On Exponential Approximation to the Hockey Stick Function*

Ian Iscoe† Ken Jackson‡ Alex Kreinin§ Xiaofang Ma¶

January 24, 2007

Abstract

The hockey stick function is a basic function in pricing and risk management of many financial derivatives. This paper considers approximating the hockey stick function by a sum of exponentials. The algorithm proposed by Beylkin and Monzón[1] is used to determine the parameters of an approximation. Theoretical properties of the approximation are studied. Numerical results are presented.

*This research was supported in part by the Natural Sciences and Engineering Research Council (NSERC) of Canada.

†Algorithmics Inc., 185 Spadina Avenue, Toronto, ON, M5T 2C6, Canada; Ian.Iscoe@algorithmics.com

‡Department of Computer Science, University of Toronto, 10 King's College Rd, Toronto, ON, M5S 3G4, Canada; krj@cs.utoronto.ca

§Algorithmics Inc., 185 Spadina Avenue, Toronto, ON, M5T 2C6, Canada; Alex.Kreinin@algorithmics.com

¶Department of Computer Science, University of Toronto, 10 King's College Rd, Toronto, ON, M5S 3G4, Canada; csxfma@cs.utoronto.ca

1 Introduction

In this paper we describe how to approximate the function

$$h(x) = \begin{cases} 1 - x & \text{if } 0 \le x \le 1; \\ 0 & \text{if } x > 1, \end{cases} \tag{1}$$

by a sum of exponentials

$$h_{\exp}(x) = \sum_{n=1}^{N} \omega_n \exp(\gamma_n x) \tag{2}$$

over $[0, \infty)$, where ω_n and γ_n are complex numbers. This function is a special case of the function

$$h(x;t) = \begin{cases} t - x & \text{if } 0 \le x \le t; \\ 0 & \text{if } x > t, \end{cases}$$

where t is a positive number. This function plays a critical role in finance, from pricing European options [7] to pricing and risk management of correlation-dependent derivatives [8]. Since, for a fixed positive t, $h(x;t) = t \cdot h(x/t)$, we can take $h(x)$ as the basic function. In this paper we call function $h(x)$ the hockey stick (HS) function.

The approximation problem considered here is in the sense of Chebyshev approximation. For such an approximation, the weights ω_n and the exponents γ_n should be chosen to solve the minimization problem

$$\min_{\omega_n, \gamma_n \in \mathbb{C}} \left\| h(x) - \sum_{n=1}^{N} \omega_n \exp(\gamma_n x) \right\|_\infty , \tag{3}$$

where $\mathbb{C}$ denotes the set of complex numbers, and $\|f\|_\infty = \sup_{x \in \mathbb{X}} |f(x)|$ is the Chebyshev norm (also known as the uniform norm or the sup-norm) of f. Theoretically, the existence of such an optimal approximation is generally not guaranteed [2, Chapters VI and VII]. Numerically, exponential fitting problems are badly-conditioned [4]. Consequently, classical optimization methods, such as Newton type methods [3], do not work well for the minimization problem (3). Fortunately, we can find numerical approximations satisfying some accuracy requirements. In this paper, we apply the method proposed by Beylkin and Monzón[1] to solve (3) approximately .

The remainder of the paper is organized as follows. Beylkin and Monzón's method and its application to the HS function are discussed in Section 2. Properties of the exponential approximation are discussed in Section 3. The paper ends with numerical results.

2 Beylkin and Monzón's method and its application to the HS function

2.1 Beylkin and Monzón's method

In a recent paper [1], Beylkin and Monzón proposed a numerical method to find a good exponential approximation to a function f. Instead of finding optimal ω_n and γ_n satisfying (3), their method finds such parameters so that the exponential approximation satisfies a given accuracy requirement. More specifically, for a given function f defined on $[0, 1]$ and a given ϵ, their method seeks the minimal (or nearly minimal) number of complex weights ω_n and nodes $\exp(\gamma_n)$ such that

$$\left| f(x) - \sum_{n=1}^{N} \omega_n \exp(\gamma_n x) \right| \leq \epsilon, \qquad \forall x \in [0, 1]. \tag{4}$$

This continuous problem is in turn approximated by a discrete problem: Given a positive integer $\mathcal{M}$, find the minimal positive integer number $N \leq \mathcal{M}$ of complex weights ω_n and complex nodes ζ_n such that

$$\left| f\left(\frac{m}{2\mathcal{M}}\right) - \sum_{n=1}^{N} \omega_n \zeta_n^m \right| \leq \epsilon, \qquad \text{for all integers } m \in [0, 2\mathcal{M}]. \tag{5}$$

Then for the continuous problem the weights and the exponents are ω_n and

$$\gamma_n = 2\mathcal{M} \log \zeta_n, \tag{6}$$

respectively, where $\log z$ is the principal value of the logarithm.

To describe their method, we introduce some additional notation. For theoretical background and a more detailed description of the method, see [1].

For a real $(2\mathcal{M}+1)$-vector $\mathbf{h} = (h_0, h_1, \ldots, h_{2\mathcal{M}})$, the $(\mathcal{M}+1) \times (\mathcal{M}+1)$ Hankel matrix $\mathbf{H_h}$ defined in terms of $\mathbf{h}$ is

$$\mathbf{H_h} = \begin{bmatrix} h_0 & h_1 & \cdots & h_{\mathcal{M}} \\ h_1 & \cdots & \cdots & h_{\mathcal{M}+1} \\ \vdots & & \iddots & \vdots \\ h_{\mathcal{M}-1} & h_{\mathcal{M}} & \cdots & h_{2\mathcal{M}-1} \\ h_{\mathcal{M}} & \cdots & h_{2\mathcal{M}-1} & h_{2\mathcal{M}} \end{bmatrix},$$

that is $\mathbf{H}_{i,j} = h_{i+j}$ for $0 \leq i, j \leq \mathcal{M}$. It is clear that $\mathbf{H_h}$ is a real symmetric matrix. By the Corollary in §4.4.4 of [6, pp. 204], there exists a unitary matrix $\mathbf{U}$ and a *nonnegative* diagonal matrix Σ such that

$$\mathbf{H_h} = \mathbf{U}\Sigma\mathbf{U}^T,$$

where the superscript T denotes transposition. This decomposition is called the Takagi factorization [6, pp. 204].

The main steps of the method are:

1. Sample the approximated function f at $2\mathcal{M}+1$ points uniformly distributed on $[0, 1]$. That is, let $h_m = f\left(\frac{m}{2\mathcal{M}}\right)$, $0 \leq m \leq 2\mathcal{M}$.

2. Form $\mathbf{h} = (h_0, h_1, \ldots, h_{2\mathcal{M}})$ and the Hankel matrix $\mathbf{H_h}$.

3. Compute the Takagi factorization of $\mathbf{H_h} = \mathbf{U}\Sigma\mathbf{U}^T$, where $\Sigma = \mathrm{diag}(\sigma_0, \sigma_1, \ldots, \sigma_{\mathcal{M}})$ and $\sigma_0 \geq \sigma_1 \geq \ldots \geq \sigma_{\mathcal{M}} \geq 0$.

4. Find the largest σ_N satisfying $\sigma_N \leq \epsilon$.

5. Let $\mathbf{u} = (u_0, u_1, \ldots, u_{\mathcal{M}})^T$ be the $(N+1)$-st column of $\mathbf{U}$.

6. Find N roots of the polynomial $\sum_{m=0}^{\mathcal{M}} u_m z^m$ with the largest moduli and denote these roots by $\zeta_1, \zeta_2, \ldots, \zeta_N$.

7. Compute the N weights ω_n, $1 \le n \le N$, by solving the linear least squares problem for the overdetermined Vandermonde system

$$h_m = \sum_{n=1}^{N} \omega_n \zeta_n^m, \qquad \text{for } 0 \le m \le 2\mathcal{M}.$$

8. Compute parameters γ_n using formula (6).

Remark 1 *This algorithm works for functions defined on* $[0,1]$. *To apply it to a function* f *defined on a finite interval* $[a,b]$, $a < b$, *we could consider the function* $\hat{f}(t) = f\left(t(b-a)+a\right)$ *for* $t \in [0,1]$. *For a function defined on an infinite interval, such as* $[0,\infty)$, *the interval could first be truncated to a finite interval, say* $[a,b] \subset [0,\infty)$, *then the finite interval could be mapped to the standard interval* $[0,1]$ *and the same approximation could be applied to* $[0,\infty)\backslash[a,b]$.

Remark 2 *For a general function the number of sample points is not known in advance. Thus* $\mathcal{M}$ *should be large enough or be increased gradually until a satisfactory accuracy is achieved. All critical points of the approximated function should be sampled. For example, for the HS function* $h(x)$, *both* $x = 0$ *and* $x = 1$ *should be sampled.*

Remark 3 *In practice it is not necessary to compute* $\mathbf{H_h}$*'s Takagi factorization explicitly. From the spectral theorem for Hermitian matrices [6, pp. 171] we know that there is a* **real** *orthogonal matrix* $\mathbf{V}$ *and a* **real** *diagonal matrix* $\Lambda = \mathrm{diag}(\lambda_0, \lambda_1, \ldots, \lambda_{\mathcal{M}})$, *with* $|\lambda_i|$ *decreasing, such that* $\mathbf{H_h} = \mathbf{V}\Lambda\mathbf{V}^T$. *Noting that generally* $\mathbf{H_h}$ *is not positive semidefinite,* Λ *may have negative element(s). Thus* $\mathbf{V}\Lambda\mathbf{V}^T$ *is not necessarily the Takagi factorization of* $\mathbf{H_h}$. *However, we could construct a Takagi factorization based on its spectral decomposition in the following way. Let* $\Sigma = \mathrm{diag}(|\lambda_0|, |\lambda_1|, \ldots, |\lambda_{\mathcal{M}}|)$ *and* $\mathbf{U} = (\mathbf{u}_0, \mathbf{u}_1, \ldots, \mathbf{u}_{\mathcal{M}})$, *where* $\mathbf{u}_m = \mathbf{v}_m$ *if* $\lambda_m \ge 0$; *and* $\mathbf{u}_m = \sqrt{-1}\mathbf{v}_m$, *if* $\lambda_m < 0$. *It is easy to check that* $\mathbf{U}$ *is a unitary matrix and* $\mathbf{H_h} = \mathbf{U}\Sigma\mathbf{U}^T$.

Remark 4 *To compute* ω_n *from the linear least squares problem in Step 7, the* N *roots determined in Step 6 must be distinct. If this condition is not met,* ω_n *should be computed by a different method [1], [9]. This condition may be difficult to verify in theory. For numerical solutions, we should check its validity, as suggested by Beylkin and Monzón.*

2.2 Application to the HS function

In this subsection we apply Beylkin and Monzón's method to the hockey stick function $h(x)$. Recall that $h(x)$ is defined on $[0,\infty)$. The infinite interval is first truncated to a finite interval $[0,b]$ for a large enough b. (In fact $b=2$ is large enough as explained below.) Then $h(b\cdot t)$ is sampled at $2\mathcal{M}+1$ points:

$$h_m = h\left(bt_m\right) = \begin{cases} 1-bt_m & \text{if } bt_m \le 1 \\ 0 & \text{otherwise} \end{cases},$$

where $t_m = \frac{m}{2\mathcal{M}}$ and $0 \le m \le 2\mathcal{M}$. To guarantee the critical point $x=1$ of $h(x)$ is sampled, it suffices that $bt_m = 1$ for some m. This implies that $\frac{2\mathcal{M}}{b}$ must be an integer. The corresponding Hankel matrix $\mathbf{H_h}$ is

$$\mathbf{H_h} = \left[\begin{array}{ccccc:ccc} 1 & 1-\frac{b}{2\mathcal{M}} & 1-2\frac{b}{2\mathcal{M}} & \cdots & \frac{b}{2\mathcal{M}} & 0 & \cdots & 0 \\ 1-\frac{b}{2\mathcal{M}} & 1-2\frac{b}{2\mathcal{M}} & \cdots & \cdots & 0 & 0 & \cdots & 0 \\ 1-2\frac{b}{2\mathcal{M}} & \cdots & \cdots & \cdots & 0 & 0 & \cdots & 0 \\ \cdots & \cdots & & & & & & \\ \frac{b}{2\mathcal{M}} & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \hdashline \cdots & \cdots & & & & & & \\ 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{array}\right].$$

To keep the neat form of $\mathbf{H_h}$ it may require that $b \ge 2$. If $b < 2$, the last nonzero row of $\mathbf{H_h}$ may have more than one nonzero elements. A direct consequence of this is that the properties of the approximation discussed in Section 3 may not hold. Thus in the remainder of this paper it is assumed that $b \ge 2$. Let $\mathcal{N} = \frac{2\mathcal{M}}{b}$ and

$$\mathbf{H}_{\mathcal{N}} = \begin{bmatrix} \mathcal{N} & \mathcal{N}-1 & \mathcal{N}-2 & \cdots & 1 \\ \mathcal{N}-1 & \mathcal{N}-2 & \cdots & \cdots & 0 \\ \mathcal{N}-2 & \cdots & \cdots & \cdots & 0 \\ \cdots & \cdots & & & \\ 1 & 0 & 0 & \cdots & 0 \end{bmatrix}. \tag{7}$$

Then we have

$$\mathbf{H_h} = \frac{b}{2\mathcal{M}} \left[\begin{array}{c:c} \mathbf{H}_{\mathcal{N}} & \mathbf{0}_{12} \\ \hdashline \mathbf{0}_{12}^T & \mathbf{0}_{22} \end{array}\right]. \tag{8}$$

where $\mathbf{0}_{12}$ and $\mathbf{0}_{22}$ are zero matrices of the proper dimensions.

Let $\mathbf{U}\Sigma\mathbf{U}^T$ be a Takagi factorization of $\mathbf{H}_\mathcal{N}$, where $\Sigma = \mathrm{diag}(\sigma_{\mathcal{N}1}, \sigma_{\mathcal{N}2}, \ldots, \sigma_{\mathcal{N}\mathcal{N}})$ and $\sigma_{\mathcal{N}1} \geq \sigma_{\mathcal{N}2} \geq \cdots \geq \sigma_{\mathcal{N}\mathcal{N}} \geq 0$. Then a Takagi factorization of $\mathbf{H}_\mathbf{h}$ can be obtained by

$$\mathbf{H}_\mathbf{h} = \frac{b}{2\mathcal{M}} \begin{bmatrix} \mathbf{U} & \mathbf{0}_{12} \\ \mathbf{0}_{12}^T & \mathbf{I}_{22} \end{bmatrix} \begin{bmatrix} \Sigma & \mathbf{0}_{12} \\ \mathbf{0}_{12}^T & \mathbf{0}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{U}^T & \mathbf{0}_{12}^T \\ \mathbf{0}_{12} & \mathbf{I}_{22} \end{bmatrix}.$$

Remark 5 *Proposition 1 in Section 3 together with Theorems 2 and 3 of [1] imply that, for a given accuracy* ϵ*,* $\mathcal{M}$ *must be large enough such that* $\frac{1}{4}\frac{b}{2\mathcal{M}} \leq \epsilon$*. From this relation and noting that* $\mathcal{N} = \frac{2\mathcal{M}}{b}$*, we can see the only requirements are* $b \geq 2$*,* $\frac{2\mathcal{M}}{b}$ *is an integer and*

$$\mathcal{N} \geq \frac{1}{4\epsilon}. \tag{9}$$

Thus we choose $b = 2$ *for simplicity and* $\mathcal{N} = \mathcal{M} \geq \frac{1}{4\epsilon}$*.*

Once $\mathbf{H}_\mathcal{N}$'s Takagi factorization is computed, we take $\mathbf{u}_\mathcal{N} = (u_0, u_1, \ldots, u_{\mathcal{N}-1})^T$ to be the last column of $\mathbf{U}$. Then find the $\mathcal{N}-1$ roots $\zeta_1, \zeta_2, \ldots, \zeta_{\mathcal{N}-1}$ of the polynomial $\sum_{n=0}^{\mathcal{N}-1} u_n z^n = 0$. Next the $\mathcal{N}-1$ weights ω_n are obtained by solving

$$h_m = \sum_{n=1}^{\mathcal{N}-1} \omega_n \zeta_n^m, \qquad \text{for } 0 \leq m \leq 2\mathcal{M},$$

in the least squares sense. Finally, parameters γ_n are obtained following formula (6).

In summary, the algorithm for determining coefficients ω_n and γ_n is (note that $b = 2$):

1. Input ϵ as given accuracy.
2. Find the smallest integer $\mathcal{N}$ such that $\mathcal{N} \geq \frac{1}{4\epsilon}$.
3. Compute the spectral decomposition of the matrix $\mathbf{H}_\mathcal{N} = \mathbf{V}\Lambda\mathbf{V}^T$.
4. Let $\mathbf{u} = (u_0, u_1, \ldots, u_{\mathcal{N}-1}^T)$ be the last column of $\mathbf{V}$.
5. Find all roots $\zeta_1, \zeta_2, \ldots, \zeta_{\mathcal{N}-1}$ of the polynomial $\sum_{n=0}^{\mathcal{N}-1} u_n z^n = 0$ and check whether they are distinct. If they are not distinct then exit.

6. Solve $h_m = \sum_{n=1}^{\mathcal{N}-1} \omega_n \zeta_n^m$, $0 \le m \le 2\mathcal{N}$, in the least squares sense for ω_n.

7. Compute $\gamma_n = 2\mathcal{N} \log \zeta_n$.

Before ending this section we want to say a little about ω_n and γ_n. As mentioned in Remark 3, $\mathbf{u}_{\mathcal{N}}$ is either a real vector or the product of a real vector and the imaginary unit $\sqrt{-1}$. In either case, the roots of $\sum_{n=0}^{\mathcal{N}-1} u_n z^n = 0$ will always be either real or pairwise conjugate. Thus ω_n are also real or pairwise complex conjugate, correspondingly. That is, if ζ_n is real, then ω_n is real too; whereas, if ζ_i and ζ_j are conjugate, then ω_i and ω_j are conjugate too, and vice versa. Furthermore, since $\gamma_n = 2\mathcal{N} \log \zeta_n$ we can see that $\exp(\gamma_n) = \zeta_n^{2\mathcal{N}}$ possesses the same conjugacy property. Thus $\omega_n \exp(\gamma_n x)$ are either real or pairwise conjugate for all real x. This result simplifies the calculation of $h_{\exp}(x) = \sum_{n=1}^{\mathcal{N}-1} \omega_n \exp(\gamma_n x)$. For real ω_n the term $\omega_n \exp(\gamma_n x)$ is evaluated as usual, whereas for the conjugate pair indexed by i and j, only one term needs to be evaluated, say $\omega_i \exp(\gamma_i x)$, and then the contribution of the complex conjugate pair of terms is $2\Re(\omega_i \exp(\gamma_i x))$, where $\Re(z)$ denotes the real part of the complex number z.

3 Properties of the approximation

In this section we discuss some properties related to this approximation. Noting that the diagonal matrix Σ is the same as the diagonal matrix of $\mathbf{H}_{\mathcal{N}}$'s singular value decomposition, we call $\sigma_{\mathcal{N}n}$, $n = 1, 2, \ldots, \mathcal{N}$, its singular value. Direct calculation shows that

$$\mathbf{H}_{\mathcal{N}}^{-1} = \begin{bmatrix} & & & & & & 1 \\ & & & & & 1 & -2 \\ & & & & 1 & -2 & 1 \\ & & & \iddots & \iddots & \iddots & \\ & & \iddots & \iddots & \iddots & & \\ & 1 & -2 & 1 & & & \\ 1 & -2 & 1 & & & & \end{bmatrix} \tag{10}$$

Proposition 1 *As $\mathcal{N}$ tends to infinity, the smallest singular value $\sigma_{\mathcal{N}\mathcal{N}}$ of the matrix $\mathbf{H}_\mathcal{N}$ tends to* $1/4$*.*

Proof Since $\mathbf{H}_\mathcal{N}$ is nonsingular, its singular values are positive. Proving that $\sigma_{\mathcal{N}\mathcal{N}}$ tends to $1/4$ as $\mathcal{N}$ tends to infinity is equivalent to proving that $\sigma_{\mathcal{N}\mathcal{N}}^{-1}$, the largest singular value of $\mathbf{H}_\mathcal{N}^{-1}$, tends to 4 as $\mathcal{N}$ tends to infinity.

From Gerschgorin's theorem [5, pp. 320] [6, pp. 344] [12, pp. 71], we know that all eigenvalues of $\mathbf{H}_\mathcal{N}^{-1}$ lie in the disc

$$D = \{z \in \mathbb{C} : |z| \le 4\}.$$

Since $\mathbf{H}_\mathcal{N}^{-1}$ is real symmetric, we can conclude that all singular values of $\mathbf{H}_\mathcal{N}^{-1}$ are bounded by 4. Therefore, it suffices to prove that $\sigma_{\mathcal{N}\mathcal{N}}^{-1}$ approaches 4. To this end, note that, if $\mathbf{A}$ is a symmetric matrix, then $\max_{\mathbf{x}\neq 0} \frac{\mathbf{x}^T\mathbf{A}\mathbf{x}}{\mathbf{x}^T\mathbf{x}} = \lambda_{\max}$, where $\lambda_{\max}$ is the largest eigenvalue of $\mathbf{A}$. Hence, if for each $\mathcal{N}$, we can find a vector $\mathbf{x}_\mathcal{N}$ such that the Rayleigh quotient $\frac{\mathbf{x}_\mathcal{N}^T\mathbf{H}_\mathcal{N}^{-T}\mathbf{H}_\mathcal{N}^{-1}\mathbf{x}_\mathcal{N}}{\mathbf{x}_\mathcal{N}^T\mathbf{x}_\mathcal{N}} \to 16$ as $\mathcal{N} \to \infty$, then we can conclude that $\sigma_{\mathcal{N}\mathcal{N}}^{-1} \to 4$ as $\mathcal{N} \to \infty$.

For even $\mathcal{N} \ge 6$, let $\mathcal{N} = 2n$. Define a vector $\mathbf{x}_\mathcal{N} = (x_1, x_2, \ldots, x_\mathcal{N})^T$ by

$$\begin{aligned}
&x_1 = 1,\\
&x_i = x_{\mathcal{N}-i+2} = (-1)^{i-1}(i-1), \text{ for } i = 2, 3, \ldots, n,\\
&x_{n+1} = -x_n.
\end{aligned}$$

Through direct calculation we obtain

$$\mathbf{x}_\mathcal{N}^T\mathbf{x}_\mathcal{N} = 1 + 2\sum_{i=1}^{n-1} i^2 + (n-1)^2, \tag{11}$$

and

$$
\mathbf{H}_{\mathcal{N}}^{-1}\mathbf{x}_{\mathcal{N}} = \left[\begin{array}{c} -1 \\ 4 \\ \hdashline -8 \\ \vdots \\ (-1)^{n-1}4(n-2) \\ \hdashline (-1)^{n}(4n-5) \\ (-1)^{n+1}4(n-1) \\ (-1)^{n}(4n-5) \\ \hdashline (-1)^{n-1}4(n-2) \\ \vdots \\ -8 \\ \hdashline 5 \end{array}\right],
$$

which implies

$$
\mathbf{x}_{\mathcal{N}}^{T}\mathbf{H}_{\mathcal{N}}^{-T}\mathbf{H}_{\mathcal{N}}^{-1}\mathbf{x}_{\mathcal{N}} = 10 + 2\cdot 4^2 \sum_{i=1}^{n-2} i^2 + 4^2(n-1)^2 + 2(4n-5)^2.
$$

This result together with (11) implies that for large even $\mathcal{N}$

$$
\frac{\mathbf{x}_{\mathcal{N}}^{T}\mathbf{H}_{\mathcal{N}}^{-T}\mathbf{H}_{\mathcal{N}}^{-1}\mathbf{x}_{\mathcal{N}}}{\mathbf{x}_{\mathcal{N}}^{T}\mathbf{x}_{\mathcal{N}}} \approx \frac{2\cdot 4^2 \sum_{i=1}^{n-1} i^2}{2\sum_{i=1}^{n-1} i^2} = 16.
$$

Therefore, the largest singular value of $\mathbf{H}_{\mathcal{N}}^{-1}$ approaches 4 as $\mathcal{N}\to\infty$.

For odd $\mathcal{N}\geq 7$, let $\mathcal{N}=2n+1$. Construct an $\mathcal{N}$-vector $\mathbf{x}_{\mathcal{N}}=(x_1,x_2,\ldots,x_{\mathcal{N}})^T$ with

$$
\begin{aligned}
&x_i = x_{\mathcal{N}-i+1} = (-1)^i i, \text{ for } i=1,2,\ldots,n,\\
&x_{n+1} = -x_n.
\end{aligned}
$$

Similar to the case for even $\mathcal{N}$, we have

$$
\mathbf{x}_{\mathcal{N}}^{T}\mathbf{x}_{\mathcal{N}} = 2\sum_{i=1}^{n} i^2 + n^2, \tag{12}
$$

$$\mathbf{H}_{\mathcal{N}}^{-1}\mathbf{x}_{\mathcal{N}} = \left[\begin{array}{c} -1 \\ 4 \\ \hdashline -8 \\ \vdots \\ (-1)^n 4(n-1) \\ \hdashline (-1)^{n+1}(4n-1) \\ (-1)^n 4n \\ (-1)^{n+1}(4n-1) \\ \hdashline (-1)^n 4(n-1) \\ \vdots \\ -8 \end{array}\right],$$

$$\mathbf{x}_{\mathcal{N}}^T \mathbf{H}_{\mathcal{N}}^{-T} \mathbf{H}_{\mathcal{N}}^{-1} \mathbf{x}_{\mathcal{N}} = 17 + 2 \cdot 4^2 \sum_{i=3}^{n} (i-1)^2 + 2(4n-1)^2 + 4^2 n^2. \tag{13}$$

Thus, from (12) and (13) we have that for large odd $\mathcal{N}$

$$\frac{\mathbf{x}_{\mathcal{N}}^T \mathbf{H}_{\mathcal{N}}^{-T} \mathbf{H}_{\mathcal{N}}^{-1} \mathbf{x}_{\mathcal{N}}}{\mathbf{x}_{\mathcal{N}}^T \mathbf{x}_{\mathcal{N}}} \approx \frac{2 \cdot 4^2 \sum_{i=1}^{n} i^2}{2 \sum_{i=1}^{n} i^2} = 16.$$

The proof is completed.

As explained before, $\mathbf{u}_{\mathcal{N}}$ is either a real vector or the product of a real vector and the imaginary unit. In either case, the next three propositions hold. For simplicity we assume in the proofs that $\mathbf{u}_{\mathcal{N}}$ is a real vector.

Proposition 2 *The smallest singular value $\sigma_{\mathcal{N}\mathcal{N}}$ satisfies the relation $\sigma_{\mathcal{N}+2,\mathcal{N}+2} < \sigma_{\mathcal{N}\mathcal{N}}$, $\mathcal{N} \geq 2$.*

Proof Let $\lambda_{\mathcal{N}}$ be the eigenvalue of $\mathbf{H}_{\mathcal{N}}$ corresponding to $\sigma_{\mathcal{N}\mathcal{N}}$ and $\mathbf{u}_{\mathcal{N}} = (u_0, u_1, \ldots, u_{\mathcal{N}-1})^T \in \mathbb{R}^{\mathcal{N}}$ be a corresponding eigenvector. Thus $\mathbf{H}_{\mathcal{N}}\mathbf{u}_{\mathcal{N}} = \lambda_{\mathcal{N}}\mathbf{u}_{\mathcal{N}}$. Without loss of generality, assume $\|\mathbf{u}_{\mathcal{N}}\|_2 = 1$. Consequently, $\mathbf{u}_{\mathcal{N}}^T \mathbf{H}_{\mathcal{N}}^{-2} \mathbf{u}_{\mathcal{N}} = \sigma_{\mathcal{N}\mathcal{N}}^{-2}$.

Now we show that u_0 and u_1 cannot be zero simultaneously. Suppose $u_0 = u_1 = 0$ for some $\mathcal{N} > 2$. Note that $\mathbf{H}_{\mathcal{N}}^{-1}\mathbf{u}_{\mathcal{N}} = \lambda_{\mathcal{N}}^{-1}\mathbf{u}_{\mathcal{N}}$. That is

$$\begin{bmatrix} & & & & & & 1 \\ & & & & & 1 & -2 \\ & & & & 1 & -2 & 1 \\ & & & \iddots & \iddots & \iddots & \\ & & \iddots & \iddots & \iddots & & \\ & 1 & -2 & 1 & & & \\ 1 & -2 & 1 & & & & \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \\ u_2 \\ \vdots \\ \vdots \\ u_{\mathcal{N}-2} \\ u_{\mathcal{N}-1} \end{bmatrix} = \lambda_{\mathcal{N}}^{-1} \begin{bmatrix} u_0 \\ u_1 \\ u_2 \\ \vdots \\ \vdots \\ u_{\mathcal{N}-2} \\ u_{\mathcal{N}-1} \end{bmatrix}. \tag{14}$$

By comparing the two sides of the system of equations, we obtain $u_{\mathcal{N}-1} = 0$ from the first row; then $u_{\mathcal{N}-2} = 0$ from the second row; and then $u_2 = u_3 = 0$ from the last two rows. Continuing this process we end with $\mathbf{u}_{\mathcal{N}} = 0$, which contradicts $\mathbf{u}_{\mathcal{N}} \neq 0$.

Let $\mathbf{u}_{\mathcal{N}+2} = (0, 0, u_0, u_1, \ldots, u_{\mathcal{N}-1})^T$, then $\|\mathbf{u}_{\mathcal{N}+2}\|_2 = 1$, and

$$\mathbf{H}_{\mathcal{N}+2}^{-1}\mathbf{u}_{\mathcal{N}+2} = \begin{bmatrix} \mathbf{H}_{\mathcal{N}}^{-1}\mathbf{u}_{\mathcal{N}} \\ -2u_0 + u_1 \\ u_0 \end{bmatrix}.$$

Furthermore we have

$$\begin{aligned} \sigma_{\mathcal{N}+2,\mathcal{N}+2}^{-2} = \max_{\|\mathbf{x}\|=1} \mathbf{x}^T \mathbf{H}_{\mathcal{N}+2}^{-2} \mathbf{x}^T &\geq \mathbf{u}_{\mathcal{N}+2}^T \mathbf{H}_{\mathcal{N}+2}^{-2} \mathbf{u}_{\mathcal{N}+2} \\ &= \mathbf{u}_{\mathcal{N}}^T \mathbf{H}_{\mathcal{N}}^{-1} \mathbf{H}_{\mathcal{N}}^{-1} \mathbf{u}_{\mathcal{N}} + (u_1 - 2u_0)^2 + u_0^2 \\ &= \sigma_{\mathcal{N}\mathcal{N}}^{-2} + (u_1 - 2u_0)^2 + u_0^2 \\ &> \sigma_{\mathcal{N}\mathcal{N}}^{-2}. \end{aligned}$$

The last inequality follows from the observation above that u_0 and u_1 cannot both be zero. This completes the proof.

Proposition 3 *$u_0 \neq 0$ and $u_{\mathcal{N}-1} \neq 0$ for $\mathcal{N} \geq 4$.*

Proof From equation (14) we know that $u_0 = 0$ implies $u_{\mathcal{N}-1} = 0$, and vice versa. We will finish the proof by way of contradiction. Suppose that for some $\mathcal{N}$ we have $u_0 = 0$. Then by equation (14) we have

$$\left[\begin{array}{c|ccccc|c} & & & & & & 1 \\ \hline & & & & & 1 & -2 \\ & & & & 1 & -2 & 1 \\ & & & \iddots & \iddots & \iddots & \\ & & \iddots & \iddots & \iddots & & \\ & 1 & -2 & 1 & & & \\ \hline 1 & -2 & 1 & & & & \end{array}\right] \left[\begin{array}{c} u_0 \\ \hline u_1 \\ u_2 \\ \vdots \\ \vdots \\ u_{\mathcal{N}-2} \\ \hline u_{\mathcal{N}-1} \end{array}\right] = \lambda_{\mathcal{N}}^{-1} \left[\begin{array}{c} u_0 \\ \hline u_1 \\ u_2 \\ \vdots \\ \vdots \\ u_{\mathcal{N}-2} \\ \hline u_{\mathcal{N}-1} \end{array}\right].$$

Since $u_0 = u_{\mathcal{N}-1} = 0$, deleting the first and the last rows (columns) of the matrix and correspondingly the first and the last elements of $\mathbf{u}_{\mathcal{N}}$, *i.e.,* u_0 and $u_{\mathcal{N}-1}$, results in a new system of equations

$$\begin{bmatrix} & & & & 1 \\ & & & 1 & -2 \\ & & \iddots & \iddots & \\ & \iddots & \iddots & & \\ 1 & -2 & 1 & & \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ \vdots \\ u_{\mathcal{N}-2} \end{bmatrix} = \lambda_{\mathcal{N}}^{-1} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ \vdots \\ u_{\mathcal{N}-2} \end{bmatrix},$$

which says that $\lambda_{\mathcal{N}}$ is an eigenvalue of $\mathbf{H}_{\mathcal{N}-2}$. Note $(u_1, \ldots, u_{\mathcal{N}-2})^T \neq 0$, otherwise $\mathbf{u}_{\mathcal{N}} = 0$, which contradicts the definition of an eigenvector. By definition, $\sigma_{\mathcal{N}-2,\mathcal{N}-2} \leq |\lambda_{\mathcal{N}}|$. However, $\sigma_{\mathcal{N}\mathcal{N}} = |\lambda_{\mathcal{N}}|$, whence $\sigma_{\mathcal{N}\mathcal{N}} \geq \sigma_{\mathcal{N}-2,\mathcal{N}-2}$, which contradicts Proposition 2. Thus we conclude that $u_0 \neq 0$. This completes the proof.

Since $u_0 \neq 0$, zero is not a root of $\sum_{n=0}^{\mathcal{N}-1} u_n z^n = 0$, *i.e.,* $\zeta_n \neq 0$. To locate ζ_n we apply Schur's Theorem:

Theorem 1 (Schur) *[11, pp. 220] [10, pp. 109] The roots of the polynomial*

$$c_0 + c_1 z + \cdots + c_{n-1} z^{n-1} + c_n z^n = 0$$

are on or within the unit circle if and only if the quadratic form

$$\sum_{i=0}^{n-1} \left[(c_n x_i + c_{n-1} x_{i+1} + \cdots + c_{i+1} x_{n-1})^2 - (c_0 x_i + c_1 x_{i+1} + \cdots + c_{n-i-1} x_{n-1})^2 \right] \tag{15}$$

is positive semidefinite.

Proposition 4 *All the roots $\zeta_1, \zeta_2, \ldots, \zeta_{\mathcal{N}-1}$ of the polynomial $\sum_{n=0}^{\mathcal{N}-1} u_n z^n = 0$ are on or within the unit circle and they are either real or pairwise complex conjugate.*

Proof Let $\hat{\mathcal{N}} = \mathcal{N} - 1$; then the polynomial of interest can be written as $\sum_{n=0}^{\hat{\mathcal{N}}} u_n z^n = 0$. As explained in Section 2.2, all its roots $\zeta_1, \zeta_2, \ldots, \zeta_{\mathcal{N}-1}$ are either real or pairwise complex conjugate. To prove that all roots are on or within the unit circle, it suffices to prove the quadratic form

$$\sum_{n=0}^{\hat{\mathcal{N}}-1} \left[\left(u_{\hat{\mathcal{N}}} x_n + u_{\hat{\mathcal{N}}-1} x_{n+1} + \cdots + u_{n+1} x_{\hat{\mathcal{N}}-1} \right)^2 - \left(u_0 x_n + u_1 x_{n+1} + \cdots + u_{\hat{\mathcal{N}}-n-1} x_{\hat{\mathcal{N}}-1} \right)^2 \right] \quad (16)$$

is positive semidefinite. Note that this claim is implied by the positive semidefiniteness of the quadratic form

$$\sum_{n=0}^{\hat{\mathcal{N}}} \left[\left(u_{\hat{\mathcal{N}}} x_n + \cdots + u_{n+1} x_{\hat{\mathcal{N}}-1} + u_n x_{\hat{\mathcal{N}}} \right)^2 - \left(u_0 x_n + \cdots + u_{\mathcal{N}-n-1} x_{\mathcal{N}-1} + u_{\hat{\mathcal{N}}-n} x_{\hat{\mathcal{N}}} \right)^2 \right],$$

or equivalently the positive semidefiniteness of the matrix

$$\mathbf{C}^T\mathbf{C} - \mathbf{D}^T\mathbf{D},$$

where

$$\mathbf{C} = \begin{bmatrix} u_{\hat{\mathcal{N}}} & u_{\hat{\mathcal{N}}-1} & \cdots & u_0 \\ & u_{\hat{\mathcal{N}}} & \cdots & u_1 \\ & & \cdots & \cdots \\ & & & u_{\hat{\mathcal{N}}} \end{bmatrix}, \qquad \mathbf{D} = \begin{bmatrix} u_0 & u_1 & \cdots & u_{\hat{\mathcal{N}}} \\ & u_0 & \cdots & u_{\hat{\mathcal{N}}-1} \\ & & \cdots & \cdots \\ & & & u_0 \end{bmatrix}. \quad (17)$$

From Proposition 3 we know that $u_{\hat{\mathcal{N}}} \neq 0$, so $\mathbf{C}^{-1}$ exists. Therefore

$$\mathbf{C}^T\mathbf{C} - \mathbf{D}^T\mathbf{D} = \mathbf{C}^T \left(\mathbf{I} - \mathbf{C}^{-T}\mathbf{D}^T\mathbf{D}\mathbf{C}^{-1} \right) \mathbf{C}.$$

Let $\mathbf{Y} = \mathbf{D}\mathbf{C}^{-1}$. It is easy to verify that

$$\mathbf{Y} = \lambda_{\mathcal{N}} \begin{bmatrix} 1 & -2 & 1 & & & \\ & \cdots & \cdots & & & \\ & & \cdots & \cdots & & \\ & & & 1 & -2 & 1 \\ & & & & 1 & -2 \\ & & & & & 1 \end{bmatrix} = \lambda_{\mathcal{N}} \mathbf{P} \mathbf{H}_{\mathcal{N}}^{-1},$$

where $\lambda_\mathcal{N} \neq 0$ is the eigenvalue of $\mathbf{H}_\mathcal{N}$ that corresponds to the eigenvector $\mathbf{u}_\mathcal{N}$, *i.e.*, $\mathbf{H}_\mathcal{N}\mathbf{u}_\mathcal{N} = \lambda_\mathcal{N}\mathbf{u}_\mathcal{N}$, and

$$\mathbf{P} = \begin{bmatrix} & & & 1 \\ & & 1 & \\ & \iddots & & \\ 1 & & & \end{bmatrix}$$

is a permutation matrix. Consequently,

$$\mathbf{Y}^T\mathbf{Y} = \lambda_\mathcal{N}^2 \mathbf{H}_\mathcal{N}^{-T}\mathbf{P} \cdot \mathbf{P}\mathbf{H}_\mathcal{N}^{-1} = \sigma_{\mathcal{N}\mathcal{N}}^2 \mathbf{H}_\mathcal{N}^{-T}\mathbf{H}_\mathcal{N}^{-1} = \sigma_{\mathcal{N}\mathcal{N}}^2 \mathbf{H}_\mathcal{N}^{-2}$$

since $|\lambda_\mathcal{N}| = \sigma_{\mathcal{N}\mathcal{N}}$ and $\mathbf{P}^{-1} = \mathbf{P}$. Let $\mathbf{H}_\mathcal{N} = \mathbf{V}\Lambda\mathbf{V}^T$ be the spectral decomposition of $\mathbf{H}_\mathcal{N}$, where $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_\mathcal{N})$, $|\lambda_n| = \sigma_{\mathcal{N}n}$, for $1 \leq n \leq \mathcal{N}$, and $\mathbf{V}^T\mathbf{V} = \mathbf{I}$. Thus

$$\mathbf{I} - \mathbf{Y}^T\mathbf{Y} = \mathbf{V}\left[\mathbf{I} - \sigma_{\mathcal{N}\mathcal{N}}^2 \Lambda^{-2}\right]\mathbf{V}^T = \mathbf{V}\left[\mathbf{I} - \sigma_{\mathcal{N}\mathcal{N}}^2 \left[\mathrm{diag}(\sigma_{\mathcal{N}1}, \sigma_{\mathcal{N}2}, \ldots, \sigma_{\mathcal{N}\mathcal{N}})\right]^{-2}\right]\mathbf{V}^T.$$

Since $\sigma_{\mathcal{N}1} \geq \sigma_{\mathcal{N}2} \geq \ldots \geq \sigma_{\mathcal{N}\mathcal{N}} > 0$, $\mathbf{I} - \sigma_{\mathcal{N}\mathcal{N}}^2 \left[\mathrm{diag}(\sigma_{\mathcal{N}1}, \sigma_{\mathcal{N}2}, \ldots, \sigma_{\mathcal{N}\mathcal{N}})\right]^{-2}$ is positive semidefinite. Thus $\mathbf{C}^T\mathbf{C} - \mathbf{D}^T\mathbf{D}$ is positive semidefinite. This completes the proof.

Since $\gamma_n = 2\mathcal{M}\log\zeta_n$, Proposition 4 implies that $\Re(\gamma_n) \leq 0$, but all our numerical results show that all roots ζ_n are strictly within the unit circle, whence $\Re(\gamma_n) < 0$. Thus $\exp(\gamma_n x)$ converges to zero as x goes to infinity. This leads to the following conjecture.

Conjecture 1 *All the roots $\zeta_1, \zeta_2, \ldots, \zeta_{\mathcal{N}-1}$ of the polynomial $\sum_{n=0}^{\mathcal{N}-1} u_n z^n = 0$ are strictly within the unit circle.*

4 Numerical results

Presented below are plots related to different numbers, N, of terms in the exponential approximation $h_{\exp}(x)$ to the hockey stick function $h(x)$. In Figure 1 we plot the parameters for the 25-term exponential approximation. Conjugacy of ω_n and also γ_n is clearly shown in the plot. In Figure 2 we present the singular values of the Hankel matrix $\frac{1}{N}\mathbf{H}_\mathcal{N}$ associated with this 25-term exponential approximation.

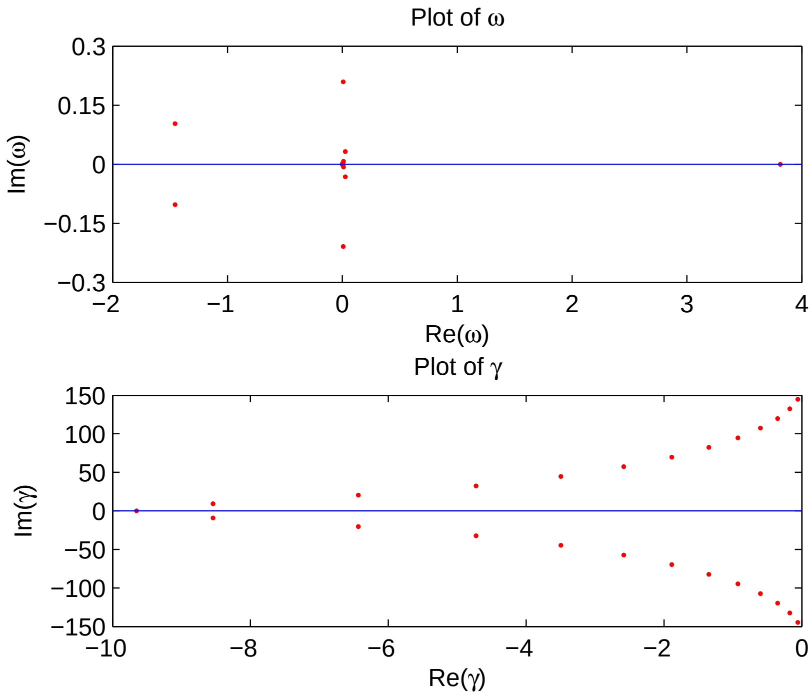

Figure 1: The parameters ω_n and γ_n for the 25-term exponential approximation

Finally in Figure 3 we plot the approximation errors of the 25-, 50-, 100-, 200-, and 400-term approximations over the interval $[0, 30]$. From these plots we can see that the approximation errors for all five choices of $\mathcal{N}$ converge to zero as the variable x increases.

Acknowledgments

The authors thank Professor Christina Christara for very helpful discussions on theoretical properties of the exponential approximation.

References

[1] Gregory Beylkin and Lucas Monzón. On approximation of functions by exponential sums. *Applied and Computational Harmonic Analysis*, 19(1):17–48, 2005.

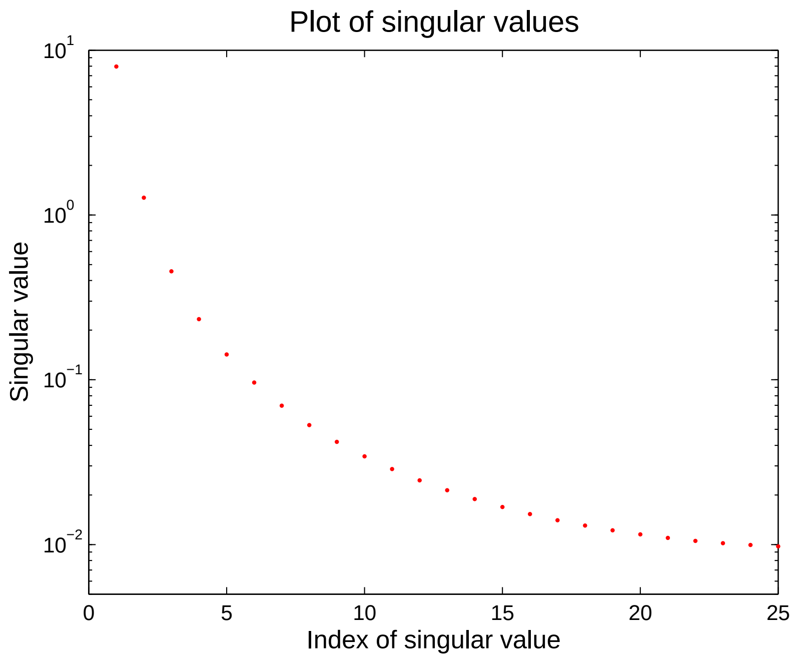

Figure 2: The singular values associated with the 25-term exponential approximation

[2] Dietrich Braess. *Nonlinear Approximation Theory*, volume 7 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin Heidelberg, 1986.

[3] Roger Fletcher. *Practical Methods of Optimization*. John Wiley & Sons Ltd, 2 edition, 1987.

[4] Gene Golub and Victor Pereyra. Separable nonlinear least squares: The variable projection method and its applications. *Inverse Problems*, 19(2):R1–R26, 2003.

[5] Gene H Golub and Charles F Van Loan. *Matrix Computations*, volume 3 of *Johns Hopkins Series in the Mathematical Sciences*. The Johns Hopkins University Press, 701 West 40th Street, Baltimore, Maryland 21211, 2 edition, 1991.

[6] Roger A Horn and Charles R Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, 1985.

[7] John Hull. *Options, Futures, and Other Derivatives*. Prentice Hall, Upper SaddleRiver, New Jersey 07458, 5th edition, 2003.

[8] Ian Iscoe, Ken Jackson, Alex Kreinin, and Xiaofang Ma. Pricing correlation-dependent derivatives based on exponential approximations to the hockey stick function. In preparation, January 2007.

[9] Liviu Gr Ixaru and Guido Vanden Berghe. *Exponential Fitting*. Kluwer Academic Publishers, 2004.

[10] Maurice Mignotte and Doru Stefanescu. *Polynomials — An Algorithmic Approach*. Springer-Verlag, 1999.

[11] Anthony Ralston. *A First Course in Numerical Analysis*. McGraw-Hill Book Company, 1965.

[12] James Hardy Wilkinson. *The Algebraic Eigenvalue Problem*. Oxford University Press, 1965.

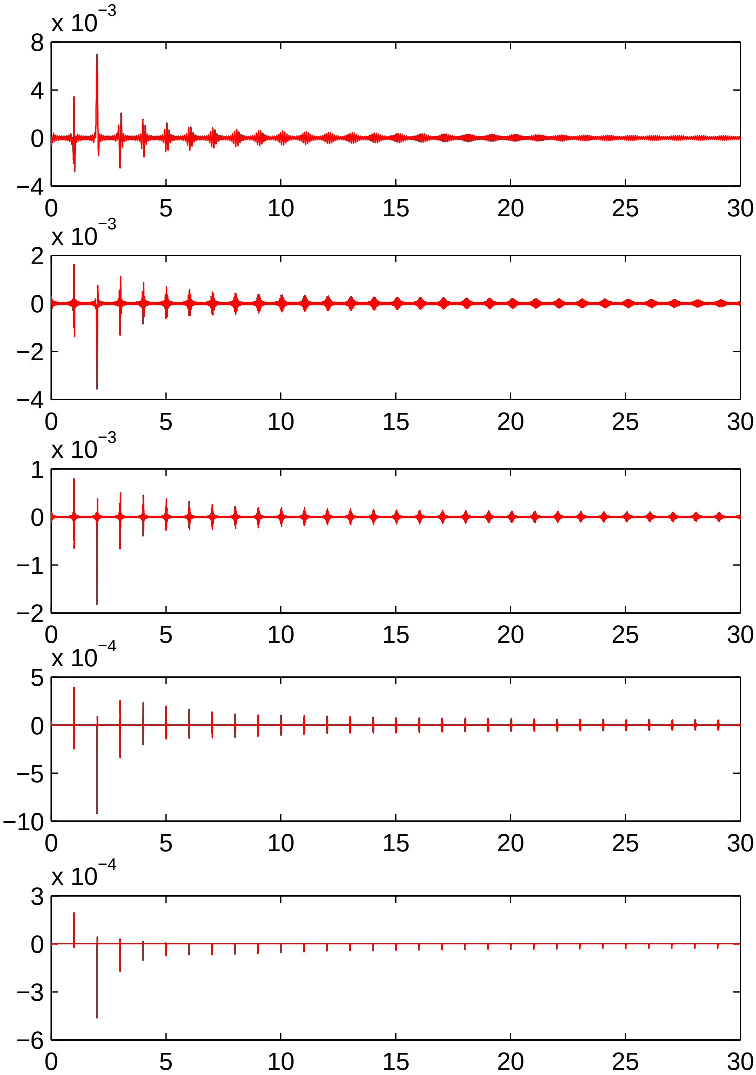

Figure 3: The panels from top to bottom are the approximation errors of the 25-term to 400-term exponential approximations to the HS function over $[0, 30]$, with the number of terms doubling in successive panels.