

Computing the Distance between Piecewise-Linear Bivariate Functions

Guillaume Moroz, Boris Aronov

► **To cite this version:**

Guillaume Moroz, Boris Aronov. Computing the Distance between Piecewise-Linear Bivariate Functions. SODA - Proceedings of the Twenty-Third Annual ACM-SIAM Symposium on Discrete Algorithms - 2012, Jan 2012, Kyoto, Japan. SIAM, 2012. <inria-00608255v2>

HAL Id: inria-00608255

<https://hal.inria.fr/inria-00608255v2>

Submitted on 12 Jul 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Computing the Distance between Piecewise-Linear Bivariate Functions*

Guillaume Moroz[†]

Boris Aronov[‡]

Abstract

We consider the problem of computing the distance between two piecewise-linear bivariate functions f and g defined over a common domain M . We focus on the distance induced by the L_2 -norm, that is $\|f - g\|_2 = \sqrt{\iint_M (f - g)^2}$. If f is defined by linear interpolation over a triangulation of M with n triangles, while g is defined over another such triangulation, the obvious naïve algorithm requires $\Theta(n^2)$ arithmetic operations to compute this distance. We show that it is possible to compute it in $\mathcal{O}(n \log^4 n)$ arithmetic operations, by reducing the problem to multi-point evaluation of a certain type of polynomials.

We also present an application to terrain matching.

*Work on this paper was initiated at the International INRIA-McGill-Victoria Workshop on Problems in Computational Geometry, held at the Bellairs Research Institute of McGill University in Barbados, West Indies.

[†]INRIA Nancy - Grand Est, 615 rue du Jardin Botanique, 54600 Villers-lès-Nancy, France.

[‡]Department of Computer Science and Engineering, Polytechnic Institute of NYU, Brooklyn, NY 11201-3840, USA; aronov@poly.edu. Work by B.A. on this paper has been supported by grant No. 2006/194 from the U.S.-Israel Binational Science Foundation, by NSF Grant CCF-08-30691, and by NSA MSP Grant H98230-10-1-0210.

1 Introduction and problem statement

In this paper we use a novel combination of tools from computational geometry and computer algebra to speed up a computation of $\|\cdot\|_2$ -norm distance between two bivariate piecewise-linear functions. Algebraic tools have already been used in other recent work in computational geometry, to seemingly defy “obvious” lower bounds: for example, Ajwani, Ray, Seidel, and Tiwary [2] use algebraic tools to compute the centroid of all vertices in an arrangement of n lines in the plane, without explicitly computing the vertices.

We feel that working on computational geometry problems using a combination of traditional and algebraic tools expands the repertoire of questions that can be approached and answered satisfactorily. Employing such combinations of methods expands the horizon of solvable problems. Indeed, in a significant recent development, several breakthrough results have been obtained by applying algebraic methods to problems of combinatorial geometry: Guth and Katz’s recent work on the joints problem [8] and on the Erdős distinct distance problem [9] has triggered an avalanche of activity; see, for example, [5, 6, 10–14]. It appears that the use of algebraic tools in geometry (both combinatorial and computational) allows one to approach problems inaccessible by more traditional methods. The present work is just one step in that direction.

Background and previous work In [1], Aronov et al. considered a quite common object in computational geometry and geographical information systems (GIS): that of a “terrain.” A *terrain* is (the graph of) a bivariate function over some planar domain, say, a rectangle or a square. It is often used to model geographic terrains, e.g., elevation in a mountainous locale, but can also be applied to storing any two-dimensional data sets, such as precipitation or snow cover data. A common (though by no means the only) way of interpolating and representing discrete two-dimensional data is a *triangulated irregular network (TIN)*: the values of a bivariate function are given at discrete points in, say, the unit square. The square is triangulated, using data points as vertices. The function is then linearly interpolated over each triangle. This produces a piecewise-linear approximation of the “real” (and unknown) function.

The problem raised in [1] was that of comparing two terrains over the same domain, say, the unit square, but given over two unrelated triangulations. One could imagine comparing the outcome of two different ways of measuring the same data, or finding correlation between, say, the elevation and the snow cover over the same geographic region. The focus of that work was on identifying linear dependence between the two functions or terrains. Three natural distance measures between the two functions were considered and several algorithms presented for computing such a distance and optimizing it, subject to vertical translation and scaling. The only observation made for the $\|\cdot\|_2$ norm (see the definitions below) in [1] is that, if the two terrains share a triangulation, both the distance computation and the optimization problem can be solved easily in linear time, while it appears that in general quadratic time seems to be needed to deal with the case of arbitrarily overlapping triangulations. The substance of the current work is disproving this assertion and describing a near-linear-time algorithm for both problems.

Problem statement and results Given bivariate functions $f, g: M \rightarrow \mathbb{R}$, one can naturally define a distance between them as

$$\|f - g\|_2 = \left(\iint_M (f(x, y) - g(x, y))^2 dx dy \right)^{1/2}.$$

Expanding the expression under the integral, we obtain $\iint (f - g)^2 = \iint f^2 - 2 \iint fg + \iint g^2$. If the two functions are piecewise linear, defined over different triangulations of M , only the middle term

presents a problem for efficient computation. Thus, in the bulk of the paper we will focus on the computation of $\iint fg$, showing the following:

Theorem 1.1. *Given piecewise-linear functions f and g defined over different triangulations of the same domain M , with n triangles each, $\iint_M f(x, y)g(x, y)dxdy$ can be computed using $\mathcal{O}(n \log^4 n)$ arithmetic operations.*

Armed with this result, as already mentioned, we can quickly compute $\|f - g\|_2$:

Theorem 1.2. *Given piecewise-linear functions f and g defined over different triangulations of the same domain M , with n triangles each, $\|f - g\|_2$ can be computed using $\mathcal{O}(n \log^4 n)$ arithmetic operations.*

Naïvely, the integral in Theorem 1.1 can be expressed as a sum of integrals over each cell appearing in the overlay of the two triangulations of f and g . Unfortunately, this overlay has a quadratic number of cells, in the worst case. The main idea of our algorithm is to reduce the computation of the integral to double sums of algebraic functions over grids, which allows us to use fast multi-point evaluation algorithms.

The paper is organized as follows. In section 2, we will show how the integral of a function over a convex polygon can be expressed as a sum of elementary algebraic functions over its vertices. Applying the process to a convex decomposition of a region M expresses the value of $\iint_M fg$ as a summation over the vertices of the decomposition. Then, in section 3, we show how the integral over the overlay of the two triangulations can be reduced to a sum of elementary functions over pairs of edges, plus some additional terms computable in linear time. In section 4, we use the bipartite clique decomposition to arrange the pairs of edges in complete grids of fairly rigid form. Finally, in section 5, we use a fast multi-point evaluation algorithm to compute the sums over each grid, completing the description of our method.

2 How to integrate over a convex subdivision

Consider a bivariate function h defined over a convex polygon C in the plane. To simplify our presentation and without loss of generality, we assume that all vertices of C lie to the right of the y -axis. For a vertex p of C we refer to the lines supporting the edges of C incident to p as $L(p, C)$ (the one with higher slope) and $U(p, C)$ (the one with lower slope); to the left of p , $L(p, C)$ is below $U(p, C)$. Let

$$\begin{aligned} L(p, C) &: y = y_l(p, C) + s_l(p, C)x, \text{ and} \\ U(p, C) &: y = y_u(p, C) + s_u(p, C)x. \end{aligned}$$

We omit the explicit dependence on C and/or p whenever it causes no confusion. Define

$$\delta(p, C) = \begin{cases} -1 & \text{if } C \text{ is above both } L(p, C) \text{ and } U(p, C), \text{ or below both of them,} \\ +1 & \text{if } C \text{ is below } U(p, C) \text{ and above } L(p, C), \text{ or vice versa.} \end{cases}$$

Finally, put

$$\mathcal{T}(p, C, h) := \delta(p, C) \int_{x=0}^{x_{p_j}} \int_{y=y_l(p, C)+s_l(p, C)x}^{y_u(p, C)+s_u(p, C)x} h(x, y) dx dy.$$

In words, \mathcal{T} is the signed integral of h over the triangle T_p delimited by the y -axis, $L(p, C)$, and $U(p, C)$.

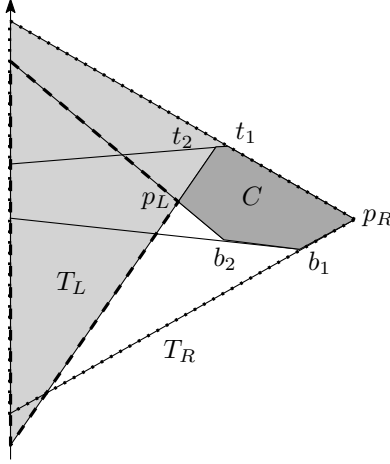


Figure 1: An illustration of the proof of Lemma 2.1. T_R is dotted; T_L is dashed; A_T is lightly shaded.

With the above notation, we express the integral of h over C in a convenient way as a sum of terms associated with its vertices:

Lemma 2.1. *Let C be a convex polygon with vertices p_1, \dots, p_k , and h a bivariate function. Then*

$$\iint_C h(x, y) dx dy = \sum_{j=1}^k \mathcal{T}(p_j, C, h). \quad (1)$$

Proof. Partition the vertices of C into four subsets V_L , V_R , V_T , and V_B as follows. $V_L := \{p_L\}$ (resp., $V_R := \{p_R\}$) consists of the unique leftmost (resp., rightmost) vertex of C . $V_T := \{t_1, \dots\}$ (resp., $V_B := \{b_1, \dots\}$) is the sequence of vertices of C from p_R to p_L in the counterclockwise (resp., clockwise) direction; refer to Fig. 1.

To each vertex p , we associate the triangle T_p as defined above. Put $T_L := T_{p_L}$ and $T_R := T_{p_R}$. Since C is convex, so for all i , line $t_i t_{i+1}$ (resp., $b_i b_{i+1}$) is below the line $t_{i-1} t_i$ (resp., above the line $b_{i-1} b_i$) left of C . Thus, the triangles T_t , $t \in V_T$, (resp., T_b , for $b \in V_B$) do not overlap. Let $A_T = \bigcup_{t \in V_T} T_t$ and $A_B = \bigcup_{b \in V_B} T_b$. By construction we have

$$A_T \cup A_B = (T_L \cup T_R) \setminus C \quad \text{and} \quad A_T \cap A_B = T_L \cap T_R.$$

Since $C \subset T_L \cup T_R$, the first equality, written in terms of characteristic functions, gives

$$1_{A_T} + 1_{A_B} - 1_{A_T \cap A_B} = 1_{T_L} + 1_{T_R} - 1_{T_L \cap T_R} - 1_C,$$

which can be simplified, using the second equality, to yield, as promised

$$1_{A_T} + 1_{A_B} = 1_{T_L} + 1_{T_R} - 1_C. \quad \square$$

We now consider a convex subdivision \mathcal{C} of some bounded region M in the plane, with each cell C associated with its own bivariate function h_C , thereby defining a function h over all of M (as it does not affect the value of the integral, the functions h_C need not agree along the common boundaries of adjacent cells). By summing eq. (1) over the cells C of \mathcal{C} , we can compute $\iint_M h(x, y) dx dy$:

Corollary 2.2.

$$\iint_M h(x, y) dx dy = \sum_{\text{cell } C \in \mathcal{C}} \iint_C h_C(x, y) dx dy = \sum_{\substack{\text{cell } C \in \mathcal{C} \\ C \text{ adjacent to } p}} \mathcal{T}(p, C, h_C). \quad (2)$$

One of the advantages of the formulation in eq. (2) for our application is that an individual integral under the sum can be expressed as a rational function when h_C is a bivariate polynomial.

Lemma 2.3. *Let p be an intersection point of $y = y_l + s_l x$ and $y = y_u + s_u x$ as above. Then*

$$x_p = -\frac{y_u - y_l}{s_u - s_l}$$

and

$$\int_{x=0}^{x_p} \int_{y=y_l+s_l x}^{y_u+s_u x} x^i y^j dx dy = \frac{P_{i,j}(y_l, y_u, s_l, s_u)}{(s_u - s_l)^{i+j+1}},$$

where $P_{i,j}$ is a polynomial of total degree $i + 2j + 2$.

Proof. Let $Q_{i,j}(u, v, x)$ be the only polynomial such that $\frac{\partial Q_{i,j}}{\partial x} = x^i(u + vx)^j$ and $Q_{i,j}(u, v, 0) = 0$ for all $u, v \in \mathbb{R}$. In particular, $Q_{i,j}$ has total degree $i + 2j + 1$, its degree in x is $i + j + 1$, and the coefficient of x^{i+j+1} in $Q_{i,j}$ is $\frac{1}{i+j+1}v^j$. Now

$$\begin{aligned} \int_{x=0}^{x_p} \int_{y=y_l+s_l x}^{y_u+s_u x} x^i y^j dx dy &= \int_{x=0}^{x_p} \frac{1}{j+1} (x^i (y_u + s_u x)^{j+1} - x^i (y_l + s_l x)^{j+1}) dx \\ &= \frac{1}{j+1} (Q_{i,j+1}(y_u, s_u, x_p) - Q_{i,j+1}(y_l, s_l, x_p)). \end{aligned}$$

Therefore the value of the integral can be expressed as a polynomial in y_l, y_u, s_l, s_u, x_p of total degree $i + 2j + 3$ of the form

$$\frac{1}{j+1} (s_u^{j+1} - s_l^{j+1}) x_p^{i+j+2} + \sum_{k=0}^{i+j+1} q_k(y_u, s_u, y_l, s_l) x_p^k.$$

After substituting $-\frac{y_u - y_l}{s_u - s_l}$ for x_p and bringing to a common denominator, we conclude that the expression can be rewritten in the form

$$\frac{(s_u - s_l)P(y_l, y_u, s_l, s_u)}{(s_u - s_l)^{i+j+2}},$$

as claimed. □

3 How to integrate over an overlay

In our problem, we are interested in computing the integral $\iint f(x, y)g(x, y) dx dy$, where f is defined over a triangulation \mathbf{T}_f of $M \subset \mathbb{R}^2$, with a separate linear function $f_\Delta(x, y)$ determining f over each triangle $\Delta \in \mathbf{T}_f$; g is defined similarly over a different triangulation \mathbf{T}_g of M . The product $h(x, y) := f(x, y)g(x, y)$ is thus naturally defined over the convex decomposition \mathcal{C} of M that is the overlay of \mathbf{T}_f and \mathbf{T}_g . By Corollary 2.2, it is sufficient to evaluate a sum over all vertices of \mathcal{C} . The

vertices of \mathcal{C} come in two flavors: the original vertices of \mathbf{T}_f and of \mathbf{T}_g , and the intersections of edges of \mathbf{T}_f and \mathbf{T}_g . Therefore,

$$\begin{aligned} \iint_M f(x,y)g(x,y)dxdy &= \sum_{\substack{p \text{ vertex of } \mathcal{C} \\ C \text{ adjacent to } p}} \sum_C \mathcal{T}(p, C, h_C) \\ &= \underbrace{\sum_{\substack{p \text{ vertex} \\ \text{of } \mathbf{T}_f \cup \mathbf{T}_g}} \sum_{\substack{C \text{ adjacent} \\ \text{to } p}} \mathcal{T}(p, C, h_C)}_{\Sigma_v} + \underbrace{\sum_{\substack{p=e_1 \cap e_2, \\ (e_1, e_2) \in \mathbf{T}_f \times \mathbf{T}_g}} \sum_{\substack{C \text{ adjacent} \\ \text{to } p}} \mathcal{T}(p, C, h_C)}_{\Sigma_e}. \end{aligned}$$

Σ_v involves $\mathcal{O}(n)$ integrals. We preprocess each of \mathbf{T}_f and \mathbf{T}_g for logarithmic-time point location queries, in $\mathcal{O}(n \log n)$ time (see, for example, [3]). For each vertex p of \mathbf{T}_f , we locate the triangle $\Delta_g \in \mathbf{T}_g$ containing it. Then, for each triangle $\Delta_f \in \mathbf{T}_f$ incident to p , we can compute $\mathcal{T}(p, \Delta_g \cap \Delta_f, f_{\Delta_f} g_{\Delta_g})$ in constant time. The treatment of vertices of \mathbf{T}_g is symmetric. We spend $\mathcal{O}(n \log n)$ total for point location. The remaining work is proportional to the sum of vertex degrees in both triangulations, which is $\mathcal{O}(n)$. Hence Σ_v can be computed in $\mathcal{O}(n \log n)$ arithmetic operations. We devote the rest of the discussion to computing Σ_e efficiently.

4 Bipartite clique decomposition

Let E_f be the set of edges of \mathbf{T}_f and E_g the set of edges of \mathbf{T}_g . In [4], it is shown that it is possible to compute a family $\mathcal{F} = \{(R_1, B_1), \dots, (R_u, B_u)\}$ where $R_k \subset E_f$ and $B_k \subset E_g$, such that

- (i) every segment in R_k intersects every segment in B_k ;
- (ii) every segment of R_k has lower slope than every segment of B_k , or vice versa;
- (iii) for every intersecting pair $(e_1, e_2) \in E_f \times E_g$ there is exactly one k such that $e_1 \in R_k, e_2 \in B_k$; no such k exists for a non-intersecting pair $(e_1, e_2) \in E_f \times E_g$;
- (iv) $\sum_k (|R_k| + |B_k|) = \mathcal{O}(n \log^2 n)$.

This family can be computed in $\mathcal{O}(n \log^2 n)$ time.

5 Multi-point evaluation

In this section, we explain how to efficiently compute

$$\sum_{\substack{p=e_1 \cap e_2, \\ (e_1, e_2) \in R \times B}} \sum_{\substack{C \text{ adjacent} \\ \text{to } p}} \mathcal{T}(p, C, h_C) = \sum_{e_1 \in R} \sum_{e_2 \in B} \sum_{\substack{C \text{ adjacent} \\ \text{to } p = e_1 \cap e_2}} \mathcal{T}(p, C, h_C), \quad (3)$$

where $(R, B) := (R_k, B_k)$ is a pair of sets of triangulation edges produced by the bipartite clique decomposition. In particular, all segments of R intersect all segments of B and, moreover, without loss of generality, the slopes of segments of R are greater than those of segments of B .

5.1 Reduction to sums of rational functions

Each triangle of \mathbf{T}_f and \mathbf{T}_g is associated with a bivariate linear function. If e is an edge of $\mathbf{T}_f \cup \mathbf{T}_g$, let $f_u(e)$ be the linear function associated to the upper triangle (in the triangulation to which e belongs) adjacent to e ; $f_l(e)$ is the corresponding function for the lower triangle; we can define the function to be identically zero for regions outside M , but it will never be used by the algorithm.

A vertex $p = e_1 \cap e_2$, with $e_1 \in R$ and $e_2 \in B$, lies on the boundary of four cells of \mathcal{C} . We focus on the cell $C_{left} = C_{left}(p)$ lying above e_1 and below e_2 , for which p is the rightmost point. Thus $h_{C_{left}} = f_u(e_1)f_l(e_2)$ and $\delta(p, C_{left}) = +1$. Suppose $f_u(e_1): (x, y) \mapsto a(e_1) + b(e_1)x + c(e_1)y$ and $f_l(e_2): (x, y) \mapsto a(e_2) + b(e_2)x + c(e_2)y$. We compute the contribution to the sum (3) of such cells, over all choices of p . (The remaining three types of cells adjacent to p are treated by an entirely symmetric argument.) Given an edge e of $\mathbf{T}_f \cup \mathbf{T}_g$, let $y = y(e) + s(e)x$ be the equation of the line supporting it, so we can write

$$\mathcal{T}(p, C_{left}, f_u(e_1)f_l(e_2)) = \int_{x=0}^{x_p} \int_{y=y(e_1)+s(e_1)x}^{y(e_2)+s(e_2)x} \begin{pmatrix} a(e_1)a(e_2) + (a(e_1)b(e_2) + a(e_2)b(e_1))x + \\ (a(e_1)c(e_2) + a(e_2)c(e_1))y + b(e_1)b(e_2)x^2 + \\ (b(e_1)c(e_2) + b(e_2)c(e_1))xy + c(e_1)c(e_2)y^2 \end{pmatrix} dx dy.$$

The above integral can be expressed as the sum of nine integrals of a function of the form $v(e_1)w(e_2)x^i y^j$, where v and w are some functions that assign a real number to each edge. Gathering all the terms and recalling that $p = e_1 \cap e_2$, eq. (3) can be rewritten as the sum of 36 expressions of the form

$$\sum_{e_1 \in R} \sum_{e_2 \in B} \int_{x=0}^{x_{e_1 \cap e_2}} \int_{y=y(e_1)+s(e_1)x}^{y(e_2)+s(e_2)x} v(e_1)w(e_2)x^i y^j = \sum_{e_1 \in R} \sum_{e_2 \in B} \frac{v(e_1)w(e_2)P_{i,j}(y(e_1), y(e_2), s(e_1), s(e_2))}{(s(e_2) - s(e_1))^{i+j+1}}, \quad (4)$$

by Lemma 2.3.

5.2 Fast multi-point evaluation

Now we will use multi-point evaluation to speed up the computation of eq. (4). To accomplish this, we will replace the values associated to the edges of R by symbolic variables, while using the actual numerical values for those for the edges of B . Then we will compute the corresponding symbolic rational function using a divide-and-conquer strategy (Lemma 5.1). Finally, we will use multi-point evaluation on the resulting polynomials (Lemma 5.2).

Lemma 5.1. *Let u and v be two functions from the edges of B to \mathbb{R} and suppose $|B| \leq n$. Then*

$$\sum_{e \in B} \frac{u(e)}{(X - v(e))^d} \quad (5)$$

can be expressed in the form

$$\frac{N(X)}{D(X)},$$

where $N(X)$ and $D(X)$ are polynomials of degree at most $(n-1)d$ and nd respectively; their coefficients can be computed explicitly in $\mathcal{O}(\mathcal{M}(nd) \log n)$ arithmetic operations, where $\mathcal{M}(q) = \mathcal{O}(q \log q)$ is the cost of multiplication of two univariate polynomials of degree at most q .

Proof. For simplicity of presentation and without loss of generality, assume that $|B|$ is a power of two. We bring eq. (5) to a common denominator by combining the fractions in pairs, reducing their

number to $|B|/2$, and repeating the process $\log |B| = \mathcal{O}(\log n)$ times. The bounds on the degree of the final numerator and denominator are immediate from examining the original fractions.

We now explain how to bring

$$\frac{N_1(X)}{D_1(X)} + \frac{N_2(X)}{D_2(X)},$$

with N_1, N_2, D_1, D_2 of degree at most kd , to a common denominator in time $3\mathcal{M}(kd) + \mathcal{O}(kd)$. Indeed, the above fraction is equal to

$$\frac{N_1(X)D_2(X) + N_2(X)D_1(X)}{D_1(X)D_2(X)},$$

so it can be computed by three calls to fast polynomial multiplication plus a linear number of additional operations, as claimed.

This completes the proof of the lemma, as the cost of one round of combining fractions with denominators and numerators of degree at most kd is $n/k \cdot (3\mathcal{M}(kd) + \mathcal{O}(kd)) = \mathcal{O}(\mathcal{M}(nd))$, since \mathcal{M} is superlinear. \square

The second lemma handles summing the values of a special kind of polynomials.

Lemma 5.2. *Let $P(X_0, X_1, \dots, X_r)$ be a polynomial of degree n in X_0 and d in X_1, \dots, X_r . Let E be a set of at most n points of \mathbb{R}^{r+1} . The values of P at the points of E can be simultaneously computed in $\mathcal{O}\left(\binom{d+r}{r}\mathcal{M}(n)\log n\right)$ time.*

Proof. P can be expanded with respect to the variables X_1, \dots, X_r , and has at most $\binom{d+r}{r}$ monomials. Each coefficient is a univariate polynomial in X_0 of degree at most n . Then, using standard multi-point evaluation algorithm for univariate polynomials [7, ch. 10], we can compute simultaneously the values of these univariate coefficients of P in $\mathcal{O}\left(\binom{d+r}{r}\mathcal{M}(n)\log n\right)$ at every point of E . Finally, we compute the values of each of the monomials of P in time $\mathcal{O}\left(\binom{d+r}{r}n\right)$. Combining all these values also costs $\mathcal{O}\left(\binom{d+r}{r}n\right)$ arithmetic operations, concluding the proof. \square

Now we are ready to efficiently evaluate eq. (4). Let $F(X, Y, V)$ be the polynomial

$$F(X, Y, V) := \sum_{e_2 \in B} \frac{Vw(e_2)P_{i,j}(Y, y(e_2), X, s(e_2))}{(s(e_2) - X)^{i+j+1}}.$$

After expanding the numerator of this fraction, we note that F has the form

$$F(X, Y, V) = \sum_{\substack{0 \leq d_X \leq i+2j+2 \\ 0 \leq d_Y \leq i+2j+2}} \left(\sum_{e_2 \in B} \frac{c_{d_X, d_Y, d_V}(w(e_2), y(e_2), s(e_2))}{(s(e_2) - X)^{i+j+1}} \right) X^{d_X} Y^{d_Y} V.$$

In our case, $i + j + 1 \leq 3$, and, using Lemma 5.1, we can compute each coefficient of $X^{d_X} Y^{d_Y} V$ and express F in the following form, in $\mathcal{O}(n \log^2 n)$ time:

$$F(X, Y, V) = \frac{N(X, Y)V}{D(X)},$$

with N a polynomial of degree $n(i + j + 1) + 1$ in X , $i + j + 2$ in Y , and D a univariate polynomial of degree $n(i + j + 1)$.

Finally, eq. (4) can be rewritten as

$$\sum_{e_1 \in R} \sum_{e_2 \in B} \frac{v(e_1)w(e_2)P_{i,j}(y(e_1), y(e_2), s(e_1), s(e_2))}{(s(e_2) - s(e_1))^{i+j+1}} = \sum_{e_1 \in R} \frac{N(s(e_1), y(e_1))v(e_1)}{D(s(e_1))}.$$

As $|R| < n$, using Lemma 5.2, we can compute simultaneously all the terms under the sum, and add them together in $\mathcal{O}(n \log^2 n)$ arithmetic operations.

5.3 Putting it together

To summarize, we can evaluate eq. (4) in $\mathcal{O}((|B_k| + |R_k|) \log^2(|B_k| + |R_k|))$ operations, for each pair (R_k, B_k) . We also saw in section 4 that $\sum_k (|R_k| + |B_k|) = \mathcal{O}(n \log^2 n)$, where n is the number of triangles appearing in our triangulations. Therefore

$$\sum_k (|R_k| + |B_k|) \log^2(|R_k| + |B_k|) \leq \sum_k (|R_k| + |B_k|) \log^2 n = \mathcal{O}(n \log^4 n),$$

which allows us to conclude that we can compute Σ_e in $\mathcal{O}(n \log^4 n)$ time. This completes the proof of Theorem 1.1.

6 Discussion

In [1], the following optimization problem was considered: given two functions f and g and a distance measure $\|f - g\|$ between them (the paper discusses $\|\cdot\|_1$, $\|\cdot\|_2$, and $\|\cdot\|_\infty$, but we only consider $\|\cdot\|_2$ here), find the values of real parameters s and t that minimize $\|f - (sg + t)\|$. If f and g are interpreted as geometric “terrains,” we are looking for the scaling and translation of the vertical coordinate of the terrain g to best match terrain f [1]. Since $\|f - (sg + t)\|_2^2 = \iint (f - (sg + t))^2$ is a degree-two polynomial in s and t with coefficients easily expressible in terms of $\iint f$, $\iint g$, $\iint f^2$, $\iint g^2$, $\iint fg$, and $\iint 1$, being able to compute $\iint fg$ efficiently immediately yields

Theorem 6.1. *Given piecewise-linear functions f and g defined over different triangulations of the same domain M , with n triangles each, the values s and t minimizing $\|f - (sg + t)\|_2$ can be computed using $\mathcal{O}(n \log^4 n)$ arithmetic operations.*

Theorems 1.1, 1.2, and 6.1 extend to piecewise-polynomial functions of constant maximum degree with essentially no modifications. What other classes of functions can be handles using similar methods?

Acknowledgments

The authors would like to thank Raimund Seidel, Christian Knauer and Sylvain Lazard for helpful discussions. The authors were also inspired to work on this problem at the Workshop on Discrete and Algebraic Geometry in September 2010, at Val d’Ajol, France.

References

- [1] P.K. Agarwal, B. Aronov, M. van Kreveld, M. Löffler, and R. Silveira, “Computing similarity between piecewise-linear functions,” *Proc. 26th Ann. Sympos. Comput. Geometry*, 2010, pp. 375–383.
- [2] D. Ajwani, S. Ray, R. Seidel, and H. Tiwary, “On computing the centroid of the vertices of an arrangement and related problems,” *Proc. Symp. Algo. Data Structures (WADS’07)*, LNCS 4619, 2007, pp. 519–528.

- [3] M. de Berg, O. Cheong, M. van Kreveld, and M. Overmars, *Computational Geometry: Algorithms and Applications*, 3rd ed., Springer-Verlag, 2008.
- [4] B. Chazelle, H. Edelsbrunner, L. J. Guibas, and M. Sharir, “Algorithms for bichromatic line segment problems and polyhedral terrains,” *Algorithmica*, 11:116–132, 1994.
- [5] Gy. Elekes, H. Kaplan, and M. Sharir, “On lines, joints, and incidences in three dimensions,” *J. Combinat. Theory, Ser. A* 118(2011):962–977. Also in [arXiv:0905.1583](#).
- [6] Gy. Elekes and M. Sharir, “Incidences in three dimensions and distinct distances in the plane,” *Proc. 26th Annu. ACM Sympos. Comput. Geom.*, pp. 413–422, 2010.
- [7] J. von zur Gathen and J. Gerhard, *Modern Computer Algebra*, 2nd ed., Cambridge University Press, 2003.
- [8] L. Guth and N.H. Katz, “Algebraic methods in discrete analogs of the Kakeya problem,” *Advances in Mathematics*, 225(5):2828–2839, 2010.
- [9] L. Guth and N.H. Katz, “On the Erdős distinct distance problem in the plane,” [arXiv:1011.4105v3 \[math.CO\]](#), 2011.
- [10] H. Kaplan, J Matoušek, and M. Sharir, “Simple proofs of classical theorems in discrete geometry via the Guth–Katz polynomial partitioning technique,” [arXiv:1102.5391v1 \[math.CO\]](#).
- [11] J. Matoušek, “The dawn of an algebraic era in discrete geometry?” *27th European Workshop on Computational Geometry (EuroCG 2011)*, extended abstract, 2011.
- [12] R. Quilodrán, “The joints problem in \mathbb{R}^n ,” [arXiv:0906.0555v3 \[math.CO\]](#).
- [13] J. Solymosi and T. Tao, “An incidence theorem in higher dimensions,” [arXiv:1103.2926v2 \[math.CO\]](#).
- [14] K. Haim, S. Micha, and E. Shustin, “On lines and joints,” *Discrete Comput. Geometry*, 44(4)(2010):838–843.