PROCEEDINGS A

rspa.royalsocietypublishing.org

Research

CrossMark
click for updates

Cite this article: Kisil AV. 2015 Stability analysis of matrix Wiener–Hopf factorization of Daniele–Khrapkov class and reliable approximate factorization. *Proc. R. Soc. A* **471**: 20150146.
http://dx.doi.org/10.1098/rspa.2015.0146

Received: 3 March 2015
Accepted: 10 April 2015

Subject Areas:
analysis

Keywords:
Wiener–Hopf, Daniele–Khrapkov, Riemann–Hilbert, rational approximation

Author for correspondence:
Anastasia V. Kisil
e-mail: a.kisil@maths.cam.ac.uk

THE ROYAL SOCIETY
PUBLISHING

Stability analysis of matrix Wiener–Hopf factorization of Daniele–Khrapkov class and reliable approximate factorization

Anastasia V. Kisil

Cambridge Centre of Analysis, University of Cambridge, Wilberforce Road, Cambridge CB3 0WA, UK

This paper presents new stability results for matrix Wiener–Hopf factorization. The first part of the paper examines conditions for stability of Wiener–Hopf factorization in the Daniele–Khrapkov class. The second part of the paper concerns the class of matrix functions which can be exactly or approximately reduced to the factorization of the Daniele–Khrapkov matrices. The results of the paper are demonstrated by numerical examples with partial indices $\{1,-1\}$, $\{0,0\}$ and $\{-1,-1\}$.

1. Introduction

This paper examines the stability of Wiener–Hopf matrix factorization [1–3] in a certain class of matrices. In essence, a factorization of a scalar or matrix function $\mathbf{G}(t)$ is its decomposition into a product

$$\mathbf{G}(t) = \mathbf{G}_+(t)\mathbf{G}_-(t) \tag{1.1}$$

with the invertible factors $\mathbf{G}_+(t)$ and $\mathbf{G}_-(t)$ analytically extendable into the upper/lower half-plane (§2a). We consider the class of Daniele–Khrapkov 2×2 matrices [4,5], which have the form

$$\mathbf{K}(t) = \mathbf{I} + f(t)\mathbf{J}(t), \tag{1.2}$$

where $f(t)$ is an arbitrary scalar function with algebraic growth at infinity, and $\mathbf{J}(t)$ is a polynomial matrix with

$$\mathbf{J}^2(t) = \Delta^2(t)\mathbf{I},$$

where $\Delta^2(t)$ is a polynomial in t and $\mathbf{I}$ is the 2×2 identity matrix. The Daniele–Khrapkov matrices can be factorized explicitly (§4).

Fundamentally, the scalar and the matrix Wiener–Hopf factorizations (1.1) are different: the former has a constructive solution in terms of a Cauchy type integral and the latter has no explicit solution in general. The existence of matrix factorization under general assumptions has been proved by Gohberg & Krein [6]. Nevertheless, up to date constructive matrix factorization remains a formidable challenge. Owing to its complexity, different classes of matrix functions have to be treated separately (see the recent review article [2]). The class of Daniele–Khrapkov matrices is very important for applications and arise naturally in a number of interesting problems in acoustic, electromagnetics, etc., see for example [4,7,8].

The Wiener–Hopf factorization (1.1) is said to be *stable* if small changes in the matrix function $\mathbf{G}(t)$ lead to small changes in the factors $\mathbf{G}_+(t)$ and $\mathbf{G}_-(t)$ (§4). Almost all implementations of Wiener–Hopf technique are performed numerically [9]; therefore, a careful analysis of stability is essential. Among popular approximate techniques are truncated pole removal [10] and rational approximations [11,12]. There are also new asymptotic methods [13–15], which also rely on stability. Even in the rare cases when explicit factorizations are known, e.g. for Daniele–Khrapkov matrices, they still require numerical computations of scalar factorizations. Those computations introduce small errors, which can lead to large errors in the Wiener–Hopf factors (§3).

A landmark theorem of Litvinchuk & Spitkovskii [3, §6.2] gives general conditions for stability of matrix Wiener–Hopf factorization (§3). The difficulty of applying these results is that the stability conditions depend on the knowledge of Wiener–Hopf factorization and hence are impractical to check. The aim of this paper is to provide direct criteria for stability of factorization in a case of Daniele–Khrapkov matrices. The conditions are demonstrated by numerical examples.

This work is a continuation of the author's paper [16], which demonstrated a novel method of approximately solving scalar Wiener–Hopf equations. In the scalar case, the formula for the solution in terms of a Cauchy type integral was used to bound the error in the factors. In this paper, the previous results are extended to the Daniele–Khrapkov matrices.

The first part of the paper establishes the stability of the Daniele–Khrapkov class under perturbations within the class. There are benefits to considering the 'near' matrices only within the class. It allows the question whether numerical implementation of the factorization is stable to be answered. This also allowed explicit error bounds to be obtained. The third advantage is that in a specific case stronger results can be obtained then in the general case.

The second part of the paper extends the class of matrix functions to these which can be approximately reduced to Daniele–Khrapkov matrices. The class of matrix functions considered by Abrahams in [17] is a special case of this construction. It is shown that the stability results could be applied to this meromorphic factorization. This is then used to show stability in an interesting example.

2. Preliminaries

Throughout the paper we are using the subscripts + and − to denote functions which admit an analytic continuation into the upper and lower half-planes, respectively. The *Wiener algebra* $W(\mathbb{R})$ over the real line [1, ex. 2.2] consists of all complex valued functions f in $\mathbb{R}$ that admit a representation of the form

$$f(\lambda) = d + \int_{-\infty}^{\infty} e^{i\lambda t} k(t)\, dt, \quad \lambda \in \mathbb{R},$$

for some $d \in \mathbb{C}$ and $k \in L_1(\mathbb{R})$.

(a) Wiener–Hopf factorization

This subsection recalls the different types of Wiener–Hopf factorization, which have their own merits, see [1] for a detailed exposition. Let $\mathbf{G}(t)$ be in the matrix Wiener algebra $W_{2\times 2}(\mathbb{R})$ [3, §5.2].

If $\det \mathbf{G}(t) \neq 0$ for all real t then there exists the full factorization

$$\mathbf{G}(t) = \mathbf{G}_+(t)\mathbf{D}(t)\mathbf{G}_-(t), \quad t \in \mathbb{R}, \tag{2.1}$$

where factors and their inverses belong to the subalgebras of analytically extendable functions to the respective half-planes

$$\mathbf{G}_+^{\pm 1} \in W^+_{2\times 2}(\mathbb{R}), \quad \mathbf{G}_-^{\pm 1} \in W^-_{2\times 2}(\mathbb{R}) \quad \text{and} \quad \mathbf{D}(t) = \operatorname{diag}\left[\left(\frac{t-i}{t+i}\right)^{\kappa_1}, \left(\frac{t-i}{t+i}\right)^{\kappa_2}\right].$$

The integer exponents κ_1 and κ_2 are called *partial indices*. Unlike factorization, the partial indices are unique. But in contrast to the scalar case, they cannot be determined *a priori* in general.

A factorization (1.1) with the invertible factors $\mathbf{G}_+(t)$ and $\mathbf{G}_-(t)$ analytically extendable into the respective half-planes and polynomially bounded growth at infinity will be called *function-theoretic* factorization. The function-theoretic factorization is useful in applications as it retains most information and is easier to find.

Remark 2.1. The partial indices are linked to the growth at infinity in function-theoretic factorization, see [18].

It is also useful to consider a *meromorphic* factorization, where the conditions are further relaxed to allow the presence of a finite number of poles and zeroes in the factors.

(b) Scalar error estimates

The *index* of a continuous non-zero function $K(t)$ on the real line is:

$$\operatorname{ind}(K(t)) = \frac{1}{2\pi}\left(\lim_{t\to+\infty} \arg K(t) - \lim_{t\to-\infty} \arg K(t)\right). \tag{2.2}$$

Note that $\operatorname{ind}((t-i)/(t+i)) = 1$. Thus, given a function $K(t)$ with index κ one can reduce it to zero index by considering

$$K(t)\left(\frac{t-i}{t+1}\right)^{-\kappa}.$$

For the rest of this subsection, it will be assumed that all functions have zero index.

We also assume that $K(t) \to 1$ for $t \to \pm\infty$, then we can normalize factors such that $K_\pm(t) \to 1$ for $t \to \pm\infty$. A non-zero Hölder continuous function $K(t)$ on the real line with $K(t) - 1$ in $L_2(\mathbb{R})$ possesses a factorization [19]

$$K(t) = K_+(t)K_-(t),$$

where $K_\pm(t)$ are limiting values of functions analytic and non-zero in the respective half-planes.

The distinctive feature of the scalar factorization is the ability to express the factors in terms of the Cauchy-type integrals. It is the existence of such expressions and the bounds in L_p on the Hilbert transform which allowed to obtain some useful estimation [16]. We adapt them here for L_2 case in the following form.

Theorem 2.2 (Additive estimates in L_2). *Let $F(t) = F_+(t) + F_-(t)$ and $\tilde{F}(t) = \tilde{F}_+(t) + \tilde{F}_-(t)$ with $\|F(t) - \tilde{F}(t)\|_2 < \epsilon$ then*

$$\|F_\pm(t) - \tilde{F}_\pm(t)\|_2 \le \epsilon.$$

Theorem 2.3 (Multiplicative estimates in L_2). *Let $K(t) = K_+(t)K_-(t)$ and $\tilde{K}(t) = \tilde{K}_+(t)\tilde{K}_-(t)$ be two functions and $m < |K| < M$. If $\|K(t) - \tilde{K}(t)\|_2 < \epsilon$ then*

$$\|K_\pm(t) - \tilde{K}_\pm(t)\|_2 < \frac{5(M+\epsilon)^{1/2}}{(m-\epsilon)}\epsilon.$$

The above results are special cases of theorems from [16] with some more explicit constants calculated.

3. Stability of matrix Wiener–Hopf

For the sake of completeness, we review here the most general results on stability of matrix factorization, as they are not widely known in the Wiener–Hopf community. The examples are adapted from a different context of a Riemann–Hilbert problem on a circle. There is a wealth of different classes of factorizations considered by different authors; for the purpose of clear exposition, we consider here only factorization in Wiener algebra (2.1).

The simplest example of instability is obtained by mapping an example [1] from the unit circle to the real line. Consider a diagonal matrix function with partial indices $\{1, -1\}$

$$\begin{pmatrix} \dfrac{t-i}{t+i} & 0 \\ 0 & \dfrac{t+i}{t-i} \end{pmatrix} = \mathbf{I} \begin{pmatrix} \dfrac{t-i}{t+i} & 0 \\ 0 & \dfrac{t+i}{t-i} \end{pmatrix} \mathbf{I}. \tag{3.1}$$

Perturbing the matrix we have

$$\begin{pmatrix} \dfrac{t-i}{t+i} & 0 \\ \epsilon & \dfrac{t+i}{t-i} \end{pmatrix} = \begin{pmatrix} 1 & \dfrac{t-i}{t+i} \\ 0 & \epsilon \end{pmatrix} \mathbf{I} \begin{pmatrix} 0 & -\dfrac{1}{\epsilon} \\ 1 & \dfrac{t+i}{\epsilon(t-i)} \end{pmatrix}. \tag{3.2}$$

This example demonstrates that a small perturbations can not only change the factors by an arbitrary amount but can also change the partial indices (from $\{1,-1\}$ to $\{0,0\}$). This is significant because the partial indices are uniquely defined. Note that the sum of the partial indices remains the same. This is true in general, which can be demonstrated if we equate the determinants of both sides to reduce the problem to scalar factorization. The partial indices add to give the index (2.2) of the determinant. In this case, the index of a function f is the winding number of the curve $(\mathrm{Re} f(t),\ \mathrm{Im} f(t))$, $t \in \mathbb{R}$. Hence, $\mathrm{ind}(f)$ and, thus the sum of partial indices, are stable under small perturbations.

Remark 3.1. It is possible to use the non-uniqueness of factorization [1] to obtain a different factorization of (3.1)

$$\begin{pmatrix} \dfrac{t-i}{t+i} & 0 \\ 0 & \dfrac{t+i}{t-i} \end{pmatrix} = \begin{pmatrix} 1 & \dfrac{t-i}{t+i} \\ 0 & \epsilon \end{pmatrix} \begin{pmatrix} \dfrac{t-i}{t+i} & 0 \\ 0 & \dfrac{t+i}{t-i} \end{pmatrix} \begin{pmatrix} 1 & -\dfrac{t+i}{\epsilon(t-i)} \\ 0 & \dfrac{1}{\epsilon} \end{pmatrix}. \tag{3.3}$$

This is more similar to (3.1).

The following surprising theorem provides the necessary and sufficient conditions for the partial indices to be invariant under sufficiently small perturbations.

Theorem 3.2 (Gohberg–Krein [3, §6.2]). *The system $\kappa_1 \geq \cdots \geq \kappa_n$ of partial indices is stable if and only if*

$$\kappa_1 - \kappa_n \leq 1.$$

In fact, this condition is also sufficient for the stability of factors in the Wiener norm.

Theorem 3.3 (Shubin [3, §6.6]). *Assume the matrix function $\mathbf{G}$ has a Wiener–Hopf factorization and the tuple of its partial indices is stable. Then, for every $\epsilon > 0$ there exists a $\delta > 0$ such that, for $\|\mathbf{F} - \mathbf{G}\| < \delta$, the matrix function $\mathbf{F}$ admits a factorization in which $\|\mathbf{F}_{\pm} - \mathbf{G}_{\pm}\| < \epsilon$.*

An obstacle in using this result in applications is that one cannot in general determine the partial indices without constructing the factorization. The next section presents new conditions for stability of factorization for Daniele–Khrapkov matrices.

4. Error estimates in Daniele–Khrapkov matrices

This section examines function-theoretic factorization of matrices of Daniele–Khrapkov class (1.2). This class was first considered by Khrapkov in connection to static stress fields induced by notches in elastic wedges [5]. There are other numerous applications, e.g. related to wave propagation [4,7,20].

Owing to this special form (1.2) $\mathbf{K}(t)$ can be re-expressed as

$$\mathbf{K}(t) = r(t)\left(\cosh[\Delta(t)\theta(t)]\mathbf{I} + \frac{1}{\Delta(t)}\sinh[\Delta(t)\theta(t)]\mathbf{J}(t)\right),$$

where

$$r(t) = \sqrt{1-\Delta^2(t)f^2(t)}, \quad \theta(t) = \frac{1}{\Delta(t)}\ln\left(\frac{1+\Delta(t)f(t)}{1-\Delta(t)f(t)}\right). \tag{4.1}$$

Multiplication of the above matrices is commutative, moreover,

$$\mathbf{K_1}(t)\mathbf{K_2}(t) = R(t)\left(\cosh[\Delta(t)\Theta(t)]\mathbf{I} + \frac{1}{\Delta(t)}\sinh[\Delta(t)\Theta(t)]\mathbf{J}(t)\right),$$

where

$$R(t) = r_1(t)r_2(t), \quad \Theta(t) = \theta_1(t) + \theta_2(t).$$

This property is enough to obtain function-theoretical factorization

$$\mathbf{K}_\pm(t) = r_\pm(t)\left(\cosh[\Delta(t)\theta_\pm(t)]\mathbf{I} + \frac{1}{\Delta(t)}\sinh[\Delta(t)\theta_\pm(t)]\mathbf{J}(t)\right), \tag{4.2}$$

where

$$r(t) = r_-(t)r_+(t), \quad \theta = \theta_-(t) + \theta_+(t).$$

The limitation is the degree of the polynomial Δ^2: if it is greater than 2 then $\cosh[\Delta(t)\theta_\pm(t)]$ and $\sinh[\Delta(t)\theta_\pm(t)]$ have exponential growth at infinity [17]. This is an obstacle to the use of the Wiener–Hopf technique.

We consider the question of stable factorization for Daniele–Khrapkov matrices in the following sense. Let $K(t)$ and $\tilde{K}(t)$ be of Daniele–Khrapkov type and suppose $\|K(t)-\tilde{K}(t)\|_2$ is small. We provide an estimate on $\|K_\pm(t)-\tilde{K}_\pm(t)\|_2$. This splits into three parts. The first part is to establish estimates for $\|r(t)-\tilde{r}(t)\|_2$ and $\|\theta(t)-\tilde{\theta}(t)\|_2$ defined by (4.1). The second is to apply the error estimates to parameters $r_\pm(t)$ and $\theta_\pm(t)$ of the factors. Lastly, $\|K_\pm(t)-\tilde{K}_\pm(t)\|_2$ can be examined.

Consider the matrix function $\mathbf{K}(t)$ and its perturbation $\tilde{\mathbf{K}}(t)$

$$\mathbf{K}(t) = \mathbf{I} + f(t)\mathbf{J}(t), \quad \tilde{\mathbf{K}}(t) = \mathbf{I} + \tilde{f}(t)\mathbf{J}(t),$$

such that $\|\Delta(t)f(t)-\Delta(t)\tilde{f}(t)\|_2 < \epsilon$. In this set-up, the perturbation of $r(t)$ can be estimated as follows.

Lemma 4.1. *Let* $r = \sqrt{1-\Delta^2(t)f^2(t)}$ *and* $\tilde{r} = \sqrt{1-\Delta^2(t)\tilde{f}^2(t)}$. *Suppose that the winding number of* $(1-\Delta^2(t)f^2(t))$ *is zero, then for* $\|\Delta(t)f(t)-\Delta(t)\tilde{f}(t)\|_2 < \epsilon$ *the following estimate holds*

$$\|r-\tilde{r}\|_2 < \frac{N}{m}\epsilon,$$

where $m = \min_{\mathbb{R}}\{|r(t)|, |\tilde{r}(t)|\} > 0$ *and* $N = \max_{\mathbb{R}}\{|\Delta(t)f(t)|, |\Delta(t)\tilde{f}(t)|\} < \infty$.

Remark 4.2. The assumptions are natural as $|r(t)|^2$ is the determinant of the matrix K which together with the determinant of its inverse is non-zero.

Proof. As winding number of $(1-\Delta^2(t)f^2(t))$ is zero and ϵ is small enough, we have winding number of $(1-\Delta^2(t)\tilde{f}^2(t))$ is also zero. The square root for $r(t)$ in (4.1) can be taken single valued. In the inequality

$$\left|\sqrt{a}-\sqrt{b}\right| = \frac{|a-b|}{\sqrt{a}+\sqrt{b}} \le \frac{|a-b|}{2\min(\sqrt{a},\sqrt{b})},$$

we substitute $a = 1-\Delta^2(t)f^2(t)$ and $b = 1-\Delta^2(t)\tilde{f}^2(t)$. We also replace $\min(\sqrt{a},\sqrt{b})$ by a smaller value $m = \min_{\mathbb{R}}\{|r(t)|, |\tilde{r}(t)|\} > 0$. Integrating squares of the both sides over the real line we obtain

$$\begin{aligned}\|r-\tilde{r}\|_2 &\le \frac{1}{2m}\|\Delta^2(f^2-\tilde{f}^2)\|_2\\ &\le \frac{1}{2m}\|(\Delta(f+\tilde{f}))(\Delta(f-\tilde{f}))\|_2\\ &\le \frac{1}{2m}\left(\int_{\mathbb{R}}|\Delta(t)(f(t)+\tilde{f}(t))\Delta(t)(f(t)-\tilde{f}(t))|^2\,\mathrm{d}t\right)^{1/2}\\ &\le \frac{2N}{2m}\left(\int_{\mathbb{R}}|\Delta(t)(f(t)-\tilde{f}(t))|^2\,\mathrm{d}t\right)^{1/2}\\ &\le \frac{N}{m}\epsilon,\end{aligned}$$

as $|\Delta(t)f(t)|$ and $|\Delta(t)\tilde{f}(t)|$ are bounded by $N = \max_{\mathbb{R}}\{|\Delta(t)f(t)|, |\Delta(t)\tilde{f}(t)|\}$. ■

Similarly, the behaviour of θ under perturbation is important.

Lemma 4.3. *Let*

$$\theta(t) = \frac{1}{\Delta(t)}\ln\left(\frac{1-\Delta(t)f(t)}{1+\Delta(t)f(t)}\right), \quad \tilde{\theta}(t) = \frac{1}{\Delta(t)}\ln\left(\frac{1-\Delta(t)\tilde{f}(t)}{1+\Delta(t)\tilde{f}(t)}\right).$$

Suppose that the winding number of $((1-\Delta(t)f(t))/(1+\Delta(t)f(t)))$ *is zero, then for small* $\|\Delta(t)f(t)-\Delta(t)\tilde{f}(t)\|_2 < \epsilon$ *the following estimate holds*

$$\|\theta-\tilde{\theta}\|_2 < \frac{2\epsilon}{cd^2L},$$

where $d = \min_{\mathbb{R}}\{|1+\Delta(t)f(t)|, |1+\Delta(t)\tilde{f}(t)|\} > 0$ *and* $L = \max_{\mathbb{R}}|(1-\Delta(t)f(t))/(1+\Delta(t)f(t))|$, $c = \min_{\mathbb{R}}|\Delta(t)| > 0$.

Remark 4.4. Since Δ has no zeroes on the real line we can assume $\min|\Delta| \ge c > 0$. Also note that $|1+\Delta(t)f(t)|$ and $|1-\Delta(t)f(t)|$ are non-zero and finite, respectively, as they are multiples of $\det K$.

Proof. From the assumption on zero winding number, the logarithms in the definition $\theta(t)$ and $\tilde{\theta}(t)$ are single-valued functions. The mean-value theorem applied to the logarithm function provides an inequality:

$$|\ln a - \ln b| \le \frac{|a-b|}{\min(a,b)}.$$

We substitute $\ln a = \Delta(t)\theta(t)$, $\ln b = \Delta(t)\tilde{\theta}(t)$ and replace $\min(a,b)$ by L defined in the statement. Then, squaring both sides and integrating over the real line, we obtain

$$\begin{aligned}\|\theta-\tilde{\theta}\|_2 &\le \frac{1}{cL}\left\|\frac{1-\Delta(t)f(t)}{1+\Delta(t)f(t)} - \frac{1-\Delta(t)\tilde{f}(t)}{1+\Delta(t)\tilde{f}(t)}\right\|_2\\ &\le \frac{2}{cL}\left\|\frac{\Delta(t)f(t)-\Delta(t)\tilde{f}(t)}{(1+\Delta(t)f(t))(1+\Delta(t)\tilde{f}(t))}\right\|_2\\ &\le \frac{2}{cd^2L}\|\Delta(t)f(t)-\Delta(t)\tilde{f}(t)\|_2,\end{aligned}$$

where c and d are defined in the statement. ■

Now we are in the position to apply the scalar error estimates. Under the assumptions of the above lemma 4.3 and using the additive error estimates theorem 2.2 we obtain

$$\|\theta_\pm - \tilde{\theta}_\pm\|_2 < \frac{2}{cd^2L}\epsilon. \tag{4.3}$$

Using lemma 4.1 and the multiplicative error estimates theorem 2.3, it follows that

$$\|r_\pm - \tilde{r}_\pm\|_2 < \frac{5MN}{m^2}\epsilon, \tag{4.4}$$

where $M = \max_{\mathbb{R}}\{|r(t)|, |\tilde{r}(t)|\} > 0$.

To simplify calculation in the next theorem, we will assume that

$$\mathbf{J} = \begin{pmatrix} 0 & k_1 \\ k_2 & 0 \end{pmatrix},$$

is a constant matrix. Then, a sufficiently small $\|\Delta(t)f(t) - \Delta(t)\tilde{f}(t)\|_2$ guarantees that $\|\mathbf{K} - \tilde{\mathbf{K}}\|_2$ is also small.

Theorem 4.5. *Let* $\mathbf{K}$ *and* $\tilde{\mathbf{K}}$ *be of the above form,* $\|\Delta(t)f(t) - \Delta(t)\tilde{f}(t)\|_2 < \epsilon$ *and* $\Delta(t) = C$*, satisfying the assumptions of lemmas* 4.1 *and* 4.3*. Then, the error* $\|\mathbf{K}_\pm - \tilde{\mathbf{K}}_\pm\|_2$ *is a linear function of* ϵ *and the exact estimates can be obtained using the above scalar estimates.*

Proof. Let a_{11} and $\tilde{a}_{11}$ are the top-left elements of $\mathbf{K}$ and $\tilde{\mathbf{K}}$, respectively. Then

$$\begin{aligned}\|a_{11} - \tilde{a}_{11}\|_2 &= \|r_\pm(t)\cosh[\Delta(t)\theta_\pm(t)] - \tilde{r}_\pm(t)\cosh[\Delta(t)\tilde{\theta}_\pm(t)]\|_2 \\ &\le \|r_\pm(\cosh[\Delta(t)\theta_\pm(t)] - \cosh[\Delta(t)\tilde{\theta}_\pm(t)])\|_2 \\ &\quad + \|\cosh[\Delta(t)\tilde{\theta}_\pm(t)](r_\pm - \tilde{r}_\pm)\|_2,\end{aligned}$$

where the triangle inequality was used. Then, using the mean-value theorem for cosh we obtain

$$\begin{aligned}\|a_{11} - \tilde{a}_{11}\|_2 \le |r_\pm|\,|\sinh[\Delta(t)\theta_\pm(t)]|\,\|\Delta(t)\theta_\pm(t) - \Delta(t)\tilde{\theta}_\pm(t)\|_2 \\ + |\cosh[\Delta(t)\tilde{\theta}_\pm(t)]|\,\|(r_\pm - \tilde{r}_\pm)\|_2.\end{aligned}$$

To complete the calculation it is enough to use the bound for $|r_\pm|$, $|\sinh[\Delta(t)\theta_\pm(t)]|$ and $|\cosh[\Delta(t)\tilde{\theta}_\pm(t)]|$. This follows from $r_\pm$ and $\theta_\pm$, being bounded, having zero winding number and tending to a constant [16]. The calculations for other entries $\|a_{ij} - \tilde{a}_{ij}\|_2$, $i, j = 1, 2$ are performed analogously. All the norms of 2×2 matrices are equivalent so it does not matter which one is chosen. ■

In the subsequent sections, we present several situations where our results may be applied. Numerical examples will be presented in §6.

5. Approximate reducing to extended Daniele–Khrapkov

(a) Exact reduction to Daniele–Khrapkov matrices

The most general class of matrix functions which can be factored using the above technique is

$$\mathbf{K} = \mathbf{S}_+(g_1\mathbf{I} + g_2\mathbf{J})\mathbf{S}_-. \tag{5.1}$$

with $\mathbf{S}_+$ and $\mathbf{S}_-$ analytic in the upper and lower half-plane, respectively, and $\mathrm{tr}\mathbf{J} = 0$.

This can be rearranged as

$$\mathbf{K} = g_1\mathbf{S_1} + g_2\mathbf{S_2}, \quad \text{where } \mathbf{S_1} = \mathbf{S}_+\mathbf{S}_- \text{ and } \mathbf{S_2} = \mathbf{S}_+\mathbf{J}\mathbf{S}_-. \tag{5.2}$$

The challenge is to work backwards from equation (5.2) to (5.1). The first step is the factorization of $\mathbf{S_1} = \mathbf{S}_+\mathbf{S}_-$ and second step is to ensure the second term satisfies the necessary conditions for $\mathbf{J} = \mathbf{S}_+^{-1}\mathbf{S}_2\mathbf{S}_-^{-1}$. To satisfy these considerations one can take $\mathbf{S_1}$ and $\mathbf{S_2}$ to be rational; this class was studied in Prössdorf & Speck [21].

Now we will outline the procedure to reduce equation (5.2) to (5.1). Initially, one must rule out the case when $\mathbf{S_1}$ has a zero on the real line. As the matrix $\mathbf{K}$ does not have any zeros, any zeros of $\mathbf{S_1}$ must be compensated either by multiplying by f_1 or by adding $f_2\mathbf{S_2}$. So by constructing a different linear combination it can be assumed that $\mathbf{S}_1$ is non-zero on the real line. Then using the rational factorization $\mathbf{S_1} = \mathbf{S}_+\mathbf{S}_-$, we obtain

$$\mathbf{K} = \mathbf{S}_+(g_1\mathbf{I} + g_2\mathbf{R})\mathbf{S}_-,$$

with $\mathbf{R} = \mathbf{S}_+^{-1}\mathbf{S_2}\mathbf{S}_-^{-1}$.

This can it can be re-written as

$$\mathbf{K} = \mathbf{S}_{1+}(f_1\mathbf{I} + f_2\mathbf{J})\mathbf{S}_{1-},$$

where $\mathbf{J} = \mathbf{R} - 1/2\,\mathrm{tr}(\mathbf{R})$ for some new functions f_1 and f_2, see [21] for further details. We will call such matrices *extended* Daniele–Khrapkov class.

(b) Approximate reduction to Daniele–Khrapkov

We give a description of a larger class of matrices which may approximately factorized through approximation by matrix functions from the extended Daniele–Khrapkov class (5.2). Those matrices have the property that every entry of the matrix has elements of the form:

$$f_1 r_{ij}^1 + f_2 r_{ij}^2,$$

with two fixed arbitrary functions f_1 and f_2 and rational functions r_{ij}^1 and r_{ij}^2. In the whole generality, it shall be discussed elsewhere. Here, we concentrate on a subclass, related to work [17] with interesting applications [7]. This subclass allows to overcome the problem of exponential growth of the factors in the Daniele–Khrapkov matrices for high degree of polynomial $\Delta(t)$. This approximate procedure is simpler than the exact one provided by Daniele [4, §4.8.5].

Let us begin with matrix

$$\mathbf{K}(t) = \mathbf{I} + f(t)\begin{pmatrix} 0 & n(t) \\ p(t) & 0 \end{pmatrix}.$$

We can rearrange it into the form

$$\mathbf{K}(t) = \mathbf{I} + g(t)\mathbf{J}(t),$$

with

$$\mathbf{J}(t) = \begin{pmatrix} 0 & \left(\dfrac{n(t)}{p(t)}\right)^{1/2} \\ \left(\dfrac{p(t)}{n(t)}\right)^{1/2} & 0 \end{pmatrix},$$

and $g(t) = f(t)(n(t)/p(t))^{1/2}$. The advantage of this rearrangement being,

$$\mathbf{J}^2(t) = \mathbf{I},$$

and the disadvantage is that now $\mathbf{J}$ has branch cut singularities. To overcome that Abrahams proposed to rationally approximate $(p(t)/n(t))^{1/2}$ by $r_N(t)$ giving

$$\mathbf{J}_N(t) = \begin{pmatrix} 0 & \dfrac{1}{r_N(t)} \\ r_N(t) & 0 \end{pmatrix}.$$

This procedure is exact when $n(t)$ and $p(t)$ have perfect squares as factors.

The approximate matrix can be decomposed as in (4.2)

$$\mathbf{K}_N(t) = \mathbf{I} + g(t)\mathbf{J}_N(t) = \mathbf{Q}_{N-}\mathbf{Q}_{N+},$$

but the factors $\mathbf{Q}_{N\pm}$ have poles. Hence, a meromorphic factorization is obtained.

Remark 5.1. Error bounds (4.3) and (4.4) on $\theta_\pm$ and $r_\pm$ still hold in this meromorphic factorization.

To remove poles, we can consider the factorization

$$\mathbf{K}_N(t) = (\mathbf{Q}_{N-}\mathbf{M})(\mathbf{M}^{-1}\mathbf{Q}_{N+}), \tag{5.3}$$

where $\mathbf{M}$ is a rational matrix, which is chosen such that the resulting factorization has no poles in the required half-planes, see [17] for further details. We are turning to illustrations of this method.

Example 5.2. This example is concerned with the earlier example of instability (3.1). The aim is to show that although the indices are 1 and −1, it is still possible to have a stable perturbation. The construction is based on the results from the previous sections

$$\begin{pmatrix} \dfrac{t-i}{t+i} & \epsilon f(t) \\ c\epsilon f(t) & \dfrac{t+i}{t-i} \end{pmatrix} = \begin{pmatrix} \dfrac{t-i}{t+i} & 0 \\ 0 & \dfrac{t+i}{t-i} \end{pmatrix} + \epsilon f(t) \begin{pmatrix} 0 & 1 \\ c & 0 \end{pmatrix},$$
$$= \begin{pmatrix} t+i & 0 \\ 0 & \dfrac{1}{t+i} \end{pmatrix} \mathbf{K} \begin{pmatrix} \dfrac{1}{t-i} & 0 \\ 0 & t-i \end{pmatrix}.$$

with

$$\mathbf{K} = \mathbf{I} + \epsilon f(t) \begin{pmatrix} 0 & (t-i)^{-1}(t+i)^{-1} \\ c(t-i)(t+i) & 0 \end{pmatrix}.$$

The matrix $\mathbf{K}$ is of Abrahams type with the ratio of the off-diagonal elements being a square. Hence, there is no need for rational approximation and the procedure is exact in this case. One can construct the factors using (4.2). Lemmas 4.1 and 4.3 can be applied when f satisfies their assumptions. Hence, a meromorphic factorization has been obtained which is stable for small ϵ.

Then, the final step is to construct a matrix **M** as in (5.3). In the case when $f(t) \equiv k$, the matrix **M** takes the form

$$\mathbf{M} = \begin{pmatrix} 1 + \dfrac{i/2}{t-i} + \dfrac{i/2k}{t+i} & \dfrac{i/2}{t-i} + \dfrac{i/2k}{t+i} \\ 1 & 1 \end{pmatrix}, \tag{5.4}$$

with $\det \mathbf{M} = 1$. This completes the factorization of the perturbed matrix.

6. Numerical results

This section presents two approximate scalar factorizations with different indices and these are used to construct two approximate Daniele–Khrapkov factorizations.

(a) Rational approximation

Rational approximation of functions has its uses in Wiener–Hopf factorization. One example was mentioned in previous section. Kisil [16] applies rational approximation to simplify the scalar factorization and avoid calculations of a Cauchy type integral.

Rational approximation is useful for Daniele–Khrapkov factorization because once the approximations for K_1 and K_2 are obtained algebraic expressions such as

$$K_1 + c, \quad K_1 + K_2, \quad K_1 K_2,$$

can be factored easily. This is not true in general as can be seen from the next two examples.

Example 6.1. Consider the function with zero index

$$F(t) = \sqrt{\frac{t^2+1}{t^2+k^2}}, \tag{6.1}$$

and with finite branch cuts from i to ki and from $-i$ to $-ki$. This function is closely associated with the matrix function factorization from problems in acoustics and elasticity [22]. The factors can easily be seen by inspection

$$F_\pm(t) = \sqrt{\frac{(t \pm i)}{(t \pm ik)}}, \quad F_+(t) = F_-(-t).$$

However, the factorization of $F(t) + 1$ cannot be achieved by inspection. Rational approximation of $\sqrt{(t^2+1)/(t^2+4)}$ had been also extensively studied in [16]. The approximation was achieved by constructing an appropriate transformation from the whole real line to the unit interval. As a result, an approximate factorization has a small global error (10^{-12} on the real line). Here, we produce figure 1, which demonstrates the closedness of approximation on the whole complex plane.

Example 6.2. Let us consider rational approximation of the function

$$K = \sqrt{\frac{(t+2i)(t+3i)}{(t-2i)(t-3i)}} \tag{6.2}$$

with the index -1. Again, the function has been chosen to have the explicit exact factorization

$$\sqrt{\frac{(t+2i)(t+ki)}{(t-2i)(t-ki)}} = \frac{\sqrt{(t+2i)(t+ki)}}{t+i} \left(\frac{t-i}{t+i}\right)^{-1} \frac{t-i}{\sqrt{(t-2i)(t-ki)}}.$$

The function-theoretic factorization has growth at infinity, making it more difficult to approximate. Nevertheless, it can be rationally approximated and the error $|K - \tilde{K}|$ is presented in figure 2. Importantly, the error of the factors $|K_\pm - \tilde{K}_\pm|$ is also small (figure 3). For more details on rational approximation of complex-valued functions see [23].

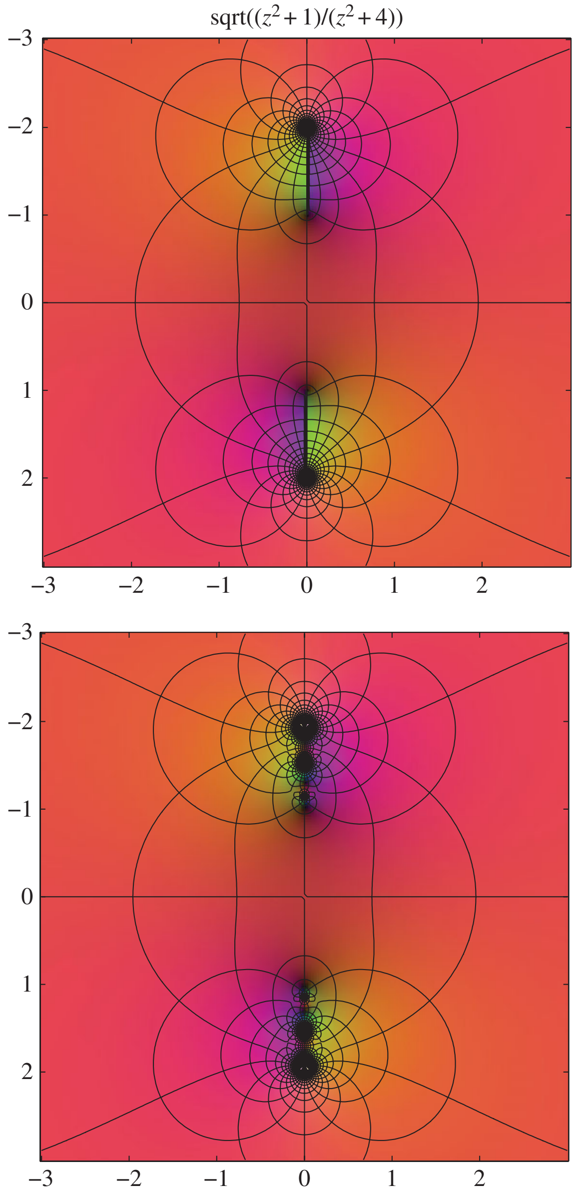

Figure 1. Contour lines for the real and imaginary parts of function F (6.1) and its rational approximation $\tilde{F}$. They are superimposed on a full colour image using a colour scheme developed by John Richardson. Red is real, blue is positive imaginary, green is negative imaginary, black is small magnitude and white is large magnitude. Branch cuts appear as colour discontinuities and coalescent contour lines. Produced using Matlab function `zviz.m`.

(b) Numerical matrix factorization

The stability result from §4 can be used in numerical computations. Two different examples are presented. For each example the Daniele–Khrapkov factorization is computed in two different ways. The first method is the direct use of a Cauchy integral to calculate the scalar factorization of $r_\pm$ and splitting $\theta_\pm$. So the initial matrix is exact, the factors have errors due to computation of Cauchy integrals. In the second method, the entries of the matrix are rationally approximated and for this matrix the exact Daniele–Khrapkov factorization is obtained. The matrix is approximate

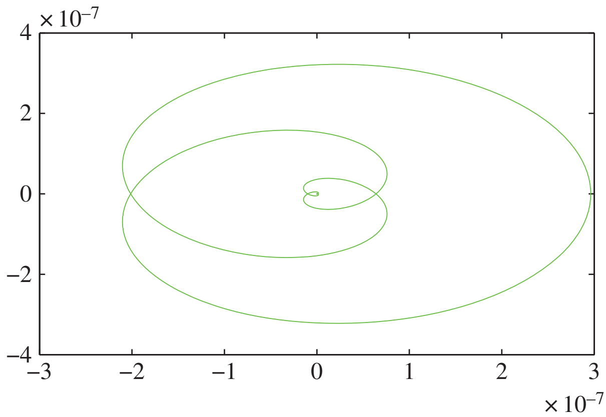

Figure 2. Error in approximating function K (6.2) by [8, 8] plotted as real against imaginary part. The accuracy of an approximation is denoted by the size of the disc the curve is contained in.

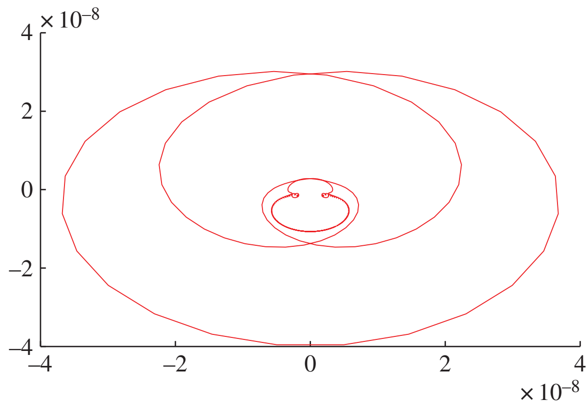

Figure 3. Error of factor $K_{\pm}$ on the real line plotted as real against imaginary part. The accuracy of an approximation is denoted by the size of the disc the curve is contained in.

but the factorization of this matrix is exact. The first method will be referred to as 'exact' and the second one as 'approximate' although the reader should note that both are approximate factorizations. The results of these two methods are then compared for each example.

The first example is

$$\mathbf{K}_1(t) = \mathbf{I} + \sqrt{\frac{t^2+1}{t^2+4}}\begin{pmatrix} 0 & 1 \\ -2 & 0 \end{pmatrix}.$$

The ideas is to rationally approximate $\sqrt{(t^2+1)/(t^2+4)}$ by f_N. Then the factorization of

$$\begin{pmatrix} 1 & f_N \\ cf_N & 1 \end{pmatrix},$$

is computed and compared with the 'exact' factorization. The advantage of such an approximation is that there is no need to use the Cauchy formula to find $r_{\pm}$ and $\theta_{\pm}$. Note that the approximate matrix has all rational entries and hence in theory factorization can be achieved using methods for rational matrix functions. But in practice the implemented

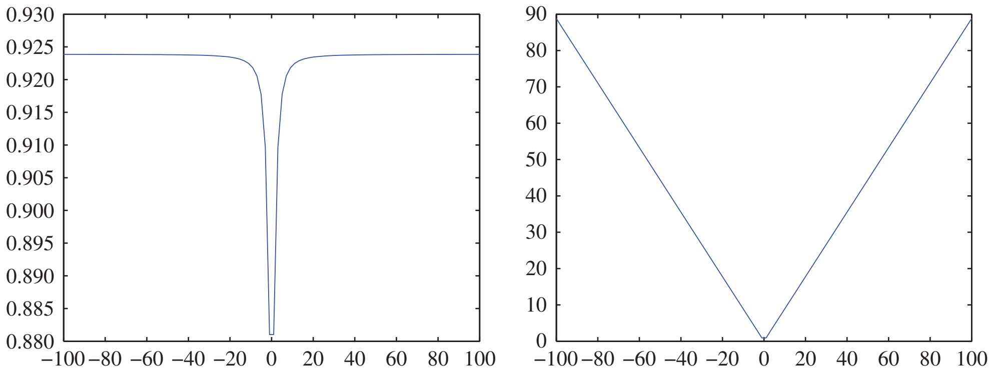

Figure 4. The modulus of K_{1+} and K_{2+} on the real line.

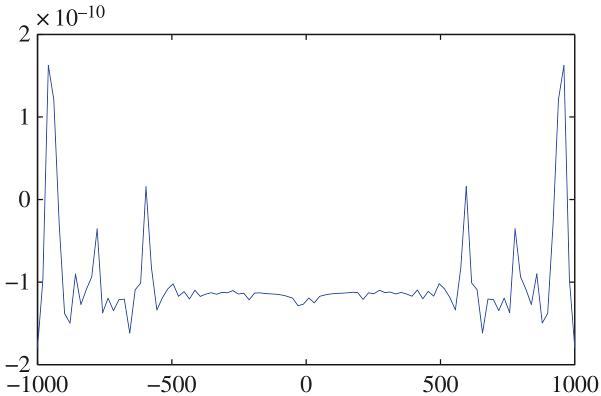

Figure 5. The modulus of the difference in a_{11} elements of 'exact' and 'approximate' factors for K_1.

procedures are unstable, making it impossible. At present, very few implemented Wiener–Hopf algorithms exist. For example, there have been some attempts recently [24] to produce numerical factorization algorithms for rational matrix functions and numerical algorithms for Riemann–Hilbert problems [25–27].

The second example is

$$\mathbf{K}_2(t) = \mathbf{I} + \sqrt{\frac{(t+2i)(t+i)}{(t-2i)(t-i)}}\begin{pmatrix} 0 & 1 \\ -2 & 0 \end{pmatrix}.$$

Similarly the approximate factorization is considered by approximating $\sqrt{((t+2i)(t+i))/((t-2i)(t-i))}$.

The difference in behaviour on the real line of the two examples can be seen in figure 4. This is because their partial indices are different. The first example have partial indices $\{0,0\}$ and the second $\{-1,-1\}$. These partial indices can be computed using the following identity

$$\begin{pmatrix} 1 & f \\ cf & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ c^{1/2} & -c^{1/2} \end{pmatrix}\begin{pmatrix} 1+c^{1/2}f & 0 \\ 0 & 1-c^{1/2}f \end{pmatrix}\begin{pmatrix} \frac{1}{2} & \frac{c^{-1/2}}{2} \\ \frac{1}{2} & \frac{-c^{-1/2}}{2} \end{pmatrix}.$$

The errors are compared in figures 5 and 6. It should be noted that the calculation of 'approximate' factors took significantly less computational time than the 'exact' factors. Besides the natural difference in magnitude of errors (due to the difference in errors of rational approximations),

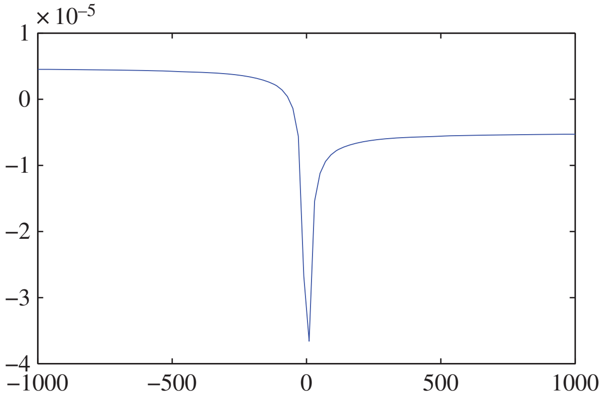

Figure 6. The modulus of the difference in a_{11} elements of 'exact' and 'approximate' factors for K_2.

the shape of the curves are dramatically different. It seems the error in figure 5 is random and in figure 6 is systemic. This suggests that in the first example, the error in 'exact' factorization is greater than 'approximate' factorization. So the accumulated errors in computing Cauchy integrals is greater than the error in once approximating entries of the matrix function. The reverse is true in the second example.

Data accessibility. Rational approximations are performed using Matlab package Chebfun. Figure 1 is produced using Matlab function `zviz.m`.

Acknowledgements. I am grateful for support from Prof. Nigel Peake. I benefited from useful discussions with Dr Rogosin, Prof. Speck and Prof. Spitkovsky. Suggestions of the anonymous referees helped to improve this paper.

Funding statement. This work was supported by the UK Engineering and Physical Sciences Research Council (EPSRC) grant no. EP/H023348/1 for the University of Cambridge Centre for Doctoral Training, the Cambridge Centre for Analysis.

Conflict of interests. I have no competing interests.

References

1. Gohberg I, Kaashoek MA, Spitkovsky IM. 2003 An overview of matrix factorization theory and operator applications. In *Factorization and integrable systems* (eds I Gohberg, N Manojlovic, A dos Santos). Operator theory: advances and applications, vol. 141, pp. 1–102. Basel, Switzerland: Birkhäuser. (doi:10.1007/978-3-0348-8003-9_1)
2. Gennady Mishuris SR. In preparation. Constructive methods for factorization of matrix-functions.
3. Litvinchuk GS, Spitkovskii IM. 1987 *Factorization of measurable matrix functions*. Operator theory: advances and applications, vol. 25. Basel, Switzerland: Birkhäuser. (Transl. from the Russian by Bernd Luderer, with a foreword by Bernd Silbermann). MR1015716 (90g:47030). (doi:10.1007/978-3-0348-6266-0)
4. Daniele V, Zich R. 2014 *The Wiener–Hopf method in electromagnetics*. Mario Boella series on electromagnetism in information and communication series. Raleigh, NC: SciTech Publishing.
5. Khrapkov AA. 1971 The first basic problem for a notch at the apex of an infinite wedge. *Int. J. Fracture Mech.* **7**, 373–382. (doi:10.1007/BF00189109)
6. Gohberg IC, Kreĭn MG. 1960 Systems of integral equations on a half line with kernels depending on the difference of arguments. *Am. Math. Soc. Transl.* (2) **14**, 217–287.
7. Abrahams ID. 1996 Radiation and scattering of waves on an elastic half-space; a non-commutative matrix Wiener-Hopf problem. *J. Mech. Phys. Solids* **44**, 2125–2154. (doi:10.1016/S0022-5096(96)00064-6)
8. Veitch BH, Abrahams ID. 2007 On the commutative factorization of $n \times n$ matrix Wiener-Hopf kernels with distinct eigenvalues. *Proc. R. Soc. A* **463**, 613–639. (doi:10.1098/rspa.2006.1780)

9. Lawrie JB, David Abrahams I. 2007 A brief historical perspective of the Wiener-Hopf technique. *J. Eng. Math.* **59**, 351–358. (doi:10.1007/s10665-007-9195-x)
10. Heaton CJ, Peake N. 2005 Acoustic scattering in a duct with mean swirling flow. *J. Fluid Mech.* **540**, 189–220. (doi:10.1017/S0022112005005719)
11. Veitch B, Peake N. 2008 Acoustic propagation and scattering in the exhaust flow from coaxial cylinders. *J. Fluid Mech.* **613**, 275–307. (doi:10.1017/S0022112008003169)
12. David Abrahams I, Davis AMJ, Llewellyn Smith SG. 2008 Matrix Wiener-Hopf approximation for a partially clamped plate. *Q. J. Mech. Appl. Math.* **61**, 241–265. (doi:10.1093/qjmam/hbn004)
13. Mishuris G, Rogosin S. 2014 An asymptotic method of factorization of a class of matrix functions. *Proc. R. Soc. A* **470**, 20140109. (doi:10.1098/rspa.2014.0109)
14. Mishuris G, Rogosin S. 2015 Factorization of a class of matrix-functions with stable partial indices. (http://arxiv.org/abs/1501.02692)
15. Crighton DG. 2001 Asymptotic factorization of Wiener-Hopf kernels. *Wave Motion* **33**, 51–65. (doi:10.1016/S0165-2125(00)00063-9)
16. Kisil AV. 2013 A constructive method for an approximate solution to scalar Wiener–Hopf equations. *Proc. R. Soc. A* **469**, 20120721. (doi:10.1098/rspa.2012.0721)
17. Abrahams DI. 1997 On the solution of Wiener-Hopf problems involving noncommutative matrix kernel decompositions. *SIAM J. Appl. Math.* **57**, 541–567. (doi:10.1137/S0036139995287673)
18. Câmara MC, Lebre AB, Speck F-O. 1992 Meromorphic factorization, partial index estimates and elastodynamic diffraction problems. *Math. Nachr.* **157**, 291–317. (doi:10.1002/mana.19921570124)
19. Gahov FD, Čerskiĭ JuI. 1978 Equations of convolution type. Moscow, Russia: Nauka. [In Russian.]
20. Rawlins AD. 1997 Two waveguide trifurcation problems. *Math. Proc. Camb. Philos. Soc.* **121**, 555–573. (doi:10.1017/S0305004196001296)
21. Prössdorf S, Speck F-O. 1990 A factorisation procedure for two by two matrix functions on the circle with two rationally independent entries. *Proc. R. Soc. Edinb. Sect. A* **115**, 119–138. (doi:10.1017/S0308210500024616)
22. David Abrahams I. 2000 The application of Padé approximations to Wiener-Hopf factorization. *IMA J. Appl. Math.* **65**, 257–281. (doi:10.1093/imamat/65.3.257)
23. Trefethen LN. 1981 Rational chebyshev approximation on the unit disc. *Numer. Math.* **37**, 297–320. (doi:10.1007/BF01398258)
24. Conceicao AC, Kravchenko VG, Pereira JC. 2012 Rational functions factorization algorithm: a symbolic computation for the scalar and matrix cases. In *Proc. 1st Natl Conf. on Symbolic Computation in Education and Research* (*CD-ROM*), *Lisbon, 2–3 April*.
25. Olver S, Trogdon T. 2014 Nonlinear steepest descent and numerical solution of Riemann-Hilbert problems. *Comm. Pure Appl. Math.* **67**, 1353–1389. (doi:10.1002/cpa.21497)
26. Fokas AS, Its AR, Kapaev AA, Novokshenov VYu. 2006 The Riemann-Hilbert approach. In *Painlevé transcendents*. Mathematical surveys and monographs, vol. 128. Providence, RI: American Mathematical Society. MR2264522 (2010e:33030). (doi:10.1090/surv/128)
27. Kisil AV. In press. On the relationship between a strip Wiener-Hopf problem and a Riemann-Hilbert problem. *IMA J. Appl. Math.*