

The Beardwood-Halton-Hammersley Theorem *Let $\{X_1, \dots, X_n\}$, $n \geq 1$, be a set of random variables in \mathbb{R}^d , independently and identically distributed with bounded support. Then the length L_n of a shortest TSP tour through the points X_i satisfies*

$$L_n/n^{(d-1)/d} \rightarrow \beta_d \int_{\mathbb{R}^d} f(x)^{(d-1)/d} dx, \quad \text{with probability 1, as } n \rightarrow \infty,$$

where $f(x)$ is the absolutely continuous part of the distribution of the X_i and β_d is a constant which depends on d but not on this distribution.

The easiest (2-dimensional) case of the theorem has the X_i uniformly distributed in the unit square: they have bounded support (the distribution is zero outside a bounded region), the integrand becomes $1^{1/2} = 1$ and the integral, being the volume of a unit cube, evaluates to 1. So the theorem says that a uniform distribution of n points in the unit square will, ‘almost surely’, be joined in a shortest Travelling Salesman Problem [TSP] circuit, or ‘tour’, whose length is asymptotic to $\beta_2 \sqrt{n}$, as $n \rightarrow \infty$. More generally, to any continuous random variable X there corresponds, by definition, a *probability density function*, f , whose integral over an interval is the probability that X takes a value in that interval. In the theorem, the ‘absolutely continuous part’ of the distribution of the X_i can be thought of, roughly, as a version of f which excludes ‘spikes’ in the distribution of the X_i , i.e., measure zero subsets of \mathbb{R}^n which occur with non-zero probability.

The well-known Concorde TSP-solving program (www.tsp.gatech.edu/concorde.html) is shown in the screen shot on the right being applied to a TSP problem based on the well-known logarithmic distribution of prime numbers: the probability that an integer n is prime being asymptotic to $1/\log n$. Specifically, 210^2 consecutive integers were arranged in rows of 210; and of these a total of 214 were found to be prime numbers and were marked as vertices (based on a depiction of the discovery of 10 consecutive primes in arithmetic progression: anthony.d.forbes.googlepages.com/10primes.htm). Concorde calculated an optimal tour with length 2287 units. Although these vertices certainly do not come from sampling a uniform distribution, we can still get some kind of an estimate for β_2 as though they were: $2287/(\sqrt{214} \times 210) \approx 0.74$. The best current estimates, based on simulations, range from 0.7119 to 0.7124.

Jillian Beardwood and John H. Halton were D.Phil. students of John M. Hammersley (1920–2004) at Oxford during the 1950s.

Web link: www.nap.edu/openbook.php?record_id=2026 (Chapter 8, by J.M. Steele)

Further reading: *The Traveling Salesman Problem: a Computational Study* by D.L. Applegate, R.E. Bixby, V. Chvátal and W.J. Cook, Princeton University Press, 2006. Chapter 1 can be previewed here: press.princeton.edu/chapters/s8451.pdf (2.44MB, but worth it for the pictures alone!).

