Discrete Comput Geom 6:343-367 (1991)

Discrete & Computational
Geometry

Higher-Dimensional Voronoi Diagrams in Linear Expected Time*

Rex A. Dwyer

Department of Computer Science, North Carolina State University,
Raleigh, NC 27695-8206, USA

Abstract. A general method is presented for determining the mathematical expectation of the combinatorial complexity and other properties of the Voronoi diagram of n independent and identically distributed points. The method is applied to derive exact asymptotic bounds on the expected number of vertices of the Voronoi diagram of points chosen from the uniform distribution on the interior of a d-dimensional ball; it is shown that in this case, the complexity of the diagram is $\Theta(n)$ for fixed d. An algorithm for constructing the Voronoi diagram is presented and analyzed. The algorithm is shown to require only $\Theta(n)$ time on average for random points from a d-ball assuming a real-RAM model of computation with a constant-time floor function. This algorithm is asymptotically faster than any previously known and optimal in the average-case sense.

Introduction

The Voronoi diagram is a natural and intuitively appealing structure, repeatedly reinvented by researchers in several fields. While computer scientists generally name it for the mathematician Voronoi [32], meteorologists associate the two-dimensional version with the name Thiessen [30], and physicists honor Wigner and Seitz [35] for the three-dimensional version. It has been used by geologists, foresters, agriculturalists, medical researchers, geographers, crystallographers, and astronomers. Within the domain of the mathematical sciences, it is applied to simulate differential equations by finite-element methods, to interpolate surfaces in geometric modeling systems, and to solve geometric problems such as finding Euclidean minimum spanning trees and largest empty circles. (Avis and Bhattacharya [2] present an extensive list of references for applications.)

* Based upon work supported by the National Science Foundation under Grant No. CCR-8658139 while the author was a student at Carnegie-Mellon University.

The Voronoi diagram of a set of points—called *sites*—is a partition of $\mathbf{R}^d$ that assigns a surrounding polytope of "nearby" points to each of the sites. Each region is a d-polytope containing the points lying nearer to the site in its interior than to any other site. More rigor is supplied by the following definition.

Definition 1. The (*nearest-site*) *Voronoi diagram* of the set $\mathcal{X}_n = \{x_1, x_2, \ldots, x_n\}$ of n sites in $\mathbf{R}^d$ is the set of n convex regions $\mathcal{V}_i = \{x | \forall j\colon \text{dist}(x, x_i) \leq \text{dist}(x, x_j)\}$ for $1 \leq i \leq n$.

The straight-line dual of the Voronoi diagram in the plane is called the *Delaunay triangulation.* In the planar case, sites x_i and x_j are joined by an edge in the Delaunay triangulation if and only if $\mathcal{V}_i$ and $\mathcal{V}_j$ share an edge. In d dimensions, sites $x_{i_0}, x_{i_1}, \ldots, x_{i_k}$ define a k-face of the Voronoi dual if and only if $\mathcal{V}_{i_0}, \mathcal{V}_{i_1}, \ldots, \mathcal{V}_{i_k}$ share a $(d-k)$-face in the Voronoi diagram. The Voronoi diagram can be constructed easily from its dual and *vice versa.* If, as is assumed in the following, no $d+2$ sites fall on the same hypersphere, the facts below are easily verified.

Fact 1. $(d+1)$ sites are vertices of a d-simplex in the Voronoi dual if and only if the inside of the hypersphere passing through these sites contains no other sites. (We call such a hypersphere *empty* or *site-free.*)

Fact 2. $(k+1)$ sites are vertices of a k-face of some Voronoi dual simplex if and only if some empty hypersphere passes through these sites.

Fact 3. Every convex-hull face is a face of some Voronoi dual simplex.

Fact 4. The Voronoi dual partitions the convex hull of $\mathcal{X}_n$ into d-simplices.

We may also define a *furthest-site* Voronoi diagram.

Definition 2. The *furthest-site Voronoi diagram* of the set $\mathcal{X}_n = \{x_1, x_2, \ldots, x_n\}$ of n sites in $\mathbf{R}^d$ is the set of n convex regions $\mathcal{W}_i = \{x | \forall j\colon \text{dist}(x, x_i) \geq \text{dist}(x, x_j)\}$ for $1 \leq i \leq n$.

Region $\mathcal{W}_i$ contains the points lying further from site x_i than from any other site. Only sites that are vertices of the convex hull of $\mathcal{X}_n$ have nonempty furthest-site Voronoi regions. The dual of the furthest-site Voronoi diagram also partitions the convex hull of $\mathcal{X}_n$ into simplices; the vertices of these simplices determine hyperspheres that each contain *all* the sites of $\mathcal{X}_n$.

Many researchers have considered Voronoi-diagram construction. Previous work on the two-dimensional case has been surveyed elsewhere [10]. Browstow *et al.* [8], Finney [14], and Tanemura *et al.* [28] have addressed the construction of three-dimensional diagrams; none of these authors present detailed analyses of running time.

Bowyer [6] and Watson [33] describe algorithms for higher dimensions, but

neither analyzes his algorithm rigorously. Bowyer gives some theoretical and empirical evidence suggesting that his algorithm requires $O(n^{1+1/d})$ time on average for points uniform in a d-dimensional hypercube. Watson claims $O(n^{2-1/d})$ time in the worst case to find the edges of the dual. Avis and Bhattacharya's algorithms for the Voronoi diagram and its dual [2] rely heavily on the simplex method for linear programming; since this method has exponential worst-case running time, they focus mainly on experimental studies of their algorithms' performance and of the expected complexity of the diagrams for points distributed uniformly in the unit hypercube.

A pleasing connection between nearest- and furthest-site Voronoi diagrams in d dimensions and convex hulls in $d+1$ dimensions allows any convex-hull algorithm to be used to construct Voronoi diagrams. This connection, first observed by Brown [7], is restated here in a form due to Edelsbrunner and Seidel [12]. If a is a d-vector and b is scalar, let $a * b$ be the $(d+1)$-vector $(a^{(1)}a^{(2)}, \ldots, a^{(d)}, b)$, and let λ: $\mathbf{R}^d \to \mathbf{R}^{d+1}$ be the "lifting function" defined by $\lambda(x) = x * \langle x, x \rangle$, where $\langle x, y \rangle$ denotes the inner product. The range of λ is the surface of a $(d+1)$-dimensional paraboloid of revolution. The image of a d-ball is the intersection of a half-space with the paraboloid and *vice versa*. This is easily verified by applying the bilinearity of the inner product. Now suppose that the function λ is applied to the points of $\mathcal{X}_n$ and the convex hull of their images is constructed. Then the simplices of the duals of the nearest- and furthest-site Voronoi diagrams are in one-to-one correspondence with the facets of the convex hull of the transformed points.

Constructing a $(d+1)$-dimensional convex hull is a viable approach to the problem of constructing a d-dimensional Voronoi diagram. For example, the gift-wrapping algorithm [5], [9], [27] may be used in $\Theta(n(S_n + S_n^*))$ time. Or Seidel's shelling algorithm [25] may be used in $O(n^2 + (S_n + S_n^*)\log n)$, where S_n is the number of nearest-site simplices and S_n^* is the number of furthest-site simplices in the result. In fact, it is not difficult to modify either algorithm to eliminate the S_n^* (S_n) term if only the nearest-site (furthest-site) diagram is required.

Like the number of facets in the case of convex hulls, S_n and S_n^* can vary wildly. Seidel [24], [26] has shown that both S_n and S_n^* can be extremely large—$\Theta(n^{\lfloor (d+1)/2 \rfloor})$—in the worst case. On the other hand, it is not difficult to construct families of problem instances for which $S_n = \Theta(n)$. Thus probabilistic estimates of the average value of the two quantities are useful.

Meijering [18] and Gilbert [15] have considered the Voronoi diagram of sites from a Poisson process of fixed intensity in $\mathbf{R}^d$; Meijering showed that the expected number of nearest-site Voronoi neighbors of a site depends only on d; in particular, it is 6 for $d = 2$ and ≈ 15.54 for $d = 3$. Such a set of sites may be thought of as an infinite set of sites drawn from a uniform distribution over all of $\mathbf{R}^d$. In computational practice, however, we must deal with finite sets of sites drawn from a particular distribution, e.g., a uniform distribution on the interior of some convex body like a hypercube or hypersphere. Two sites are neighbors if and only if they lie on the surface of some ball that contains no other site. In the Poisson case, a pair of distant neighbors is always unlikely since it implies the existence of a large empty ball. In the case of a bounded set of sites, it can still be shown that sites far from the boundary of the body probably have only nearby neighbors, but some pairs of

distant neighbors always occur near the boundary of the body, where most of the empty ball may lie outside the support of the distribution. Thus results dealing only with the Poisson case are inadequate for the average-case analysis of algorithms.

In the next section we present a general method for determining ES_n and ES_n^*, the expected number of simplices in the dual of the nearest- and furthest-site diagrams. In Section 2 this method is applied to the analysis of the asymptotic behavior of ES_n for sites drawn independently from the uniform distribution in the unit d-ball. In Section 3 we present a variation of the gift-wrapping algorithm that uses standard bucketing techniques. In Section 4 we apply the methods of Section 1 to show that the algorithm requires only linear time on average for independent and uniformly distributed points in the unit d-ball. The last section mentions some open problems and possible extensions to the work.

1. A General Method for Bounding the Expected Complexity of Voronoi Diagrams

In this section we describe a general method for bounding ES_n and ES_n^*, the expected number of simplices in the duals of nearest- and furthest-site Voronoi diagrams of random point sets. The approach taken is similar to that of the seminal work of Rényi and Sulanke [22], [23] and Efron [13] to evaluate convex-hull expectations.

The first $d+1$ points $x_1, \ldots, x_{d+1}$ define a d-simplex with probability one. Let us first reckon the probability P_n that they also define a simplex in the dual of the nearest-site Voronoi diagram. This is just the probability that the other $n-d-1$ points lie outside the hypersphere passing through the $d+1$ points. Writing $g(\cdot)$ for the density function of the x_i and Γ for the probability content of interior of the hypersphere, we see that this probability is

$$P_n = \int_{\mathbf{R}^d} \cdots \int_{\mathbf{R}^d} (1-\Gamma)^{n-d-1} g(x_1) \cdots g(x_{d+1})\, dx_1 \cdots dx_{d+1},$$

and that the expected number of simplices is therefore

$$\begin{aligned} ES_n &= \binom{n}{d+1} P_n \\ &= \binom{n}{d+1} \int_{\mathbf{R}^d} \cdots \int_{\mathbf{R}^d} (1-\Gamma)^{n-d-1} g(x_1) \cdots g(x_{d+1})\, dx_1 \cdots dx_{d+1}. \end{aligned}$$

We next carry out a transformation of coordinates. The $d+1$ points x_1, $x_2, \ldots, x_{d+1}$ can be expressed in terms of a d-vector p representing the center of the d-sphere they define, a scalar r representing the radius of that sphere, and $d-1$ angles $\psi_{i1}, \psi_{i2}, \ldots, \psi_{i,d-1}$ for each x_i. The angles satisfy $0 \le \psi_{i1}, \psi_{i2}, \ldots, \psi_{i,d-2} \le \pi$ and $0 \le \psi_{i,d-1} \le 2\pi$. Let $y_i = x_i - p$. Then the two systems of coordinates are

related by the equations

$$
\begin{aligned}
x_i^{(1)} &= p^{(1)} + y_i^{(1)} = p^{(1)} + rc_{i,d-1}c_{i,d-2}\cdots c_{i3}c_{i2}c_{i1},\\
x_i^{(2)} &= p^{(2)} + y_i^{(2)} = p^{(2)} + rc_{i,d-1}c_{i,d-2}\cdots c_{i3}c_{i2}s_{i1},\\
x_i^{(3)} &= p^{(3)} + y_i^{(3)} = p^{(3)} + rc_{i,d-1}c_{i,d-2}\cdots c_{i3}s_{i2},\\
&\vdots\\
x_i^{(d)} &= p^{(d)} + y_i^{(d)} = p^{(d)} + rs_{i,d-1},
\end{aligned}
$$

where c_{ij} and s_{ij} represent $\cos\psi_{ij}$ and $\sin\psi_{ij}$, respectively.

The Jacobian of this transformation is given without proof by Miles [20], and Affentranger [1] has recently given a derivation. However, it is worthwhile to give another derivation here, since both assume a more-than-passing familiarity with the concepts and notations of integral geometry. In three dimensions, the Jacobian, expressed in tabular form, is

	$p^{(1)}$	$p^{(2)}$	$p^{(3)}$	r	ψ_{11}	ψ_{12}	ψ_{21}	ψ_{22}	ψ_{31}	ψ_{32}	ψ_{41}	ψ_{42}
$x_1^{(1)}$	1	0	0	$y_1^{(1)}/r$	$t'_{11}y_1^{(1)}$	$t'_{12}y_1^{(1)}$	0	0	0	0	0	0
$x_2^{(1)}$	1	0	0	$y_2^{(1)}/r$	0	0	$t'_{21}y_2^{(1)}$	$t'_{22}y_2^{(1)}$	0	0	0	0
$x_3^{(1)}$	1	0	0	$y_3^{(1)}/r$	0	0	0	0	$t'_{31}y_3^{(1)}$	$t'_{32}y_3^{(1)}$	0	0
$x_4^{(1)}$	1	0	0	$y_4^{(1)}/r$	0	0	0	0	0	0	$t'_{41}y_4^{(1)}$	$t'_{42}y_4^{(1)}$
$x_1^{(2)}$	0	1	0	$y_1^{(2)}/r$	$k_{11}y_1^{(2)}$	$t'_{12}y_1^{(2)}$	0	0	0	0	0	0
$x_1^{(3)}$	0	0	1	$y_1^{(3)}/r$	0	$k_{12}y_1^{(3)}$	0	0	0	0	0	0
$x_2^{(2)}$	0	1	0	$y_2^{(2)}/r$	0	0	$k_{21}y_2^{(2)}$	$t'_{22}y_2^{(2)}$	0	0	0	0
$x_2^{(3)}$	0	0	1	$y_2^{(3)}/r$	0	0	0	$k_{22}y_2^{(3)}$	0	0	0	0
$x_3^{(2)}$	0	1	0	$y_3^{(2)}/r$	0	0	0	0	$k_{31}y_3^{(2)}$	$t'_{32}y_3^{(2)}$	0	0
$x_3^{(3)}$	0	0	1	$y_3^{(3)}/r$	0	0	0	0	0	$k_{32}y_3^{(3)}$	0	0
$x_4^{(2)}$	0	1	0	$y_4^{(2)}/r$	0	0	0	0	0	0	$k_{41}y_4^{(2)}$	$t'_{42}y_4^{(2)}$
$x_4^{(3)}$	0	0	1	$y_4^{(3)}/r$	0	0	0	0	0	0	0	$k_{42}y_4^{(3)}$

where $t'_{ij} = -\tan\psi_{ij}$ and $k_{ij} = \cot\psi_{ij}$. The generalization to higher dimensions is straightforward.

If the row for $x_i^{(1)}$, denoted by $\rho_i^{(1)}$, is replaced by $\sigma_i = \sum_{1\le j\le d}(y_i^{(j)}/y_i^{(1)})\rho_i^{(j)}$, it becomes

$$
y_i^{(1)}/y_i^{(1)} \quad y_i^{(2)}/y_i^{(1)} \quad y_i^{(3)}/y_i^{(1)} \quad r/y_i^{(1)} \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0
$$

and the matrix is in quasi-triangular form. The entire determinant is the product of a $(d+1)\times(d+1)$ determinant in the upper left and a $(d^2-1)\times(d^2-1)$ determinant in the lower right. The lower-right determinant can be decomposed into $d+1$ similar $(d-1)\times(d-1)$ determinants. These are upper triangular, and the ith is easily seen to be

$$
r^{d-1}c_{i,d-1}^{d-1}c_{i,d-2}^{d-2}\cdots c_{i1}^{1} = y_i^{(1)}r^{d-2}c_{i,d-1}^{d-2}c_{i,d-2}^{d-3}\cdots c_{i2}^{1}c_{i1}^{0}.
$$

If the factors of $(1/y_i^{(1)})$ are removed from the rows of the upper left matrix, and then the factor of r is removed from the last column, the remaining determinant is

$$D = \begin{vmatrix} y_1^{(1)} & y_1^{(2)} & y_1^{(3)} & 1 \\ y_2^{(1)} & y_2^{(2)} & y_2^{(3)} & 1 \\ y_3^{(1)} & y_3^{(2)} & y_3^{(3)} & 1 \\ y_4^{(1)} & y_4^{(2)} & y_4^{(3)} & 1 \end{vmatrix}$$

in the three-dimensional case, and in general its magnitude is just $d!$ times the volume of the simplex formed by $x_1, x_2, \ldots, x_{d+1}$, denoted by $\mathrm{simp}(x_1, x_2, \ldots, x_{d+1})$. Thus the determinant of the original Jacobian matrix is

$$\mathrm{sgn}(D) d! \, \mathrm{simp}(x_1, x_2, \ldots, x_{d+1}) r^{(d+1)(d-2)+1} \prod_{1 \le i \le d+1} c_{i,d-1}^{d-2} c_{i,d-2}^{d-3} \cdots c_{i2}^{1}.$$

Since this is essentially a probability density, it must be positive, so

$$P_n = d! \int_{\mathbf{R}^d} \int_0^\infty r^{-d} (1 - \Gamma)^{n-d-1} \times \left[\underbrace{\int_0^\pi \cdots \int_0^{2\pi}}_{(d+1)(d-1)} h(x_1) \cdots h(x_{d+1}) \, \mathrm{simp}(x_1, \ldots, x_{d+1}) \, d\psi_{1,1} \cdots d\psi_{d+1,d-1} \right] dr \, dp \tag{1.1}$$

with

$$h(x_i) = |r^{d-1} c_{i,d-1}^{d-2} c_{i,d-2}^{d-3} \cdots c_{i2}^{1} g(x_i)|.$$

If we define

$$\hat{g}(r, p) = \underbrace{\int_0^\pi \cdots \int_0^{2\pi}}_{(d-1)} h(x_1) \, d\psi_{1,1} \cdots d\psi_{1,d-1},$$

then, since x_1 through x_{d+1} are i.i.d.,

$$\underbrace{\int_0^\pi \cdots \int_0^{2\pi}}_{(d+1)(d-1)} h(x_1) \cdots h(x_{d+1}) d\psi_{1,1} \cdots d\psi_{d+1,d-1} = (\hat{g}(r, p))^{d+1}. \tag{1.2}$$

Since the quotient of the bracketed quantity of (1.1) and (1.2) is

$$\mathrm{esimp}(r, p) := E(\mathrm{simp}(x_1, x_2, \ldots, x_{d+1}) \mid \| x_i - p \| = r \text{ for } 1 \le i \le d+1),$$

the bracketed quantity equals $(\hat{g}(r, p))^{d+1}$ esimp(r, p), and

$$P_n = d! \int_{\mathbf{R}^d} \int_0^\infty r^{-d}(1 - \Gamma)^{n-d-1}(\hat{g}(r, p))^{d+1}\text{esimp}(r, p)\, dr\, dp.$$

It is clear that $\hat{g}$, Γ, and esimp depend only on r and $\|p\|$ if g is spherically symmetric. To exploit this symmetry, we express p in generalized spherical coordinates $(q, \theta_1, \theta_2, \ldots, \theta_{d-1})$ defined by

$$\begin{aligned}
\|p\| &= q, \\
p^{(1)} &= qc_{d-1}c_{d-2}\cdots c_3c_2c_1, \\
p^{(2)} &= qc_{d-1}c_{d-2}\cdots c_3c_2s_1, \\
p^{(3)} &= qc_{d-1}c_{d-2}\cdots c_3s_2, \\
&\vdots \\
p^{(d)} &= qs_{d-1},
\end{aligned}$$

where c_i and s_i represent $\cos\theta_i$ and $\sin\theta_i$. The Jacobian of this transformation is known to be $q^{d-1}c_{d-1}^{d-2}c_{d-2}^{d-3}\cdots c_2^1$ [16, p. 17], thus P_n is

$$d!\Bigg(\underbrace{\int_0^\pi \cdots \int_0^\pi}_{(d-2)} \int_0^{2\pi} c_{d-1}^{d-2}c_{d-2}^{d-3}\cdots c_2^1 d\theta_1 \cdots d\theta_{d-1}\Bigg)$$
$$\times \left(\int_0^\infty \int_0^\infty q^{d-1}r^{-d}(1 - \Gamma)^{n-d-1}\hat{g}^{d+1}\text{esimp}\, dr\, dq\right).$$

The integral in the θ_i can be shown to be $\kappa_d = (2\pi^{d/2})/\Gamma(d/2)$, the surface area of the unit d-ball, by repeatedly applying a standard definite-integral identity [4, p. 287]. We estimate $(1 - \Gamma)^{n-d-1} \approx \exp(-n\Gamma)$; to be strictly correct, we should apply the standard inequalities [34, p. 242]

$$\exp(-n\Gamma)(1 - n\Gamma^2) \le (1 - \Gamma)^n \le \exp(-n\Gamma)$$

to derive

$$\exp(-n\Gamma)(1 - n\Gamma^2) \le (1 - \Gamma)^{n-d-1} \le \exp(-(n - d - 1)\Gamma)$$

and then perform separate calculations to show that the upper and lower bounds give asymptotically equivalent results. Finally, since

$$\binom{n}{d+1} \sim \frac{n^{d+1}}{(d+1)!},$$

$$\begin{aligned}
ES_n &\sim \frac{\kappa_d n^{d+1}}{d+1} \cdot \int_0^\infty \int_0^\infty q^{d-1}r^{-d}\hat{g}^{d+1}\text{esimp}\exp(-n\Gamma)\, dr\, dq \\
&= \frac{\kappa_d n^{d+1}}{d+1} \cdot \int_0^\infty \int_0^\infty I(q, r)\, dr\, dq. \qquad (1.3)
\end{aligned}$$

Similarly

$$ES_n^* \sim \frac{\kappa_d n^{d+1}}{d+1} \cdot \int_0^\infty \int_0^\infty q^{d-1} r^{-d} \hat{g}^{d+1} \operatorname{esimp} \Gamma^{n-d-1} \, dr \, dq.$$

2. Bounds for the Uniform Distribution in a d-Ball

In this section we turn to the uniform distribution in the unit d-ball in particular and prove the following theorem.

Theorem 1. *Let $\mathcal{X}_n = \{X_1, X_2, \ldots, X_n\}$ be a set of n sites drawn independently from the uniform distribution on the interior of the unit d-ball. Then ES_n, the expected number of simplices of the dual of the Voronoi diagram of $\mathcal{X}_n$, is $\Theta(n)$ for fixed d as $n \to \infty$.*

We actually show that $ES_n \sim C_d n$. and give C_d exactly in (2.3).

Proof. Recall that κ_d is the surface area of the unit d-ball. We write $\mu_d = \kappa_d/d$ for its volume, and v_d for the expected volume of a d-simplex with random vertices on the unit d-sphere. (The exact value of v_d is given later.)

Let $\mathcal{U}$ denote the unit d-ball, $\mathcal{B}$ the ball defined by the points x_1 through x_{d+1}, and $\partial\mathcal{B}$ the surface of $\mathcal{B}$. We have $g(x) = 1/\mu_d$ when $\|x\| \leq 1$ and $g(x) = 0$ otherwise. Thus probabilities and densities are proportional to volumes and surface areas of intersections with the unit ball $\mathcal{U}$. In particular,

$$\begin{aligned} \Gamma &= \mu_d^{-1} \operatorname{vol}(\mathcal{B} \cap \mathcal{U}), \\ \hat{g} &= \mu_d^{-1} \operatorname{vol}_{d-1}(\partial\mathcal{B} \cap \mathcal{U}), \\ \operatorname{esimp} &\leq \operatorname{vol}(\operatorname{conv}(\partial\mathcal{B} \cap \mathcal{U})). \end{aligned}$$

The last bound is based on the observation that the convex hull of points chosen on $\partial\mathcal{B} \cap \mathcal{U}$ is contained in the convex hull of $\partial\mathcal{B} \cap \mathcal{U}$. The crude but useful bounds

$$\begin{aligned} \hat{g} &\leq \mu_d^{-1} \kappa_d r^{d-1} = O(r^{d-1}), \\ \hat{g} &\leq \mu_d^{-1} \kappa_d = O(1), \\ \operatorname{esimp} &\leq \mu_d r^d = O(r^d), \end{aligned}$$

and

$$\operatorname{esimp} \leq \mu_d = O(1)$$

are immediate.

We proceed by dividing the domain of integration of the integral of (1.3) into eight regions corresponding to the possible patterns of intersection of the two balls $\mathcal{B}$ and $\mathcal{U}$. We apply different estimates of Γ, $\hat{g}$, and esimp in each region.

Case 1: $q \le 1$ and $0 \le r \le 1 - q$. In this case, $\mathscr{B} \subset \mathscr{U}$,

$$\hat{g} = \frac{\kappa_d r^{d-1}}{\mu_d} = dr^{d-1} \quad \text{and} \quad \Gamma = \frac{\mu_d r^d}{\mu_d} = r_d.$$

Also

$$\text{esimp} = v_d r^d.$$

Thus

$$\begin{aligned}\int_0^1 \int_0^{1-q} I(q, r)\, dr\, dq &\sim d^{d+1} v_d \int_0^1 \int_0^{1-q} q^{d-1} r^{d^2-1} \exp(-nr^d)\, dr\, dq \\ &\sim d^{d+1} v_d \left(\int_0^1 q^{d-1}\, dq \right) \left(\frac{1}{d} \int_0^{n(1-q)^d} \left(\frac{t}{n} \right)^d \frac{e^{-t}}{t}\, dt \right) \\ &\sim d^{d-2} d!\, v_d n^{-d}.\end{aligned}$$

Case 2: $q \le 1$ and $1 - q \le r \le \sqrt{1 - q^2}$. In this case (Fig. 1) at least half of $\mathscr{B}$ and $\partial\mathscr{B}$ lie inside $\mathscr{U}$, and

$$\Gamma \ge \frac{\mu_d r^d}{2\mu_d} = \frac{r^d}{2}.$$

By the crude bounds $\hat{g} = O(r^{d-1})$ and esimp $= O(r^d)$ along with Tricomi's formula $\int_x^\infty t^{-a} e^{-t}\, dt \sim x^{-a} e^{-x}$ for the incomplete gamma function [31, Section 4.3],

$$\begin{aligned}\int_0^1 \int_{1-q}^{\sqrt{1-q^2}} I(q, r)\, dr\, dq &= O(1) \cdot \int_0^1 \int_{1-q}^{\sqrt{1-q^2}} q^{d-1} r^{d^2-1} \exp\left(\frac{-nr^d}{2} \right) dr\, dq \\ &= O(1) \cdot \int_0^1 q^{d-1} \left(\int_{n(1-q)^d}^\infty \left(\frac{t}{n} \right)^d \frac{e^{-t}}{t}\, dt \right) dq \qquad (2.1) \\ &= O(n^{-d}) \cdot \int_0^1 q^{d-1} (n(1-q)^d)^{d-1} \exp(-n(1-q)^d)\, dq \\ &= O(n^{-d}) \cdot \int_0^n n^{-1/d} u^{d-2+(1/d)} e^{-u}\, du = O(n^{-d-(1/d)}).\end{aligned}$$

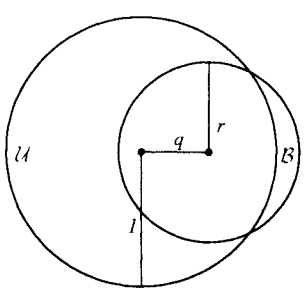

Fig. 1. Case 2: $q \le 1$ and $1 - q \le r \le \sqrt{1 - q^2}$.

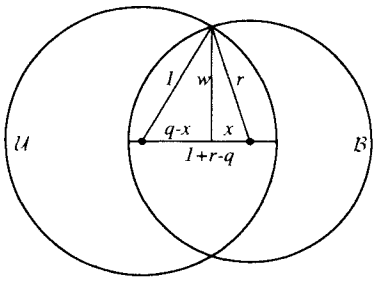

Fig. 2. Case 3: $q \leq 1$ and $\sqrt{1-q^2} \leq r \leq 1$.

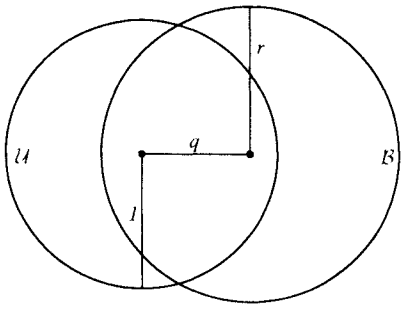

Fig. 3. Case 4: $q \leq 1$ and $r \geq 1$.

Case 3: $q \leq 1$ *and* $\sqrt{1-q^2} \leq r \leq 1$. Referring to Fig. 2, where w is the radius of the $(d-1)$-sphere $\partial\mathscr{B} \cap \partial\mathscr{U}$ and x is the distance between the centers of $\mathscr{B}$ and $\partial\mathscr{B} \cap \partial\mathscr{U}$, we see that the volume of the bipyramid with base $\partial\mathscr{B} \cap \partial\mathscr{U}$ and height $1 + r - q$ gives a lower bound on Γ. From the triangle inequality $(r - x) \leq w \leq r$. By first solving the equation $1 - (q - x)^2 = r^2 - x^2$ for x, it is easy to verify that

$$(r - x) = \frac{(q + 1 - r)(r + 1 - q)}{2q}.$$

Since $r \geq \sqrt{1-q^2} \geq 1 - q$, it follows that

$$r \leq (r + 1 - q) \leq 2r,$$
$$q \leq (q + 1 - r) \leq 2q;$$

thus $(r - x) = \Theta(r)$ and also $w = \Theta(r)$. This implies that

$$\Gamma \geq \frac{(r + 1 - q)\mu_{d-1} w^{d-1}}{d\mu_d} = \Theta(r^d).$$

Applying the crude bounds $\hat{g} = O(r^{d-1})$ and esimp $= O(r^d)$ gives

$$\int_0^1 \int_{\sqrt{1-q^2}}^1 I(q, r)\, dr\, dq = O(1) \int_0^1 \int_{\sqrt{1-q^2}}^1 q^{d-1} r^{-d} r^{(d+1)(d-1)/2} r^d \exp(-nr^d)\, dr\, dq$$
$$= O(1) \int_0^1 q^{d-1} \left(\int_{n(1-q^2)^{d/2}}^n \left(\frac{t}{n}\right)^d e^{-t} \frac{dt}{t} \right) dq;$$

this is clearly dominated by (2.1) of Case 2.

Case 4: $q \leq 1$ *and* $r \geq 1$. In this case (Fig. 3) we apply $\hat{g} = O(1)$, esimp $= O(1)$, and

$$\text{vol}(\mathscr{B} \cap \mathscr{U}) \geq \text{vol}(\mathscr{U} \cap \{x | x^{(1)} \geq \tfrac{1}{2}\}) = \Omega(1),$$

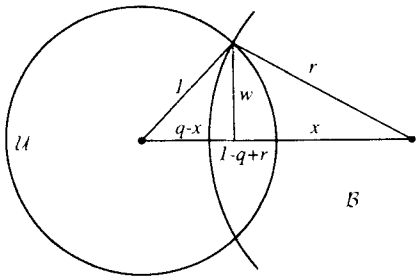

Fig. 4. Case 6: $q \le 1$ and $q - 1 \le r \le q$.

giving $\Gamma = \Omega(1)$ and

$$\int_0^1 \int_1^\infty I(q, r)\, dr\, dq \le O(1) \int_0^1 \int_1^\infty q^{d-1} r^{-d} \exp(-\Omega(n))\, dr\, dq = O(e^{-\Omega(n)}).$$

Case 5: $q > 1$ and $0 \le r \le q - 1$. In this case $\mathscr{B} \cap \mathscr{U} = \varnothing$, $\hat{g} = \Gamma = \text{esimp} = 0$, and the integral vanishes.

Case 6: $q > 1$ and $q - 1 \le r \le q$. Let w be the radius of $\partial\mathscr{B} \cap \partial\mathscr{U}$, and let x be the distance between the centers of $\partial\mathscr{B} \cap \partial\mathscr{U}$ and $\mathscr{B}$ (Fig. 4) as in Case 3. We can bound Γ/μ_d above and below by the volumes of a cylinder and a bipyramid, respectively, each with base $\text{vol}_{d-1}(\partial\mathscr{B} \cap \partial\mathscr{U}) = \mu_{d-1} w^{d-1}$ and height $h = 1 - q + r$. esimp can be bounded above by a cylinder with the same base but height $r - x$. Since $1 - (q - x)^2 = w^2 = r^2 - x^2$, we have

$$x = \frac{r^2 + q^2 - 1}{2q} = r - \frac{(2 - h)h}{2q} = r - \Theta\left(\frac{h}{q}\right),$$

$$w = \frac{1}{2q}\sqrt{-(1 + q + r)(1 + q - r)(1 - q + r)(1 - q - r)}$$

$$= \frac{1}{2q}\sqrt{(2q + h)(2 - h)h(r + q - 1)}$$

$$= \Theta\left(\sqrt{\frac{hr}{q}}\right),$$

since

$$0 \le h \le 1,$$
$$1 \le (2 - h) \le 2,$$
$$2q \le (2q + h) \le 3q,$$
$$r \le (r + q - 1) \le 2r.$$

Thus

$$\begin{aligned}\Gamma &= \Theta(hw^{d-1}) = \Theta(h^{(d+1)/2}(r/q)^{(d-1)/2}),\\ \text{esimp} &\le \mu_{d-1}w^{d-1}(r-x) = \Theta(\Gamma/q),\\ \hat{g} &= \mu_d^{-1}\kappa_d r^{d-1}K_d(x/r) = \Theta(r^{d-1}(1-(x/r))^{(d-1)/2})\\ &= \Theta((hr/q)^{(d-1)/2}) = \Theta(\Gamma/h),\end{aligned}$$

where $K_d(y)$ is the fraction of the surface of the unit d-ball cut off by a hyperplane at distance y from the origin. Now

$$\begin{aligned}\int_1^\infty\int_{q-1}^q I(q,r)\,dr\,dq &= \Theta(1)\int_1^\infty\int_{q-1}^q \left(\frac{q^{d-1}}{r^d}\right)\left(\frac{\Gamma}{h}\right)^{d+1}\left(\frac{\Gamma}{q}\right)e^{-n\Gamma}\,dr\,dq\\ &= \Theta(1)\int_1^\infty\int_0^1 q^{-2}h^{-1}\left(\frac{q}{hr}\right)^d\Gamma^{d+2}e^{-n\Gamma}\,dh\,dq\\ &= \Theta(1)\int_1^\infty\int_0^1 q^{-2}h^{-1}w^{-2d}\Gamma^{d+2}e^{-n\Gamma}\,dh\,dq\\ &= \Theta(1)\int_1^\infty\int_0^1 q^{-2}hw^{-2}\Gamma^d e^{-n\Gamma}\,dh\,dq.\end{aligned}$$

Substituting $t = n\Gamma$; $dt = \Theta(t/h)\,dh$ gives

$$\begin{aligned}\int_1^\infty\int_{q-1}^q I(q,r)\,dr\,dq &= \Theta(1)\int_1^\infty\int_0^n q^{-2}\left(\frac{h}{w}\right)^2\left(\frac{t}{n}\right)^d e^{-t}\frac{dt}{t}\,dq \qquad (2.2)\\ &= \Theta(n^{-d})\int_1^\infty\int_0^n q^{-1}\,hr^{-1}t^{d-1}e^{-t}\,dt\,dq.\end{aligned}$$

Since the Γ-region contains a ball of radius $h/2$, $\Gamma = \Omega(h^d)$ and $h = O(\Gamma^{1/d}) = O((t/n)^{1/d})$. Also $r^{1/2} \ge h^{1/2}$ and $r^{1/2} \ge (q-1)^{1/2}$, so

$$hr^{-1} = h^{1/2}\cdot(h^{1/2}\cdot r^{-1/2})\cdot r^{-1/2} = O((t/n)^{1/2d}\cdot 1\cdot(q-1)^{-1/2}),$$

and

$$\begin{aligned}&\int_1^\infty\int_{q-1}^q I(q,r)\,dr\,dq\\ &\quad= O(n^{-d-(1/2d)})\left(\int_1^\infty q^{-1}(q-1)^{-1/2}\,dq\right)\left(\int_0^n t^{d-1-(1/2d)}e^{-t}\,dt\right)\\ &\quad= O(n^{-d-(1/2d)})B(\tfrac{1}{2},\tfrac{1}{2})\Gamma(d-1-(1/2d))\\ &\quad= O(n^{-d-(1/2d)}).\end{aligned}$$

Case 7: $q > 1$ and $q \le r \le q+1$. Let x be the distance of the center of $\partial\mathscr{B}\cap\partial\mathscr{U}$ from $\partial\mathscr{B}$. We can bound esimp by the volume of a cylinder with a unit $(d-1)$-ball

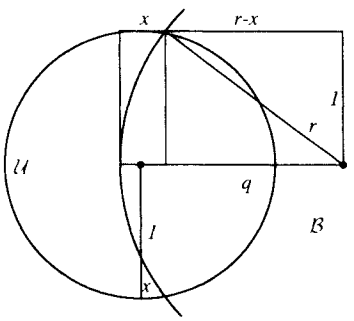

Fig. 5. Case 7: $q \leq 1$ and $q \leq r \leq q + 1$.

as base and x as height. From Fig. 5 we see that $(r - x)^2 + 1^2 = r^2$, or $x = r - \sqrt{r^2 - 1} = O(r^{-1})$, thus esimp $= O(r^{-1})$. We have the crude bound $\hat{g} = O(1)$ and the bound $\Gamma = \Omega(1)$ as in Case 4. Thus, since $r^{-1} \leq q^{-1}$,

$$\begin{aligned}\int_1^\infty \int_q^{q+1} I(q, r)\, dr\, dq &= \int_1^\infty \int_q^{q+1} q^{d-1} r^{-d} O(1) O(r^{-1}) \exp(-n\Omega(1))\, dr\, dq \\ &= \left(\int_1^\infty q^{-2}\, dq\right)\left(\int_q^{q+1} dr\right) O(e^{-\Omega(n)}) \\ &= O(e^{-\Omega(n)}).\end{aligned}$$

Case 8: $q > 1$ and $q + 1 \leq r$. In this case $\mathscr{U} \subset \mathscr{B}$ and $\partial\mathscr{B} \cap \mathscr{U} = \varnothing$, thus $\hat{g} = 0$ and the integral vanishes.

Examining all eight cases, we see that Case 1 dominates, and that

$$ES_n \sim \frac{d! d^{d-1} \mu_d v_d n}{(d + 1)} \tag{2.3}$$

with $\mu_d = (2\pi^{d/2}\Gamma(d/2))$ and, according to Miles [19],

$$v_d = \frac{\Gamma((d^2 + 1)/2)\Gamma(d/2)^{d+1}}{\sqrt{\pi} d!\, \Gamma(d^2/2)\Gamma((d + 1)/2)^d}. \qquad \square$$

3. A Fast Algorithm for the Unit d-Ball

It is immediate from Theorem 1 and the discussion of the lifting function in the introduction that the Voronoi diagram of random points from a d-ball can be constructed in $O(n^2)$ time on average by either the shelling algorithm or the gift-wrapping algorithm. In this section we describe a faster algorithm requiring only $O(n)$ time on average.

Our algorithm enumerates the d-simplices of the Voronoi dual. It is similar to Maus' planar algorithm [17]: it employs standard bucketing techniques, and its operation in $\mathbf{R}^d$ corresponds to the operation of the gift-wrapping algorithm in $\mathbf{R}^{d+1}$. (Bentley *et al.* [3] employed similar bucketing techniques in a quite different algorithm.) It is convenient to call the d-simplices of the Voronoi dual *cells* and the $(d-1)$-simplices *facets* (since they are facets of the cells); likewise, we call an empty d-sphere defined by the vertices of a cell a *cell sphere* and a $(d-1)$-sphere defined by the vertices of a facet a *facet sphere*. The algorithm proceeds by repeatedly finding a new cell adjacent to a known facet. Except for facets that are also facets of the (d-dimensional) convex hull, every facet belongs to exactly two cells. We maintain a dictionary of facets for which only one cell is known. At each step a facet is removed from the dictionary and its unknown cell (if it exists) is found by

```
Algorithm A
—Box(x) is the bucket for the box containing x.
—f_dict is the facet dictionary and contains (facet,half-space) pairs.
for x ∈ 𝒳_n do Box(x) := Box(x) ∪ {x};
Find an initial (facet,half-space) pair (ℱ, ℋ) by gift-wrapping;
Insert(ℱ, ℋ, f_dict);
while f_dict ≠ ∅ do
  (ℱ, ℋ) := any pair from f_dict;
  new_v := Find_site(ℱ, ℋ);
  Delete(ℱ, f_dict);
  if new_v ≠ nil then
    output(ℱ ∪ {new_v});
    for v ∈ ℱ do
      ℱ' := (ℱ\{v}) ∪ {new_v};
      ℋ' := the half-space defined by ℱ' not containing v:
      if ℱ' ∈ f_dict then Delete(ℱ', f_dict)
      else Insert(ℱ', ℋ', f_dict);
  end
end

function Find_site(ℱ, ℋ)
—Site_level(x, ℱ) is the signed distance of the center
—   of the d-ball defined by x and ℱ above the hyperplane of ℱ.
—   We define Site_level(nil, ℱ) = +∞.
—Box_level(B, ℱ) = min_{y ∈ (𝒮(B) ∩ ℋ)} Site_level(y, ℱ).
—q is a priority queue of boxes ordered by Box_level, initially empty.
Insert(Box(center of facet sphere of ℱ), q);
answer := nil;
while (q ≠ ∅) ∧ (Box_level(Find_min(q), ℱ) < Site_level(answer, ℱ)) do
  currentBox := Delete_min(q);
  for x ∈ currentBox do
    if (x ∈ ℋ) ∧ (Site_level(x, ℱ) < Site_level(answer, ℱ)) then answer := x;
  for newBox adjoining currentBox do
    if ((newBox ∩ ℋ ∩ 𝒰) ≠ ∅) ∧ (newBox ∉ q) then Insert(newBox, q);
return answer;
end Find_site
```

Fig. 6. Algorithm A.

searching for the unknown $(d+1)$st vertex (the *site search*). The remaining facets of the new cell are searched for in the dictionary (*facet searches*). Each that is found is deleted, since both of its cells are already known. Each that is not found is inserted so that its unknown cell will be searched for in some later step. The algorithm is described more formally in Fig. 6.

The facet dictionary is organized as a linear array of n buckets; a random facet falls into a particular bucket with probability $1/n$. Within each bucket facets may be organized in a balanced search tree to ensure good (logarithmic) worst-case performance, but a simple linear list is sufficient to achieve a linear bound on expected time.

Pseudocode describing the searching function *Find_site* is also found in Fig. 6. To speed the site searches, we partition the hypercube $[-1, 1]^d$ into approximately $2^d n/\mu_d$ hypercubic "boxes" of side $\lfloor(\mu_d/n)^{1/d}\rfloor$ and volume about μ_d/n. We assign each site to the bucket for the box in which it lies. Boxes lying completely outside the unit ball will always be empty. Boxes lying inside the unit ball will contain in expectation about one site each. An asymptotically vanishing fraction of the boxes will intersect the boundary of the unit ball and will contain less than one site each in expectation. On a particular cell to *Find_site*, let $\mathscr{F}$ and $\mathscr{H}$ be the actual arguments (a facet and a half-space to search), let $\mathscr{U}$ be the unit d-ball, and let $\mathscr{B}$ be the cell ball of the unknown cell if it exists. If the cell exists, it is necessary and sufficient to examine those boxes intersecting $(\mathscr{B} \cap \mathscr{H} \cap \mathscr{U})$ to determine it (Fig. 7). If it does not exist, it is necessary and sufficient to examine the boxes intersecting $(\mathscr{H} \cap \mathscr{U})$ to determine this (Fig. 8). The function *Find_site* examines almost exactly those boxes.

Simp_ctr$(x, \mathscr{F})$ is the center of the d-sphere determined by the point x and the d vertices of the facet $\mathscr{F}$. If $\mathscr{H}$ is the half-space $\langle a, x\rangle > 0$, then the goal of the site search is to find the site $x \in \mathscr{H}$ for which *Site_level*$(x, \mathscr{F}) = \langle a,$ *Simp_ctr*$(x, \mathscr{F})\rangle$ is minimized. The priority queue is primed by inserting the box containing the center of the facet sphere of $\mathscr{F}$. When a box is removed from the priority queue and its contents are examined, the $2d$ boxes that share a facet with it are entered into the queue for possible examination later. The boxes in the priority queue are ordered by the quantity *Box_level*$(B, \mathscr{F})$. Ideally, we would like to have

$$\textit{Box\_level}(B, \mathscr{F}) = \min_{y \in (B \cap \mathscr{H})} \textit{Site\_level}(y, \mathscr{F})$$

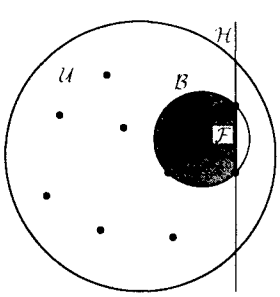

Fig. 7. A successful *Find_site* search.

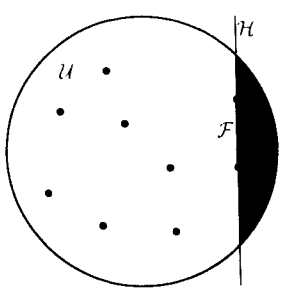

Fig. 8. An unsuccessful *Find_site* search.

(where the minimum is taken over all *points*, not sites, in B). Unfortunately, this ideal is difficult to achieve because the minimum may occur at any point on the surface of the hypercube B. Apparently it would be necessary to find minima separately on each of the $3^d - 1$ faces of the box. The time required to do this, while independent of n, obviously grows rapidly with the dimension.

Fortunately we can approximate this ideal sufficiently well in time proportional to d. The definition

$$Box\_level(B, \mathscr{F}) = \min_{y \in (\mathscr{S}(B) \cap \mathscr{H})} Site\_level(y, \mathscr{F}),$$

where $\mathscr{S}(B)$ is the circumsphere of the hypercube B, guarantees that the correct site will be found, while perhaps causing a few boxes to be examined unnecessarily. The boxes examined are exactly those with circumspheres intersecting $\mathscr{B} \cap \mathscr{H} \cap \mathscr{U}$ or $\mathscr{H} \cap \mathscr{U}$.

Now let us consider the computation of *Box_level*. Let $C_{\mathscr{F}}$ and $r_{\mathscr{F}}$ be the center and radius of the facet sphere of $\mathscr{F}$, let $C_{\mathscr{S}}$ and $r_{\mathscr{S}}$ be the center and radius of $\mathscr{S}(B)$, let $C'_{\mathscr{S}}$ be the projection of $C_{\mathscr{S}}$ onto the hyperplane defined by $\mathscr{F}$, let $d_{\mathscr{S}} = \text{dist}(C_{\mathscr{S}}, C'_{\mathscr{S}})$ and $d_{\mathscr{F}} = \text{dist}(C_{\mathscr{F}}, C'_{\mathscr{S}})$. A typical situation in two dimensions is depicted in Fig. 9. Without loss of generality, we may assume that $C_{\mathscr{F}}$ and $C_{\mathscr{S}}$ lie in the $(x^{(1)}, x^{(2)})$-plane, that $\mathscr{H}$, the half-space being searched, is $x^{(1)} > 0$, and further that $C_{\mathscr{F}}$ is the origin and that $C_{\mathscr{S}} = (d_{\mathscr{S}}, d_{\mathscr{F}}, 0, 0, \ldots, 0)$. We seek a point equidistant from all points on the facet sphere and the circumsphere of B. Such a point clearly lies on the $x^{(1)}$-axis. Formally, we require a point $P = (x. 0, 0, \ldots, 0)$ for which

$$\text{dist}(P, C_{\mathscr{S}}) = \text{dist}(P, (0, r_{\mathscr{F}}, 0, 0, \ldots, 0)) + r_{\mathscr{S}},$$

i.e., one for which

$$\sqrt{(x - d_{\mathscr{S}})^2 + d_{\mathscr{F}}^2} = r_{\mathscr{S}} + \sqrt{x^2 + r_{\mathscr{F}}^2}.$$

Squaring and rearranging terms, we have

$$2r_{\mathscr{S}}\sqrt{x^2 + r_{\mathscr{F}}^2} = d_{\mathscr{S}}^2 + d_{\mathscr{F}}^2 - r_{\mathscr{S}}^2 - r_{\mathscr{F}}^2 - (2d_{\mathscr{S}})x.$$

Writing q for $(d_{\mathscr{S}}^2 + d_{\mathscr{F}}^2 - r_{\mathscr{S}}^2 - r_{\mathscr{F}}^2)$, squaring again, and rearranging gives

$$4(r_{\mathscr{S}}^2 - d_{\mathscr{S}}^2)x^2 + (4qd_{\mathscr{S}})x + (4r_{\mathscr{S}}^2 r_{\mathscr{F}}^2 - q^2) = 0.$$

This quadratic equation in x always has two real roots. Which root to return as the value of *Box_level* depends on the relationship of the circumsphere $\mathscr{S}(B)$ and the half-space $\mathscr{H}$.

- If $\mathscr{S}(B)$ lies entirely inside $\mathscr{H}$, we choose the smaller root as the value of $Site\_level(B, \mathscr{F})$; the larger root *maximizes* the value of *Site_level* on the circumsphere (Fig. 9).

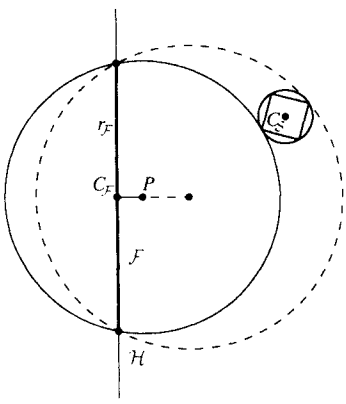

Fig. 9. Computing *Box_level* for a box lying completely in the half-space $\mathcal{H}$.

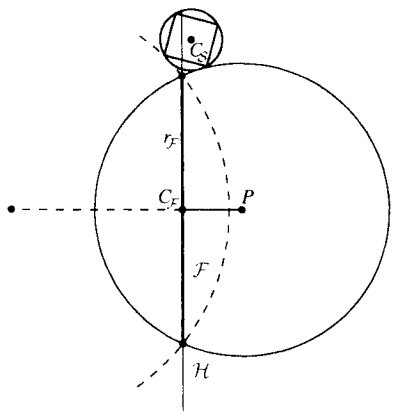

Fig. 10. Computing *Box_level* for a box lying only partly in the half-space $\mathcal{H}$.

- If $\mathcal{S}(B)$ lies entirely outside $\mathcal{H}$, we choose $+\infty$ for *Site_level*, eliminating B from further consideration.
- If $\mathcal{S}(B)$ intersects the facet sphere of $\mathcal{F}$, we choose $-\infty$ for *Site_level*, causing B to be searched immediately.
- If $\mathcal{S}(B)$ intersects the hyperplane defined by $\mathcal{F}$, but not the facet sphere, we choose the larger root of the equation; the smaller root is *Site_level* for a point on the circumsphere lying outside of $\mathcal{H}$ (Fig. 10).

The priority queue operations *Insert*, *Find_min*, and *Delete_min* can be implemented so that only $O(\log n)$ time is required for each [29], but a naïve linked-list implementation in which each operation requires time proportional to the length of the list suffices for the purposes of our average-case analysis.

4. Analysis of the Algorithm

In this section we show that all the facet and sites searches made by Algorithm A in the previous section can be completed in $O(n)$ time on average.

Lemma 1. *The facet searches can be completed in $O(n)$ expected time.*

Proof. To manage the facet dictionary, we maintain n buckets and hash facets into buckets by computing the exclusive-or of the binary representations of the indices of the sites defining the facet. Within each bucket we maintain a linked list of the facets in the bucket; each search or insertion in the bucket takes time proportional to the length of this list. It is not hard to see that this scheme hashes an equal number of the d-subsets of $\{1, 2, \ldots, n\}$ to each bucket, but we must show that those d-subsets actually defining facets are not correlated in a way that causes them to be hashed to a small number of buckets.

Let $C(n, d) = \binom{n}{d}$, let $\mathscr{D}_1, \mathscr{D}_2, \ldots, \mathscr{D}_{C(n,d)}$ be the d-subsets of the input set $\mathscr{X}_n$, and let M_k be the number of facets hashed into the kth bucket. In the kth bucket, $O(M_k)$ facet searchs are performed, each in time $O(M_k)$. Thus, by the linearity of expectation, the average amount of work done for the facet searches is

$$O\left(E\left(\sum_{0 \le k \le n} M_k^2\right)\right) = O\left(\sum_{0 \le k \le n} E(M_k^2)\right);$$

hence it will suffice to show that $E(M_k^2) = O(1)$ for any fixed k.

Let F_i represent the condition "$\mathscr{D}_i$ is a dual facet," and let I_i be the indicator variable of F_i, i.e., $I_i = 1$ if $\mathscr{D}_i$ is a dual facet, otherwise $I_i = 0$. Let $\mathscr{B}_k$ be the set of indices i for which $\mathscr{D}_i$ hashes into the kth bucket. It is easy to verify that $|\mathscr{B}_k| = n^{-1}C(n, d)$. Without loss of generality let us consider the bucket containing $\mathscr{D}_1$. Then

$$\begin{aligned}
E(M_k^2) &= E\left(\sum_{i \in \mathscr{B}_k} I_i\right)^2 \\
&= E\left(\sum_{i \in \mathscr{B}_k} I_i^2\right) + E\left(\sum_{i \in \mathscr{B}_k} \sum_{\substack{j \in \mathscr{B}_k \\ j \ne i}} I_i I_j\right) \\
&= EM_k + \sum_{i \in \mathscr{B}_k} \sum_{\substack{j \in \mathscr{B}_k \\ j \ne i}} \Pr\{F_i\} \cdot \Pr\{F_j | F_i\} \\
&= O(1) + n^{-1}C(n, d)\Pr\{F_1\} \cdot \sum_{\substack{j \in \mathscr{B}_k \\ j \ne 1}} \Pr\{F_j | F_1\}.
\end{aligned}$$

By Theorem 1 the expected number of facets is $\Theta(n)$, so $\Pr\{F_1\} = \Theta(n/C(n, d))$. The conditional probability $\Pr\{F_j | F_1\}$ depends only on $|\mathscr{D}_j \backslash \mathscr{D}_1|$. If $|\mathscr{D}_j \backslash \mathscr{D}_1| = m > 0$, then there are at most $C(n - d, m - 1)/C(m, m - 1) = O(n^{m-1})$ ways to choose the m elements of $\mathscr{D}_j \backslash \mathscr{D}_1$ so that $\mathscr{D}_1$ and $\mathscr{D}_j$ fall into the same bucket. This holds because choosing the first $m - 1$ elements fixes the mth. (We must say "at most" because the mth element required by the first $m - 1$ sometimes belongs to $\mathscr{D}_1$ and cannot be chosen. This always occurs when $m = 1$, for example.) So

$$\begin{aligned}
E(M_k^2) &= O(1) + O(1) \cdot \sum_{1 \le m \le d} O(n^{m-1})\Pr\{F_j | F_1 \wedge (|\mathscr{D}_j \backslash \mathscr{D}_i| = m)\} \\
&= O(1) + O(1/n) \cdot \sum_{1 \le m \le d} \binom{n-d}{m} \Pr\{F_j | F_1 \wedge (|\mathscr{D}_j \backslash \mathscr{D}_i| = m)\} \\
&= O(1) + O(1/n)E(\text{number of facets} \mid F_1) \\
&= O(1) + O(1/n)E(\text{number of facets}) \\
&= O(1),
\end{aligned}$$

since the number of d-subsets (in all buckets together) satisfying $|\mathscr{D}_j \backslash \mathscr{D}_i| = m$ is exactly $\binom{n-d}{m}$. □

We now turn to the search for the $(d+1)$st site completing a cell with a known facet. We call a site search "successful" if a site is found, and "unsuccessful" if no site is found because the facet lies on the boundary of the convex hull.

Lemma 2. *The successful site searches can be completed in $O(n)$ expected time.*

Proof. If we define the distance between a point x and a set $\mathscr{Y}$ by

$$\operatorname{dist}(x, \mathscr{Y}) = \min_{y \in \mathscr{Y}} \operatorname{dist}(x, y),$$

all the boxes with circumspheres intersecting $\mathscr{B} \cap \mathscr{H} \cap \mathscr{U}$ are completely contained by the set

$$\mathscr{A} = \{x \mid \operatorname{dist}(x, \mathscr{B} \cap \mathscr{U}) \leq \sqrt{d}(\mu_d/n)^{1/d}\}. \tag{4.1}$$

Intuitively, $\mathscr{A}$ is the intersection of $\mathscr{U}$ with $\mathscr{B}$ inflated by the diameter of a box, just enough so that the inflated ball contains completely all the boxes that intersect $\mathscr{B}$. The use of a priority queue in *Find_site* guarantees that exactly these boxes are examined. Assuming the naïve linked-list implementation, the cost of each priority-queue operation is at most proportional to the total number of boxes examined. The cost of examining the sites in a box is $O(1)$ in expectation. The expected total cost of the site search is therefore proportional to the square of the number of boxes examined, or

$$\operatorname{ecost} = O((n \cdot \operatorname{vol} \mathscr{A})^2) \tag{4.2}$$

If we write C_n for the total cost of all successful site searches needed to compute the Voronoi diagram of $\mathscr{X}_n$, we have

$$EC_n \leq \binom{n}{d+1} \int_{\mathbf{R}^d} \cdots \int_{\mathbf{R}^d} \operatorname{ecost}(x_1, \ldots, x_{d+1})(1 - \Gamma)^{n-d-1} \times g(x_1) \cdots g(x_{d+1})\, dx_1 \cdots dx_{d+1};$$

the $(d+1)$-fold integral represents an upper bound on the expected cost of a successful site search to complete the cell $x_1 x_2, \ldots, x_{d+1}$. Proceeding as in Section 1, we eventually obtain

$$EC_n = O(1) \cdot \int_0^\infty \int_0^\infty \operatorname{ecost}(q, r) I(q, r)\, dr\, dq.$$

We continue as in Section 1, dividing the domain of integration into eight regions. However, we ignore constant factors this time.

Case 1, 2, and 3. In all these cases,

$$\hat{g} = \Theta(r^{d-1}), \qquad \Gamma = \Theta(r^d), \qquad \text{esimp} = \Theta(r^d).$$

The set $\mathscr{A}$ of (4.1) is contained in a d-ball of radius $(r + \sqrt{d}(\mu_d/n)^{1/d})^d$ and, by (4.2), with $t = n\Gamma = \Theta(nr^d)$ as in Section 2,

$$\sqrt{\text{ecost}(q, r)} = O(n(r + n^{-1/d})^d) = O\left(\sum_{i=0}^{d} r^i n^{i/d}\right) = O\left(\sum_{i=0}^{d} t^{i/d}\right) = O(1 + t).$$

It follows that

$$\begin{aligned}\int_0^1 \int_0^1 \text{ecost}(q, r) I(q, r)\, dr\, dq &= \Theta(1)\left(\int_0^1 q^{d-1}\, dq\right)\left(\int_0^n (1 + t)^2 \left(\frac{t}{n}\right)^d \frac{e^{-t}}{t}\, dt\right) \\ &= \Theta(n^{-d}).\end{aligned}$$

Cases 4 and 7. For these cases we apply the trivial bound $\text{ecost}(r, q) = O(n^2)$, and

$$\begin{aligned}\int\int \text{ecos}(q, r) I(q, r)\, dr\, dq &= O(n^2) \int\int I(q, r)\, dr\, dq \\ &= O(e^{-\Omega(n)}).\end{aligned}$$

Cases 5 and 8. In these cases $\hat{g} = 0$ and the integral vanishes as before.

Case 6. In this case the set $\mathscr{A}$ of (4.1) is contained in a $(d-1)$-spherical cylinder with height $(h + 2\sqrt{d}(\mu_d/n)^{1/d})$ and base radius $(w + \sqrt{d}(\mu_d/n)^{1/d})$, so

$$\begin{aligned}\sqrt{\text{cost}(q, r)} &= O(n(h + n^{-1/d})(w + n^{-1/d})^{d-1}) \\ &= O(n) \cdot \left(\sum_{i=0}^{d-1} h n^{-i/d} w^{d-1-i} + \sum_{i=1}^{d} n^{-i/d} w^{d-i}\right) \\ &= O(n) \cdot \left(hw^{d-1} + \sum_{i=1}^{d-1} (h + w) n^{-i/d} w^{d-1-i} + n^{-1}\right) \\ &= O\left(nhw^{d-1} + \sum_{i=1}^{d-1} w^i n^{i/d} + 1\right) \qquad (4.3) \\ &= O\left(n\Gamma + \sum_{i=0}^{d-1} \left(\frac{wn\Gamma}{h}\right)^{i/d}\right) \\ &= O\left(\left(\frac{w}{h}\right)(n\Gamma + 1)\right),\end{aligned}$$

since $h = O(w)$ and $w = (w\Gamma/h)^{1/d}$. Computation is similar to the first subcase of Case 6 in the proof of Theorem 1, substituting $t = n\Gamma$, to the point of (2.2), where this time we have

$$\begin{aligned}\int_0^\infty \int_{q-1}^\infty \text{ecost}(q, r)I(q, r)\, dr\, dq &= \Theta(n^{-d}) \int_0^n \text{ecost}(q, r)\left(\frac{h}{w}\right)^2 t^{d-1}e^{-t}\, dt \\ &= O(n^{-d}) \int_0^n (1 + t)^2 t^{d-1}e^{-t}\, dt \\ &= O(n^{-d}).\end{aligned}$$

Summing over the eight cases, we see that

$$\int_0^\infty \int_0^\infty \text{ecost}(q, r)I(q, r)\, dr\, dq = O(n^{-d})$$

and

$$EC_n = O(n^{d+1}) \cdot O(n^{-d}) = O(n). \qquad \square$$

Finally, we consider the unsuccessful site searches for facets of the convex hull.

Lemma 3. *The unsuccessful site searches can be completed in $o(n)$ expected time.*

Proof. An unsuccessful search involves examination of all the boxes with circumspheres intersecting $\mathscr{H} \cap \mathscr{U}$; these are completely contained in

$$\mathscr{A} = \{x | \text{dist}(x, \mathscr{H} \cap \mathscr{U}) \leq \sqrt{d}(\mu_d/n)^{1/d}\};$$

here, we shift the hyperplane $\partial\mathscr{H}$ slightly to enclose completely all intersected boxes. As before,

$$\text{ecost} = O((n \cdot \text{vol}\, \mathscr{A})^2)).$$

Let us write U_n for the total cost of all unsuccessful site searches. These occur only for facets of the convex hull. We have

$$EU_n \leq \binom{n}{d} \int_{\mathbf{R}^d} \cdots \int_{\mathbf{R}^d} \text{ecost}(x_1, x_2, \ldots, x_d)(1 - \Gamma)^{n-d} g(x_1) \cdots g(x_d) dx_1 \cdots dx_d.$$

Here Γ is the probability content of the smaller half-space defined by x_1 through x_d, and $(1 - \Gamma)^{n-d}$ is the probability that all the other sites lie in the larger half-space. We apply a standard change of variables used by Efron [13], Raynaud [21], the author [11], and others. Each of the d vertices of $\mathscr{F}$ is expressed in terms of p, the

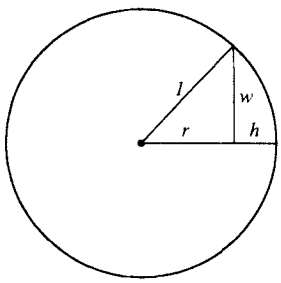

Fig. 11. Dimensions of an unsuccessful search.

projection of the origin onto the hyperplane $\mathscr{H}$, and an orthonormal basis for $\mathscr{H} - p$. After some computation similar to that of Section 1, we obtain

$$EU_n = O(n^d) \cdot \int_0^1 \text{ecost}(r)\,\text{esimp}(r)(\hat{g}(r))^d \exp(-n\Gamma)\,dr,$$

where r is the distance of the hyperplane $\mathscr{H}$ from the origin, esimp(r) is the expected volume of the simplex $x_1 x_2 \cdots x_d$ given the distance r, ecost(r) is the expected cost of the unsuccessful search given r, $\hat{g}(r)$ is the density of the probability that a random point falls on $\mathscr{H}$, and $\exp(-n\Gamma)$ estimates the probability that $\mathscr{H}$ defines an empty halfspace.

Let $h = 1 - r$ and $w = \sqrt{1 - r^2} = \Theta(\sqrt{h})$ as in Fig. 11. Then

$$\Gamma(r) = \Theta(hw^{d-1}) = \Theta(w^{d+1}).$$

With $t = n\Gamma$ or $w = (t/n)^{1/(d+1)}$, we have

$$dr = -dh = \Theta(1)\left(\frac{t}{n}\right)^{2/(d+1)} \frac{dt}{t},$$

$$\text{esimp}(r) = \Theta(w^{d-1}) = \Theta\left(\left(\frac{t}{n}\right)^{(d-1)/(d+1)}\right),$$

$$\hat{g}(r) = \Theta(w^{d-1}) = \Theta\left(\left(\frac{t}{n}\right)^{(d-1)/(d+1)}\right),$$

$$\begin{aligned}
\sqrt{\text{ecost}(r)} &= \Theta(n(h + n^{-1/d})(w + n^{-1/d})^{d-1}) \\
&= \Theta\left(nhw^{d-1} + \sum_{i=0}^{d-1} w^i n^{i/d}\right) \qquad \text{as in (4.3)} \\
&= \Theta\left(t + \sum_{i=0}^{d-1} (tn^{1/d})^{i/(d+1)}\right) \\
&= O(n^{(d-1)/(d^2+d)}(1 + t)),
\end{aligned}$$

and

$$\begin{aligned} EU_n &= O(n^d) \int_0^n (n^{2(d-1)/(d^2+d)}(1+t)^2) \\ &\quad \times \left(\frac{t}{n}\right)^{(d-1)/(d+1)} \left(\frac{t}{n}\right)^{d(d-1)/(d+1)} e^{-t\left(\frac{t}{n}\right)^{2/(d+1)}} \frac{dt}{t} \\ &= O(n^{1-(2/(d^2+d))}) \int_0^n (1+t)^2 t^{d-2+(2/(d+1))} e^{-t}\, dt \\ &= o(n). \end{aligned}$$

□

Lemmata 1, 2, and 3 together imply the following theorem.

Theorem 2. *Let $\mathcal{X}_n = \{X_1, X_2, \ldots, X_n\}$ be a set of n sites drawn independently from the uniform distribution on the interior of the unit d-ball. Then for fixed d, Algorithm A constructs the Voronoi diagram of $\mathcal{X}_n$ in $O(n)$ time on average.*

5. Discussion

Applying (2.3) for $d = 2$, we obtain $ES_n \sim 2n$. This is confirmed by well-known combinatorial results. We also have

$$ES_n \sim \frac{24\pi^2}{35} n \approx 6.77n \qquad \text{for} \quad d = 3,$$

$$ES_n \sim \frac{286}{9} n \approx 31.78n \qquad \text{for} \quad d = 4.$$

These values are not obviously inconsistent with the values 6.31 and 25.6 found empirically by Avis and Bhattacharya for (rather small) samples of 1000 points chosen from the unit hypercube [2, Table 1]; it is reasonable to conjecture that (2.3) in fact holds for point set chosen from a uniform distribution on *any* convex body.

Algorithm A is clearly optimal in the average-case sense for fixed dimension, and is asymptotically faster than any other known. If a balanced-tree implementation of priority queues is used, its worst-case running time is $O(S_n n \log n)$, only a factor of $\Theta(\log n)$ worse that the standard gift-wrapping algorithm. Worst-case performance can be improved to $O(nS_n)$ if the use of buckets is abandoned on any site search that examines $\sqrt{n}$ buckets. It should not be difficult to show that this occurs so infrequently that average performance is not affected.

It is easy to show that linear performance is preserved if the distribution is "quasi-uniform" in the unit d-ball, i.e., if its density is bounded above and below by two positive constants everywhere in the d-ball. It is an open question whether the same approach yields an $O(n)$ algorithm—or even linear bounds on S_n—for other distributions.

Acknowledgments

D. Sleator, K. Clarkson, W. Eddy, D. Scott, and D. Tygar, the members of my thesis committee at Carnegie-Mellon University, all gave useful comments on drafts of this paper. W. Eddy in particular encouraged me to pursue this topic more seriously. I. Kolodner gave useful tips on determining the asymptotic behavior of the integrals.

References

1. F. Affentranger, Random spheres in a convex body, *Arch. der Math.* **55** (1990), 74–81.
2. D. Avis and B. K. Bhattacharya, Algorithms for computing d-dimensional Voronoi diagrams and their duals, in *Advances in Computing Research: Computational Geometry*, F. P. Preparata, ed., pp. 159–180, JAI Press, Greenwich, CT, 1983.
3. J. L. Bentley, B. W. Weide, and A. C. Yao, Optimal expected-time algorithms for closest-point problems *ACM Trans. Math. Software* **6** (1980), 563–580.
4. W. H. Beyer, *CRC Standard Mathematical Tables*, 27th ed., CRC Press, Boca Raton, FL, 1984.
5. B. Bhattacharya, Worst-case Analysis of a Convex Hull Algorithm, Technical Report, Simon Fraser University, 1982.
6. A. Bowyer, Computing Dirichlet tessellations, *Comput. J.* **24** (1981), 162–166.
7. K. Q. Brown, Geometric Transforms for Fast Geometric Algorithms, Ph.D. thesis, Carnegie-Mellon University, Pittsburgh, PA, 1979.
8. W. Browstow, J.-P. Dussault, and B. L. Fox, Construction of Voronoi polyhedra, *J. Comput. Phys.* **29** (1978), 81–97.
9. D. R. Chand and S. S. Kapur, An algorithm for convex polytopes, *J. Assoc. Comput. Mach.* **17** (1970), 78–86.
10. R. A. Dwyer, A faster divide-and-conquer algorithm for constructing Delaunay triangulations, *Algorithmica* **2** (1987), 137–151.
11. R. A. Dwyer. On the convex hull of random points in a polytope, *J. Appl. Probab.* **25** (1988), 688–699.
12. H. Edelsbrunner and R. Seidel, Voronoi diagrams and arrangements, *Discrete Comput. Geom.* **1** (1986), 25–44.
13. B. Efron, The convex hull of a random set of points, *Biometrika* **52** (1965), 331–342.
14. J. L. Finney, A procedure for the construction of Voronoi polyhedra, *J. Comput. Phy.* **32** (1979), 137–143.
15. E. N. Gilbert, Random subdivisions of space into crystals, *Ann. Math. Statist.* **33** (1962), 958–972.
16. M. G. Kendall, *A Course in the Geometry of n Dimensions*, Hafner, New York, 1961.
17. A. Maus, Delaunay triangulation and the convex hull of n points in expected linear time, *BIT* **24** (1984), 151–163.
18. J. L. Meijering, Interface area, edge lenght, and number of vertices in crystal aggregates with random nucleation, *Philips Res. Rep.* **8** (1953), 270–290.
19. R. E. Miles, Isotropic random simplices, *Adv. Appl. Probab.* **3** (1971), 353–382.
20. R. E. Miles, A synopsis of "Poisson flats in Euclidean spaces", in *Stochastic Geometry*, E. F. Harding and D. G. Kendall, eds., pp. 202–227, Wiley, New York, 1974.
21. H. Raynaud, Sur l'envelope convexe des nuages des points aléatoires dans $\mathbf{R}^n$, I, *J. Appl. Probab.* **7** (1970), 35–48.
22. A. Rényi and R. Sulanke, Über die konvexe Hülle von n zufällig gewählten Punkten, *Z. Wahrsch. Verw. Gebiete* **2** (1963), 75–84.
23. A. Rényi and R. Sulanke, Über die konvexe Hülle von n zufällig gewählten Punkten, II, *Z. Wahrsch. Verw. Gebiete* **3** (1964), 138–147.
24. R. Seidel, The complexity of Voronoi diagrams in higher dimensions, in *Proc. 20th Ann. Allerton*

Conf. on Communication, Control, and Computing, pp. 94–95, University of Illinois at Urbana-Champaign, 1982.
25. R. Seidel, Constructing higher-dimensional convex hulls at logarithmic cost per face, in *Proc. 18th ACM Symp. on Theory of Computing*, pp. 404–413, ACM, 1986.
26. R. Seidel, On the number of faces in higher-dimensional Voronoi diagrams, in *Proc. 3rd Ann. Symp. on Computational Geometry*, pp. 181–185, ACM, 1987.
27. G. Swart, Finding the convex hull facet by facet, *J. Algorithms* **6** (1985), 17–48.
28. M. Tanemura, T. Ogawa, and N. Ogita, A new algorithm for three-dimensional Voronoi tessellation, *J. Comput. Phys.* **51** (1983), 191–207.
29. R. E. Tarjan, *Data Structures and Network Algorithms*, CBMS–NSF Regional Conference Series in Applied Mathematics, Vol. 44, Society of Industrial and Applied Mathematics, Philadelphia, PA, 1983.
30. A. H. Thiessen, Precipitation averages for large areas, *Monthly Weather Rev.* **39** (1911), 1082–1084.
31. F. G. Tricomi, *Funzioni Ipergeometriche Confluente*, Edizioni Cremonese, Rome, 1954.
32. G. Voronoi, Nouvelles applications des parametres continus à la theorie des formes quadratique, deuxième mémoire: recherches sur les parallélloèdres primitifs, *J. Reine Angew. Math.* **134** (1908), 198–287.
33. D. F. Watson, Computing the n-dimensional Delaunay tessellation with application to Voronoi polytopes, *Comput. J.* **24** (1981), 167–172.
34. E. T. Whittaker and G. N. Watson, *A Course of Modern Analysis*, 4th edn., Cambridge University Press, Cambridge, 1927.
35. E. Wigner and F. Seitz, On the constitution of metallic sodium, *Phy. Rev.* **43** (1933), 804–810.

Received November 9, 1988, and in revised form November 14, 1989.