Algorithmica (1998) 20: 61–76

Algorithmica

Minkowski-Type Theorems and Least-Squares Clustering[1]

F. Aurenhammer,[2] F. Hoffmann,[3] and B. Aronov[4]

Abstract. Dissecting Euclidean d-space with the power diagram of n weighted point sites partitions a given m-point set into clusters, one cluster for each region of the diagram. In this manner, an assignment of points to sites is induced. We show the equivalence of such assignments to constrained Euclidean least-squares assignments. As a corollary, there always exists a power diagram whose regions partition a given d-dimensional m-point set into clusters of prescribed sizes, no matter where the sites are placed. Another consequence is that constrained least-squares assignments can be computed by finding suitable weights for the sites. In the plane, this takes roughly $O(n^2 m)$ time and optimal space $O(m)$, which improves on previous methods. We further show that a constrained least-squares assignment can be computed by solving a specially structured linear program in $n+1$ dimensions. This leads to an algorithm for iteratively improving the weights, based on the gradient-descent method. Besides having the obvious optimization property, least-squares assignments are shown to be useful in solving a certain transportation problem, and in finding a least-squares fitting of two point sets where translation and scaling are allowed. Finally, we extend the concept of a constrained least-squares assignment to continuous distributions of points, thereby obtaining existence results for power diagrams with prescribed region volumes. These results are related to Minkowski's theorem for convex polytopes. The aforementioned iterative method for approximating the desired power diagram applies to continuous distributions as well.

Key Words. Power diagrams, Least-squares clustering, Point partitioning.

1. Introduction. The purpose of this paper is to discuss a relationship between power diagrams and so-called constrained least-squares assignments.

A *power diagram* is a generalization of the classical Voronoi diagram of a set S of n points in Euclidean d-space. The points in S (called *sites* in what follows) have individual weights expressing their capability to influence their neighborhood. The regions of a power diagram define a convex polyhedral partition of d-space. (See, e.g., the survey paper [6] for properties of Voronoi-type diagrams in general and power diagrams in particular.) If we fix, in addition to the sites, according to their containment in the regions. This naturally defines an assignment of points in X to sites in S. This assignment depends on the weighting of S.

---

[1] Part of the work by Franz Aurenhammer and Boris Aronov was carried out at the Institut für Informatik, Freie Universität Berlin. The research of Franz Aurenhammer was partially supported by the Deutsche Forschungsgemeinschaft under Grant A1 253/1-4, Schwerpunktoprogramm "Datenstrukturen und effziente Algorithmen." Work by Boris Aronov was supported in part by NSF Grant CCR-92-11541. The present paper is an extended version of the conference article [7].
[2] Institute for Theoretical Computer Science, Graz University of Technology, A-8010 Graz, Austria.
[3] Institut für Informatik, Technische Universität München, München, Germany.
[4] Computer and Information Science Department, Polytechnic University, New York, NY 11201, USA.

May 30, 1995; revised June 25, 1996. Communicated by K. Mehlhorn.

A *least-squares assignment* of a set X of points to a set S of sites in d-space is defined to minimize the total square distance between the sites and their associated points. Our interest is in *constrained* least-squares assignments where the number of associated points per site is prescribed by the so-called *capacity* if a site.

The main result of this paper shows that power diagrams and constrained least-squares assignments are equivalent in the following sense. Any assignment $X \to S$ induced by a power diagram of S is a least-squares assignment subject to the capacities resulting from the power diagram. Conversely, a least-squares assignment $X \to S$, for any choice of capacities, can be realized by the power diagram of S for appropriate weights.

This result contributes to the known list of optimal clusterings of a point set X induced by Voronoi-type diagrams. For example, the *k-centroid problem* asks for a set S of k sites such that the unconstrained least-squares assignment $X \to S$ has minimum value. (Each site then is the centroid of its cluster, hence the name.) Boros and Hammer [9] observed that unconstrained least-squares $X \to S$ are always induced by the Voronoi diagram of S, which by definition assigns each point in d space to the closest site in S. Similarly, in the *k-center problem* a set S of k sites is required such that X can be covered by k disks centered at the sites and having the minimum sum, or maximum, of radii. The corresponding clusterings are realized by power diagrams of S; see [10]. Both the k-centroid and the k-center problem are known to be NP-hard. However, realizability by Voronoi-type diagrams restricts the number of candidate clusterings to $|X|^{O(dk)}$ and thus leads to polynomial-time algorithms for fixed k; see [10] and [15]. Computing least-squares assignments for a *given* set of sites, as is discussed in this paper, is a computationally simpler problem as only the assignment needs to be optimized.

The equivalence of constrained least-squares assignments and power diagrams extends to the case where, instead of a finite set X, a continuous, nonvanishing probability distribution in the unit cube is considered. Proofs for both the finite and continuous version are given in Section 2.

The remainder of the paper is concerned with several consequences of these equivalence results. Partition theorems for power diagrams are direct corollaries: there always exists a power diagram whose regions partition a given d-dimensional finite point set X into clusters of prescribed sizes, no matter where the sites there exists a power diagram that partitions the unit hypercube into n convex polyhedral regions of prescribed volumes. These results and their relation to Minkowski-type theorems for convex polyhedra are discussed in Section 3.

Exploiting the machinery of power diagrams, we propose two algorithms for computing constrained least-squares assignments in the plane. The first algorithm, described in Section 4, applies to finite point sets X. It proceeds by inserting the m pints of X, one by one, at each step adjusting the weights of the n sites such that the capacities are not exceeded. Time complexity of $O(n^2 m \log m + nm \log^2 m)$ and optimal space complexity $O(m)$ are achieved. As $m \geq n$ can be assumed, this is an improvement over the $O(nm^2 + m^2 \log m)$-time and $O(nm)$-space algorithm that results from transforming the problem into a minimum-cost flow problem; see [13]. (For a discussion of the general assignment problem, see also [20].) Alberts [4] recently reduced the running time to $O(n^2 m \log m)$, by generalizing the Hungarian method. The space requirement is still $O(n, m)$, as is the case for the randomized algorithm of Tokuyama and Nakano [21] that achieves expected time $O(nm + n^3\sqrt{nm})$. On the other hand, the cited algorithms are more general than ours in that they find the optimum constrained assignment on a general weighted bipartite graph. We are able to exploit the geometric interpretation of least-squares assignments to reduce the space requirement to $O(m)$.

The second algorithm, outlined in Section 5, is applicable to both the finite and the continuous version of the problem. We show that finding a weight vector that yields the desired least-squares assignment is equivalent to finding a maximum of a concave n-variate function whose domain is the weight space. For the continuous case, we propose a gradient method for iteratively improving the weight vector. This method has superlinear convergence and optimal $O(n)$ space requirement. In the finite case, on the other hand, the n-variate function to be optimized is piecewise-linear. Finding a point in the maximum is now a linear programming problem whose number of constraints is, however, exponential in n. Our iterative algorithm can still be used. Experiments have shown that it approximates the maximum quite fast. Again, the space requirement is optimal, $O(m)$.

Section 6 sketches some applications of our results. We show that a certain transportation problem can be solved by computing a constrained least-squares assignment. Another application exploits the fact that constrained least-squares assignments are invariant under translation and scaling of the set S of sites. We obtain an algorithm that finds the best least-squares fitting of two n-point sets S and X, under translation and scaling, in $O(n^3)$ time and $O(n)$ space. The time complexity matches that of the matching algorithm for general weighted bipartite graphs [16] which requires $O(n^2)$ space. Vaidya [23] described an $O(n^2\sqrt{n}\log n)$-time and $O(n \log n)$-space bipartite matching algorithm for a version of the problem in which weights are Euclidean distances. His algorithm seems to generalize directly to least-squares matchings, with additively weighted Voronoi diagrams replaced in his data structure by power diagrams. Recent developments reduce the running time of Vaidya's algorithm to $O(n^{2+\epsilon})$ [1]; its version for power diagrams follows from [2]. To summarize, a least-squares fitting of two n-point sets can be computed either in $O(n^{2+\epsilon})$ time and $O(n^{1+\epsilon})$ space or, with our algorithm, in $O(n^3)$ time and optimal $O(n)$ space.

2. Equivalence Theorems.

This section establishes an equivalence between constrained least-squares assignments and power diagrams.

Consider a set S of n point sites in Euclidean d-space $\mathbb{E}^d$. S induces a partition of $\mathbb{E}^d$ into polyhedral regions in the following natural way. The *region* of a site $s \in S$, $reg(s)$, consists of all points x which are closer to s than to the remaining $n - 1$ sites. This partition is known as the *Voronoi diagram* of S. If we fix, in addition to the sites, a set X of m points in $\mathbb{E}^d$, then this set is partitioned by the Voronoi diagram of S into subsets. More precisely, the diagram defines an *assignment function* $A: X \to S$, given by

$$A(x) = s \quad \Leftrightarrow \quad x \in reg(s).$$

Equivalently, $A^{-1}(s) = X \cap reg(s)$ for all $s \in S$. Note that points of X that have more than one closest site in S are not covered by this definition. By convention, A assigns each such point to an arbitrary but fixed closest site. The total number of points assigned to a particular site s, $|A^{-1}(s)|$, is called the *capacity* of s. The capacities of all sites add up to

$m = |X|$. The assignment A has an obvious optimization property: it minimizes the sum of the distances between sites and their assigned points, over all possible assignments of X to S.

Given S and X, we would like to be able to change the assignment by varying the distance function that underlies the Voronoi diagram of S. To this end, we attach a set $W = \{w(s) \mid s \in S\}$ of real numbers, called *weights*, to the sites and replace the Euclidean distance $\delta(x, s)$ between a point x and a site s by the *power function*

$$\mathrm{pow}_W(x, s) = \delta^2(x, s) - w(s).$$

The resulting partition of $\mathbb{E}^d$ is known as the *power diagram* of S with weights W. Each region is still a convex polyhedron, and has the property of shrinking (resp. expanding) when the weight of its defining site is decreased (resp. increased). As above, we obtain an assignment function $A_W\colon X \to S$ which now clearly depends on the choice of weights. In particular, the site capacities depend on W.

Power diagrams also given rise, in the obvious way, to mappings of the entire d-space to the set of sites. Let $A_W\colon \mathbb{E}^d \to S$ be the assignment induced by the power diagram of S with weights W. That is, $A_W^{-1}(s) = reg_W(s)$, the region of site s in the diagram. The capacity of a site can now be defined as the fraction of the unit hypercube contained in its region. Formally, let ϱ be a continuous and nonvanishing probability distribution on $[0, 1]^d$, and let $\mu(X) = \int_X \varrho(x)\,dx$ denote the measure of a set $X \subset \mathbb{E}^d$ with respect to ϱ. Then $\mu(A_W^{-1}(s))$ is the capacity of s that results from A_W. The capacities of all sites add up to 1.

We prove the following general result.

THEOREM 1. *Let S be a finite set of sites in $\mathbb{E}^d$. Any (finite or continuous) assignment induced by a power diagram of S is a least-squares assignment, subject to the resulting capacities. Conversely, a least-squares assignment for S, subject to any given capacities (whose sum is the total number of assigned points in the finite case, and* 1 *in the continuous case) exists and can be realized by a power diagram of S.*

Theorem 1 contains several assertions which are now stated separately (and more precisely) and proved. We start by showing that assignments defined by power diagrams are constrained least-squares assignments. We consider the finite case first.

LEMMA 1. *Let S and X be finite sets of sites and points in $\mathbb{E}^d$, respectively, and fix a set W of weights for S. The assignments A_W minimizes*

$$\sum_{x \in X} \delta^2(x, A(x))$$

over all assignments $A\colon X \to S$ with capacity constraints $|A^{-1}(s)| = |A_W^{-1}(s)|$ for all $s \in S$.

PROOF. From the definition of A_W it is evident that A_W minimizes the expression

$$\sum_{x \in X} \mathrm{pow}_W(x, A(x)) = \sum_{x \in X} \delta^2(x, A(x)) - \sum_{x \in X} w(A(x))$$

over all possible assignments $A\colon X \to S$, regardless of the capacity constraints. The last sum, being equal to $\sum_{s\in S} |A^{-1}(s)| \cdot w(s)$, is a fixed constant for all assignments A with capacities $|A^{-1}(s)| = |A_W^{-1}(s)|$, and the lemma follows. □

The following continuous version of Lemma 1 can be proved in a similar fashion.

LEMMA 2. *Let S be a finite set of sites in $\mathbb{E}^d$ with weights W, let ϱ be some probability distribution on $[0,1]^d$, and let μ be the measure defined by ϱ. The assignment $A_W\colon [0,1]^d \to S$ minimizes*

$$\int_{[0,1]^d} \varrho(x) \cdot \delta^2(x, A(x))\, dx$$

over all assignments $A\colon [0,1]^d \to S$ with capacities $\mu(A^{-1}(s))$ for all $s \in S$.

We proceed to prove the existence and realizability of least-squares assignments with prescribed capacities. For finite point sets X, the existence of a constrained least-squares assignment $X \leftarrow S$ is trivial. Its realizability by power diagrams is proved in Section 4 by giving an algorithm that constructs such a power diagram. So, in the rest of this section, we concentrate on the continuous case only.

Fix a set S of sites in $\mathbb{E}^d$, a capacity function $c\colon S \to [0,1]$ with $\sum_{s\in S} c(s) = 1$, and a probability distribution ϱ in $[0,1]^d$. We now require that ϱ be continuous and nonvanishing in $[0,1]^d$. We assume that $c(s) > 0$, for any $s \in S$. Suppose that a least-squares assignment $L\colon [0,1]^d \to S$ subject to c exists. To simplify notation, let $R(s) = L^{-1}(s)$. We first show that L has to satisfy the following property. For any two sites $s, t \in S$, there is a hyperplane separating $R(s)$ from $R(t)$. More precisely, we have:

OBSERVATION 1. *Let $s, t \in S$, $s \neq t$. There exists a hyperplane H orthogonal to $t - s$ such that $\mu(H_{ts} \cap R(s)) = 0$ and $\mu(H_{st} \cap R(t)) = 0$, where H_{ts} is the half-space bounded by H and containing $H + (t - s)$, and H_{st} is the complementary half-space.*

PROOF. Suppose that there is no such hyperplane. Then there is a hyperplane H orthogonal to $t - s$ and such that $\mu(H_{ts} \cap R(s)) > 0$ and $\mu(H_{st} \cap R(t)) > 0$. Now we use the fact that, if a point $x \in R(s)$ is in H_{ts} and a point $y \in R(t)$ is in H_{st}, then x can be reassigned to t and y reassigned to s, thereby reducing the sum of squared distances. Indeed, applying the Pythagorean theorem gives

$$\delta^2(x,t) + \delta^2(y,s) < \delta^2(y,t) + \delta^2(x,s).$$

Integration over two subsets of $R(s)$ and $R(t)$ of equal positive measure that were assumed to exist on the wrong sides of H thus shows that these subsets could be reassigned, obtaining an assignment better than L but subject to the same capacities. This contradicts the assumed minimality of L and thus proves that there exists a hyperplane H that separates $R(s)$ from $R(t)$. □

Observation 1 implies that, if L exists, it can be realized by a family of convex polyhedra $\{P(s)\}$, one for each site s. These polyhedra are given by $P(s) = \bigcap_{t \neq s} H_{st}$ and therefore have pairwise disjoint interiors. In fact, their intersections with $[0, 1]^d$ induce a partition of $[0, 1]^d$, because we assumed the distribution ϱ to be nonvanishing.

Note that Observation 1 still holds for finite point sets X. We take ϱ to be the indicator function of X in $[0, 1]^d$ and replace integrals by sums. Degenerate positions of X may be handled by defining both H_{st} and H_{ts} in the statement of Observation 1 as open half-spaces. The resulting polyhedra, however, will not necessarily define a partition of $[0, 1]^d$ in the finite case. One could try to enforce a partition by considering X as the limit of a series of continuous and nonvanishing distributions. Such an approach seems feasible also for objects (of equal dimension) in $\mathbb{E}^d$ more general than points. We do not elaborate on this idea in this paper but rather deduce realizability by power diagrams in the finite case from the algorithm in Section 4.

Based on the convex partition property, we now show the existence of a constrained least-squares assignment L in the continuous case, i.e., when ϱ is a continuous and nonvanishing probability distribution.

LEMMA 3. *Consider the class of assignments* $[0, 1]^d \to S$ *realized by the family of convex polyhedra* $\{P(s)\}_{s \in S}$ *with the following properties*: (1) $\{P(s) \cap [0, 1]^d\}_{s \in S}$ *defines a partition of* $[0, 1]^d$, (2) $\mu(P(s)) = c(s)$ *for all* $s \in S$, *and* (3) *each* $P(s)$ *has fewer than* $|S|$ *facets. Then this class contains a least-squares assignment* L *subject to* c.

PROOF. Let $n = |S|$, and let P_i be a polyhedron with at most $(n - 1)$ facets associated with the ith site s_i. P_i is the intersection of $n - 1$ half-spaces in $\mathbb{E}^d$, each of which can be specified by the vector extending from s_i to its defining hyperplane and normal to it. Hence $P_1, \ldots, P_n$ are completely determined by a k-tuple of real numbers, for $k = n(n - 1)d$. For simplicity, we do not distinguish between P_i and its describing $((n - 1)d)$-tuple in the remainder of the proof. Now consider the continuous function

$$\varphi : \mathbb{R}^k \to \mathbb{R}^n, \qquad \varphi(P_1, \ldots, P_n) = (\mu(P_1), \ldots, \mu(P_n)).$$

Let $\prod_\varphi = \varphi^{-1}(c) \subset \mathbb{R}^k$. $\prod_\varphi$ corresponds to the set of all n-tuples of $(n - 1)$-facet polyhedra whose measures fulfill the capacity constraints. $\prod_\varphi$ is a closed set, being the inverse image of a closed set under a continuous function. Since ϱ is zero outside $[0, 1]^d$, attention may be restricted to a bounded subset of $\prod_\varphi$: there is a number b such that, for all i, if all entries of P_i are between $-b$ and b, then $\mu(P_i) = c(s_i)$ can still be achieved for all possible directions of half-space normals for P_i. Hence we need only consider tuples $(P_1, \ldots, P_n) \in \prod_\varphi \cap [-b, b]^k$. Next, take the continuous function

$$\psi: \mathbb{R}^k \to \mathbb{R}, \qquad \psi(P_1, \ldots, P_n) = \sum_{i \neq j} V(P_i \cap P_j \cap [0, 1]^d),$$

where V denotes the d-dimensional volume. Let $\prod_\psi = \psi^{-1}(0) \subset \mathbb{R}^k$. Again, $\prod_\psi$ is a closed set. It corresponds to the set of all n-tuples of $(n - 1)$-facet polyhedra yielding a packing in $[0, 1]^d$.

We now consider the compact set $\prod = \prod_\varphi \cap \prod_\psi \cap [-b, b]^k$. Each element $(P_1, \ldots, P_n)$ in $\prod$ fulfills both the capacity and the packing constraints. Recall that the sum of all capacities is 1, and that the probability distribution ϱ was assumed to be nowhere zero in $[0,1]^d$. We conclude that each $(P_1, \ldots, P_n) \in \prod$ induces a partition of $[0,1]^d$. We further know from Observation 1 that, if the constrained least-squares assignment L exists, it is realized by an element of $\prod$. Note that $\prod$ is nonempty; for example, take n parallel slices of $[0,1]^d$ with measures $c(s_i)$, $i = 1, \ldots, n$. Finally, consider the function $Q: \prod \to \mathbb{R}$,

$$
\begin{aligned}
Q(P_1, \ldots, P_n) &= \sum_{i=1}^{n} \int_{P_i} \varrho(x)\delta^2(x, s_i)\, dx \\
&= \int_{[0,1]^d} \varrho(x)\delta^2(x, A(x))\, dx,
\end{aligned}
$$

where $A: [0,1]^d \to S$ denotes the assignment defined by $(P_1, \ldots, P_n)$. Q expressed the value of the assignment A. Q is a continuous and nonnegative function whose domain is compact, so it must attain its minimum, the value of L. This proves the existence of L. □

We have recently learned that the existence and further the uniqueness of L can be deduced from properties of the so-called Monge–Kantorovich mass transference problem considered by Cuesta-Albertos and Tuero-Diaz in [12]. Their proof is in terms of probability theory and more general and involved than needed for the purposes of this paper.

Finally we show that constrained least-squares assignments—for the continuous case—can always be realized by power diagrams.

LEMMA 4. *The polyhedral family $\{P(s)\}_{s \in S}$ that realizes L has the property that, within $[0,1]^d$, for some choice W of weights for S, $P(s) = reg_W(s)$ for all $s \in S$.*

PROOF. From the proof of Lemma 3 we know that $\{P(s)\}_{s \in S}$ defines a partition $\{P'(s)\}_{s \in S}$ of $[0,1]^d$. Observation 1 implies that, for each pair of sites $s, t \in S$, if $P'(s)$ and $P'(t)$ share a facet F, then F is orthogonal to the vector $t - s$. Moreover, $F + (t - s)$ lies on the same side of the hyperplane through F as $R'(t)$ does. It is known [5] that these two conditions are necessary and sufficient for a convex partition of a given polytope Q to be the power diagram of S for some suitable set W of weights, restricted to Q. □

3. Partition Theorems.

The existence of least-squares assignments, together with their realizability by power diagrams, immediately implies several partitioning results for power diagrams.

THEOREM 2. *Let S and X be a set of n sites and m points in Euclidean d-space $\mathbb{E}^d$, respectively. For any choice of integer site capacities $c(s)$ with $\sum_{s \in S} c(s) = m$, there exists a set W of weights such that $|A_W^{-1}(s)| = c(s)$, for all sites $s \in S$.*

In other words, there always exists a power diagram whose regions partition a given d-dimensional finite point-set X into clusters of prescribed sizes, no matter where the sites of the power diagram are chosen. Moreover, we have the following continuous version of Theorem 2.

THEOREM 3. *Let S be a set of n sites in $\mathbb{E}^d$, let ϱ be a continuous and nonvanishing probability distribution on $[0,1]^d$, and let μ denote the measure with respect to ϱ. For any capacity function c: $S \to [0,1]$ with $\sum_{s \in S} c(s) = 1$, there is a set W of weights such that $\mu(reg_W(s)) = c(s)$, for all sites $s \in S$.*

By taking, for instance, ϱ be the uniform distribution in $[0,1]^d$ we get:

COROLLARY 1. *For any set of n sites in $\mathbb{E}^d$ there exists a power diagram that partitions the unit hypercube into n polyhedral regions of prescribed volumes.*

This seems surprising, as the placement of the sites determines the normals of the facets separating the power regions.

Corollary 1 is related to Minkowski's theorem for convex polytopes (see, e.g., [14]) which, for the purposes of this paper, can be stated as follows. Let V be any collection of n nonzero nonparallel vectors that span $\mathbb{E}^{d+1}$ and sum up to zero. Then there exists a $(d+1)$-polytope with n facets in one-to-one correspondence with vectors of V so that each facet is normal to its corresponding vector and has d-dimensional volume equal to the vector length.

It is well known that any power diagram for n sites in $\mathbb{E}^d$ is a projection of an unbounded $(d+1)$-polyhedron formed as the lower envelope of n hyperplanes, one for each site; see, for example, [6]. The orientation of the hyperplanes is determined by the placement of the sites, while their position is given by the corresponding weights. As facet orientations are fixed, giving their d-volume is equivalent to fixing the volume of their projection onto the hyperplane $s_{d+1} = 0$. Thus Corollary 1 is equivalent to the statement that an unbounded polyhedron whose facets have prescribed orientation and d-volume (within the prism $[0,1]^d \times \mathbb{R}$) always exists. This statement differs from Minkowski's theorem because of the presence of unbounded faces and a restricting prism.

A related and more general theorem was proven by Pogorelov [19, p. 476]. It is stated in three dimensions. A real-valued function σ on convex polygons in $\mathbb{E}^3$ is *monotone* if (1) it is positive for polygons with positive area, (2) if polygon Q is properly contained in polygon Q', then $\sigma(Q') > \sigma(Q)$, and (3) if Q' is obtained from Q by an upward translation, $\sigma(Q') \geq \sigma(Q)$. For example, area is a monotone function. Now let P be an unbounded polyhedron formed as the lower envelope of planes in $\mathbb{E}^3$. Define $\Omega(P)$ to be the set of all polyhedra that coincide with P outside a sufficiently large ball and whose bounded facets $f_1, \ldots, f_n$ are parallel to the corresponding bounded facets of P. Now fix a positive numbers a_i for each f_i. The conclusion of the theorem of Pogorolev is that, provided that there is (1) a polyhedron in $\Omega(P)$ with $\sigma(f_i) \leq a_i$ for all i, and (2) a plane so that any polyhedron in $\Omega(P)$ lying fully above it satisfies $\sum_{i=1}^n \sigma(f_i) \geq \sum_{i=1}^n a_i$, there exists a polyhedron in $\Omega(P)$ with $\sigma(f_i) = a_i$.

The parallel between Pogorelov's theorem and Theorem 3 is that power diagrams correspond to unbounded polyhedra and capacities correspond to the condition that

$\sigma(f_i) = a_i$, for all i. Once again, placement of sites determines the orientation of polyhedron facets. The most natural definition of σ would be as the measure of the projection of a facet onto the xy-plane, restricted to the unit square, so there is an exact correspondence between conclusions of the theorems. However, with this definition, σ does not satisfy some of the above technical conditions. Namely, σ need not be strictly positive for nondegenerate facets and need not strictly increase if a facet is enlarged. Finally, Pogorelov's theorem only constrains bounded facets, whereas Theorem 3 covers all facets. So, despite apparent similarity, neither of the two theorems immediately implies the other.

4. Computing Least-Squares Assignments. In this section we describe an algorithm that, for a set S of n sites and a set X of m points in the plane, computes a least-squares assignment $L\colon X \to S$ subject to a given integer capacity vector c with $\sum_{s\in S} c(s) = m$. By Lemma 1, it is sufficient to compute a weight vector $W = (w(s)_{s\in S}$ such that $|X \cap reg_W(s)| c(s)$ for all $s \in S$. The algorithm below computes such a weight vector in time $O(n^2 m \log m + nm \log^2 m)$ and optimal space $O(m)$ and, as a by-product, also determines the desired assignment $L = A_W$. Note that correctness of this algorithm implies the realizability of $L\colon X \to S$ by a power diagram, claimed in Theorem 1.

The algorithm starts with the weight vector $W = 0$, for which the power diagram is just the classical Voronoi diagram of S. It proceeds in m phases. During each phase, one point of X is inserted into the current diagram. W, and with it the power diagram, is then recomputed such that the invariant $b(s) \le c(s)$ for all $s \in S$ is maintained, where $b(s)$ denotes current number of points in $reg_W(s)$. More specifically, the algorithm carries out the following steps for each point x to be inserted:

1. Determine the region $reg_W(s)$ of the current power diagram containing x. Add x to the set of points contained in $reg_W(s)$. If $b(s) \le c(s)$ the phase ends—there is no need to change W. Otherwise, let $D = \{s\}$. Intuitively, D will contain the sites whose regions are too large and must be shrunk.
2. Repeat the following two steps:
   (a) Shrink all D-regions by simultaneously decreasing their weights. More formally, find the smallest positive number Δ so that decreasing the weights of all D-sites simultaneously by more than Δ causes one of the shrinking regions to lose a point, say p'. Notice that in the process a site in $S\backslash D$ cannot lose a point to a D-site, and that no point can move between two D-regions or between two non-D-regions.
   (b) Decrease the weights of all D-sites by Δ. Consider the region $reg(s')$ where p' would end up, had we shrunk the weights by more than Δ. If $b(s') < c(s')$, then go to 3, as we found a region which is not full. Else add s' to D and repeat (a).
3. We have found a region $reg(s')$ that is not full and a point p' on its boundary. Assign p' to s'. This makes some region $reg(s'')$ with $s'' \in D$ less than full. However, s'' was added to D because of some point p'' that it shared with site s''' that had already been in D. So assign p'' to s'' and follow the chain back, until the original site s is encountered and relieved of one point. This restores the invariant that was violated in the beginning of the current phase, and the phase ends.

To implement and analyze this procedure, we must specify how to store points belonging to a region, and how to detect the smallest weight change that makes a set of regions lose a point. We store the points of $reg(s)$ as a dynamic convex hull structure that allows $O(\log^2 m)$ time insertion and deletion. We use the data structure of Overmars and van Leeuwen [18] that can return, in logarithmic time, the two hull points of tangency for the two tangents to the hull with given slope. Each time D changes we recompute the power the diagram and determine the list of edges separating D-regions from the non-D-regions. Those are the $O(n)$ edges that will move by translation as Δ varies. For each edge, the convex hull data structure is used to determine the first time (i.e., the value of Δ) at which the line supporting the edge will strike a point contained in the D-region that it bounds. This requires $O(n \log m)$ time. The smallest such Δ is the one we are looking for. At this point, one region has shrunk so much as to lose a point. Check if the new region of nonfull. If it is, we are done—reshuffling $O(n)$ points clearly takes only $O(n)$ updates to the convex hulls (and thus $O(n \log^2 m)$ time) and the phase is complete. If not, the new region joins D and we again recompute the power diagram, identify moving edges, find the first time each edge hits a point in a D-region, etc. Growing D by one requires $O(n \log n + n \log m)$ time, hence one phase requires $O(n^2 \log m + n \log^2 m)$ time. As there are m phases, the running time claimed at the beginning of this section follows. The space requirement is dominated by the convex hull structure and is $O(m)$.

It is not necessary to recompute the power diagram anew after each shrinking step, as it can be maintained dynamically. However, we did not succeed in proving a better than $O(n^2)$ upper bound on the number of combinatorial changes the diagram undergoes during one phase of the algorithm. In fact, we suspect that the number of changes is $\Omega(n^2)$ in the worst case.

We already mentioned in the Introduction the connection of least-squares assignments to network flow problems. In the terminology of network flows, the chain-like process of reassigning points to sites at the end of a phase corresponds exactly to an augmenting path.

5. Iteratively Improving the Weight Vector. We now propose a method for iteratively improving the weight vector W. The method relies on the fact that, for a fixed set of sites, the value of the assignment induced by the power diagram with weights W is a concave function of W. Finding a weight vector such that the resulting assignment fulfills the capacity requirements is then equivalent to finding the maximum of a related function whose domain is the weight space. The method can be used to compute finite as well as continuous least-squares assignments. The continuous case is treated first.

Let ϱ be a continuous and nonvanishing probability distribution in $[0, 1]^d$. For an arbitrary but fixed assignment A: $[0, 1]^d \to S$, define the function f_A: $\mathbb{R}^n \to \mathbb{R}$ by

$$f_A(W) = \int_{[0,1]^d} \varrho(x) \cdot \mathrm{pow}_W(x, A(x))\, dx.$$

Let $B(A) = (\mu(A^{-1}(s)))_{s \in S}$ be the vector of capacities resulting from the assignment A, and put

$$Q(A) = \int_{[0,1]^d} \varrho(x) \cdot \delta^2(x, A(x))\, dx,$$

the value of A. With this notation, f_A can be written as

$$f_A(W) = -\langle B(A), W\rangle + Q(A),$$

where $\langle\cdot,\cdot\rangle$ denotes the inner product. Hence f_A is a linear function of W. Now consider the function $f\colon W \mapsto f_{A_W}(W)$; recall that A_W is the assignment induced by the power diagram with weights W. We claim that f is the pointwise minimum of the class of functions f_A. Indeed, for fixed W, the assignment A_W minimizes the value $f_A(W)$ by definition of the power diagram of S with weights W. In other words, the graph of f is the lower envelope of a set of hyperplanes in $\mathbb{R}^{n+1}$. Hence f is a concave function. By the choice of properties of ϱ, $B(A_W)$ and $Q(A_W)$ depend continuously on W. Hence, for each $W = W'$, the graph of f has at point $(W', f(W'))$ a unique tangent hyperplane $x_{n+1} = -\langle B(A'_W), W\rangle + Q(A'_W)$ that changes continuously with W. That is, f describes a smooth surface. Note that the gradient $\nabla f(W)$ of f at W is given by $-B(A_W)$.

Recall that we aim to find a weight vector W^* such that $B(A_{W^*}) = C$, the given capacity vector. Consider the function

$$\begin{aligned} g(W) &= f(W) + \langle C, W\rangle \\ &= \langle C - B(A_W), W\rangle + Q(A_W). \end{aligned}$$

Its gradient $\nabla g(W)$ is $C - B(A_W)$, hence our requirement $B(A_{W^*}) = C$ just means $\nabla g(W^*) = 0$. This corresponds to a global maximum of the smooth concave function g. So the problem we want to solve is: Find W^* such that $g(W^*)$ is maximized.

Finding the maximum of a concave and smooth n-variate function is a well-studied problem. In our case, we can exploit the fact that, for any given weight vector W, we can compute $g(W)$ and $\nabla g(W)$ from the power diagram with weights W. So a gradient method (see, e.g., [8]) for iteratively approximating W^* can be used. Starting, for example, with the weight vector $W_0 = 0$ (corresponding to the Voronoi diagram of S), we use the iteration scheme

$$W_{k+1} = W_k + t_k \nabla g(W_k).$$

If the step sizes t_k are chosen properly, then W_k converges to the solution W^* at a superlinear rate. Intuitively, what happens is that weights of sites whose region measures are too small (resp. large) are increased (resp. decreased) at each step.

If S is a set of n sites in the plane, and ϱ is the uniform distribution in the square, each step can be carried out in $O(n \log n)$ time. For the current weight vector W_k, we need $O(n \log n)$ time to construct the power diagram of S and W_k, and time $O(n)$ is needed in addition to calculate the area and the integral of squared distances for each region within the unit square. The space requirement is optimal, $O(n)$.

The method just described was inspired by the algorithm for "inverting" Minkowski's theorem, i.e., computing a three-dimensional polytope given by normals and areas of all its facets, proposed by Little [17].

In the case of a finite assignment $A_W\colon X \to S$, the graphs of the aforementioned functions f and g are lower envelopes of finitely many hyperplanes, and thus are concave polyhedral surfaces in $\mathbb{R}^{n+1}$. The gradient of g at W is orthogonal to the facet that lies vertically above W. The gradient is given by $C - B(A_W)$, where $B(A_W) = (|A_W^{-1}(s)|)_{s\in S}$

counts the numbers of points of X in the regions $reg_W(s)$. By Lemma 1, the number of hyperplanes defining g is equal to the number of different vectors $B(A)$ for all possible assignments A: $X \to S$, which is $\binom{m+n-1}{n-1}$ for $|X| = m$. Theorem 2 implies that the surface g actually realizes as many facets. Except for degenerate sets X, the maximum of g is attained by a facet; the set $\{W^* \mid \nabla g(W^*) = 0\}$ has dimension n.

Finding a maximum of g can now be seen as a linear programming problem. Its number of constraints is, however, at least exponential in n. On the other hand, this linear program has a very special structure; in Section 4, we have described a strongly polynomial algorithm for solving it.

Concerning the gradient method, the full-dimensionality of the maximum can be exploited in the choice of step sizes. We proceed as follows. Initially, an overestimate $\bar{g}$ of g is determined. Let $Q(A) = \sum_{x \in X} \delta^2(x, A(x))$. Since $g(W^*) = Q(A_{W^*})$, Lemma 1 implies that $\bar{g} = Q(A)$ will do for any assignment A with $B(A) = C$. We identify the horizontal hyperplane $\bar{H}$: $x_{n+1} = \bar{g}$ in $\mathbb{R}^{n+1}$ with the weight space. Let $B_k = \nabla g(W_k)$ denote the gradient of g at W_k. In geometric terms, the $(k+1)$st step now moves along the ray, r_k, from point $(W_k, g(W_k))$ in direction $(B_k, 1)$ until $\bar{H}$ is hit. This corresponds to the step size

$$t_k = \frac{\bar{g} - g(W_k)}{\langle B_k, B_k \rangle}$$

in the preceding iteration scheme. This step is iterated until either the maximum is reached, which means $B_{k+1} = C$, or the maximum is missed, meaning that $g(W_{k+1} + \varepsilon B_k) < g(W_{k+1})$ for all sufficiently small $\varepsilon > 0$, or, equivalently, $\langle B_{k+1}, B_k \rangle > 0$. If the latter happens, the overestimate $\bar{g}$ is lowered. Geometrically speaking, $\bar{H}$ is translated so as to pass through the point $r_k \cap H_{k+1}$, where H_{k+1} is the hyperplane through the facet below W_{k+1}. That is, $\bar{g}$ is taken such that

$$\frac{\bar{g} - g(W_k)}{\langle B_k, B_k \rangle} = \frac{\bar{g} - g(W_{k+1})}{\langle B_{k+1}, B_{k+1} \rangle}.$$

Starting once more at W_k, but now with the lower estimate, the iteration is continued. Note that—whether or not the estimate has to be lowered—the facet of g below W_{k+1} is different from the facet below W_k.

For sites and points in the plane, the cost for each step is $O(m \log n)$. $O(n \log n)$ time suffices for computing the power diagram and preprocessing it for point location, and $O(\log n)$ time is needed for locating each of the $m \geq n$ points in X. The space requirement is $O(m)$ which is optimal.

We have tested the method for the planar case.[5] For sets of 100 sites and 1000 points uniformly distributed in a square, the procedure always stopped after less than 10 steps. Due to numerical errors, however, only an approximation B_k of C was reached. We observed $\sum_{s \in S} |C(s) - B_k(s)| \approx n$, which indicates that a combination of this method with the insertion algorithm in Section 4 might lead to a fast exact procedure for computing the least-squares assignment. After identifying a good approximation W_k of W^*, a modified insertion algorithm should be started with W equal to W_k rather than

[5] Thanks to David Alberts for implementing the gradient algorithm and the insertion algorithm of Section 4, and to Juan Cuesta-Albertos for independently implementing the latter algorithm.

zero. Note that the insertion method of Section 4 cannot be used directly, as it assumes that no power region is ever filled beyond its capacity.

6. Some Applications. The constrained least-squares assignment problem, as considered in this paper, is a quite natural concept and has several applications. It is a special constrained optimum assignment problem for weighted bipartite graphs, which is a special case of the Hitchcock transportation problem; see, e.g., [22]. We first mention two interpretations of the problem. Throughout the section, let S and X be finite sets of sites and points in $\mathbb{E}^d$, respectively.

For $Y \subset X$ and $s \in S$, define the *variance* of the cluster Y with respect to the site s as $\sum_{x\in Y} \delta^2(x, s)$. Then a constrained least-squares assignment $L\colon X \to S$ is just a clustering for X such that the clusters have prescribed sizes and the sum of cluster variances is minimized. Besides being optimum in the above sense, these clusters have the important property that their convex hulls are pairwise disjoint: By Lemma 4, distinct clusters are contained in different regions of a power diagram, and power regions are convex. Hull-disjointness is a natural and desirable property of clusters which, for instance, eases the classification of new points. Simple examples show that replacing variance by the sum of distances destroys hull-disjointness.

If we define the *profit* of cluster Y with respect to site s as $\sum_{x\in Y}\langle x, s\rangle$, then L maximizes the sum of cluster profits for given cluster sizes: $\delta^2(x, s) = \langle x, x\rangle + \langle s, s\rangle - 2\langle x, s\rangle$, and the sum of the first two terms is independent of the assignment, provided capacity constraints are satisfied. This definition is motivated by the following transportation problem. interpret a point $x = (x_1, \ldots, x_d)$ as a truck loaded with x_i units of the ith good, and a site $s = (s_1, \ldots, s_d)$ as a market that sells the ith good at price s_i per unit. Choose the site capacities according to the attractiveness of the markets, and L will tell you where each truck should go in order to achieve maximal profit for these capacities.

The next application makes use of the property that constrained least-squares assignments are invariant under translation and scaling.

OBSERVATION 2. *Let $\sigma \in \mathbb{R}^+$ and $\tau \in \mathbb{E}^d$, and consider a least-squares assignment $L\colon X \to S$ with capacities c. Then $\sigma L + \tau$ is a least-squares assignment of X to $\sigma S + \tau$ subject to c.*

PROOF. L maximizes $\sum_{x\in X}\langle x, A(x)\rangle$ over all assignments A with capacities c. A least-squares assignment $L'\colon X \to \sigma S + \tau$ maximizes

$$\sum_{x\in X}\langle x, \sigma A(x) + \tau\rangle = \sigma \sum_{x\in X}\langle x, A(x)\rangle + \sum_{x\in X}\langle x, \tau\rangle.$$

Since the last sum does not depend on A and since $\sigma > 0$, L' must also maximize $\sum_{x\in X}\langle x, A(x)\rangle$. □

Consider the special case that S and X are of equal cardinality n and let $L\colon X \to S$ be a least-squares assignment subject to $c(s) = 1$ for all $s \in S$. L is called a *least-squares matching* in this case. Define a (*one-to-one*) *least-squares fitting* as the least-squares matching L_*: $X \to \sigma S + \tau$ such that the value of L_* is minimal over all positive scaling factors σ and all translation vectors τ. Observation 2 tells us that $L_*^{-1}(\sigma s + \tau) = L^{-1}(s)$ for all $s \in S$. Thus, when computing the least-squares fitting L_*, we can first calculate and fix the matching L, as a least-squares matching of X to S, and then determine the optimizing values of σ and τ for this matching.

Indeed, the latter task is easy when L is fixed. Let $S = \{s_1, \ldots, s_n\}$ and $L^{-1}(s_i) = x_i$. We want to find σ and $\tau = (\tau_1, \ldots, \tau_d)$ such that

$$Q(\sigma, \tau) = \sum_{i=1}^{n} \delta^2(x_i, \sigma s_i + \tau)$$

is minimal. Setting the partial derivatives Q_σ and Q_{τ_j}, $1 \le j \le d$, to zero shows that the minimum is achieved for

$$\sigma = \frac{an - \langle \alpha, \beta \rangle}{bn - \langle \beta, \beta \rangle}, \qquad \tau = \frac{1}{n}(\alpha - \sigma\beta),$$

where we have put

$$a = \sum \langle x_i, s_i \rangle, \qquad b = \sum \langle s_u, s_i \rangle,$$

and

$$\alpha = \sum x_i, \qquad \beta = \sum s_i.$$

$O(n)$ time suffices for calculating σ and τ if d is considered a constant. In the special cases $\sigma = 1$ (translation only) and $\tau = 0$ (scaling only) the minimum is achieved for $\tau = (1/n)(\alpha - \beta)$ and $\sigma = a/b$, respectively.

The insertion algorithm in Section 4 can be modified to run in time $O(n^3)$ and space $O(n)$ if all capacities are 1. In particular, no dynamic convex hull structure for the points of X within the current power regions is needed as each region can contain at most one point. This saves the $\log m$ factors in the runtime of the original algorithm, leading to time $O(n^2 \log n)$ per inserted point. The remaining $\log n$ factor stems from computing the power diagram anew in $O(n \log n)$ time after each reduction of weights for the so-called D-regions. Alternatively, the diagram can be updated dynamically in $O(n)$ time as follows.

For each power region R we store the partition of R that results from shrinking R to zero area. This is just the part of the order-2 power diagram within R and thus does not harm the $O(n)$-space complexity; see [6]. The boundary of the shrunk D-regions, as well as the extensions of the neighboring non-D-regions, can be found in $O(n)$ time by walking through the edges of the order-2 diagram. For reconstructing the order-2 diagram for the new weights, we use the $O(n)$-time deterministic algorithm of Aggarwal *et al.* [3], or the more practical $O(n)$-time randomized algorithm of Chew [11]. It suffices to apply the algorithm to those non-D-regions that expanded as the result of adjustment of weights, as the other non-D-regions remain unchanged, and the D-regions will contain the correct part of the order-2 diagram after having been shrunk.

For computing a least-squares fitting of m points to n sites with given capacities, the algorithm of Section 4 applies in its original form. The optimizing values σ and τ can be calculated in $O(m)$ time using respective generalizations of the formulas given above. We summarize:

LEMMA 5. *A (one-to-one) least-squares fitting of size n can be computed in* $O(n^3)$ *time and* $O(n)$ *space. A least-squares fitting of m points to n sites subject to given capacities can be computed in* $O(n^2 m \log m + nm \log^2 m)$ *time and* $O(m)$ *space.*

Observe also that the constrained assignment problem for m points and n sites can be reduced to matching in a trivial manner, by replicating each site according to its capacity. According to the fastest-known least-squares matching algorithm (the variant of Vaidya's algorithm mentioned in the Introduction), this yields an $O(m^{2+\epsilon})$-time and $O(m^{1+\epsilon})$-space algorithm for the least-squares assignment problem. Aside from being less practical than ours, this algorithm is not space-optimal, and is beaten in running time by our insertion method in Section 4 for $m > n^2$.

7. Concluding Remarks. In conclusion, we mention several open problems. Clearly, the existence of a better combinatorial algorithm for least-squares assignments is one of them. We would also like to be able to estimate the running time and increase numerical stability of the iterative procedure. Finally, the fastest least-squares fitting algorithm mentioned above (based on Vaidya's matching algorithm modified for power diagrams) requires tools which are difficult to implement. Is there a simpler algorithm with comparable performance?

Acknowledgments. The authors gratefully acknowledge numerous helpful discussions with P.K. Agarwal, H. Edelsbrunner, B. Gärtner, F. Johannes, H. Ranke, E. Welzl, and L. Wernisch. They also wish to thank I. Bárány for a useful pointer to Pogorelov's work, and J.-A. Cuesta-Albertos for drawing our attention to the Monge–Kantorovich mass transference problem.

References

[1] Agarwal, P.K. Private communication, 1992.

[2] Agarwal, P.K., Eppstein, D., Matoušek, J. Dynamic half-space reporting, geometric optimization, and minimum spanning trees. *Proc.* 33*rd Ann. IEEE Symp. on Foundations of Computer Science* (1992), pp. 80–89.

[3] Aggarwal, A., Guibas, L.J., Saxe, J., Shor, P.W. A linear time algorithm for the Voronoi diagram of a convex polygon. *Discrete Comput. Geom.* 4 (1989), 591–604.

[4] Alberts, D. A generalization of the Hungarian method for the *c*-assignment problem. Diplomarbeit, Freie Universität Berlin, 1992.

[5] Aurenhammer, F. A criterion for the affine equivalence of cell complexes in $\mathbb{R}^d$ and convex polyhedra in $\mathbb{R}^{d+1}$. *Discrete Comput. Geom.* 2 (1987), 49–64.

[6] Aurenhammer, F. Voronoi diagrams—a survey of a fundamental geometric data structure. *ACM Comput. Surv.* 23 (1991), 345–405.

[7] Aurenhammer, F., Hoffman, F., Aronov, B. Minkowski-type theorems and least-squares partitioning. *Proc.* 8*th ACM Symp. on Computational Geometry* (1992), pp. 350–357.

[8] Bertsekas, D.P. *Constrained Optimization and Lagrange Multiplier Methods*. Academic Press, New York, 1982.

[9] Boros, E., Hammer, P.L. On clustering problems with connected optima in Euclidean spaces. Tech. Report, RUTCOR, Rutgers University, New Brunswick, NJ, 1988.

[10] Capoyleas, V., Rote, G., Wöginger, G. Geometric clustering. *J. Algorithms* 12 (1991), 341–356.

[11] Chew, L.P. Building Voronoi diagrams for convex polygons in linear expected time. Tech. Report PCS-TR90-147, Dept. Math. Comput. Sci., Dartmouth College, Hanover, NH, 1986.

[12] Cuesta-Albertos, J.A., Tuero-Diaz, A. A characterization for the solution of the Monge–Kantorovich mass transference problem. *Statist*. *Probab. Lett*. 16 (1993), 147–152.

[13] Fredman, M., Tarjan, R. Fibonacci heaps and their uses in improved network optimization algorithms. *J*. *Assoc*. *Comput*. *Mach*. 34 (1987), 596–615.

[14] Grünbaum, B. *Convex Polytopes*. Interscience, New York, 1967.

[15] Inaba, M., Katoh, N., Imai, H. Applications of weighted Voronoi diagrams and randomization to variance-based *k*-clustering. *Proc*. 10*th ACM Symp*. *on Computational Geometry* (1994), pp. 332–339.

[16] Lawler, E. *Combinatorial Optimization*: *Networks and Matroids*. Holt Rinehart and Winston, New York, 1976.

[17] Little, J.J. Extended gaussian images, mixed volumes, and shape reconstruction. *Proc*. 1*st ACM Symp*. *on Computational Geometry* (1985), pp. 15–23.

[18] Overmars, M.H., van Leeuwen, J. Maintenance of configurations in the plane. *J*. *Comput*. *System Sci*. 23 (1981), 166–204.

[19] Pogorelov, A.V. *Extrinsic Geometry of Convex Surfaces*. American Mathematical Society, Providence, RI, 1973. (Translation of Pogorelov, A.V. VnexnØØ geometriØ vypuklyh poverhnost$ Izd. Nauka, Moskva, 1969.)

[20] Tarjan, R.E. *Data Structures and Network Algorithms*. SIAM, Philadelphia, PA, 1987.

[21] Tokuyama, T., Nakano, J. Geometric algorithms for a minimum cost assignment problem. *Proc*. 7*th Ann*. *ACM Symp*. *on Computational Geometry* (1991), pp. 262–271.

[22] Tokuyama, T., Nakano, J. Efficient algorithms for the Hitchcock transportation problem. *Proc*. *Symp*. *on Data Structures and Algorithms* (1992), pp. 175–184.

[23] Vaidya, P. Geometry helps in matching. *SIAM J*. *Comput*. 18 (1989), 1201–1225.