AN ELEMENTARY AND UNIFIED PROOF OF GROTHENDIECK’S INEQUALITY

SHMUEL FRIEDLAND, LEK-HENG LIM, AND JINJIE ZHANG

Abstract. We present an elementary, self-contained proof of Grothendieck’s inequality that unifies the real and complex cases and yields both the Krivine and Haagerup bounds, the current best-known explicit bounds for the real and complex Grothendieck constants respectively. This article is intended to be pedagogical, combining and streamlining known ideas of Lindenstrauss-Pelczyński, Krivine, and Haagerup into a proof that need only univariate calculus, basic complex variables, and a modicum of linear algebra as prerequisites.

1. Introduction

We will let $F = \mathbb{R}$ or $\mathbb{C}$ throughout this article. In 1953, Grothendieck proved a powerful result that he called “the fundamental theorem in the metric theory of tensor products” [9]; he showed that there exists a finite constant $K > 0$ such that for every $l, m, n \in \mathbb{N}$ and every matrix $M = (M_{ij}) \in F^{m \times n}$,

$$\max_{\|x_i\|=\|y_j\|=1} \left| \sum_{i=1}^{m} \sum_{j=1}^{n} M_{ij} \langle x_i, y_j \rangle \right| \leq K \max_{\|\varepsilon_i\|=\|\delta_j\|=1} \left| \sum_{i=1}^{m} \sum_{j=1}^{n} M_{ij} \varepsilon_i \delta_j \right|$$

where $\langle \cdot, \cdot \rangle$ is the standard inner product in $F^l$, the maximum on the left is taken over all $x_i, y_j \in F^l$ of unit 2-norm, and the maximum on the right is taken over all $\varepsilon_i, \delta_j \in F$ of unit absolute value (i.e., $\varepsilon_i = \pm 1$, $\delta_j = \pm 1$ over $\mathbb{R}$; $\varepsilon_i = e^{i\theta_i}$, $\delta_j = e^{i\phi_j}$ over $\mathbb{C}$). The inequality (1) has since been christened Grothendieck’s inequality and the smallest possible constant $K$ Grothendieck’s constant.

The value of Grothendieck’s constant depends on the choice of $F$ and we will denote it by $K^F_G$.

Over the last 65 years, there have been many attempts to improve and simplify the proof of Grothendieck’s inequality, and also to obtain better bounds for the Grothendieck constant $K^F_G$, whose exact value remains unknown. The following are some major milestones:

(i) The central result of Grothendieck’s original paper [9] is that his eponymous inequality holds with $\pi/2 \leq K^\mathbb{R}_G \leq \sinh(\pi/2) \approx 2.301$ and $1.273 \approx 4/\pi \leq K^\mathbb{C}_G$. Grothendieck relied on the sign function for the real case and obtained the complex case from the real case via a complexification argument.

(ii) The power of Grothendieck’s inequality was not generally recognized until the work of Lindenstrauss and Pełczyński [10] 15 years later, which connected the inequality to absolutely $p$-summing operators. They elucidated and improved Grothendieck’s proof in the real case by computing expectations of sign functions and using Taylor expansions, although they did not get better bounds for $K^\mathbb{R}_G$.

(iii) Rietz [21] obtained a slightly smaller bound $K^\mathbb{R}_G \leq 2.261$ in 1974 by averaging over $\mathbb{R}^n$ with normalized Gaussian measure and using a variational argument to determine an optimal scalar map corresponding to the sign function.
(iv) Our current best known upper bounds for $K^R_G$ and $K^C_G$ are due to Krivine [14], who in 1979 used Banach space theory and ideas in [16] to get

$$K^R_G \leq \frac{\pi}{2 \log(1 + \sqrt{2})} \approx 1.78221;$$

and Haagerup [10], who in 1987 extended Krivine’s ideas to $C$ to get

$$K^C_G \leq \frac{8}{\pi(x_0 + 1)} \approx 1.40491,$$

where $x_0 \in [0, 1]$ is the unique solution to:

$$x \int_0^{\pi/2} \frac{\cos^2 t}{\sqrt{1 - x^2 \sin^2 t}} dt = \frac{\pi}{8} (x + 1).$$

(v) Our current best known lower bounds for $K^R_G$ and $K^C_G$ are due to Davie [4, 5], who in 1984 used spherical integrals to get

$$K^R_G \geq \sup_{x \in (0, 1)} \frac{1 - \rho(x)}{\max(\rho(x), f(x))} \approx 1.67696,$$

where

$$\rho(x) := \sqrt{\frac{2}{\pi}} xe^{-x^2/2}, \quad f(x) := \frac{2}{\pi} e^{-x^2} + \rho(x) \left[ 1 - \sqrt{\frac{8}{\pi}} \int_x^\infty e^{-t^2/2} dt \right];$$

and

$$K^C_G \geq \sup_{x > 0} \frac{1 - \theta(x)}{g(x)} \approx 1.33807,$$

where

$$\theta(x) := \frac{1}{2} \left[ 1 - e^{-x^2} + x \int_x^\infty e^{-t^2} dt \right],$$

$$g(x) := \left[ \frac{1}{x} (1 - e^{-x^2}) + \int_x^\infty e^{-t^2} dt \right]^2 + \theta(x) \left[ 1 - \frac{2}{x} (1 - e^{-x^2}) \right].$$

(vi) Progress on improving the aforementioned bounds halted for many years. Believing that Krivine’s bound is the exact value of $K^R_G$, some were spurred to find matrices that yield it as the lower bound of $K^R_G$ [13]. The belief was dispelled in 2011 in a landmark paper [3], which demonstrated the existence of a positive constant $\varepsilon$ such that $K^R_G < \pi/(2 \log(1 + \sqrt{2})) - \varepsilon$ but the authors did not provide an explicit better bound. To date, Krivine’s and Haagerup’s bounds remain the best known explicit upper bounds for $K^R_G$ and $K^C_G$ respectively.

(vii) There have also been many alternate proofs of Grothendieck’s inequality employing a variety of techniques, among them factorization of Hilbert spaces [18, 11, 19], absolutely summing operators [7, 16, 20], geometry of Banach spaces [1, 17], metric theory of tensor product [6], basic probability theory [2], bilinear forms on $C^*$-algebra [12].

In this article, we will present a proof of Grothendieck’s inequality that unifies both the (a) real and (b) complex cases; and yields both the (c) Krivine and (d) Haagerup bounds [14, 10]. It is also elementary in that it requires little more than standard college mathematics. Our proof will rely on Lemma 2.1, which is a variation of known ideas in [16, 10, 11]. In particular, the idea of using the sign function to establish (1) in the real case was due to Grothendieck himself [9] and later also appeared in [16, 14]; whereas the use of the sign function in the complex case first appeared in [10]. To be clear, all the key ideas in our proof were originally due to Lindenstrauss–Pełczyński, Krivine, and Haagerup [16, 14, 10], our only contribution is pedagogical — combining, simplifying, and streamlining their ideas into what we feel is a more palatable proof. To understand the proof, readers need only know univariate calculus, basic complex variables, and a small amount of linear
algebra. We will use some basic Hilbert space theory and tensor product constructions in Section 4 but both notions will be explained in a self-contained and elementary way.

2. GAUSSIAN INTEGRAL OF SIGN FUNCTION

Throughout this article, our inner product over \( \mathbb{C} \) will be sesquilinear in the second argument, i.e.,

\[
\langle x, y \rangle := y^* x \quad \text{for all } x, y \in \mathbb{C}^n.
\]

For \( z \in F = \mathbb{R} \) or \( \mathbb{C} \), the sign function is

\[
\text{sgn}(z) = \begin{cases} 
\frac{z}{|z|} & \text{if } z \neq 0, \\
0 & \text{if } z = 0;
\end{cases}
\]

and for \( z \in F^n \), the Gaussian function is

\[
G_n^F(z) = \begin{cases} 
(2\pi)^{-n/2} \exp\left(-\frac{\|z\|^2}{2}\right) & \text{if } F = \mathbb{R}, \\
\pi^{-n} \exp\left(-\frac{\|z\|^2}{2}\right) & \text{if } F = \mathbb{C}.
\end{cases}
\]

Lemma 2.1 below is based on [11,10]; the complex version in particular is a slight variation of [10, Lemma 3.2]. It plays an important role in our proof because the right side of (3) depends only on the inner product \( \langle u, v \rangle \) and not (explicitly) on the dimension \( n \). In addition, the functions on the right are homeomorphisms and admit Taylor expansions, making it possible to expand them in powers \( (u, v)^d \), which will come in useful when we prove Theorem 4.1.

**Lemma 2.1.** Let \( u, v \in F^n \) with \( \|u\|_2 = \|v\|_2 = 1 \). Then

\[
\int_{\mathbb{R}^n} \text{sgn}(u, z) \text{sgn}(z, v) G_n^F(z) \, dz = \begin{cases} 
\frac{2}{\pi} \arcsin \langle u, v \rangle & \text{if } F = \mathbb{R}, \\
\langle u, v \rangle \int_0^{\pi/2} \frac{\cos^2 t}{(1 - |\langle u, v \rangle|^2 \sin^2 t)^{1/2}} \, dt & \text{if } F = \mathbb{C}.
\end{cases}
\]

**Proof.** Case I: \( F = \mathbb{R} \). Let \( \arccos \langle u, v \rangle = \theta \), so that \( \theta \in [0, \pi] \) and \( \arcsin \langle u, v \rangle = \pi/2 - \theta \). Choose \( \alpha, \beta \) such that \( 0 < \beta - \alpha < \pi \) and define

\[
E(\alpha, \beta) = \{(r \cos \theta, r \sin \theta, x_3, \ldots, x_n) : r \geq 0, \alpha \leq \theta \leq \beta\}.
\]

The Gaussian measure of a measurable set \( A \) is the integral of \( G_n^F(x) \) over \( A \). Upon integrating with respect to \( x_3, \ldots, x_n \), the following term remains:

\[
\frac{1}{2\pi} \int_{E(\alpha, \beta)} e^{-\frac{1}{2}(x_1^2 + x_2^2)} \,dx_1 \,dx_2 = \frac{1}{2\pi} \int_\alpha^\beta d\theta \int_0^\infty re^{-\frac{1}{2}r^2} \,dr = (\beta - \alpha)/2\pi.
\]

Hence the Gaussian measure of \( E(\alpha, \beta) \) is \((\beta - \alpha)/2\pi\). Since there is an isometry \( T \) of \( \mathbb{R}^n \) such that \( Tu = e_1 \) and \( Tv = (\cos \theta, \sin \theta, 0, \ldots, 0) \), the left side of (3) may be expressed as

\[
\int_{\mathbb{R}^n} \text{sgn}(Tu, x) \text{sgn}(x, Tv) G_n^F(x) \, dx.
\]

The set of \( x \) where \( \langle Tu, x \rangle > 0 \) and \( \langle Tv, x \rangle > 0 \) is \( E(\theta - \pi/2, \pi/2) \), which has Gaussian measure \((\pi - \theta)/2\pi\); ditto for \( \langle Tu, x \rangle < 0 \) and \( \langle Tv, x \rangle < 0 \). The set of \( x \) where \( \langle Tu, x \rangle < 0 \) and \( \langle Tv, x \rangle > 0 \) is \( E(\pi/2, \theta + \pi/2) \), which has Gaussian measure \( \theta/2\pi \); ditto for \( \langle Tu, x \rangle > 0 \) and \( \langle Tv, x \rangle < 0 \). The set of \( x \) where \( \langle Tu, x \rangle = 0 \) has zero Gaussian measure. Hence the value of this integral is \((\pi - \theta)/2\pi + (\pi - \theta)/2\pi - \theta/2\pi - \theta/2\pi = 2 \arcsin \langle u, v \rangle / \pi\).

Case II: \( F = \mathbb{C} \). We define vectors \( \alpha, \beta \in \mathbb{R}^{2n} \) with \( \alpha_{2i-1} = \text{Re}(u_i), \alpha_{2i} = \text{Im}(u_i), \beta_{2i-1} = \text{Re}(v_i), \beta_{2i} = \text{Im}(v_i), i = 1, \ldots, n \). Then \( \alpha \) and \( \beta \) are unit vectors in \( \mathbb{R}^{2n} \). For any \( z = (z_1, \ldots, z_n) \in \mathbb{C}^n \), we write

\[
x = (\text{Re}(z_1), \text{Im}(z_1), \ldots, \text{Re}(z_n), \text{Im}(z_n)) \in \mathbb{R}^{2n}.
\]
Then,
\[
\text{Re}(\langle u, z \rangle) = \sum_{i=1}^{n} \text{Re}(u_i z_i) = \frac{1}{n} \sum_{i=1}^{n} \left( \text{Re}(u_i) \text{Re}(z_i) + \text{Im}(u_i) \text{Im}(z_i) \right) = \langle \alpha, x \rangle = \langle x, \alpha \rangle,
\]
and likewise \(\text{Re}(\langle z, v \rangle) = \langle x, \beta \rangle\). By a change-of-variables and Case I, we have
\[
\int_{\mathbb{C}^n} \text{sgn}(\text{Re}(\langle u, z \rangle)) \text{sgn}(\text{Re}(\langle z, v \rangle)) G_{n}^{\mathbb{C}}(z) \, dz = \int_{\mathbb{R}^2} \text{sgn}(x, \alpha) \text{sgn}(x, \beta) G_{2n}^{\mathbb{R}}(x) \, dx
\]
(4)
\[
= \frac{2}{\pi} \arcsin(\alpha, \beta) = \frac{2}{\pi} \arcsin(\text{Re}(\langle u, v \rangle)).
\]
It is easy to verify that for any \(z \in \mathbb{C}\),
(5) \quad \text{sgn}(z) = \frac{1}{4} \int_{0}^{2\pi} \text{sgn}(\text{Re}(e^{-i\theta} z)) e^{i\theta} \, d\theta.

By (4), (5), and Fubini’s theorem,
\[
\int_{\mathbb{C}^n} \text{sgn}(\langle u, z \rangle) \text{sgn}(\langle z, v \rangle) G_{n}^{\mathbb{C}}(z) \, dz
\]
\[
= \frac{1}{16} \int_{0}^{2\pi} \int_{0}^{2\pi} \int_{\mathbb{C}^n} \text{sgn}(\langle e^{-i\theta} u, z \rangle) \text{sgn}(\langle z, e^{-i\varphi} v \rangle) e^{i(\theta + \varphi)} G_{n}^{\mathbb{C}}(z) \, d\theta \, d\varphi
\]
\[
= \frac{1}{8\pi} \int_{0}^{2\pi} \int_{0}^{2\pi} \arcsin(\text{Re}(\langle e^{-i\theta} u, e^{i\varphi} v \rangle)) e^{i(\theta + \varphi)} \, d\theta \, d\varphi.
\]
CASE II(a): \(\langle u, v \rangle \in \mathbb{R}\). The integral above becomes
\[
= \frac{1}{8\pi} \int_{0}^{2\pi} \left[ \int_{0}^{2\pi} \arcsin(\cos(\theta + \varphi) \langle u, v \rangle) e^{i(\theta + \varphi)} \, d\theta \right] d\varphi
\]
\[
= \frac{1}{8\pi} \int_{0}^{2\pi} \left[ \int_{0}^{2\pi + \varphi} \arcsin(\langle u, v \rangle \cos t) e^{it} \, dt \right] d\varphi
\]
(6)
\[
= \frac{1}{8\pi} \int_{0}^{2\pi} \left[ \int_{0}^{2\pi} \arcsin(\langle u, v \rangle \cos t) e^{it} \, dt \right] d\varphi
\]
Since \(\arcsin(\langle u, v \rangle \cos t)\) is an even function with period \(2\pi\),
\[
\int_{0}^{2\pi} \arcsin(\langle u, v \rangle \cos t) \sin t \, dt = 0,
\]
the last integral in (6) becomes
\[
\frac{1}{4} \int_{0}^{2\pi} \arcsin(\langle u, v \rangle \cos t) \cos t \, dt,
\]
and as \(\arcsin(\langle u, v \rangle \cos t)\) \(\cos t\) is an even function with period \(\pi\), it becomes
\[
\int_{0}^{\pi/2} \arcsin(\langle u, v \rangle \cos t) \cos t \, dt = \int_{0}^{\pi/2} \arcsin(\langle u, v \rangle \sin t) \sin t \, dt,
\]
which, upon integrating by parts, becomes
(7) \quad \langle u, v \rangle \int_{0}^{\pi/2} \frac{\cos^2 t}{(1 - |\langle u, v \rangle|^2 \sin^2 t)^{1/2}} \, dt.
CASE II(b): \( \langle u, v \rangle \notin \mathbb{R} \). This reduces to Case II(a) by setting \( c \in \mathbb{C} \) of unit modulus so that \( c\langle u, v \rangle = |\langle u, v \rangle| \) and \( \langle cu, v \rangle \in \mathbb{R} \), then by (7),

\[
\int_{\mathbb{C}^n} \operatorname{sgn}(u, z) \operatorname{sgn}(z, v) G_n^C(z) \, dz = \bar{c} \int_{\mathbb{C}^n} \operatorname{sgn}(cu, z) \operatorname{sgn}(z, v) G_n^C(z) \, dz
\]

\[
= \bar{c}\langle cu, v \rangle \int_0^{\pi/2} \frac{\cos^2 t}{(1 - |\langle cu, v \rangle|^2 \sin^2 t)^{1/2}} \, dt = \langle u, v \rangle \int_0^{\pi/2} \frac{\cos^2 t}{(1 - |\langle u, v \rangle|^2 \sin^2 t)^{1/2}} \, dt.
\]

We will make a simple but useful observation\footnote{This of course follows from other well-known results but we would like to keep our exposition self-contained.} about the quantities in (8) that we will need for the proof of Corollary 2.3 later.

**Lemma 2.2.** Let \( F = \mathbb{R} \) or \( \mathbb{C} \) and \( d, m, n \in \mathbb{N} \). For any \( M \in F^{m \times n} \), we have

\[
(8) \quad \max_{|\varepsilon_i| \leq 1} \sum_{i=1}^{m} \sum_{j=1}^{n} M_{ij} \varepsilon_i \delta_j = \max_{|\varepsilon_i| = |\delta_j| = 1} \left| \sum_{i=1}^{m} \sum_{j=1}^{n} M_{ij} \varepsilon_i \delta_j \right|
\]

and for any \( x_1, \ldots, x_m, y_1, \ldots, y_n \in F^d \),

\[
(9) \quad \max_{\|x_i\| \leq 1, \|y_j\| \leq 1} \left| \sum_{i=1}^{m} \sum_{j=1}^{n} M_{ij} \langle x_i, y_j \rangle \right| = \max_{\|x_i\| = \|y_j\| = 1} \left| \sum_{i=1}^{m} \sum_{j=1}^{n} M_{ij} \langle x_i, y_j \rangle \right|
\]

**Proof.** We will start with (8). Suppose there exists \( M \in F^{m \times n} \) such that the left-hand side of (8) exceeds the right-hand side. Let the maximum of the left-hand side be attained by \( \varepsilon_1^*, \ldots, \varepsilon_m^* \) and \( \delta_1^*, \ldots, \delta_n^* \). By our assumption, at least one \( \varepsilon_i^* \) or \( \delta_j^* \) must be less than 1 in absolute value and so let \( |\varepsilon_1^*| < 1 \) without loss of generality. Fix \( \varepsilon_i = \varepsilon_i^*, i = 2, \ldots, m \) and \( \delta_j = \delta_j^*, j = 1, \ldots, n \), but let \( \varepsilon_1 \) vary with \( |\varepsilon_1| \leq 1 \) and consider the maximum of the left-hand side over \( \varepsilon_1 \). Since \( \max\{ |a\varepsilon_1 + b| : |\varepsilon_1| \leq 1 \} \) is always attained on the boundary \( |\varepsilon_1| = 1 \) for any \( a, b \in F \), this contradicts our assumption. The proof for (9) is similar with norm in place of absolute value.

In the corollary below, the inequality on the left is the “original Grothendieck inequality,” i.e., as first stated by Grothendieck\footnote{The better known modern version \cite{16} is in fact due to Lindenstrauss and Pelczyński in [19].} in [9], and the inequality on the right is due to Haagerup\footnote{[10].}.

**Corollary 2.3.** Let \( F = \mathbb{R} \) or \( \mathbb{C} \) and \( d, m, n \in \mathbb{N} \). For any \( M \in F^{m \times n} \) with

\[
(10) \quad \max_{|\varepsilon_i| = |\delta_j| = 1} \left| \sum_{i=1}^{m} \sum_{j=1}^{n} M_{ij} \varepsilon_i \delta_j \right| \leq 1,
\]

any \( x_1, \ldots, x_m, y_1, \ldots, y_n \in F^d \) of unit 2-norm, we have

\[
\left| \sum_{i=1}^{m} \sum_{j=1}^{n} M_{ij} \arcsin \langle x_i, y_j \rangle \right| \leq \frac{\pi}{2} \quad \text{if } F = \mathbb{R}, \quad \left| \sum_{i=1}^{m} \sum_{j=1}^{n} M_{ij} H(\langle x_i, y_j \rangle) \right| \leq 1 \quad \text{if } F = \mathbb{C},
\]

where \( H \) denotes the function on the right side of (3) for \( F = \mathbb{C} \).

**Proof.** The condition (10) implies that

\[
\left| \sum_{i=1}^{m} \sum_{j=1}^{n} M_{ij} \operatorname{sgn}(x_i, x) \operatorname{sgn}(y_j, x) G_d^R(z) \right| \leq G_d^R(z),
\]

\[
\left| \sum_{i=1}^{m} \sum_{j=1}^{n} M_{ij} \operatorname{sgn}(z, \overline{x_i}) \operatorname{sgn}(z, \overline{y_j}) G_d^C(z) \right| \leq G_d^C(z),
\]
for any \( x \in \mathbb{R}^d, z \in \mathbb{C}^d \) respectively. Integrating over \( \mathbb{R}^d \) or \( \mathbb{C}^d \) respectively and applying Lemma 2.1 give the required results. Note that we have implicitly relied on (8) in Lemma 2.2 as the \( \text{sgn} \) function is not always of absolute value one and may be zero. \( \square \)

Corollary 2.3 already looks a lot like the Grothendieck inequality \([1]\) but the nonlinear functions \( \arcsin \) and \( H \) are in the way. To obtain the Grothendieck inequality, we linearize them: First by using Taylor series to replace these functions by polynomials; and then using a ‘tensor trick’ to express the polynomials as linear functions on a larger space. This is the gist of the proof in Section 4.

3. Haagerup function

We will need to make a few observations regarding the functions on the right side of (3) for the proof of Grothendieck’s inequality. Let the complex Haagerup function of a complex variable \( z \) be

\[
H(z) := z \int_0^{\pi/2} \frac{\cos^2 t}{(1 - |z|^2 \sin^2 t)^{1/2}} \, dt, \quad |z| \leq 1,
\]

and the real Haagerup function \( h \) as the restriction of \( H \) to \([-1, 1] \subseteq \mathbb{R} \). Observe that \( h : [-1, 1] \to [-1, 1] \) and is a strictly increasing continuous bijection. Since \([-1, 1]\) is compact, \( h \) is a homeomorphism of \([-1, 1]\) onto itself. By the Taylor expansion

\[
(1 - x^2 \sin^2 t)^{-1/2} = \sum_{k=0}^{\infty} \frac{(2k - 1)!!}{(2k)!!} x^{2k} \sin 2k t, \quad |x| \leq 1, \quad 0 \leq t < \pi/2,
\]

and

\[
\int_0^{\pi/2} \cos^2 t \sin 2k t \, dt = \frac{\pi}{4(k + 1)} \frac{(2k - 1)!!}{(2k)!!},
\]

thus we get

\[
(11) \quad h(x) = \sum_{k=0}^{\infty} \frac{\pi}{4(k + 1)} \left[ \frac{(2k - 1)!!}{(2k)!!} \right]^2 x^{2k+1}, \quad x \in [-1, 1].
\]

Since \( h \) is analytic at \( x = 0 \) and \( h'(0) \neq 0 \), its inverse function \( h^{-1} : [-1, 1] \to [-1, 1] \) can be expanded in a power series in some neighborhood of 0

\[
(12) \quad h^{-1}(x) = \sum_{k=0}^{\infty} b_{2k+1} x^{2k+1}.
\]

One may in principle determine the coefficients using the Lagrange inversion formula:

\[
b_{2k+1} = \frac{1}{(2k + 1)!} \lim_{t \to 0} \left[ \frac{d^{2k}}{d t^{2k}} \left( \frac{t}{h(t)} \right)^{2k+1} \right].
\]

For example,

\[
b_1 = \frac{4}{\pi}, \quad b_3 = -\frac{1}{8} \left( \frac{4}{\pi} \right)^3, \quad b_5 = 0, \quad b_7 = -\frac{1}{1024} \left( \frac{4}{\pi} \right)^7.
\]

But determining \( b_{2k+1} \) explicitly becomes difficult as \( k \) gets larger. A key step in Haagerup’s proof \([10]\) requires the nonpositivity of the coefficients beyond the first:

\[
(13) \quad b_{2k+1} \leq 0, \quad \text{for all} \ k \geq 1.
\]

This step is in our view the most technical part of \([10]\). We have no insights on how it may be avoided but we simplified Haagerup’s proof of (13) in Section 5 to keep to our promise of an elementary proof — using only calculus and basic complex variables.

It follows from (13) that \( \tilde{h}(z) := b_1 z - h^{-1}(z) \) has nonnegative Taylor coefficients. Pringsheim’s theorem implies that if the radius of convergence of the Taylor series of \( \tilde{h}(z) \) is \( r \), then \( \tilde{h}(z) \), and
thus $h^{-1}(z)$, has a singular point at $z = r$. As $h'(t) > 0$ on $(0, 1)$ and $h(1) = 1$, we must have $r \geq 1$. It also follows from (13) that $h^{-1}(t) \leq \sum_{k=0}^{N} b_{2k+1} t^{2k+1}$ for any $t \in (0, 1)$ and $N \in \mathbb{N}$. So \[ \sum_{k=1}^{N} |b_{2k+1}| t^{2k+1} \leq b_1 t - h^{-1}(t) \] for any $t \in (0, 1)$ and $N \in \mathbb{N}$. So $\sum_{k=1}^{N} |b_{2k+1}| \leq b_1 - 1$ for any $N \in \mathbb{N}$ and we have $\sum_{k=0}^{\infty} |b_{2k+1}| \leq 2b_1 - 1$. As $h^{-1}(1) = 1$ we deduce that $\sum_{k=0}^{\infty} b_{2k+1} = h^{-1}(1) = 1$, and therefore

(14) \[ \sum_{k=0}^{\infty} |b_{2k+1}| = 2b_1 - 1. \]

We now turn our attention back to the complex Haagerup function. Observe that $|H(z)| = h(|z|)$ for all $z \in D := \{z \in \mathbb{C} : |z| \leq 1\}$ and $\arg(H(z)) = \arg(z)$ for $0 \neq z \in D$. So $H : D \to D$ is a homeomorphism of $D$ onto itself. Let $H^{-1} : D \to D$ be its inverse function. Since $H(z) = \text{sgn}(z) h(|z|)$, we get

(15) \[ H^{-1}(z) = \text{sgn}(z) h^{-1}(|z|) = \text{sgn}(z) \sum_{k=0}^{\infty} b_{2k+1} |z|^{2k+1}. \]

Dini’s theorem shows that the function $\varphi(x) := \sum_{k=0}^{\infty} b_{2k+1} |x|^{2k+1}$ is a strictly increasing and continuous on $[0, 1]$, with $\varphi(0) = 0$ and $\varphi(1) = \sum_{k=0}^{\infty} |b_{2k+1}| \geq b_1 = 4/\pi > 1$; note that $\varphi(1)$ is finite by (14). Thus there exists a unique $c_0 \in (0, 1)$ such that $\varphi(c_0) = 1$. So

\[ 1 = \varphi(c_0) = \sum_{k=0}^{\infty} |b_{2k+1}| c_0^{2k+1} = \frac{8}{\pi} c_0 - h^{-1}(c_0), \]

where the last equality follows from $b_1 = 4/\pi$ and (13). Therefore we obtain $h^{-1}(c_0) = 8c_0/\pi - 1$, and if we let $x_0 := h^{-1}(c_0) \in (0, 1)$, then $h(x_0) - \pi(x_0 + 1)/8 = 0$. From the Taylor expansion of $h(x)$, the function $x \mapsto h(x) - \pi(x + 1)/8$ is increasing and continuous on $[0, 1]$. Hence $x_0$ is the unique solution in $[0, 1]$ to

(16) \[ h(x) - \frac{\pi}{8}(x + 1) = 0 \]

and $c_0 = \pi(x_0 + 1)/8$.

As Corollary 2.3 indicates, the Haagerup function $H$ plays the analogue of arcsin in the complex case. Unlike arcsin, $H$ is a completely obscure function\(^3\) and any of its properties that we require will have to be established from scratch. The goal of this section is essentially to establish (11)–(16), which we will need later.

4. A unified proof of Grothendieck’s inequality

In this section we will need the notions of (i) tensor product and (ii) Hilbert space, but just enough to make sense of $\mathcal{H}_n(\mathbb{F}) = \bigoplus_{k=0}^{\infty} (\mathbb{F}^n)^{\otimes (2k+1)}$ where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. In keeping to our promise of an elementary proof, we will briefly introduce these notions in a simple manner. For our purpose, it suffices to regard the tensor product of $k$ copies of $\mathbb{F}^n$, denoted

\[ (\mathbb{F}^n)^{\otimes k} = \bigotimes_{k \text{ copies}} \mathbb{F}^n, \]

as the $\mathbb{F}$-vector space of $k$-dimensional hypermatrices,

\[ (\mathbb{F}^n)^{\otimes k} := \{ [a_{i_1 \ldots i_k}] : a_{i_1 \ldots i_k} \in \mathbb{F}, \ i_1, \ldots, i_k \in \{1, \ldots, n\}\}, \]

where scalar multiplication and vector addition of hypermatrices are defined coordinatewise. For $k$ vectors $x, y, \ldots, z \in \mathbb{F}^n$, their tensor product is the $k$-dimensional hypermatrix given by

\[ x \otimes y \otimes \cdots \otimes z := [x_{i_1} y_{i_2} \cdots z_{i_k}]_{i_1, i_2, \ldots, i_k=1}^{n} \in (\mathbb{F}^n)^{\otimes k}. \]

\(^3\)We are unaware of any other occurrence of $H$ outside its use in Haagerup’s proof of his bound in [10].
We write
\[ x^{\otimes k} := x \otimes \cdots \otimes x. \]

If \( \langle \cdot, \cdot \rangle \) is an inner product on \( \mathbb{F}^n \), then defining
\[ \langle x \otimes y \otimes \cdots \otimes z, x' \otimes y' \otimes \cdots \otimes z' \rangle := \langle x, x' \rangle \langle y, y' \rangle \cdots \langle z, z' \rangle \]
and extending bilinearly (if \( \mathbb{F} = \mathbb{R} \)) or sesquilinearly (if \( \mathbb{F} = \mathbb{C} \)) to all of \((\mathbb{F}^n)^{\otimes k}\) yields an inner product on the \( k \)-dimensional hypermatrices. In particular we have
\[ \langle x^{\otimes k}, y^{\otimes k} \rangle = \langle x, y \rangle^k. \]

If \( \{e_1, \ldots, e_n\} \) is the standard orthonormal basis of \( \mathbb{F}^n \), then
\[ \{e_{i_1} \otimes \cdots \otimes e_{i_k} \in (\mathbb{F}^n)^{\otimes k} : i_1, \ldots, i_k \in \{1, \ldots, n\} \} \]
is an orthonormal basis of \((\mathbb{F}^n)^{\otimes k}\). For more information about hypermatrices see [15] and for a more formal definition of tensor products see [8].

If an \( \mathbb{F} \)-vector space \( \mathcal{H} \) is equipped with an inner product \( \langle \cdot, \cdot \rangle \) such that every Cauchy sequence in \( \mathcal{H} \) converges with respect to the induced norm \( \|v\| = |\langle v, v \rangle|^{1/2} \), we call \( \mathcal{H} \) a Hilbert space. Hilbert spaces need not be finite-dimensional; we call \( \mathcal{H} \) separable if there is a countable set of orthonormal vectors \( \{e_j \in \mathcal{H} : j \in J\} \), i.e., \( J \) is a countable index set, such that every \( v \in \mathcal{H} \) satisfies
\[ \|v\|^2 = \sum_{j \in J} |\langle v, e_j \rangle|^2. \]

Let \( \langle \cdot, \cdot \rangle_k \) be the inner product on \((\mathbb{F}^n)^{\otimes (2k+1)}\) as defined in (17), \( \|\cdot\|_k \) be its induced norm, and \( B_k \) be the orthonormal basis in (18). Let \( n \in \mathbb{N} \). The \( \mathbb{F} \)-vector space
\[ \mathcal{H}_n(\mathbb{F}) := \bigoplus_{k=0}^{\infty} (\mathbb{F}^n)^{\otimes (2k+1)} = \{ (v_0, v_1, v_2, \ldots) : v_k \in (\mathbb{F}^n)^{\otimes (2k+1)}, \sum_{k=0}^{\infty} \|v_k\|_k^2 < \infty \} \]
equipped with the inner product
\[ \langle u, v \rangle_* := \sum_{k=0}^{\infty} \langle u_k, v_k \rangle_k \]
is a separable Hilbert space since \( \bigcup_{k=0}^{\infty} B_k \) is a countable set of orthonormal vectors satisfying (19). We write \( \|\cdot\|_* \) for the norm induced by (21).

**Theorem 4.1** (Grothendieck inequality with Krivine and Haagerup bounds). Let \( \mathbb{F} = \mathbb{R} \) or \( \mathbb{C} \) and \( l, m, n \in \mathbb{N} \). For any \( M \in \mathbb{F}^{m \times n} \), any \( x_1, \ldots, x_m, y_1, \ldots, y_n \in \mathbb{F}^l \) of unit 2-norm, we have
\[ \max_{\|x_i\| = \|y_j\| = 1} \left| \sum_{i=1}^{m} \sum_{j=1}^{n} M_{ij} \langle x_i, y_j \rangle \right| \leq K_{\mathbb{F}}^{\|\cdot\|} \max_{|\varepsilon_i| = |\delta_j| = 1} \left| \sum_{i=1}^{m} \sum_{j=1}^{n} M_{ij} \varepsilon_i \delta_j \right|, \]
where
\[ K_{\mathbb{R}} := \frac{\pi}{2 \log(1 + \sqrt{2})} \quad \text{and} \quad K_{\mathbb{C}} := \frac{8}{\pi(x_0 + 1)} \]
are Krivine’s and Haagerup’s bounds respectively. Recall that \( x_0 \) is as defined in (16).

**Proof.** As we described at the end of Section [2], we will ‘linearize’ the nonlinear functions arcsin and \( H \) in Corollary [2.3] by using Taylor series to replace these functions by polynomials, followed by a ‘tensor trick’ to express polynomials as linear functions on an infinite-dimensional space.

---

The direct sum in (20) is a Hilbert space direct sum, i.e., it is the closure of the vector space direct sum.
Case I: \( \mathbb{F} = \mathbb{R} \). Let \( c := \arcsinh(1) = \log(1 + \sqrt{2}) \). Taylor expansion gives

\[
\sin(c \langle x_i, y_j \rangle) = \sum_{k=0}^{\infty} (-1)^k \frac{c^{2k+1}}{(2k+1)!} (x_i, y_j)^{2k+1} = \sum_{k=0}^{\infty} (-1)^k \frac{c^{2k+1}}{(2k+1)!} (x_i^{(2k+1)} \otimes y_j^{(2k+1)})^k.
\]

For any \( l \in \mathbb{N} \), let \( \mathcal{H}_l(\mathbb{R}) \) be as in (20), and \( S, T : \mathbb{R}^l \to \mathcal{H}_l(\mathbb{R}) \) be nonlinear maps defined by

\[
S(x) := \left( S_k(x) \right)_{k=0}^{\infty}, \quad S_k(x) = (-1)^k \sqrt{\frac{c^{2k+1}}{(2k+1)!}} x^{(2k+1)},
\]

\[
T(x) := \left( T_k(x) \right)_{k=0}^{\infty}, \quad T_k(x) = \sqrt{\frac{c^{2k+1}}{(2k+1)!}} x^{(2k+1)},
\]

for any \( x \in \mathbb{R}^l \). To justify that \( S \) and \( T \) are indeed maps into \( \mathcal{H}_l(\mathbb{R}) \), we need to demonstrate that \( \|S(x)\|, \|T(x)\| < \infty \) but this follows from

\[
\|S(x)\|_*^2 = \sum_{k=0}^{\infty} \|S_k(x)\|_*^2 = \sum_{k=0}^{\infty} \frac{c^{2k+1}}{(2k+1)!} \|x\|^{2(2k+1)} = \sum_{k=0}^{\infty} \|T_k(x)\|_*^2 = \|T(x)\|_*^2
\]

and

\[
\sum_{k=0}^{\infty} \frac{c^{2k+1}}{(2k+1)!} \|x\|^{2(2k+1)} = \sinh(c \|x\|^2) < \infty
\]

for all \( x \in \mathbb{R}^l \). Note that

\[
\langle S(x), T(y) \rangle_* = \sum_{k=0}^{\infty} (-1)^k \frac{c^{2k+1}}{(2k+1)!} (x, y)^{2k+1} = \sin(c \langle x, y \rangle).
\]

Hence (23) becomes:

\[
\sin(c \langle x_i, y_j \rangle) = \langle S(x_i), T(y_j) \rangle_* \quad \text{or} \quad c \langle x_i, y_j \rangle = \arcsin \langle S(x_i), T(y_j) \rangle_*.
\]

Moreover, since \( x_i \) and \( y_j \) are unit vectors in \( \mathbb{R}^l \), we get

\[
\|S(x_i)\|_*^2 = \sinh(c \|x_i\|^2) = 1 \quad \text{and} \quad \|T(y_j)\|_*^2 = \sinh(c \|y_j\|^2) = 1.
\]

As the \( m + n \) vectors \( S(x_1), \ldots, S(x_m), T(y_1), \ldots, T(y_n) \) in \( \mathcal{H}_l(\mathbb{R}) \) span a subspace \( \mathcal{S} \subseteq \mathcal{H}_l(\mathbb{R}) \) of dimension \( d \leq m + n \); and since any two finite-dimensional inner product spaces are isometric, \( \mathcal{S} \) is isometric to \( \mathbb{R}^d \) with the standard inner product. So we may apply Corollary 2.3 to obtain

\[
\left| \sum_{i=1}^{m} \sum_{j=1}^{n} M_{ij} \langle x_i, y_j \rangle \right| = \frac{1}{c} \left| \sum_{i=1}^{m} \sum_{j=1}^{n} M_{ij} \arcsin \langle S(x_i), T(y_j) \rangle_* \right| \leq \frac{\pi}{2c},
\]

which is Krivine’s bound since \( \pi/2c = \pi/(2 \log(1 + \sqrt{2})) = K^\mathbb{R} \).
Case II: $F = \mathbb{C}$. Let $c_0 \in (0, 1)$ be the unique constant defined in [16] such that $\varphi(c_0) = 1$. By the Taylor expansion in [15] and noting that $\text{sgn}(z)|z|^{2k+1} = \pi^k z^{k+1}$,

$$H^{-1}(c_0\langle x, y \rangle) = \text{sgn}(c_0\langle x, y \rangle) \sum_{k=0}^{\infty} b_{2k+1}c_0^{2k+1}\langle x, y \rangle^{2k+1}$$

$$= \sum_{k=0}^{\infty} b_{2k+1}c_0^{2k+1}\langle x, y \rangle^{k+1}$$

$$= \sum_{k=0}^{\infty} b_{2k+1}c_0^{2k+1}\langle x_i, y_j \rangle^{k}\langle x_i, y_j \rangle^{k+1}$$

$$= \sum_{k=0}^{\infty} b_{2k+1}c_0^{2k+1}\langle x_i \otimes y_j \otimes y_j \otimes y_j \rangle_k.$$  

(24)

For any $l \in \mathbb{N}$, let $D_l = \{x \in \mathbb{C}^l : \|x\| \leq 1\}$ be the unit ball, let $H_l(\mathbb{C})$ be as in [20], and let $S, T : D_l \to H_l(\mathbb{C})$ be nonlinear maps defined by

$$S(x) = (S_k(x))_{k=0}^{\infty}, \quad S_k(x) := \text{sgn}(b_{2k+1})|b_{2k+1}|c_0^{2k+1}\langle x \rangle^{(k)} \otimes x^{(k+1)},$$

$$T(x) = (T_k(x))_{k=0}^{\infty}, \quad T_k(x) := |b_{2k+1}|c_0^{2k+1}\langle x \rangle^{(k)} \otimes x^{(k+1)},$$

for any $x \in D_l$. Then $S$ and $T$ are maps into $H_l(\mathbb{C})$ since

$$\|S(x)\|_k^2 = \sum_{k=0}^{\infty} \|S_k(x)\|_k^2 = \sum_{k=0}^{\infty} |b_{2k+1}|c_0^{2k+1}\|x\|^{2(2k+1)} = \sum_{k=0}^{\infty} \|T_k(x)\|_k^2 = \|T(x)\|_k^2$$

and, as $b_1 > 0$ and $b_{2k+1} \leq 0$ for all $k \geq 1$ by [13],

$$\sum_{k=0}^{\infty} |b_{2k+1}|c_0^{2k+1}\|x\|^{2(2k+1)} = 2b_1c_0\|x\|^2 - H^{-1}(c_0\|x\|^2) < \infty.$$  

As in Case I, we may rewrite (24) as

$$H^{-1}(c_0\langle x, y \rangle) = \langle S(x), T(y) \rangle_{\ast} \quad \text{or} \quad c_0\langle x, y \rangle = H(\langle S(x), T(y) \rangle_{\ast}).$$

Moreover, since $x_i$ and $y_j$ are unit vectors in $\mathbb{C}^l$, we get

$$\|S(x_i)\|^2 = \sum_{k=0}^{\infty} |b_{2k+1}|c_0^{2k+1} = \varphi(c_0) = 1,$$

and similarly $\|T(y_j)\| = 1$. So we may apply Corollary 2.3 to get

$$\left| \sum_{i=1}^{m} \sum_{j=1}^{n} M_{ij}\langle x_i, y_j \rangle \right| \leq \frac{1}{c_0} \left| \sum_{i=1}^{m} \sum_{j=1}^{n} M_{ij}H(\langle S(x_i), T(y_j) \rangle_{\ast}) \right| \leq \frac{1}{c_0},$$

which is Haagerup’s bound since $1/c_0 = 8/\pi(x_0 + 1) = KC$.

5. Nonpositivity of $b_{2k+1}$

To make the proof in this article entirely self-contained, we present Haagerup’s proof of the nonpositivity of $b_{2k+1}$ that we used earlier in [13]. While the main ideas are all due to Haagerup, our small contribution here is that we avoided the use of any known results of elliptic integrals in order to stay faithful to our claim of an elementary proof, i.e., one that uses only calculus and basic complex variables. To be clear, while the functions

$$K(x) := \int_0^{\pi/2} (1 - x^2 \sin^2 t)^{-1/2} \, dt, \quad E(x) := \int_0^{\pi/2} (1 - x^2 \sin^2 t)^{1/2} \, dt$$

...
Lemma 5.1. Let $h_1, h_2 : [1, \infty) \to \mathbb{R}$ be defined by

\[
h_1(x) := \int_0^{\pi/2} \frac{\sin^2 t}{\sqrt{1 - x^{-2} \sin^2 t}} dt,
\]

\[
h_2(x) := (1 - x^{-2}) \int_0^{\pi/2} \frac{\sin^2 t}{\sqrt{1 - (1 - x^{-2}) \sin^2 t}} dt,
\]

which are clearly strictly increasing functions on $[1, \infty)$ with

\[
h_1(1) = 1, \quad \lim_{x \to \infty} h_1(x) = \pi/2, \quad h_2(1) = 0, \quad \lim_{x \to \infty} h_2(x) = \infty.
\]

Then

\[
\omega_1(x) := x(h_1(x)h_2'(x) - h_1'(x)h_2(x)) = \frac{\pi}{2} \quad \text{for } x \geq 1,
\]

\[
\omega_2(x) := x(h_1(x)h_1'(x) + h_2(x)h_2'(x)) \geq 2h_1(\sqrt{2})h_2(\sqrt{2}) > \frac{\pi}{4} \quad \text{for } 1 \leq x \leq \sqrt{2}.
\]

Proof. We start by observing some properties of $h_1'$ and $h_2'$. As

\[
h_1'(x) = \frac{1}{x^3} \int_0^{\pi/2} \frac{\sin^2 t}{\sqrt{1 - x^{-2} \sin^2 t}} dt = \frac{1}{x^2} \int_0^{\pi/2} \frac{\sin^2 t}{\sqrt{x^2 - \sin^2 t}} dt,
\]

$h_1'$ is strictly decreasing on $(1, \infty)$. As $\int_0^{\pi/2} \cos^{-1} t dt = \infty$, $\lim_{x \to 1^+} h_1'(x) = \infty$. Clearly $\lim_{x \to \infty} h_1'(x) = 0$. Furthermore, when $x > 1$, since $\sqrt{x^2 - \sin^2 t} \geq \sqrt{x^2 - 1}$, we have

\[
0 < h_1'(x) \leq \frac{\pi}{4x^2\sqrt{x^2 - 1}} \quad \text{for } x > 1.
\]

It is straightforward to see that the functions $E$ and $K$ in (25) have derivatives given by

\[
E'(y) = \frac{1}{y}(E(y) - K(y)), \quad K'(y) = \frac{1}{y(1 - y^2)}(E(y) - (1 - y^2)K(y)).
\]

Clearly, $h_2(x) = K(y) - E(y)$, where $y = y(x) = \sqrt{1 - x^{-2}}$. So by chain rule,

\[
h_2'(x) = y'(x) \frac{d}{dy}(K - E)(y(x)) = \frac{1}{x} \int_0^{\pi/2} (1 - (1 - x^{-2}) \sin^2 t)^{1/2} dt.
\]

Hence $h_2'$ is strictly decreasing on $[1, \infty)$, $h_2'(1) = \pi/2$, and $\lim_{x \to \infty} h_2'(x) = 0$.

To show (26), observe that

\[
h_1(x) = E(1/x), \quad xh_1'(x) = K(1/x) - E(1/x), \quad h_2(x) = K(y) - E(y), \quad xh_2'(x) = E(y),
\]

where again $y = \sqrt{1 - x^{-2}}$. Hence

\[
\omega_1(x) = E(1/x)E(y) - [K(1/x) - E(1/x)][K(y) - E(y)]
\]

\[
= E(1/x)K(y) + K(1/x)E(y) - K(1/x)K(y).
\]

Computing $\omega_1'$, we see from (29) that $\omega_1' \equiv 0$. So $\omega_1$ is a constant function. By (28), $\lim_{x \to 1} h_1'(x)(1 - x^{-2}) = 0$, and so $\lim_{x \to 1} \omega_1(x) = \pi/2$. Thus $\omega_1(x) = \pi/2$ for all $x > 1$ and we may set $\omega_1(1) = \pi/2$. 


We now show (3) following Haagerup’s arguments. Note that 
\[ \omega_2(x) = E(1/x)(K(1/x) - E(1/x)) + E(y)(K(y) - E(y)). \]
Let \( g(x) := E(\sqrt{x})(K(\sqrt{x}) - E(\sqrt{x})) \). A straightforward calculation using (29) shows that 
\[ g''(x) = \frac{1}{2} \left[ \frac{E(\sqrt{x}) - K(\sqrt{x})}{x} \right]^2 \geq 0, \quad x \in [0, 1]. \]
So \( g \) is convex on \([0, 1] \). Hence \( g(1-x) \) is also convex on \([0, 1] \). Let \( f(x) := g(x) + g(1-x) \). Then \( f \) is convex on \([0, 1] \) and \( f'(1/2) = 0 \). Therefore \( f(x) \geq f(1/2) \geq 2g(1/2) \). This yields the first inequality in (27): \( \omega_2(x) \geq 2h_1(\sqrt{2})h_2(\sqrt{2}) \) for \( x \in [1, \sqrt{2}] \).

The Taylor expansions of \( h_1 \) and \( h_2 \) may be obtained as that in (11),

\begin{align*}
(30) & \quad h_1(x) = \frac{\pi}{2} \sum_{k=0}^{\infty} \frac{(2k)!}{2^{2k}(k!)^2} \frac{1}{1-2k} x^{-2k}, \\
(31) & \quad h_2(x) = \frac{\pi}{2} \sum_{k=0}^{\infty} \frac{(2k)!}{2^{2k}(k!)^2} \frac{2k}{2k-1} (1-x^{-2})^k.
\end{align*}

Approximate numerical values of \( h_1 \) and \( h_2 \) at \( x = \sqrt{2} \) and 4 are calculated\(^{\text{5}}\) to be:

\begin{align*}
(32) & \quad h_1(\sqrt{2}) \approx 1.3506438, \quad h_2(\sqrt{2}) \approx 0.5034307, \quad h_1(4) \approx 1.5459572, \quad h_2(4) \approx 1.7289033.
\end{align*}

The second inequality in (27) then follows from \( 2h_1(\sqrt{2})h_2(\sqrt{2}) \approx 2 \times 1.35064 \times 0.50343 > \pi/4 \). \( \square \)

In the next two lemmas and their proofs, \( \text{Arg} \) will denote principal argument.

**Lemma 5.2.** Let \( h : [-1, 1] \rightarrow [-1, 1] \) be the real Haagerup function as defined in Section 3. Then \( h \) can be extended to a function \( h_+ : \overline{\mathbb{H}} \rightarrow \mathbb{C} \) that is continuous on the closed upper half-plane \( \overline{\mathbb{H}} = \{ z \in \mathbb{C} : \text{Im}(z) \geq 0 \} \) and analytic on the upper half-plane \( \mathbb{H} = \{ z \in \mathbb{C} : \text{Im}(z) > 0 \} \). In addition, \( h_+ \) has the following properties:

(i) \( \text{Im}(h_+(z)) \geq \text{Im}(h_+(|z|)) \) for all \( z \in \mathbb{H} \cap \{ z \in \mathbb{C} : \text{Re}(z) \geq 1 \} \) and \( h_+(z) \neq 0 \) for all \( z \in \mathbb{H} \\setminus \{0\} \).

(ii) For \( x \in [1, \infty) \), 
\[ \text{Re}(h_+(x)) = h_1(x), \quad \text{Im}(h_+(x)) = h_2(x), \]
where \( h_1, h_2 \) are as defined in Lemma 5.1.

(iii) For all \( k \in \mathbb{N} \) and all real \( \alpha > 1 \),
\[ b_{2k+1} = \frac{2}{\pi(2k+1)} \int_{1}^{\alpha} \text{Im}(h_+(x))^{-2k+1} dx + r_k(\alpha) \]
where 
\[ |r_k(\alpha)| \leq \frac{\alpha}{2k+1} \left( \text{Im}(h_+(\alpha)) \right)^{-(2k+1)}. \]

**Proof.** Integrating by parts, we obtain 
\[ h(x) = \int_{0}^{\pi/2} \cos t \cdot d(\text{arcsin}(x \sin t)) = \int_{0}^{\pi/2} \sin t \arcsin(x \sin t) \, dt, \quad x \in [-1, 1]. \]

The analytic function \( \sin z \) is a bijection of \( [-\pi/2, \pi/2] \times [0, \infty) \) onto \( \overline{\mathbb{H}} \) and it maps the line segment \( \{ t + ia : -\pi/2 \leq t \leq \pi/2 \} \) onto the half ellipsoid \( \{ z \in \overline{\mathbb{H}} : |z-1| + |z+1| = 2 \cosh a \} \). Let \( \text{arcsin}_+ \)

\(^{5}\)For example, using [http://www.wolframalpha.com](http://www.wolframalpha.com), which is freely available. Such numerical calculations cannot be completely avoided — Haagerup’s proof implicitly contains them as he used tabulated values of elliptic integrals.
be the inverse of this mapping. Then $\arcsin_+$ is continuous in $\mathbb{H}$ and analytic in $\mathbb{H}$. In addition, we have:

$$
\arcsin_+ x = \begin{cases} 
\arcsin x & \text{if } x \in [-1, 1], \\
\frac{\pi}{2} \text{sgn } x + i \text{arccosh } |x| & \text{if } x \in (-\infty, -1) \cup (1, \infty),
\end{cases}
$$

Therefore for $z \in \mathbb{H} \cap \{z \in \mathbb{C} : |z| \geq 1\}$,

$$
\text{Im}(\arcsin_+ z) = \text{arccosh}\left(\frac{1}{2}(|z - 1| + |z + 1|)\right), \quad z \in \mathbb{H}.
$$

If we define

$$
h_+(z) := \int_0^{\pi/2} \sin t \arcsin_+(z \sin t) \, dt, \quad z \in \mathbb{H},
$$

then $h_+$ is a continuous extension of $h$ to $\mathbb{H}$ and is analytic in $\mathbb{H}$.

(i) Since $\text{arccosh}$ is increasing on $[1, \infty)$, we have

$$
\text{Im}(\arcsin_+ z) = \text{arccosh}\left(\frac{1}{2}(|z - 1| + |z + 1|)\right) \geq \begin{cases} 
\text{arccosh } |z| & \text{if } |z| \geq 1, \\
0 & \text{if } |z| < 1.
\end{cases}
$$

Therefore for $z \in \mathbb{H} \cap \{z \in \mathbb{C} : |z| \geq 1\}$,

$$
\text{Im}(h_+(z)) = \int_0^{\pi/2} \sin t \cdot \text{Im}(\arcsin_+(z \sin t)) \, dt \\
\geq \int_0^{\pi/2} \text{arccosh}\left(|z| \sin t\right) \, dt = \text{Im}(h_+(|z|)).
$$

As $\text{Im}(\arcsin_+ z) > 0$ on $\mathbb{H}$, we have $\text{Im}(h_+(z)) > 0$ on $\mathbb{H}$. For $x \in [-1, 1]$, $h_+(x) = h(x)$ is zero only at $x = 0$. For $x \in (-\infty, -1) \cup (1, \infty)$,

$$
\text{Im}(h_+(x)) = \int_0^{\pi/2} \sin t \text{arccosh}(|x| \sin t) \, dt > 0.
$$

Hence $h_+$ has no zero in $\mathbb{H} \setminus \{0\}$.

(ii) Let $x \in (1, \infty)$. Integrating by parts followed by a change-of-variables $\sin u = x \sin t$ in the next-to-last equality gives us:

$$
\text{Re}(h_+(x)) = \int_0^{\arcsin(1/x)} \sin t \arcsin(x \sin t) \, dt + \frac{\pi}{2} \int_{\arcsin(1/x)}^{\pi/2} \sin t \, dt \\
= x \int_0^{\arcsin(1/x)} \frac{\cos^2 t}{\sqrt{1 - x^2 \sin^2 t}} \, dt = \int_0^{\pi/2} \sqrt{1 - x^{-2} \sin^2 u} \, du = h_1(x).
$$

A change-of-variables $\sin v = (1 - x^{-2})^{-1/2} \cos t$ in the next-to-last equality gives us:

$$
\text{Im}(h_+(x)) = \int_0^{\pi/2} \sin t \text{arccosh}(x \sin t) \, dt = x \int_0^{\pi/2} \frac{\cos^2 t}{\sqrt{x^2 \sin^2 t - 1}} \, dt \\
= (1 - x^{-2}) \int_0^{\pi/2} \frac{\sin^2 v}{\sqrt{1 - (1 - x^{-2}) \sin^2 v}} \, dv = h_2(x).
$$

(iii) The power series (11) shows that $h$ defines an analytic function $h(z)$ in the open unit disk that is identically equal to $h_+(z)$ on $\{z \in \mathbb{C} : |z| < 1\} \cap \mathbb{H}$. Since $h(0) = 0$ and $h'(0) \neq 0$, we can find some $\delta_0 \in (0, 1]$ such that $h(z)$ has an analytic inverse function (12) in $\{z \in \mathbb{C} : |z| < \delta_0\}$. For $0 < \delta < \delta_0$, let $C_\delta$ be a counterclockwise orientated circle with radius $\delta$. It follows
that \( h(C'_k) \) is a simple closed curve with winding number +1. Integrating by parts with a change-of-variables, we have

\[
b_{2k+1} = \frac{1}{2\pi i} \int_{h(C'_k)} h^{-1}(z) \frac{dz}{z^{2k+2}} = \frac{1}{2\pi} \int_{C'_k} \frac{z}{h(z)^{2k+2}} h'(z) \, dz.
\]

Note that \( b_{2k+1} \in \mathbb{R} \) and

\[
-(2k+1) \int_{C'_k} \frac{z h'(z)}{h(z)^{2k+2}} \, dz + \int_{C'_k} \frac{1}{h(z)^{2k+1}} \, dz = \int_{C'_k} \frac{d}{dz} \left[ \frac{z}{h(z)^{2k+1}} \right] \, dz = 0.
\]

Then we get

\[
b_{2k+1} = \frac{1}{2\pi(2k+1)} \int_{C'_k} h(z)^{-(2k+1)} \, dz = \frac{1}{2\pi(2k+1)} \int_{C'_k} \operatorname{Im}(h(z)^{-(2k+1)}) \, dz
\]

where \( C'_k \) is the quarter circle \( \{\delta e^{i\theta} : 0 \leq \theta \leq \pi/2\} \). Since \( h(z) \) identically equals \( h_+(z) \) on \( C'_k \) and \( h_+(z) \) has no zeros in the set \( \{z \in \mathbb{C} : \delta \leq |z| \leq \alpha, \ 0 \leq \operatorname{Arg} z \leq \pi/2\} \) by [1], Cauchy’s integral formula yields

\[
b_{2k+1} = \frac{2}{\pi(2k+1)} \operatorname{Im} \left[ \int_{C'_k} h_+(z)^{-(2k+1)} \, dz + \int_{C'_k} h_+(z)^{-(2k+1)} \, dz + \int_{\alpha}^{i\delta} h_+(z)^{-(2k+1)} \, dz \right].
\]

Moreover, since \( h_+(z) \) is real on \([\delta, 1]\) and its real part vanishes on the imaginary axis, we are left with

\[
b_{2k+1} = \frac{2}{\pi(2k+1)} \int_{1}^{\alpha} \operatorname{Im}(h_+(z)^{-(2k+1)}) \, dz + \frac{2}{\pi(2k+1)} \operatorname{Im} \left[ \int_{C'_k} h_+(z)^{-(2k+1)} \, dz \right].
\]

By [1], \( h_+(z) \geq \operatorname{Im}(h_+(z)) \geq \operatorname{Im}(h_+(|z|)) \). Thus

\[
\left| \int_{C'_k} h_+(z)^{-(2k+1)} \, dz \right| \leq \frac{\pi \alpha}{2} \left( \operatorname{Im}(h_+(\alpha)) \right)^{-(2k+1)}.
\]

\( \square \)

The integral expression of \( b_{2k+1} \) in (33) will be an important ingredient in the proof that \( b_{2k+1} \leq 0 \) for \( k \geq 1 \). We establish some further approximations for this integral in the next and final lemma.

**Lemma 5.3.** Let \( \alpha = 4 \) throughout⁶ Let \( \theta(x) := \operatorname{Arg}(h_+(x)) \) for \( x \in [1, \infty) \). Then \( \theta : [1, \infty) \to [0, 2\pi] \) is strictly increasing on \( x \geq 1 \), \( \theta(1) = 0 \), and \( \lim_{x \to \infty} \theta(x) = \pi/2 \). In addition, we have the following:

(i) Let \( p := \lfloor (2k+1)\theta(\alpha)/\pi \rfloor \). Let

\[
I_r := \frac{2}{\pi(2k+1)} \int_{\theta(x) = \pi r/(2k+1)}^{\pi r + (2k+1)/(2k+1)} |h_+(x)|^{-(2k+1)} \sin((2k+1)\theta(x)) \, dx
\]

for \( r = 1, 2, \ldots, p \), and

\[
J := \frac{2}{\pi(2k+1)} \int_{\theta(x) = \pi p/(2k+1)}^{\pi} |h_+(x)|^{-(2k+1)} \sin((2k+1)\theta(x)) \, dx.
\]

Then

\[
\frac{2}{\pi(2k+1)} \int_{1}^{\alpha} \operatorname{Im}(h_+(x)^{-(2k+1)}) \, dx = -I_1 + I_2 - \ldots + (-1)^p I_p + (-1)^{p+1} J.
\]

(ii) Let \( k \geq 4 \). Then \( p \geq 2 \) and \( I_1 > I_2 > \cdots > I_p > J \).

---

⁶To avoid confusion, we write ‘\( \alpha \)’ for the upper limit of our integrals instead of ‘4’ as the same number will also appear in an unrelated context ‘\( k \geq 4 \).’
(iii) Let \( k \geq 4 \) and \( c = |h_+(\sqrt{2})|e^{-\theta(\sqrt{2})/2} \). Then \( I_1 > 0.57e^{-(2k+1)/(2k+1)^2} \) and \( I_2 < 0.85I_1 \).

**Proof.** Since \( \theta(x) = \arctan(h_2(x)/h_1(x)) \), by \( (26) \), we get

\[
\frac{d\theta(x)}{dx} = \frac{h_1(x)h_2'(x) - h_1'(x)h_2(x)}{|h_+(x)|^2} > 0, \quad x > 1.
\]

So \( \theta(x) \) is strictly increasing on for \( x \geq 1 \). It is clear that \( \theta(1) = 0 \). By Lemma 5.1, \( \lim_{x \to \infty} h_1(x) = \pi/2 \) and \( \lim_{x \to \infty} h_2(x) = +\infty \), so \( \lim_{x \to \infty} \theta(x) = \pi/2 \).

(i) This follows from dividing the interval of the integral \([1, \alpha]\) into \( p+1 \) subsets:

\[
\frac{2}{\pi(2k+1)} \int_1^\alpha \text{Im}(h_+(x)^{(2k+1)}) \, dx = -\frac{2}{\pi(2k+1)} \int_1^\alpha |h_+(x)|^{-(2k+1)} \sin((2k+1)\theta(x)) \, dx
\]

\[
= -I_1 + I_2 - \ldots + (-1)^p I_p + (-1)^{p+1}J.
\]

(ii) We write \( x = x(\theta) \), \( \theta \in [0, \pi/2] \), for the inverse function of \( \theta = \theta(x) \). By \( (35) \), we have

\[
I_r = \frac{4}{\pi^2(2k+1)} \int_{\pi(r-1)/(2k+1)}^{\pi r/(2k+1)} x(\theta)|h_+(x(\theta))|^{-(2k+1)}|\sin((2k+1)\theta)| \, d\theta,
\]

\[
J = \frac{4}{\pi^2(2k+1)} \int_{\pi p/(2k+1)}^{\theta(\alpha)} x(\theta)|h_+(x(\theta))|^{-(2k+1)}|\sin((2k+1)\theta)| \, d\theta.
\]

By Lemma 5.1, \( h_1(x) \) and \( h_2(x) \) are strictly increasing function of \( x \in [1, \infty) \), therefore, so is \( |h_+(x)|^2 = h_1(x)^2 + h_2(x)^2 \). With this, we deduce that \( x|h_+(x)|^{2k+1} \) is strictly decreasing on \([1, \alpha]\) for \( k \geq 4 \) as

\[
\frac{d}{dx}(x|h_+(x)|^{-2k+1}) = |h_+(x)|^{-2k+1} + \frac{(-2k+1)x}{2} |h_+(x)|^{-2k-1} \frac{d}{dx}|h_+(x)|^2
\]

\[
= |h_+(x)|^{-2k-1} \left( |h_+(x)|^2 - (2k-1)x(h_1(x)h_1'(x) + h_2(x)h_2'(x)) \right)
\]

\[
\leq |h_+(x)|^{-2k-1} \left( |h_+(x)|^2 - \frac{7\pi}{4} \right)
\]

\[
\leq |h_+(x)|^{-2k-1} \left( |h_+(\alpha)|^2 - \frac{7\pi}{4} \right) \approx -0.1187 < 0,
\]

where we have used the fact that \( |h_+(x)|^2 \) is increasing on \([1, \alpha]\) in the next-to-last inequality and the numerical value is calculated from those of \( h_1(4) \) and \( h_2(4) \) in \((32)\). Since \( |\sin((2k+1)\theta)| \) is periodic with period \( \pi/(2k+1) \), we obtain \( I_1 > I_2 > \ldots > I_p \). In addition,

\[
J = \frac{4}{\pi^2(2k+1)} \int_{\pi p/(2k+1)}^{\theta(\alpha)} x(\theta)|h_+(x(\theta))|^{-(2k+1)}|\sin((2k+1)\theta)| \, d\theta
\]

\[
\leq \frac{4}{\pi^2(2k+1)} \int_{(p-1)\pi/(2k+1)}^{\theta(\alpha)-\pi/(2k+1)} x(\theta)|h_+(x(\theta))|^{-(2k+1)}|\sin((2k+1)\theta)| \, d\theta < I_p.
\]

Finally, we have \( \theta(\alpha) = \arctan(h_1(\alpha)/h_2(\alpha)) \approx 0.8412 > \pi/4 = \arctan(1) \), and so \( p = \lfloor(2k+1)\theta(\alpha)/\pi\rfloor \geq \lfloor9\theta(\alpha)/\pi\rfloor = 2 \) for \( k \geq 4 \).

(iii) Since \( x(\theta) \geq 1 \) for \( \theta \in [0, \pi/2] \), we have

\[
I_1 \geq \frac{4}{\pi^2(2k+1)} \int_0^{\pi/(2k+1)} x(\theta)|h_+(x(\theta))|^{-(2k+1)}|\sin((2k+1)\theta)| \, d\theta.
\]

Recall that \( \theta = \theta(x) \) and \( x = x(\theta) \) are inverse functions of one another. For \( \theta \in [0, \theta(\sqrt{2})] \),

\[
\frac{d}{d\theta} \log |h_+(x(\theta))| = \frac{1}{2} \frac{d}{dx} \log |h_+(x)|^2 \cdot \left( \frac{d\theta}{dx} \right)^{-1}
\]

\[
= \frac{h_1(x)h_1'(x) + h_2(x)h_2'(x)}{h_1(x)h_2'(x) - h_1'(x)h_2(x)} = \frac{\omega_2(x)}{\omega_1(x)} > \frac{1}{2},
\]
for \( x \in [1, \sqrt{2}] \), where we have used (26), (27), and the fact that \( \theta(x) \) is strictly increasing for \( x \geq 1 \). Hence \( \log |h_+(x(\theta))| \leq \log |h_+(\sqrt{2})| - (\theta(\sqrt{2}) - \theta)/2 \) which is equivalent to

\[
|h_+(x(\theta))| \leq ce^{\theta/2}, \quad \theta \in [0, \theta(\sqrt{2})]
\]

where \( c = |h_+(\sqrt{2})| e^{-\theta(\sqrt{2})/2} \approx 1.2059 \) and \( \theta(\sqrt{2}) > \pi/9 \), using values of \( h_1(\sqrt{2}) \) and \( h_2(\sqrt{2}) \) in (32).

It follows that for \( k \geq 4 \), we have

\[
I_1 \geq \frac{4}{\pi^2(2k+1)^2} \int_0^{\pi/(2k+1)} (ce^{\theta/2})^{-2k+1} \sin((2k+1)\theta) \, d\theta
\]

\[
= \frac{4c^{-2k+1}}{\pi^2(2k+1)^2} \int_0^{\pi} e^{-(k-1)\theta/(2k+1)} \sin \theta \, d\theta \geq \frac{4c^{-2k+1}}{\pi^2(2k+1)^2} \int_0^{\pi} e^{-\theta/2} \sin \theta \, d\theta
\]

\[
= \left( \frac{2c}{\pi} \right)^{2} \left( \frac{1}{1+1/4} \right) \frac{c^{-2k+1}}{(2k+1)^2} > 0.57 c^{-2k+1}.
\]

Since \( \frac{d}{d\theta} \log |h_+(x(\theta))| \geq 1/2 \), we get

\[
|h_+(x(\theta + \pi/(2k+1)))|^{-2k+1} \leq e^{-(k-1/2)\pi/(2k+1)} |h_+(x(\theta))|^{-2k+1}
\]

Moreover, since \( \theta(5/\sqrt{3}) > 2\pi/9 \), we know that \( x(\theta) \leq 5/\sqrt{3} \) on \([0, 2\pi/9]\). Hence for \( k \geq 4 \), it follows from the above results that

\[
I_r = \frac{4}{\pi^2(2k+1)^2} \frac{5}{\sqrt{3}} \int_0^{2\pi/(2k+1)} x(\theta) |h_+(x(\theta))|^{-2k+1} \sin((2k+1)\theta) \, d\theta
\]

\[
= \frac{4}{\pi^2(2k+1)^2} \frac{5}{\sqrt{3}} \int_0^{\pi/(2k+1)} x(\theta) |h_+(x(\theta + \pi/(2k+1)))|^{-2k+1} \sin((2k+1)\theta) \, d\theta
\]

\[
\leq \frac{4}{\pi^2(2k+1)^2} \frac{5}{\sqrt{3}} e^{-(k-1/2)\pi/(2k+1)} \int_0^{\pi/(2k+1)} x(\theta) |h_+(x(\theta))|^{-2k+1} \sin((2k+1)\theta) \, d\theta
\]

\[
\leq \frac{5}{\sqrt{3}} e^{-7\pi/18} I_1 < 0.85 I_1.
\]

The fact that \( x|h_+(x)|^{-2k+1} \) is strictly decreasing on \([1, 4]\) for \( k \geq 4 \), established in the proof of (ii) above, is a crucial observation for establishing the nonpositivity of \( b_{2k+1} \) for \( k \geq 4 \). Observe that since \( |h_+(x)| \) is strictly increasing for \( x > 1 \), it is enough to show that \( x|h_+(x)|^{-7} \) is strictly decreasing on \([1, 4]\), which is what we did. Note that for a fixed \( k \geq 1 \), \( x|h_+(x)|^{-2k+1} \) is increasing for large enough \( x \), as \( |h_+(x)| \) behaves like \( C \log x \) for \( x \gg 1 \).

**Theorem 5.4.** Let the Taylor expansion of \( h^{-1}(x) \) be as in (12). Then \( b_{2k+1} \leq 0 \) for \( k \geq 1 \).

**Proof.** Let \( k \geq 4 \) and let \( I_1, I_2, \ldots, I_p, J \) be as defined in Lemma 5.3. By (33) with \( \alpha = 4 \) and Lemma 5.3(ii) and (ii), we have

\[
b_{2k+1} = I_1 - I_2 + \ldots + (-1)^{p-1} I_p + (-1)^p J - r_{2k+1}(5\sqrt{2}) > I_1 - I_2 - r_{2k+1}(5\sqrt{2}).
\]

By (34) and Lemma 5.3(iii) with \( c \approx 1.2059 \) (established in its proof), we get

\[
I_1 - I_2 > \frac{0.0855}{(2k+1)^2} (1.206)^{-2k+1}, \quad |r_{2k+1}(4)| \leq \frac{4}{2k+1} (1.728)^{-2k+1}.
\]

Since \( -b_{2k+1} > I_1 - I_2 - r_{2k+1}(4) \), we get \( b_{2k+1} \leq 0 \) for \( k \geq 9 \). Direct computation using the Lagrange inversion formula gives us \( b_3, b_5, \ldots, b_{17} \leq 0 \), proving nonpositivity for \( k \leq 8 \). \( \square \)
Acknowledgment

We thank the anonymous referee for his careful reading and many helpful suggestions that greatly improved our article. The work in this article is generously supported by DARPA D15AP00109 and NSF IIS 1546413. LHL acknowledges additional support from a DARPA Director’s Fellowship and the Eckhardt Faculty Fund.

References