Journal of Mathematics Research; Vol. 7, No. 4; 2015
ISSN 1916-9795 E-ISSN 1916-9809
Published by Canadian Center of Science and Education

Residual Symmetry for Linear Algebraic Equations and the Lorentz-Group Algorithm

Chein-Shan Liu[1,2]

[1] Department of Civil Engineering, National Taiwan University, Taipei 106-17, Taiwan

[2] International Center for Simulation Software in Engineering and Sciences, College of Mechanics and Materials, Hohai University, Nanjing, Jiangsu 210098, China

Correspondence: Chein-Shan Liu. E-mail: liucs@ntu.edu.tw

Received: September 18, 2015 Accepted: October 13, 2015 Online Published: October 27, 2015

doi:10.5539/jmr.v7n4p74 URL: http://dx.doi.org/10.5539/jmr.v7n4p74

Abstract

In the iterative solution of n linear algebraic equations $\mathbf{Bx} = \mathbf{b}$ by using the steepest descent method, i.e., $\mathbf{x}_{k+1} = \mathbf{x}_k - \alpha_k \mathbf{B}^{\mathrm{T}}\mathbf{r}_k$, it is known that the steplength $\alpha_k := \mathbf{r}_k^{\mathrm{T}}\mathbf{A}\mathbf{r}_k/\|\mathbf{A}\mathbf{r}_k\|^2$ causes a slow convergence, where $\mathbf{r} = \mathbf{Bx} - \mathbf{b}$ is the residual vector and $\mathbf{A} = \mathbf{B}\mathbf{B}^{\mathrm{T}}$. In this paper we study the residual symmetry of the residual dynamics for a scaled residual vector $\mathbf{y} \in \mathbb{S}^{n-1}_{\|\mathbf{r}_0\|}$, which as expressed in the augmented space is a nonlinear Lorentzian dynamical system, and is endowed with a cone structure in the Minkowski space with the Lorentz group $SO_o(n,1)$ being the internal symmetry group. Consequently, we can modify the steplength to $\alpha_k = \mathbf{y}_k^{\mathrm{T}}\mathbf{A}\mathbf{y}_k/\|\mathbf{A}\mathbf{y}_k\|^2$ with $\mathbf{y}_k$ being computed by a Lorentz group algorithm (LGA) based on $SO_o(n,1)$, which can significantly improve the convergence speed and enhance the stability. Several linear inverse problems are used to assess the numerical performance of the LGA.

Keywords: ill-posed linear system, residual dynamics, future cone, internal symmetry, Jordan dynamics, linear inverse problems, the Lorentz group algorithm

1. Introduction

In this paper we study the *internal symmetry* of the *residual dynamics*, abbreviated as *residual symmetry*, which is defined in terms of the residual vector:

$$\mathbf{r}(\mathbf{x}) = \mathbf{Bx} - \mathbf{b}, \tag{1}$$

for a system of linear algebraic equations (LAEs):

$$\mathbf{Bx} = \mathbf{b}, \tag{2}$$

where $\mathbf{x} \in \mathbb{R}^n$ is an unknown vector to be determined from a given coefficient matrix $\mathbf{B} \in \mathbb{R}^{n\times n}$ and the input vector $\mathbf{b} \in \mathbb{R}^n$. Equation (2) is in general an ill-posed system if it is derived from the numerical solution of linear inverse problems.

The relaxed steepest descent method (RSDM) used to solve Equation (2) is given by Liu (2011a, 2012a):
(i) Give an initial guess of $\mathbf{x}_0$, and then $\mathbf{R}_0 = \mathbf{C}\mathbf{x}_0 - \mathbf{b}_1$.
(ii) For $k = 0, 1, 2 \ldots$, we repeat the following computations:

$$\mathbf{x}_{k+1} = \mathbf{x}_k - (1-\gamma)\frac{\|\mathbf{R}_k\|^2}{\mathbf{R}_k^{\mathrm{T}}\mathbf{C}\mathbf{R}_k}\mathbf{R}_k, \tag{3}$$

$$\mathbf{R}_{k+1} = \mathbf{C}\mathbf{x}_{k+1} - \mathbf{b}_1. \tag{4}$$

If $\|\mathbf{R}_{k+1}\| < \varepsilon$ for a prescribed convergence criterion ε then stop; otherwise, go to step (ii). In the above, $\mathbf{C} = \mathbf{B}^{\mathrm{T}}\mathbf{B}$, $\mathbf{b}_1 = \mathbf{B}^{\mathrm{T}}\mathbf{b}$, $\mathbf{R}_k = \mathbf{B}^{\mathrm{T}}\mathbf{r}_k = \mathbf{B}^{\mathrm{T}}\mathbf{r}(\mathbf{x}_k)$, $0 \le \gamma < 1$ is a relaxed parameter, and

$$\alpha_k := \frac{\|\mathbf{R}_k\|^2}{\mathbf{R}_k^{\mathrm{T}}\mathbf{C}\mathbf{R}_k} = \frac{\mathbf{r}_k^{\mathrm{T}}\mathbf{A}\mathbf{r}_k}{\|\mathbf{A}\mathbf{r}_k\|^2} \tag{5}$$

is the steplength for the steepest descent method (SDM), where $\mathbf{A} = \mathbf{B}\mathbf{B}^{\mathrm{T}}$.

To solve the ill-posed linear problems, there are several methods developed after the pioneering work of Tikhonov & Arsenin (1977). To account of the sensitivity to noise it is often used a regularization method to solve the ill-posed problem [Kunisch & Jou (1998); Wang & Xiao (2001); Xie & Jou (2002); Resmerita (2005)], where a suitable regularization parameter is used to depress the bias in the computed solution by a better balance of approximation error and propagated data error. Liu (2012b) has developed an optimally scaled vector regularization method, the dynamic Tikhonov regularization

method [Liu (2013a)], a two-side conditioning method [Liu (2013b)], as well as the Krylov subspace double-optimization method [Liu (2013c); Liu (2014)] to solve the ill-posed linear problems.

There are many methods that converge significantly faster than the SDM, unlike that of the conjugate gradient method (CGM), they insist their search directions to be the gradient vector at each iteration [Barzilai & Borwein (1988); Friedlander et al. (1999); Raydan & Svaiter (2002); Dai & Yuan (2003); Dai et al. (2006)]. The SDM performs poorly, yielding iteration counts that grow linearly with Cond(**C**) [Akaike (1959); Forsythe (1968); Nocedal et al. (2002)]. Several modifications to the SDM have been made, and they have stimulated a new interest in the SDM because it is recognized that the gradient vector itself is not a bad choice of the solution direction, but rather that the steplength in Equation (5) originally used by the SDM is to blame for the slow convergence behavior.

An interesting problem is that there exists a certain symmetry of the iterative dynamics of the LAEs when we solve them by the iterative method. What is the underlying space of the residual dynamics and what sort Lie-group is acting on that space. In this paper we try to answer these problems and propose a theoretical foundation to modify the steplength in Equation (5) from the Lorentz-group symmetry which is acting on a future cone in the Minkowski space.

The remaining part of this paper is arranged as follows. In Section 2 we start from a future cone in the Minkowski space to derive a system of nonlinear ordinary differential equations (ODEs) for the numerical solution of Equation (2). Then, the residual dynamics on the future cone is constructed in Section 3, resulting to a Jordan dynamics and a nonlinear Lorentzian dynamical system, of which the residual symmetry is developed. Accordingly, the Lorentz-group algorithm (LGA) used to solve Equation (2) is developed in Section 4. The numerical examples of linear inverse problems are given in Section 5 to display some advantages of the LGA. Finally, the conclusions are drawn in Section 6.

2. A Future Cone in the Minkowski Space

2.1 Nonlinear ODEs for **x**

For Equation (1) we can introduce an invariant manifold:

$$h(\mathbf{x},t)=\frac{1}{2}Q(t)\|\mathbf{r}(\mathbf{x})\|^2-\frac{1}{2}\|\mathbf{r}_0\|^2=0, \tag{6}$$

where we let $\mathbf{x}$ be a function of a time-like variable t, with the initial values of $\mathbf{x}(0)=\mathbf{x}_0$ and $\mathbf{r}_0=\mathbf{r}(\mathbf{x}_0)$ being given, and $Q(t)>0$, with $Q(0)=1$, is a monotonically increasing function of t. In terms of

$$\mathbf{X}=\left[\begin{array}{c}\frac{\mathbf{r}}{\|\mathbf{r}_0\|}\\ \frac{1}{\sqrt{Q(t)}}\end{array}\right], \tag{7}$$

Equation (6) represents a positive cone:

$$\mathbf{X}^{\mathrm{T}}\mathbf{g}\mathbf{X}=0 \tag{8}$$

in the Minkowski space $\mathbb{M}^{n+1}$, which is endowed with an indefinite Minkowski metric tensor:

$$\mathbf{g}=\left[\begin{array}{cc}\mathbf{I}_n & \mathbf{0}_{n\times 1}\\ \mathbf{0}_{1\times n} & -1\end{array}\right], \tag{9}$$

where $\mathbf{I}_n$ is the $n\times n$ identity matrix. Because the last component $1/\sqrt{Q(t)}$ of $\mathbf{X}$ is positive, the cone in Equation (8) is a future cone [Liu (2001)] as shown in Fig. 1.

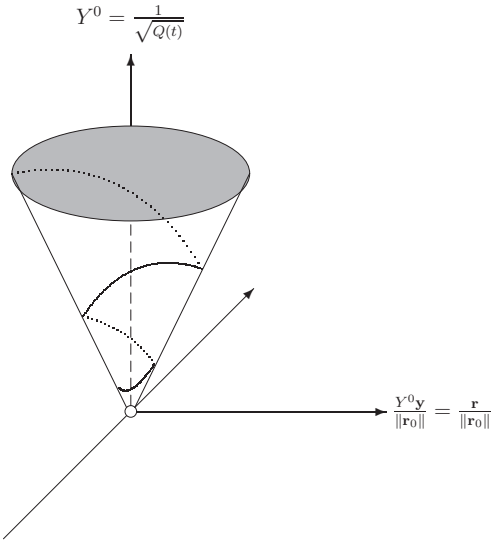

Figure 1. The construction of a future cone in the Minkowski space and the Lorentzian residual dynamics for solving the ill-posed linear system signifies a conceptual breakthrough.

The cone structure about the residual vector was first pointed out by Liu (2012c). However, the residual dynamics is not yet clearly developed in that paper. In Section 3 we will derive a nonlinear Lorentzian dynamical system for $\mathbf{X}$.

When $Q > 0$, the manifold defined by Equation (6) is continuous and differentiable, and by the consistency condition we have

$$\frac{1}{2}\dot{Q}(t)\|\mathbf{r}(\mathbf{x})\|^2 + Q(t)\mathbf{R}\cdot\dot{\mathbf{x}} = 0, \tag{10}$$

where

$$\mathbf{R} := \mathbf{B}^{\mathrm{T}}\mathbf{r} \tag{11}$$

is the steepest descent vector. Equation (10) is obtained by taking the differential of Equation (6) with respect to t and considering $\mathbf{x} = \mathbf{x}(t)$ and $h(\mathbf{x}, t) = 0$ for all t.

We suppose that the evolution of $\mathbf{x}$ is driving by $\mathbf{R}$ as like the SDM:

$$\dot{\mathbf{x}} = \lambda\mathbf{R}, \tag{12}$$

inserting which into Equation (10) we can derive a nonlinear ODEs system for $\mathbf{x}$:

$$\dot{\mathbf{x}} = -q(t)\frac{\|\mathbf{r}\|^2}{\mathbf{r}^{\mathrm{T}}\mathbf{A}\mathbf{r}}\mathbf{B}^{\mathrm{T}}\mathbf{r}, \tag{13}$$

where

$$\mathbf{A} := \mathbf{B}\mathbf{B}^{\mathrm{T}}, \tag{14}$$

$$q(t) := \frac{\dot{Q}(t)}{2Q(t)} > 0. \tag{15}$$

Hence, if $Q(t)$ can be guaranteed to be a monotonically increasing function of t, we have an absolutely convergent algorithm to solve Equation (2):

$$\|\mathbf{r}(\mathbf{x})\|^2 = \frac{\|\mathbf{r}_0\|^2}{Q(t)} =: \frac{C}{Q(t)},\ C := \|\mathbf{r}_0\|^2 > 0, \tag{16}$$

wherein we can observe that the path of $\mathbf{X}$ gradually moves down to the vertex point along the cone defined by Equation (8) as schematically shown in Fig. 1. However, it is still a great challenge by developing a suitable numerical integrator to solve the nonlinear ODEs in Equation (13), such that the orbit of $\mathbf{x}$ can really retain on the future cone in the Minkowski space. This important issue is addressed in the next two sections.

2.2 Discretizing and Keeping $\mathbf{x}$ *on the Manifold*

From Equation (13) we can observe that it is utmost important to study the residual dynamics of $\mathbf{r}$ in order to have a better understanding of the solution behavior of $\mathbf{x}$. In order to keep $\mathbf{x}$ on the manifold (16) we can consider the evolution of $\mathbf{r}(t) := \mathbf{r}(\mathbf{x}(t))$ along the path $\mathbf{x}(t),\ t \geq 0$ generated from Equation (13) by

$$\dot{\mathbf{r}} = \mathbf{B}\dot{\mathbf{x}} = -q(t)\frac{\|\mathbf{r}\|^2}{\mathbf{r}^{\mathrm{T}}\mathbf{A}\mathbf{r}}\mathbf{A}\mathbf{r}. \tag{17}$$

Now we discretize both Equations (13) and (17) into discrete dynamics by applying the forward Euler scheme:

$$\mathbf{x}(t+\Delta t) = \mathbf{x}(t) - \beta\frac{\|\mathbf{r}\|^2}{\mathbf{r}^{\mathrm{T}}\mathbf{A}\mathbf{r}}\mathbf{B}^{\mathrm{T}}\mathbf{r}, \tag{18}$$

$$\mathbf{r}(t+\Delta t) = \mathbf{r}(t) - \beta\frac{\|\mathbf{r}\|^2}{\mathbf{r}^{\mathrm{T}}\mathbf{A}\mathbf{r}}\mathbf{A}\mathbf{r}, \tag{19}$$

where

$$\beta = q(t)\Delta t \tag{20}$$

is a time-varying steplength.

In order to keep $\mathbf{x}$ on the invariant manifold (16), we can take the square-norm of both sides of Equation (19) and use Equation (16) to obtain

$$\frac{C}{Q(t+\Delta t)} = \frac{C}{Q(t)} - 2\beta\frac{C}{Q(t)} + \beta^2\frac{C}{Q(t)}\frac{\|\mathbf{r}\|^2}{(\mathbf{r}^{\mathrm{T}}\mathbf{A}\mathbf{r})^2}\|\mathbf{A}\mathbf{r}\|^2. \tag{21}$$

Thus, by dividing both sides by $C/Q(t)$ and reminding $Q(t) > 0$ the following scalar equation can be derived:

$$a^0\beta^2 - 2\beta + 1 - \frac{Q(t)}{Q(t+\Delta t)} = 0, \tag{22}$$

where

$$a^0 := \frac{\|\mathbf{r}\|^2\|\mathbf{A}\mathbf{r}\|^2}{(\mathbf{r}^{\mathrm{T}}\mathbf{A}\mathbf{r})^2} \geq 1. \tag{23}$$

Upon letting

$$\frac{Q(t)}{Q(t+\Delta t)} = 1 - \frac{1-\gamma^2}{a^0}, \tag{24}$$

the solution of β in Equation (22) is found to be

$$\beta = \frac{1-\gamma}{a^0}, \tag{25}$$

where $0 \leq \gamma < 1$ is a relaxed parameter, and $0 < \beta \leq 1$.

3. The Residual Dynamics on the Future Cone

In this section we are going to disclose some essential structures of the residual dynamics in terms of a scaled residual vector.

3.1 A Perpendicular Operator

Let

$$\mathbf{y} = \mathbf{r}\sqrt{Q(t)} \tag{26}$$

be a scaled residual vector, and from Equation (16) we know that $\mathbf{y} \in \mathbb{S}^{n-1}_{\|\mathbf{r}_0\|}$ with a radius $\|\mathbf{r}_0\|$, i.e.,

$$\|\mathbf{y}(t)\| = \|\mathbf{r}_0\|. \tag{27}$$

This equation is an invariant condition of the residual dynamical system. At the same time, by using Equation (26), Equations (18) and (23) become

$$\mathbf{x}(t+\Delta t) = \mathbf{x}(t) - \beta\frac{\|\mathbf{y}\|^2}{\mathbf{y}^{\mathrm{T}}\mathbf{A}\mathbf{y}}\mathbf{B}^{\mathrm{T}}\mathbf{r}, \tag{28}$$

$$a^0 := \frac{\|\mathbf{y}\|^2\|\mathbf{A}\mathbf{y}\|^2}{(\mathbf{y}^{\mathrm{T}}\mathbf{A}\mathbf{y})^2}, \tag{29}$$

where we do not modify the steepest descent vector, keeping it to be $\mathbf{R} = \mathbf{B}^{\mathrm{T}}\mathbf{r}$, but we modify $\|\mathbf{r}\|^2/\mathbf{r}^{\mathrm{T}}\mathbf{A}\mathbf{r}$ to $\|\mathbf{y}\|^2/\mathbf{y}^{\mathrm{T}}\mathbf{A}\mathbf{y}$.

Now we derive the governing ODEs for $\mathbf{y}$. From Equations (17) and (26) we can derive

$$\dot{\mathbf{y}} = q(t)\left[\mathbf{I}_n - \frac{\|\mathbf{y}\|^2}{\mathbf{y}^{\mathrm{T}}\mathbf{A}\mathbf{y}}\mathbf{A}\right]\mathbf{y}, \tag{30}$$

which is the first form of the residual dynamics. If we define the following operator:

$$\mathbf{D} = \mathbf{I}_n - \frac{\|\mathbf{y}\|^2}{\mathbf{y}^{\mathrm{T}}\mathbf{A}\mathbf{y}}\mathbf{A}, \tag{31}$$

we have

$$\dot{\mathbf{y}} = q(t)\mathbf{D}\mathbf{y}, \tag{32}$$

where $\mathbf{D}$ satisfies the following properties:

$$\mathbf{D}^{\mathrm{T}} = \mathbf{D}, \quad \mathbf{y}^{\mathrm{T}}\mathbf{D}\mathbf{y} = 0, \tag{33}$$

due to $\mathbf{A}^{\mathrm{T}} = \mathbf{A}$ and

$$\mathbf{y}^{\mathrm{T}}\left[\mathbf{I}_n - \frac{\|\mathbf{y}\|^2}{\mathbf{y}^{\mathrm{T}}\mathbf{A}\mathbf{y}}\mathbf{A}\right]\mathbf{y} = 0.$$

Indeed, $\mathbf{D}$ is a symmetric perpendicular operator, because it maps $\mathbf{y}$ to a new vector $\mathbf{D}\mathbf{y}$ which is perpendicular to $\mathbf{y}$ itself in view of the second equation in Equation (33).

3.2 The Jordan Dynamics

Here we give a new aspect of the residual dynamics (30) derived for Equation (2) from the Jordan structure [Iordanescu (2007)]. There exist three kinds of Jordan structures: algebras, triple systems, and pairs. We will use the Jordan algebra as developed by Liu (2000) which with the use of triplet vectors was developed for a general purpose of dynamical system, including a conservative force and a dissipative force on the right-hand side. Since the creation of Jordan algebras by Pasqual Jordan, an improvement of the mathematical foundation of quantum mechanics was made. In the meantime, the Jordan structures have been intensively studied by many mathematicians, and a lot of important results have been obtained. The study of Jordan structures and their applications is at the present time a wide-spreading and interesting field of mathematical researches [Iordanescu (2009)].

Liu (2000) has derived a system of ODEs based on the Jordan algebra:

$$\dot{\mathbf{x}} = [\mathbf{w}, \mathbf{z}, \mathbf{u}] = \mathbf{w} \cdot \mathbf{z}\mathbf{u} - \mathbf{u} \cdot \mathbf{z}\mathbf{w}. \tag{34}$$

The triplet vectors $\mathbf{w}$, $\mathbf{z}$ and $\mathbf{u}$ are functions of $\mathbf{x}$ and t.

In terms of the Jordan dynamics in Equation (34) we can rewrite Equation (30) as

$$\dot{\mathbf{y}} = q(t)\left[\frac{\mathbf{A}\mathbf{y}}{\mathbf{y} \cdot \mathbf{A}\mathbf{y}}, \mathbf{y}, \mathbf{y}\right], \tag{35}$$

which is the second form of the residual dynamics.

3.3 The Lorentz Group $SO_o(n, 1)$

Obviously, by letting

$$\dot{\mathbf{q}}(t) = -\frac{q(t)\mathbf{A}\mathbf{y}}{\mathbf{y} \cdot \mathbf{A}\mathbf{y}}, \tag{36}$$

Equation (35) can be rearranged to

$$\dot{\mathbf{y}} + (\dot{\mathbf{q}} \cdot \mathbf{y})\mathbf{y} = \|\mathbf{y}\|^2\dot{\mathbf{q}}. \tag{37}$$

By using Equation (27) we can further write

$$\frac{d}{dt}\left(\frac{\mathbf{y}}{\|\mathbf{r}_0\|}\right) + (\dot{\mathbf{q}} \cdot \mathbf{y})\frac{\mathbf{y}}{\|\mathbf{r}_0\|} = \|\mathbf{r}_0\|\dot{\mathbf{q}}. \tag{38}$$

Defining the integrating factor Y^0 by

$$Y^0 := \exp\left[\int_0^t [\dot{\mathbf{q}}(\xi) \cdot \mathbf{y}(\xi)]d\xi\right], \tag{39}$$

and from Equation (38) we have

$$\frac{d}{dt}\left(Y^0\frac{\mathbf{y}}{\|\mathbf{r}_0\|}\right) = \|\mathbf{r}_0\|Y^0\dot{\mathbf{q}}. \tag{40}$$

On the other hand, by taking the time differential of Equation (39) we have

$$\dot{Y}^0 = Y^0\dot{\mathbf{q}}\cdot\mathbf{y}. \tag{41}$$

Equations (40) and (41) can be written together into a matrix form:

$$\dot{\mathbf{Y}} = \mathbf{C}\mathbf{Y}, \tag{42}$$

where

$$\mathbf{Y} = \begin{bmatrix} \mathbf{Y}^s \\ Y^0 \end{bmatrix} := \begin{bmatrix} Y^0\frac{\mathbf{y}}{\|\mathbf{r}_0\|} \\ Y^0 \end{bmatrix} \tag{43}$$

is an augmented vector, and the following system coefficient matrix:

$$\mathbf{C} := \begin{bmatrix} \mathbf{0}_n & \mathbf{C}_0^s \\ (\mathbf{C}_0^s)^{\mathrm{T}} & 0 \end{bmatrix} = \begin{bmatrix} \mathbf{0}_n & \|\mathbf{r}_0\|\dot{\mathbf{q}} \\ \|\mathbf{r}_0\|\dot{\mathbf{q}}^{\mathrm{T}} & 0 \end{bmatrix} \tag{44}$$

belongs to a Lie algebra of the proper orthochronous Lorentz-group $SO_o(n,1)$, satisfying

$$\mathbf{C}^{\mathrm{T}}\mathbf{g} + \mathbf{g}\mathbf{C} = \mathbf{0}, \tag{45}$$

with $\mathbf{g}$ defined by Equation (9). In this sense, Equation (42) is a nonlinear Lorentzian dynamical system [Liu (2002)], which is the third form of the residual dynamics.

3.4 The Cone Condition

Now we prove that

$$Y^0 = \frac{1}{\sqrt{Q(t)}}, \quad \mathbf{Y} = \mathbf{X}, \tag{46}$$

where $\mathbf{X}$ is defined by Equation (7).

From Equations (39), (36) and (15) it follows that

$$Y^0 = \exp\left[\int_0^t -q(\xi)d\xi\right] = \exp\left[\int_0^t \frac{-\dot{Q}(\xi)}{2Q(\xi)}d\xi\right] = \frac{1}{\sqrt{Q(t)}}, \tag{47}$$

where we have used $Q(0) = 1$. Then inserting it into Equation (43) and by Equations (26) and (7) we can obtain

$$\mathbf{Y} = \begin{bmatrix} \frac{\mathbf{y}}{\sqrt{Q(t)}\|\mathbf{r}_0\|} \\ \frac{1}{\sqrt{Q(t)}} \end{bmatrix} = \begin{bmatrix} \frac{\mathbf{r}}{\|\mathbf{r}_0\|} \\ \frac{1}{\sqrt{Q(t)}} \end{bmatrix} = \mathbf{X}. \tag{48}$$

Inserting $\mathbf{Y}$ into the cone condition (8) we can prove that

$$\mathbf{Y}^{\mathrm{T}}\mathbf{g}\mathbf{Y} = (Y^0)^2\left[\frac{\|\mathbf{y}\|^2}{\|\mathbf{r}_0\|^2} - 1\right] = 0, \tag{49}$$

and thus

$$\mathbf{Y}^{\mathrm{T}}\mathbf{g}\mathbf{Y} = 0 \Leftrightarrow \|\mathbf{y}\| = \|\mathbf{r}_0\|. \tag{50}$$

It means that the cone condition in the Minkowski space $\mathbb{M}^{n+1}$ and the invariant condition of $\|\mathbf{y}\| = \|\mathbf{r}_0\|$ in the Euclidean space $\mathbb{E}^n$ are equivalent. The above cone condition is further shown in Fig. 1 by the augmented vector $\mathbf{Y}$. Consequently, we have derived a genuine residual dynamics as expressed by Equations (42) and (48) on the future cone in the Minkowski space with $\mathbf{Y} \in \mathbb{M}^{n+1}$.

3.5 The Fixed Point

We can prove that the eigenvector $\mathbf{y}$ of the matrix $\mathbf{A}$, i.e.,

$$\mathbf{A}\mathbf{y} = \lambda\mathbf{y}, \tag{51}$$

is a fixed point of the dynamical system (35). Inserting the above equation into Equation (35) and using Equation (34) we can derive

$$\begin{aligned}\dot{\mathbf{y}} &= q(t)\left[\frac{\lambda\mathbf{y}}{\lambda\|\mathbf{y}\|^2},\mathbf{y},\mathbf{y}\right]\\ &= q(t)\left[\frac{\mathbf{y}}{\|\mathbf{y}\|^2},\mathbf{y},\mathbf{y}\right]\\ &= q(t)\left[\frac{\mathbf{y}\cdot\mathbf{y}}{\|\mathbf{y}\|^2}\mathbf{y}-\mathbf{y}\cdot\mathbf{y}\frac{\mathbf{y}}{\|\mathbf{y}\|^2}\right]=\mathbf{0}.\end{aligned} \tag{52}$$

Thus when the orbit of $\mathbf{y}$ tends to the eigenvector, we have $\dot{\mathbf{y}}=\mathbf{0}$, which means that the eigenvector is a critical point of the dynamical system (35).

4. The Lorentz-Group Algorithm Based on $SO_o(n,1)$

As that done by Liu (2001), we can develop the group-preserving scheme for the nonlinear Lorentzian system (42). The numerical scheme provides a medium to calculate the value of $\mathbf{Y}$ at time $t=t_{k+1}$ when $\mathbf{Y}$ is already known at time $t=t_k$. The evolution of $\mathbf{Y}$ is governed by the dynamical law (42) with the coefficient matrix $\mathbf{C}$ given by Equation (44). Unluckily, due to the presence of $\dot{\mathbf{q}}$ in the coefficient matrix $\mathbf{C}$, which as shown in Equation (36) is a function of t and $\mathbf{y}=\|\mathbf{r}_0\|\mathbf{Y}^s/Y^0$, we need to approximate the solution of the dynamical law (42) by considering Y^0 and $\mathbf{Y}^s$ to be constant in each single time step with a stepsize $\Delta t=t_{k+1}-t_k$. So the solution of Equation (42) within one-step is known to be

$$\mathbf{Y}(k+1)=\mathbf{G}(k)\mathbf{Y}(k), \tag{53}$$

where

$$\mathbf{G}(k):=\exp[\Delta t\mathbf{C}(k)]=\begin{bmatrix}\mathbf{I}_n+\frac{a(k)-1}{\|\mathbf{C}_0^s(k)\|^2}\mathbf{C}_0^s(k)(\mathbf{C}_0^s)^{\mathrm{T}}(k) & \frac{b(k)}{\|\mathbf{C}_0^s(k)\|}\mathbf{C}_0^s(k)\\ \frac{b(k)}{\|\mathbf{C}_0^s(k)\|}(\mathbf{C}_0^s)^{\mathrm{T}}(k) & a(k)\end{bmatrix}, \tag{54}$$

in which

$$a(k):=\cosh(\Delta t\|\mathbf{C}_0^s(k)\|),\quad b(k):=\sinh(\Delta t\|\mathbf{C}_0^s(k)\|). \tag{55}$$

A numerical algorithm to solve Equation (42) is called a group-preserving scheme (GPS) if for every time increment the mapping $\mathbf{G}(k)$ from $\mathbf{Y}(k)$ to $\mathbf{Y}(k+1)$ can preserve the following group properties [Liu (2001)]:

$$\mathbf{G}^{\mathrm{T}}\mathbf{g}\mathbf{G}=\mathbf{g},\quad \det\mathbf{G}=1,\quad G_0^0\geq 1, \tag{56}$$

where det is the shorthand of determinant, and G_0^0 is the 00-th mixed component of $\mathbf{G}$. In Equation (54), $G_0^0=a(k)>1$.

Based on the above results and Equations (28), (29) and (25) we can develop the Lorentz group algorithm (LGA) to solve Equation (2):
(i) Give $0\leq\gamma<1$ and an initial guess of $\mathbf{x}_0$.
(ii) Calculate $\mathbf{r}_0=\mathbf{B}\mathbf{x}_0-\mathbf{b}$ and $\mathbf{y}_0=\mathbf{r}_0$.
(iii) For $k=1,2\ldots$, we repeat the following computations:

$$\begin{aligned}\mathbf{x}_k &= \mathbf{x}_{k-1}-(1-\gamma)\frac{\mathbf{y}_{k-1}^{\mathrm{T}}\mathbf{A}\mathbf{y}_{k-1}}{\|\mathbf{A}\mathbf{y}_{k-1}\|^2}\mathbf{B}^{\mathrm{T}}\mathbf{r}_{k-1},\\ a_{k-1}^0 &= \frac{\|\mathbf{y}_{k-1}\|^2\|\mathbf{A}\mathbf{y}_{k-1}\|^2}{(\mathbf{y}_{k-1}^{\mathrm{T}}\mathbf{A}\mathbf{y}_{k-1})^2},\\ \beta_{k-1} &= \frac{1-\gamma}{a_{k-1}^0},\\ a_{k-1} &= \cosh(\beta_{k-1}\sqrt{a_{k-1}^0}),\\ b_{k-1} &= \sinh(\beta_{k-1}\sqrt{a_{k-1}^0}),\\ \mathbf{y}_k &= \frac{\mathbf{y}_{k-1}+[(a_{k-1}-1)(\mathbf{y}_{k-1}^{\mathrm{T}}\mathbf{A}\mathbf{y}_{k-1})/\|\mathbf{A}\mathbf{y}_{k-1}\|^2-b_{k-1}\|\mathbf{y}_{k-1}\|/\|\mathbf{A}\mathbf{y}_{k-1}\|]\mathbf{A}\mathbf{y}_{k-1}}{a_{k-1}-b_{k-1}/\sqrt{a_{k-1}^0}},\\ \mathbf{r}_k &= \mathbf{B}\mathbf{x}_k-\mathbf{b},\\ \mathbf{R}_k &= \mathbf{B}^{\mathrm{T}}\mathbf{r}_k.\end{aligned} \tag{57}$$

If $\mathbf{x}_k$ converges according to a given stopping criterion $\|\mathbf{r}_k\| < \varepsilon_1$ (or $\|\mathbf{R}_k\| < \varepsilon_2$) then stop; otherwise, go to step (iii). From Liu (2011a), $\beta_k = q(t_k)\Delta t$ is found to be

$$\beta_k = \frac{1-\gamma}{a_k^0} = (1-\gamma)\frac{(\mathbf{y}_k^{\mathrm{T}}\mathbf{A}\mathbf{y}_k)^2}{\|\mathbf{y}_k\|^2\|\mathbf{A}\mathbf{y}_k\|^2}, \tag{58}$$

where $0 \le \gamma < 1$ is a relaxed parameter.

When the iterative sequence of $\mathbf{y}_k$ approaches the eigenvector of $\mathbf{A}$ the steplength tends to

$$\frac{\mathbf{y}_k^{\mathrm{T}}\mathbf{A}\mathbf{y}_k}{\|\mathbf{A}\mathbf{y}_k\|^2} = \frac{1}{\lambda} > 0. \tag{59}$$

Below we will use two numerical examples to demonstrate this phenomenon, which is quite different from that of the RSDM about the behavior of adapting steplength.

5. Numerical Examples

In order to assess the performance of the newly developed method of the Lorentz-group algorithm (LGA), let us investigate the following examples. Especially, we emphasize the numerical solutions of linear inverse problems.

5.1 Example 1

The following example is used to illustrate the fixed point behavior of the LGA:

$$\begin{bmatrix} 10 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 10 \\ 1 \end{bmatrix}. \tag{60}$$

The exact solution is $(x, y) = (1, 1)$. Denote the coefficient matrix in Equation (60) by $\mathbf{B}$; hence, Cond($\mathbf{A}$) =Cond($\mathbf{B}\mathbf{B}^{\mathrm{T}}$) = 100 and the eigenvalues of $\mathbf{A}$ are $\lambda_1 = 100$, and $\lambda_2 = 1$.

In the computation of $\mathbf{y}_k$ we have discretized it by an iterative sequence as shown in Equation (57), which is a discretized model of the LGA of the residual dynamics. Instead of we can also integrate it by using a time stepsize Δt, which is a continuous model of the residual dynamics; by setting $\beta_{k-1} = q_{k-1}\Delta t$ in Equation (57) we can obtain the continuous model.

Then we compare the residual errors obtained by the LGA with $\gamma = 0.05$ and the continuous model with $q = 1$ and $\Delta t = 0.02$ in Fig. 2(a). Under the convergence criterion $\varepsilon_2 = 10^{-10}$, the LGA is convergence with 227 iterations, while the continuous model is convergence with 101 iterations. As shown in Fig. 2(b), starting from the same point $(y_1, y_2) = (-10, -2)$, the continuous model approaches to the fixed point $(y_1, y_2) = (0, -10)$, while the LGA does not exactly converge to the fixed point, and it oscillates near the fixed point. Because the LGA is obtained by a discretization to a purely iterative algorithm with β_k given by Equation (58), it does not fully obey the dynamical law of a continuous time dynamical system. The above two methods both give very accurate solutions with the maximum errors being smaller than 10^{-11}.

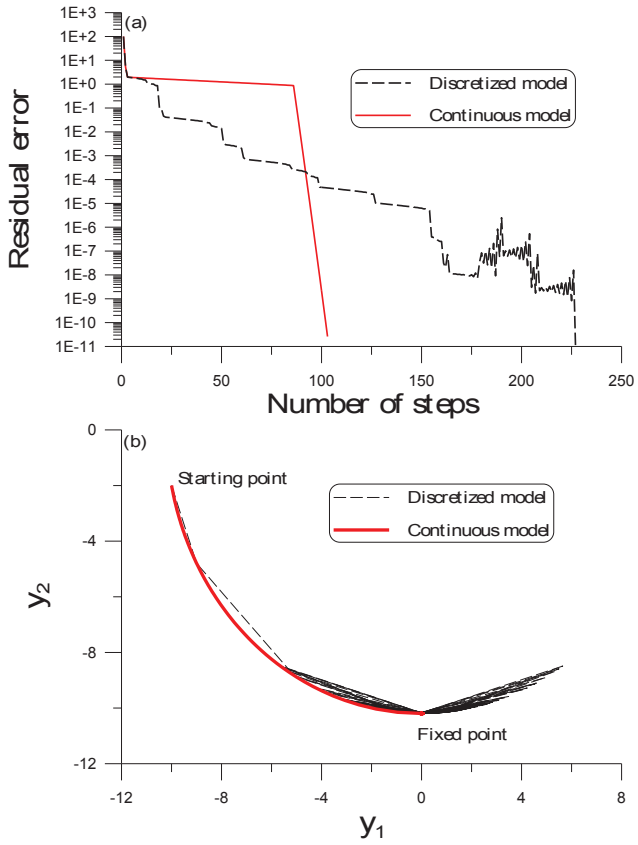

Figure 2. For example 1, (a) comparing the residual errors, and (b) comparing the iterative orbit of y, which tends to a fixed point for a continuous model, but oscillate near the fixed point for the discretized model.

There have two main drawbacks of the continuous model that one needs to specify the function $q(t)$ and determine the stepsize Δt. In the followings, we only use the discretized model of the LGA to solve linear problems.

5.2 Example 2

This example is used to illustrate that the presented LGA is better than the RSDM, where we give a simple example:

$$\begin{bmatrix} 2 & 2 \\ 6 & 6.0001 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4 \\ 12.0001 \end{bmatrix}. \tag{61}$$

The exact solution is $(x, y) = (1, 1)$. It is interesting to note that the condition number is Cond($\mathbf{A}$) =Cond($\mathbf{B}\mathbf{B}^{\mathrm{T}}$) = 1.5957×10^{11}, where $\mathbf{B}$ denotes the coefficient matrix in Equation (61). The eigenvalues are $\lambda_1 = 80.0012$, and $\lambda_2 = 5.01352 \times 10^{-10}$.

Then, we add a random noise with an intensity $\sigma = 0.05$ on the data of $(4, 12.0001)^{\mathrm{T}}$. By using the LGA with $\gamma = 0$ we obtain a solution of $(x, y) = (0.99981, 0.99945)$ under a convergence criterion $\varepsilon_2 = 10^{-7}$ or with the maximum number 1000 of iterations. In Fig. 3 we display the iterative orbit of (y_1, y_2) starting from $(-4.01326, -11.9907)$ and ending at $(-11.9958, 3.99819)$, which is an eigenvector of $\mathbf{A}$ corresponding to the eigenvalue λ_1. Then we apply the RSDM with $\gamma = 0.01$ to solve this problem, whose solution is found to be $(x, y) = (1.0007, 0.9986)$. It can be seen that the LGA is more accurate than the RSDM with one order. In Fig. 4(a) we compare the residual errors of RSDM and LGA, while the steplengths are compared in Fig. 4(b). The two thick solid red lines composed of squares present the iterative steplengths of which the lower part is just the inverse of the first eigenvalue $1/\lambda_1 = 0.0124998$. However, due to the matrix $\mathbf{A}$ being highly ill-conditioned, the calculated steplengths are jumping between two values of 1.53160 and 0.0124998, which are sensitive to the slight difference of the computed values of (y_1, y_2). On the other hand, the steplengths of the RSDM as shown by the blue dashed line are quite regularly varying in a large range, which are drastically different from the ones of the LGA.

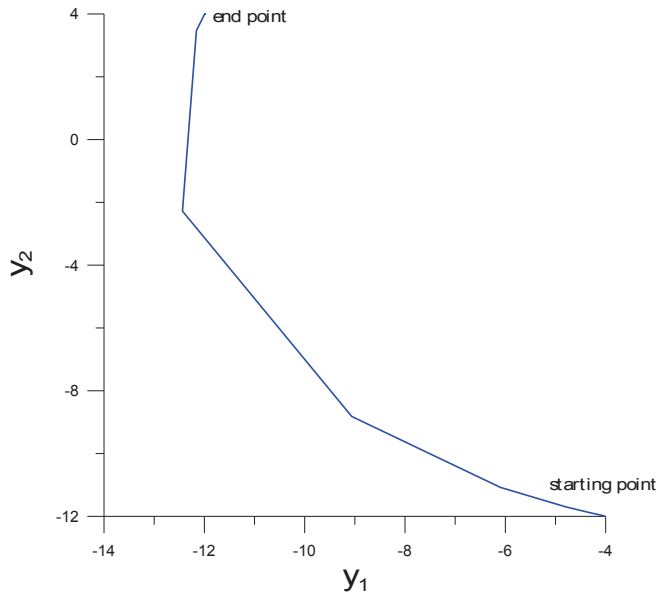

Figure 3. For example 2 showing the iterative orbit of $\mathbf{y}$, which tends to a fixed point.

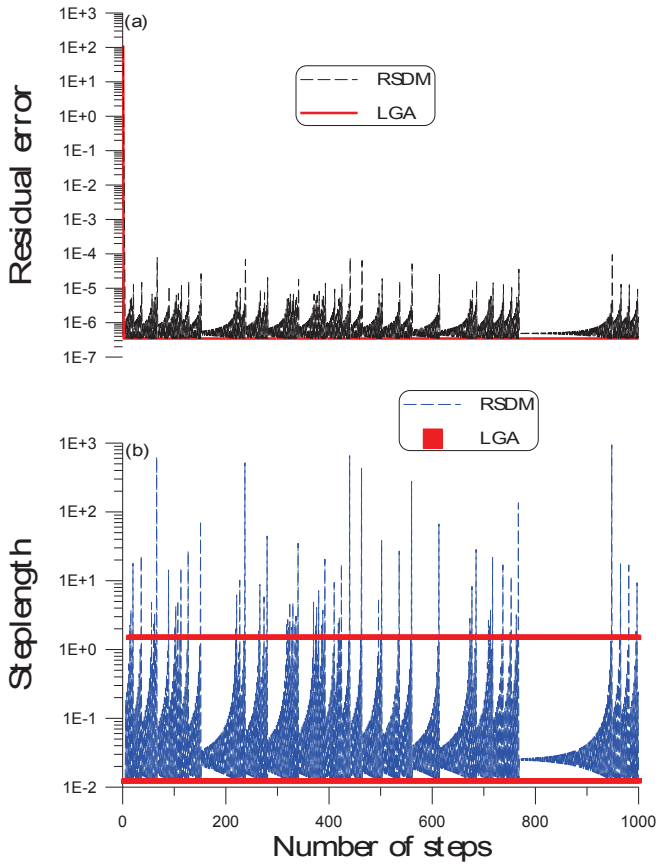

Figure 4. For example 2: (a) comparing the residual errors of RSDM and LGA, and (b) showing the steplengths of RSDM and LGA.

The existence of a fixed point of the residual dynamics for the LGA is very important, which can enhance the long-term stability of the LGA. For example we increase the maximum number of iterations to 30000. The above RSDM leads to an incorrect solution of $(x, y) = (2.0427, -0.0434)$; however, the LGA with $\gamma = 0.05$ gives a rather accurate solution of $(x, y) = (0.9952, 1.004)$, with the accuracy being three orders.

5.3 Example 3

When the backward heat conduction problem (BHCP) is considered in a spatial interval of $0 < x < \ell$ by subjecting to the boundary conditions at two ends of a slab:

$$u_t(x,t) = \alpha u_{xx}(x,t), \quad 0 < t < T, \quad 0 < x < \ell, \tag{62}$$

$$u(0,t) = u_0(t), \quad u(k,t) = u_\ell(t), \tag{63}$$

we solve u under a final time condition:

$$u(x,T) = u^T(x). \tag{64}$$

The fundamental solution to Equation (62) is given as follows:

$$K(x,t) = \frac{H(t)}{2\sqrt{\alpha\pi t}} \exp\left(\frac{-x^2}{4\alpha t}\right), \tag{65}$$

where $H(t)$ is the Heaviside function.

The method of fundamental solutions (MFS) has a broad application in engineering computations. In the MFS the solution of u at the field point $\mathbf{z} = (x,t)$ can be expressed as a linear combination of the fundamental solutions $U(\mathbf{z}, \mathbf{s}_j)$ at different sources:

$$u(\mathbf{z}) = \sum_{j=1}^{n} c_j U(\mathbf{z}, \mathbf{s}_j), \quad \mathbf{s}_j = (\eta_j, \tau_j) \in \Omega^c, \tag{66}$$

where n is the number of source points, c_j are unknown coefficients, and $\mathbf{s}_j$ are source points being located in the complement Ω^c of $\Omega = [0,\ell] \times [0,T]$. For the heat conduction equation we have

$$U(\mathbf{z}, \mathbf{s}_j) = K(x - \eta_j, t - \tau_j). \tag{67}$$

It is known that the distribution of source points in the MFS has a great influence on the accuracy and stability. In a practical application of MFS to solve the BHCP, the source points are uniformly located on two vertical straight lines parallel to the t-axis, and one horizontal line over the final time $t = T$, which was adopted by Hon & Li (2009) and Liu (2011b), showing a large improvement than the line location of source points below the initial time. After imposing the boundary conditions and the final time condition on Equation (66) we can obtain a linear equations system (2), where

$$\begin{aligned} &B_{ij} = U(\mathbf{z}_i, \mathbf{s}_j), \quad \mathbf{x} = (c_1, \cdots, c_n)^{\mathrm{T}}, \\ &\mathbf{b} = (u_\ell(t_i),\ i = 1, \ldots, m_1; u^T(x_j),\ j = 1, \ldots, m_2; u_0(t_k),\ k = m_1, \ldots, 1)^{\mathrm{T}}, \end{aligned} \tag{68}$$

and $n = 2m_1 + m_2$.

Here we compare the numerical solution with an exact solution:

$$u(x,t) = \cos(\pi x) \exp(-\pi^2 t). \tag{69}$$

For the case with $T = 1$ the value of final time data is in the order of 10^{-4}, which is much small in a comparison with the value of the initial temperature $u(x,0) = \cos(\pi x)$ to be retrieved, which is $O(1)$. We solve this problem by the Lorentz group algorithm (LGA). As shown in Fig. 5(a), the LGA with $\gamma = 0.04$ converges very fast with 867 iterations under the convergence criterion $\varepsilon_2 = 10^{-4}$, while the RSDM with $\gamma = 0.05$ converges with 1195 iterations. We add a relative random noise with an intensity $\sigma = 10\%$ on the final time data. The numerical errors are shown in Fig. 5(b), of which the maximum errors of RSDM and LGA are, respectively, 0.036, 0.029. It can be seen that the performance of LGA is better than that of the RSDM.

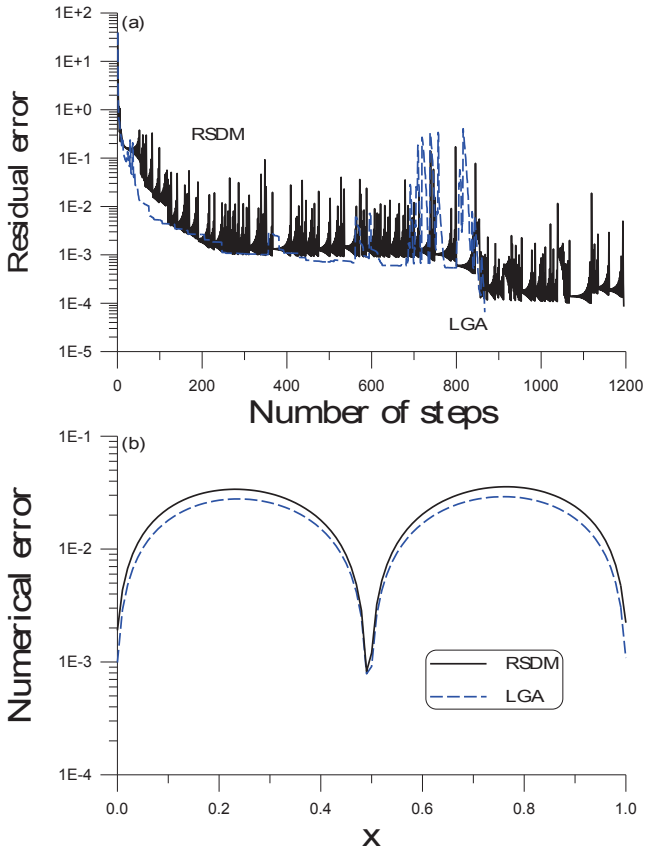

Figure 5. For example 3, (a) comparing the residual errors of RSDM and LGA, and (b) showing the numerical errors.

5.4 Example 4

We solve the Cauchy problem of the Laplace equation under boundary conditions:

$$\begin{aligned} &\Delta u = u_{rr} + \frac{1}{r}u_r + \frac{1}{r^2}u_{\theta\theta} = 0, \quad r < \rho, \quad 0 \le \theta \le 2\pi, \\ &u(\rho,\theta) = h(\theta), \quad 0 \le \theta \le \pi, \\ &u_n(\rho,\theta) = g(\theta), \quad 0 \le \theta \le \pi, \end{aligned} \tag{70}$$

where $h(\theta)$ and $g(\theta)$ are given functions, and $\rho = \rho(\theta)$ is a given contour to describe the boundary shape. The contour in the polar coordinates is specified by $\Gamma = \{(r,\theta)|r = \rho(\theta),\ 0 \le \theta \le 2\pi\}$, which is the boundary of the problem domain Ω, and n denotes the normal direction.

In the potential theory, it is well known that the method of fundamental solutions (MFS) can be used to solve the Laplacian problems when a fundamental solution is known [Kupradze (1964)]. In the MFS the trial solution of u at the field point $\mathbf{z} = (r\cos\theta, r\sin\theta)$ can be expressed as a linear combination of the fundamental solutions $U(\mathbf{z}, \mathbf{s}_j)$:

$$u(\mathbf{z}) = \sum_{j=1}^{n} c_j U(\mathbf{z}, \mathbf{s}_j), \quad \mathbf{s}_j \in \Omega^c, \tag{71}$$

where n is the number of source points, c_j are the unknown coefficients, $\mathbf{s}_j$ are the source points, and Ω^c is the complementary set of Ω. For the Laplace equation (70) we have the fundamental solutions:

$$U(\mathbf{z}, \mathbf{s}_j) = \ln r_j, \quad r_j = \|\mathbf{z} - \mathbf{s}_j\|. \tag{72}$$

In the practical application of the MFS, frequently the source points are uniformly located on a circle with a radius R, such that after imposing the boundary conditions in Equation (70) on Equation (71) we can obtain a linear equations system

(2), where

$$\begin{aligned}
&\mathbf{z}_i = (z_i^1, z_i^2) = (\rho(\theta_i)\cos\theta_i, \rho(\theta_i)\sin\theta_i),\\
&\mathbf{s}_j = (s_j^1, s_j^2) = (R\cos\theta_j, R\sin\theta_j),\\
&B_{ij} = \begin{cases} \ln\|\mathbf{z}_i - \mathbf{s}_j\|, & \text{if } i \text{ is odd},\\ \frac{\eta(\theta_i)}{\|\mathbf{z}_i - \mathbf{s}_j\|^2}\Big(\rho(\theta_i) - s_j^1\cos\theta_i - s_j^2\sin\theta_i \\ \quad - \frac{\rho'(\theta_i)}{\rho(\theta_i)}[s_j^1\sin\theta_i - s_j^2\cos\theta_i]\Big), & \text{if } i \text{ is even},\end{cases}\\
&\mathbf{x} = (c_1, \ldots, c_n)^{\mathrm{T}}, \quad \mathbf{b} = (h(\theta_1), g(\theta_1), \ldots, h(\theta_m), g(\theta_m))^{\mathrm{T}},
\end{aligned} \tag{73}$$

in which $n = 2m$, and

$$\eta(\theta) = \frac{\rho(\theta)}{\sqrt{\rho^2(\theta) + [\rho'(\theta)]^2}}. \tag{74}$$

We fix $n = 38$ and employ a circle with a constant radius $R = 15$ to distribute the source points. Then we apply the LGA with $\gamma = 0.02$ to solve the linear system (2) under a convergence criterion $\varepsilon_2 = 10^{-3}$, where a noise with an intensity $\sigma = 10\%$ is imposed on the given data. The residual error is plotted in Fig. 6(a), which is convergence with 4624 iterations.

Along the lower half contour $\rho(\theta) = \sqrt{10 - 6\cos(2\theta)}$, $\pi \le \theta < 2\pi$, in Fig. 6(b) we show the numerical error obtained by the LGA with that given by $u = \rho^2\cos(2\theta)$, $\pi \le \theta < 2\pi$. For the purpose of comparison with the regularized MFS proposed by Wei et al. (2007), we also display its numerical error in Fig. 6(b), whose maximum error is found to be 0.39. We can observe that the result obtained by the LGA with the maximum error being 0.111 is better than that of the MFS.

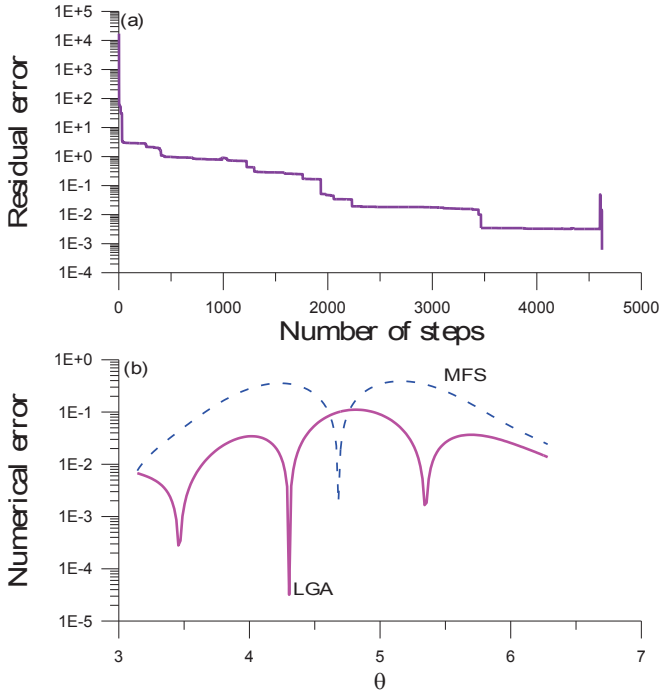

Figure 6. (a) the residual error of LGA, and (b) showing the numerical errors of LGA and MFS.

5.5 Example 5

When we apply a central difference scheme to the following two-point boundary value problem:

$$\begin{aligned}
&-u''(x) = f(x), \quad 0 < x < 1,\\
&u(0) = a, \quad u(1) = b,
\end{aligned} \tag{75}$$

we can derive a linear equations system

$$\mathbf{Bu} = \begin{bmatrix} 2 & -1 & & & & \\ -1 & 2 & -1 & & & \\ & \cdot & \cdot & \cdot & & \\ & & \cdot & \cdot & \cdot & \\ & & & \cdot & \cdot & \cdot \\ & & & & -1 & 2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} = \begin{bmatrix} (\Delta x)^2 f(\Delta x) + a \\ (\Delta x)^2 f(2\Delta x) \\ \vdots \\ (\Delta x)^2 f((n-1)\Delta x) \\ (\Delta x)^2 f(n\Delta x) + b \end{bmatrix}, \tag{76}$$

where $\Delta x = 1/(n+1)$ is the spatial length, and $u_i = u(i\Delta x)$, $i = 1, \ldots, n$, are unknown values of $u(x)$ at the grid points $x_i = i\Delta x$. $u_0 = a$ and $u_{n+1} = b$ are the given boundary conditions.

In this numerical test we fix $n = 30$. Let us consider the boundary value problem in Equation (75) with $f(x) = \sin \pi x$. The exact solution is

$$u(x) = a + (b-a)x + \frac{1}{\pi^2} \sin \pi x, \tag{77}$$

where we fix $a = 1$ and $b = 2$.

A relative random noise with intensity $\sigma = 0.01$ is added into the data on the right-hand side of Equation (76). Under a moderate convergence criterion with $\varepsilon_2 = 10^{-8}$, we find that the RSDM with $\gamma = 0.15$ converges with 6111 iterations as shown in Fig. 7(a) by dashed line, and the maximum error as shown in Fig. 7(b) is 5.3×10^{-5}. It can be seen that the presented LGA with $\gamma = 0.04$ converges with 4140 iterations and with the maximum error being 5.04×10^{-5}, which is better than the algorithm RSDM.

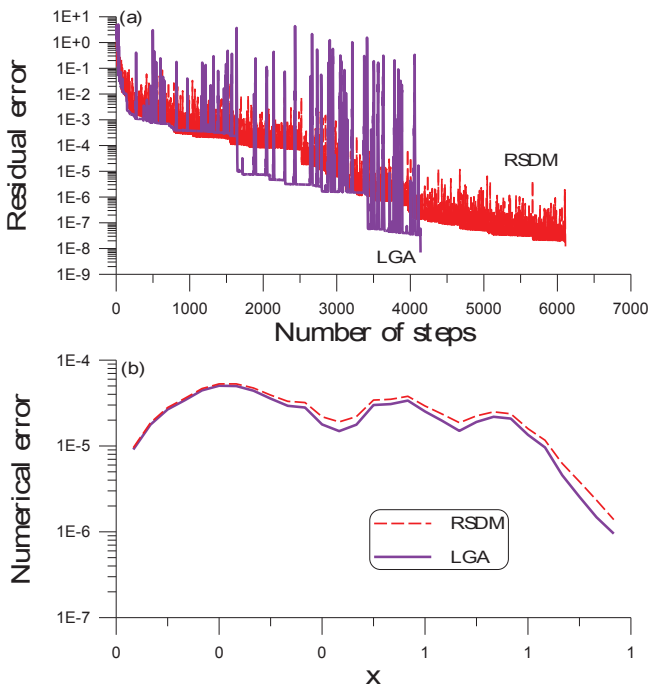

Figure 7. For example 5, (a) comparing the residual errors of RSDM and LGA, and (b) showing the numerical errors.

5.6 Example 6

Let us consider the following inverse problem to recover the external force $F(t)$ for an ODE:

$$\ddot{y}(t) + \dot{y}(t) + y(t) = F(t). \tag{78}$$

In a time interval of $t \in [0, t_f]$ the discretized data $y_i = y(t_i)$ are supposed to be measurable, which are subjected to the random noise with an intensity $\sigma = 0.01$. Usually, it is very difficult to recover the external force $F(t_i)$ from Equation (78) by the direct differentials of the noisy data of the displacements, because the differential is an ill-posed linear operator.

To approach this inverse problem by the polynomial interpolation, we begin with

$$p_m(x) = c_0 + \sum_{k=1}^{m} c_k x^k. \tag{79}$$

Now, the coefficient c_k is projected into two coefficients a_k and b_k to absorb more interpolation points; in the meanwhile, $\cos(k\theta_k)$ and $\sin(k\theta_k)$ are introduced to reduce the condition number of the coefficient matrix [Liu (2011c)]. We suppose that

$$c_k = \frac{a_k \cos(k\theta_k)}{R_{2k}^k} + \frac{b_k \sin(k\theta_k)}{R_{2k+1}^k}, \tag{80}$$

and

$$\theta_k = \frac{2k\pi}{m}, \quad k = 1, \ldots, m. \tag{81}$$

The considered problem domain is $[a, b]$, and the interpolating points are:

$$a = x_0 < x_1 < x_2 < \ldots < x_{2m-1} < x_{2m} = b. \tag{82}$$

Substituting Equation (80) into Equation (79), we can obtain

$$p(x) = a_0 + \sum_{k=1}^{m} \left[a_k \left(\frac{x}{R_{2k}} \right)^k \cos(k\theta_k) + b_k \left(\frac{x}{R_{2k+1}} \right)^k \sin(k\theta_k) \right], \tag{83}$$

where we let $c_0 = a_0$, and a_k and b_k are unknown coefficients. In order to obtain them, we impose the following n interpolated conditions:

$$p(x_i) = y_i, \quad i = 0, \ldots, n-1. \tag{84}$$

Thus, we obtain a linear equations system to determine a_k and b_k:

$$\begin{bmatrix} 1 & \frac{x_0 \cos\theta_1}{R_2} & \frac{x_0 \sin\theta_1}{R_3} & \ldots & \left(\frac{x_0}{R_{2m}}\right)^m \cos m\theta_m & \left(\frac{x_0}{R_{2m+1}}\right)^m \sin m\theta_m \\ 1 & \frac{x_1 \cos\theta_1}{R_2} & \frac{x_1 \sin\theta_1}{R_3} & \ldots & \left(\frac{x_1}{R_{2m}}\right)^m \cos m\theta_m & \left(\frac{x_1}{R_{2m+1}}\right)^m \sin m\theta_m \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \frac{x_{2m-1} \cos\theta_1}{R_2} & \frac{x_{2m-1} \sin\theta_1}{R_3} & \ldots & \left(\frac{x_{2m-1}}{R_{2m}}\right)^m \cos m\theta_m & \left(\frac{x_{2m-1}}{R_{2m+1}}\right)^m \sin m\theta_m \\ 1 & \frac{x_{2m} \cos\theta_1}{R_2} & \frac{x_{2m} \sin\theta_1}{R_3} & \ldots & \left(\frac{x_{2m}}{R_{2m}}\right)^m \cos m\theta_m & \left(\frac{x_{2m}}{R_{2m+1}}\right)^m \sin m\theta_m \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ b_1 \\ \vdots \\ a_m \\ b_m \end{bmatrix} = \begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ \vdots \\ y_{2m-1} \\ y_{2m} \end{bmatrix}. \tag{85}$$

We note that the norm of the first column of the above coefficient matrix is $\sqrt{2m+1}$. According to the concept of equilibrated matrix [Liu (2012d)], we can derive the optimal scales for the current interpolation with a half-order technique as

$$R_{2k} = \beta_0 \left(\frac{1}{2m+1} \sum_{j=0}^{2m} x_j^{2k} (\cos k\theta_k)^2 \right)^{1/(2k)}, \quad k = 1, 2, \ldots, m, \tag{86}$$

$$R_{2k+1} = \beta_0 \left(\frac{1}{2m+1} \sum_{j=0}^{2m} x_j^{2k} (\sin k\theta_k)^2 \right)^{1/(2k)}, \quad k = 1, 2, \ldots, m, \tag{87}$$

where β_0 is a stability factor. The improved method uses m order polynomial to interpolate $n = 2m + 1$ data nodes, while regular method with a full-order can only interpolate $m + 1$ data points.

Now we fix $m = 10$ and $t_f = 5$ and consider the exact solution of $F(t) = \cos t$, which is obtained by inserting the exact $y(t) = \sin t$ into Equation (78). The parameters used are $\beta_0 = 1.2$, $\gamma = 0.025$ for the LGA, and $\beta_0 = 2.2$ and $\gamma = 0.35$ for the RSDM. Under the same convergence criterion $\varepsilon_2 = 10^{-5}$, the LGA does not converge over 10000 iterations, while the RSDM is not convergent over 50000 iterations, as compared the residual errors in Fig. 8(a). We compare the numerical solutions and the exact solution in Fig. 8(b), where the maximum error for the RSDM is a large value 0.452, and that for the LGA is 0.067. Obviously, the LGA is much better than the RSDM. The result calculated by the LGA is also better than that calculated by Liu (2012a).

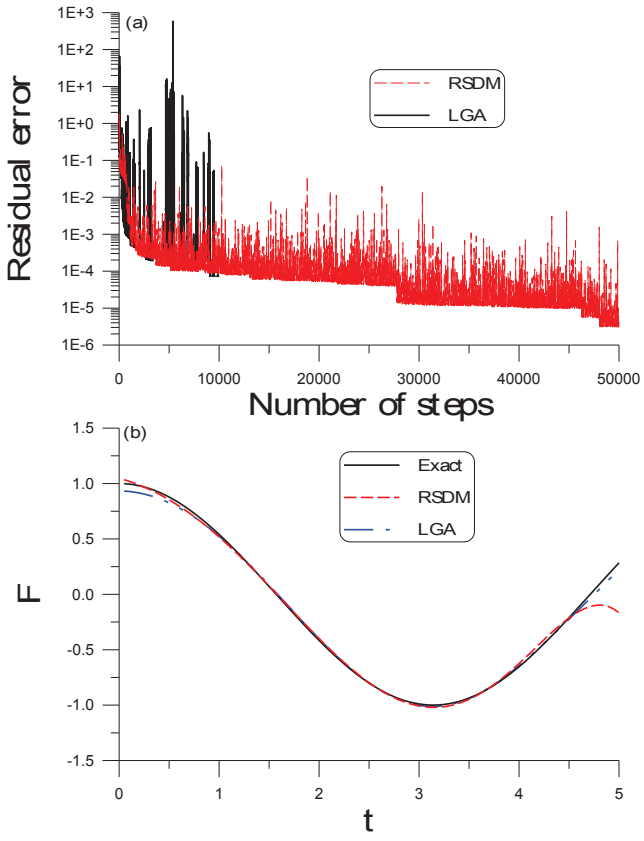

Figure 8. For example 6, (a) comparing the residual errors of RSDM and LGA, and (b) comparing the numerical results with exact one.

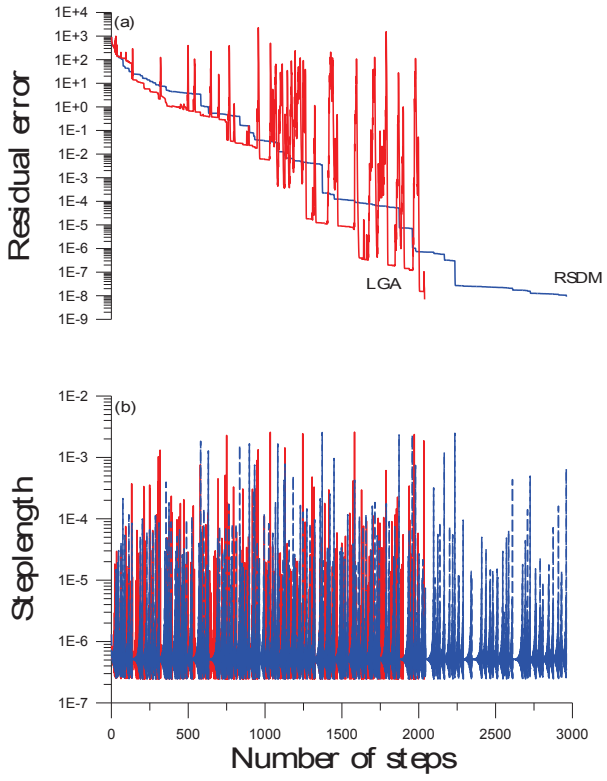

Figure 9. For example 7, comparing (a) the residual errors and (b) the steplengths of RSDM and LGA.

5.7 Example 7

In order to compare the LGA with the RSDM for a well-posed problem, we solve the following Laplace equation:

$$\begin{aligned} &u_{xx} + u_{yy} = 0, \quad 0 < x < 1, \quad 0 < y < 1, \\ &u(x, y) = \sin x \cosh y. \end{aligned} \tag{88}$$

The boundary conditions can be derived from the exact solution. A standard five point finite difference is used to discretize the above equation. Under the discretization with $\Delta x = \Delta y = 1/16$, and with $\gamma = 0.05$ for both the RSDM and the LGA, and under the convergence criterion $\varepsilon_1 = 10^{-8}$, the RSDM converges with 2961 iterations, but the LGA spends 2038 iterations, as shown in Fig. 9. The above two algorithms obtain very accurate solutions with the maximum numerical error being 2.73×10^{-5}.

6. Conclusions

The present paper in the framework of the future cone in the Minkowski space has explored two essential structures about the residual dynamics, namely the Jordan dynamics, as well as the Lie-group $SO_o(n, 1)$. It is a residual symmetry group of the residual dynamics as being expressed by a nonlinear Lorentzian system. We also proved that the eigenvector of the coefficient matrix is a critical point of the residual dynamics, which is important for a long term computation by the Lorentz-group algorithm (LGA). We have proposed a theoretical modification of the steplength used in the steepest descent method by using the LGA, which can significantly accelerate the convergence speed and is more stable in the numerical solution of ill-posed linear problems. Because all formulas required in the new steplength were derived explicitly, the presented LGA can be implemented easily and used effectively to solve the ill-posed linear inverse problems.

Acknowledgements

The project NSC-102-2221-E-002-125-MY3 from the Ministry of Science and Technology of Taiwan, and the 2014 ISI Highly Cited Researcher Award granted to the author are highly appreciated.

References

Akaike, H. (1959). On a successive transformation of probability distribution and its application to the analysis of the optimum gradient method. *Ann. Inst. Stat. Math. Tokyo 11*, 1-16. http://dx.doi.org/10.1007/BF01831719

Barzilai, J., & Borwein, J. M. (1988). Two point step size gradient methods. *IMA J. Numer. Anal. 8*, 141-148. http://dx.doi.org/10.1093/imanum/8.1.141

Dai, Y. H., & Yuan, Y. (2003). Alternate minimization gradient method. *IMA J. Numer. Anal. 23*, 377-393. http://dx.doi.org/10.1093/imanum/23.3.377

Dai, Y. H., Hager, W., Schittkowsky, K., & Zhang, H. (2006). A cyclic Barzilai-Borwein method for unconstrained optimization. *IMA J. Numer. Anal. 26*, 604-627. http://dx.doi.org/10.1093/imanum/drl006

Forsythe, G. E. (1968). On the asymptotic directions of the s-dimensional optimum gradient method. *Numer. Math. 11*, 57-76. http://dx.doi.org/10.1007/BF02165472

Friedlander, A., Martinez, J., Molina, B., & Raydan, M. (1999). Gradient method with retard and generalizations. *SIAM J. Num. Anal. 36*, 275-289. http://dx.doi.org/10.1137/S003614299427315X

Hon, Y. C., & Li, M. (2009). A discrepancy principle for the source points location in using the MFS for solving the BHCP. *Int. J. Comput. Meth. 6*, 181-197. http://dx.doi.org/10.1142/S0219876209001759

Iordanescu, R. (2007). Dynamical systems and Jordan structures. *Int. J. Pure Appl. Math. 35*, 125-143.

Iordanescu, R. (2009). Jordan Structures in Analysis, Geometry and Physics. Editura Academiei Romane, Bucuresti.

Kunisch, K., & Zou, J. (1998). Iterative choices of regularization parameters in linear inverse problems. *Inverse Problems 14*, 1247-1264. http://iopscience.iop.org/article/10.1088/0266-5611/14/5/010/pdf

Kupradze, V. D., & Aleksidze, M. A. (1964). The method of functional equations for the approximate solution of certain boundary value problems. *USSR Comput. Math. Math. Phy. 4*, 82-126. http://dx.doi.org/10.1016/0041-5553(64)90006-0

Liu, C.-S. (2000). A Jordan algebra and dynamic system with associator as vector field. *Int. J. Non-Linear Mech. 35*, 421-429. http://dx.doi.org/10.1016/S0020-7462(99)00027-X

Liu, C.-S. (2001). Cone of non-linear dynamical system and group preserving schemes. *Int. J. Non-Linear Mech. 36*, 1047-1068. http://dx.doi.org/10.1016/S0020-7462(00)00069-X

Liu, C.-S. (2002). Nonlinear Lorentzian and Hamiltonian formulations and their relations. *Int. J. Appl. Math. 10*, 59-97.

Liu, C.-S. (2011a). A revision of relaxed steepest descent method from the dynamics on an invariant manifold. *Comput. Model. Eng. Sci. 80*, 57-86.

Liu, C.-S. (2011b). The method of fundamental solutions for solving the backward heat conduction problem with conditioning by a new post-conditioner. *Num. Heat Transfer, B: Fundamentals 60*, 57-72. http://dx.doi.org/10.1080/10407790.2011.588134

Liu, C.-S. (2011c). A highly accurate multi-scale full/half-order polynomial interpolation. *Comput. Mater. Continua 25*, 239-263.

Liu, C.-S. (2012a). Modifications of steepest descent method and conjugate gradient method against noise for ill-posed linear systems. *Commu. Num. Anal. 2012*, Article ID cna-00115, 24 pages. http://dx.doi.org/10.5899/2012/cna-00115

Liu, C.-S. (2012b). Optimally scaled vector regularization method to solve ill-posed linear problems. *Appl. Math. Comp. 218*, 10602-10616. http://dx.doi.org/10.1016/j.amc.2012.04.022

Liu, C.-S. (2012c). The concept of best vector used to solve ill-posed linear inverse problems. *Comput. Model. Eng. Sci. 83*, 499-525.

Liu, C.-S. (2012d). An equilibrated method of fundamental solutions to choose the best source points for the Laplace equation. *Eng. Anal. Bound. Elem. 36*, 1235-1245. http://dx.doi.org/10.1016/j.enganabound.2012.03.001

Liu, C.-S. (2013a). A dynamical Tikhonov regularization for solving ill-posed linear algebraic systems. *Acta Appl. Math. 123*, 285-307. http://dx.doi.org/10.1007/s10440-012-9766-3

Liu, C.-S. (2013b). A two-side equilibration method to reduce the condition number of an ill-posed linear system. *Comput. Model. Eng. Sci. 91*, 17-42.

Liu, C.-S. (2013c). Discussing a more fundamental concept than the minimal residual method for solving linear system in a Krylov subspace. *J. Math. Research 5(4)*, 58-70. http://dx.doi.org/10.5539/jmr.v5n4p58

Liu, C.-S. (2014). A doubly optimized solution of linear equations system expressed in an affine Krylov subspace. *J. Comp. Appl. Math. 260*, 375-394. http://dx.doi.org/10.1016/j.cam.2013.10.013

Nocedal, J., Sartenar, A., & Zhu, C. (2002). On the behavior of the gradient norm in the steepest descent method. *Comp. Optim. Appl. 22*, 5-35. http://dx.doi.org/10.1023/A:1014897230089

Raydan, M., & Svaiter, B. F. (2002). Relaxed steepest descent and Cauchy-Barzilai-Borwein method. *Comp. Optim. Appl. 21*, 155-167. http://dx.doi.org/10.1023/A:1013708715892

Resmerita, E. (2005). Regularization of ill-posed problems in Banach spaces: convergence rates. *Inverse Problems 21*, 1303-1314. http://dx.doi.org/10.1088/0266-5611/21/4/007

Tikhonov, A. N., & Arsenin, V. Y. (1977). Solutions of Ill-Posed Problems. John-Wiley & Sons, New York.

Wang, Y., & Xiao, T. (2001). Fast realization algorithms for determining regularization parameters in linear inverse problems. *Inverse Problems 17*, 281-291. http://iopscience.iop.org/article/10.1088/0266-5611/17/2/308/pdf

Wei, T., Hon, Y. C., & Ling, L. (2007). Method of fundamental solutions with regularization techniques for Cauchy problems of elliptic operators. *Eng. Anal. Bound. Elem. 31*, 373-385. http://dx.doi.org/10.1016/j.enganabound.2006.07.010

Xie, J., & Zou, J. (2002). An improved model function method for choosing regularization parameters in linear inverse problems. *Inverse Problems 18*, 631-643. http://iopscience.iop.org/article/10.1088/0266-5611/18/3/307/pdf