{P}^n\mathfrak{E}$ denotes the subset of $P_1, P_2, \ldots, P_n$ which lies in $\mathfrak{E}$; whereas $(\mathbf{P}\mathfrak{E})^n$ denotes the first n points of that subset of $\mathbf{P}$ which lies in $\mathfrak{E}$. $N(\mathbf{P}\mathfrak{E})$ denotes the (possibly infinite) number of points in the sequence $\mathbf{P}\mathfrak{E}$; and $T(\mathbf{P}\mathfrak{E})$ and $l(\mathbf{P}\mathfrak{E})$ denote the tour and the tour-length of $\mathbf{P}\mathfrak{E}$. If $\mathbf{P}\mathfrak{E}$ is empty, $T(\mathbf{P}\mathfrak{E})$ is taken to be the empty set of $\mathfrak{R}$ and $l(\mathbf{P}\mathfrak{E}) = 0$. To illustrate a fairly complicated use of this symbolism, which occurs in Lemma 2,

$$\sup_{\mathbf{Q}_j} l(\mathbf{Q}_j^{2N(\mathbf{P}^n\mathfrak{E}_j)}\overline{\mathfrak{E}}_j\,\overline{\mathfrak{R} - \mathfrak{E}_j})$$

denotes the supremum of all tour-lengths through not more than ν points of the boundary of the set $\mathfrak{E}_j$, where ν is twice the number of points which are both in $\mathfrak{E}_j$ and amongst the first n points of the sequence $\mathbf{P}$. To avoid tiresome repetition of the phrase 'if $\mathfrak{E}$ is a Lebesgue set', we adopt the convention that the mere writing of $v(\mathfrak{E})$ to denote the k-dimensional Lebesgue measure of $\mathfrak{E}$ carries with it the automatic postulate that $\mathfrak{E}$ is in fact a Lebesgue set. Notice, however, that the writing of $v(\overline{\mathfrak{E}})$ merely confirms

† We have only a very vague idea of what we mean here by 'bounded', and part of the problem is to give a suitable definition to this word.

https://doi.org/10.1017/S0305004100034095 Published online by Cambridge University Press

that a closed set is a Lebesgue set and does not necessarily imply that $\mathfrak{E}$ itself is a Lebesgue set. We write $\delta(\mathfrak{E})$ for the diameter of $\mathfrak{E}$: this is necessarily finite because $\mathfrak{E}$ is bounded. $\mathfrak{O}$ will denote the empty set of $\mathfrak{R}$. $\mathfrak{C}$ will always denote a semi-closed unit hypercube of k dimensions, i.e. a hypercube of unit length side, open on its lower-left faces and closed on its upper-right faces. $\mathfrak{C}_1, \mathfrak{C}_2, \ldots$ will denote semiclosed unit hypercubes in different positions in $\mathfrak{R}$. If ξ is a positive real number, we write $\xi\mathfrak{E}$ for the set of all points with coordinates $(\xi x_1, \xi x_2, \ldots, \xi x_k)$ such that $(x_1, x_2, \ldots, x_k)$ is a point of $\mathfrak{E}$. Thus $\xi\mathfrak{E}$ is a ξ-fold linear magnification of $\mathfrak{E}$ without rotation, which leaves the origin of $\mathfrak{R}$ invariant; and $\delta(\xi\mathfrak{E}) = \xi\delta(\mathfrak{E})$ and $v(\xi\mathfrak{E}) = \xi^k v(\mathfrak{E})$. The symbol ξ conventionally operates on all subsequent sets in a product: thus $\xi\mathbf{P}\mathfrak{E}$ is the magnification of $\mathbf{P}\mathfrak{E}$ whereas $\mathbf{P}\xi\mathfrak{E}$ is the intersection of the unmagnified $\mathbf{P}$ with the magnified $\mathfrak{E}$. However, brackets will override this convention: for instance, in Lemma 5, we have $(\eta\mathfrak{C})(\xi\mathfrak{E})$ to denote the intersection of the η-fold magnification of $\mathfrak{C}$ with the ξ-fold magnification of $\mathfrak{E}$; and this contrasts with $\eta\mathfrak{C}\xi\mathfrak{E}$, which would be the η-fold magnification of the intersection of $\mathfrak{C}$ with ξ-fold magnification of $\mathfrak{E}$. The phrase '$\mathbf{P} \in p(\mathfrak{E})$' carries the following meaning: $\mathfrak{E}$ is a Lebesgue set of strictly positive measure and $\mathbf{P} = P_1, P_2, \ldots$ is a sample of random points independently distributed over $\mathfrak{E}$ with a common Lebesgue-measurable probability density function p, where $\int_{\mathfrak{E}} p\,dv = 1$. In the special case when this distribution is uniform, i.e. when $p = 1/v(\mathfrak{E})$, we write $\mathbf{P} \in u(\mathfrak{E})$ in place of $\mathbf{P} \in p(\mathfrak{E})$. The phrase '$\mathbf{P} \in W_\xi$' means that $\mathbf{P} = P_1, P_2, \ldots$ is a sample from a Poisson process of density ξ over $\mathfrak{R}$; that is to say, for arbitrary disjoint Lebesgue sets $\mathfrak{E}_1, \mathfrak{E}_2, \ldots, \mathfrak{E}_m$,

$$\text{prob}\{N(\mathbf{P}\mathfrak{E}_j) = N_j; \quad j = 1, 2, \ldots, m\} = \prod_{j=1}^{m} \frac{\{\xi v(\mathfrak{E}_j)\}^{N_j}}{N_j!} \exp\{-\xi v(\mathfrak{E}_j)\}. \tag{15}$$

In most of the proof k is fixed and it is convenient to suppress the dependence of constants and functions upon k. In particular we adopt the abbreviations and notation

$$\left.\begin{aligned} &\alpha = k^{\frac{1}{2}}\alpha_k; \quad \beta = k^{\frac{1}{2}}\beta_k; \quad q = (k-1)/k; \\ &v^{1-q}(\mathfrak{E}) = [v(\mathfrak{E})]^{1/k}; \quad \log^z x = (\log x)^z. \end{aligned}\right\} \tag{16}$$

With these notational conventions in mind, the reader will be able to get precise statements of our full results by reading the enunciations of the various lemmas.

Suppose that we break a given set $\mathfrak{E}$ into a number of disjoint parts $\mathfrak{E}_j (j = 1, 2, \ldots, m)$. Then Lemmas 1 and 2 taken together state that, for any given $\mathbf{P}^n = P_1, P_2, \ldots, P_n$, the tour-length through all points of $\mathbf{P}^n$ in $\mathfrak{E}$ (namely $l(\mathbf{P}^n\mathfrak{E})$) is roughly equal to the sum of the corresponding tour-lengths in the parts of $\mathfrak{E}$ (namely $\sum_{j=1}^{m} l(\mathbf{P}^n\mathfrak{E}_j)$), provided n is large. The explanation of this rough statement is that, whereas each of these two quantities is of magnitude $n^{(k-1)/k}$ in view of the ultimate results (2) and (3), their difference is bounded on one side by a fixed quantity, depending on the diameter of $\mathfrak{E}$, and on the other side by the sum of tour-lengths through points on the boundaries of the $\mathfrak{E}_j$. If, as we shall eventually arrange, these boundaries are sufficiently well-behaved to be sets of dimensionality $k-1$, the tour-lengths through such boundary points will be of magnitude $n^{(k-2)/(k-1)}$, which is small in comparison with $n^{(k-1)/k}$. We therefore seek

https://doi.org/10.1017/S0305004100034095 Published online by Cambridge University Press

the convolution of the distributions of the several $l(\mathbf{P}^n\mathfrak{E}_j)$. However, it is not easy to handle this convolution when the respective $l(\mathbf{P}^n\mathfrak{E}_j)$ are not independent of one another; and clearly they are dependent when n is fixed, inasmuch as an excess of points of $\mathbf{P}^n$ in $\mathfrak{E}_1$, say, will tend to reduce the number in $\mathfrak{E}_2$, so that one may expect $l(\mathbf{P}^n\mathfrak{E}_1)$ and $l(\mathbf{P}^n\mathfrak{E}_2)$ to be negatively correlated. To overcome this difficulty we make n a random variable. Specifically, instead of sampling $\mathbf{P}$ from $p(\mathfrak{E})$ or $u(\mathfrak{E})$, we sample $\mathbf{P}$ from W_ξ for large ξ. Then, by (15), the random point sets $\mathbf{P}\mathfrak{E}_j$, and therefore the random tour-lengths $l(\mathbf{P}^n\mathfrak{E}_j)$, are statistically independent for $j = 1, 2, \ldots, m$. This Poisson process has the useful property that $n = N(\mathbf{P}\mathfrak{E})$ is a Poisson variable with parameter $\xi v(\mathfrak{E})$, and the conditional distribution of $\mathbf{P}\mathfrak{E}$ given n is identical with the unconditional distribution of $\mathbf{P}^n\mathfrak{E}$ under the assumption that $\mathbf{P} \in u(\mathfrak{E})$. When ξ is large, n will be nearly equal to $\xi\, v(\mathfrak{E})$ with high probability; so that the distribution of $l(\mathbf{P}\mathfrak{E})$ for $\mathbf{P} \in W_\xi$ should be identifiable with the distribution of $l(\mathbf{P}^x\mathfrak{E})$ for $\mathbf{P} \in u(\mathfrak{E})$, where x is the integer part of $\xi v(\mathfrak{E})$. This identification can be made rigorous by a Tauberian argument, in which the Tauberian condition is that $l(\mathbf{P}^n\mathfrak{E})$, regarded as a function of n, cannot increase or decrease too rapidly. The foregoing remarks exhibit the essence of our line of attack; but it is convenient to make a further modification to deal with the fact that the limit in (2) is independent of the shape of $\mathfrak{E}$. To this end we transform the problem so that $l(\mathbf{P}\mathfrak{E})$ behaves rather like Lebesgue measure: in brief, we have to remove the power $1/k$ in (2) before we can hope to get additivity in $\mathfrak{E}$. Now the distribution of $\mathbf{P}\mathfrak{E}$ when $\mathbf{P} \in W_\xi$ is, apart from a simple change of scale, the same as the distribution of $\mathbf{P}\xi\mathfrak{E}$ when $\mathbf{P} \in W_1$. Moreover, when $\mathbf{P} \in W_1$, $n = N(\mathbf{P}\xi\mathfrak{E})$ will be roughly equal to $v(\xi\mathfrak{E})$; and hence the expression $n^{(k-1)/k}\, v^{1/k}(\xi\mathfrak{E})$ will behave approximately like $\xi^k v(\mathfrak{E})$, which is additive in $\mathfrak{E}$ as required. This motivates the consideration of $l(\mathbf{P}\xi\mathfrak{E})$ for $\mathbf{P} \in W_1$ in Lemmas 5 and 6.

As already stated, $l(\mathbf{P}\xi\mathfrak{E})$ is nearly equal to the sum of the independent random quantities $l(\mathbf{P}\xi\mathfrak{E}_j)$, where the $\mathfrak{E}_j$ are disjoint parts composing $\mathfrak{E}$. One might therefore expect $l(\mathbf{P}\xi\mathfrak{E})$ to be asymptotically normal in view of the central limit theorem. But we have not succeeded in proving this, although we do show that, for real positive t, the cumulant generating function of $l(\mathbf{P}\xi\mathfrak{E})$ satisfies

$$K_t(\mathbf{P}\xi\mathfrak{E}) = \log \mathscr{E} \exp t l(\mathbf{P}\xi\mathfrak{E}) \sim w(t)\, \xi^k v(\mathfrak{E}) \quad \text{as} \quad \xi \to \infty, \tag{17}$$

where $w(t)$ is a function of k and t only. This in itself does not imply normality, nor even does it provide information about the magnitude of the cumulants of $l(\mathbf{P}\xi\mathfrak{E})$. The following simple counter-example is instructive. Let x be a random variable which takes the values 0 and ξ with equal probability $\frac{1}{2}$. The cumulant generating function of x is

$$K_t(x) = \log(\tfrac{1}{2} + \tfrac{1}{2}e^{t\xi}) \sim t\xi \quad \text{as} \quad \xi \to \infty \quad (t > 0); \tag{18}$$

but the asymptotic behaviour of the cumulants of x cannot be deduced from the coefficients in $t\xi$ regarded as a power series in t: the mean of x is $\frac{1}{2}\xi$, not $\xi + o(\xi)$ and the variance of x is $\frac{1}{4}\xi^2$, not $o(\xi)$. There is every reason to believe that, when $\mathbf{P} \in W_1$,

$$\mathscr{E}(\mathbf{P}\xi\mathfrak{E}) = \beta\xi^k v(\mathfrak{E}) + O(\xi^{k-\{1/(k-1)\}}) \quad \text{as} \quad \xi \to \infty, \tag{19}$$

in which case we could deduce results about the magnitude of the error term implicit in (2) and (3). Indeed, by following the type of argument used in Lemma 6, we can

https://doi.org/10.1017/S0305004100034095 Published online by Cambridge University Press

prove (19) under rather severe restrictions about the nature of $\mathfrak{E}$: for instance, (19) holds in the particular case $\mathfrak{E} = \mathfrak{C}$. Although we cannot even prove (19) for general $\mathfrak{E}$, we believe that the error term in (19) is capable of still further reduction, probably as far as $O(\xi^{k-1})$. Similarly, we believe that the order of magnitude in Lemma 6 could be reduced.

As one may easily see, when $k = 1$ the corresponding results depend upon the shape of $\mathfrak{E}$, in particular upon how pathologically disconnected $\mathfrak{E}$ is. The analysis for $k \geqslant 2$ fails for $k = 1$ because $\delta(\xi\mathfrak{E})$ is then not small compared with $v(\xi\mathfrak{E})$ when ξ is large. There may, however, be some justification for comparing the results for $k = 1$ with those for $k > 1$ in the special case $\mathfrak{E} = \mathfrak{C}$. When $k = 1$ and $\mathbf{P} \in W_1$, we find $l(\mathbf{P}\xi\mathfrak{C}) \sim \xi v(\mathfrak{C})$ with probability one as $\xi \to \infty$, and $\operatorname{var} l(\mathbf{P}\xi\mathfrak{C}) = O(1)$, but $l(\mathbf{P}\xi\mathfrak{C})$ is *not* asymptotically normal.

APPENDIX

The appendix is wholly concerned with formalities of proof.

LEMMA 1. *If* $\mathfrak{E}_{m+1} = \mathfrak{E}_1$, *then*

$$l\left(\mathbf{P}^n \sum_{j=1}^{m} \mathfrak{E}_j\right) \leqslant \sum_{j=1}^{m} \{l(\mathbf{P}^n\mathfrak{E}_j) + \delta(\mathfrak{E}_j + \mathfrak{E}_{j+1})\} \leqslant \sum_{j=1}^{m} \left\{l(\mathbf{P}^n\mathfrak{E}_j) + \delta\left(\sum_{j=1}^{m} \mathfrak{E}_j\right)\right\}. \tag{1·1}$$

We may suppose $\mathbf{P}^n\mathfrak{E}_j \neq \mathfrak{O}$ and take $Q_j \in \mathbf{P}^n\mathfrak{E}_j$. The central term of (1·1) is at least as large as the length of a closed path of straight segments joining $Q_1, Q_2, \ldots, Q_m, Q_1$ together with digressions at each Q_j around $T(\mathbf{P}^n\mathfrak{E}_j)$. But this path passes through $\mathbf{P}^n \sum \mathfrak{E}_j$ and so cannot be shorter than the first term of (1·1).

LEMMA 2. *If* $\mathfrak{E}_1, \mathfrak{E}_2, \ldots, \mathfrak{E}_m$ *are disjoint, then*

$$l\left(\mathbf{P}^n \sum_{j=1}^{m} \mathfrak{E}_j\right) \geqslant \sum_{j=1}^{m} \{l(\mathbf{P}^n\mathfrak{E}_j) - 2 \sup_{\mathbf{Q}_j} l(\mathbf{Q}_j^{2N(\mathbf{P}^n\mathfrak{E}_j)} \overline{\mathfrak{E}_j} \overline{\mathfrak{R} - \mathfrak{E}_j})\}. \tag{2·1}$$

We may suppose $\mathbf{P}^n\mathfrak{E}_j \neq \mathfrak{O}$ and $m > 1$. $\mathfrak{E}_j T\left(\mathbf{P}^n \sum_{j=1}^{m} \mathfrak{E}_j\right)$ consists of one or more connected pieces, some of which may degenerate to points. Discard any such piece which does not contain some point of $\mathbf{P}^n$. The closure of any surviving piece will be called a j-piece. The number ν of j-pieces satisfies $\nu \leqslant N(\mathbf{P}^n\mathfrak{E}_j)$. Let

$$\mathbf{Q}_j^{2\nu} = Q_{1j}, Q_{2j}, \ldots, Q_{2\nu,j}$$

be the set of end-points of the j-pieces, in which there are two coincident Q's for any surviving degenerate j-piece. Let l'_j denote the total length of all j-pieces. Since $\mathbf{Q}_j^{2\nu} \in \overline{\mathfrak{E}_j}\,\overline{\mathfrak{R} - \mathfrak{E}_j}$, and since $l\left(\mathbf{P}^n \sum_{j=1}^{m} \mathfrak{E}_j\right) \geqslant \sum_{j=1}^{m} l'_j$, (2·1) will follow as soon as we have proved

$$l(\mathbf{P}^n\mathfrak{E}_j) \leqslant l'_j + 2l(\mathbf{Q}_j^{2\nu}). \tag{2·2}$$

We shall establish (2·2) by constructing a closed path consisting only of the j-pieces, each traversed just once, together with a part of a double circuit of $T = T(\mathbf{Q}_j^{2\nu})$. Choose a particular $\mathfrak{E}_j$ and suppress the suffix j. We may suppose the Q's are labelled so that $Q_1, Q_2, \ldots, Q_{2\nu}, Q_1$ occur in that order when passing around T. We interpret suffices modulo 2ν; and write $Q_{i'}$ for the point at the other end of the j-piece (denoted by J_i)

https://doi.org/10.1017/S0305004100034095 Published online by Cambridge University Press

terminated by Q_i. We construct the closed path in stages. At any particular stage certain of the Q_i will have been visited by the path, the remainder unvisited. Suppose that Q_a, Q_b, Q_c have been visited, whilst $Q_{a+1}, Q_{a+2}, \ldots, Q_{b-1}, Q_{b+1}, Q_{b+2}, \ldots, Q_{c-1}$ have not yet been visited. We say that Q_{b+1} is an odd neighbour of Q_b if $c-1-b$ is odd. Similarly Q_{b-1} is an odd neighbour of Q_b if $b-1-a$ is odd. We start the path at Q_1 and proceed as follows. From Q_1 pass to Q_2 along T; thence to $Q_{2'}$ along J_2; thence to the (unique) odd neighbour of $Q_{2'}$ (say Q_x) along T; thence to $Q_{x'}$ along J_x; thence to the (unique) odd neighbour of $Q_{x'}$ along T; etc, ... Ultimately this procedure of alternate steps along j-pieces and along T to the unique odd neighbour will bring the path back to Q_1. The reason why the required odd neighbours are unique is because 2ν is even. If, upon return to Q_1, the path has traversed all j-pieces, it is the required path. If not, start a fresh similar closed path from any Q-point not yet visited, the first step of the fresh path being along T to the odd neighbour of the starting point of the fresh path; and continue to construct such closed paths, starting from unvisited points, until all j-pieces are traversed. Thereafter we get the required closed path by travelling around T with digressions around each of the already-constructed closed paths whenever such a path is encountered for the first time on T.

In following the above argument, it may be helpful to consider some particular case. For instance, when $\nu = 6$ and $1' = 6$, $2' = 5$, $3' = 10$, $4' = 7$, $8' = 12$ and $9' = 11$, the ultimately constructed path is

$$Q_1Q_2Q_5Q_6Q_1Q_2Q_3Q_4Q_7Q_8Q_{12}Q_{11}Q_9Q_{10}Q_3Q_4Q_5Q_6Q_7Q_8Q_9Q_{10}Q_{11}Q_{12}Q_1.$$

LEMMA 3. *If* $\mathbf{P} \in u(\mathfrak{E})$ *and* $n \geqslant 3$, *then* (12) *holds.*

Let x_{ij} denote the distance of P_j from the ith nearest of $P_1, P_2, \ldots, P_{j-1}, P_{j+1}, \ldots, P_n$. Since $T(\mathbf{P}^n)$ does not visit any point twice, the sum of the lengths of the two segments of $T(\mathbf{P}^n)$ which terminate at P_j is at least $x_{1j}+x_{2j}$. Hence

$$\mathscr{E}2l(\mathbf{P}^n) \geqslant \sum_{j=1}^{n} \mathscr{E}_j \mathscr{E}'_j(x_{1j}+x_{2j}), \tag{3·1}$$

where $\mathscr{E}_j$ is the expectation over P_j, and $\mathscr{E}'_j$ is the conditional expectation over $P_1, P_2, \ldots, P_{j-1}, P_{j+1}, \ldots, P_n$ given P_j. Let $\mathfrak{S}_x$ denote a k-dimensional hypersphere of radius x centred at P_j and write

$$V = v(\mathfrak{E}); \quad V_x = v(\mathfrak{S}_x \mathfrak{E}) \leqslant v(\mathfrak{S}_x) = \pi^{\frac{1}{2}k} x^k / (\tfrac{1}{2}k)!. \tag{3·2}$$

The conditional probability, given P_j, that x_{ij} exceeds x is the probability that not more than $i-1$ of the $n-1$ independently and uniformly random $P_1, P_2, \ldots, P_{j-1}, P_{j+1}, \ldots, P_n$ lie in $\mathfrak{S}_x \mathfrak{E}$. Thus

$$\text{prob}\,(x_{ij} > x | P_j) = \sum_{r=0}^{i-1} \binom{n-1}{r} \left(\frac{V_x}{V}\right)^r \left(1 - \frac{V_x}{V}\right)^{n-1-r}. \tag{3·3}$$

Hence

$$\begin{aligned} \mathscr{E}'_j(x_{1j}+x_{2j}) &= \int_0^\infty \text{prob}\,(x_{1j} > x | P_j) + \text{prob}(x_{2j} > x | P_j)\, dx \\ &= \int_0^\infty (1 - V_x/V)^{n-2} \{2 + (n-3)\, V_x/V\}\, dx. \end{aligned} \tag{3·4}$$

https://doi.org/10.1017/S0305004100034095 Published online by Cambridge University Press

It is an elementary matter to verify that $(1-z)^{n-2}\{2+(n-3)z\}$ is a non-increasing non-negative function of z for $0 \leqslant z \leqslant 1$. Hence we shall not increase the right-hand side of (3·4) if we replace V_x by $\min\{v(\mathfrak{E}), v(\mathfrak{S}_x)\}$. This gives, from (3·2),

$$\mathfrak{E}'_j(x_{1j}+x_{2j}) \geqslant \int_0^\rho (1-x^k/\rho^k)^{n-2}\{2+(n-3)\,x^k/\rho^k\}\,dx, \tag{3·5}$$

where

$$\rho^k = V(\tfrac{1}{2}k)!/\pi^{\frac{1}{2}k}. \tag{3·6}$$

Since the right-hand side of (3·5) is independent of P_j, (3·1) provides

$$\mathscr{E}l(\mathbf{P}^n) \geqslant \tfrac{1}{2}n\int_0^\rho (1-z)^{n-2}\{2+(n-3)\,z\}\,dx, \tag{3·7}$$

where $z = x^k/\rho^k$. The right-hand side of (3·7) is the sum of two complete beta functions. Expressing these in terms of gamma functions and simplifying, we get the central term of (12). To prove the rest of (12), we write $a_n = n!\,n^{-q}/(n-q)!$ [see (16) for the definition of q] and combine the well-known fact that $a_n \to 1$ as $n \to \infty$ with the easily verified fact that

$$\frac{a_n}{a_{n+1}} = \left(1+\frac{1}{kn}\right)\left(1+\frac{1}{n}\right)^{-1/k} \geqslant 1. \tag{3·8}$$

LEMMA 4. *There exist an absolute constant* $\alpha_k \leqslant k^{\frac{1}{2}}(k-1)^{-(k-1)/2k}$ *and a set function* $a(\mathfrak{E}, \epsilon)$ *corresponding to every given* $\epsilon > 0$, *such that*

$$l(\mathbf{P}^n\mathfrak{E}) \leqslant \alpha n^q[v(\overline{\mathfrak{E}})+\epsilon]^{1-q} + a(\mathfrak{E}, \epsilon) \quad (\alpha = \alpha_k k^{\frac{1}{2}}) \tag{4·1}$$

for all $\mathbf{P}$. *Consequently*

$$L(\mathbf{P}\mathfrak{E}) = \limsup_{n\to\infty} n^{-q} l(\mathbf{P}^n\mathfrak{E}) \leqslant \alpha v^{1-q}(\overline{\mathfrak{E}}). \tag{4·2}$$

A sequence $\mathbf{P}$ *is called maximal if equality holds in* (4·2). *A necessary and sufficient condition for* $\mathfrak{E}$ *to contain a maximal sequence is that* $\overline{\mathfrak{E}}$ *should contain one. A sufficient condition for* $\overline{\mathfrak{E}}$ *to contain a maximal sequence is that the boundary of* $\overline{\mathfrak{E}}$ *should be of zero* k*-dimensional measure. Each maximal sequence (if any) in* $\mathfrak{E}$ *is uniformly dense in the open subsets (if any) of* $\overline{\mathfrak{E}}$. *If* $\mathfrak{E}$ *is open, then*

$$L(\mathbf{P}\mathfrak{E}) \geqslant \alpha v^{1-q}(\mathfrak{E}) \tag{4·3}$$

for some $\mathbf{P}$ *in* $\mathfrak{E}$.

First, take $\mathfrak{E} = \mathfrak{C}$; and for any positive integer m, dissect $\mathfrak{C}$ into m^{k-1} disjoint identical cylinders, each of unit height and having as cross-section a $(k-1)$-dimensional semi-closed hypercube of side m^{-1}. Each point of $\mathbf{P}^n\mathfrak{C}$ lies not further than $\frac{1}{2}(k-1)^{\frac{1}{2}}m^{-1}$ from the axis of the cylinder containing it. Consider the closed path starting from one end of an axis nearest to a corner of $\mathfrak{C}$ and passing in turn along the axes of each cylinder, digressing at right-angles to and from any point of $\mathbf{P}^n$ in the same cylinder as that axis, going from axis to neighbouring axis in an obvious regular fashion on reaching the end of an axis, and finally returning direct to its starting point. The length of this path does not exceed the right-hand side of

$$l(\mathbf{P}^n\mathfrak{C}) \leqslant m^{k-1}.1 + n.(k-1)^{\frac{1}{2}}m^{-1} + m^{k-1}.m^{-1} + 1.k^{\frac{1}{2}}. \tag{4·4}$$

https://doi.org/10.1017/S0305004100034095 Published online by Cambridge University Press

Now take m to be the least integer exceeding $(k-1)^{-1/2k}n^{1/k}$, substitute in (4·4), divide through by $n^q k^{\frac{1}{2}}$, and let $n \to \infty$. This gives

$$k^{-\frac{1}{2}}L(\mathbf{P}\mathfrak{C}) \leqslant k^{\frac{1}{2}}(k-1)^{-(k-1)/2k}. \tag{4·5}$$

Define α_k to be the supremum of the left-hand side of (4·5) taken over all $\mathbf{P}$. This establishes (4·1) for $\mathfrak{E} = \mathfrak{C}$ and also proves the right-hand side of (4).

In accordance with a standard theorem (see, for example, (25), Theorem 1·6) we may cover $\overline{\mathfrak{E}}$ with a finite number, say $m = m(\mathfrak{E}, \epsilon)$, of disjoint semi-closed hypercubes whose measures V_j satisfy

$$\sum_{j=1}^{m} V_j \leqslant v(\overline{\mathfrak{E}}) + \tfrac{1}{2}\epsilon. \tag{4·6}$$

Let $\mathfrak{E}_j$ be the intersection of the jth hypercube with $\mathfrak{E}$, and let $n_j = N(\mathbf{P}^n\mathfrak{E}_j)$. Then

$$\sum_{j=1}^{m} n_j = N\left(\mathbf{P}^n \sum_{j=1}^{m} \mathfrak{E}_j\right) \leqslant n. \tag{4·7}$$

Making due allowance for a simple change of scale of linear magnitude V_j^{1-q}, we see that the definition of α ensures the existence of a number $a^*(V_j, \epsilon)$ such that

$$l(\mathbf{P}^n\mathfrak{E}_j) \leqslant \alpha n_j^q (V_j + \tfrac{1}{2}\epsilon m^{-1})^{1-q} + a^*(V_j, \epsilon), \tag{4·8}$$

because $l(\mathbf{P}^n\mathfrak{E}_j)$ is the tour-length through n_j points contained in a semi-closed hypercube of measure V_j and we have proved (4·1) for $\mathfrak{E} = \mathfrak{C}$. Define

$$a(\mathfrak{E}, \epsilon) = \sum_{j=1}^{m} \{a^*(V_j, \epsilon) + \delta(\mathfrak{E})\}. \tag{4·9}$$

Then Lemma 1, (4·8), Hölder's inequality, (4·6), (4·7) and (4·9) yield

$$\begin{aligned} l(\mathbf{P}^n\mathfrak{E}) &\leqslant \sum_{j=1}^{m} \{l(\mathbf{P}^n\mathfrak{E}_j) + \delta(\mathfrak{E})\} \leqslant \sum_{j=1}^{m} \{\alpha n_j^q (V_j + \tfrac{1}{2}\epsilon m^{-1})^{1-q} + a^*(V_j, \epsilon) + \delta(\mathfrak{E})\} \\ &\leqslant \alpha \left\{\sum_{j=1}^{m} n_j\right\}^q \left\{\sum_{j=1}^{m} (V_j + \tfrac{1}{2}\epsilon m^{-1})\right\}^{1-q} + a(\mathfrak{E}, \epsilon) \\ &\leqslant \alpha n^q \{v(\overline{\mathfrak{E}}) + \epsilon\}^{1-q} + a(\mathfrak{E}, \epsilon), \end{aligned} \tag{4·10}$$

which establishes (4·1) and hence (4·2).

Before proceeding to the remainder of Lemma 4, we make two remarks. First, $\alpha > 0$; this follows from Lemma 3. Secondly, if we define

$$\Lambda(\mathfrak{E}) = \sup_{\mathbf{P}} L(\mathbf{P}\mathfrak{E}), \tag{4·11}$$

then $\mathfrak{E}$ contains a sequence $\mathbf{P}$ such that this supremum is attained; that is to say

$$L(\mathbf{P}\mathfrak{E}) = \Lambda(\mathfrak{E}). \tag{4·12}$$

For, by (4·11), we can find $\mathbf{P}_j$ such that

$$\Lambda(\mathfrak{E}) \leqslant n^{-q} l(\mathbf{P}_j^n\mathfrak{E}) + j^{-1} \quad (n = n_{1j}, n_{2j}, \ldots), \tag{4·13}$$

where $n_{1j}, n_{2j}, \ldots$ is an increasing sequence of integers for each $j = 1, 2, \ldots$; whereupon we may take for the desired sequence $\mathbf{P}$ the sequence

$$P_{11}, P_{21}, \ldots, P_{A1}, P_{12}, P_{22}, \ldots, P_{B2}, P_{13}, P_{23}, \ldots,$$

where $A, B, \ldots$ is a sufficiently rapidly increasing sequence of the form $n_{X1}, n_{Y2}, \ldots$.

https://doi.org/10.1017/S0305004100034095 Published online by Cambridge University Press

We now prove (4·3). Since changes of scale are trivially allowed for, we may suppose that the given open set $\mathfrak{E}$ satisfies $\mathfrak{E} \subseteq \mathfrak{C}$. By the definition of α, (4·12) ensures the existence of $\mathbf{P}$ such that

$$L(\mathbf{P}\mathfrak{C}) = \alpha. \tag{4·14}$$

Then, since $L(\mathbf{P}\overline{\mathfrak{C}}) \geqslant L(\mathbf{P}\mathfrak{C})$, (4·14) and (4·2) yield

$$L(\mathbf{P}\overline{\mathfrak{C}}) = \alpha. \tag{4·15}$$

For brevity, write $V = v(\mathfrak{E})$. Since $\mathfrak{E}$ is open, $\overline{\mathfrak{C}} - \mathfrak{E}$ is a closed set of measure $1 - V$. Write $n_1 = N(\mathbf{P}^n\mathfrak{E})$, $n_2 = N[\mathbf{P}^n(\overline{\mathfrak{C}} - \mathfrak{E})]$, $l_1 = l(\mathbf{P}^n\mathfrak{E})$, and $l_2 = l(\mathbf{P}^n(\overline{\mathfrak{C}} - \mathfrak{E})]$. By Lemma 1 and (4·1) applied to l_2,

$$l(\mathbf{P}^n\overline{\mathfrak{C}}) \leqslant l_1 + l_2 + 2\delta(\overline{\mathfrak{C}}) \leqslant l_1 + \alpha n_2^q (1 - V + \epsilon)^{1-q} + a(\overline{\mathfrak{C}} - \mathfrak{E}, \epsilon) + 2\delta(\overline{\mathfrak{C}}). \tag{4·16}$$

Then (4·15) and (4·16) yield

$$\alpha \leqslant \limsup_{n\to\infty} (n_1/n)^q L(\mathbf{P}\mathfrak{E}) + \alpha(1 - V + \epsilon)^{1-q} \limsup_{n\to\infty} (n_2/n)^q. \tag{4·17}$$

Suppose, in order to obtain a contradiction, that

$$\limsup_{n\to\infty} (n_1/n)^q = 0. \tag{4·18}$$

By (4·2), $L(\mathbf{P}\mathfrak{E}) \leqslant \alpha v^{1-q}(\overline{\mathfrak{E}}) \leqslant \alpha$. Also $n_2/n \leqslant 1$. Hence (4·17) and (4·18) give

$$\alpha \leqslant \alpha(1 - V + \epsilon)^{1-q}; \tag{4·19}$$

and because $\alpha > 0$ and ϵ is arbitrary, (4·19) is a contradiction unless $V = 0$, in which case (4·3) is trivial. Hence we suppose (4·18) is false. By Hölder's inequality

$$(n_1/n)^q (V - \epsilon)^{1-q} + (n_2/n)^q (1 - V + \epsilon)^{1-q} \leqslant 1. \tag{4·20}$$

From (4·17) and (4·20),

$$\begin{aligned} \limsup_{n\to\infty} (n_1/n)^q L(\mathbf{P}\mathfrak{E}) &\geqslant \alpha\{1 - (1 - V + \epsilon)^{1-q} \limsup_{n\to\infty} (n_2/n)^q\} \\ &\geqslant \alpha(V - \epsilon)^{1-q} \limsup_{n\to\infty} (n_1/n)^q. \end{aligned} \tag{4·21}$$

Thereupon (4·3) follows from (4·21), the falsity of (4·18), and the arbitrariness of ϵ.

By (4·2), since $\overline{\overline{\mathfrak{E}}} = \overline{\mathfrak{E}}$, any maximal sequence in $\mathfrak{E}$ is a maximal sequence in $\overline{\mathfrak{E}}$. Conversely, suppose $\overline{\mathfrak{E}}$ contains a maximal sequence $\mathbf{P}$. Then we can find a $\mathbf{Q}$ in $\mathfrak{E}$ such that the distance between P_j and Q_j does not exceed 2^{-j}. Next we can find a closed path through $P_1, P_2, \ldots, P_n$ of length not exceeding $l(\mathbf{Q}^n) + 2\sum_{j=1}^{n} 2^{-j}$ by touring around $Q_1, Q_2, \ldots, Q_n$ and making digressions at each Q_j to and from P_j. Thus $l(\mathbf{Q}^n) + 2 \geqslant l(\mathbf{P}^n)$; and this implies that $\mathbf{Q}$ is a maximal sequence in $\mathfrak{E}$.

If $\overline{\mathfrak{E}}$ has a boundary of measure zero and $v(\overline{\mathfrak{E}}) > 0$, we can find in $\overline{\mathfrak{E}}$ an open set of measure $v(\overline{\mathfrak{E}})$, namely, the complement of the closure of the complement of $\overline{\mathfrak{E}}$. By (4·3) this open set contains a maximal sequence. This sequence is *a fortiori* a maximal sequence in $\overline{\mathfrak{E}}$.

https://doi.org/10.1017/S0305004100034095 Published online by Cambridge University Press

Finally, suppose that $\mathfrak{E}$ contains a maximal sequence $\mathbf{P}$; and suppose that $\overline{\mathfrak{E}}$ contains an open interval $\mathfrak{J}$ of (necessarily positive) measure V. Let

$$n_1 = N(\mathbf{P}^n\mathfrak{J}), \quad n_2 = N[\mathbf{P}^n(\mathfrak{E}-\mathfrak{J})], \quad l_1 = l(\mathbf{P}^n\mathfrak{J}), \quad l_2 = l[\mathbf{P}^n(\mathfrak{E}-\mathfrak{J})].$$

Since $\overline{\mathfrak{J}}$ also has measure V, Lemma 1 and (4·2) give

$$\alpha v^{1-q}(\overline{\mathfrak{E}}) \leqslant \limsup_{n\to\infty} \{\alpha(n_1/n)^q V^{1-q} + \alpha(n_2/n)^q [v(\overline{\mathfrak{E}}) - V]^{1-q}\}. \tag{4·22}$$

By Hölder's inequality, the reverse inequality (and therefore equality) holds in (4·22). But equality holds in Hölder's inequality only if the relevant terms are proportional. Thus

$$\limsup_{n\to\infty} n_1/n = V/v(\overline{\mathfrak{E}}). \tag{4·23}$$

Since (4·23) holds for every open interval $\mathfrak{J}$ in $\overline{\mathfrak{E}}$, $\mathbf{P}$ is uniformly dense in the open subsets of $\overline{\mathfrak{E}}$.

LEMMA 5. *There exists an absolute constant β_k such that for every Lebesgue set $\mathfrak{E}$*

$$\mathscr{E}l(\mathbf{P}\xi\mathfrak{E}) \sim \beta_k k^{\frac{1}{2}} v(\mathfrak{E}) = \beta \xi^k v(\mathfrak{E}) \quad as \quad \xi \to \infty \tag{5·1}$$

when $\mathbf{P} \in W_1$. Moreover for $t \geqslant 0$, the cumulant-generating function

$$K_t^*(\xi\mathfrak{E}) = \log \mathscr{E}\{\exp[tl(\mathbf{P}\xi\mathfrak{E})]\} \sim w(t)\, \xi^k v(\mathfrak{E}) \quad \text{as} \quad \xi \to \infty, \tag{5·2}$$

where $w(t)$ is a function of k and t only, such that $w(t) = o(t^k)$ as $t \to \infty$.

In the proof of Lemma 5 we assume without further mention that all sets are Lebesgue sets. It is convenient to prove (5·1) and (5·2) simultaneously by defining

$$K_t^{**}(\xi\mathfrak{E}) = \mathscr{E}tl(\mathbf{P}\xi\mathfrak{E}), \tag{5·3}$$

and writing K for K^* and K^{**} in all statements true for both K^* and K^{**}. Since e^x is a convex increasing function of x, we have

$$0 \leqslant K_t^{**}(\xi\mathfrak{E}) \leqslant K_t^*(\xi\mathfrak{E}). \tag{5·4}$$

Let $\mathfrak{D}_1, \mathfrak{D}_2, \ldots, \mathfrak{D}_m$ be disjoint sets; and write $\mathfrak{D}_{m+1} = \mathfrak{D}_1$. The disjointness together with the hypothesis $\mathbf{P} \in W_1$ implies that $\mathbf{P}\mathfrak{D}_1, \mathbf{P}\mathfrak{D}_2, \ldots, \mathbf{P}\mathfrak{D}_m$ are stochastically independent. Hence $l(\mathbf{P}\mathfrak{D}_1), l(\mathbf{P}\mathfrak{D}_2), \ldots, l(\mathbf{P}\mathfrak{D}_m)$ are independent. Hence, by Lemma 1,

$$\begin{aligned} K_t^*\left(\sum_{j=1}^m \mathfrak{D}_j\right) &= \log \mathscr{E} \exp\left[tl\left(\mathbf{P}\sum_{j=1}^m \mathfrak{D}_j\right)\right] \\ &\leqslant \log \mathscr{E} \exp\left[t\sum_{j=1}^m \{l(\mathbf{P}\mathfrak{D}_j) + \delta(\mathfrak{D}_j + \mathfrak{D}_{j+1})\}\right] \\ &= \sum_{j=1}^m \{K_t^*(\mathfrak{D}_j) + t\delta(\mathfrak{D}_j + \mathfrak{D}_{j+1})\}. \end{aligned} \tag{5·5}$$

A similar result holds for K_t^{**}; and so we have

$$K_t\left(\sum_{j=1}^m \mathfrak{D}_j\right) \leqslant \sum_{j=1}^m \{K_t(\mathfrak{D}_j) + t\delta(\mathfrak{D}_j + \mathfrak{D}_{j+1})\}. \tag{5·6}$$

If $n = N(\mathbf{P}\mathfrak{E}), l(\mathbf{P}\mathfrak{E}) \leqslant n\delta(\mathfrak{E})$. Hence

$$K_t^*(\mathfrak{E}) \leqslant \log \sum_{n=0}^{\infty} e^{-v(\mathfrak{E})} \frac{[v(\mathfrak{E})]^n}{n!} e^{tn\delta(\mathfrak{E})} = v(\mathfrak{E})\,(e^{t\delta(\mathfrak{E})} - 1). \tag{5·7}$$

https://doi.org/10.1017/S0305004100034095 Published online by Cambridge University Press

By (5·4) and (5·7),
$$0 \leqslant K_t(\mathfrak{E}) \leqslant v(\mathfrak{E})\,(e^{t\delta(\mathfrak{E})}-1). \tag{5·8}$$

Define
$$\chi_t(\mathfrak{E}) = \limsup_{\xi\to\infty} \xi^{-k} K_t(\xi\mathfrak{E}), \tag{5·9}$$

so that
$$\chi_t(a\mathfrak{E}) = \limsup_{\xi\to\infty} \xi^{-k} K_t(\xi a\mathfrak{E}) = a^k \limsup_{\xi a\to\infty} (\xi a)^{-k} K_t(\xi a\mathfrak{E}) = a^k \chi_t(\mathfrak{E}). \tag{5·10}$$

This definition is valid provided that we work with extended real numbers (i.e. including $+\infty$) as far as (5·14), where we prove that χ_t is actually finite. From (5·5) and (5·9)
$$0 \leqslant \chi_t\Big(\sum_{j=1}^{m} \mathfrak{D}_j\Big) \leqslant \sum_{j=1}^{m} \chi_t(\mathfrak{D}_j). \tag{5·11}$$

Choose a such that $\mathfrak{E} \subseteq a\mathfrak{C}$. Cover $a\xi\mathfrak{C}$ with disjoint $\eta\mathfrak{C}_1, \eta\mathfrak{C}_2, \ldots, \eta\mathfrak{C}_m$, where $m^{1/k}$ is the integer part of $1+a\xi/\eta$; and arrange for $\eta\mathfrak{C}_j$ and $\eta\mathfrak{C}_{j+1}$ to abut facially for $j = 1, 2, \ldots, m-1$. From (5·6), we have, with $\mathfrak{C}_{m+1} = \mathfrak{C}_1$,

$$\begin{aligned}
\chi_t(\mathfrak{E}) &= \limsup_{\xi\to\infty} \xi^{-k} K_t\Big[\sum_{j=1}^{m} (\eta\mathfrak{C}_j)\,(\xi\mathfrak{E})\Big] \\
&\leqslant \limsup_{\xi\to\infty} \xi^{-k} \sum_{j=1}^{m} \{K_t[(\eta\mathfrak{C}_j)\,(\xi\mathfrak{E})] + t\delta[(\eta\mathfrak{C}_j + \eta\mathfrak{C}_{j+1})\,(\xi\mathfrak{E})]\} \\
&\leqslant \Big\{\limsup_{\xi\to\infty} \xi^{-k} \sum_{j=1}^{m} K_t[(\eta\mathfrak{C}_j)\,(\xi\mathfrak{E})]\Big\} + a^k(k+3)^{\frac{1}{2}}\eta^{1-k}t.
\end{aligned} \tag{5·12}$$

From (5·8)
$$\begin{aligned}
\sum_{j=1}^{m} K_t[(\eta\mathfrak{C}_j)\,(\xi\mathfrak{E})] &\leqslant \sum_{j=1}^{m} v[(\eta\mathfrak{C}_j)\,(\xi\mathfrak{E})]\,\{\exp(t\delta[(\eta\mathfrak{C}_j)\,(\xi\mathfrak{E})]) - 1\} \\
&\leqslant (e^{t\eta\sqrt{k}}-1) \sum_{j=1}^{m} v[(\eta\mathfrak{C}_j)\,(\xi\mathfrak{E})] = (e^{t\eta\sqrt{k}}-1)\,v(\xi\mathfrak{E}).
\end{aligned} \tag{5·13}$$

From (5·12) and (5·13)
$$0 \leqslant \chi_t(\mathfrak{E}) \leqslant (e^{t\eta\sqrt{k}}-1)\,v(\mathfrak{E}) + a^k(k+3)^{\frac{1}{2}}\eta^{1-k}t, \tag{5·14}$$

which shows that χ_t is finite.

Now consider the special case $\mathfrak{E} = a\mathfrak{C}$. We have
$$K_t[(\eta\mathfrak{C}_j)\,(\xi\mathfrak{C})] \leqslant K_t(\eta\mathfrak{C}_j), \tag{5·15}$$

because $l[\mathbf{P}(\eta\mathfrak{C}_j)\,(\xi\mathfrak{C})] \leqslant l(\mathbf{P}\eta\mathfrak{C}_j)$. (This is also true for l^{**} in place of l, because $(\eta\mathfrak{C}_j)\,(\xi\mathfrak{C})$ is a convex set.) Hence (5·12) and (5·15) give

$$\chi_t(a\mathfrak{C}) \leqslant \Big\{\limsup_{\xi\to\infty} \xi^{-k} \sum_{j=1}^{m} K_t(\eta\mathfrak{C}_j)\Big\} + a^k(k+3)^{\frac{1}{2}}\eta^{1-k}t = (a/\eta)^k K_t(\eta\mathfrak{C}) + a^k(k+3)^{\frac{1}{2}}\eta^{1-k}t. \tag{5·16}$$

Let $\eta \to \infty$ in (5·16), and apply (5·9) and (5·10), with the result
$$\chi_t(a\mathfrak{C}) \leqslant a^k \liminf_{\eta\to\infty} \eta^{-k} K_t(\eta\mathfrak{C}) \leqslant a^k \chi_t(\mathfrak{C}) = \chi_t(a\mathfrak{C}). \tag{5·17}$$

This implies the existence of
$$\lim_{\xi\to\infty} \xi^{-k} K_t(\xi a\mathfrak{C}) = a^k \chi_t(\mathfrak{C}) = a^k w(t), \quad \text{say.} \tag{5·18}$$

Of course, $w(t)$ is one function when $K = K^*$ and is another function when $K = K^{**}$.

https://doi.org/10.1017/S0305004100034095 Published online by Cambridge University Press

Next let $\mathfrak{E}$ be once more an arbitrary Lebesgue set. Prescribe ϵ such that $0 < \epsilon < 1$. Because $\mathfrak{E}$ is a Lebesgue set we can find disjoint $a_1\mathfrak{C}_1, a_2\mathfrak{C}_2, \ldots, a_\mu\mathfrak{C}_\mu$ such that

$$\sum_{j=1}^{\mu} a_j^k = \sum_{j=1}^{\mu} v(a_j\mathfrak{C}_j) \leqslant v(\mathfrak{E}) + \epsilon; \quad v\left(\mathfrak{E} - \sum_{j=1}^{\mu} a_j\mathfrak{C}_j\right) \leqslant \epsilon. \tag{5·19}$$

By (5·11), (5·18), (5·14) and (5·19)

$$\begin{aligned}\chi_t(\mathfrak{E}) &\leqslant \left\{\sum_{j=1}^{\mu} \chi_t(a_j\mathfrak{C}_j)\right\} + \chi_t\left(\mathfrak{E} - \sum_{j=1}^{\mu} a_j\mathfrak{C}_j\right) = w(t)\sum_{j=1}^{\mu} a_j^k + \chi_t\left(\mathfrak{E} - \sum_{j=1}^{\mu} a_j\mathfrak{C}_j\right) \\ &\leqslant w(t)\,[v(\mathfrak{E}) + \epsilon] + (e^{t\eta\sqrt{k}} - 1)\,\epsilon + a^k(k+3)^{\frac{1}{2}}\eta^{1-k}t \\ &\leqslant w(t)\,[v(\mathfrak{E}) + \epsilon] + \epsilon^{\frac{1}{2}} + (2ak^{\frac{1}{2}}t)^k(-\log\epsilon)^{1-k},\end{aligned} \tag{5·20}$$

upon choosing $\eta = -\frac{1}{2}k^{-\frac{1}{2}}t^{-1}\log\epsilon$. Let $\epsilon \to 0$ in (5·20), with the result

$$\chi_t(\mathfrak{E}) \leqslant v(\mathfrak{E})\,w(t). \tag{5·21}$$

By (5·6)
$$K_t(a\mathfrak{C}) \leqslant K_t(\mathfrak{E}) + K_t[(a\mathfrak{C}) - \mathfrak{E}] + 2ta\sqrt{k}. \tag{5·22}$$

Hence, from (5·22), (5·18) and (5·21),

$$\begin{aligned}\liminf_{\xi\to\infty} \xi^{-k} K_t(\xi\mathfrak{E}) &\geqslant \lim_{\xi\to\infty} \xi^{-k} K_t(a\xi\mathfrak{C}) - \limsup_{\xi\to\infty} \xi^{-k} K_t[\xi\{(a\mathfrak{C}) - \mathfrak{E}\}] \\ &\geqslant a^k w(t) - w(t)\,v[(a\mathfrak{C}) - \mathfrak{E}] = w(t)\,v(\mathfrak{E}).\end{aligned} \tag{5·23}$$

Thus, by (5·9), (5·21), and (5·23)

$$\lim_{\xi\to\infty} \xi^{-k} K_t(\xi\mathfrak{E}) = w(t)\,v(\mathfrak{E}). \tag{5·24}$$

To prove $w(t) = o(t^k)$ as $t \to \infty$, put $\mathfrak{E} = \mathfrak{C}$ in (5·14) and choose $\eta = k^{-\frac{1}{2}}t^{-1}\log t$. The required result (5·2) follows from (5·24) with K^* in place of K. Similarly (5·1) follows from (5·24) with K^{**} in place of K and $t = 1$.

LEMMA 6. *If* $\mathbf{P} \in W_1$ *and* $\mathfrak{E}$ *is a Lebesgue set satisfying* $\mathfrak{E} \subseteq \mathfrak{C}$, *then*

$$\operatorname{var} l(\mathbf{P}\xi\mathfrak{E}) = \left.\begin{cases} O(\xi^{2k-2/(k-1)}) & (k > 2) \\ O(\xi^2\log^2\xi) & (k = 2) \end{cases}\right\} \quad as \quad \xi \to \infty \tag{6·1}$$

uniformly in $\mathfrak{E}$.

By Lemma 4, there exists an absolute constant c_k such that

$$l(\mathbf{P}^n\overline{\mathfrak{C}}) \leqslant c_k n^q = c_k n^{(k-1)/k}, \tag{6·2}$$

for any sequence $\mathbf{P}$ whatsoever; and hence

$$l(\mathbf{P}^n\xi\mathfrak{E}) \leqslant l(\mathbf{P}^n\xi\mathfrak{C}) \leqslant l(\mathbf{P}^n\xi\overline{\mathfrak{C}}) \leqslant c_k n^q \xi. \tag{6·3}$$

(The conclusion $l^{**}(\mathbf{P}^n\xi\mathfrak{E}) \leqslant c_k^{**} n^q \xi$ follows from (6·3) and (14).) The inequality (6·3) also holds trivially when $k = 1$ if we take $c_1 = 1$.

Now let $\mathbf{P} \in W_1$. By (6·3),

$$\begin{aligned}0 \leqslant \operatorname{var} l(\mathbf{P}\xi\mathfrak{E}) \leqslant \mathscr{E}[l(\mathbf{P}\xi\mathfrak{E})]^2 &\leqslant \sum_{n=0}^{\infty} e^{-v(\xi\mathfrak{E})}\frac{[v(\xi\mathfrak{E})]^n}{n!} c_k^2 n^{2q}\xi^2 \leqslant \sum_{n=0}^{\infty} e^{-v(\xi\mathfrak{E})}\frac{[v(\xi\mathfrak{E})]^n}{n!} c_k^2 n^2\xi^2 \\ &= c_k^2\{[v(\xi\mathfrak{E})]^2 + v(\xi\mathfrak{E})\}\,\xi^2 \leqslant c_k^2(\xi^{2k} + \xi^k)\,\xi^2.\end{aligned} \tag{6·4}$$

https://doi.org/10.1017/S0305004100034095 Published online by Cambridge University Press

Hence we may define the finite function

$$w(\xi) = \sup_{\mathfrak{E} \subseteq \mathfrak{C}} [\operatorname{var} l(\mathbf{P}\xi(\mathfrak{E})]^{\frac{1}{2}}. \tag{6·5}$$

Cover $\mathfrak{C}$ with 2^k equal disjoint hypercubes $\frac{1}{2}\mathfrak{C}_1, \frac{1}{2}\mathfrak{C}_2, \ldots, \frac{1}{2}\mathfrak{C}_{2^k}$; and write

$$L_1 = l(\mathbf{P}\xi(\mathfrak{E}) = l\left(\mathbf{P}\xi \sum_{j=1}^{2^k} \mathfrak{E}\tfrac{1}{2}\mathfrak{C}_j\right); \quad L_2 = \sum_{j=1}^{2^k} l(\mathbf{P}\xi(\mathfrak{E}\tfrac{1}{2}\mathfrak{C}_j). \tag{6·6}$$

We have

$$\begin{aligned} \operatorname{var} L_1 &= \operatorname{var} L_2 + 2\operatorname{cov}(L_2, L_1 - L_2) + \operatorname{var}(L_1 - L_2) \\ &\leqslant \{[\operatorname{var} L_2]^{\frac{1}{2}} + [\operatorname{var}(L_1 - L_2)]^{\frac{1}{2}}\}^2 \leqslant \{[\operatorname{var} L_2]^{\frac{1}{2}} + [\mathscr{E}(L_1 - L_2)^2]^{\frac{1}{2}}\}^2. \end{aligned} \tag{6·7}$$

The lengths $l(\mathbf{P}\xi(\mathfrak{E}\frac{1}{2}\mathfrak{C}_j)$ are stochastically independent because $\mathbf{P} \in W_1$. Thus

$$\operatorname{var} L_2 = \sum_{j=1}^{2^k} \operatorname{var} l(\mathbf{P}\xi(\mathfrak{E}\tfrac{1}{2}\mathfrak{C}_j) = \sum_{j=1}^{2^k} \operatorname{var} l[\mathbf{P}\tfrac{1}{2}\xi 2(\mathfrak{E}\tfrac{1}{2}\mathfrak{C}_j)] \leqslant \sum_{j=1}^{2^k} [w(\tfrac{1}{2}\xi)]^2 = 2^k[w(\tfrac{1}{2}\xi)]^2, \tag{6·8}$$

because $2(\mathfrak{E}\frac{1}{2}\mathfrak{C}_j)$ is a subset of some semi-closed unit hypercube.

By Lemma 1 we have

$$L_1 - L_2 \leqslant \sum_{j=1}^{2^k} \delta(\xi\mathfrak{E}) \leqslant 2^k k^{\frac{1}{2}}\xi. \tag{6·9}$$

To get an inequality in the opposite direction to (6·9), we write $n = N(\mathbf{P}\xi(\mathfrak{E})$ and consider the situation conditional upon n being fixed. We may then number the points of $\mathbf{P}$ so that the first n of them are the ones in $\xi\mathfrak{E}$. In this case

$$L_1 = l(\mathbf{P}\xi(\mathfrak{E}) = l(\mathbf{P}^n\xi(\mathfrak{C}); \quad L_2 = \sum_{j=1}^{2^k} l(\mathbf{P}^n\xi\tfrac{1}{2}\mathfrak{C}_j). \tag{6·10}$$

Applying Lemma 2 to (6·10), we have

$$L_1 - L_2 \geqslant -2\sum_{j=1}^{2^k} \sup_{\mathbf{Q}_j} l[\mathbf{Q}_j^{2N(\mathbf{P}^n\frac{1}{2}\mathfrak{C}_j)}(\overline{\tfrac{1}{2}\xi(\mathfrak{C}_j)})\,(\overline{\mathfrak{R} - \tfrac{1}{2}\xi(\mathfrak{C}_j)})] \geqslant -2^{k+1}\sup_{\mathbf{Q}} l[\mathbf{Q}^{2n}(\overline{\tfrac{1}{2}\xi(\mathfrak{C})})\,(\overline{\mathfrak{R} - \tfrac{1}{2}\xi(\mathfrak{C})})]. \tag{6·11}$$

Now $l[\mathbf{Q}^{2n}(\overline{\frac{1}{2}\xi(\mathfrak{C})})\,(\overline{\mathfrak{R} - \frac{1}{2}\xi(\mathfrak{C})})]$ is the length of a path through $2n$ points on the boundary of $\frac{1}{2}\xi\mathfrak{C}$; and is therefore, by Lemma 1, at most equal to $2k[l' + \delta(\overline{\frac{1}{2}\xi(\mathfrak{C})})]$, where l' is the maximum length of a path through $2n$ points on one of the $2k$ faces of $\frac{1}{2}\xi\mathfrak{C}$. Since each face is a $(k-1)$-dimensional hypercube of side $\frac{1}{2}\xi$, we have $l' \leqslant c_{k-1}\frac{1}{2}\xi(2n)^{(k-2)/(k-1)}$ by (6·3). Thus

$$L_1 - L_2 \geqslant -2^{k+1} \cdot 2k[c_{k-1}\tfrac{1}{2}\xi(2n)^{(k-2)/(k-1)} + \tfrac{1}{2}k^{\frac{1}{2}}\xi]. \tag{6·12}$$

By (6·9) and (6·12)

$$(L_1 - L_2)^2 \leqslant 2^{2k+2}k^3\xi^2[k^{-\frac{1}{2}}c_{k-1}(2n)^{(k-2)/(k-1)} + 1]^2. \tag{6·13}$$

Now if ρ and ϵ are positive constants, we have a standard result ((10), Theorem 137 (6))

$$\sum_{n=0}^{\infty} e^{-\eta}\frac{\eta^n}{n!}n^\rho \leqslant [(1+\epsilon)\eta]^\rho + \sum_{n \geqslant (1+\epsilon)\eta} e^{-\eta}\frac{\eta^n}{n!}n^\rho = [(1+\epsilon)\eta]^\rho + O(e^{-\frac{1}{2}\epsilon^2\eta}) \tag{6·14}$$

as $\eta \to \infty$. Hence (6·13) and (6·14) give

$$\begin{aligned} \mathscr{E}(L_1 - L_2)^2 &\leqslant \sum_{n=0}^{\infty} e^{-v(\xi\mathfrak{E})}\frac{[v(\xi(\mathfrak{E})]^n}{n!}2^{2k+2}k^3\xi^2[k^{-\frac{1}{2}}c_{k-1}(2n)^{(k-2)/(k-1)} + 1]^2 \\ &\leqslant \xi^2(A\xi^{k(k-2)/(k-1)} + B)^2 \quad (\xi \geqslant 1), \end{aligned} \tag{6·15}$$

https://doi.org/10.1017/S0305004100034095 Published online by Cambridge University Press

where A and B are absolute constants (depending only on k) since we may replace $\mathfrak{E}$ by $\mathfrak{C}$ after summation without decreasing the sum. From (6·7), (6·8) and (6·15), we get

$$[\operatorname{var} L_1]^{\frac{1}{2}} \leqslant 2^{k/2} w(\tfrac{1}{2}\xi) + \xi(A\xi^{k(k-2)/(k-1)} + B) \quad (\xi \geqslant 1). \tag{6·16}$$

In (6·16) the right-hand side is independent of $\mathfrak{E}$, and hence taking the supremum of the left-hand side for all $\mathfrak{E} \subseteq \mathfrak{C}$, we get

$$w(\xi) \leqslant 2^{k/2} w(\tfrac{1}{2}\xi) + \xi(A\xi^{k(k-2)/(k-1)} + B) \quad (\xi \geqslant 1). \tag{6·17}$$

Now define
$$x(\xi) = w(\xi)/\xi^{k-1/(k-1)}. \tag{6·18}$$

Then (6·17) and (6·18) give

$$x(\xi) \leqslant 2^{-\frac{1}{2}k+1/(k-1)} x(\tfrac{1}{2}\xi) + A + B\xi^{1-k+1/(k-1)} \leqslant 2^{-\frac{1}{2}k+1/(k-1)} x(\tfrac{1}{2}\xi) + C$$
$$(C = A + B, \xi \geqslant 1). \tag{6·19}$$

If $k > 2$, (6·19) gives
$$x(\xi) \leqslant \tfrac{1}{2} x(\tfrac{1}{2}\xi) + C; \tag{6·20}$$

and hence

$$x(2^n\xi) \leqslant C + \tfrac{1}{2}x(2^{n-1}\xi) \leqslant C(1+\tfrac{1}{2}) + \tfrac{1}{4}x(2^{n-2}\xi) \leqslant \ldots \leqslant 2C + 2^{-n}x(\xi). \tag{6·21}$$

Thus
$$x(\eta) \leqslant 2C + ([\log_2 \eta])^{-1} \sup_{1 \leqslant \xi \leqslant 2} x(\xi), \tag{6·22}$$

where $[z]$ is the integer part of z and $\eta = 2^n\xi$. Thus, in the case $k > 2$, (6·1) follows from (6·18), (6·4) and (6·22).

When $k = 2$, (6·19) gives
$$x(\xi) \leqslant x(\tfrac{1}{2}\xi) + C, \tag{6·23}$$

in which we make the substitution

$$y(\xi) = x(\xi)/\log \xi, \tag{6·24}$$

with the result
$$[y(\tfrac{1}{2}\xi) - y(\xi)] \log \tfrac{1}{2}\xi \geqslant y(\xi) \log 2 - C. \tag{6·25}$$

We shall show that (6·1) follows from (6·25), inasmuch as the assumption that $y(\xi)$ is unbounded leads to a contradiction in the following way. If $y(\xi)$ is unbounded there exists $\xi_0 > 2$ such that

$$y(\xi_0) > \max\left\{\frac{C}{\log 2}, \sup_{2 \leqslant \xi < 4} y(\xi)\right\}. \tag{6·26}$$

Substitution of (6·26) into the right-hand side of (6·25) gives

$$y(\tfrac{1}{2}\xi_0) \geqslant y(\xi_0) > \frac{C}{\log 2}, \tag{6·27}$$

since $\log \frac{1}{2}\xi_0 > 0$. If $\frac{1}{2}\xi_0 > 2$, we may repeat the process with $\xi = \frac{1}{2}\xi_0$ in (6·25); and so on until we get

$$y(2^{-n}\xi_0) \geqslant y(2^{-n+1}\xi_0) \geqslant \ldots \geqslant y(\xi_0) > \sup_{2 \leqslant \xi < 4} y(\xi), \tag{6·28}$$

where $n+1$ is the integer part of $\log_2 \xi_0$. However (6·28) is a contradiction because $2 \leqslant 2^{-n}\xi_0 < 4$.

LEMMA 7. *If* $\mathbf{P} \in u(\mathfrak{E})$, (2) *holds with probability one.*

Since we may make a simple change of scale, there is no essential loss of generality in assuming $\mathfrak{E} \subseteq \mathfrak{C}$. Let $\mathbf{P}_0 \in W_1$ and $\mathbf{P}_1 \in W_{\xi t}$. Since the distribution of $l(\mathbf{P}_0\xi\mathfrak{E})$ is the same as the distribution of $\xi l(\mathbf{P}_1\mathfrak{E})$, Lemmas 5 and 6 give

$$\mathscr{E} l(\mathbf{P}_1\mathfrak{E}) = \xi^{-1}\mathscr{E} l(\mathbf{P}_0\xi\mathfrak{E}) \sim \beta\xi^{k-1} v(\mathfrak{E}) \quad \text{as} \quad \xi \to \infty \tag{7·1}$$

https://doi.org/10.1017/S0305004100034095 Published online by Cambridge University Press

and $$\operatorname{var} l(\mathbf{P}_1\mathfrak{E}) = \xi^{-2}\operatorname{var} l(\mathbf{P}_0\xi\mathfrak{E}) = O(\xi^{2(k-1)-\{2/(k-1)\}+2\epsilon}) \quad \text{as} \quad \xi\to\infty, \tag{7·2}$$

for any prescribed $\epsilon > 0$. Hence for all sufficiently large ξ

$$|\mathscr{E} l(\mathbf{P}_1\mathfrak{E}) - \beta\xi^{k-1}v(\mathfrak{E})| < \tfrac{1}{2}\epsilon\xi^{k-1}v(\mathfrak{E}); \tag{7·3}$$

and therefore, by Chebyshev's inequality,

$$\begin{aligned}\operatorname{prob}[|l(\mathbf{P}_1\mathfrak{E}) - \beta\xi^{k-1}v(\mathfrak{E})| \leqslant \epsilon\xi^{k-1}v(\mathfrak{E})] &\geqslant \operatorname{prob}[|l(\mathbf{P}_1\mathfrak{E}) - \mathscr{E} l(\mathbf{P}_1\mathfrak{E})| \leqslant \tfrac{1}{2}\epsilon\xi^{k-1}v(\mathfrak{E})] \\ &\geqslant 1 - O[\xi^{(k-1)-\{1/(k-1)\}+\epsilon}/\tfrac{1}{2}\epsilon\xi^{k-1}v(\mathfrak{E})]^2 \geqslant 1 - \tfrac{1}{2}\xi^{-2/k},\end{aligned} \tag{7·4}$$

provided ξ is sufficiently large and $0 < \epsilon < 1/k(k-1)$.

Next, writing $\xi^k v(\mathfrak{E}) = \eta$, we have

$$\begin{aligned}\operatorname{prob}[(1-\epsilon)\eta \leqslant N(\mathbf{P}_1\mathfrak{E}) \leqslant (1+\epsilon)\eta] &= \sum_{(1-\epsilon)\eta\leqslant n\leqslant(1+\epsilon)\eta} e^{-\eta}\eta^n/n! \\ &= 1 - O(e^{-\frac{1}{3}\epsilon^2\eta}) \geqslant 1 - \tfrac{1}{2}\xi^{-2/k}\end{aligned} \tag{7·5}$$

for large enough ξ. Since the pair of relations

$$\left.\begin{aligned}(\beta-\epsilon)\,\xi^{k-1}v(\mathfrak{E}) \leqslant l(\mathbf{P}_1\mathfrak{E}) \leqslant (\beta+\epsilon)\,\xi^{k-1}v(\mathfrak{E}), \\ (1-\epsilon)\,\xi^k v(\mathfrak{E}) \leqslant N(\mathbf{P}_1\mathfrak{E}) \leqslant (1+\epsilon)\,\xi^k v(\mathfrak{E})\end{aligned}\right\} \tag{7·6}$$

imply the relation $$\frac{(\beta-\epsilon)}{(1+\epsilon)^q} \leqslant \frac{l(\mathbf{P}_1\mathfrak{E})}{[N(\mathbf{P}_1\mathfrak{E})]^q v^{1-q}(\mathfrak{E})} \leqslant \frac{(\beta+\epsilon)}{(1-\epsilon)^q}, \tag{7·7}$$

it follows from (7·4) and (7·5) that the probability of (7·7) is at least $1-\xi^{-2/k}$ for large enough ξ.

Next define $f(n)$ by

$$1 - f(n) = \operatorname{prob}\left[\frac{(\beta-\epsilon)}{(1+\epsilon)^q} \leqslant \frac{l(\mathbf{P}^n)}{n^q v^{1-q}(\mathfrak{E})} \leqslant \frac{(\beta+\epsilon)}{(1-\epsilon)^q}\right], \tag{7·8}$$

in which $\mathbf{P}\in u(\mathfrak{E})$. Since the conditional distribution of $l(\mathbf{P}_1\mathfrak{E})$ given that $N(\mathbf{P}_1\mathfrak{E}) = n$ is the unconditional distribution of $l(\mathbf{P}^n)$, we have

$$\sum_{n=0}^{\infty} e^{-\eta}\frac{\eta^n}{n!}[1-f(n)] = \operatorname{prob}[(7\cdot7)\text{ holds}] \geqslant 1 - \xi^{-2/k}. \tag{7·9}$$

Since $0 \leqslant 1-f(n) \leqslant 1$, (7·9) gives

$$\sup_{(1-\epsilon)\eta\leqslant n\leqslant(1+\epsilon)\eta}[1-f(n)] \geqslant \frac{1-\xi^{-2/k} - \sum_{|n-\eta|>\epsilon\eta} e^{-\eta}\eta^n/n!}{\sum_{|n-\eta|\leqslant\epsilon\eta} e^{-\eta}\eta^n/n!} \geqslant 1 - 2\xi^{-2/k} = 1 - A\eta^{-2/k^2}, \tag{7·10}$$

where A is a constant (depending upon ϵ and $\mathfrak{E}$) and ξ and η are sufficiently large.

Now let J_m denote the set of integers n satisfying

$$\left(\frac{1+\epsilon}{1-\epsilon}\right)^m \leqslant n \leqslant \left(\frac{1+\epsilon}{1-\epsilon}\right)^{m+1} \quad (m = 1, 2, \ldots). \tag{7·11}$$

Then (7·10) gives, for large enough m,

$$\sup_{n\in J_m}[1-f(n)] \geqslant 1 - A\left(\frac{1+\epsilon}{1-\epsilon}\right)^{-2m/k^2}. \tag{7·12}$$

https://doi.org/10.1017/S0305004100034095 Published online by Cambridge University Press

Since J_m contains only a finite number of integers, it contains an integer n_m such that $1-f(n_m) = \sup_{n \in J_m}[1-f(n)]$ for each large enough integer m. Consequently, there is an integer m_0 such that

$$\sum_{m=m_0}^{\infty}\left\{1-\text{prob}\left[\frac{(\beta-\epsilon)}{(1+\epsilon)^q} \leqslant \frac{l(\mathbf{P}^{n_m})}{n_m^q v^{1-q}(\mathfrak{E})} \leqslant \frac{(\beta+\epsilon)}{(1-\epsilon)^q}\right]\right\} = \sum_{m=m_0}^{\infty} f(n_m) \leqslant \sum_{m=m_0}^{\infty} A\left(\frac{1+\epsilon}{1-\epsilon}\right)^{-2m/k^2} < \infty. \tag{7·13}$$

By the Borel–Cantelli lemma, (7·13) implies that with probability one

$$\frac{(\beta-\epsilon)}{(1+\epsilon)^q} \leqslant \liminf_{m\to\infty} \frac{l(\mathbf{P}^{n_m})}{n_m^q v^{1-q}(\mathfrak{E})} \leqslant \limsup_{m\to\infty} \frac{l(\mathbf{P}^{n_m})}{n_m^q v^{1-q}(\mathfrak{E})} \leqslant \frac{(\beta+\epsilon)}{(1-\epsilon)^q}. \tag{7·14}$$

Since ϵ is arbitrary, (7·14) yields

$$\lim_{m\to\infty} \frac{l(\mathbf{P}^{n_m})}{n_m^q} = \beta v^{1-q}(\mathfrak{E}) \tag{7·15}$$

with probability one.

Next let n be any member of J_m and write $\mu_n = n_{m-1}$, $\nu_n = n_{m+1}$. We have

$$\left(\frac{1+\epsilon}{1-\epsilon}\right)^{m-1} \leqslant \mu_n \leqslant \left(\frac{1+\epsilon}{1-\epsilon}\right)^{m} \leqslant n \leqslant \left(\frac{1+\epsilon}{1-\epsilon}\right)^{m+1} \leqslant \nu_n \leqslant \left(\frac{1+\epsilon}{1-\epsilon}\right)^{m+2}; \tag{7·16}$$

so that, when ϵ is sufficiently small,

$$0 \leqslant n-\mu_n \leqslant 5\epsilon n; \quad 0 \leqslant \nu_n - n \leqslant 5\epsilon n. \tag{7·17}$$

Thus $\mathbf{P}^{\nu_n} - \mathbf{P}^n$ consists of a set of not more than $5\epsilon n$ points in $\mathfrak{E}$; so that Lemmas 1 and 4 give

$$\begin{aligned} l(\mathbf{P}^{\nu_n}) &\leqslant l(\mathbf{P}^n) + l(\mathbf{P}^{\nu_n} - \mathbf{P}^n) + 2\delta(\mathfrak{E}) \\ &\leqslant l(\mathbf{P}^n) + (5\epsilon n)^q [v(\overline{\mathfrak{E}}) + \epsilon]^{1-q} + a(\mathfrak{E}, \epsilon) + 2\delta(\mathfrak{E}). \end{aligned} \tag{7·18}$$

Hence $$\nu_n^{-q} l(\mathbf{P}^{\nu_n}) \leqslant n^{-q} l(\mathbf{P}^{\nu_n}) \leqslant n^{-q} l(\mathbf{P}^n) + (5\epsilon)^q [v(\overline{\mathfrak{E}}) + \epsilon]^{1-q} + O(n^{-q}). \tag{7·19}$$

Since ν_n belongs to the sequence $\{n_m\}$, (7·15) gives

$$\beta v^{1-q}(\mathfrak{E}) \leqslant \liminf_{n\to\infty} n^{-q} l(\mathbf{P}^n) + (5\epsilon)^q [v(\overline{\mathfrak{E}}) + \epsilon]^{1-q} \tag{7·20}$$

with probability one. Similarly

$$\begin{aligned} l(\mathbf{P}^{\mu_n}) &\geqslant l(\mathbf{P}^n) - l(\mathbf{P}^n - \mathbf{P}^{\mu_n}) - 2\delta(\mathfrak{E}) \\ &\geqslant l(\mathbf{P}^n) - (5\epsilon n)^q [v(\overline{\mathfrak{E}}) + \epsilon]^{1-q} - a(\mathfrak{E}, \epsilon) - 2\delta(\mathfrak{E}), \end{aligned} \tag{7·21}$$

whence $$\mu_n^{-q} l(\mathbf{P}^{\mu_n}) \geqslant n^{-q} l(\mathbf{P}^{\mu_n}) \geqslant n^{-q} l(\mathbf{P}^n) - (5\epsilon)^q [v(\overline{\mathfrak{E}}) + \epsilon]^{1-q} - O(n^{-q});$$

and so, by (7·15),

$$\beta v^{1-q}(\mathfrak{E}) \geqslant \limsup_{n\to\infty} n^{-q} l(\mathbf{P}^n) - (5\epsilon)^q [v(\overline{\mathfrak{E}}) + \epsilon]^{1-q} \tag{7·22}$$

with probability one. Since ϵ is arbitrary, (7·20) and (7·22) imply (2).

LEMMA 8. *If* $P \in p(\mathfrak{E})$, *then* (3) *holds with probability one.*

Assume first that $\mathfrak{E}$ has a boundary of measure zero and that p is a simple non-zero function; that is to say, p takes only a finite number of distinct values

https://doi.org/10.1017/S0305004100034095 Published online by Cambridge University Press

$p_j > 0$ $(j = 1, 2, \ldots, m)$. Since p is, by hypothesis, Lebesgue integrable, the set $\mathfrak{E}_j$, defined as the set of all points at which $p = p_j$, is a Lebesgue set, and

$$\sum_{j=1}^{m} \mathfrak{E}_j = \mathfrak{E}. \tag{8·1}$$

We also have

$$\int_{\mathfrak{E}} p\, dv = 1. \tag{8·2}$$

Write

$$\omega = \max_{1 \leqslant j \leqslant m} p_j. \tag{8·3}$$

We propose to prove by induction on m that with probability one

$$\liminf_{n \to \infty} n^{-q} l(\mathbf{P}^n) \geqslant \beta \int_{\mathfrak{E}} p^q\, dv. \tag{8·4}$$

This is certainly true when $m = 1$, by Lemma 7. Hence we take $m \geqslant 2$, and assume that (8·4) holds with probability one for any set $\mathfrak{E}$ with a boundary of measure zero and any simple non-zero function p which does not take more than $m - 1$ distinct values.

Since $\mathfrak{E}_1$ is a Lebesgue set, we can find a finite number of intervals whose union $\mathfrak{F}$ satisfies

$$v[(\mathfrak{E}_1 - \mathfrak{F}) + (\mathfrak{F} - \mathfrak{E}_1)] \leqslant \epsilon/2\omega, \tag{8·5}$$

for any prescribed $\epsilon > 0$. Now define a probability density function r on $\mathfrak{E}$ by means of

$$r = \begin{cases} \theta p & \text{in} \quad \mathfrak{E} - (\mathfrak{E}_1 + \mathfrak{F}), \\ \theta p_1 & \text{in} \quad \mathfrak{F}, \\ \theta p_2 & \text{in} \quad \mathfrak{E}_1 - \mathfrak{F}, \end{cases} \tag{8·6}$$

where θ is a constant, given accordingly by

$$\begin{aligned} 1 = \int_{\mathfrak{E}} r\, dv &= \theta \left\{ \int_{\mathfrak{E} - [(\mathfrak{E}_1 - \mathfrak{F}) + (\mathfrak{F} - \mathfrak{E}_1)]} p\, dv + \int_{\mathfrak{E}_1 - \mathfrak{F}} p_2\, dv + \int_{\mathfrak{F} - \mathfrak{E}_1} p_1\, dv \right\} \\ &= \theta \left\{ 1 + \int_{\mathfrak{E}_1 - \mathfrak{F}} (p_2 - p)\, dv + \int_{\mathfrak{F} - \mathfrak{E}_1} (p_1 - p)\, dv \right\}, \end{aligned} \tag{8·7}$$

by virtue of (8·2) and (8·6). From (8·3), (8·5) and (8·7),

$$|\theta^{-1} - 1| \leqslant \int_{\mathfrak{E}_1 - \mathfrak{F}} |p_2 - p|\, dv + \int_{\mathfrak{F} - \mathfrak{E}_1} |p_1 - p|\, dv \leqslant \tfrac{1}{2}\epsilon. \tag{8·8}$$

From the given random sequence $\mathbf{P} \in p(\mathfrak{E})$ we derive another sequence $\mathbf{R}$ in the following way. If $p = r$ almost everywhere in $\mathfrak{E}$, we simply take $\mathbf{R} = \mathbf{P}$. Otherwise, by (8·2) and (8·7), the set $\mathfrak{D}$ of all points, at which $r > p$, has positive measure. Let p^* be the probability density function on $\mathfrak{D}$ defined by

$$p^* = (r - p) \Big/ \int_{\mathfrak{D}} (r - p)\, dv. \tag{8·9}$$

Take a random sequence $\mathbf{Q} \in p^*(\mathfrak{D})$. If P_i lies in $\mathfrak{D}$, take $R_i = P_i$. If P_i lies in $\mathfrak{E} - \mathfrak{D}$, take (independently for each relevant value of i)

$$R_i = \begin{cases} P_i \text{ with probability } r(P_i)/p(P_i), \\ Q_i \text{ with probability } 1 - r(P_i)/p(P_i), \end{cases} \tag{8·10}$$

https://doi.org/10.1017/S0305004100034095 Published online by Cambridge University Press

where $r(P_i)$ and $p(P_i)$ are the values of r and p, respectively, at P_i. It is not hard to verify from (8·9) and (8·10) that the resulting sequence $\mathbf{R}$ satisfies $\mathbf{R} \in r(\mathfrak{E})$. From (8·10), (8·6), (8·7), (8·3), (8·5) and (8·8), together with the fact that $r > 0$, we deduce

$$\begin{aligned}
\text{prob}\,[R_i \neq P_i] &= \text{prob}\,[R_i = Q_i \mid P_i \in \mathfrak{E}-\mathfrak{D}]\,\text{prob}\,[P_i \in \mathfrak{E}-\mathfrak{D}] = \int_{\mathfrak{E}-\mathfrak{D}} \left[1-\frac{r}{p}\right] p\,dv \\
&= \int_{\{\mathfrak{E}-\mathfrak{D}\}\{\mathfrak{E}-[(\mathfrak{E}_1-\mathfrak{F})+(\mathfrak{F}-\mathfrak{E}_1)]\}} [(p/r)-1]r\,dv + \int_{\{\mathfrak{E}-\mathfrak{D}\}\{(\mathfrak{E}_1-\mathfrak{F})+(\mathfrak{F}-\mathfrak{E}_1)\}} \cdot (p-r)\,dv \\
&\leqslant \int_{\mathfrak{E}-[(\mathfrak{E}_1-\mathfrak{F})+(\mathfrak{F}-\mathfrak{E}_1)]} |(p/r)-1|\,r\,dv + \int_{(\mathfrak{E}_1-\mathfrak{F})+(\mathfrak{F}-\mathfrak{E}_1)} p\,dv \\
&= \int_{\mathfrak{E}-[(\mathfrak{E}_1-\mathfrak{F})+(\mathfrak{F}-\mathfrak{E}_1)]} |\theta^{-1}-1|\,r\,dv + \int_{(\mathfrak{E}-\mathfrak{F})+(\mathfrak{F}-\mathfrak{E}_1)} p\,dv \\
&\leqslant |\theta^{-1}-1| + \omega v[(\mathfrak{E}_1-\mathfrak{F})+(\mathfrak{F}-\mathfrak{E}_1)] \leqslant \epsilon. \qquad (8{\cdot}11)
\end{aligned}$$

Now (8·11) implies the existence of a number $n_0 = n_0(\epsilon)$ such that with probability at least $1-\epsilon$

$$N(\mathbf{R}^n - \mathbf{P}^n) \leqslant 2\epsilon n \quad (n \geqslant n_0). \qquad (8{\cdot}12)$$

By Lemma 1, Lemma 4, and (8·12),

$$\begin{aligned}
l(\mathbf{R}^n) &= l[\mathbf{P}^n + (\mathbf{R}^n - \mathbf{P}^n)] \leqslant l(\mathbf{P}^n) + l(\mathbf{R}^n - \mathbf{P}^n) + 2\delta(\mathfrak{E}) \\
&\leqslant l(\mathbf{P}^n) + \alpha(2\epsilon n)^q [v(\overline{\mathfrak{E}}) + \epsilon]^{1-q} + a(\mathfrak{E}, \epsilon) + 2\delta(\mathfrak{E}) \qquad (8{\cdot}13)
\end{aligned}$$

with probability at least $1-\epsilon$; and hence, with probability at least $1-\epsilon$,

$$\liminf_{n\to\infty} n^{-q} l(\mathbf{R}^n) \leqslant \{\liminf_{n\to\infty} n^{-q} l(\mathbf{P}^n)\} + \alpha(2\epsilon)^q [v(\mathfrak{E}) + \epsilon]^{1-q}. \qquad (8{\cdot}14)$$

Letting $\epsilon \to 0$ in (8·14), we deduce

$$\liminf_{n\to\infty} n^{-q} l(\mathbf{P}^n) \geqslant \liminf_{\epsilon\to 0} \liminf_{n\to\infty} n^{-q} l(\mathbf{R}^n) \qquad (8{\cdot}15)$$

with probability one.

According to Lemma 2 and Lemma 4, we have, for prescribed $\epsilon_0 > 0$,

$$\begin{aligned}
l(\mathbf{R}^n) &\geqslant l(\mathbf{R}^n\mathfrak{F}) + l[\mathbf{R}^n(\mathfrak{E}-\mathfrak{F})] - 2\sup_{\mathbf{Q}_1} l(\mathbf{Q}_1^{2n}\overline{\mathfrak{F}}\,\overline{\mathfrak{R}-\mathfrak{F}}) - 2\sup_{\mathbf{Q}_2} l[\mathbf{Q}_2^{2n}\overline{\mathfrak{E}-\mathfrak{F}}\,\overline{\mathfrak{R}-(\mathfrak{E}-\mathfrak{F})}] \\
&\geqslant l(\mathbf{R}^n\mathfrak{F}) + l[\mathbf{R}^n(\mathfrak{E}-\mathfrak{F})] - 2\alpha(2n)^q\{v(\overline{\mathfrak{F}}\,\overline{\mathfrak{R}-\mathfrak{F}}) + \epsilon_0\}^{1-q} \\
&\qquad - 2\alpha(2n)^q\{v[\overline{\mathfrak{E}-\mathfrak{F}}\,\overline{\mathfrak{R}-(\mathfrak{E}-\mathfrak{F})}] + \epsilon_0\}^{1-q} + O(1). \qquad (8{\cdot}16)
\end{aligned}$$

However, $\mathfrak{E}$ has a boundary of measure zero and $\mathfrak{F}$ is the sum of a finite number of intervals. Hence $\mathfrak{F}$ and $\mathfrak{E}-\mathfrak{F}$ have boundaries of measure zero, i.e.

$$v(\overline{\mathfrak{F}}\,\overline{\mathfrak{R}-\mathfrak{F}}) = v[\overline{\mathfrak{E}-\mathfrak{F}}\,\overline{\mathfrak{R}-(\mathfrak{E}-\mathfrak{F})}] = 0. \qquad (8{\cdot}17)$$

From (8·16) and (8·17), we have

$$\liminf_{n\to\infty} n^{-q} l(\mathbf{R}^n) \geqslant \liminf_{n\to\infty} \{n^{-q} l(\mathbf{R}^n\mathfrak{F}) + n^{-q} l[\mathbf{R}^n(\mathfrak{E}-\mathfrak{F})] - 2^{2+q}\alpha\epsilon_0^{1-q}\}. \qquad (8{\cdot}18)$$

Letting $\epsilon_0 \to 0$ in (8·18) we have

$$\liminf_{n\to\infty} n^{-q} l(\mathbf{R}^n) \geqslant \liminf_{n\to\infty} n^{-q} l(\mathbf{R}^n\mathfrak{F}) + \liminf_{n\to\infty} n^{-q} l[\mathbf{R}^n(\mathfrak{E}-\mathfrak{F})]. \qquad (8{\cdot}19)$$

https://doi.org/10.1017/S0305004100034095 Published online by Cambridge University Press

By (8·6), the function r is a non-zero simple function taking at most $m-1$ distinct values on $\mathfrak{E}-\mathfrak{F}$, whose boundary has measure zero. By the inductive assumption that (8·4) holds in these circumstances with probability one [i.e. by writing $\mathfrak{E}-\mathfrak{F}$ for $\mathfrak{E}$ in (8·4) and taking the p of (8·4) to the conditional distribution of r on $\mathfrak{E}-\mathfrak{F}$], we have with probability one

$$\liminf_{N[\mathbf{R}^n(\mathfrak{E}-\mathfrak{F})]\to\infty} \{N[\mathbf{R}^n(\mathfrak{E}-\mathfrak{F})]\}^{-q} l[\mathbf{R}^n(\mathfrak{E}-\mathfrak{F})] \geqslant \int_{\mathfrak{E}-\mathfrak{F}} \left\{\frac{r}{\int_{\mathfrak{E}-\mathfrak{F}} r\,dv}\right\}^q dv. \tag{8·20}$$

Similarly, r is constant on $\mathfrak{F}$; so that Lemma 7 gives

$$\liminf_{N(\mathbf{R}^n\mathfrak{F})\to\infty} [N(\mathbf{R}^n\mathfrak{F})]^{-q} l(\mathbf{R}^n\mathfrak{F}) \geqslant \beta\int_{\mathfrak{F}} \left\{\frac{r}{\int_{\mathfrak{F}} r\,dv}\right\}^q dv, \tag{8·21}$$

with probability one. The strong law of large numbers gives

$$\lim_{n\to\infty} N(\mathbf{R}^n\mathfrak{F})/n = \int_{\mathfrak{F}} r\,dv, \quad \lim_{n\to\infty} N[\mathbf{R}^n(\mathfrak{E}-\mathfrak{F})]/n = \int_{\mathfrak{E}-\mathfrak{F}} r\,dv \tag{8·22}$$

with probability one. Assembling (8·19), (8·20), (8·21), and (8·22), we deduce that

$$\liminf_{n\to\infty} n^{-q} l(\mathbf{R}^n) \geqslant \beta\int_{\mathfrak{E}} r^q\,dv \tag{8·23}$$

with probability one. However (8·5), (8·6) and (8·8) imply that $r \to p$ almost everywhere in $\mathfrak{E}$ as $\epsilon \to 0$. Moreover, p and r are bounded. Therefore (8·15) and (8·23) imply (8·4) with probability one. This closes the inductive cycle from $m-1$ to m, and establishes (8·4) for all finite m.

Next, relax the restriction that $\mathfrak{E}$ should have a boundary of zero measure. Let $n_j = N(\mathbf{P}^n\mathfrak{E}_j)$. The strong law of large numbers gives with probability one

$$\lim_{n\to\infty} n_j/n = p_j v(\mathfrak{E}_j). \tag{8·24}$$

From Lemma 1, we have $$l(\mathbf{P}^n) \leqslant \sum_{j=1}^{m} \{l(\mathbf{P}^n\mathfrak{E}_j) + \delta(\mathfrak{E})\}. \tag{8·25}$$

From Lemma 7, we have $$\lim_{n_j\to\infty} n_j^{-q} l(\mathbf{P}^n\mathfrak{E}_j) = \beta[v(\mathfrak{E}_j)]^{1-q} \tag{8·26}$$

with probability one. Taking (8·24), (8·25) and (8·26) together we have with probability one

$$\limsup_{n\to\infty} n^{-q} l(\mathbf{P}^n) \leqslant \sum_{j=1}^{m} [p_j v(\mathfrak{E}_j)]^q \beta[v(\mathfrak{E}_j)]^{1-q} = \beta\int_{\mathfrak{E}} p^q\,dv. \tag{8·27}$$

We may suppose without loss of generality that $\mathfrak{E} \subseteq \mathfrak{C}$. Let r be a simple non-zero probability density function on $\mathfrak{C}$ chosen so that p is the conditional probability density function of r in $\mathfrak{E}$. Let $\mathbf{Q} \in r(\mathfrak{C})$. Then $\mathbf{P} = \mathbf{Q}\mathfrak{E} \in p(\mathfrak{E})$. Let $\nu = N(\mathbf{Q}^n\mathfrak{E})$; so that $\mathbf{P}^\nu = \mathbf{Q}^n\mathfrak{E}$. The strong law of large numbers gives

$$\lim_{n\to\infty} \nu/n = \int_{\mathfrak{E}} r\,dv, \quad \lim_{n\to\infty} (n-\nu)/n = \int_{\mathfrak{C}-\mathfrak{E}} r\,dv \tag{8·28}$$

https://doi.org/10.1017/S0305004100034095 Published online by Cambridge University Press

with probability one. By Lemma 1,

$$l(\mathbf{P}^\nu) \geqslant l(\mathbf{Q}^n) - l(\mathbf{Q}^n - \mathbf{P}^\nu) - 2\delta(\mathfrak{C}). \tag{8·29}$$

Divide (8·29) by n^q and let $n \to \infty$. We get with probability one

$$\left(\int_{\mathfrak{C}} r\,dv\right)^q \liminf_{\nu\to\infty} \nu^{-q} l(\mathbf{P}^\nu) \geqslant \liminf_{n\to\infty} n^{-q} l(\mathbf{Q}^n) - \left(\int_{\mathfrak{C}-\mathfrak{E}} r\,dv\right)^q \limsup_{(n-\nu)\to\infty} (n-\nu)^{-q} l(\mathbf{Q}^n - \mathbf{P}^\nu)$$

$$\geqslant \beta\int_{\mathfrak{C}} r^q\,dv - \left(\int_{\mathfrak{C}-\mathfrak{E}} r\,dv\right)^q \beta\int_{\mathfrak{C}-\mathfrak{E}} \left(\frac{r}{\int_{\mathfrak{C}-\mathfrak{E}} r\,dv}\right)^q dv = \beta\int_{\mathfrak{E}} r^q\,dv \tag{8·30}$$

by (8·4) and (8·27), because $\mathfrak{C}$ has a boundary of zero measure. Hence with probability one

$$\liminf_{\nu\to\infty} \nu^{-q} l(\mathbf{P}^\nu) \geqslant \beta\int_{\mathfrak{E}} \left(\frac{r}{\int_{\mathfrak{E}} r\,dv}\right)^q dv = \beta\int_{\mathfrak{E}} p^q\,dv. \tag{8·31}$$

Finally, relax the condition that p should be a non-zero simple function. According to a standard device in the theory of Lebesgue integration (see, for example, (9) p. 115, Ex. 7) we can find a non-negative simple function f such that, for prescribed $\epsilon > 0$,

$$0 \leqslant f \leqslant p, \quad \int_{\mathfrak{E}} f\,dv \geqslant 1-\epsilon, \quad \int_{\mathfrak{E}} f^q\,dv \geqslant \int_{\mathfrak{E}} p^q\,dv - \epsilon. \tag{8·32}$$

We place the points P_i of $\mathbf{P} = P_1, P_2, \ldots$ in one of two classes ($\mathscr{C}_1$ and $\mathscr{C}_2$) by the following random procedure which deals with each value of i independently (in the stochastic sense). If $f(P_i) = 0$, we place P_i in $\mathscr{C}_1$. If $0 < f(P_i) \leqslant p(P_i)$ we place P_i in $\mathscr{C}_1$ with probability $1 - f(P_i)/p(P_i)$, and in $\mathscr{C}_2$ with probability $f(P_i)/p(P_i)$. It follows that the conditional distribution of each P_i, given that P_i belongs to $\mathscr{C}_2$, has a density function

$$f \Big/ \int_{\mathfrak{E}_0} f\,dv, \tag{8·33}$$

where $\mathfrak{E}_0$ is the set of all points in $\mathfrak{E}$ at which $f > 0$. Let n_1 and n_2 be the numbers of points from $\mathbf{P}^n$ which belong to $\mathscr{C}_1$ and $\mathscr{C}_2$, respectively, and let l_1 and l_2 be the corresponding tour-lengths. The strong law of large numbers and (8·32) give

$$\lim_{n\to\infty} (n_2/n) = \text{prob}\,[P_i \in \mathscr{C}_2] = \int_{\mathfrak{E}_0} \frac{f}{p} p\,dv = \int_{\mathfrak{E}_0} f\,dv = \int_{\mathfrak{E}} f\,dv \geqslant 1-\epsilon \tag{8·34}$$

with probability one. By Lemma 4,

$$\limsup_{n_1\to\infty} n_1^{-q} l_1 \leqslant \alpha v^{1-q}(\overline{\mathfrak{E}}); \tag{8·35}$$

and now, because $n_1 = n - n_2$, (8·34) gives with probability one

$$\limsup_{n\to\infty} n^{-q} l_1 \leqslant \alpha\epsilon^q v^{1-q}(\overline{\mathfrak{E}}). \tag{8·36}$$

https://doi.org/10.1017/S0305004100034095 Published online by Cambridge University Press

Since (8·33) exhibits a non-zero simple probability density function, (8·27) and (8·31) together give

$$\lim_{n_2\to\infty} n_2^{-q} l_2 = \beta \int_{\mathfrak{E}_0} \left(\frac{f}{\int_{\mathfrak{E}_0} f\, dv} \right)^q dv \tag{8·37}$$

with probability one; and then (8·34) gives

$$\lim_{n\to\infty} n^{-q} l_2 = \beta \int_{\mathfrak{E}_0} f^q\, dv = \beta \int_{\mathfrak{E}} f^q\, dv, \tag{8·38}$$

with probability one. However, by Lemma 1,

$$l(\mathbf{P}^n) \leqslant l_1 + l_2 + 2\delta(\mathfrak{E}). \tag{8·39}$$

Also, since we can obtain a closed path through the points of $\mathbf{P}^n$ in $\mathscr{C}_2$ by connecting together the tours through $\mathbf{P}^n$ and through the points of $\mathbf{P}^n$ in $\mathscr{C}_1$,

$$l_2 \leqslant l(\mathbf{P}^n) + l_1 + 2\delta(\mathfrak{E}). \tag{8·40}$$

From (8·36), (8·38), (8·39) and (8·32)

$$\begin{aligned} \limsup_{n\to\infty} n^{-q} l(\mathbf{P}^n) &\leqslant \alpha\epsilon^q v^{1-q}(\overline{\mathfrak{E}}) + \beta \int_{\mathfrak{E}} f^q\, dv \\ &\leqslant \alpha\epsilon^q v^{1-q}(\overline{\mathfrak{E}}) + \beta \int_{\mathfrak{E}} p^q\, dv \end{aligned} \tag{8·41}$$

with probability one. From (8·36), (8·38), (8·40) and (8·32)

$$\begin{aligned} \liminf_{n\to\infty} n^{-q} l(\mathbf{P}^n) &\geqslant \beta \int_{\mathfrak{E}} f^q\, dv - \alpha\epsilon^q v^{1-q}(\overline{\mathfrak{E}}) \\ &\geqslant \beta \int_{\mathfrak{E}} p^q\, dv - \beta\epsilon - \alpha\epsilon^q v^{1-q}(\overline{\mathfrak{E}}) \end{aligned} \tag{8·42}$$

with probability one. The required result follows from (8·41) and (8·42) on letting $\epsilon \to 0$.

LEMMA 9. *If the points of* $\mathbf{P}$ *are drawn independently from an arbitrary probability distribution over* $\mathfrak{E}$, *whose continuous component is represented almost everywhere by the density function* p, *then* (3) *holds with probability one.*

We shall prove Lemma 9 in the particular case $p = 0$. The general case is an immediate consequence of combining this particular case with the result of Lemma 8, using the arguments of the last part of the proof of Lemma 8 (especially two inequalities like (8·39) and (8·40)) to effect this combination.

When the distribution possesses discrete and singular components only, it is confined to a subset $\mathfrak{E}_0$ of $\mathfrak{E}$, where $v(\mathfrak{E}_0) = 0$. Given $\epsilon > 0$, we can find a closed set $\overline{\mathfrak{F}}$ such that $v(\overline{\mathfrak{F}}) < \epsilon$ and such that the total probability measure of $\mathfrak{E}_0 - \overline{\mathfrak{F}}$ does not exceed ϵ. Applying Lemma 4 to $\overline{\mathfrak{F}}$, the strong law of large numbers to $\mathfrak{E}_0 - \overline{\mathfrak{F}}$, and Lemma 1 to combine these two parts, we get

$$l(\mathbf{P}^n) \leqslant \alpha n^q [v(\overline{\mathfrak{F}}) + \epsilon]^{1-q} + a(\overline{\mathfrak{F}}, \epsilon) + \alpha(2\epsilon n)^q [v(\overline{\mathfrak{E}_0 - \overline{\mathfrak{F}}}) + \epsilon]^{1-q} + a(\mathfrak{E}_0 - \overline{\mathfrak{F}}, \epsilon) + 2\delta(\mathfrak{E}) \tag{9·1}$$

https://doi.org/10.1017/S0305004100034095 Published online by Cambridge University Press

with probability one for all sufficiently large n. Hence

$$\limsup_{n\to\infty} n^{-q} l(\mathbf{P}^n) \leqslant \alpha[v(\overline{\mathfrak{F}})+\epsilon]^{1-q} + \alpha(2\epsilon)^q[v(\overline{\mathfrak{E}_0-\mathfrak{F}})+\epsilon]^{1-q}$$
$$\leqslant \alpha(2\epsilon)^{1-q} + \alpha(2\epsilon)^q[v(\overline{\mathfrak{E}})+\epsilon]^{1-q} \tag{9·2}$$

with probability one. Since ϵ is arbitrary, the left-hand side of (9·2) must be zero, as required.

LEMMA 10. *$\beta_k \leqslant 6^{-\frac{1}{2}}12^{1/2k}$, and $\beta_2 \leqslant 0{\cdot}6508$.*

Suppose $\mathbf{P} \in W_1$ and consider $l(\mathbf{P}\xi\mathfrak{C})$. Since $N(\mathbf{P}\xi\mathfrak{C}) = \xi^k + o(\xi^k)$ as $\xi \to \infty$, with probability one, β is the limiting value (as $\xi \to \infty$) of the expected distance between successive points on $T(\mathbf{P}\xi\mathfrak{C})$, by virtue of Lemma 5. Now dissect $\xi\mathfrak{C}$ into $(\xi/\eta)^{k-1}$ disjoint hypercubical cylinders, as in the proof of Lemma 4, where η is the length of the side of the hypercubical cross-section. We shall keep η finite while $\xi \to \infty$. Consider the closed path which visits the points of $\mathbf{P}\xi\mathfrak{C}$ in the same order as the path constructed in Lemma 4, but which passes directly from point to point instead of going by way of the axis of a cylinder. The expected length of the join between the last point in one cylinder and the first point in the next is finite, and there are only $(\xi/\eta)^{k-1}-1$ such joins. These particular joins therefore contribute only a negligible amount to the total path length, which must be at least $\beta\xi^k + o(\xi^k)$ for large ξ by Lemma 5. Likewise the length of the return path from the end of the last cylinder to the beginning of the first, being less than $k^{\frac{1}{2}}\xi$, is also negligible. Thus

$$\beta_k k^{\frac{1}{2}} = \beta \leqslant \lambda, \tag{10·1}$$

where λ is the expected distance between two successive points in an infinite hypercubical cylinder whose cross-section has side η. Let x_i, x_i' $(i = 1, 2, \ldots, k-1)$ be coordinates of two successive points in their cross-sectional hyperplanes $(0 \leqslant x_i, x_i' \leqslant \eta)$ and let y be the distance between them measured along the axis of the cylinder. The probability that $y \geqslant y_0$ is the probability that a length y_0 of the cylinder contains no points of $\mathbf{P}$; i.e.

$$\text{prob}\,(y \geqslant y_0) = \exp\{-\eta^{k-1}y_0\}; \tag{10·2}$$

and hence

$$\lambda = \left\{\prod_{i=1}^{k-1}\int_0^\eta \frac{dx_i}{\eta}\int_0^\eta \frac{dx_i'}{\eta}\right\}\int_{y=0}^\infty d\{1-\exp(-\eta^{k-1}y)\}\left\{\sum_{i=1}^{k-1}(x_i-x_i')^2+y^2\right\}^{\frac{1}{2}}, \tag{10·3}$$

since x_i, x_i' are uniformly distributed between 0 and η. Since the expected value of a random variable does not exceed its root-mean-square expectation, (10·3) gives

$$\lambda^2 \leqslant \left\{\prod_{i=1}^{k-1}\int_0^\eta \frac{dx_i}{\eta}\int_0^\eta \frac{dx_i'}{\eta}\right\}\int_{y=0}^\infty d\{1-\exp(-\eta^{k-1}y)\}\left\{\sum_{i=1}^{k-1}(x_i-x_i')^2+y^2\right\} = \eta^2\left(\frac{k-1}{6}+\frac{2}{\eta^{2k}}\right). \tag{10·4}$$

When $\eta = 12^{1/2k}$, the right-hand side of (10·4) attains its minimum value, namely $12^{1/k}k/6$. Hence (10·1) and (10·4) yield $\beta_k \leqslant 6^{-\frac{1}{2}}12^{1/2k}$.

When $k = 2$, we can evaluate (10·3) in terms of the Struve and Neumann functions $\mathscr{S}_n(p)$ and $N_n(p)$ which satisfy

$$I_n(p) = \int_0^\infty e^{-pt}(1+t^2)^{n-\frac{1}{2}}\,dt = 2^{n-1}\pi^{\frac{1}{2}}(n-\tfrac{1}{2})!\,p^{-n}[\mathscr{S}_n(p)-N_n(p)]. \tag{10·5}$$

https://doi.org/10.1017/S0305004100034095 Published online by Cambridge University Press

Equation (10·5) is given in (6), p. 138, with another notation. The notation used in (10·5) conforms to that of (14), which we used as a source of numerical values. Setting $u = x_1/\eta$, $u' = x_1'/\eta$, $v = y/\eta$, $w = u - u'$, in (10·3), we have

$$\begin{aligned}
\lambda &= \eta^3 \int_0^1 du \int_0^1 du' \int_0^\infty dv\, e^{-\eta^2 v}[(u-u')^2 + v^2]^{\frac{1}{2}} \\
&= 2\eta^3 \int_0^1 dw(1-w) \int_0^\infty dv\, e^{-\eta^2 v}(w^2+v^2)^{\frac{1}{2}} \\
&= 2\eta \int_0^1 dw(1-w)\left\{w + \int_0^\infty dv\, e^{-\eta^2 v} v(w^2+v^2)^{-\frac{1}{2}}\right\} \\
&= \tfrac{1}{3}\eta + 2\eta \int_0^1 dw \int_0^\infty dv\, e^{-\eta^2 v} v(w^2+v^2)^{-\frac{1}{2}} - 2\eta \int_0^\infty dv\, e^{-\eta^2 v} v \int_0^1 dw\, w(w^2+v^2)^{-\frac{1}{2}} \\
&= \tfrac{1}{3}\eta + 2\eta \int_0^1 dw \int_0^\infty dz\, e^{-\eta^2 wz} wz(1+z^2)^{-\frac{1}{2}} - 2\eta \int_0^\infty dv\, e^{-\eta^2 v} v\{(1+v^2)^{\frac{1}{2}} - v\} \\
&= \tfrac{1}{3}\eta - 2\eta^{-1} \int_0^1 dw\, w \frac{\partial}{\partial w} \int_0^\infty dz\, e^{-\eta^2 wz}(1+z^2)^{-\frac{1}{2}} - 2\eta \int_0^\infty dv\, e^{-\eta^2 v} v(1+v^2)^{\frac{1}{2}} + 4\eta^{-5} \\
&= \tfrac{1}{3}\eta + 4\eta^{-5} - 2\eta^{-1} \int_{w=0}^1 w\, dI_0(\eta^2 w) - 2\eta^{-1} \int_0^\infty dv\, e^{-\eta^2 v}\{2(1+v^2)^{\frac{1}{2}} - (1+v^2)^{-\frac{1}{2}}\} \\
&= \tfrac{1}{3}\eta + 4\eta^{-5} - 2\eta^{-1} I_0(\eta^2) + 2\eta^{-1} \int_0^1 I_0(\eta^2 w)\, dw - 4\eta^{-1} I_1(\eta^2) + 2\eta^{-1} I_0(\eta^2) \\
&= \tfrac{1}{3}\eta + 4\eta^{-5} - 2\pi\eta^{-3}[\mathscr{S}_1(\eta^2) - N_1(\eta^2)] + \pi\eta^{-3} \int_0^{\eta^2} [\mathscr{S}_0(z) - N_0(z)]\, dz. \qquad (10{\cdot}6)
\end{aligned}$$

This is a convenient form for computation, since the last term is a running integral in η^2. Employing Simpson's formula with an interval of 0·01 in η^2 we have the following values of λ as a function of η^2:

η^2	2·0	2·5	2·9	3·0	3·1	3·5	4·0
λ	0·948	0·924	0·9218	0·9204	0·9223	0·924	0·934.

Hence $\beta_2 \leqslant 0{\cdot}9204/\sqrt{2} = 0{\cdot}6508$.

We are indebted to Dr D. R. Cox for bringing this problem to our notice, and to Mrs M. Swaffield of the Inland Telecommunications Department for drawing our attention to reference (13). One of the authors (J. E. B.) acknowledges the support of a postgraduate scholarship at St Hugh's, and another (J. H. H.) an Atomic Energy Authority Studentship.

REFERENCES

(1) BORŮVKA, O. On a minimal problem. *Práce Moravské Pridovedecké Spolecnosti* 3 (1926).
(2) COURANT, R. and ROBBINS, H. *What is mathematics?* (Oxford, 1941).
(3) COURANT, R. and SCHIFFER, M. *Dirichlet's principle, conformal mapping, and minimal surfaces.* (New York, 1950).
(4) DANTZIG, G., FULKERSON, R. and JOHNSON, S. Solution of a large-scale traveling salesman problem. *J. Oper. Res. Soc. Amer.* 2 (1954), 393–410.
(5) DOUGLAS, J. Minimal surfaces of higher topological structure. *Ann. Math., Princeton,* 40 (1939), 205–98.

https://doi.org/10.1017/S0305004100034095 Published online by Cambridge University Press

(6) Erdélyi, A., Magnus, W., Oberhettinger, F. and Tricomi, F. G. *Tables of integral transforms. Vol. I.* (New York, 1954).
(7) Flood, M. M. The traveling salesman problem. *J. Oper. Res. Soc. Amer.* 4 (1956), 61–75.
(8) Ghosh, H. W. Expected travel among random points. *Bull. Calcutta Statist. Ass.* 2 (1948), 83–7.
(9) Halmos, P. R. *Measure theory.* (New York, 1950).
(10) Hardy, G. H. *Divergent series.* (Oxford, 1949).
(11) Heller, I. On the traveling salesman's problem. Part I. Basic Facts. *George Washington Univ. Logistic Res. Project.* (1954).
(12) Heller, I. On the traveling salesman's problem. *Proc. Symp. Linear Programming.* (Washington, 1955), 643–65.
(13) Holmes, H. S. Telegraph hand delivery. *P.O. Telecomm. J.* 9 (1957), 66–71.
(14) Jahnke, E. and Emde, F. *Tables of functions with formulae and curves.* 4th ed. (New York, 1945).
(15) Jessen, R. J. Statistical investigation of a sample farm survey. *Bull. Iowa St. Coll. Agric. Res.* 304 (1942).
(16) Kruskal, J. B. On the shortest spanning subtree of a graph and the traveling salesman problem. *Proc. Amer. Math. Soc.* 7 (1956), 48–50.
(17) Loberman, H. and Weinberger, A. Formal processes for connecting terminals with a minimum total wire length. *J. Ass. Comp. Machy.* 4 (1957), 428–37.
(18) Mahalanobis, P. C. A sample survey of the acreage under jute in Bengal. *Sankhyā,* 4 (1940), 511–31.
(19) Marks, E. S. A lower bound for the expected travel among m random points. *Ann. Math. Statist.* 19 (1948), 419–22.
(20) Morton, G. and Land, A. H. A contribution to the travelling-salesman problem. *J. Roy. Statist. Soc.* (B), 17 (1955), 185–94.
(21) Motzkin, T. S. The assignment problem. *Proc. Symp. Appl. Math., Vol.* 6., *Numerical Analysis.* (New York, 1956), 109–25.
(22) *Oxford Atlas* (Revised ed.) (Oxford, 1952).
(23) Tompkins, C. Permutation problems. *Proc. Symp. Appl. Math., Vol.* 6. *Numerical Analysis* (New York, 1956), 195–211.
(24) Verblunsky, S. On the shortest path through a number of points. *Proc. Amer. Math. Soc.* (2) 6 (1951), 904–13.
(25) von Neumann, J. *Functional operators, Vol. I. Annals of Math. Studies,* 21 (Princeton, 1950).

St Hugh's College
Oxford

Balliol College
Oxford

Trinity College
Oxford

https://doi.org/10.1017/S0305004100034095 Published online by Cambridge University Press