SIAM REVIEW
Vol. 37, No. 2, pp. 235–250, June 1995

PROBLEMS AND SOLUTIONS

EDITED BY CECIL C. ROUSSEAU (The University of Memphis)
AND OTTO G. RUEHR (Michigan Technological University)

All problems and solutions should be sent, typewritten in duplicate, with complete address, to Cecil C. Rousseau, Department of Mathematical Sciences, The University of Memphis, Memphis, Tennessee 38152 [E-mail: rousseac@hermes.msci.memphis.edu. Fax: (901) 678-2480]. *Preference is given to problems in applied mathematics, particularly those motivated by questions from science or industry. Whenever possible, a brief description of the problem background should be included. An asterisk placed beside a problem number indicates that the problem was submitted without solution. Proposers and solvers whose solutions are published will receive five reprints of the corresponding problem section. Other solvers will receive one reprint provided an address label is enclosed. Proposers and solvers desiring acknowledgment of their contributions should include a self-addressed stamped postcard or an e-mail address. (No stamps necessary for outside the U.S.A. and Canada.) Solutions should be received by September* 30, 1995.

PROBLEMS

Expected Wire Length Between Two Randomly Chosen Terminals

Problem 95-6, *by* DAVID M. LAZOFF (The Johns Hopkins Applied Physics Laboratory and the Computer Science Department, University of Maryland Baltimore County *and* ALAN T. SHERMAN (University of Maryland Baltimore County).

What is the expected Euclidean distance beween two independent, randomly chosen points uniformly distributed in an arbitrary rectangle? This problem arises in VLSI layout, analysis of rectangle heuristics for minimum-weighted Euclidean matchings, computation of the expected cost of random minimum-cost spanning trees, and in cognitive science in the analysis of experiments assessing human memory of spatial relations. Overwhelmed by algebraic difficulties of the problem, previous computer science researchers have resorted to special cases, asymptotic bounds, and numerical approximations.

1. Give exact, closed-form expressions for the first two moments of the distance in simple form as elementary functions of the two rectangle dimensions a and b.

2. Express the formulae from Part 1 in terms of the *aspect ratio* $r = a/b$ and *area* $A = ab$ of the rectangle. For any fixed aspect ratio r, what is the asymptotic behavior of the expected distance in terms of A? For any fixed area A, what is the asymptotic behavior of the expected distance in terms of r?

3. Simplify your answers to Part 1 for the special cases when the two points are uniformly distributed within a square or along a line segment.

An Integral Arising in Computing the Energy of Crystals

Problem 95-7, *by* M. L. GLASSER (Clarkson University).

In a study of the total electronic energy of crystals using the tight binding approximation [1] one encounters the integral

$$K(a) = \int_0^{\infty} [(e^{-x} + e^{x})^a - e^{-ax} - e^{ax}]dx,$$

Solution

Expected Wire Length Between Two Randomly Chosen Terminals

Problem 95-6, *by* DAVID M. LAZOFF (The Johns Hopkins Applied Physics Laboratory and the Computer Science Department, University of Maryland Baltimore County) *and* ALAN T. SHERMAN (University of Maryland Baltimore County).

What is the expected Euclidean distance beween two independent, randomly chosen points uniformly distributed in an arbitrary rectangle? This problem arises in VLSI layout, the analysis of rectangle heuristics for minimum-weighted Euclidean matchings, the computation of the expected cost of random minimum-cost spanning trees, and in cognitive science in the analysis of experiments assessing human memory of spatial relations. Overwhelmed by algebraic difficulties of the problem, previous computer science researchers have resorted to special cases, asymptotic bounds, and numerical approximations.

1. Give exact, closed-form expressions for the first two moments of the distance in simple form as elementary functions of the two rectangle dimensions a and b.

2. Express the formulae from Part 1 in terms of the *aspect ratio* $r = a/b$ and *area* $A = ab$ of the rectangle. For any fixed aspect ratio r what is the asymptotic behavior of the expected distance in terms of A? For any fixed area A what is the asymptotic behavior of the expected distance in terms of r?

3. Simplify your answers to Part 1 for the special cases when the two points are uniformly distributed within a square or along a line segment.

Solution by W. BOEHM (University of Economics, Vienna, Austria).

It is sufficient to consider the generic rectangle with sides having lengths t and 1, where t is an arbitrary positive real number, because the uniform distribution has a scale invariance property: if X is uniform in $(0, 1)$ and Y is uniform in $(0, t)$, then Y has the same distribution as tX. So general rectangles may always be reduced to generic ones by choosing an appropriate scaling constant $\lambda > 0$ and t, such that the sides have lengths λt and λ. Observe that for the generic rectangle both aspect ratio and area are equal to t.

If D_λ denotes the distance of two randomly chosen points in the rectangle with sides λt and λ, then by scale invariance it follows immediately that

$$E(D_\lambda) = \lambda E(D) \quad \text{and} \quad E(D_\lambda^2) = \lambda^2 E(D^2),$$

where D is the distance of two points in the generic rectangle.

Now let the points chosen be (X_0, Y_0) and (X_1, Y_1), where X_0, X_1 are uniform in $(0, t)$ and Y_0, Y_1 are uniform in $(0, 1)$. Furthermore, let $D_X = |X_1 - X_0|$ and $D_Y = |Y_1 - Y_0|$, and denote the densities of D_X and D_Y by f and g, respectively. These densities are well known (see, e.g., [1, p. 22]) and are given by

$$f(x) = \frac{2}{t^2}(t - x), \qquad g(x) = 2(1 - y), \tag{1}$$

where $0 \le x \le t$ and $0 \le y \le 1$.

From (1) we get the answer to the first part of Point 3:

$$E(D_X) = \int_0^t \frac{2}{t^2}(t - x)x\,dx = \frac{t}{3} \tag{2}$$

and

$$E(D_X^2) = \int_0^t \frac{2}{t^2}(t - x)x^2\,dx = \frac{t^2}{6}. \tag{3}$$

Now consider the conditional distribution of $D = \sqrt{D_X^2 + D_Y^2}$, given $D_X = x$. This distribution is

$$P(D \le z \mid D_X = x) = 2\int_0^{\sqrt{z^2-x^2}} (1-y)\,dy = 2\left[\sqrt{z^2-x^2} - \frac{1}{2}(z^2-x^2)\right].$$

Differentiating with respect to z yields the conditional density

$$h(z|x) = 2\left(\frac{z}{\sqrt{z^2-x^2}} - z\right).$$

If we multiply $h(z|x)$ by z and integrate with respect to z on the interval $(x, \sqrt{1+x^2})$, we obtain the conditional expectation of D:

$$\text{(4)} \qquad E(D|D_X = x) = 2\int_x^{\sqrt{1+x^2}} \left[\frac{z^2}{\sqrt{z^2-x^2}} - z^2\right] dz$$

$$= \frac{1}{3}\sqrt{1+x^2}(1-2x^2) + x^2 \ln\left(\frac{1}{x} + \sqrt{1+\frac{1}{x^2}}\right) + \frac{2}{3}x^3.$$

If we multiply (4) by the density of D_X, as given in (1), and integrate with respect to x over $(0, t)$, then we get the unconditional expectation of D,

$$\text{(5)} \qquad E(D) = \int_0^t \frac{2}{t^2}(t-x)E(D \mid D_X = x)\,dx$$

$$= \frac{\sinh^{-1} t}{6t} + \frac{t^2 \sinh^{-1}(1/t)}{6} - \frac{\sqrt{1+t^2}(t^4 - 3t^2 + 1)}{15t^2} + \frac{1}{15t^2} + \frac{t^3}{15},$$

which is the desired formula for generic rectangles. All integrals occurring in (5) are, by the way, more or less elementary or at least standard ones.

If we set $t = 1$, then we get $E(D)$ for the unit square:

$$\text{(6)} \qquad E(D) = \frac{1}{3}\ln(1+\sqrt{2}) + \frac{1}{15}(2+\sqrt{2}) = 0.5214\ldots .$$

Multiply (6) by $\lambda > 0$ to obtain the corresponding result for an arbitrary square with side length λ.

The second moment of D is easy; since $D^2 = D_X^2 + D_Y^2$, we have from (3)

$$\text{(7)} \qquad E(D^2) = E(D_X^2) + E(D_Y^2) = \frac{1}{6}(1+t^2).$$

Now consider an arbitrary rectangle with aspect ratio t and sides having lengths λt and λ. Its area equals $A = \lambda^2 t$, and therefore the mean distance expressed in terms of the aspect ratio is

$$\text{(8)} \qquad E(D_\lambda) = \lambda E(D) = \sqrt{\frac{A}{t}}\,E(D).$$

Thus for fixed t and large A, we have

$$E(D_\lambda) = O(\sqrt{A}).$$

Now we fix A and let t grow large. Using the fact that

$$\sinh^{-1}(1/t) \sim \frac{1}{t} \quad \text{and} \quad \sqrt{1+t^2} \sim t + \frac{1}{2t},$$

we find from (5) that

$$E(D) \sim \frac{t}{3},$$

which is not surprising since λ must tend to zero, and as a result, asymptotically, the rectangle collapses to the line segment $(0, t)$. However, in this case necessarily

$$E(D) \to E(D_X) = \frac{t}{3}.$$

Thus for large t and fixed A,

$$E(D_\lambda) \sim \frac{\sqrt{At}}{3}. \tag{9}$$

REFERENCE

[1] W. FELLER, *An Introduction to Probability Theory and Its Applications*, Vol. 2, 2nd ed., John Wiley and Sons, New York, 1971.

Also solved by CARL C. GROSJEAN (University of Ghent, Belgium), W. WESTON MEYER (General Motors R&D Center, Warren, MI), *and the proposers.*

Editorial note. W. W. MEYER obtains a general formula for the moments of D. His result is

$$E(D^n) = 4\sum_{k=0}^{\lfloor n/2\rfloor} \binom{\frac{n}{2}}{k}\binom{\frac{n-1}{2}}{k}^{-1} \frac{(a^{2k}+b^{2k})d^{n-2k}}{(n+1)(n+2)(n+3)} + \frac{4}{(ab)^2}\cdot\frac{a^{n+4}+b^{n+4}-d^{n+4}}{(n+2)(n+3)(n+4)}$$
$$+ \frac{4|\sin(n\pi/2)|}{(n+2)(n+3)}\binom{\frac{n}{2}}{-\frac{1}{2}}\left[a^{n+1}b^{-1}\ln\left(\frac{d+b}{a}\right) + b^{n+1}a^{-1}\ln\left(\frac{d+a}{b}\right)\right],$$

where $d = \sqrt{a^2+b^2}$.

As pointed out by J. BOERSMA (Eindhoven University of Technology, The Netherlands), the first four moments are contained in a 1951 paper of Ghosh [5], and this paper is referenced as part of the published solution of Problem 75-12 in this journal [6]. The authors knew of Ghosh's work. In a discussion section of their problem proposal, they note that the results in question have been rediscovered by various workers:

> "The first solution for an arbitrary rectangle may be due to Ghosh [4, 5], with relevant previous work dating back to Crofton's second formula of 1885 [2]. Over the past century, several people have rediscovered the solution to this and related problems. For example, as part of his 1972 thesis, Ehlers [3] rediscovered the formula for the case of an arbitrary rectangle, and in 1976 Alagar [1] computed the expected distance when the points fall in two adjacent rectangles. We can now add our names to the list."

In submitting their proposal to *Problems and Solutions*, the authors sought to elicit more elegant and/or general solutions and provide an accessible source for the moment formulas and references to previous work. The editors agreed with these objectives. [C.C.R.]

REFERENCES

[1] V. S. ALAGAR, *The distribution of the distance between random points*, J. Appl. Probab., 13 (1976), pp. 558–566.
[2] M. CROFTON, *Probability*, in Encyclopedia Britannica, 9th ed., 19 (1885), pp. 768–788.
[3] P. EHLERS AND E. G. ENNS, *Random secants of a convex body generated by randomness*, J. Appl. Probab., 18 (1981), pp. 157–166.
[4] B. GHOSH, *On the distribution of random distances in a rectangle*, Science and Culture, 8 (1943), p. 388.
[5] ———, *Random distances within a rectangle and between two rectangles*, Bull. Calcutta Math. Soc., 43 (1951), pp. 17–24.
[6] J. D. MURCHLAND, *Comment on Problem* 75-12, SIAM Rev., 18 (1976), p. 498.

An Integral Arising in Computing the Energy of Crystals

Problem 95-7, *by* M. L. GLASSER (Clarkson University).

In a study of the total electronic energy of crystals using the tight binding approximation [1], one encounters the integral

$$K(a) = \int_0^\infty [(e^{-x} + e^{x})^a - e^{-ax} - e^{ax}]dx,$$

and by numerical integration Harrison obtained the value $K(5/3) \approx 4.45$. Evaluate $K(a)$ in terms of Gamma functions.

REFERENCE

[1] W. HARRISON, *Total energies in the tight-binding theory*, Phys. Rev., B23 (1981), pp. 5230–5245.

Solution by MOURAD E. H. ISMAIL (University of South Florida, Tampa, FL).

From the large x behavior of $(e^x + e^{-x})^a - e^{ax}$ it is clear that the integral converges only for $0 < a < 2$. When $0 < a < 2$, the binomial theorem gives

$$\begin{aligned} K(a) &= \int_0^\infty \left[e^{ax} \sum_{n=1}^{\infty} \frac{(-a)_n}{n!} e^{-2nx}(-1)^n - e^{-ax} \right] dx \\ &= \sum_{n=0}^{\infty} \frac{(-a)_n}{n!} \frac{(-1)^n}{2n - a} \\ &= -\frac{1}{a}\, {}_2F_1\left(-a, -\frac{a}{2}; 1 - \frac{a}{2}; -1\right) = -\frac{\Gamma^2(1 - a/2)}{a\Gamma(1 - a)} \end{aligned}$$

by Theorem 26, p. 68 in Rainville's book on special functions.

Also solved by J. BOERSMA (Eindhoven University of Technology, The Netherlands), PAUL BRACKEN (University of Waterloo), DAVID BRADLEY (Simon Fraser University), R. G. BUSHMAN (University of Wyoming), ROBIN CHAPMAN (University of Exeter, UK), CARL C. GROSJEAN (University of Ghent, Belgium), W. B. JORDAN (Scotia, NY), JOHN C. MALVIDO (Interet Corporation, Millburn, NJ), W. WESTON MEYER (General Motors R&D Center, Warren, MI), ALLEN R. MILLER (Washington, DC), MICHAEL RENARDY (Virginia Tech), BILLY SPRATT (University of Paisley, Scotland), PETER WAGNER (University of Innsbruck, Austria), JAMES A. WILSON (Iowa State University), PETER N. ZHEVANDROV (Universidad Michoacana, Mexico), *and the proposer*.

Several solvers noted that $K(5/3) \approx 4.626291112$.

A Conditional Trace Inequality

Problem 95-8*, *by* PAUL A. ROEDIGER (U.S. Army Armament Research, Picatinny Arsenal, NJ).

Our submission

(wirelength)

Submitted to SIAM Review

An Exact Formula for the Expected Wire Length Between Two Randomly Chosen Terminals

David M. Lazoff[1] and Alan T. Sherman
Computer Science Department
University of Maryland Baltimore County
Baltimore, Maryland 21228-5398
email: lazoff@umbc8.umbc.edu, sherman@cs.umbc.edu

August 2, 1994

Keywords. Combinatorial problems, computational geometry, computer aided design, geometric probability, geometric Steiner tree problem, layout algorithms, minimum-cost spanning trees, very large scale integration (VLSI).

Problem

What is the expected Euclidean distance between two independent, randomly-chosen points uniformly distributed in an arbitrary rectangle? This problem arises in VLSI layout, analysis of rectangle heuristics for minimum weighted Euclidean matchings, computing the expected cost of random minimum-cost spanning trees, and in cognitive science in the analysis of experiments assessing human memory of spatial relations. Overwhelmed by algebraic difficulties of the problem, previous computer science researchers have resorted to special cases, asymptotic bounds, and numerical approximations.

1) Give exact, closed-form expressions for the first two moments of the distance in simple form as elementary functions of the two rectangle dimensions a and b.

2) Reexpress the formulae from Part 1 in terms of the *aspect* ratio $r = a/b$ and *area* $A = ab$ of the rectangle. For any fixed aspect ratio r, what is the asymptotic behavior of the expected distance in terms of A? For any fixed area A, what is the asymptotic behavior of the expected distance in terms of r?

3) Simplify your answers to Part 1 for the special cases when the two points are uniformly distributed within a square or along a line segment.

[1]Member, Applied Physics Laboratory, The Johns Hopkins University.

Solution

Let the random variable D denote the Euclidean distance between two independently-chosen points, uniformly distributed over a rectangle of dimensions $a > 0$ and $b > 0$.

1 The First Two Moments of D

Using elementary calculus, we compute closed-form expressions for the first two moments of D in terms of elementary functions of a and b. By the definition of expected value, and because the two points are uniformly distributed, the expected value of D can be expressed as the quadruple integral

$$E[D] = \frac{1}{a^2b^2} \int_0^b \int_0^b \int_0^a \int_0^a \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}\, dx_1\, dx_2\, dy_1\, dy_2. \tag{1}$$

In Equation 1, each integral corresponds to one of the four coordinates needed to describe the two points. The difficulty of the problem arises from the square root in the integrand.

After applying the variable substitutions $u = x_1 - x_2$ and $v = y_1 - y_2$, we solve the resulting simpler quadruple integral, one integral at a time. In this calculation, the six integrals $\int x^k \sqrt{x^2 + a^2}$ with $0 \le k \le 3$ and $\int \ln(x + \sqrt{x^2 + a^2})$ and $\int x^2 \ln(a + \sqrt{x^2 + a^2})$ account for most of the calculus. Since the amount of algebra is extensive, we leave the detailed calculations to Appendix B of our technical report [10]. As we expected, the resulting formula is symmetrical in a and b:

$$E[D] = \frac{a^5 + b^5 - (a^4 - 3a^2b^2 + b^4)\sqrt{a^2 + b^2}}{15a^2b^2} + \frac{a^2}{6b} \ln\left(\frac{b + \sqrt{a^2 + b^2}}{a}\right) + \frac{b^2}{6a} \ln\left(\frac{a + \sqrt{a^2 + b^2}}{b}\right). \tag{2}$$

By evaluating Equation 2, practitioners can solve instances of the problem exactly, rather than resorting to upper and lower bounds or to numerical approximations.

Because squaring the integrand in Equation 1 eliminates the square root, an easy calculation determines the second moment of D to be

$$E[D^2] = \frac{a^2 + b^2}{6}. \tag{3}$$

2 Solution Expressed in Terms of Aspect Ratio and Area

To interpret our solution further, we express $E[D]$ in terms of the *aspect ratio* $r = a/b$ and *area* $A = ab$ of the rectangle. Substituting $a = \sqrt{Ar}$ and $b = \sqrt{A/r}$ into Equations 2 and 3, and using the inverse hyperbolic substitution $\sinh^{-1} r = \ln(r + \sqrt{r^2 + 1})$, we obtain

$$E[D] = \sqrt{Ar} \left[\frac{3\sqrt{1 + r^{-2}} + r(r - \sqrt{r^2 + 1}) + r^{-3}(1 - \sqrt{r^2 + 1})}{15} + \frac{r}{6} \sinh^{-1}(1/r) + \frac{r^{-2}}{6} \sinh^{-1} r \right] \tag{4}$$

and

$$E[D^2] = \frac{A}{6}(r + r^{-1}). \tag{5}$$

Thus, for any fixed aspect ratio, $E[D]$ grows linearly with the square root of the area. Conversely, for any fixed area, $E[D] \in \Theta(\sqrt{r})$, because $\lim_{r\to\infty} E[D]/\sqrt{r} = \sqrt{A}[(3-0.5+0)/15+(1/6)+0] = \sqrt{A}/3$. This asymptotic behavior can be seen in Figure 1, which shows a 3-dimensional graph of Equation 4 produced by *Maple* [2].

3 Special Cases: Square and Line Segment

We simplify Equations 2 and 3 for the special cases when the points are uniformly distributed within a square or along a line segment. Let D_s denote the random variable D when the rectangle is a square, and let D_l denote D for the degenerate case when the two points are uniformly distributed along a line segment of length a. Substituting $b = a$ in Equations 2 and 3 yields

$$E[D_s] = a\left[\frac{2+\sqrt{2}}{15} + \frac{\ln(1+\sqrt{2})}{3}\right] \approx 0.521405a \qquad \text{and} \qquad E[D_s^2] = \frac{a^2}{3}, \tag{6}$$

which agrees with Gilbert's calculation [7, p. 387]. A simple integration of polynomials yields

$$E[D_l] = \frac{1}{a^2}\int_0^a\int_0^a |x_1 - x_2|\, dx_2\, dx_1 = \frac{a}{3} \qquad \text{and} \qquad E[D_l^2] = \frac{a^2}{6}. \tag{7}$$

As a final check of Equation 2, we verified that $\lim_{b\to 0} E[D] = E[D_l]$ and $\lim_{b\to 0} E[D^2] = E[D_l^2]$.

4 Monte Carlo Simulations

To check our solution, we ran Monte Carlo simulations and compared the resulting sample mean and standard deviations of D with the corresponding exact values given by Equations 2 and 3. We implemented our simulations in a straightforward fashion and ran them on a Silicon Graphics workstation, using L'Ecuyer's [13, p. 282] pseudorandom number generator with the Bays-Durham shuffle.

Table 1 summarizes the results of some of our simulations. In these simulations, we worked with rectangles of constant area 1 with selected aspect ratios $1 \leq a/b \leq 1024$. Using 10^6 trials per rectangle, the empirical values agree closely with their corresponding theoretical values; moreover, the extent of this agreement increased with the number of trials. Also, as we expected, the standard deviations in Table 1 become large for long, narrow rectangles. For additional simulations, see [10].

5 Discussion

Our interest in this problem grew out of a need to compute the expected wire length between two randomly-chosen terminals on a VLSI chip. Only after deriving Equation 2 did we learn of the 1985 work of Sheng [15], who used the advanced Borel overlap technique to solve this and other related problems. In particular, Equation 2 follows after a page of routine algebra from Sheng's Equation 2.5.1, which also yields all higher moments of the distance. The first solution for an arbitary rectangle may be due to Ghosh [5, 6] in 1943, with relevant previous work dating back to Crofton's [3] second formula of 1885. Over the past century several people have rediscovered the solution to this and related problems. For example, as part of his 1972 thesis, Ehlers [4] rediscovered the formula for the case of an arbitrary rectangle, and in 1976 Alagar [1] computed the expected distance when the points fall in two adjacent rectangles. We can now add our names to the list.

Despite the intuitive nature, apparent simplicity, and wide applicability of the problem, the general case has eluded exact analytical solution by many researchers. For example, in 1965, motivated by the challenge of computing the expected cost of minimum-cost Euclidean spanning trees, Gilbert [7] rediscovered the special case of our problem for the unit square. Similarly, in their 1983 analysis of the ∞-version rectangle heuristic for Minimum Weighted Euclidean Matchings, Reingold and Supowit [14, p. 53] solved a variation only for the $\sqrt{2} \times 1$ rectangle, in which the two points are required to lie on opposite sides of the rectangle. And in his 1990 calculation of the quantization error coefficient of a 2-dimensional lattice for the Euclidean Traveling Salesman Problem [8, pp. 72–74], Goddyn [8] solved a much simpler variation for which one point is fixed at the lower-left corner of a unit square. Apparently, none of these researchers were aware of the relevant work by Ghosh and others. Several researchers have studied the more general problem of computing the expected costs of random minimum-cost spanning trees on $n > 2$ vertices, but we are aware only of experimental or asymptotic bounds (for example, see [7, 9, 11]).

The difficulty of the problem is not conceptual—it is simple to express the solution as a multiple integral and to sketch a high-level plan for solving the integral. Rather, the difficulty lies solely in the excruciatingly toilsome algebra necessary to solve the integral to produce a usable formula. For example, in his 1993 probability textbook, Pitman [12, Prob. 21, p. 356] considers the problem on the unit square and challenges the reader to bound its solution from above and below, admonishing that even for the unit square, the calculation "is hard to do exactly by calculus." It would be natural to solve the integral with a symbolic math package, such as *Macsyma*, *Maple*, or *Mathematica*. Although we did use these packages to verify some of our intermediate calculations, none of these tools were able to solve Equation 1 without extensive hints from our manual solution. Meticulously applying elementary techniques, we relentlessly overcame all algebraic difficulties.

a/b	a	b	Exact D	Empirical D
1	1.000000	1.000000	0.521405 ± 0.247931	0.521598 ± 0.247950
2	1.414214	0.707107	0.569060 ± 0.304693	0.569337 ± 0.304542
4	2.000000	0.500000	0.713743 ± 0.445987	0.712539 ± 0.445795
8	2.828427	0.353553	0.964207 ± 0.651514	0.963945 ± 0.651531
16	4.000000	0.250000	1.342640 ± 0.935094	1.342688 ± 0.934585
32	5.656854	0.176777	1.889535 ± 1.329736	1.888086 ± 1.328148
64	8.000000	0.125000	2.668275 ± 1.884032	2.667324 ± 1.884120
128	11.313708	0.088388	3.771884 ± 2.665994	3.771421 ± 2.665872
256	16.000000	0.062500	5.333591 ± 3.770959	5.340682 ± 3.772555
512	22.627417	0.044194	7.542573 ± 5.333221	7.529635 ± 5.331104
1024	32.000000	0.031250	10.666706 ± 7.542428	10.646558 ± 7.537217

Table 1: Results of Monte Carlo simulations for unit-area rectangles of dimensions a and b, for selected aspect ratios $1 \leq a/b \leq 1024$. Values of D are listed as means $\pm$ standard deviations. Exact values are computed from Equations 2 and 3; empirical values are based on 10^6 trials for each rectangle.

6 Open Problem

We conclude by posing the following related open problem: Derive an exact, closed-form expression for the expected cost of minimum-cost spanning trees (respectively, minimum Steiner trees) connecting n randomly chosen points in an arbitrary rectangle, for small values of $n > 2$. For $n = 3$, we conjecture that this problem is solvable by case analysis.

Acknowledgments

We thank Peter Ehlers for pointing out the prior work of Ghosh [5, 6] and Sheng [15], and we thank Konstantinos Kalpakis for helpful comments.

References

[1] Alagar, Vangalur S., "The distribution of the distance between random points," *Journal of Applied Probability*, **13**:3 (September 1976), 558–566.

[2] Char, Bruce W.; K. Geddes; G. Gonnet; B. Leong; M. Monagan; and S. Watt, *First Leaves: A Tutorial Introduction to Maple V*, Springer-Verlag (1992).

[3] Crofton, M., "Probability" in *Encyclopaedia Britannica*, 9th edition, **19** (1885), 768–788.

[4] Ehlers, P. Frank; and Ernest G. Enns, "Random secants of a convex body generated by randomness," *Journal of Applied Probability*, **18**:1 (March 1981), 157–166.

[5] Ghosh, B., "On the distribution of random distances in a rectangle," *Science and Culture*, **8** (1943), 388.

[6] Ghosh, B., "Random distances within a rectangle and between two rectangles," *Bulletin of the Calcutta Mathematical Society*, **43** (1951), 17–24.

[7] Gilbert, E. N., "Random minimal trees," *J. SIAM*, **13**:2 (June 1965), 376–387.

[8] Goddyn, Luis A., "Quantizers and the worst-case Euclidean Traveling Salesman Problem," *Journal of Combinatorial Theory*, Series B, **50**:1 (October 1990), 65–81.

[9] Jain, Anjani; and John W. Mamer, "Approximations for the random minimal spanning tree with application to network provisioning," *Operations Research*, **36**:4 (July-August 1988), 575–584.

[10] Lazoff, David M.; and Alan T. Sherman, "An exact formula for the expected wire length between two randomly chosen terminals," Technical Report TR CS-94-08, Computer Science Department, University of Maryland Baltimore County (July 20, 1994). 14 pages.

[11] Lueker, George S., "Optimization problems on graphs with independent random edge weights," *SIAM Journal on Computing*, **10**:2 (May 1981), 338–351.

[12] Pitman, Jim, *Probability*, Springer-Verlag (New York, 1993).

[13] Press, William H.; Saul A. Teukolsky; William T. Vetterling; and Brian P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*, 2nd edition, Cambridge University Press (Cambridge, 1992).

[14] Reingold, Edward M.; and Kenneth P. Supowit, "Probabilistic analysis of divide-and-conquer heuristics for minimum weighted Euclidean matching," *Networks*, **13**:1 (spring 1983), 49–66.

[15] Sheng, T. K., "The distance between two random points in plane regions," *Advances in Applied Probability*, **17**:4 (December 1985), 748–773.

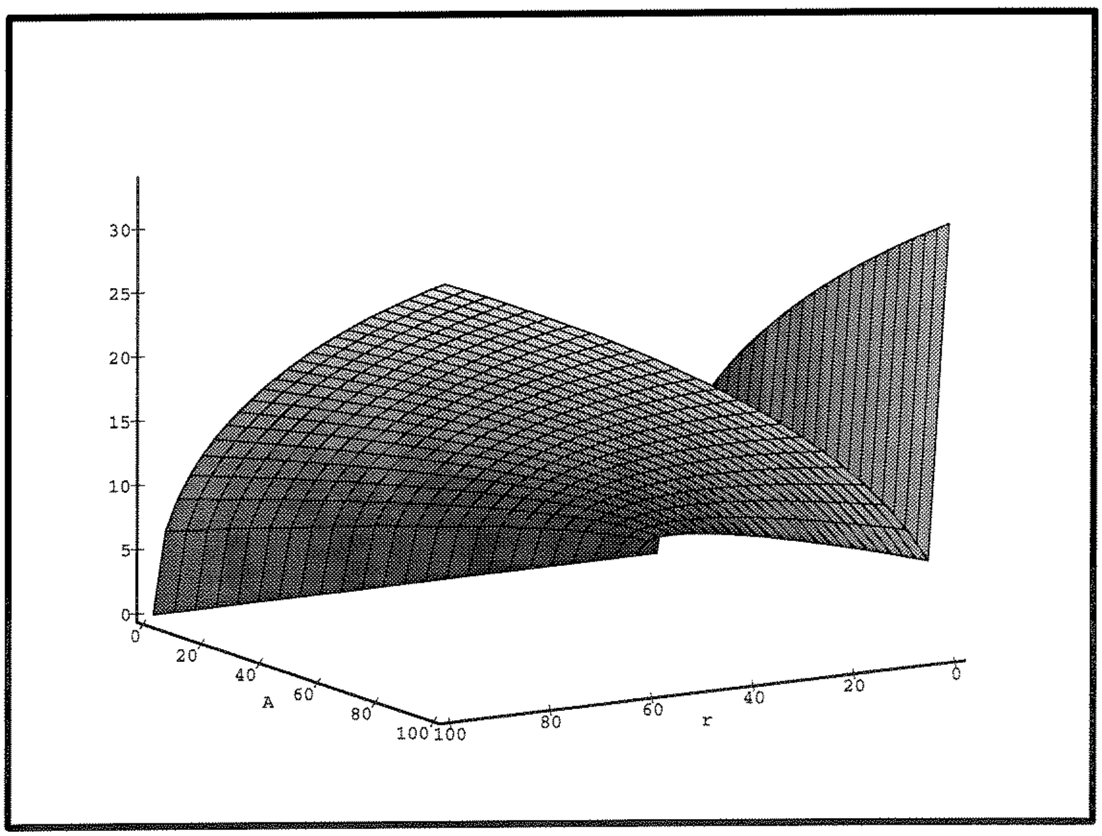

Figure 1: Expected value of D for rectangles of dimensions a and b, as function of aspect ratio $r = a/b$ and rectangle area $A = ab$.