Journal of Approximation Theory **119**, 252–270 (2002)
doi:10.1006/jath.2002.3731

On the Asymptotics of Fekete-Type Points for Univariate Radial Basis Interpolation

L. P. Bos

Department of Mathematics and Statistics, University of Calgary, Calgary T2N 1N4, Alberta Canada
E-mail: lpbos@math.ucalgary.ca

and

U. Maier

Mathematisches Institut, Universität Gießen, Arndtstr. 2, 35392 Gießen, Germany

Received November 26, 2001; accepted August 19, 2002

Suppose that $K \subset \mathbb{R}^d$ is compact and that we are given a function $f \in C(K)$ together with distinct points $x_i \in K$, $1 \leqslant i \leqslant n$. Radial basis interpolation consists of choosing a fixed (basis) function $g : \mathbb{R}^+ \to \mathbb{R}$ and looking for a linear combination of the translates $g(|x - x_j|)$ which interpolates f at the given points. Specifically, we look for coefficients $c_j \in \mathbb{R}$ such that

$$F(x) = \sum_{j=1}^{n} c_j g(|x - x_j|)$$

has the property that $F(x_i) = f(x_i)$, $1 \leqslant i \leqslant n$. The Fekete-type points of this process are those for which the associated interpolation matrix $[g(|x_i - x_j|)]_{1 \leqslant i,j \leqslant n}$ has determinant as large as possible (in absolute value). In this work, we show that, in the univariate case, for a broad class of functions g, among all point sequences which are (strongly) asymptotically distributed according to a weight function, the equally spaced points give the asymptotically largest determinant. This gives strong evidence that the Fekete points themselves are indeed asymptotically equally spaced. © 2002 Elsevier Science (USA)

1. THE CASE OF $g(x) = x$

In the case of classical polynomial interpolation, Fekete points have been much studied and are known to be nearly "optimal" (see e.g. [11] or [10] (where they are referred to as extremal fundamental systems)). In contrast, very little is known about optimal points for radial basis interpolation, a

0021-9045/02 $35.00

practical method of data fitting that has found much success in a myriad of applications (see e.g. [2–5]).

To make our problem more precise, suppose that $K \subset \mathbb{R}^d$ is compact and that we are given a function $g : \mathbb{R}^+ \to \mathbb{R}$. By Fekete-type points (of order n) for g (on K) we mean a set of n points $x_i \in K,\ 1 \leqslant i \leqslant n$ for which

$$|\det[g(|x_i - x_j|)]|$$

is as large as possible.

In [1] we began a study of Fekete-type points for radial basis interpolation. From this earlier work it appeared that a strong case could be made for the conjecture that, at least in the univariate case, the Fekete points for a rather broad class of functions g, were asymptotically uniformly distributed. However, as opposed to the classical Fekete points which maximize the polynomial Vandermonde determinant, there is no known analytic characterization of such Radial Basis Fekete-type points. Hence their asymptotic analysis seems to be a rather difficult problem. In this present work we give further strong evidence that these points are indeed asymptotically uniformly distributed by proving that among all point sequences which are (strongly) asymptotically distributed according to a weight function (see Definition 1.2), the equally spaced points give the asymptotically largest determinant (see Corollary 2.2 for a precise statement). The arguments are based on a rather remarkable relationship with the entropy of the weight function, given in Theorem 2.1. Loosely speaking, the entropy of the weight plays the same role as does the logarithmic energy of the limiting measure in the classical case.

We begin by discussing the special (and as it turns out, fundamental) case of $g(x) = x$.

Suppose then that $K = [0, 1]$ and that $0 \leqslant x_1 < x_2 < \cdots < x_{n+1} \leqslant 1$ (we take $n + 1$ points for technical simplicity). We are concerned with the determinant

$$D_n := \det[|x_i - x_j|]_{1 \leqslant i,j \leqslant n+1}.$$

If we set $h_i := x_{i+1} - x_i,\ 1 \leqslant i \leqslant n$, then we have

$$D_n = \begin{vmatrix} 0 & h_1 & h_1 + h_2 & \cdots & \cdots & h_1 + h_2 + \cdots + h_n \\ h_1 & 0 & h_2 & \cdots & \cdots & h_2 + \cdots + h_n \\ h_1 + h_2 & h_2 & 0 & h_3 & \cdots & \vdots \\ \vdots & & \ddots & \ddots & & h_n \\ h_1 + h_2 + \cdots + h_n & \cdots & & h_n & & 0 \end{vmatrix}$$

which, as is not hard to see, is

$$D_n = (-1)^n 2^{n-1} \left(\prod_{i=1}^{n} h_i \right) \left(\sum_{i=1}^{n} h_i \right) \tag{1}$$

so that

$$|D_n| = 2^{n-1} \left(\prod_{i=1}^{n} h_i \right) \left(\sum_{i=1}^{n} h_i \right).$$

Clearly, for this latter value to be maximized, $\sum_{i=1}^{n} h_i$ must be as large as possible. It follows that $\sum_{i=1}^{n} h_i = 1$ and that $x_1 = 0$ and $x_{n+1} = 1$. Consequently, we are left with the problem of maximizing $\prod_{i=1}^{n} h_i$ subject to the constraint that $\sum_{i=1}^{n} h_i = 1$. But, as is well-known, this maximum is uniquely attained for

$$h_i = \frac{1}{n}, \quad 1 \leqslant i \leqslant n,$$

i.e. precisely for the equally spaced points

$$x_i = \frac{i-1}{n}, \quad 1 \leqslant i \leqslant n+1,$$

which are clearly uniformly distributed on $[0, 1]$.

Thus the Fekete points for $g(x) = x$ are indeed equally spaced. In the sequel, for other functions g, we will compare the determinants for equally spaced points with those for competing distributions.

Definition 1.1. We will say that $w \in C[0, 1]$ with $w(x) > 0$, $\forall x \in [0, 1]$ and

$$\int_0^1 w(x)\, dx = 1$$

is an *allowable* weight function.

Now, to see the behaviour of the determinant D_n when the points are asymptotically distributed according to an allowable weight function, consider first the special case when

$$x_i := W^{-1}\left(\frac{i-1}{n} \right), \quad 1 \leqslant i \leqslant n+1,$$

where

$$W(x) := \int_0^x w(t)\, dt.$$

It is not difficult to see that, in this case,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n+1} f(x_i) = \int_0^1 f(x)w(x)\,dx$$

for every $f \in C[0,1]$, justifying our calling these points equally spaced with respect to $w(x)$.

Continuing, we may calculate

$$\begin{aligned} nh_i = n(x_{i+1} - x_i) &= \frac{W^{-1}\left(\frac{i}{n}\right) - W^{-1}\left(\frac{i-1}{n}\right)}{\frac{i}{n} - \frac{i-1}{n}} \\ &= \frac{d}{dx} W^{-1}(c_i) \quad \text{for some } c_i, \ \frac{i-1}{n} \leqslant c_i \leqslant \frac{i}{n}. \end{aligned}$$

But

$$\begin{aligned} \frac{d}{dx} W^{-1}(x) &= \frac{1}{\frac{dW}{dx}(W^{-1}(x))} \\ &= \frac{1}{w(W^{-1}(x))} \end{aligned}$$

and so,

$$nh_i = \frac{1}{w(W^{-1}(c_i))}, \quad \frac{i-1}{n} \leqslant c_i \leqslant \frac{i}{n}. \tag{2}$$

It follows from (1) that

$$\begin{aligned} \log((n^n|D_n|)^{1/n}) &= \log(2^{(n-1)/n}) + \frac{1}{n} \sum_{i=1}^{n} \log(nh_i) \\ &= \log(2^{(n-1)/n}) + \frac{1}{n} \sum_{i=1}^{n} \log\left(\frac{1}{w(W^{-1}(c_i))}\right). \end{aligned}$$

The second term is a Riemann sum for $\int_0^1 \log(\frac{1}{w(W^{-1}(x))})\,dx$ which, by the change of variables, $x' = W^{-1}(x)$, is easily seen to be

$$\int_0^1 \log\left(\frac{1}{w(x)}\right) w(x)\,dx =: \log(\mathscr{E}(w)),$$

the log of the *entropy* of the weight function w. Hence, we have the already interesting formula

$$\lim_{n\to\infty} (n^n|D_n|)^{1/n} = 2\exp\left(\int_0^1 \log\left(\frac{1}{w(x)}\right) w(x)\,dx\right) = 2\mathscr{E}(w).$$

We will use property (2) in the following definition.

Definition 1.2. Suppose that $w(x)$ is an allowable weight function and that $0 \leqslant x_1 < x_2 < \cdots < x_{n+1} \leqslant 1$ is a triangular array of points. We say that the points are (strongly) asymptotically distributed with weight $w(x)$ if for each $1 \leqslant i \leqslant n$,

$$nh_i = \frac{1}{w(W^{-1}(c_i))} \quad \text{for some } \frac{i-1}{n} \leqslant c_i \leqslant \frac{i}{n}.$$

Remark. For technical simplicity we do not consider a weaker form of asymptotic distribution. Weak-$*$ convergence, for example, allows point repetitions for which our determinants would be 0 and hence a result such as Proposition 1.3 would not hold.

With this definition then, we have immediately

Proposition 1.3. *Suppose that* $0 \leqslant x_1 < x_2 < \cdots < x_n \leqslant 1$ *is a distinct set of points, asymptotically distributed with respect to the allowable weight* $w(x)$. *Setting* $D_n = \det[|x_i - x_j|]_{1\leqslant i,j\leqslant n}$, *we have,*

$$\lim_{n\to\infty} (n^n|D_n|)^{1/n} = 2\exp\left(\int_0^1 \log\left(\frac{1}{w(x)}\right) w(x)\,dx\right) = 2\mathscr{E}(w).$$

Proposition 1.3 can be used to give an alternative, although somewhat weaker, explanation of why the equally spaced points yield a larger determinant than points asymptotically distributed according to any other competing allowable weight function. Precisely, since the limiting expression

$$2\exp\left(\int_0^1 \log\left(\frac{1}{w(x)}\right) w(x)\,\mathrm{d}x\right) = 2\mathscr{E}(w) \tag{3}$$

is twice the exponential of the entropy of w, we may recall that, as is well-known (cf. [9]), the function of maximum entropy is just $w \equiv 1$. Alternatively, note that this limiting expression is also twice the Geometric

Mean of the function $1/w(x)$ with respect to the weight $w(x)$. Then by the Arithmetic–Geometric Mean inequality,

$$2\exp\left(\int_0^1 \log\left(\frac{1}{w(x)}\right)w(x)\,dx\right) \leqslant 2\left(\int_0^1 \frac{1}{w(x)}w(x)\,dx\right) = 2.$$

But 2 is just the value of the Geometric Mean (3) for the weight $w(x) \equiv 1$, i.e., for that corresponding to *equally spaced* points. If we then write D_n^w for the determinant of points asymptotically distributed according to weight w, we have

$$\lim_{n\to\infty}\left(\frac{|D_n^w|}{|D_n^1|}\right)^{1/n} = \lim_{n\to\infty}\left(\frac{n^n|D_n^w|}{n^n|D_n^1|}\right)^{1/n} = \frac{2\exp(\int_0^1 \log(\frac{1}{w(x)})w(x)\,dx)}{2\exp(\int_0^1 \log(\frac{1}{1})1\,dx)} < 1 \tag{4}$$

for $w \not\equiv 1$. It follows that for $w \not\equiv 1$ and large n, $|D_n^w|$ is exponentially smaller than $|D_n^1|$.

2. THE CASE OF GENERAL g

To repeat, in the first section we showed that for $g(x) = x$, if the points are asymptotically distributed according to the allowable weight function $w(x)$ then

$$\lim_{n\to\infty}(n^n|D_n|)^{1/n} = 2\exp\left(\int_0^1 \log\left(\frac{1}{w(x)}\right)w(x)\,dx\right).$$

Clearly then, for $g(x) = ax$, since each entry in the matrix is multiplied by the factor a,

$$\lim_{n\to\infty}(n^n|D_n|)^{1/n} = 2|a|\exp\left(\int_0^1 \log\left(\frac{1}{w(x)}\right)w(x)\,dx\right) = 2|g'(0)|\exp\left(\int_0^1 \log\left(\frac{1}{w(x)}\right)w(x)\,dx\right)$$

seeing that $a = g'(0)$.

The remarkable fact is that this latter formula holds (essentially) in general.

THEOREM 2.1. *Suppose that* $g \in C^2[0,1]$ *with* $g(x) = g'(0)x + r(x)$ *where* $g'(0) \neq 0$ *and* $r \in C^2[0,1]$ *is such that* $r'(0) = 0$. *Suppose further that g satisfies the technical condition that* $-g'(0)$ *is* not *an eigenvalue of the operator* $A : C[0,1] \to C[0,1]$ *given by*

$$(Af)(x) = f(0)\frac{r_x(1) + r_x(0) + r_x'(0)}{2} + f(1)\frac{r_x(0) + r_x(1) - r_x'(1)}{2} + \frac{1}{2}\int_0^1 r_x''(y)f(y)\,dy, \tag{5}$$

where

$$r_x(y) := r(|y - x|). \tag{6}$$

Let $w(x)$ *be an allowable weight function. Then if* $0 \leqslant x_1 < x_2 < \cdots < x_{n+1} \leqslant 1$ *are asymptotically distributed according to w* (*in the sense of Definition* 1.2), *we have*

$$\lim_{n\to\infty} (n^n|D_n|)^{1/n} = 2|g'(0)| \exp\left(\int_0^1 \log\left(\frac{1}{w(x)}\right) w(x)\,dx\right)$$

$$= 2|g'(0)|\mathscr{E}(w),$$

where

$$D_n := \det\,[g(|x_i - x_j|)]_{1 \leqslant i,j \leqslant n+1}. \tag{7}$$

Remark. (i) The operator A is compact and hence the technical condition will always be satisfied by a slight perturbation of $r(x)$. Hence the condition does not substantially reduce the generality of the theorem. (ii) The (continuous) differentiability of r_x at $y = x$ can be seen with the aid of the Taylor expansion of r at 0 using $r'(0) = 0$.

The discussion following Proposition 1.3 gives the immediate

COROLLARY 2.2. *Suppose that* $g \in C^2[0,1]$ *with* $g'(0) \neq 0$ *satisfies the technical condition of Theorem* 2.1. *Let* $w(x)$ *be an allowable weight function. Let* D_n^w *denote the determinant* (7) *for points asymptotically distributed according to w. Then*

$$\lim_{n\to\infty} \left(\frac{|D_n^w|}{|D_n^1|}\right)^{1/n} = \exp\left(\int_0^1 \log\left(\frac{1}{w(x)}\right) w(x)\,dx\right). \tag{8}$$

Consequently, by the Arithmetic–Geometric Mean inequality, if $w \not\equiv 1$, D_n^w *is exponentially smaller than* D_n^1.

Proof. Using Theorem 2.1,

$$\begin{aligned}
\lim_{n\to\infty}\left(\frac{|D_n^w|}{|D_n^1|}\right)^{1/n} &= \lim_{n\to\infty}\left(\frac{n^n|D_n^w|}{n^n|D_n^1|}\right)^{1/n} \\
&= \frac{2|g'(0)|\exp(\int_0^1 \log(\frac{1}{w(x)})w(x)\,dx)}{2|g'(0)|\exp(\int_0^1 \log(\frac{1}{1})1\,dx)} \\
&= \frac{\exp(\int_0^1 \log(\frac{1}{w(x)})w(x)\,dx)}{\exp(\int_0^1 \log(\frac{1}{1})1\,dx)} \\
&= \exp\left(\int_0^1 \log\left(\frac{1}{(x)}\right)w(x)\,dx\right) \\
&< 1
\end{aligned}$$

if $w \not\equiv 1$. ∎

Proof of Theorem 2.1. Without loss of generality we may assume that $g'(0) = 1$. Let $G_n := [g(|x_i - x_j|)]_{1\leqslant i,j\leqslant n+1}$ and $F_n := [|x_i - x_j|]_{1\leqslant i,j\leqslant n+1}$. Our theorem, restated, is that

$$\lim_{n\to\infty}(n^n|\det(G_n)|)^{1/n} = \lim_{n\to\infty}(n^n|\det(F_n)|)^{1/n}. \tag{9}$$

We will actually show the stronger statement that

$$\lim_{n\to\infty}\frac{|\det(G_n)|}{|\det(F_n)|} =: c \tag{10}$$

exists and is strictly positive. For, from (10) it follows easily that

$$\lim_{n\to\infty}\frac{(n^n|\det(G_n)|)^{1/n}}{(n^n|\det(F_n)|)^{1/n}} = 1$$

and hence (9).

Now to show (10). Set $h_i := x_{i+1} - x_i$ as before. Since $\det(G_n)/\det(F_n) = \det(G_nF_n^{-1})$ we consider the matrix $G_nF_n^{-1}$. An elementary calculation reveals that

$$F_n^{-1} = \frac{1}{2}\begin{pmatrix} \frac{h_1 - S}{h_1 S} & \frac{1}{h_1} & 0 & 0 & \cdots & 0 & \frac{1}{S} \\ \frac{1}{h_1} & -\frac{h_1 + h_2}{h_1 h_2} & \frac{1}{h_2} & 0 & \cdots & 0 & 0 \\ 0 & \frac{1}{h_2} & -\frac{h_2 + h_3}{h_2 h_3} & \frac{1}{h_3} & 0 & \cdots & 0 \\ \cdot & & & & & & \\ \cdot & & & & & & \\ \cdot & & & & & & \\ 0 & 0 & & & \frac{1}{h_{n-1}} & -\frac{h_{n-1} + h_n}{h_{n-1} h_n} & \frac{1}{h_n} \\ \frac{1}{S} & 0 & & & 0 & \frac{1}{h_n} & \frac{h_n - S}{h_n S} \end{pmatrix},$$

where $S \coloneqq \sum_{i=1}^{n} h_i = x_{n+1} - x_1$. We may then express F_n^{-1} as the sum of a rank one matrix and a tridiagonal matrix. Specifically,

$$F_n^{-1} = \frac{1}{2S}\begin{pmatrix} 1 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 0 & 0 \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ 0 & & & 0 & 0 \\ 1 & 0 & \cdots & 0 & 1 \end{pmatrix}$$
$$+ \frac{1}{2}\begin{pmatrix} -\frac{1}{h_1} & \frac{1}{h_1} & 0 & 0 & \cdots & 0 & 0 \\ \frac{1}{h_1} & -\frac{h_1 + h_2}{h_1 h_2} & \frac{1}{h_2} & 0 & \cdots & 0 & 0 \\ 0 & \frac{1}{h_2} & -\frac{h_2 + h_3}{h_2 h_3} & \frac{1}{h_3} & 0 & \cdots & 0 \\ \cdot & & & & & & \\ \cdot & & & & & & \\ \cdot & & & & & & \\ 0 & 0 & & & \frac{1}{h_{n-1}} & -\frac{h_{n-1} + h_n}{h_{n-1} h_n} & \frac{1}{h_n} \\ 0 & 0 & & & 0 & \frac{1}{h_n} & -\frac{1}{h_n} \end{pmatrix}.$$

Note that the first and last columns of the tridiagonal part of F_n^{-1} are the coefficients of *first* divided differences at x_1, x_2 and x_n, x_{n+1}, respectively, and that the interior columns are (essentially) the coefficients of second divided differences at x_{j-1}, x_j, x_{j+1}. Thus F_n^{-1} may be considered as the discretization of a certain differential operator.

Note further that, by Definition 1.2,

$$S = \sum_{i=1}^{n} h_i = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{w(W^{-1}(c_i))}$$

and hence

$$\lim_{n\to\infty} S = \int_0^1 \frac{1}{w(W^{-1}(x))}\, dx = 1,$$

by the change of variables $t = W^{-1}(x)$.

For the sake of simplicity of presentation, let us assume that, in fact, $S = 1$, i.e., that $x_1 = 0$ and $x_{n+1} = 1$.

Now, since by assumption, $g(x) = x + r(x)$,

$$G_n = [g(|x_i - x_j|)] = [|x_i - x_j| + r(|x_i - x_j|)] = F_n + R_n,$$

where

$$R_n := [r(|x_i - x_j|)]_{1 \leqslant i,j \leqslant n+1}.$$

It will be convenient for us to write R_n in the form

$$R_n = \begin{pmatrix} r_1(x_1) & r_1(x_2) & \cdots & r_1(x_{n+1}) \\ r_2(x_1) & r_2(x_2) & \cdots & r_2(x_{n+1}) \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ r_{n+1}(x_1) & r_{n+1}(x_2) & \cdots & r_{n+1}(x_{n+1}) \end{pmatrix},$$

where

$$r_i(x) := r(|x - x_i|).$$

Hence, $G_nF_n^{-1} = I_n + R_nF_n^{-1}$ and we calculate

$$R_nF_n^{-1} = \frac{1}{2}\begin{pmatrix} r_1(x_1)+r_1(x_{n+1}) & 0 & \cdots & 0 & r_1(x_1)+r_1(x_{n+1}) \\ r_2(x_1)+r_2(x_{n+1}) & 0 & \cdots & 0 & r_2(x_1)+r_2(x_{n+1}) \\ \cdot & 0 & \cdots & 0 & \cdot \\ \cdot & 0 & \cdots & 0 & \cdot \\ r_{n+1}(x_1)+r_{n+1}(x_{n+1}) & 0 & \cdots & 0 & r_{n+1}(x_1)+r_{n+1}(x_{n+1}) \end{pmatrix}$$
$$+ \frac{1}{2}\big(r_i[x_1,x_2] \quad \ldots \quad (h_j+h_{j-1})r_i[x_{j-1},x_j,x_{j+1}] \quad \ldots \quad -r_i[x_n,x_{n+1}] \big).$$

Here, and in the sequel, with an abuse of notation, I_n denotes the $(n+1)(n+1)$ identity matrix. The last displayed matrix represents the entries of the ith row and jth column with $1 \leqslant i \leqslant n+1$ and $2 \leqslant j \leqslant n$. The expressions $r_i[x_1,x_2]$ and $r_i[x_n,x_{n+1}]$ denote the first divided difference of r_i at x_1,x_2 and x_n,x_{n+1}, respectively, whereas $r_i[x_{j-1},x_j,x_{j+1}]$ denote the second divided differences of r_i at x_{j-1},x_j,x_{j+1}.

Now, since $r'(0)=0$, each $r_i \in C^2[0,1]$ and hence we may write the divided differences of r_i in terms of first and second derivatives, i.e.

$$R_nF_n^{-1} = \frac{1}{2}\begin{pmatrix} r_1(x_1)+r_1(x_{n+1}) & 0 & \cdots & 0 & r_1(x_1)+r_1(x_{n+1}) \\ r_2(x_1)+r_2(x_{n+1}) & 0 & \cdots & 0 & r_2(x_1)+r_2(x_{n+1}) \\ \cdot & 0 & \cdots & 0 & \cdot \\ \cdot & 0 & \cdots & 0 & \cdot \\ r_{n+1}(x_1)+r_{n+1}(x_{n+1}) & 0 & \cdots & 0 & r_{n+1}(x_1)+r_{n+1}(x_{n+1}) \end{pmatrix}$$
$$+ \frac{1}{2}\left(r_i'(d_i) \quad \ldots \quad (h_j+h_{j-1})\frac{r_i''(c_{ij})}{2} \quad \ldots \quad -r_i'(e_i) \right) \tag{11}$$

for some $c_{ij} \in [x_{j-1},x_{j+1}]$, $d_i \in [x_1,x_2]$ and $e_i \in [x_n,x_{n+1}]$.

Applying $R_nF_n^{-1}$ to the vector of function evaluations $(f(x_1),f(x_2),\ldots,f(x_{n+1}))^T$ we get

$$R_nF_n^{-1}\begin{pmatrix} f(x_1) \\ f(x_2) \\ \vdots \\ f(x_{n+1}) \end{pmatrix}$$

$$= \frac{1}{2}\begin{pmatrix} (r_1(x_1)+r_1(x_{n+1}))(f(x_1)+f(x_{n+1})) \\ (r_2(x_1)+r_2(x_{n+1}))(f(x_1)+f(x_{n+1})) \\ \vdots \\ (r_{n+1}(x_1)+r_{n+1}(x_{n+1}))(f(x_1)+f(x_{n+1})) \end{pmatrix}$$

$$+ \frac{1}{2}\begin{pmatrix} r_1'(d_1)f(x_1) - r_1'(e_1)f(x_{n+1}) \\ +\sum_{j=2}^{n} \frac{(h_j+h_{j-1})}{2} r_1''(c_{1j})f(x_j) \\ \vdots \\ r_{n+1}'(d_{n+1})f(x_1) - r_{n+1}'(e_{n+1})f(x_{n+1}) \\ +\sum_{j=2}^{n} \frac{(h_j+h_{j-1})}{2} r_{n+1}''(c_{n+1,j})f(x_j) \end{pmatrix}.$$

Note that according to our assumption on S, $x_1 = 0$ and $x_{n+1} = 1$. Now the sums in the second matrix again can be interpreted as Riemann sums, so that $R_n F_n^{-1}$ is a discrete approximation of the operator $A : C[0,1] \to C[0,1]$ given by

$$\begin{aligned}(Af)(x) = &\frac{1}{2}(r_x(0)+r_x(1))(f(0)+f(1)) \\ &+ \frac{1}{2}\left(r_x'(0)f(0) + \int_0^1 r_x''(y)f(y)dy - r_x'(1)f(1)\right), \qquad (12)\end{aligned}$$

where

$$r_x(y) := r(|y-x|),$$

i.e., that given in the statement of Theorem 2.1, Eq. (5).

We may consider A as a (generalized) integral operator

$$(Af)(x) = \int_0^1 K(x,y)f(y)\,dy$$

with kernel

$$\begin{aligned}K(x,y) := &\tfrac{1}{2}\{[r_x(0)+r_x(1)+r_x'(0)]\delta_0(y) \\ &+ r_x''(y) + [r_x(0)+r_x(1)-r_x'(1)]\delta_1(y)\}. \qquad (13)\end{aligned}$$

Here $\delta_0(y)$ denotes the Dirac delta function at $y = 0$ and δ_1 that at $y = 1$.

For simplicity's sake, denote

$$A_n \coloneqq R_n F_n^{-1}$$

and consider the matrix $I_n + \lambda A_n$, where $\lambda \in \mathbb{C}$. Its determinant

$$D_n(\lambda) \coloneqq \det(I_n + \lambda A_n)$$

may be expanded as

$$D_n(\lambda) = \sum_{k=0}^{n+1} \frac{a_k(n)}{k!} \lambda^k$$

for certain coefficients $a_k(n)$ (with $a_0(n) = 1$) which (cf. [8, p. 237]) by standard determinantal identities may be expressed as

$$a_k(n) = \sum_{i_1,\dots,i_k} A_n(i_1, \dots, i_k), \tag{14}$$

where $A_n(i_1, \dots, i_k)$ denotes the determinant obtained by deleting from A_n all of its rows and columns except those labelled $i_1, \dots, i_k$. Explicitly,

$$A_n(i_1, \dots, i_k) = \begin{vmatrix} (A_n)_{i_1 i_1} & \cdot & \cdot & (A_n)_{i_1 i_k} \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ (A_n)_{i_k i_1} & \cdot & \cdot & (A_n)_{i_k i_k} \end{vmatrix}. \tag{15}$$

The sum in (14) is taken over all distinct $1 \leqslant i_1, i_2, \dots, i_k \leqslant n+1$. Note also that determinant (15) is invariant under permutations of the tuple $(i_1, i_2, \dots, i_k)$.

LEMMA 2.3. *Suppose that K is given by* (13) *and $a_k(n)$ by* (14). *Then for each fixed k,*

$$\lim_{n\to\infty} a_k(n) = \alpha_k \coloneqq \int_0^1 \cdots \int_0^1 K\begin{pmatrix} s_1, s_2, \dots, s_k \\ s_1, s_2, \dots, s_k \end{pmatrix} ds_1 \cdots ds_k,$$

where, as is standard,

$$K\begin{pmatrix} s_1, s_2, \dots, s_k \\ t_1, t_2, \dots, t_k \end{pmatrix} \coloneqq \begin{vmatrix} K(s_1, t_1) & \cdot & \cdot & K(s_1, t_k) \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ K(s_k, t_1) & \cdot & \cdot & K(s_k, t_k) \end{vmatrix}.$$

Proof. Were it not for the presence of the delta functions in K, this result would be completely standard in the theory of integral equations (cf. [8, Chapter 6]). Even then, it is but a minor extension but we outline a proof for the sake of completeness.

First, note that α_k is indeed well-defined (despite the delta functions), since, as can be seen by expanding along the rows of $K\binom{s_1,s_2,\ldots,s_k}{s_1,s_2,\ldots,s_k}$, a delta of a certain variable s_i is only possibly multiplied by deltas of *different* variables. Perhaps the easiest way to understand the proof is by considering the first few $a_k's$.

Now,

$$\begin{aligned}
a_1(n) &= \sum_{i_1=1}^{n+1} (A_n)_{i_1 i_1} \\
&= \frac{r_1(x_1) + r_1(x_{n+1})}{2} + \frac{1}{2}\sum_{i=2}^{n} (h_i + h_{i-1}) r_i[x_{i-1}, x_i, x_{i+1}] \\
&\quad + \frac{r_{n+1}(x_1) + r_{n+1}(x_{n+1})}{2} + \frac{1}{2} r_1[x_1, x_2] - \frac{1}{2} r_{n+1}[x_n, x_{n+1}] \\
&= \frac{r_1(x_1) + r_1(x_{n+1})}{2} + \frac{1}{2}\sum_{i=2}^{n} (h_i + h_{i-1}) \frac{r_i''(c_{ii})}{2} \\
&\quad + \frac{r_{n+1}(x_1) + r_{n+1}(x_{n+1})}{2} + \frac{1}{2} r_1'(d_1) - \frac{1}{2} r_{n+1}'(e_{n+1})
\end{aligned}$$

for some $c_{ii} \in [x_{i-1}, x_{i+1}]$, $d_1 \in [x_1, x_2]$, and $e_{n+1} \in [x_n, x_{n+1}]$ by the mean value property of divided differences.

But, explicitly,

$$r_i''(c_{ii}) = \frac{d^2 r}{dy^2}(|x_i - y|)\Big|_{y=c_{ii}}$$

so that $\sum_{i=2}^{n} (h_i + h_{i-1}) \frac{r_i''(c_{ii})}{2}$ is but a Riemann sum for the integral $\int_0^1 r_x''(x)\,dx$. Further, $r_1'(d_1) = r'(d_1) \to r'(0) = 0$ and $r_{n+1}'(e_{n+1}) = -r'(1 - e_{n+1}) \to -r'(0) = 0$. Hence, since we have set $x_1 = 0$ and $x_{n+1} = 1$,

$$\lim_{n\to\infty} a_1(n) = \frac{r(0) + r(1)}{2} + \frac{1}{2}\int_0^1 r_x''(x)\,dx + \frac{r(1) + r(0)}{2}. \tag{16}$$

We wish to show that this expression is equal to

$$\alpha_1 = \int_0^1 K(x, x)\,dx.$$

But, from (13),

$$\int_0^1 K(x,x)\,dx = \frac{1}{2}\int_0^1 [r_x(0)+r_x(1)+r_x'(0)]\delta_0(x)+r_x''(x)$$
$$+\ [r_x(0)+r_x(1)-r_x'(1)]\delta_1(x)\,dx$$
$$=\frac{1}{2}\left\{\left[r_0(0)+r_0(1)+r_0'(0)]+\int_0^1 r_x''(x)\,dx.\right.\right.$$
$$\left.\left.+\,[r_1(0)+r_1(1)-r_1'(1)\right]\right\}.$$

This is easily seen to equal the right-hand side of (16) by noting that $r_0(0)=r(|0-0|)=r(0)$, $r_0(1)=r(|0-1|)=r(1)$, $r_0'(0)=r'(0)=0$, $r_1'(1)=r'(|1-1|)=r'(0)=0$, $r_1(0)=r(1)$ and $r_1(1)=r(0)$.

Similarly,

$$a_2(n)=\sum_{i_1=1}^{n+1}\sum_{i_2=1}^{n+1}\begin{vmatrix}(A_n)_{i_1i_1} & (A_n)_{i_1i_2}\\ (A_n)_{i_2i_1} & (A_n)_{i_2i_2}\end{vmatrix}$$

is a tensor-product Riemann sum for

$$\int_0^1\int_0^1\begin{vmatrix}K(s_1,s_1) & K(s_1,s_2)\\ K(s_2,s_1) & K(s_2,s_2)\end{vmatrix}ds_1\,ds_2$$

and so on. ∎

The following estimate is also standard in the theory of integral equations, complicated only slightly by the presence of the "deltas" in the first and last columns of A_n.

LEMMA 2.4. *Under our assumptions on* $r(x)$, *there is a constant* $M>0$ *such that*

$$|a_k(n)|\leqslant M^k k^{(k+6)/2},\quad k,n=1,2,\ldots\ .$$

Proof. First, let

$$M_1 := \max_{0\leqslant y\leqslant 1}\max\{|r(y)|,|r'(y)|,|r''(y)|\}$$

so that the entries in the first and last columns of A_n (cf. (11)) are bounded by $3M_1/2$ and the entries of column j, $2\leqslant j\leqslant n$, are bounded by $\frac{M_1}{2}(\frac{h_j+h_{j-1}}{2})$. Further, our assumptions on the points x_i imply that there is some constant

$C > 0$ such that

$$h_j \leqslant \frac{C}{n}, \quad 1 \leqslant j \leqslant n$$

and hence the entries of column j, $2 \leqslant j \leqslant n$, are bounded by $\frac{M_1}{2}\frac{C}{n}$.

Now set

$$M_2 := \max\left\{\frac{3M_1}{2}, \frac{M_1}{2}C, 1\right\}$$

so that

$$|(A_n)_{ij}| \leqslant \begin{cases} M_2 & \text{if } j = 1 \text{ or } j = n+1, \\ \frac{M_2}{n} & \text{if } 2 \leqslant j \leqslant n. \end{cases}$$

We now proceed to bound $|A_n(i_1, i_2, \ldots, i_k)|$, for which we must distinguish three cases:

Case 1: *The indices* 1 *and* $n+1$ *both do not appear among* $\{i_1, i_2, \ldots, i_k\}$. In this case, the 2-norm of each row of $A_n(i_1, i_2, \ldots, i_k)$ is bounded by $\frac{M_2}{n}k^{1/2}$ so that, by Hadamard's inequality,

$$|A_n(i_1, i_2, \ldots, i_k)| \leqslant \frac{M_2^k k^{k/2}}{n^k}. \tag{17}$$

Case 2: *Exactly one of the indices* 1 *and* $n+1$ *appear among* $\{i_1, i_2, \ldots, i_k\}$. Suppose that $i_s = 1$ or $n+1$. Then we may expand down the column corresponding to i_s to obtain k determinants of dimension $(k-1) \times (k-1)$, each of which does not involve the first or last columns of A_n, and hence are each bounded by $\frac{M_2^{k-1}(k-1)^{(k-1)/2}}{n^{k-1}}$. Thus, in this case,

$$|A_n(i_1, i_2, \ldots, i_k)| \leqslant k\frac{M_2^{k-1}(k-1)^{(k-1)/2}}{n^{k-1}}. \tag{18}$$

Case 3: *Both* 1 *and* $n+1$ *appear among* $\{i_1, i_2, \ldots, i_k\}$. In this case we expand as before but down both the corresponding columns to obtain k^2 determinants of dimension $(k-2) \times (k-2)$, each of which is bounded by

$$\frac{M_2^{k-2}(k-2)^{(k-2)/2}}{n^{k-2}}$$

so that, in this case,

$$|A_n(i_1, i_2, \ldots, i_k)| \leqslant k^2\frac{M_2^{k-2}(k-2)^{(k-2)/2}}{n^{k-2}}. \tag{19}$$

Now, to bound $|a_k(n)|$ we simply note that there are $(n-1)^k$ possibilities for case (1), $2\binom{k}{1}(n-1)^{k-1}$ possibilities for case (2) and $\binom{k}{2}(n-1)^{k-2}$ for case (3). Hence, by (17)–(19) we have

$$\begin{aligned}|a_k(n)| &\leqslant (n-1)^k \frac{M_2^k k^{k/2}}{n^k} + 2\binom{k}{1}(n-1)^{k-1} k \frac{M_2^{k-1}(k-1)^{(k-1)/2}}{n^{k-1}} \\ &\quad + \binom{k}{2}(n-1)^{k-2} k^2 \frac{M_2^{k-2}(k-2)^{(k-2)/2}}{n^{k-2}} \\ &\leqslant M^k k^{(k+6)/2}\end{aligned}$$

for suitably chosen M. ∎

In particular, it follows that we may also bound

$$\alpha_k = \lim_{n\to\infty} a_k(n)$$

by

$$|\alpha_k| \leqslant M^k k^{(k+6)/2}.$$

Then, using the fact that $k! \geqslant k^k/e^k$, we have

$$\left|\frac{\alpha_k}{k!}\right| \leqslant M^k e^k k^{(6-k)/2}$$

and hence

$$D(\lambda) := \sum_{k=0}^{\infty} \frac{\alpha_k}{k!} \lambda^k$$

has infinite radius of convergence and is an entire function. $D(\lambda)$ is nothing more than the Fredholm determinant of our operator A.

We now quote a lemma of Hilbert, in slightly modified form.

LEMMA 2.5. (Hilbert [7, Hilfsatz 1, p. 9]). *With the above notation and associated assumptions*

$$\lim_{n\to\infty} D_n(\lambda) = D(\lambda),$$

uniformly in λ *on compact subsets of* $\mathbb{C}$.

Proof. Consider $|\lambda| \leqslant R$ and let $\varepsilon > 0$ be given.

First, choose m so large that

$$\left|\sum_{k=m+1}^{\infty} M^k e^k k^{(k-6)/2} \lambda^k\right| \leqslant \varepsilon/3$$

for all $|\lambda| \leqslant R$ so that, by Lemma 2.4,

$$\left|\sum_{k=m+1}^{\infty} \frac{\alpha_k}{k!} \lambda^k\right| \leqslant \varepsilon/3 \tag{20}$$

and

$$\left|D_n(\lambda) - \sum_{k=0}^{m} \frac{a_k(n)}{k!} \lambda^k\right| = \left|\sum_{k=m+1}^{n+1} \frac{a_k(n)}{k!}\right| \leqslant \varepsilon/3 \tag{21}$$

for $n > m$.

Then, for such a fixed m, choose n so large that

$$\left|\sum_{k=0}^{m} \frac{a_k(n)}{k!} \lambda^k - \sum_{k=0}^{m} \frac{\alpha_k}{k!} \lambda^k\right| \leqslant \varepsilon/3 \tag{22}$$

(which we may do, since by Lemma 2.3, $a_k(n) \to \alpha_k$ for each k). Then, from (20) to (22) it follows that

$$\begin{aligned} |D_n(\lambda) - D(\lambda)| &= \left|D_n(\lambda) - \sum_{k=0}^{\infty} \frac{\alpha_k}{k!} \lambda^k\right| \\ &\leqslant \left|D_n(\lambda) - \sum_{k=0}^{m} \frac{a_k(n)}{k!} \lambda^k\right| + \left|\sum_{k=0}^{m} \frac{a_k(n)}{k!} \lambda^k - \sum_{k=0}^{m} \frac{\alpha_k}{k!} \lambda^k\right| \\ &\quad + \left|\sum_{k=m+1}^{\infty} \frac{\alpha_k}{k!} \lambda^k\right| \\ &\leqslant \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \qquad \blacksquare \end{aligned}$$

Returning now to the proof of Theorem 2.1, note that

$$\begin{aligned} \frac{\det(G_n)}{\det(F_n)} &= \det(G_n F_n^{-1}) \\ &= \det(I_n + R_n F_n^{-1}) \\ &= D_n(\lambda) \end{aligned}$$

for $\lambda = 1$. Thus, by Lemma 2.5,

$$\lim_{n\to\infty} \frac{\det(G_n)}{\det(F_n)} = D(1),$$

which is not zero precisely when -1 ($-g'(0)$ in our case) is *not* an eigenvalue of the operator A. ∎

Remark. It is natural to ask whether our main result also holds for functions g with $g'(0) = 0$. However, our argument is that for $g'(0) \neq 0$ we may reduce to the basic case $g(x) = x$. The function $g(x) = x^3$, for example, is already fundamentally different. If we consider even the case of equally spaced points, the entries of $[|x_i - x_j|^3]$ are of order n^{-3} while those of $[|x_i - x_j|]$ are of order n^{-1}. Hence we cannot reduce the case of $g(x) = x^3$ to that of $g(x) = x$ using the arguments presented here. Nevertheless, numerical experiments indicate that the Fekete type points for $g(x) = x^3$ (and other functions with $g'(0) = 0$) do remain asymptotically equally spaced; it is just that our arguments do not directly apply to this case. Hence an analysis of the $g'(0) = 0$ case remains an intriguing open problem, which we hope to be the subject of a subsequent paper.

REFERENCES

1. L. P. Bos and U. Maier, On the asymptotics of points which maximize determinants of the form $det[g(|x_i - x_j|)]$, *in* "Advances in Multivariate Approximation", Mathematical Research, (W. Haussmann *et al.* Eds.), Vol. 107, pp. 107–128, Wiley-Vch, Berlin, 1999.
2. M. D. Buhmann, Multivariate cardinal interpolation with radial basis functions, *Constr. Approx.* **6** 1990, 225–255.
3. M. D. Buhmann, "Radial Basis Functions, Acta Numerica," Vol. 9, pp. 1–38, Cambridge Univ. Press, Cambridge, UK, 2000.
4. N. Dyn, Interpolation and Approximation by radial and related functions, *in* "Approximation Theory VI" (C. K. Chui, L. L. Schumaker, J. D. Ward, Eds.), pp. 211–234, Academic Press, New York, 1989.
5. N. Dyn, W. A. Light, and E. W. Cheney, Interpolation by piecewise-linear radial basis functions I, *J. Approx. Theory* **59** 1989, 202–223.
6. W. Hackbusch, "Integralgleichungen, Theorie und Numerik," B. G. Teubner, Stuttgart, 1997.
7. D. Hilbert, "Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen," B. G. Teubner, Leipzig, 1912. `http://cdl.library.cornell.edu/math_H.html`.
8. H. Hochstadt, "Integral Equations," Wiley, New York, 1973.
9. A. Lasota and M. C. Mackey, "Chaos, Fractals and Noise," Springer, Berlin, 1994.
10. M. Reimer, "Constructive Theory of Multivariate Functions," Wissenschaftverlag, Mannheim/Wien/Zürich, 1990.
11. E. B. Saff and V. Totik, "Logarithmic Potentials with External Fields," Grundlehren der math, Wissenschaften **316**, Springer-Verlag, Berlin Heidelberg, 1997.