TR CS-94-08

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

David M. Lazoff* 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

July 20, 1994

Abstract

We give a formula for the expected Euclidean distance between two randomly-chosen points uniformly distributed in an arbitrary rectangle. This problem arises in VLSI layout, analysis of rectangle heuristics for minimum weighted Euclidean matchings, and computing the expected cost of random minimum-cost spanning trees. Overwhelmed by algebraic difficulties of the problem, previous researchers have resorted to special cases, asymptotic bounds, and numerical approximations. Using elementary techniques, we overcome these difficulties to derive an exact closed-form solution involving square roots, natural logarithms, and rational functions of the two rectangle dimensions.

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

*Member, Applied Physics Laboratory, The Johns Hopkins University.

1 Introduction

The *Two-Terminal Wire Length Problem (WLP_2)* is to compute the expected Euclidean distance between two randomly-chosen points uniformly distributed in an arbitrary rectangle. This problem arises in VLSI layout design: it characterizes the average wire length needed to connect two randomly chosen terminals on a VLSI chip. This problem also arises as the simplest ($n = 2$) case of the difficult problems of computing the expected cost of minimum-cost Euclidean spanning trees, minimum Euclidean Steiner trees, and traveling salesman paths, connecting n randomly chosen vertices in an arbitrary rectangle. Variations of WLP_2 also arise in the analysis of the ∞-version rectangle heuristic for Minimum Weighted Euclidean Matchings [10, p. 53] and in calculating the quantization error coefficient of a 2-dimensional lattice for the Euclidean Traveling Salesman Problem [4, pp. 72–74]. In addition, WLP_2 arises in cognitive science in the analysis of experiments assessing human memory of spatial relations,[1] and in military simulations of air defense suppression and aircraft survivability [5].

Despite the intuitive nature, apparent simplicity, and wide applicability of WLP_2, the general case of this problem has eluded exact analytical solution. For example, in 1965, Gilbert [3] solved only the special case of WLP_2 for the unit square. Similarly, in 1983, Reingold and Supowit [10] solved a variation of WLP_2 only for the $\sqrt{2}\times 1$ rectangle, in which the two points are required to lie on opposite sides of the rectangle. In 1990, Goddyn [4] analytically solved a much simpler variation of WLP_2 for which one point is fixed at the lower-left corner of a unit square. In his probability textbook, Pitman [8, Prob. 21, p. 356] considers WLP_2 on the unit square and challenges the reader to bound its solution from above and below, admonishing that even for the unit square, WLP_2 "is hard to do exactly by calculus." Several researchers have studied the more general problem of computing the expected costs of random minimum-cost spanning trees, but we are aware only of experimental or asymptotic bounds (for example, see [3, 6, 7]).

Our main result is the following closed-form solution to WLP_2

$$\frac{a^5+b^5-(a^4-3a^2b^2+b^4)\sqrt{a^2+b^2}}{15a^2b^2}+\frac{a^2}{6b}\ln\left(\frac{b+\sqrt{a^2+b^2}}{a}\right)+\frac{b^2}{6a}\ln\left(\frac{a+\sqrt{a^2+b^2}}{b}\right), \qquad (1)$$

where a and b are the rectangle dimensions. By evaluating Equation 1, practitioners can now solve instances of WLP_2 exactly, rather than resorting to upper and lower bounds or to numerical approximations.

The difficulty of WLP_2 is not conceptual—it is simple to understand the problem; it is straightforward to express the solution as a multiple integral; and it is easy to outline a high-level plan for solving the integral. Rather, the difficulty of WLP_2 lies solely in the excruciatingly toilsome algebra necessary to solve the integral to produce a usable formula. Simply stated, the extraordinary *grunge* of WLP_2 has thwarted all previous attempts to solve the problem. 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 the integral without extensive hints from our manual solution. Meticulously applying elementary techniques, we relentlessly overcame all algebraic difficulties.

The rest of this paper is organized in five sections. Section 2 precisely states Problem WLP_2. Section 3 outlines our solution of WLP_2. Section 4 verifies our solution through computer simula-

[1] David Uttal, Department of Psychology, Northwestern University, private correspondence (July 1994).

tions. Section 5 restates the solution in terms of aspect ratio and area, and analyzes two special cases. Section 6 summarizes our conclusions and states some related open problems. In addition, Appendix A lists several fundamental integrals used in our solution; Appendix B gives a detailed proof of our main result; and Appendix C presents graphs of the solution.

2 Problem WLP_2

Let a and b be any positive real constants and let $R = ([0,a]\times[0,b])\cap\mathbb{R}^2$ be a rectangle of dimensions a and b. Let P and Q be independent random variables, each with uniform distribution over R, and let the random variable $D = d(P,Q)$ be the Euclidean distance between P and Q. Thus, if $P = (x_1, y_1)$ and $Q = (x_2, y_2)$, then $d(P,Q) = \sqrt{(x_1-x_2)^2 + (y_1-y_2)^2}$. The *Two-Terminal Wire Length Problem (WLP_2)* is to compute the expected value of D.

3 Exact Analytical Solution

Using elementary calculus, we characterize the distribution of D by computing exact closed-form expressions for its first two moments.

3.1 Expected Value of D

By the definition of expected value, and because the points P and Q are uniformly distributed over R, the expected value of D can be expressed as the quadruple integral

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

In Equation 2, each integral corresponds to one of the four coordinates needed to describe points P and Q. The difficulty of WLP_2 arises from the square root in the integrand.

Applying the variable substitutions $u = x_1 - x_2$ and $v = y_1 - y_2$, Equation 2 simplifies to

$$E[D] = \frac{1}{a^2b^2}\int_0^b\int_{-y_2}^{b-y_2}\int_0^a\int_{-x_2}^{a-x_2} \sqrt{u^2+v^2}\,du\,dx_2\,dv\,dy_2. \tag{3}$$

Since $0 \le x_1 \le a$, the new variable $u = x_1 - x_2$ in Equation 3 ranges from $-x_2$ to $a - x_2$; similarly, v ranges from $-y_2$ to $b - y_2$. Also, because $\partial u/\partial x_1 = \partial v/\partial y_1 = 1$, no additional constant factors are introduced.

Although more elegant methods may be possible, for simplicity, we solve Equation 3 in the most straightforward fashion, one integral at a time. Thus, $E[D] = I_4/(a^2b^2)$, where

$$I_1 = \int_{-x_2}^{a-x_2}\sqrt{u^2+v^2}\,du, \qquad I_2 = \int_0^a I_1\,dx_2, \qquad I_3 = \int_{-y_2}^{b-y_2} I_2\,dv, \qquad \text{and} \qquad I_4 = \int_0^b I_3\,dy_2. \tag{4}$$

To solve each of these four integrals, we apply several fundamental integrals given in Appendix A. In particular, 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. Also, it is helpful to observe that $(x^2+y^2)^{3/2} = x^2\sqrt{x^2+y^2} + y^2\sqrt{x^2+y^2}$. Because the amount of algebra is extensive, we state the results of these four integrations and leave the detailed calculations to Appendix B:

$$I_1 = \frac{1}{2}(a-x_2)\sqrt{(a-x_2)^2+v^2} + \frac{1}{2}v^2\ln(a-x_2+\sqrt{(a-x_2)^2+v^2}) + \frac{1}{2}x_2\sqrt{x_2^2+v^2} - \frac{1}{2}v^2\ln(-x_2+\sqrt{x_2^2+v^2}), \quad (5)$$

$$I_2 = \frac{2}{3}v^3 + \frac{1}{3}a^2\sqrt{a^2+v^2} - \frac{2}{3}v^2\sqrt{a^2+v^2} + \frac{1}{2}av^2\ln(a+\sqrt{a^2+v^2}) - \frac{1}{2}av^2\ln(-a+\sqrt{a^2+v^2}), \quad (6)$$

$$\begin{aligned} I_3 = {} & \frac{1}{6}y_2^4 + \frac{1}{6}t^4 + \frac{1}{4}a^2y_2\sqrt{a^2+y_2^2} - \frac{1}{4}a^2t\sqrt{a^2+t^2} - \frac{1}{6}y_2^3\sqrt{a^2+y_2^2} - \frac{1}{6}t^3\sqrt{a^2+t^2} \\ & - \frac{1}{12}a^4\ln(-y_2+\sqrt{a^2+y_2^2}) + \frac{1}{12}a^4\ln(t+\sqrt{a^2+t^2}) + \frac{1}{6}ay_2^3\ln(a+\sqrt{a^2+y_2^2}) \\ & + \frac{1}{6}at^3\ln(a+\sqrt{a^2+t^2}) - \frac{1}{6}ay_2^3\ln(-a+\sqrt{a^2+y_2^2}) - \frac{1}{6}at^3\ln(-a+\sqrt{a^2+t^2}), \quad (7) \end{aligned}$$

where $t = b - y_2$, and

$$I_4 = \frac{a^5+b^5-(a^4-3a^2b^2+b^4)\sqrt{a^2+b^2}}{15} + \frac{a^4b}{6}\ln\left(\frac{b+\sqrt{a^2+b^2}}{a}\right) + \frac{ab^4}{6}\ln\left(\frac{a+\sqrt{a^2+b^2}}{b}\right). \quad (8)$$

Dividing Equation 8 by a^2b^2 yields $E[D]$, given in Equation 1. As we expected, our solution is symmetrical in a and b.

3.2 Second Moment of D

Because squaring the integrand in Equation 2 eliminates the square root, an easy calculation determines the second moment of D to be

$$E[D^2] = \frac{a^2+b^2}{6}. \quad (9)$$

4 Monte Carlo Simulations

To check our solution, we ran two Monte Carlo simulations and compared the resulting sample mean and standard deviations of D with the corresponding exact values given by Equations 1 and 9. We implemented our simulations in a straightforward fashion and ran them on a Silicon Graphics workstation. For both simulations, we used L'Ecuyer's [9, p. 282] pseudorandom number generator with the Bays-Durham shuffle; we also achieved similar results with the Irix pseudorandom number generator "random()".

Tables 1 and 2 summarize the results of our simulations. In Simulation I, we worked with rectangles with all possible integral dimensions $1 \le a \le b \le 5$. In Simulation II, we worked with rectangles of constant area 1 with selected aspect ratios $1 \le a/b \le 1024$. Using 10^6 trials per rectangle, the empirical values agree closely with their corresponding theoretical values (through at least two decimal places in Simulation I); moreover, the extent of this agreement increased with the number of trials. Also, as we expected, the standard deviations in Table 2 become large for long, narrow rectangles. Thus, our simulations confirm our theoretical results.

a	b	Exact D	Empirical D	a	b	Exact D	Empirical D
1	1	0.521405 ± 0.247931	0.521598 ± 0.247949	3	3	1.564216 ± 0.743793	1.561784 ± 0.743461
1	2	0.804772 ± 0.430901	0.804837 ± 0.430419	3	4	1.834592 ± 0.894952	1.834641 ± 0.894863
1	3	1.111272 ± 0.657070	1.110910 ± 0.657219	3	5	2.119541 ± 1.083611	2.118021 ± 1.083225
1	4	1.427486 ± 0.891974	1.427213 ± 0.892606	4	4	2.085622 ± 0.991723	2.084713 ± 0.991904
1	5	1.748797 ± 1.129177	1.749467 ± 1.128266	4	5	2.353845 ± 1.136990	2.351359 ± 1.136787
2	2	1.042811 ± 0.495862	1.042117 ± 0.495223	5	5	2.607027 ± 1.239654	2.608112 ± 1.240330
2	3	1.317067 ± 0.657267	1.316955 ± 0.657041	$\sqrt{2}$	1	0.634206 ± 0.312702	0.634355 ± 0.312645
2	4	1.609544 ± 0.861802	1.610139 ± 0.862019	$\sqrt{2}$	$\sqrt{2}$	0.737379 ± 0.350627	0.737652 ± 0.350652
2	5	1.912644 ± 1.084033	1.915311 ± 1.085424	$\sqrt{2}$	2	0.896903 ± 0.442228	0.897351 ± 0.442397

Table 1: Results of Monte Carlo Simulation I for selected rectangle dimensions a and b. Values of D are listed as means ± standard deviations. Exact values are computed from Equations 1 and 9; empirical values are based on 10^6 trials for each rectangle.

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 2: Results of Monte Carlo Simulation II 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 ± standard deviations. Exact values are computed from Equations 1 and 9; empirical values are based on 10^6 trials for each rectangle.

5 Alternate Form and Special Cases

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. In addition, we simplify our formula for the special cases when the points are uniformly distributed within a square or along a line segment.

5.1 Solution Expressed in Terms of Aspect Ratio and Area

Substituting $a = \sqrt{Ar}$ and $b = \sqrt{A/r}$ into Equations 1 and 9 yields the symmetrical formulae

$$E[D] = \sqrt{A}\left[\frac{r^{5/2} + r^{-5/2} - (r^2 - 3 + r^{-2})\sqrt{r + r^{-1}}}{15} + \frac{r^{3/2}}{6}\ln(r^{-1} + \sqrt{r^{-2}+1}) + \frac{r^{-3/2}}{6}\ln(r + \sqrt{r^2+1})\right] \tag{10}$$

and

$$E[D^2] = \frac{A}{6}(r + r^{-1}). \tag{11}$$

Figure 1 shows a 3-dimensional graph of Equation 10 produced by *Maple* [2].

Using the inverse hyperbolic substitution $\sinh^{-1} r = \ln(r + \sqrt{r^2+1})$, and rearranging the numerator terms in the first fraction, Equation 10 can be reexpressed more intuitively as

$$E[D] = \sqrt{Ar}\left[\frac{3\sqrt{1 + r^{-2}} + r(r - \sqrt{r^2+1}) + r^{-3}(1 - \sqrt{r^2+1})}{15} + \frac{r}{6}\sinh^{-1}(1/r) + \frac{r^{-2}}{6}\sinh^{-1} r\right]. \tag{12}$$

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 $\lim_{r\to\infty} E[D]/\sqrt{r} = \sqrt{A}[(3-0.5+0)/15+(1/6)+0] = \sqrt{A}/3$.

5.2 Special Case: Square

Let D_s denote D when the rectangle is a square. Substituting $b = a$ in Equations 1 and 9 yields

$$E[D_s] = a\left[\frac{2+\sqrt{2}}{15} + \frac{\ln(1+\sqrt{2})}{3}\right] \approx 0.521405a \qquad \text{and} \qquad E[D_s^2] = \frac{a^2}{3}, \tag{13}$$

which agrees with Gilbert's formula [3, p. 387].

5.3 Special Case: Line Segment

We now consider the degenerate case when the two points are uniformly distributed along a line segment of length a. Let D_l denote the random variable D for this case. A simple integration of polynomials yields

$$E[D_l] = \frac{1}{a^2}\int_0^a\int_0^a |x_1 - x_2|\, dx_2\, dx_1 = \frac{a}{3} \qquad \text{and} \qquad E[D_l^2] = \frac{a^2}{6}. \tag{14}$$

As a final check of Equation 1, we verified that $\lim_{b\to 0} E[D] = E[D_l]$ and $\lim_{b\to 0} E[D^2] = E[D_l^2]$.

6 Conclusion and Open Problems

We have given an exact closed-form solution to the general case of WLP_2. Our solution has applications in a variety of geometric settings drawn from VLSI design, computational geometry, analysis of algorithms, and cognitive science.

We conclude by stating two related open problems: 1) Can the extensive algebraic manipulations of our solution be circumvented by using more sophisticated methods? 2) 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 second problem is solvable by case analysis.

Acknowledgment

We thank Konstantinos Kalpakis for helpful comments.

References

[1] Beyer, William H., ed., *CRC Standard Mathematical Tables and Formulae, 29th Edition*, CRC Press (Boca Raton, 1991).

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

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

[5] Hoeber, Francis P., *Military Applications of Modeling: Selected Case Studies*, Gordon and Breach Science Publishers (New York, 1982).

[6] 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.

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

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

[9] 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).

[10] 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.

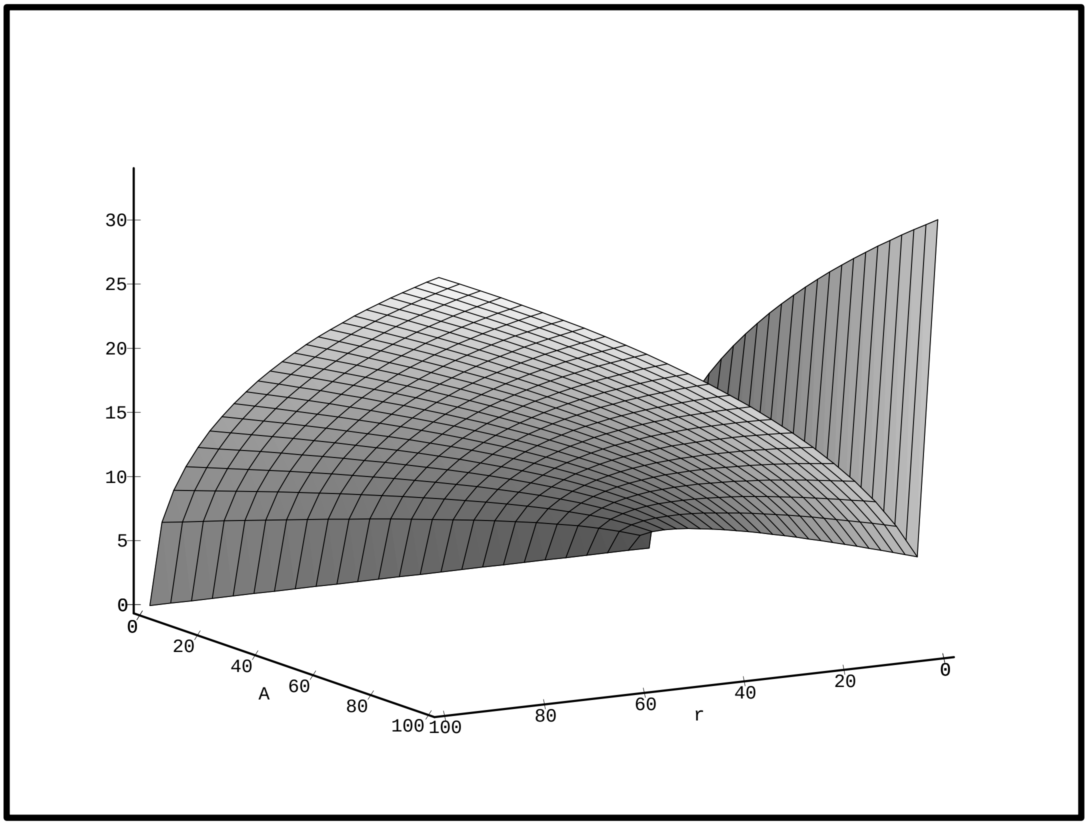

Figure 1: Expected value of D for rectangles of dimensions a and b, as function of aspect ratio $r = a/b$ and rectangle area $A = ab$.

Appendix A: Fundamental Integrals and Identities

This appendix lists six fundamental indefinite integrals and one algebraic identity that we apply to derive Equation 1. The first five integrals are given in Beyer [1]; we also point out how they can be computed. Throughout, let $a \in \mathbb{R}$ be any constant.

Lemma 1 *(by the trigonometric substitution $x = a\tan\theta$ [1, No. 156, p. 230])*

$$\int \sqrt{x^2+a^2}\,dx = \frac{1}{2}\left[x\sqrt{x^2+a^2} + a^2\ln(x+\sqrt{x^2+a^2})\right]. \tag{15}$$

Lemma 2 *(by Lemma 1 and integration by parts [1, No. 163, p. 230])*

$$\int x\sqrt{x^2+a^2}\,dx = \frac{1}{3}(x^2+a^2)^{3/2}. \tag{16}$$

Lemma 3 *(by Lemma 1 and integration by parts [1, No. 168, p. 231])*

$$\int x^2\sqrt{x^2+a^2}\,dx = \frac{x}{4}(x^2+a^2)^{3/2} - \frac{a^2}{8}x\sqrt{x^2+a^2} - \frac{a^4}{8}\ln(x+\sqrt{x^2+a^2}). \tag{17}$$

Lemma 4 *(by Lemma 1 and integration by parts [1, No. 169, p. 231])*

$$\int x^3\sqrt{x^2+a^2}\,dx = (\frac{1}{5}x^2 - \frac{2}{15}a^2)(x^2+a^2)^{3/2}. \tag{18}$$

Lemma 5 *(by the trigonometric substitution $x = a\tan\theta$ [1, No. 511, p. 258])*

$$\int \ln(x+\sqrt{x^2+a^2})\,dx = x\ln(x+\sqrt{x^2+a^2}) - \sqrt{x^2+a^2}. \tag{19}$$

Lemma 6 *(by the substitution $x = a\tan\theta$ and integration by parts, or by Mathematica)*

$$\int x^2\ln(a+\sqrt{x^2+a^2})\,dx = -\frac{x^3}{9} + \frac{ax\sqrt{x^2+a^2}}{6} + \frac{x^3\ln(a+\sqrt{x^2+a^2})}{3} - \frac{a^3\ln(x+\sqrt{x^2+a^2})}{6}. \tag{20}$$

Lemma 7 *(immediate)*

$$(x^2+y^2)^{3/2} = x^2\sqrt{x^2+y^2} + y^2\sqrt{x^2+y^2}. \tag{21}$$

Appendix B: Derivation of General Solution

To fill in the details of our derivation of Equation 1, we apply the lemmas from Appendix A to compute each of the four key integrals I_1, I_2, I_3, I_4 defined in Section 3. For clarity we omit much of the intermediate algebra.

B.1 First Integral

By Lemma 1,

$$\begin{aligned} I_1 &= \int_{-x_2}^{a-x_2} \sqrt{u^2+v^2}\,du \\ &= \frac{1}{2}(a-x_2)\sqrt{(a-x_2)^2+v^2} + \frac{1}{2}v^2 \ln(a-x_2+\sqrt{(a-x_2)^2+v^2}) \\ &\quad + \frac{1}{2}x_2\sqrt{x_2^2+v^2} - \frac{1}{2}v^2\ln(-x_2+\sqrt{x_2^2+v^2}). \end{aligned} \tag{22}$$

B.2 Second Integral

We compute the second integral $I_2 = \int_0^a I_1\,dx_2$ by separately computing each of its four summands $I_2 = I_{2_1} + I_{2_2} + I_{2_3} + I_{2_4}$. By Lemma 2 and the variable substitution $w = a - x_2$,

$$\begin{aligned} I_{2_1} &= \frac{1}{2}\int_0^a (a-x_2)\sqrt{(a-x_2)^2+v^2}\,dx_2 \\ &= \frac{1}{6}(a^2+v^2)^{\frac{3}{2}} - \frac{1}{6}v^3. \end{aligned} \tag{23}$$

Similarly, by Lemma 5,

$$\begin{aligned} I_{2_2} &= \frac{1}{2}v^2\int_0^a \ln(a-x_2+\sqrt{(a-x_2)^2+v^2})\,dx_2 \\ &= \frac{1}{2}av^2\ln(a+\sqrt{a^2+v^2}) - \frac{1}{2}v^2\sqrt{a^2+v^2} + \frac{1}{2}v^3. \end{aligned} \tag{24}$$

Again, using Lemmas 2 and 5,

$$I_{2_3} = \frac{1}{2}\int_0^a x_2\sqrt{x_2^2+v^2}\,dx_2 = \frac{1}{6}(a^2+v^2)^{\frac{3}{2}} - \frac{1}{6}v^3 \tag{25}$$

and

$$\begin{aligned} I_{2_4} &= -\frac{1}{2}v^2\int_0^a \ln(-x_2+\sqrt{x_2^2+v^2})\,dx_2 \\ &= -\frac{1}{2}av^2\ln(-a+\sqrt{a^2+v^2}) - \frac{1}{2}v^2\sqrt{a^2+v^2} + \frac{1}{2}v^3. \end{aligned} \tag{26}$$

Combining terms and applying Lemma 7 yields

$$\begin{aligned} I_2 &= I_{2_1} + I_{2_2} + I_{2_3} + I_{2_4} \\ &= \frac{2}{3}v^3 + \frac{1}{3}a^2\sqrt{a^2+v^2} - \frac{2}{3}v^2\sqrt{a^2+v^2} \\ &\quad + \frac{1}{2}av^2\ln(a+\sqrt{a^2+v^2}) - \frac{1}{2}av^2\ln(-a+\sqrt{a^2+v^2}). \end{aligned} \tag{27}$$

B.3 Third Integral

We compute $I_3 = \int_{-y_2}^{b-y_2} I_2\,dv$ by separately computing each of its five summands $I_3 = I_{3_1} + I_{3_2} + I_{3_3} + I_{3_4} + I_{3_5}$. Throughout, let $t = b - y_2$. The first summand is the polynomial

$$I_{3_1} = \int_{-y_2}^{t} \frac{2}{3}v^3\,dv = \frac{1}{6}t^4 + \frac{1}{6}y_2^4. \tag{28}$$

By Lemma 1,

$$\begin{aligned} I_{3_2} &= \int_{-y_2}^{t} \frac{1}{3}a^2\sqrt{a^2+v^2}\,dv \\ &= \frac{1}{6}a^2t\sqrt{a^2+t^2} + \frac{1}{6}a^4\ln(t+\sqrt{a^2+t^2}) + \frac{1}{6}a^2y_2\sqrt{a^2+y_2^2} - \frac{1}{6}a^4\ln(-y_2+\sqrt{a^2+y_2^2}). \end{aligned} \tag{29}$$

By Lemma 3,

$$\begin{aligned} I_{3_3} &= \int_{-y_2}^{t} -\frac{2}{3}v^2\sqrt{a^2+v^2}\,dv \\ &= -\frac{1}{6}t(a^2+t^2)^{\frac{3}{2}} + \frac{1}{12}a^2t\sqrt{a^2+t^2} + \frac{1}{12}a^4\ln(t+\sqrt{a^2+t^2}) \\ &\quad - \frac{1}{6}y_2(a^2+y_2^2)^{\frac{3}{2}} + \frac{1}{12}a^2y_2\sqrt{a^2+y_2^2} - \frac{1}{12}a^4\ln(-y_2+\sqrt{a^2+y_2^2}), \end{aligned} \tag{30}$$

which by Lemma 7 yields

$$\begin{aligned} I_{3_3} &= -\frac{1}{6}t^3\sqrt{a^2+t^2} - \frac{1}{12}a^2t\sqrt{a^2+t^2} + \frac{1}{12}a^4\ln(t+\sqrt{a^2+t^2}) \\ &\quad - \frac{1}{6}y_2^3\sqrt{a^2+y_2^2} - \frac{1}{12}a^2y_2\sqrt{a^2+y_2^2} - \frac{1}{12}a^4\ln(-y_2+\sqrt{a^2+y_2^2}). \end{aligned} \tag{31}$$

Next, using Lemma 6,

$$\begin{aligned} I_{3_4} &= \int_{-y_2}^{t} \frac{1}{2}av^2\ln(a+\sqrt{a^2+v^2})\,dv \\ &= -\frac{1}{18}at^3 + \frac{1}{12}a^2t\sqrt{a^2+t^2} + \frac{1}{6}at^3\ln(a+\sqrt{a^2+t^2}) - \frac{1}{12}a^4\ln(t+\sqrt{a^2+t^2}) \\ &\quad - \frac{1}{18}ay_2^3 + \frac{1}{12}a^2y_2\sqrt{a^2+y_2^2} + \frac{1}{6}ay_2^3\ln(a+\sqrt{a^2+y_2^2}) + \frac{1}{12}a^4\ln(-y_2+\sqrt{a^2+y_2^2}). \end{aligned} \tag{32}$$

Since $I_{3_5} = -aI_{3_4}$, it follows that

$$
\begin{aligned}
I_{3_5} &= \frac{1}{18}at^3 + \frac{1}{12}a^2t\sqrt{a^2+t^2} - \frac{1}{6}at^3\ln(-a+\sqrt{a^2+t^2}) \\
&\quad - \frac{1}{12}a^4\ln(t+\sqrt{a^2+t^2}) + \frac{1}{18}ay_2^3 + \frac{1}{12}a^2y_2\sqrt{a^2+y_2^2} \\
&\quad + \frac{1}{6}ay_2^3\ln(-a+\sqrt{a^2+y_2^2}) + \frac{1}{12}a^4\ln(-y_2+\sqrt{a^2+y_2^2}). \qquad (33)
\end{aligned}
$$

Combining terms, we obtain

$$
\begin{aligned}
I_3 &= I_{3_1} + I_{3_2} + I_{3_3} + I_{3_4} + I_{3_5} \\
&= \frac{1}{6}y_2^4 + \frac{1}{6}t^4 + \frac{1}{4}a^2y_2\sqrt{a^2+y_2^2} - \frac{1}{4}a^2t\sqrt{a^2+t^2} - \frac{1}{6}y_2^3\sqrt{a^2+y_2^2} - \frac{1}{6}t^3\sqrt{a^2+t^2} \\
&\quad - \frac{1}{12}a^4\ln(-y_2+\sqrt{a^2+y_2^2}) + \frac{1}{12}a^4\ln(t+\sqrt{a^2+t^2}) + \frac{1}{6}ay_2^3\ln(a+\sqrt{a^2+y_2^2}) \\
&\quad + \frac{1}{6}at^3\ln(a+\sqrt{a^2+t^2}) - \frac{1}{6}ay_2^3\ln(-a+\sqrt{a^2+y_2^2}) - \frac{1}{6}at^3\ln(-a+\sqrt{a^2+t^2}). \qquad (34)
\end{aligned}
$$

B.4 Fourth Integral

Finally, we compute $I_4 = \int_0^b I_3\,dy_2$ by separately computing its twelve summands $I_4 = I_{4_1} + I_{4_2} + \cdots + I_{4_{12}}$, five of which are duplicates. Let $t = b - y_2$. The first summand is the polynomial

$$
I_{4_1} = \int_0^b \frac{1}{6}y_2^4\,dy_2 = \frac{1}{30}b^5; \qquad (35)
$$

similarly, $I_{4_2} \int_0^b t^4/6\,dt = I_{4_1}$.

By Lemma 2,

$$
I_{4_3} = \int_0^b \frac{1}{4}a^2y_2\sqrt{a^2+y_2^2}\,dy_2 = \frac{1}{12}a^2(a^2+b^2)^{3/2} - \frac{1}{12}a^5, \qquad (36)
$$

and $I_{4_4} = \int_0^b (a^2t/4)\sqrt{a^2+t^2}\,dt = I_{4_3}$.

By Lemma 4,

$$
I_{4_5} = \int_0^b -\frac{1}{6}y_2^3\sqrt{a^2+y_2^2}\,dy_2 - \frac{1}{30}b^2(a^2+b^2)^{3/2} + \frac{1}{45}a^2(a^2+b^2)^{3/2} - \frac{1}{45}a^5, \qquad (37)
$$

and $I_{4_6} = \int_0^b -(t^3/6)\sqrt{a^2+t^2}\,dt = I_{4_5}$.

By Lemma 5,

$$
\begin{aligned}
I_{4_7} &= \int_0^b -\frac{1}{12}a^4\ln(-y_2+\sqrt{a^2+y_2^2})\,dy_2 \\
&= -\frac{1}{12}a^4b\ln(-b+\sqrt{a^2+b^2}) - \frac{1}{12}a^4\sqrt{a^2+b^2} + \frac{1}{12}a^5; \qquad (38)
\end{aligned}
$$

similarly,

$$\begin{aligned} I_{4_8} &= \int_0^b \frac{1}{12} a^4 \ln(t + \sqrt{a^2+t^2})\, dt \\ &= \frac{1}{12} a^4 b \ln(b + \sqrt{a^2+b^2}) - \frac{1}{12} a^4 \sqrt{a^2+b^2} + \frac{1}{12} a^5. \end{aligned} \tag{39}$$

Next, applying Lemma 4,

$$\begin{aligned} I_{4_9} &= \int_0^b \frac{1}{6} a y_2^3 \ln(a + \sqrt{a^2+y_2^2})\, dy_2 \\ &= -\frac{1}{96} ab^4 - \frac{1}{36} a^4 \sqrt{a^2+b^2} + \frac{1}{72} a^2 b^2 \sqrt{a^2+b^2} + \frac{1}{24} ab^4 \ln(a + \sqrt{a^2+b^2}) + \frac{1}{36} a^5, \end{aligned} \tag{40}$$

and $I_{4_{10}} = \int_0^b (at^3/6) \ln(a + \sqrt{a^2+t^2})\, dt = I_{4_9}$. Lemma 4 also implies

$$\begin{aligned} I_{4_{11}} &= \int_0^b -\frac{1}{6} a y_2^3 \ln(-a + \sqrt{a^2+y_2^2})\, dy_2 \\ &= \frac{1}{96} ab^4 - \frac{1}{36} a^4 \sqrt{a^2+b^2} + \frac{1}{72} a^2 b^2 \sqrt{a^2+b^2} - \frac{1}{24} ab^4 \ln(-a + \sqrt{a^2+b^2}) + \frac{1}{36} a^5 \end{aligned} \tag{41}$$

and $I_{4_{12}} = \int_0^b -(at^3/6) \ln(-a + \sqrt{a^2+t^2})\, dt = I_{4_{11}}$.

Finally, we combine the summands to obtain (after extensive algebraic manipulations)

$$I_4 = \frac{a^5 + b^5 - (a^4 - 3a^2b^2 + b^4)\sqrt{a^2+b^2}}{15} + \frac{a^4 b}{6} \ln\left(\frac{b + \sqrt{a^2+b^2}}{a}\right) + \frac{ab^4}{6} \ln\left(\frac{a + \sqrt{a^2+b^2}}{b}\right). \tag{42}$$

Appendix C: Visualization of Solution

To visualize how the expected value of D varies with the dimensions and aspect ratio of the rectangle, we plotted Equation 1 using the math package *Maple* [2]. Figure 2 shows $E[D]$ as a function of the rectangle dimensions; Figure 3 shows $E[D]$ as a function of aspect ratio for rectangles of constant area 1. In addition, Figure 1 of Section 5 shows $E[D]$ as a function of aspect ratio and rectangle area.

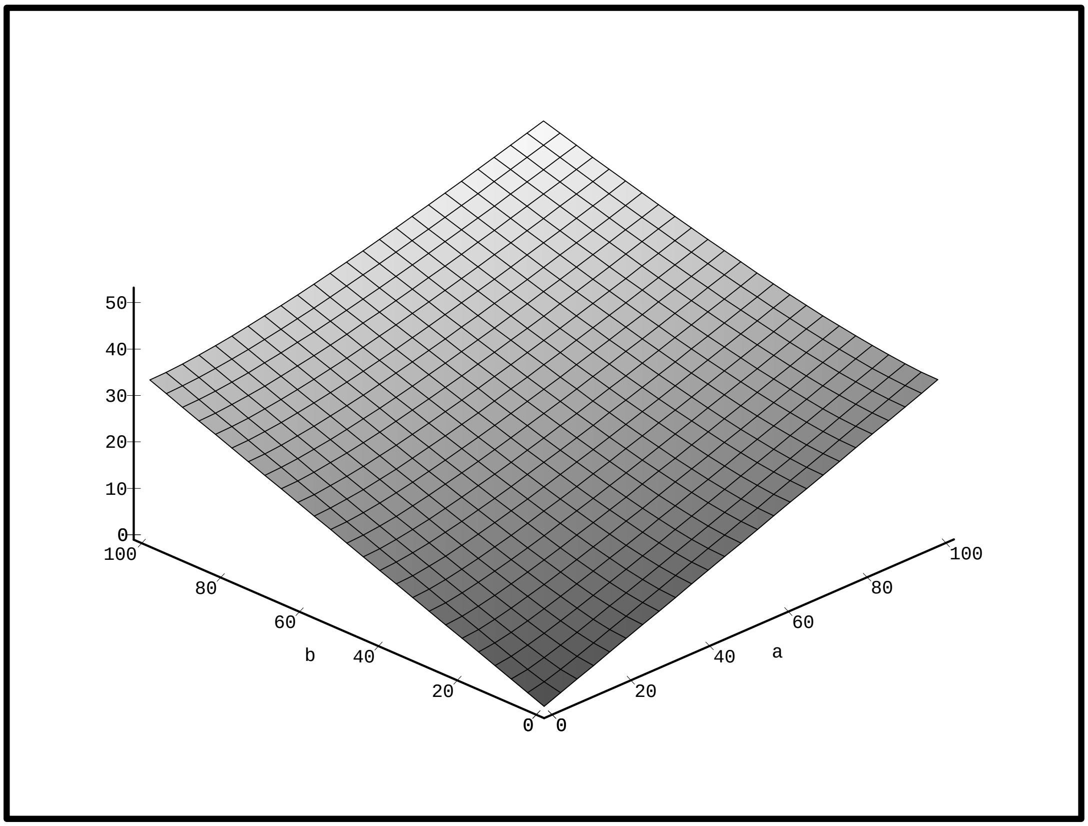

Figure 2: Expected value of D, as function of rectangle dimensions a and b.

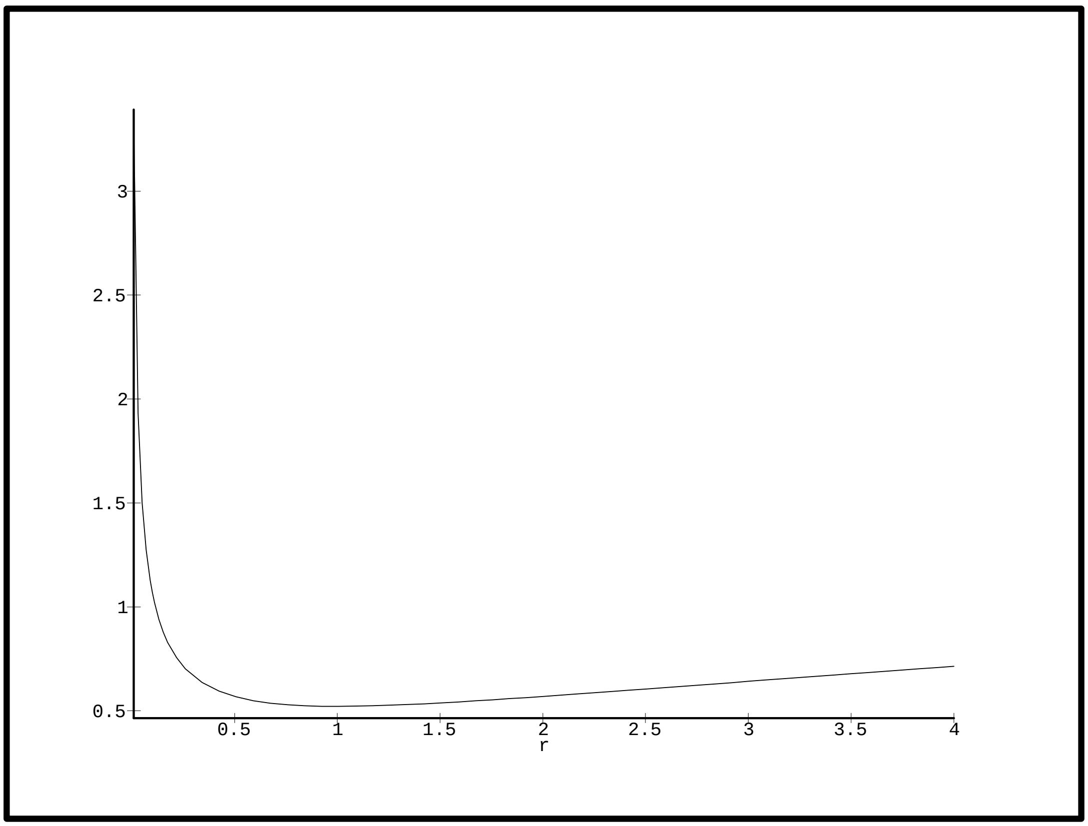

Figure 3: Expected value of D for unit-area rectangles of dimensions a and b, as function of aspect ratio $r = a/b$.