ISSN 1066-5307, Mathematical Methods of Statistics, 2018, Vol. 27, No. 4, pp. 245–267. © Allerton Press, Inc., 2018.

On Optimal Cardinal Interpolation

B. Levit[1*]

[1] *Dept. Math. and Statist., Queen's Univ., Kingston ON, Canada*

Received June 19, 2018; in final form, October 2, 2018

Abstract—For the Hardy classes of functions analytic in the strip around real axis of a size 2β, an optimal method of cardinal interpolation has been proposed within the framework of Optimal Recovery [12]. Below this method, based on the Jacobi elliptic functions, is shown to be optimal according to the criteria of Nonparametric Regression and Optimal Design.

In a stochastic non-asymptotic setting, the maximal mean squared error of the optimal interpolant is evaluated explicitly, for all noise levels away from 0. A pivotal role is played by the interference effect, in which the oscillations exhibited by the interpolant's bias and variance mutually cancel each other. In the limiting case $\beta \to \infty$, the optimal interpolant converges to the well-known Nyquist—Shannon cardinal series.

Keywords: cardinal interpolation, Optimal Recovery, Hardy classes, Jacobi elliptic functions, infinite Blaschke product, interference effect, sinc filter.

2000 Mathematics Subject Classification: primary 62G08; secondary 62K05, 33E05, 42A15.

DOI: 10.3103/S1066530718040014

1. INTRODUCTION

In cardinal interpolation, one seeks to recover a function $f(x)$, $x \in \mathbb{R}$, from its values at equidistant nodes $x_j = jh$, $h > 0$, $j \in \mathbb{Z}$. The values $f(x_j)$ may be corrupted by random errors. The set of nodes $\mathcal{X} = \{x_j\}$ is referred to as *cardinal design*. In the past, the area of *cardinal interpolation* contributed generously to developing effective mathematical tools and ideas, in the area at a crossroads of *Approximation Theory, Signal Analysis, and Statistics*.

Early on, such ideas centered around the famous *cardinal series*, or *sinc filter*, variously named after E. Borel (1898), E.T. Whittaker (1915), H. Nyquist (1928), V.A. Kotel'nikov (1933), C.E. Shannon (1949), and others. Its central result, the celebrated *sampling theorem*, is still popular in the communication theory and signal processing; see, e.g., [15], or [9], Section 20.2, for a quick reference. Connections between optimal cardinal interpolation for Hardy classes and the *sampling theorem* will be discussed below in Sections 3.2 and 4.3.

Later on, cardinal interpolation gained additional momentum due to the ground breaking book [17] focusing on *cardinal spline interpolation*. The book, explicitly linked to the cardinal series in the Introduction, became a driving force in developing the general theory of splines in the 70s and 80s. Towards the end of the century, powerful tools of *Optimal Recovery* were introduced to the field of cardinal interpolation. They allowed to find linear methods of interpolation, optimal among all linear as well as non-linear recovery methods, for various functional classes, including *Hardy classes* of analytic functions [14].

Cardinal design figured prominently in non-parametric regression as well; see, e.g., [2], [5], [6]. In such models, interpolation methods are often preceded by a smoothing procedure to reduce the noise level; cf. [8]. In the 80s and 90s, such statistical methods have been primarily used in the asymptotic setting $h \to 0$.

At that time, asymptotic approach seemed very attractive, offering a venue for studying various nonparametric tools, such as kernel estimates, splines, wavelets, polynomial and rational interpolants, in various settings including both *equidistant* and *non-equidistant* design. It turned out, however, that,

[*] E-mail: blevit@queensu.ca

with properly chosen parameters, all of the above methods of estimation were *asymptotically efficient*. Thus, a practitioner wondering which of the estimators to use in any particular problem, often had to rely primarily on their numerical comparison.

Gradually, the enthusiasm for the asymptotic approach started to lessen at the turn of the century. Two developments were affecting such a change in the general approach to nonparametric problems. The already mentioned progress in Optimal Recovery allowed to determine optimal interpolants explicitly for some selected classes, including classes of analytic functions. Although in such problems special *elliptic functions* were often used, there is hope that the general approach can be applied to more mundane functional classes as well.

Independently, it has been demonstrated that the exact variance of the optimal interpolants could also be found explicitly [10], [11]. As a result of comparing the exact expressions of the bias and the variance, a very useful phenomenon of *interference* sprang to life: the oscillations exhibited by the interpolant's bias and variance can mutually cancel each other [12]. This offers a new way of resolving the long standing problem of balancing variance vs. bias, which had previously all but dictated an asymptotic approach. Thus, finding non-asymptotically optimal methods of interpolation became possible, in the models of non-parametric regression.

Although [14] did not treat any statistical estimation problems, a large portion of the present paper is based on it. Therefore, for reader's convenience, the author thought it worthwhile to discuss some of its results below in Sections 2.4 and 3. This seems to make sense, at least for two reasons. First, the corresponding results, in the author's view, exhibit a hidden jewel of classical Analysis. Second, since [14] was originally eyeing a different audience, an interested statistician could have — as the author has experienced himself — a rather hard time trying to fill in all the intermediate details.

The paper carries on with some ideas and definitions of [12]. The reader is presumed to be familiar with the classical Jacobi functions. Various sources provide an excellent introduction to elliptic functions; see, e.g., [1], [3]. User-oriented overviews of the Jacobi elliptic functions have been offered in [12], Ch. 3, and [14], Ch. 5.4. Additionally, a brief summary of the necessary definitions and results is appended below in Section 5.

2. INTERPOLATION WITH COUNTABLE NODES

2.1. Cardinal Interpolation in the Strip

Consider the following model of random data,

$$y_j = f(j) + e_j, \qquad j \in \mathbb{Z}, \tag{2.1}$$

where the values $f_j =: f(j)$ of an unknown function f are observed at given nodes $x_j = j$, $j \in \mathbb{Z}$, in the presence of a discrete *white noise*,

$$\mathbf{E}\, e_j = 0, \tag{2.2}$$

$$\mathrm{Cov}(e_i, e_j) = \begin{cases} \sigma^2 & \text{if} \quad i = j, \\ 0 & \text{if} \quad i \neq j. \end{cases} \tag{2.3}$$

Denote $\mathbf{y} = \{y_j\}$, $\mathbf{f} = \{f_j\}$, and $\mathbf{e} = \{e_j\}$, $j \in \mathbb{Z}$. The parameter $\sigma \geq 0$ is assumed unknown. In the case $\sigma = 0$, one is dealing with a deterministic setting well known in the Approximation Theory. The more general cardinal design $x_j = jh$, $h > 0$, will be transformed to the present case later, see Remark 4.3.

Let $\mathcal{F}_\beta$ denote the class of functions $f(w)$ real on the real line and analytic in the strip $S_\beta = \{w \in \mathbb{C} \colon |\mathrm{Im}\, w| < \beta\}$. For Q, $\beta > 0$, the Hardy class $F_\beta(Q)$ is defined by

$$F_\beta(Q) = \{f \in \mathcal{F}_\beta \colon \sup_{S_\beta} |f(w)| \leq Q\}. \tag{2.4}$$

The pair $(\mathcal{X}, F_\beta(Q))$ consisting of the design set $\mathcal{X} = \{x_j = j\}$ and the functional class $F_\beta(Q)$ defines the model at hand.

The paper will be dealing primarily with linear methods of *cardinal interpolation* which can be applied equally well to deterministic and random data. Let us describe such methods in more detail. By definition, *fundamental interpolating functions* $L_j(w)$, $w \in S_\beta$, $j \in \mathbb{Z}$, satisfy, for all $i, j \in \mathbb{Z}$,

$$L_j(i) = \begin{cases} 1, & i = j, \\ 0, & i \neq j, \end{cases} \tag{2.5}$$

$$L_j(\cdot) \text{ is real} - \text{valued for } w \in \mathbb{R}, \tag{2.6}$$

$$\sum_{j\in\mathbb{Z}} L_j(\cdot) \quad \text{converges locally uniformly in } S_\beta. \tag{2.7}$$

It is clear from the definition, that the following linear operator,

$$(I\mathbf{f})(w) \equiv I\mathbf{f}(w) = \sum_{j\in\mathbb{Z}} L_j(w) f_j \tag{2.8}$$

interpolates a function f on $\mathbb{Z}$, i.e., $I\mathbf{f}(i) \equiv f(i)$. Any such operator will be referred to as a cardinal linear interpolation formula, or shortly, linear *interpolant*.

Applying the interpolant I to the random responses $\mathbf{y}$ in (2.1), one gets

$$(I\mathbf{y})(w) = \sum_{j\in\mathbb{Z}} L_j(w) y_j. \tag{2.9}$$

Equations (2.8)–(2.9) will be viewed as the same operator applied to different sets of data: deterministic, $\mathbf{f}$, or random, $\mathbf{y}$. Denote $\mathcal{I}$ the class of all linear interpolants I in (2.8)–(2.9) satisfying (2.5)–(2.7). Some basic general results concerning infinite interpolation will be briefly discussed in Section 2.4.

Note for now that for all $I \in \mathcal{I}$, $f \in F_\beta(Q)$, and $x \in \mathbb{R}$,

$$\mathbf{E}(I\mathbf{y}(x)) \equiv (I\mathbf{f})(x) = \sum_{j\in\mathbb{Z}} L_j(x) f_j,$$

and

$$\mathrm{Var}(I\mathbf{y}(x)) = \sigma^2 \sum_{j\in\mathbb{Z}} L_j^2(x) =: \sigma^2 s(x). \tag{2.10}$$

The bias of an interpolant $I \in \mathcal{I}$,

$$b(x) =: \mathbf{E}(I\mathbf{y}(x)) - f(x) = I\mathbf{f}(x) - f(x),$$

coincides with the *interpolation error* in the deterministic problem corresponding to $\sigma = 0$. The *mean squared error* (MSE) of an interpolant $I \in \mathcal{I}$ satisfies

$$\mathbf{E}(I\mathbf{y}(x) - f(x))^2 = \mathrm{Var}\,(I\mathbf{y}(x)) + b^2(x) = \sigma^2 \sum_{j\in\mathbb{Z}} L_j^2(x) + (I\mathbf{f}(x) - f(x))^2. \tag{2.11}$$

For a better part of the paper, the fundamental interpolating functions will be of the from $L_j(w) = L(w - j)$, where $L(w)$, $w \in \mathbb{C}$, is the so-called *interpolating kernel* satisfying

$$L(i) = \begin{cases} 1, & i = 0, \\ 0, & i \neq 0. \end{cases} \tag{2.12}$$

Remark 2.1. Often, a kernel $L(w)$ will itself be an analytic function,

$$L \in \mathcal{F}_\beta, \tag{2.13}$$

satisfying the following mild restriction: for any compact set $K \subset S_\beta$, there are $C > 0$ and $\delta > 0$ such that

$$|L(w + x)| \leq \frac{C}{1 + |x|^{3/2+\delta}}, \qquad w \in K, \quad x \in \mathbb{R}. \tag{2.14}$$

Under these assumptions, the interpolant (2.9) satisfies (2.7) and defines a function belonging to $\mathcal{F}_\beta$ almost surely. Indeed, by (2.4), (2.3), and the *Markov inequality*, for some $c > 0$ and all j large enough,

$$\mathbf{P}(|y_j| > j^{1/2+\delta/2}) \leq \mathbf{P}\Big(|e_j| > \frac{1}{2} j^{1/2+\delta/2}\Big) \leq \frac{c}{j^{1+\delta}}.$$

Thus, by the *Borel–Cantelli lemma*, see, e.g., [18], Chap. 2, Sect. 10, for all but finitely many indices j, $|y_j| \leq j^{1/2+\delta/2}$ almost surely. Together with (2.14), this implies that the interpolant (2.9) converges locally uniformly in S_β. It remains to use the *Weierstrass theorem* on uniformly converging families of analytic functions; see, e.g., [13], Section **I**. 76.

Note that the variance $\mathrm{Var}\, I\mathbf{y}(x)$ in (2.10) does not depend on f, unlike the MSE (2.11). Obviously, by (2.5),

$$\sup_{x\in\mathbb{R}} \mathbf{E}(I\mathbf{y}(x) - f(x))^2 \geq \sup_{x\in\mathbb{R}} \mathrm{Var}\ I\mathbf{y}(x) \geq \sigma^2. \tag{2.15}$$

This natural lower bound motivates the following definitions, cf. [12].

Definition 2.1. An interpolant $I\mathbf{y}(x)$ in (2.9) is called **(a)** D-optimal if

$$\sup_{x\in\mathbb{R}} \mathrm{Var}\ I\mathbf{y}(x) = \sigma^2, \qquad \sigma \geq 0;$$

(b) R-optimal for a given $\sigma \geq 0$, w.r.t. the model $(\mathcal{X}, F_\beta(Q))$ if

$$\sup_{f\in F_\beta(Q)} \sup_{x\in\mathbb{R}} \mathbf{E}(I\mathbf{y}(x) - f(x))^2 = \sigma^2.$$

Note that by (2.10) the property of D-optimality is equivalent to

$$s(x) \leq 1, \qquad x \in \mathbb{R}, \tag{2.16}$$

and, therefore, is independent of σ. Obviously, R-optimality, for any given $\sigma > 0$, automatically implies D-optimality. The inverse is not true, as will be seen below. Constructing R-optimal interpolants for Hardy classes is the main goal of this paper. Below it will be referred to as

Problem 1. Find an R-optimal (hence, necessarily, a D-optimal) linear interpolant (2.9) for the model $(\mathcal{X}, F_\beta(Q))$ given by (2.1), (2.4).

First, the deterministic case $\sigma = 0$ will be treated drawing on some recent results from Optimal Recovery [14]. The following definition is essentially again borrowed from [12].

Definition 2.2. An interpolant $I^* \in \mathcal{I}$ is called **(a)** $(\mathcal{X}, F_\beta(Q))$-optimal at a given point $w \in S_\beta$ if

$$\inf_{I\in\mathcal{I}} \sup_{f\in F_\beta(Q)} |I\mathbf{f}(w) - f(w)| = \sup_{f\in F_\beta(Q)} |I^*\mathbf{f}(w) - f(w)|;$$

(b) A-optimal, for the given model $(\mathcal{X}, F_\beta(Q))$, if

$$r_0 =: \inf_{I\in\mathcal{I}} \sup_{f\in F_\beta(Q)} \sup_{x\in\mathbb{R}} |I\mathbf{f}(x) - f(x)| = \sup_{f\in F_\beta(Q)} \sup_{x\in\mathbb{R}} |I^*\mathbf{f}(x) - f(x)|.$$

A-optimal interpolants, for any $Q, \beta > 0$, will be discussed below in Section 3.2. This deterministic problem will be referred to as

Problem $1_\circ$. Find an A-optimal interpolant, with respect to the model $(\mathcal{X}, F_\beta(Q))$.

Solution to **Problem $1_\circ$** is based on some recent developments in Optimal Recovery [14]. As a tribute to its powerful tools, some highlights of this theory will be briefly discussed in Section 2.4.

2.2. Countable Interpolation in the Disk

It turns out that optimal interpolation in a strip S_β is closely related to the optimal interpolation in the unit disk S of the complex plain $\mathbb{C}$. Since the general setting is quite similar, it will be discussed here very briefly.

Consider the unit disk

$$S = \{z \in \mathbb{C} \colon |z| < 1\}. \tag{2.17}$$

Following [14], denote $\mathcal{H}$ the *Hardy space* of all functions $h = h(z)$ analytic in S and real on $(-1, 1)$ such that

$$\sup_{z \in S} |h(z)| < \infty. \tag{2.18}$$

Let $H(Q)$ be the *Hardy ball* of functions $h \in \mathcal{H}$ such that

$$\sup_{z \in S} |h(z)| \le Q. \tag{2.19}$$

Consider the model of random data similar to (2.1),

$$y_j = h(z_j) + e_j, \qquad j \in \mathbb{Z}, \tag{2.20}$$

in which the values $h(z_j)$ of an unknown function $h \in H(Q)$ are observed at a given countable set $\mathcal{Z}$ of nodes $z_j \in (-1, 1)$, in the presence of *white noise*. Denote the corresponding sequences $\mathbf{y} = \{y_j\}$ and $\mathbf{h} = \{h(z_j)\}$.

Let $l_j(z)$, $z \in S$, be a sequence of *fundamental interpolating functions* satisfying the following assumptions:

$$l_j(z_i) = \delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \ne j, \end{cases} \tag{2.21}$$

$$l_j(z) \text{ are real for } z \in (-1, 1), \tag{2.22}$$

$$\sum_{j \in \mathbb{Z}} l_j(z) \text{ converges locally uniformly in } S. \tag{2.23}$$

A linear method of interpolation i is given by

$$(\mathsf{i}\,\mathbf{h})(z) = \sum_{j \in \mathbb{Z}} l_j(z) h(z_j), \qquad w \in S. \tag{2.24}$$

Denote $\Upsilon = \{\mathsf{i}\}$ the class of all interpolants (2.24) satisfying (2.21)–(2.23).

Again, for $z \in (-1, 1)$,

$$\mathrm{Var}(\mathsf{i}\,\mathbf{y}(z)) = \sigma^2 \sum_{j \in \mathbb{Z}} l_j^2(z) \qquad \text{and} \tag{2.25}$$

$$\mathbf{E}(\mathsf{i}\,\mathbf{y}(z) - h(z))^2 = \sigma^2 \sum_{j \in \mathbb{Z}} l_j^2(z) + (\mathsf{i}\,\mathbf{h}(z) - f(z))^2. \tag{2.26}$$

By (2.21),

$$\sup_{-1<z<1} \mathbf{E}(\mathsf{i}\,\mathbf{y}(z) - h(z))^2 \ge \sup_{-1<z<1} \mathrm{Var}\,(\mathsf{i}\,\mathbf{y}(z)) \ge \sigma^2.$$

Definition 2.3. *A method of interpolation $i \in \Upsilon$ is called* **(a)** *D-optimal if*

$$\sup_{z \in (-1,1)} \mathrm{Var}\,(i\mathbf{y}(z)) = \sigma^2, \qquad \sigma \ge 0;$$

(b) *R-optimal for a given $\sigma \ge 0$, w.r.t. the model $(\mathcal{Z}, H(Q))$ if*

$$\sup_{h \in H(Q)} \sup_{z \in (-1,1)} \mathbf{E}(i\mathbf{y}(z) - h(z))^2 = \sigma^2.$$

Definition 2.4. *An interpolant* $\mathfrak{i}^*\mathbf{h}$ *in* (2.24) *is called* **(a)** $(\mathcal{Z}, H(Q))$*-optimal at a given point* $z \in D$ *if*

$$\inf_{\mathfrak{i}\in\Upsilon} \sup_{h\in H(Q)} |\mathfrak{i}\,\mathbf{h}(z) - h(z)| = \sup_{h\in H(Q)} |\mathfrak{i}^*\,\mathbf{h}(z) - h(z)|;$$

(b) *A-optimal, w.r.t. the model* $(\mathcal{Z}, H(Q))$ *if*

$$r_0 =: \inf_{\mathfrak{i}\in\Upsilon} \sup_{h\in H(Q)} \sup_{z\in(-1,1)} |\mathfrak{i}\,\mathbf{h}(z) - h(z)| = \sup_{h\in H(Q)} \sup_{z\in(-1,1)} |\mathfrak{i}^*\,\mathbf{h}(z) - h(z)|.$$

The two problems in the unit disk S, with random and deterministic data, will be referred to, respectively, as

Problem 2. Find an R-optimal method of interpolation (2.24) for the model $(\mathcal{Z}, H(Q))$ given by (2.19)–(2.20).

Problem 2$_\circ$. Find an A-optimal interpolant with respect to the model $(\mathcal{Z}, H(Q))$.

2.3. Equivalence of the Two Interpolation Problems

As will be seen below, for some particular choices of the interpolating nodes z_j in (2.20), Problems 1 and $1_\circ$ are equivalent to the corresponding Problems 2 and $2_\circ$. Let us discuss the situation in some more detail.

Suppose $w = g(z)$ is a conformal mapping of the unit disk S onto the strip S_β such that $(-1, 1)$ is mapped onto $\mathbb{R}$ and $g(z_j) = j$, $j \in \mathbb{Z}$. Such a mapping establishes a bijection between the Hardy classes $H(Q)$ and $F_\beta(Q)$. One can view the two interpolation models, $(\mathcal{X}, F_\beta(Q))$ and $(\mathcal{Z}, H(Q))$, as mutually transformed one into another, while the sequences $\mathbf{h} = \mathbf{f}$ and $\mathbf{y}$ remain *invariant* under such mapping.

Let $l_j(z) = l_j(g^{-1}(w)) =: L_j(w)$, so that the fundamental interpolating functions $L_j(w)$ satisfy (2.5)–(2.7). The relation $\mathfrak{i}\,\mathbf{h}(z) = \mathfrak{i}(g^{-1}(w)) = I\mathbf{f}(w)$ is a bijection between the two classes of interpolants, Υ for the disk S, and $\mathcal{I}$ for the strip S_β. Essentially, the two models, (2.1) and (2.20), are the same. The only difference between them is in the manner of labeling the independent variable: by $z \in S$, or by $w \in S_\beta$.

A moment's thought leads to the conclusion that the notions of D-optimality, R-optimality, optimality at a given point $z \in D$, and A-optimality of an interpolant $\mathfrak{i}$ in (2.24) are equivalent, respectively, to the D-optimality, R-optimality, optimality at a given point $w = g(z) \in S_\beta$, and A-optimality of the corresponding interpolant I in (2.8)–(2.9). Thus, under the above assumptions, Problems 1 and $1_\circ$ are equivalent, respectively, to Problems 2 and $2_\circ$. Moreover, finding an optimal interpolant for the disk automatically leads to a corresponding optimal interpolant for the strip, and vice versa.

2.4. Optimal Recovery: Highlights

Methods of *Optimal Recovery* have been successfully applied to various nonparametric estimation problems. The goal of this section is to review some of its results which will be useful in the present paper. Generally, Optimal Recovery is dealing with approximation of functionals of indirectly observed objects, using only a limited or imprecise information. This includes, in particular, recovery of an unknown function $f(x)$ based on a countable set of point evaluations $f(x_j)$, $j \in \mathbb{Z}$. Below, we state some basic results of Optimal Recovery applicable to the Problems $1_\circ$ and $2_\circ$ discussed in Sections 2.1–2.2.

Let $\mathcal{F} = \{f\}$ and $\mathbf{F} = \{\mathbf{f}\}$ be two linear spaces over real numbers, $\gamma\colon \mathcal{F} \to \mathbb{C}$ a linear functional of interest (target value), and $\mathbf{G}\colon \mathcal{F} \to \mathbf{F}$ a linear mapping called *information operator*. In the deterministic setting of Optimal Recovery, the value $\mathbf{f} = \mathbf{G}f \in \mathbf{F}$, containing all the available information about an element $f \in \mathcal{F}$, represents non-random *data*.

A *method of recovery* of the given linear functional $\gamma(f)$ is *any* map $\varphi\colon \mathbf{F} \to \mathbb{C}$, while a *linear method of recovery* is a linear functional $I\colon \mathbf{F} \to \mathbb{C}$. A method of recovery φ^* is *optimal*, on a subset $F \subset \mathcal{F}$, if it achieves

$$\inf_{\varphi} \sup_{f\in F} |\varphi(\mathbf{f}) - \gamma(f)|.$$

An element $f^*_\gamma \in F$ is called *extremal*, w.r.t. the pair (F, γ), if $\mathbf{G} f^*_\gamma = 0$ and

$$\sup_{f \in F:\, \mathbf{G} f = 0} |\gamma(f)| = |\gamma(f^*_\gamma)|.$$

A set F is called *convex balanced* if for any $n \geq 1$,

$$\Big\{ f\colon f = \sum_{i=1}^{n} \lambda_j f_j,\ f_j \in F,\ \sum_{i=1}^{n} |\lambda_j| \leq 1 \Big\} \subseteq F.$$

Obviously, such a set is convex and symmetric.

Proposition 2.1 (cf. [14], Section 1.3, Theorems 1.6–1.7). *Let $F \subset \mathcal{F}$ be a non-empty convex balanced set. Then there exists an optimal linear method of recovery I and an extremal element f^*_γ such that*

$$\inf_{\varphi} \sup_{f \in F} |\varphi(\mathbf{f}) - \gamma(f)| = \sup_{f \in F} |I(\mathbf{f}) - \gamma(f)| = |\gamma(f^*_\gamma)|. \tag{2.27}$$

The following corollary will be useful below. Let $\Gamma_0 = \{\gamma\}$ be a set of linear functionals over $\mathcal{F}$. Consider the family $\Phi = \{\varphi_\gamma\}$ of all recovery methods and the family $\mathcal{I} = \{I_\gamma\}$ of all linear recovery methods, indexed by $\gamma \in \Gamma_0$.

Corollary 2.1. *Let F satisfy the assumptions of Proposition* 2.1. *Then*

$$\inf_{\Phi} \sup_{f \in F} \sup_{\gamma \in \Gamma_0} |\varphi_\gamma(\mathbf{f}) - \gamma(f)| = \sup_{\gamma \in \Gamma_0} |\gamma(f^*_\gamma)|\,. \tag{2.28}$$

Indeed, by using (2.27),

$$\begin{aligned} \inf_{\Phi} \sup_{f \in F} \sup_{\gamma \in \Gamma_0} |\varphi_\gamma(\mathbf{f}) - \gamma(f)| &= \inf_{\Phi} \sup_{\gamma \in \Gamma_0} \sup_{f \in F} |\varphi_\gamma(\mathbf{f}) - \gamma(f)| \\ &= \inf_{\mathcal{I}} \sup_{\gamma \in \Gamma_0} \sup_{f \in F} |I_\gamma(\mathbf{f}) - \gamma(f)| = \sup_{\gamma \in \Gamma_0} |\gamma(f^*_\gamma)|\,. \end{aligned}$$

Below the elements $f \in \mathcal{F}$ will be complex-valued functions defined on a given subset $\mathcal{D} \subset \mathbb{C}$, real on $\mathcal{D}_0 = \mathcal{D} \cap \mathbb{R}$, while $\mathbf{F} = \ell_\infty$ will be the linear space of all bounded real sequences $\{f_j \in \mathbb{R},\ j \in \mathbb{Z}\}$. The information operator $\mathbf{f} = \mathbf{G} f \in \mathbf{F}$ represents the values $f(x_j)$, $j \in \mathbb{Z}$, at a given set of distinct nodes $x_j \in \mathcal{D}_0$. The functional of interest γ is the point evaluation $\gamma = f(z)$ at a given point $z \in \mathcal{D}$, while $\Gamma_0 = \{f(x),\ x \in \mathcal{D}_0\}$.

The set $F \subset \mathcal{F}$ will be a given convex balanced subset. Additionally, it will be assumed that for any $i \in \mathbb{Z}$, F contains a function f such that

$$f(x_j) \begin{cases} \neq 0 & \text{if } j = i, \\ = 0 & \text{if } j \neq i. \end{cases}$$

These assumptions will automatically hold in Problems $1_\circ$ and $2_\circ$ discussed in Sections 2.1–2.2.

Noting that Proposition 2.1 applies to every functional $\gamma(f) = f(z)$, $z \in \mathcal{D}$, denote the corresponding optimal linear method of recovery $I\mathbf{f}(z) = \sum_{j \in \mathbb{Z}} L_j(z) f(x_j)$, where

$$\sum_{j \in \mathbb{Z}} |L_j(z)| < \infty, \qquad z \in \mathcal{D}.$$

Thus, for any $z \in \mathcal{D}$,

$$\inf_{\varphi} \sup_{f \in F} |\varphi\, \mathbf{f}(z) - f(z)| = \sup_{f \in F} |I\, \mathbf{f}(z) - f(z)|.$$

Obviously, if z coincides with one of the nodes x_i, the optimal linear method of recovery automatically restores $f(z)$ without error:

$$I\mathbf{f}(x_i) = \sum_{j \in \mathbb{Z}} L_j(x_i) f(x_j) = f(x_i), \qquad i \in \mathbb{Z}.$$

Thus, the optimal linear method of recovery $I\,\mathbf{f}$ is necessarily an *interpolation method*. This can be true for any $f \in \mathcal{F}$ only if

$$L_j(x_i) = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$$

Thus, $L_j(z)$ are *fundamental interpolating functions*. A linear method of recovery $I\mathbf{f}(z)$, viewed as a function of $z \in \mathcal{D}$, will be called linear *interpolant*. Denote $\mathcal{I}$ the collection of all such interpolants. In recovering the linear functional $\gamma = f(z)$, the corresponding extremal element in F will be denoted f_z^*.

Now, let $\varphi\,\mathbf{f}(z)$ be an arbitrary collection of recovery methods, for all $z \in \mathcal{D}$. The set of such methods of recovery will be again denoted Φ. In agreement with the definitions given in Sections 2.1—2.2, a linear method of interpolation I^* is called A-optimal, if

$$r_0 =: \inf_{\varphi\in\Phi} \sup_{f\in F} \sup_{x\in\mathcal{D}_0} |\varphi\mathbf{f}(x) - f(x)| = \inf_{I\in\mathcal{I}} \sup_{f\in F} \sup_{x\in\mathcal{D}_0} |I\mathbf{f}(x) - f(x)| = \sup_{f\in F} \sup_{x\in\mathcal{D}_0} |I^*\mathbf{f}(x) - f(x)|.$$

The following result follows directly from Corollary 2.1.

Proposition 2.2. *Let the interpolant $I^*\mathbf{f}(x)$ be an optimal linear method of recovery, for every $x \in \mathcal{D}_0$. Then the interpolant I^* is A-optimal and*

$$r_0 = \sup_{x\in\mathcal{D}_0} |f_x^*(x)|\,.$$

In particular, if the extremal element $f_z^* \equiv f^*$ does not depend on $z \in \mathcal{D}$,

$$r_0 = \sup_{x\in\mathcal{D}_0} |f^*(x)|.$$

3. A-OPTIMAL INTERPOLANTS

This section describes A-optimal interpolants for the corresponding Hardy classes in the unit disk S and in the strip S_β. The results are an adaptation from [14], Ch. 2.3.

3.1. Interpolation in the Disk

Let us start with the problem of countable interpolation in the disk, see (2.20). To this end, some technical tools from Complex Analysis will be reviewed first.

3.1.1. The Hardy space $\mathcal{H}$. The space $\mathcal{H}$ was defined above as the set of all bounded analytic functions in the unit disk S, see (2.17)–(2.18). It is well known that functions $h \in \mathcal{H}$ have boundary values $h(\zeta)$ almost everywhere on ∂S (see [14], Section 2.1, and further references therein); moreover, the Cauchy integral formula holds:

$$h(z) = \frac{1}{2\pi i}\int_{|\zeta|=1} \frac{h(\zeta)\,d\zeta}{\zeta - z}, \qquad z \in S. \tag{3.1}$$

From (3.1), yet another convenient representation can be derived involving a positive kernel. Indeed, the substitution $\zeta = e^{i\theta}$, $0 \le \theta \le 2\pi$, in (3.1) results in

$$h(z) = \frac{1}{2\pi}\int_0^{2\pi} \frac{h(\zeta)\zeta\,d\theta}{\zeta - z} = \frac{1}{2\pi}\int_0^{2\pi} \frac{h(\zeta)\,d\theta}{1 - \bar{\zeta}z}, \qquad z \in S.$$

Let $z \in S$ be fixed. Replacing $h(\zeta)$ by the function $g(\zeta) = \frac{1-|z|^2}{1-\zeta\bar{z}}h(\zeta)$ such that $g(z) = h(z)$, gives for any $h \in \mathcal{H}$

$$h(z) = \frac{1}{2\pi}\int_0^{2\pi} \frac{(1-|z|^2)h(\zeta)\,d\theta}{(1-\zeta\bar{z})(1-\bar{\zeta}z)} = \frac{1}{2\pi}\int_0^{2\pi} \frac{1-|z|^2}{|1-\bar{z}\zeta|^2}h(\zeta)\,d\theta = \frac{1}{2\pi}\int_0^{2\pi} \frac{1-|z|^2}{|\zeta - z|^2}h(\zeta)\,d\theta. \tag{3.2}$$

In particular,

$$\frac{1}{2\pi}\int_0^{2\pi}\frac{1-|z|^2}{|\zeta-z|^2}\,d\theta = 1, \qquad z\in S. \tag{3.3}$$

Function $P(z,\zeta)=\frac{1-|z|^2}{|\zeta-z|^2}$ is the classical *Poisson kernel*. In the polar coordinates $\zeta=e^{i\theta}$, $z=re^{i\vartheta}$, it can be represented as

$$P(z,\zeta)=\frac{1-r^2}{1-2r\cos(\theta-\vartheta)+r^2}.$$

The relation (3.3) reduces to the well-known standard integral, cf. [7], **3.613.**2,

$$\frac{1}{\pi}\int_0^{\pi}\frac{d\theta}{1-2r\cos\theta+r^2}=\frac{1}{1-r^2}.$$

In some cases, the representation (3.2) is more convenient than (3.1). In particular, the famous *maximum modulus principle* for the disk S follows immediately from (3.2)–(3.3).

3.1.2. Infinite Blaschke products. Let a sequence of interpolating nodes $\{z_j\}\subset S$ satisfy the so-called *Blaschke condition*,

$$\sum_{j=1}^{\infty}(1-|z_j|)<\infty. \tag{3.4}$$

For $a,z\in S$, define

$$B(a,z)=\operatorname{sgn}a\,\frac{z-a}{1-\bar a z},\quad \text{where}\quad \operatorname{sgn}a=\begin{cases}\frac{|a|}{a}, & a\neq 0,\\ 1, & a=0.\end{cases}$$

It can be shown that, subject to (3.4), the product

$$W(z)=\prod_{j=1}^{\infty}B(z_j,z)$$

converges for every $z\in S$, $W(z)\in\mathcal{H}$, and $W(z)=1$ almost everywhere on ∂S; cf. [16], Theorem 15.21. The function $W(z)$, having simple roots at the nodes z_j, is called *infinite Blaschke product*. By the *maximum modulus principle*,

$$\sup_S|W(z)|=1. \tag{3.5}$$

In particular, let $-1<z_j<z_{j+1}<1$, $j\in\mathbb{Z}$, satisfy the Blaschke condition (3.4). Then the corresponding infinite Blaschke product is given by

$$W(z)=\prod_{j=-\infty}^{\infty}\operatorname{sign}z_j\,\frac{z-z_j}{1-z_jz}. \tag{3.6}$$

Assume additionally that for some $\alpha_j\in(z_j,z_{j+1})$ and $c>0$,

$$|W(\alpha_j)|\geq c. \tag{3.7}$$

Under the assumptions (3.4), (3.7), the following generalization of the *residue theorem* follows from [14], Lemma 2.5: for any function $h\in\mathcal{H}$,

$$\frac{1}{2\pi i}\int_{|z|=1}\frac{h(z)}{W(z)}\,dz=\sum_{j=-\infty}^{\infty}\frac{h(z_j)}{W'(z_j)}. \tag{3.8}$$

3.1.3. Pointwise optimal interpolants. Suppose a countable set of nodes $\mathcal{Z} = \{z_j\}$ in $(-1, 1)$ satisfies the assumptions (3.4), (3.7). Let $W(z)$ be the corresponding Blaschke product (3.6). For a given $z \in S$, consider the following linear functional on $\mathcal{H}$,

$$Jh(z) = \frac{1}{2\pi i}\int_{|\zeta|=1} \frac{W(z)(1-|z|^2)h(\zeta)}{W(\zeta)(\zeta - z)(1-\bar{z}\zeta)}\,d\zeta.$$

Suppose first that $z \notin \mathcal{Z}$. Then, by the generalized *residue theorem* (3.8),

$$Jh(z) = h(z) - \sum_{j=-\infty}^{\infty} \frac{W(z)(1-|z|^2)}{W'(z_j)(z-z_j)(1-\bar{z}z_j)} h(z_j). \tag{3.9}$$

By continuity, the equation holds for all $z \in S$. This representation (3.9) points to the following interpolant:

$$\mathsf{i}^*\mathbf{h}(z) = \sum_{j=-\infty}^{\infty} l_j(z)h(z_j) =: \sum_{j=-\infty}^{\infty} \frac{W(z)(1-|z|^2)}{W'(z_j)(z-z_j)(1-\bar{z}z_j)} h(z_j). \tag{3.10}$$

Note that

$$l_j(z_i) = \delta_{ij}, \qquad i, j \in \mathbb{Z},$$

thus implying

$$\mathsf{i}^*h(z_i) \equiv h(z_i), \qquad i \in \mathbb{Z}.$$

Obviously, the interpolant i^* satisfies (2.22), while its property (2.23) will become evident later in Section 3.2. For a function $h \in \mathcal{H}$, the corresponding interpolation error, at a given point $z \in S$, can be easily evaluated using (3.9), (3.3). Indeed, by (3.3) for any $h \in H(Q)$,

$$\begin{aligned}
|\mathsf{i}^*\mathbf{h}(z) - h(z)| = |Jh(z)| &= \left| \frac{1}{2\pi i}\int_{|\zeta|=1} \frac{W(z)(1-|z|^2)h(\zeta)}{W(\zeta)(\zeta-z)(1-\bar{z}\zeta)}\,d\zeta \right| \\
&= \left| \frac{1}{2\pi}\int_0^{2\pi} \frac{W(z)(1-|z|^2)h(\zeta)\zeta}{W(\zeta)(\zeta-z)(1-\bar{z}\zeta)}\,d\theta \right| = \left| \frac{1}{2\pi}\int_0^{2\pi} \frac{W(z)(1-|z|^2)h(\zeta)}{W(\zeta)(1-\bar{\zeta}z)(1-\bar{z}\zeta)}\,d\theta \right| \\
&\le \frac{Q|W(z)|}{2\pi}\int_0^{2\pi} \frac{1-|z|^2}{|1-\bar{\zeta}z|^2}\,d\theta = \frac{Q|W(z)|}{2\pi}\int_0^{2\pi} \frac{1-|z|^2}{|\zeta-z|^2}\,d\theta = Q|W(z)|.
\end{aligned}$$

Moreover, by choosing $h^*(z) = QW(z) \in H(Q)$, cf. (3.5), one finds that for any $z \in S$,

$$\sup_{h \in H(Q)} |\mathsf{i}^*\mathbf{h}(z) - h(z)| = |h^*(z)| = Q|W(z)|.$$

The optimal properties of the interpolant i^* are summarized in the following

Proposition 3.1 (cf. [14], Corollary 2.7, p. 49)). *For a given design $\mathcal{Z} = \{z_j\}$ satisfying* (3.4)*,* (3.7)*, and a given $z \in D$, the interpolant i^* is $(\mathcal{Z}, H(Q))$-optimal and $h^*(z)$ is the extremal element, i.e.,*

$$\inf_{i \in \Upsilon} \sup_{h \in H(Q)} |i\mathbf{h}(z) - h(z)| = \sup_{h \in H(Q)} |i^*\mathbf{h}(z) - h(z)| = Q|W(z)|.$$

Remark 3.1. Generally, verification of the assumptions (3.4), (3.9) can be difficult. The problem, however, becomes straightforward for special interpolating nodes discussed next.

3.1.4. A special choice of interpolating nodes. Suppose that for some $c > 0$ the interpolating nodes $z_j \in (-1, 1)$ are chosen as

$$z_j = \tanh j/c, \qquad j \in \mathbb{Z}. \tag{3.11}$$

The following results show that in this case the assumptions (3.4), (3.9) hold automatically. Moreover, the corresponding infinite Blaschke product (3.6) can be explicitly expressed in terms of the Jacobi elliptic function $\operatorname{sn}(z,k)$, or *"elliptic sinus"*, of a *modulus* $k \in (0,1)$. For a summary of elliptic functions see Section 5 below.

The Jacobi functions can be characterized by various sets of parameters: by the modulus $k \in (0,1)$ (in the normal case); by the *complementary modulus* $k' = \sqrt{1-k^2}$; by the quarter- and half-periods, $\mathbf{K} = \mathbf{K}(k)$ and $\mathbf{K}' = \mathbf{K}(k')$, respectively; by the so-called *nome*

$$q = q(k) = \exp\Big(-\frac{\pi\mathbf{K}'}{\mathbf{K}}\Big); \tag{3.12}$$

or by the *complementary nome*

$$q' = q(k') = \exp\Big(-\frac{\pi\mathbf{K}}{\mathbf{K}'}\Big).$$

The expression of the modulus k in terms of the nome q is given below in (5.11).

Given the nodes (3.11), let us first select the modulus k such that

$$\frac{\mathbf{K}'}{\mathbf{K}} = \pi c. \tag{3.13}$$

By (3.13),

$$z_j = \tanh\frac{\pi\mathbf{K}}{\mathbf{K}'}j, \qquad j \in \mathbb{Z}.$$

Note that replacing the modulus k by k' is equivalent to interchanging $\mathbf{K}$ and $\mathbf{K}'$. The corresponding *complementary nome* satisfies

$$q' = q(k') = \exp\Big(-\frac{\pi\mathbf{K}}{\mathbf{K}'}\Big) = \exp\Big(-\frac{1}{c}\Big).$$

For typographical reasons, q' will be denoted q_1 below. In terms of q_1,

$$z_j = \frac{1-q_1^{2j}}{1+q_1^{2j}}.$$

Proposition 3.2. *Let the set of interpolating nodes* $\mathcal{Z} = \{z_j\}$ *satisfy* (3.11) *for some* $c > 0$*, the modulus* k *be selected according to* (3.13)*, and* $\operatorname{sn}(z,k)$ *denote the Jacobi elliptic function of modulus* k*. Let* $W(z)$ *be the infinite Blaschke product* (3.6)*. Then:*

(a)

$$W(z) = \sqrt{k}\,\operatorname{sn}(2c\mathbf{K}\ \operatorname{arctanh} z, k), \qquad z \in S. \tag{3.14}$$

(b) *The assumptions* (3.4)*,* (3.9) *are satisfied.*

(c) *For any* $z \in S$*, the interpolant* (3.10) *is* $(\mathcal{Z}, H(Q))$*-optimal,* $h^*(z) = QW(z)$ *is the extremal element in* $H(Q)$*, and*

$$\sup_{-1<z<1}\ \sup_{h\in H(Q)} |i^*\mathbf{h}(z) - h(z)| = Q\sqrt{k}.$$

(d) *The interpolant* i^* *is A-optimal on* $(-1,1)$*, w.r.t.* $(\mathcal{Z}, H(Q))$.

For reader's convenience, a proof of this proposition, adapted from [14] with some minor changes, is given below.

Proof. By the symmetry of the nodes z_j, the Blaschke product (3.6) becomes

$$W(z) = z\prod_{j=1}^{\infty}\frac{z_j^2 - z^2}{1 - z_j^2 z^2}. \tag{3.15}$$

It will be convenient to work with the new variable

$$v = \frac{i}{\pi}\operatorname{arctanh} z, \tag{3.16}$$

so that $z = -\tanh i\pi v = -i\tan\pi v$, and

$$z^2 = \frac{\cos 2\pi v - 1}{\cos 2\pi v + 1}. \tag{3.17}$$

By substituting (3.17) in (3.15), and a bit of algebra, the Blaschke product $W(z)$ becomes

$$W(z) = -i\tan\pi v \prod_{j=1}^{\infty} \frac{1 - 2q_1^{2j}\cos 2\pi v + q_1^{4j}}{1 + 2q_1^{2j}\cos 2\pi v + q_1^{4j}}.$$

Using the representation of the elliptic sine and cosine by means of the *Jacobi theta functions* $\theta_i(v),\ i = 1, 2$, and the product expansions of the latter, see (5.13)–(5.18) below, one obtains

$$W(z) = -i\frac{\sin\pi v}{\cos\pi v}\prod_{j=1}^{\infty}\frac{1 - 2q_1^{2j}\cos 2\pi v + q_1^{4j}}{1 + 2q_1^{2j}\cos 2\pi v + q_2^{4j}} = -i\frac{\theta_1(v)}{\theta_2(v)} = -i\sqrt{k}\frac{\operatorname{sn}(2\mathbf{K}'v, k')}{\operatorname{cn}(2\mathbf{K}'v, k')}.$$

Now, by (3.16), (3.13),

$$W(z) = -i\sqrt{k}\frac{\operatorname{sn}(2ic\mathbf{K}\ \operatorname{arctanh} z, k')}{\operatorname{cn}(2ic\mathbf{K}\ \operatorname{arctanh} z, k')}.$$

Finally, using the *second principal first degree transform*, see (5.19) below, one gets

$$W(z) = \sqrt{k}\operatorname{sn}(2c\mathbf{K}\operatorname{arctanh} z, k).$$

(b) The fact that the special nodes (3.11) satisfy the Blaschke assumption (3.4) is elementary. Next, let

$$\alpha_j = \tanh\Big((2j-1)\frac{\pi\mathbf{K}}{2\mathbf{K}'}\Big), \qquad j \in \mathbb{Z}.$$

By Part **(a)** and (5.6),

$$W(\alpha_j) = (-1)^{j+1}\sqrt{k}. \tag{3.18}$$

Thus, the system of nodes $\{z_j\}_{-\infty}^{\infty}$ satisfies the assumption (3.7).

(c) By Part **(a)**, one gets

$$\begin{aligned}\sup_{-1<z<1}|\mathrm{i}^*\mathbf{h}(z) - h(z)| &= Q\sup_{-1<z<1}|W(z)| = Q\sqrt{k}\sup_{-1<z<1}|\operatorname{sn}(2c\mathbf{K}\operatorname{arctanh} z, k)| \\ &= Q\sqrt{k}\max_{x\in\mathbb{R}}|\operatorname{sn}(2c\mathbf{K}x, k)| = Q\sqrt{k}.\end{aligned} \tag{3.19}$$

(d) This is an easy consequence of Propositions 2.2 and 3.1, in view of (3.18)–(3.19). □

3.2. Interpolation in the Strip

In this section, A-optimal methods of cardinal interpolation will be described for the Hardy classes $F_\beta(Q)$. These results are based on the equivalence of the cardinal interpolation in the strip and countable interpolation in the disk, for a special choice of the parameter c determining the sequence of interpolating nodes (3.11); cf. [14], Ch. 3.

3.2.1. Mapping the strip onto the disk. The following conformal mapping

$$z = z(w) = \tanh\frac{\pi w}{4\beta} \qquad (z \in S,\ w \in S_\beta) \tag{3.20}$$

of the strip S_β onto the unit disk S is well known; see, e.g., [13], Section **I**.10.52. The mapping (3.20) can be viewed as a composition of two transformations: the first one,

$$w \longrightarrow w_1 = \exp \frac{\pi w}{2\beta},$$

mapping the strip S_β into the right half-plane $\operatorname{Re} w_1 > 0$ (the horizontal lines $\operatorname{Im} w = \pm\beta$ are mapped into the corresponding half-lines $w_1 = \pm it\,(t > 0)$) followed by the *linear rational transformation*

$$w_1 \longrightarrow z = \frac{w_1 - 1}{w_1 + 1},$$

mapping the right half-plane into the unit disk. Note that under the second transformation, the point $w_1 = 1$ is mapped into 0, while the imaginary axis $w_1 = it$ is mapped onto the unit circle, since

$$\left|\frac{it-1}{it+1}\right|^2 = \left|\frac{(1-t^2)+2it}{1+t^2}\right|^2 = 1.$$

The diffeomorphism (3.20) is, in fact, a conformal mapping, since

$$\frac{dz}{dw} = \frac{\pi}{4\beta}\Big(1 - \tanh^2 \frac{\pi w}{4\beta}\Big) \neq 0.$$

The mapping (3.20) establishes the following bijections, $S \longleftrightarrow S_\beta$ and $H(Q) \longleftrightarrow F_\beta(Q)$, so that for any $h(z) \in H(Q)$, $f(w) = h(z(\omega)) \in F_\beta(Q)$. Moreover, if the parameter c in (3.11) is chosen such that

$$c = \frac{4\beta}{\pi}, \tag{3.21}$$

the interpolating nodes (3.11) are transformed into

$$z_j = \tanh j/c \quad \longleftrightarrow \quad j,$$

so that $h(z_j) = f(j)$, $j \in \mathbb{Z}$.

Note that since by (3.13) and (3.21),

$$\mathbf{K}' = 4\beta\mathbf{K}, \tag{3.22}$$

the *nome* (3.12) satisfies

$$q = q(k) = \exp -4\pi\beta, \tag{3.23}$$

while the *modulus* k is given by the equation (5.11) below.

3.2.2. The A-optimal interpolant I_β in S_β. Under the mapping (3.20), the optimal interpolant (3.10) translates into

$$\mathrm{i}^*\mathbf{h}(z) = \sum_{j=-\infty}^{\infty} l_j(z)h(z_j) = \sum_{j=-\infty}^{\infty} L_j(w)f(j) = I_\beta\mathbf{f}(w),$$

where

$$L_j(w) =: l_j(z(w)), \qquad w \in S_\beta.$$

Below, the fundamental interpolating functions $L_j(w)$ are described explicitly through a combination of *elliptic* and *hyperbolic* sinuses.

Proposition 3.3. *Let the modulus k satisfy* (3.23). *Under the mapping* (3.20), $L_j(w) = L(w - j)$ *where the kernel $L(\omega) = L(w,\beta)$ is given by*

$$L(w) = \frac{\pi}{4\beta\mathbf{K}} \frac{\operatorname{sn}(2\mathbf{K}w, k)}{\sinh \frac{\pi w}{2\beta}}. \tag{3.24}$$

Proof. By (3.14), (3.22),

$$W(z) = \sqrt{k}\,\mathrm{sn}(2c\mathbf{K}\ \operatorname{arctanh} z, k) = \sqrt{k}\,\mathrm{sn}(2\mathbf{K}w, k).$$

Using the *addition formulas* for the hyperbolic functions, one gets for $z_j \in (-1,1)$ and $z \in S$,

$$l_j(z) = \frac{W(z)(1-z^2)}{W'(z_j)(z-z_j)(1-z_jz)} = \frac{\sqrt{k}\,\mathrm{sn}(2\mathbf{K}w,k)}{W'(z_j)\big(\tanh\frac{\pi w}{4\beta} - \tanh\frac{\pi j}{4\beta}\big)} \cdot \frac{1-\tanh^2\frac{\pi w}{4\beta}}{1-\tanh\frac{\pi w}{4\beta}\tanh\frac{\pi j}{4\beta}}$$
$$= \frac{\sqrt{k}\,\mathrm{sn}(2\mathbf{K}w,k)\cosh^2\frac{\pi j}{4\beta}}{W'(z_j)\big(\sinh\frac{\pi w}{4\beta}\cosh\frac{\pi j}{4\beta} - \cosh\frac{\pi w}{4\beta}\sinh\frac{\pi j}{4\beta}\big)\big(\cosh\frac{\pi w}{4\beta}\cosh\frac{\pi j}{4\beta} - \sinh\frac{\pi w}{4\beta}\sinh\frac{\pi j}{4\beta}\big)}$$
$$= \frac{\sqrt{k}\,\mathrm{sn}(2\mathbf{K}w,k)\cosh^2\frac{\pi j}{4\beta}}{W'(z_j)\sinh\big(\frac{\pi}{4\beta}(w-j)\big)\cosh\big(\frac{\pi}{4\beta}(w-j)\big)} = \frac{2\sqrt{k}\,\mathrm{sn}(2\mathbf{K}w,k)\cosh^2\frac{\pi j}{4\beta}}{W'(z_j)\sinh\frac{\pi}{2\beta}(w-j)} = L_j(w).$$

Note that $L_j(i) = 0$ for $i \neq j$, and by continuity, cf. (5.6),

$$L_j(j) = \frac{8\beta\sqrt{k}\mathbf{K}(-1)^j\cosh^2\frac{\pi j}{4\beta}}{\pi W'(z_j)}.$$

On the other hand,

$$L_j(j) = l_j(z_j) = 1.$$

Thus,

$$\frac{2\sqrt{k}\cosh^2\frac{\pi j}{4\beta}}{W'(z_j)} = \frac{(-1)^j\pi}{4\beta\mathbf{K}},$$

and by (5.7) the result follows. □

By the equivalence of the two interpolation problems, cf. Section 2.2, one gets from Proposition 3.2 the following

Corollary 3.1. **(a)** *For every $Q > 0$ and $w \in S_\beta$, the interpolant*

$$I_\beta\,\mathbf{f}(w) = \frac{\pi\,\mathrm{sn}(2\mathbf{K}w,k)}{4\beta\mathbf{K}}\sum_{j=-\infty}^{\infty}\frac{(-1)^j f(j)}{\sinh\big(\frac{\pi}{2\beta}(w-j)\big)} = \frac{\pi}{4\beta\mathbf{K}}\sum_{j=-\infty}^{\infty}\frac{\mathrm{sn}\big(2\mathbf{K}(w-j),k\big)}{\sinh\big(\frac{\pi}{2\beta}(w-j)\big)}f(j) \tag{3.25}$$

is $(\mathcal{X}, F_\beta(Q))$-optimal.

(b) *For every $w \in S_\beta$, $f^*(w) = Q\sqrt{k}\,\mathrm{sn}(2\mathbf{K}w,k)$ is an extremal element in $F_\beta(Q)$, and*

$$|I_\beta\,\mathbf{f}(w) - f(w)| \le |f^*(w)|. \tag{3.26}$$

(c) *The interpolant $I_\beta\,\mathbf{f}$ is A-optimal on $\mathbb{R}$ w.r.t. $(\mathcal{X}, F_\beta(Q))$, and*

$$r_0 = \inf_{I\in\mathcal{I}}\sup_{f\in F_\beta(Q)}\sup_{x\in\mathbb{R}}|I\mathbf{f}(w) - f(w)| = \sup_{f\in F_\beta(Q)}\sup_{x\in\mathbb{R}}|I_\beta\,\mathbf{f}(w) - f(w)| = Q\sqrt{k}. \tag{3.27}$$

3.2.3. The limiting case $\beta\to\infty$: sinc interpolation. This section deals with the limiting properties of the optimal interpolant I_β for large β. It is not difficult to see that in this case the kernel (3.24) converges to the ever-present *sinc function*. Indeed, by (5.2), (5.4),

$$\lim_{\beta\to\infty} L(w;\beta) = \frac{\sin\pi w}{\pi w}.$$

Thus, the interpolant I_β is expected to approach the *cardinal series*

$$I_\infty\mathbf{f}(x) = \sum_{j=-\infty}^{\infty}\frac{\sin\pi(x-j)}{\pi(x-j)}f(j). \tag{3.28}$$

In fact, for any $f \in \mathbf{L}_2(\mathbb{R})$ uniformly in $x \in \mathbb{R}$,

$$\lim_{\beta\to\infty} I_\beta \mathbf{f}(x) = I_\infty \mathbf{f}(x). \tag{3.29}$$

Named variously after E. Borel, H. Nyquist, V.A. Kotel'nikov, C.E. Shannon, and others, the interpolant (3.28) became quite popular in *signal processing*. Chapter 6 of [15] provides a quick introduction to the topic; for a broader context see [9], Section 20.2.

The optimal variance and mean squared error properties of the interpolants I_β and I_∞ will be discussed in Section 4. In the following remarks, we focus on their bias related properties.

Remark 3.2. According to (3.26), the stochastic interpolant $I_\beta \mathbf{y}(x)$ is *asymptotically unbiased* when $\beta \to \infty$. Indeed, by (3.23) and (5.12), uniformly over $F_\beta(Q)$,

$$\sup_x |b(x)| = \sup_x |\mathbf{E} b I_\beta \mathbf{y}(x) - f(x)| \le Q\sqrt{k} = 2Qe^{-\pi\beta}(1 + O(e^{-4\pi\beta})), \qquad \beta \to \infty.$$

This result is not particularly interesting in the case of classes $F_\beta(Q)$ with a fixed Q, since in the limit $\beta \to \infty$ they would consist of entire bounded functions, i.e., constants. However, one can sharpen the result as follows. Consider Hardy classes $F_\beta(Q)$, with $Q = Q(\beta)$. If for some $A > 0$ and $\delta > 0$,

$$Q(\beta) \le Ae^{(\pi-\delta)\beta}, \tag{3.30}$$

then the interpolant $I_\beta \mathbf{y}(x)$ is asymptotically unbiased in $F_\beta(Q)$, for $\beta \to \infty$.

Remark 3.3. Denote by $\mathcal{E}_\alpha$ the linear space of *finite energy bandlimited* functions $f \in \mathbf{L}_2(\mathbb{R})$, real on the real line, whose Fourier transforms vanish outside $[-\alpha, \alpha]$. Endow $\mathcal{E}_\alpha$ with the $\mathbf{L}_2$ norm. Due to the celebrated *Paley–Wiener theorem*, $\mathcal{E}_\alpha$ coincides with the class of *entire functions of exponential type* α, i.e., for any $f \in \mathcal{E}_\alpha$ there are $A, \delta > 0$ such that the continuation of f into $\mathbb{C}$ satisfies

$$|f(z)| \le Ae^{(\alpha+\delta)|z|}, \qquad z \in \mathbb{C}. \tag{3.31}$$

By the classical *sampling theorem*, see e.g., [15], Ch. 6, or [9], Section 20.2, the interpolant $I_\infty \mathbf{y}(x)$ is unbiased in $\mathcal{E}_\pi$,

$$I_\infty \mathbf{f}(x) \equiv f(x), \qquad f \in \mathcal{E}_\pi. \tag{3.32}$$

Remark 3.4. There is a further connection between the sampling theorem and the asymptotic unbiasedness of $I_\beta\, \mathbf{y}(x)$. Denote $\mathcal{E}_\pi(M)$ the class of all functions $f \in \mathcal{E}_\pi$ such that

$$\int_{-\infty}^{\infty} |f(x)|\, dx \le M.$$

One can then deduce the sampling theorem for $\mathcal{E}_\pi(M)$ from the asymptotic unbiasedness of the interpolant $I_\beta \mathbf{y}(x)$ discussed in Remark 3.2. Indeed, by continuity of $I_\infty \mathbf{f}(x)$ in $\mathcal{E}_\alpha$, it is sufficient to demonstrate (3.32) for the classes $\mathcal{E}_{\pi-\varepsilon}$, $\varepsilon > 0$. Now, by [4], Section 6.7, for $\beta \ge 1$, a function $f \in \mathcal{E}_{\pi-\varepsilon}(M)$ satisfies

$$|f(z)| \le \frac{M}{2\pi} e^{(\pi-\varepsilon)\beta}.$$

Thus, by Remark 3.2, the asymptotic unbiasedness of $I_\beta\, \mathbf{y}(x)$ in $\mathcal{E}_\pi(M)$ – and with it the unbiasedness of $I_\infty \mathbf{y}(x)$ – follow.

4. R-OPTIMALITY OF THE CARDINAL INTERPOLANT I_β

We are ready to discuss D- and R-optimality of the interpolant I_β in the strip S_β. From this, optimal properties of the corresponding interpolants in the unit ball S can be immediately derived according to Section 2.3. By (3.25), the interpolant $I_\beta \mathbf{y}$ is defined for $w \in S_\beta$ as

$$I_\beta \mathbf{y}(w) = \frac{\pi \operatorname{sn}(2\mathbf{K}w, k)}{4\beta\mathbf{K}} \sum_{j=-\infty}^{\infty} \frac{(-1)^j y_j}{\sinh\left(\frac{\pi}{2\beta}(w-j)\right)} = \frac{\pi}{4\beta\mathbf{K}} \sum_{j=-\infty}^{\infty} \frac{\operatorname{sn}(2\mathbf{K}(w-j), k)}{\sinh\left(\frac{\pi}{2\beta}(w-j)\right)} y_j, \tag{4.1}$$

with the modulus k satisfying (3.22). Note that by (5.6), (5.7), for $i, j \in \mathbb{Z}$,

$$\frac{\pi}{4\beta\mathbf{K}} \frac{\operatorname{sn}\left(2\mathbf{K}(x-j), k\right)}{\sinh\left(\frac{\pi}{2\beta}(x-j)\right)}\Bigg|_{x=i} = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases} \tag{4.2}$$

4.1. Variance Function of I_β

In studying D- and R-optimality of the interpolant I_β, the focus turns to

$$\operatorname{Var} I_\beta \mathbf{y}(x) =: \sigma^2 s(x), \qquad x \in \mathbb{R},$$

where the *variance function* $s(x)$ is given by

$$s(x) = \left(\frac{\pi}{4\beta\mathbf{K}}\right)^2 \operatorname{sn}^2(2\mathbf{K}x, k) \sum_{j=-\infty}^{\infty} \frac{1}{\sinh^2\left(\frac{\pi}{2\beta}(x-j)\right)}. \tag{4.3}$$

By (4.3), function $s(w)$ is well defined in the whole complex plane $\mathbb{C}$ and represents there a doubly periodic meromorphic function, i.e., an elliptic function. By (4.2), $s(j) = 1$, $j \in \mathbb{Z}$.

In fact, the variance function $s(w)$ can be described in terms of $\operatorname{sn}(2\mathbf{K}w, k)$ and $\operatorname{cn}(2\mathbf{K}w, k)$ only. To this end, consider the function

$$u(w) = 1 + [a_1 + a_2 \operatorname{cn}^2(2\mathbf{K}w, k)] \operatorname{sn}^2(2\mathbf{K}w, k), \qquad w \in \mathbb{C}, \tag{4.4}$$

whose coefficients a_1, a_2 are explicitly described in the following

Theorem 4.1. *Let the coefficients a_1, a_2 in* (4.4) *be defined as*

$$a_1 = \frac{2k^2 - 1}{3} + \left(\frac{\pi}{4\beta\mathbf{K}}\right)^2 \left(-\frac{1}{3} + \sum_{j=1}^{\infty} \frac{2}{\sinh^2 \frac{\pi j}{2\beta}}\right), \qquad a_2 = -k^2. \tag{4.5}$$

Then

$$s(w) = u(w), \qquad w \in \mathbb{C}.$$

Proof. First, using (5.5), let us represent (4.4) as

$$u(w) = 1 + [b_1 + b_2 \operatorname{sn}^2(2\mathbf{K}w, k)] \operatorname{sn}^2(2\mathbf{K}z, k), \tag{4.6}$$

with coefficients b_1, b_2 satisfying $a_1 = b_1 + b_2$, $a_2 = -b_2$.

Recall that $\operatorname{sn}(\cdot, k)$ is a second order $(4\mathbf{K}, 2i\mathbf{K}')$ doubly periodic elliptic function such that $\operatorname{sn}(x + 2\mathbf{K}) = -\operatorname{sn} x$. It has simple zeros at $2j\mathbf{K}$ and simple poles at $2j\mathbf{K} + i\mathbf{K}'$, $j \in \mathbb{Z}$. Thus, $\operatorname{sn}^2(\cdot, k)$ is a second order elliptic $(2\mathbf{K}, 2i\mathbf{K}')$-periodic function, with double poles at $2j\mathbf{K} + i\mathbf{K}'$. By (3.22), $\operatorname{sn}^2(2\mathbf{K}x, k)$ is $(1, i\mathbf{K}'/\mathbf{K}) = (1, 4i\beta)$ doubly periodic, with real zeros at $j \in \mathbb{Z}$ and a double pole at $2i\beta$, the only pole within its *period parallelogram* $\Pi = \{z\colon 0 \le \operatorname{Re} w < 1,\ 0 \le \operatorname{Im} w < 4i\beta\}$.

Similarly, since

$$\sinh(w + \pi i) = -\sinh w, \tag{4.7}$$

$\sinh^2 w$ is πi-periodic. Thus, $\sinh^2 \frac{\pi w}{2\beta}$ has period $2i\beta$ and a double root at $2i\beta$.

It follows that both functions $s(w)$ and $u(w)$, respectively, in (4.3) and (4.6), are $(1, 4i\beta)$ doubly periodic meromorphic functions, with the same period parallelogram Π, and that their only singularities within Π are fourth order poles at $2i\beta$.

Now, to prove Theorem 4.1, it is sufficient to show that, with the coefficients a_1, a_2 defined by (4.5), the principal parts of the Laurent expansions of both functions, in a vicinity of the pole $w = 2i\beta$, coincide. Indeed, were it the case, their difference would be an elliptic function, with no poles in the whole plane $\mathbb{C}$. By the Liouville theorem, $s(w) - u(w)$ would be constant in $\mathbb{C}$. Since both functions are equal to 1 at $w = 0$, they would then coincide identically. Thus, let us take a closer look at the local behavior of these functions in a vicinity of $w = 2i\beta$.

1°. **Behavior of $\operatorname{sn}^2(2\mathbf{K}w, k)$ in a vicinity of its double pole $w = 2i\beta$.** By (3.22) and (5.8)–(5.9), for $\delta \to 0$, $\delta \neq 0$,

$$\operatorname{sn}(2\mathbf{K}(2i\beta + \delta)) = \operatorname{sn}(4i\mathbf{K}\beta + 2\mathbf{K}\delta) = \operatorname{sn}(i\mathbf{K}' + 2\mathbf{K}\delta) = \frac{1}{k \operatorname{sn} 2\mathbf{K}\delta} = \frac{1}{2\mathbf{K}k\delta\big(1 - \frac{1+k^2}{6}(2\mathbf{K}\delta)^2 + O(\delta^4)\big)}.$$

Thus,

$$\operatorname{sn}^2(2\mathbf{K}(2i\beta + \delta)) = \frac{1}{(2\mathbf{K})^2 k^2 \delta^2}\Big(1 + \frac{1+k^2}{3}(2\mathbf{K}\delta)^2 + O(\delta^4)\Big). \tag{4.8}$$

2°. **Behavior of $\frac{1}{\sinh^2(\frac{\pi w}{2\beta})}$ in a vicinity of the double pole $w = 2i\beta$.** By (4.7),

$$\sinh(i\pi + w) = -\sinh w = -w\Big(1 + \frac{w^2}{6} + O(w^4)\Big), \qquad w \to 0.$$

Thus, for $\delta \to 0$, $\delta \neq 0$,

$$\frac{1}{\sinh^2\big(\frac{\pi}{2\beta}(2i\beta + \delta)\big)} = \frac{1}{\sinh^2\big(\pi i + \frac{\pi\delta}{2\beta}\big)} = \frac{1}{\sinh^2 \frac{\pi\delta}{2\beta}} = \frac{1}{\big(\frac{\pi\delta}{2\beta}\big)^2\big(1 + \frac{1}{3}\big(\frac{\pi\delta}{2\beta}\big)^2 + O(\delta^4)\big)} = \frac{4\beta^2}{\pi^2\delta^2}\Big(1 - \frac{1}{3}\cdot\frac{\pi^2\delta^2}{4\beta^2} + O(\delta^4)\Big). \tag{4.9}$$

Also by (4.7),

$$\sum_{j\neq 0} \frac{1}{\sinh^2(\frac{\pi}{2\beta}(2i\beta + \delta - j)} = \sum_{j\neq 0} \frac{1}{\sinh^2\big(i\pi - \frac{\pi(j+\delta)}{2\beta}\big)} = \sum_{j\neq 0} \frac{1}{\sinh^2 \frac{\pi(j+\delta)}{2\beta}}$$
$$= \sum_{j\neq 0} \frac{1}{\sinh^2 \frac{\pi j}{2\beta}} + \frac{\pi\delta}{\beta}\sum_{j\neq 0} \frac{\cosh \frac{\pi j}{2\beta}}{\sinh^3 \frac{\pi j}{2\beta}} + O(\delta^2)\sum_{j\neq 0} \frac{1}{\sinh^2 \frac{\pi j}{2\beta}} = \sum_{j=1}^{\infty} \frac{2}{\sinh^2 \frac{\pi j}{2\beta}} + O(\delta^2). \tag{4.10}$$

3°. **Laurent expansion of the variance function $s(w)$ in a vicinity of $w = 2i\beta$.** By (4.8)–(4.10), for $\delta \to 0,\ \delta \neq 0$,

$$\Big(\frac{\pi}{4\beta\mathbf{K}}\Big)^2 \frac{\operatorname{sn}^2\big(2\mathbf{K}(2i\beta + \delta), k\big)}{\sinh^2\big(\frac{\pi}{2\beta}(2i\beta + \delta)\big)} = \frac{1}{(2\mathbf{K})^4 k^2 \delta^4}\Big(1 + \frac{\delta^2}{3}\Big((1+k^2)(2\mathbf{K})^2 - \frac{\pi^2}{4\beta^2}\Big) + O(\delta^4)\Big)$$
$$= \frac{1}{(2\mathbf{K})^4 k^2 \delta^4} + \frac{1}{3(2\mathbf{K})^4 k^2 \delta^2}\Big((1+k^2)(2\mathbf{K})^2 - \frac{\pi^2}{4\beta^2}\Big) + O(1),$$

and

$$\Big(\frac{\pi}{4\beta\mathbf{K}}\Big)^2 \operatorname{sn}^2\big(2\mathbf{K}(2i\beta + \delta), k\big) \sum_{j\neq 0} \frac{1}{\sinh^2(\frac{\pi}{2\beta}(2i\beta + \delta - j)} = \frac{\pi^2}{\beta^2 (2\mathbf{K})^4 k^2 \delta^2} \sum_{j=1}^{\infty} \frac{2}{\sinh^2 \frac{\pi j}{2\beta}} + O(1).$$

Thus, by (4.3),

$$s(2i\beta+\delta)=\frac{1}{(2\mathbf{K})^4k^2\delta^4}+\frac{1}{(2\mathbf{K})^4k^2\delta^2}\Big(\frac{1+k^2}{3}(2\mathbf{K})^2+\frac{\pi^2}{4\beta^2}\Big(-\frac{1}{3}+\sum_{j=1}^{\infty}\frac{1}{\sinh^2\frac{\pi j}{2\beta}}\Big)\Big)+O(1). \quad (4.11)$$

4°. **Laurent expansion of $u(w)$ in a vicinity of $w=2i\beta$.** Using (4.8) and arranging the resulting terms in (4.6) in powers of δ^{-1}, one easily gets

$$u(2i\beta+\delta)=\frac{b_2}{(2\mathbf{K}^4)k^4\delta^4}+\frac{1}{(2\mathbf{K})^2k^2\delta^2}\Big(b_1+2b_2\frac{1+k^2}{2k^2}\Big)+O(1). \quad (4.12)$$

5°. **Comparing Laurent expansions of $s(w)$ and $u(w)$ in a vicinity of $w=2i\beta$.** One observes that the principal parts of the Laurent expansions (4.11), (4.12) coincide if the coefficients b_1, b_2 satisfy the relations

$$b_2=k^2, \qquad b_1=-\frac{1+k^2}{3}+\Big(\frac{\pi}{2\beta\mathbf{K}}\Big)^2\Big(-\frac{1}{3}+\sum_{j=1}^{\infty}\frac{2}{\sinh^2\frac{\pi j}{2\beta}}\Big),$$

which are equivalent to (4.5). From this, Theorem 4.1 follows directly, in view of the *Liouville theorem*. □

4.2. D-Optimality

The main result of this section is the following

Proposition 4.1. *For all $\beta>0$ the interpolants $I_\beta\mathbf{y}$ are D-optimal.*

Proof. According to Definition 2.1, D-optimality of $I_\beta\mathbf{y}$ is equivalent to the property (2.16) of the variance function $s(x)$. By Theorem 4.1, (2.16) holds if $a_1\le 0$ in (4.5). This and more will be proved next. □

Lemma 4.1. **(a)** *For all $\beta>0$, the coefficient a_1 in* (4.5) *satisfies $a_1<0$.*
(b) *For $\beta\to\infty$,*

$$a_1\le-\frac{1}{2\sqrt{6}\beta}+O\Big(\frac{1}{\beta^2}\Big).$$

Proof. **(a)** Let us represent the coefficient a_1 in (4.5) as $a_1=(A-1/3)+B$, where

$$A=\Big(\frac{\pi}{4\beta\mathbf{K}}\Big)^2\sum_{j=1}^{\infty}\frac{2}{\sinh^2\frac{\pi j}{2\beta}} \qquad \text{and} \qquad B=\frac{1}{3}\Big(2k^2-\Big(\frac{\pi}{4\beta\mathbf{K}}\Big)^2\Big).$$

To prove Part **(a)**, it is sufficient to show that $A-1/3<0$ and $B<0$. Using the inequality $\sinh x> x, x>0$, the classical series [7] **0.233**(3)

$$\sum_{j=1}^{\infty}\frac{1}{j^2}=\frac{\pi^2}{6}, \quad (4.13)$$

and (5.1), one gets

$$A-\frac{1}{3}=\Big(\frac{\pi}{4\beta\mathbf{K}}\Big)^2\sum_{j=1}^{\infty}\frac{2}{\sinh^2\frac{\pi j}{2\beta}}-\frac{1}{3}\le\Big(\frac{1}{2\mathbf{K}}\Big)^2\sum_{j=1}^{\infty}\frac{2}{j^2}-\frac{1}{3}=\frac{1}{3}\Big(\Big(\frac{\pi}{2\mathbf{K}}\Big)^2-1\Big)<0.$$

Next, by (3.22)

$$B=\frac{1}{3}\Big(2k^2-\Big(\frac{\pi}{\mathbf{K}'}\Big)^2\Big)=\frac{1}{3}\Big(2k^2-\Big(\frac{\pi}{\mathbf{K}(k')}\Big)^2\Big).$$

Note that (5.3) implies the following relations

$$\mathbf{K}(k') \le \frac{\pi}{2} + \log\frac{1}{k} < \frac{\pi}{\sqrt{2}} + \Big(\frac{1}{k} - 1\Big) = \Big(\frac{\pi}{\sqrt{2}} - 1\Big) - \frac{1}{k}\Big(\frac{\pi}{\sqrt{2}} - 1\Big) + \frac{\pi}{\sqrt{2}k} < \frac{\pi}{\sqrt{2}k},$$

thus proving that $B < 0$.

(b) To shorten the notation, let $u = \pi/(2\beta)$, $v = u^2/6$. Using a stronger inequality $\sinh x > x + \frac{x^3}{6}$, $x > 0$, and (5.1), one gets for $\beta \to \infty$,

$$\begin{aligned} A &= \frac{u^2}{2\mathbf{K}^2}\sum_{j=1}^{\infty}\frac{1}{\sinh^2 uj} < \frac{u^2}{2\mathbf{K}^2}\sum_{j=1}^{\infty}\frac{1}{\big(uj + \frac{(uj)^3}{6}\big)^2} < \frac{2}{\pi^2}\sum_{j=1}^{\infty}\frac{1}{j^2\big(1+vj^2\big)^2} \\ &< \frac{2}{\pi^2}\sum_{j=1}^{\infty}\Big(\frac{1}{j^2} - \frac{v}{1+vj^2}\Big) < \frac{2}{\pi^2}\Big(\frac{\pi^2}{6} - \int_1^{\infty}\frac{v\,dx}{1+vx^2}\Big) \\ &= \frac{2}{\pi^2}\Big(\frac{\pi^2}{6} - \sqrt{v}\arctan\sqrt{v}x\Big|_{x=1}^{\infty}\Big) = \frac{2}{\pi^2}\Big(\frac{\pi^2}{6} - \frac{\pi}{2}\sqrt{v} + O(v)\Big) \\ &= \frac{2}{\pi^2}\Big(\frac{\pi^2}{6} - \frac{\pi^2}{4\sqrt{6}\beta} + O\Big(\frac{1}{\beta^2}\Big)\Big) = \frac{1}{3} - \frac{1}{2\sqrt{6}\beta} + O\Big(\frac{1}{\beta^2}\Big). \end{aligned}$$

□

4.3. R-Optimality

Recall that by definition $\mathcal{I}$ is the class of all cardinal interpolants (2.8) satisfying (2.5)–(2.7), and I_β is the A-optimal interpolant (3.25), with approximation error satisfying (3.26) and maximal error $r_0 = Q\sqrt{k}$, where k satisfies (3.22). Denote

$$g = -\frac{k}{a_1} \qquad \text{and} \qquad \sigma_0^2 = Q^2 g,$$

where $a_1 < 0$ was defined by (4.5). Below $x_+ = \max(0, x)$. We start with the following

Theorem 4.2. *For any $\sigma^2 \ge 0$,*

$$\begin{aligned} \max(\sigma^2, r_0^2) &\le \min_{I\in\mathcal{I}}\ \sup_{f\in F_\beta(Q)}\ \sup_{x\in\mathbb{R}}\mathbf{E}(I\mathbf{y}(x) - f(x))^2 \\ &\le \sup_{f\in F_\beta(Q)}\ \sup_{x\in\mathbb{R}}\mathbf{E}(I_\beta\mathbf{y}(x) - f(x))^2 \le \sigma^2 + \frac{r_0^2}{\sigma_0^2}(\sigma_0^2 - \sigma^2)_+\,. \end{aligned} \tag{4.14}$$

Proof. Combining Corollary 3.1 **(b)** with Theorem 4.1, one gets for any $f \in F_\beta(Q)$ and $x \in \mathbb{R}$,

$$\begin{aligned} \mathbf{E}(I_\beta\mathbf{y}(x) - f(x))^2 &\le \sigma^2(1 + [a_1 - k^2\,\mathrm{cn}^2(2\mathbf{K}x, k)]\,\mathrm{sn}^2(2\mathbf{K}x)) + Q^2k\,\mathrm{sn}^2(2\mathbf{K}x, k) \\ &\le \sigma^2 + (\sigma^2 a_1 + Q^2 k)\,\mathrm{sn}^2(2\mathbf{K}x, k). \end{aligned}$$

Thus, for $\sigma^2 \ge \sigma_0^2$,

$$\sup_{f\in F_\beta(Q)}\ \sup_{x\in\mathbb{R}}\mathbf{E}\big(I_\beta\mathbf{y}(x) - f(x)\big)^2 = \sigma^2. \tag{4.15}$$

For $0 \le \sigma^2 \le \sigma_0^2$, one gets similarly

$$\begin{aligned} \sup_{f\in F_\beta(Q)}\ \sup_{x\in\mathbb{R}}\mathbf{E}(I_\beta\,\mathbf{y}(x) - f(x))^2 &\le \sigma^2 + \sup_{x\in\mathbb{R}}(\sigma^2 a_1 + Q^2 k)\,\mathrm{sn}^2(2\mathbf{K}x, k) \\ &= \sigma^2 + (\sigma^2 a_1 + Q^2 k) = \sigma^2 + r_0^2\Big(1 - \frac{\sigma^2}{\sigma_0^2}\Big) = \sigma^2 + \frac{r_0^2}{\sigma_0^2}(\sigma_0^2 - \sigma^2), \end{aligned}$$

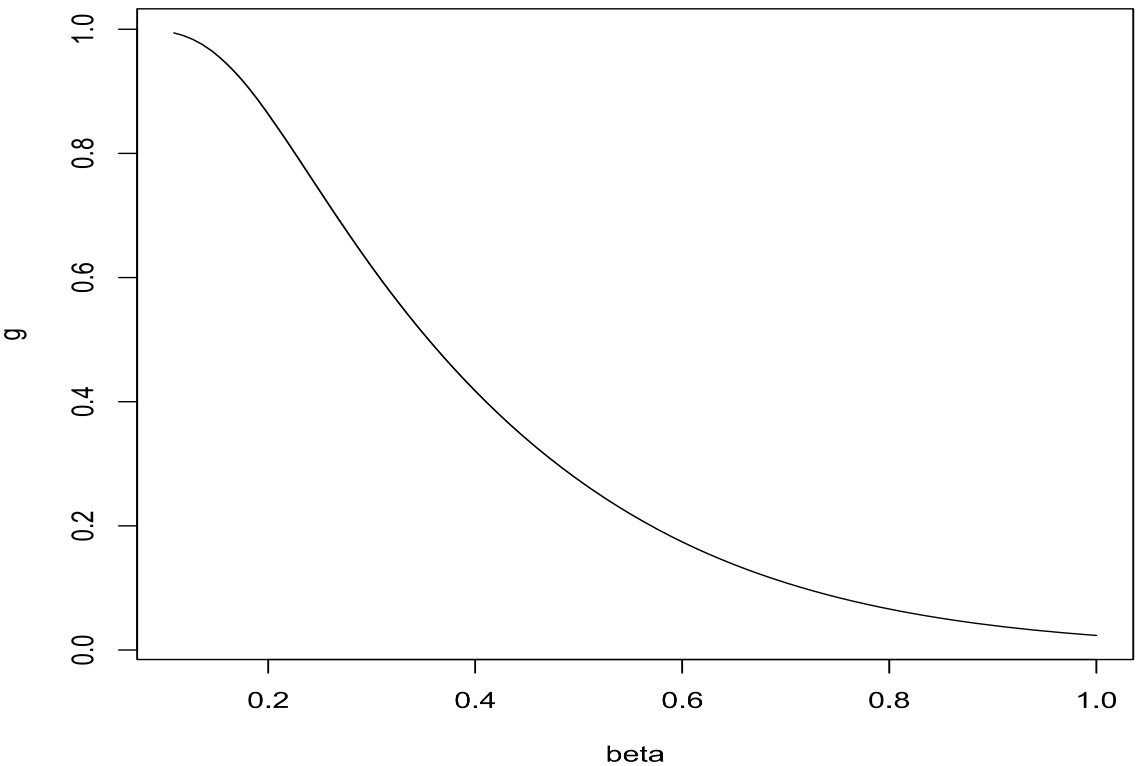

Fig. 1. $g = g(\beta)$, $0.1 \le \beta \le 1$.

and the upper bound (4.14) follows. The lower bound is obvious, since

$$\min_{I\in\mathcal{I}} \sup_{f\in F_\beta(Q)} \sup_{x\in\mathbb{R}} \mathbf{E}(I\mathbf{y}(x) - f(x))^2$$
$$\ge \max\Big(\min_{I\in\mathcal{I}} \sup_{x\in\mathbb{R}} \mathrm{Var}\ I\mathbf{y}(x),\ \min_{I\in\mathcal{I}} \sup_{f\in F_\beta(Q)} \sup_{x\in\mathbb{R}} (I\mathbf{f}(x) - f(x))^2\Big) \ge \max(\sigma^2, r_0^2).$$

□

Corollary 4.1. *For $\sigma^2 \ge \sigma_0^2$, the interpolant $I_\beta\mathbf{y}$ is R-optimal, i.e.,*

$$\sup_{f\in F_\beta(Q)} \sup_{x\in\mathbb{R}} \mathbf{E}\big(I_\beta\mathbf{y}(x) - f(x)\big)^2 = \sigma^2.$$

Note that the equation (4.15) illustrates the effect of *interference* which eliminates altogether the risk invoked by the bias. Of course, no interpolant can possibly be R-optimal if σ^2 is smaller than the (exponentially small) lower bound r_0^2. Incidentally, this implies that $\sigma_0^2 \ge r_0^2$. Nevertheless, the above Corollary claims that I_β is R-optimal for all σ^2 exceeding the (exponentially small) value σ_0^2; cf., the following

Remark 4.1. By (3.23), (5.12), and Lemma 4.1 (b), for $\beta \to \infty$,

$$g \le 2\sqrt{6}\beta e^{-2\pi\beta}\big(1 + O(\beta^{-1})\big).$$

A plot of $g = g(\beta)$ for $0.1 \le \beta \le 1$ is shown in Fig. 1.

Remark 4.2. As mentioned in Section 3.2.3, the limiting version of the interpolant I_β for $\beta \to \infty$ is the famous *sinc filter* I_∞, see (3.28). Thus, one may expect I_∞ to have similar optimality properties. In fact, these properties are almost obvious (with $\sigma_0^2 = 0$).

Proposition 4.2. *The sinc interpolant* $I_\infty \mathbf{y}$ *is* **(a)** *D-optimal;* **(b)** *R-optimal for all* $\sigma \geq 0$*, w.r.t. the model* $(\mathcal{X}, \mathcal{E}_\pi)$.

Indeed, D-optimality of $I_\infty \mathbf{y}$ follows from the equation, cf. [7] **1.422**(4),

$$\operatorname{Var} I_\infty \mathbf{y}(x) = \frac{\sigma^2 \sin^2 \pi x}{\pi^2} \sum_{j=-\infty}^{\infty} \frac{1}{(x-j)^2} \equiv \sigma^2.$$

The R-optimality (as well as A-optimality) of $I_\infty \mathbf{y}$ follows trivially from its unbiasedness in $\mathcal{E}_\pi$, cf. (3.32).

Remark 4.3 (A further generalization). The above results can be easily extended to a more general class of cardinal designs $\mathcal{X}_h =: \{x_j = jh, j \in \mathbb{Z}\}$. To this end, note first that for any $h > 0$, the relation

$$f(w) = g(wh) \tag{4.16}$$

establishes a bijection between the classes $F_\beta(Q) = \{f\}$ and $F_\alpha(Q) = \{g\}$, with $\alpha = \beta h$, as well as between the classes $\mathcal{E}_\pi$ and $\mathcal{E}_\gamma$, with $\gamma = \pi/h$.

Suppose the values $\mathbf{g} = \{g(jh),\ j \in \mathbb{Z}\}$ of a function g in (4.16) are given at the nodes $x_j = jh \in \mathcal{X}_h$, possibly contaminated by a white noise, $y_j = g(jh) + e_j \equiv f(j) + e_j$. To get an interpolant for $g(w)$, one can, in principle, interpolate $f(w)$ first, by using I_β in (4.1), and then apply (4.16). This leads to the following more general interpolant,

$$I_{\alpha,h}\, \mathbf{g}(x) = \frac{\pi h}{4\alpha \mathbf{K}} \operatorname{sn}\Big(\frac{2\mathbf{K}x}{h}, k\Big) \sum_{j=-\infty}^{\infty} \frac{(-1)^j g(jh)}{\sinh\big(\frac{\pi}{2\alpha}(x - jh)\big)}.$$

It has the same optimality properties as those described for the interpolant I_β in Proposition 4.1 and Theorem 4.2, where β should be replaced by α/h; see, in particular, (3.23), (4.5), and (5.11).

In the limiting case $\alpha \to \infty$, $I_{\alpha,h}\, \mathbf{g}(x)$ converges to the *cardinal series*

$$I_{\infty,h}\mathbf{g}(x) = \sum_{j=-\infty}^{\infty} \frac{\sin \gamma(x - jh) g(jh)}{\gamma(x - jh)} = \sin \gamma x \sum_{j=-\infty}^{\infty} \frac{(-1)^j g(jh)}{\gamma(x - jh)}, \qquad \gamma = \frac{\pi}{h}.$$

It has the same optimality properties with respect to the model $(\mathcal{X}_h, \mathcal{E}_{\pi/h})$ as those described by Proposition 4.2 with respect to $(\mathcal{X}, \mathcal{E}_\pi)$.

5. APPENDIX

This section summarizes for the reader's convenience some properties of the Jacobi elliptic functions used in this paper. A similar review is provided in [14]. The underlying theory can be found in [1], [3]. All the formulas below are cited from [3], Ch. 13.

The integral

$$\mathbf{K} = \mathbf{K}(k) = \int_0^1 \frac{dx}{\sqrt{(1 - x^2)(1 - k^2 x^2)}}$$

is called the *complete elliptic integral of the 1st kind of modulus* k. Only the so-called normal case $0 < k < 1$ is considered in this paper. The *complementary modulus* is defined as $k' = \sqrt{1 - k^2}$. Let $\mathbf{K}' = \mathbf{K}(k')$. Obviously, for all $0 < k < 1$,

$$\mathbf{K}(k) > \int_0^1 \frac{dx}{\sqrt{1 - x^2}} = \frac{\pi}{2}, \tag{5.1}$$

and as $k \to 0$,

$$\mathbf{K}(k) = \frac{\pi}{2} + O(k^2). \tag{5.2}$$

By [3], Section 13.8,

$$\mathbf{K}(k) \leq \frac{\pi}{2} + \log \frac{1}{k'}. \tag{5.3}$$

For $-1 \le y \le 1$, the integral

$$x = \int_0^y \frac{du}{\sqrt{(1-u^2)(1-k^2u^2)}}$$

defines a monotone continuous odd function $x = x(y)$ increasing from $-\mathbf{K}$ to $\mathbf{K}$. Its inverse is the Jacobi *elliptic sinus*, $y = \operatorname{sn} x = \operatorname{sn}(x,k)$ − an odd monotone function mapping $[-\mathbf{K}, \mathbf{K}]$ into $[-1,1]$, [3], Section 13.16. It can be continued to a $(4\mathbf{K}, 2i\mathbf{K}')$ doubly periodic meromorphic function $\operatorname{sn}(w,k)$, $w \in \mathbb{C}$. Clearly,

$$\operatorname{sn}(w,0) = \sin w. \tag{5.4}$$

The other *main Jacobi function*, $\operatorname{cn} x = \operatorname{cn}(x,k)$, satisfies [3], Section 13.17

$$\operatorname{cn}^2 w + \operatorname{sn}^2 w = 1. \tag{5.5}$$

Some properties of the Jacobi functions are:

$$\operatorname{sn}(2j\mathbf{K},k) = 0, \quad \operatorname{sn}'(2j\mathbf{K},k) = \operatorname{sn}((2j+1)\mathbf{K},k) = (-1)^j, \qquad j \in \mathbb{Z}, \tag{5.6}$$

$$\operatorname{sn}(w - 2j\mathbf{K},k) = (-1)^j \operatorname{sn}(w,k), \qquad j \in \mathbb{Z}, \tag{5.7}$$

$$\operatorname{sn}(w,k) = w\Big(1 - \frac{1+k^2}{6}w^2 + O(w^4)\Big), \qquad w \to 0. \tag{5.8}$$

$$\operatorname{sn}(w + i\mathbf{K}',k) = \frac{1}{k \operatorname{sn}(w,k)}, \tag{5.9}$$

$$\operatorname{res}_{2j\mathbf{K}+i\mathbf{K}'} \operatorname{sn}(\cdot,k) = \frac{(-1)^j}{k}. \tag{5.10}$$

The parameter $q = q(k) = \exp(-\pi\mathbf{K}'/\mathbf{K})$ is called *nome*. In terms of the nome, cf. [12], Eq. (3.27), or [14], Eq. (A.38), p. 200,

$$k = k(q) = 4q^{1/2}\left(\frac{\sum_{m=0}^{\infty} q^{m(m+1)}}{1 + 2\sum_{m=1}^{\infty} q^{m^2}}\right)^2. \tag{5.11}$$

Note that by (5.11),

$$k = 4q^{1/2}(1 + O(q)), \qquad q \to 0. \tag{5.12}$$

The Jacobi elliptic functions can be expressed in terms of the *Jacobi theta functions*. These functions can be defined as

$$\theta_1(v) = 2q^{1/4}\sum_{n=0}^{\infty}(-1)^n q^{n(n+1)} \sin(2n+1)\pi v,$$

$$\theta_2(v) = 2q^{1/4}\sum_{n=0}^{\infty} q^{n(n+1)} \cos(2n+1)\pi v,$$

$$\theta_3(v) = 1 + 2\sum_{n=1}^{\infty} q^{n^2} \cos 2n\pi v,$$

$$\theta_4(v) = 1 + 2\sum_{n=1}^{\infty}(-1)^n q^{n^2} \cos 2n\pi v.$$

Let

$$q_0 = \prod_{n=1}^{\infty}(1 - q^{2n}).$$

The Jacobi theta functions can be represented by the following infinite products,

$$\theta_1(v) = 2q_0 q^{1/4} \sin \pi v \prod_{n=1}^{\infty}(1 - 2q^{2n}\cos 2nv + q^{4n}), \tag{5.13}$$

$$\theta_2(v) = 2q_0 q^{1/4} \cos \pi v \prod_{n=1}^{\infty} (1 + 2q^{2n} \cos 2nv + q^{4n}), \tag{5.14}$$

$$\theta_3(v) = q_0 \prod_{n=1}^{\infty} (1 + 2q^{2n-1} \cos 2nv + q^{4n-2}), \tag{5.15}$$

$$\theta_4(v) = q_0 \prod_{n=1}^{\infty} (1 - 2q^{2n-1} \cos 2nv + q^{4n-2}). \tag{5.16}$$

One gets, see [3], Section 13.20,

$$k^{1/2} = \frac{\theta_2(0)}{\theta_3(0)}, \qquad k'^{1/2} = \frac{\theta_4(0)}{\theta_3(0)}, \tag{5.17}$$

$$\operatorname{sn}(2\mathbf{K}v, k) = \frac{\theta_3(0)\theta_1(v)}{\theta_2(0)\theta_4(v)} = \frac{1}{\sqrt{k}} \frac{\theta_1(v)}{\theta_4(v)}, \qquad \operatorname{cn}(2\mathbf{K}v, k) = \frac{\theta_4(0)\theta_2(v)}{\theta_2(0)\theta_4(v)} = \sqrt{\frac{k'}{k}} \frac{\theta_2(v)}{\theta_4(v)}. \tag{5.18}$$

The following relation is known as the *second principal first degree transform*, [3], Section 13.22, Table 11,

$$\frac{\operatorname{sn}(iu, k')}{\operatorname{cn}(iu, k')} = i \operatorname{sn}(u, k). \tag{5.19}$$

REFERENCES

1. N. I. Akhiezer, *Elements of the Theory of Elliptic Functions* (AMS, Providence, R.I., 1970).
2. L. Artiles and B. Levit, "Adaptive Regression on the Real Line in Classes of Smooth Functions", Austrian J. Statist. **32**, 99–129 (2003).
3. H. Bateman and A. Erdélyi, *Higher Transcendental Functions* (McGraw-Hill, New York, 1955), Vol. 2.
4. R. P. Boas, *Entire Functions* (Academic Press, New York, 1954).
5. J. Cho and B. Levit, "Asymptotic Optimality of Periodic Spline Interpolation in Nonparametric Regression", J. Statist. Theory Pract. **2**, 465–474 (2008).
6. Y. Golubev, B. Levit, and A. Tsybakov, "Asymptotically Efficient Estimation of Analytic Functions in Gaussian Noise", Bernoulli **2**, 167–181 (1996).
7. I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products* (Academic Press, New York, 1994), 5th ed.
8. I. Ibragimov, "Estimation of Analytic Functions", in *IMS Lecture Notes Monogr. Ser.* **36** (2001), pp. 359–383.
9. B. Ya. Levin, Yu. Lyubarski, M. Sodin, and V. Tkachenko, *Lectures on Entire Functions* (AMS, Providence, RI, 1996).
10. B. Levit, "Some New Perspectives in Best Approximation and Interpolation of Random Data", Math. Meth. Statist. **22**, 165–192 (2013).
11. B. Levit, "Variance-Optimal Data Approximation Using the Abel–Jacobi Functions", Math. Meth. Statist. **24**, 37–54 (2015).
12. B. Levit, "Optimal Methods of Interpolation in Nonparametric Regression", Math. Meth. Statist. **25**, 235–261 (2016).
13. A. I. Markushevich, *The Theory of Functions of a Complex Variable* (Chelsea Publ., New York, 1985).
14. K. Yu. Osipenko, *Optimal Recovery of Analytic Functions* (Nova Science Publ., New York, 2000).
15. A. Papoulis, *Signal Analysis* (McGraw-Hill, New York, 1977).
16. W. Rudin, *Real and Complex Analysis* (McGraw-Hill, New York, 1987), 3d ed.
17. I. J. Schoenberg, *Cardinal Spline Interpolation* (SIAM, Philadelphia, PA, 1973).
18. A. N. Shiryaev, *Probability-1*, in *Graduate Texts in Mathematics* (Springer, New York, 2016), 3rd ed.