Riemannian Newton and conjugate gradient algorithm for computing Lagrangian invariant subspaces

U. Helmke† and M. Kleinsteuber†

Abstract. The computation of Lagrangian invariant subspaces of a Hamiltonian matrix, or the closely related task of solving algebraic Riccati equations, is an important issue in linear optimal control, stochastic control and $H^\infty$-design. We propose a new class of Riemannian Newton methods that allows to compute isolated Lagrangian invariant subspaces of a Hamiltonian matrix. The algorithm implements a variant of the Newton method for a quadratic vector field on the Lagrange Grassmann manifold. It yields new methods to solve the algebraic Riccati equation in linear optimal control. In addition, an intrinsic conjugate gradient algorithm on the Lagrangian Grassmanian is introduced.

Keywords: Hamiltonian matrices, Riccati equations, Linear optimal control, Lagrange Grassmann manifold, Riemannian Newton method, Riemannian conjugate gradient method.

1 Introduction

In this paper, we propose a Riemannian geometry approach to computing Lagrangian invariant subspaces of an arbitrary Hamiltonian matrix. Our method is in particular applicable to the algebraic Riccati equation from linear optimal control. It yields a modification of the familiar Kleinman method [7] that did proceed by applying standard Newton steps to the algebraic Riccati equation. Our approach differs from that by using a geometrically more natural covariant derivative in order to define the Newton direction. The algorithm also has the advantage of being defined on the entire Lagrange Grassmannian, while Kleinman’s method works only on a local coordinate chart. Therefore, our method avoids potential numerical difficulties occurring on the boundary of the coordinate chart. We begin by extending the Riccati vector field to a quadratic vector field on the Lagrange Grassmannian.

*This research was partially supported by the German Research Foundation, DFG-Project KONNEW HE 1858/10-2.
†Institute of Mathematics, University of Würzburg, 97074 Würzburg, Germany
Actually, this defines a generalization of Brockett’s double bracket flow [4] on the Lagrange Grassmannian, whose critical points yield exactly the Lagrangian invariant subspaces of a given Hamiltonian matrix. We then formulate a family of Riemannian Newton algorithms for this vector field and establish local quadratic convergence to an isolated Lagrangian invariant subspace. The algorithms require significantly only solving linear Sylvester equations and computing approximations to the matrix exponentials of a skew-symmetric matrix. The computational burden in the last step can be reduced, using Gram-Schmidt orthogonalizations. In the last section we also propose a conjugate gradient algorithm for the generalized double bracket vector field. Numerical simulations demonstrate the suitability of the approach.

2 Geometry of the Lagrange Grassmannian

The Lagrange Grassmann manifold is defined as the set of all \( n \)-dimensional Lagrangian subspaces of \( \mathbb{R}^{2n} \) with respect to the standard symplectic form

\[
J := \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}.
\] (1)

Recall, that an \( n \)-dimensional subspace \( V \subset \mathbb{R}^{2n} \) is called Lagrangian, if \( v^\top J v = 0 \) holds for all \( v \in V \). Equivalently, one can identify a Lagrangian subspace with a corresponding self adjoint projection operator \( P \) satisfying \( PJP = 0 \). Thus we can identify the Lagrange Grassmann manifold with the Lagrange Grassmannian

\[
\mathbb{L}G_n := \{ P \in \text{Sym}_{2n} | P^2 = P, \text{tr} P = n, PJP = 0 \}.
\] (2)

Note, that \( \mathbb{L}G_n \) is a compact, connected manifold of dimension \( \frac{1}{2}n(n + 1) \). The manifold \( \mathbb{L}G_n \) can also be regarded as a homogeneous space for the action of the orthogonal symplectic group

\[
\text{OSp}_{2n} := \{ T \in \text{GL}_{2n} | T^\top JT = J, T \in \text{SO}_{2n} \}
\] (3)

with Lie algebra

\[
\mathfrak{osp}_{2n} = \{ X \in \mathfrak{so}_{2n} | X^\top J + JX = 0 \}
\] (4)

of skew-symmetric Hamiltonian \( 2n \times 2n \)-matrices. Thus the compact Lie group \( \text{OSp}_{2n} \) acts transitively on \( \mathbb{L}G_n \) via \( \sigma : \text{OSp}_{2n} \times \mathbb{L}G_n \to \mathbb{L}G_n, \ (T, P) \mapsto T^\top PT \). Therefore, \( \mathbb{L}G_n \) is a homogeneous space that can be identified with \( \text{OSp}_{2n} / \text{O}_n \).

**Theorem 1 ([3]).**

(a) The tangent space of \( \mathbb{L}G_n \) at an element \( P \in \mathbb{L}G_n \) is given as

\[
T_P \mathbb{L}G_n = \{ [P, \Omega] | \Omega \in \mathfrak{osp}_{2n} \}.
\] (5)

(b) Let \( P \in \mathbb{L}G_n \) be arbitrary. The linear map

\[
\pi : \text{Sym}_{2n} \to \text{Sym}_{2n}, \quad X \mapsto \frac{1}{2} [P, [P, JXJ + X]]
\] (6)

is a self-adjoint projection operator onto \( T_P \mathbb{L}G_n \).
In the sequel, we consider $\mathbb{LG}_n$ as a Riemannian submanifold of the symmetric matrix space $\text{Sym}_2^n$; i.e. for any two tangent elements $\xi, \eta \in T_p\mathbb{LG}_n$, the Riemannian metric is defined by the Frobenius inner product $\left<\xi, \eta\right> = \text{tr}(\xi \eta)$. Given this Riemannian manifold structure on $\mathbb{LG}_n$, we can explicitly determine the canonical Levi-Civita connection on the Lagrange Grassmannian via the Gauss formula.

3 Generalized double bracket flows

From the geometric theory of Riccati equations it is well-known [5, 9], that the differential Riccati equation from linear optimal control

$$\dot{K} = -F^T K - KF + KQK - S =: R(K),$$

with $Q, S$ symmetric, extends to a linear induced flow on the Lagrange Grassmannian, defined via the associated Hamiltonian matrix

$$A = \begin{bmatrix} F & -Q \\ -S & -F^T \end{bmatrix}.$$  \hfill (8)

In particular, the symmetric solutions $K$ of the algebraic Riccati equation

$$F^T K + KF - KQK + S = 0$$

define Lagrange invariant subspaces $V = \text{span} \left[ \begin{bmatrix} I \\ K \end{bmatrix} \right]$ of the Hamiltonian $A$. By identifying $V$ with the associated projection operator

$$P = \begin{bmatrix} I \\ K \end{bmatrix} (I + K^2)^{-1} \left[ \begin{bmatrix} I \\ K \end{bmatrix} \right] \in \mathbb{LG}_n$$

it follows, that (7) defines a vector field on an open and dense coordinate chart of $\mathbb{LG}_n$. We show that this vector field is actually globally defined on $\mathbb{LG}_n$ and has a pleasant Lie bracket form. Thus, let $A = \Omega + S$ denote an arbitrary Hamiltonian matrix, with $\Omega = \frac{1}{2}(A - A^\top)$ and $S = \frac{1}{2}(A + A^\top)$ the skew-symmetric and symmetric parts of $A$, respectively. Note, that $\Omega \in \mathfrak{osp}_{2n}$ and $S$ are skew-symmetric Hamiltonian and symmetric Hamiltonian matrices, respectively. More precisely, let

$$\Phi: \text{Sym}_n \to \mathbb{LG}_n$$

denote the smooth map defined as

$$\Phi(K) := \begin{bmatrix} I \\ K \end{bmatrix} (I + K^2)^{-1} \left[ \begin{bmatrix} I \\ K \end{bmatrix} \right].$$  \hfill (12)

It is easily seen, that $\Phi$ maps the set of symmetric $n \times n$ matrices diffeomorphically onto an open, dense subset $\mathcal{C} \subset \mathbb{LG}_n$ and

$$D\Phi(K)R(K) = (I - P)AP + PA^\top (I - P)$$

$$= [\Omega, P] + [P, [P, S]].$$  \hfill (13)
The subset $\mathcal{C}$ is called the big coordinate cell of $LG_n$. In particular, we have the following result.

**Proposition 2.**

(a) The solutions of the differential Riccati equation (7) correspond bijectively via (10) to those solutions of

$$\dot{P} = [\Omega, P] + [P, [P, S]]$$

that start in the big coordinate cell of $LG_n$.

(b) The critical points of the smooth vector field $F_A : LG_n \to T LG_n$

$$F_A(P) = [\Omega, P] + [P, [P, S]]$$

are exactly the projection operators of Lagrangian invariant subspaces of $A$.

Note, that the above quadratic vector field defines generalizes Brockett’s double bracket flow $\dot{P} = [P, [P, S]]$. Therefore we refer to (15) as the **generalized double bracket flow**.

## 4 Newton’s method

Riemannian normal coordinates of the Lagrangian Grassmanian are defined through the matrix exponential map

$$\exp_P : T_P LG_n \to LG_n, \quad \exp_P(\xi) = e^{[\xi, P]} P e^{-[\xi, P]}.$$ (16)

Thus, given any $\Theta \in OSp_{2n}$ with

$$P = \Theta^T \begin{bmatrix} I_n & 0 \\ 0 & 0 \end{bmatrix} \Theta$$ (17)

and

$$[\xi, P] = \Theta^T \begin{bmatrix} 0 & -Z \\ Z & 0 \end{bmatrix} \Theta, \quad Z \in Sym_n$$ (18)

we obtain

$$\exp_P(\xi) = \Theta^T \begin{bmatrix} \cos Z & \sin Z \\ \sin Z & \cos Z \end{bmatrix} \Theta.$$ (19)

Given an arbitrary smooth vector field $F$ on $LG_n$ let $\nabla_\xi F$ denote the covariant derivative of $F$ at a point $P$ in direction $\xi \in T_P LG_n$. Assuming that $\xi \mapsto \nabla_\xi F(P)$ is invertible, there exists a unique tangent vector $\xi = N_F(P)$ with

$$\nabla_\xi F(P) = -F(P).$$ (20)
We refer to \( N_F \) as the **Newton vector field** for \( F \). In order to implement efficient versions of the Newton algorithm, we also consider more general *local parametrizations* for the Lagrange Grassmannian, namely the smooth maps

\[
\nu_P : T_P LG_n \to LG_n
\]

satisfying

\[
\nu_P(0) = P \text{ and } D \nu_P(0) = \text{id}.
\]  

We now introduce a class of Newton-like algorithms, following [3]. Let \( \nu_P, P \in LG_n \) be a family of smooth local parametrizations of \( LG_n \). Extending a result from [10, 3], we can prove the following theorem.

**Theorem 3.** Let \( P^* \in LG_n \) be an isolated equilibrium point of a vector field \( F \) on \( LG_n \) and let \( \nu_P : T_P LG_n \to LG_n \) be local parametrizations. Then the *generalized Riemannian Newton method*

\[
P_0 \in LG_n, \ P_{k+1} = \nu_{P_k}(N_F(P_k))
\]

converges locally quadratically to \( P^* \).

For the special choice, where \( \{\nu_P\}_{P \in LG_n} \) are the Riemannian normal coordinates, our iteration (23) is precisely the so-called Riemannian Newton method along geodesics

\[
P_{k+1} = \exp_{P_k}(N_F(P_k)) \in LG_n.
\]

5 **Algorithm Implementation**

We proceed by specifying the algorithm for the generalized double bracket vector field \( F_A \). In order to determine the Newton direction of \( F_A(P) \) we need the following formula for the covariant derivative. It is canonically defined by the Levi-Civita connection on the Lagrange Grassmannian and follows from a straightforward but lengthy computation.

**Proposition 4.** Let \( \Omega \in \mathfrak{osp}_{2n} \) be skew-symmetric Hamiltonian, \( S \in \text{Sym}_{2n} \cap \mathfrak{sp}_{2n} \) be symmetric Hamiltonian and let an arbitrary tangent element \( \xi \in T_P LG_n \) be given. Then the covariant derivative of the generalized double bracket vector field \( F_A(P) = [\Omega, P] + [P, [P, S]] \) is

\[
\nabla_\xi F_A(P) = \text{ad}_P^2 \circ \text{ad}_\Omega(\xi) - \text{ad}_P \circ \text{ad}_S(\xi), \quad \xi \in T_P LG_n.
\]  

In order to obtain efficient local parameterizations, we use an appropriate approximation of the matrix exponential function. Here, we focus on the *QR*-coordinates, defined by

\[
\nu_{P_k}^{QR}(\xi) := (I + [\xi, P_k])QP_k(I + [\xi, P_k])Q^\top,
\]
where $X_Q$ denotes the $Q$-factor of the $QR$ decomposition of $X$. Our main result now reads as follows.

**Theorem 5.** The generalized Riemannian Newton algorithm

$$P_{k+1} = \nu_{P_k}^{QR}(\xi) \quad \text{with}$$

$$[P_k, [P_k, \Omega, \xi]] - [P_k, [S, \xi]] = -[\Omega, P_k] - [P_k, [P_k, S]]$$

(26)

converges locally quadratically to projectors onto stable Lagrangian invariant subspaces of an arbitrary Hamiltonian matrix $A$.

Of course, the above algorithm is in a rather implicit form and does not detail how to solve the linear matrix equation, that defines the Newton direction $\xi$. In the sequel, we therefore derive a more simpler looking implementation. Theorem 5 shows local quadratic convergence for this algorithm.

**Newton Lagrangian invariant subspace algorithm**

**Step 1.** Pick an orthogonal matrix $\Theta_0 \in \text{OSp}_{2n}$ corresponding to

$$P_0 = \Theta_0^T \begin{bmatrix} I_n & 0 \\ 0 & 0 \end{bmatrix} \Theta_0 \in \mathbb{L}^n,$$

and set $j = 0$.

**Step 2.** Compute

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & -A_{11}^T \end{bmatrix} = \Theta_j A \Theta_j^T.$$

**Step 3.** Solve the Lyapunov equation $A_{11}^T Z_j + Z_j A_{11} = A_{21}$ for the symmetric matrix $Z_j \in \text{Sym}_n$.

**Step 4.** Compute

$$\Theta_{j+1}^T = \Theta_j^T \begin{bmatrix} I_n & -Z_j \\ Z_j & I_n \end{bmatrix}_Q \quad \text{and} \quad P_{j+1} = \Theta_{j+1}^T \begin{bmatrix} I_n & 0 \\ 0 & 0 \end{bmatrix} \Theta_{j+1}.$$

**Step 5.** Set $j = j + 1$ and goto Step 2.

The above algorithm can, of course, also be applied to solve the algebraic Riccati equation from linear optimal control $F^T K + K F - K G G^T K + H^T H = 0$, simply by applying it to the associated Hamiltonian matrix (8). For the sake of easier comparison, we present Kleinman’s algorithm [7] on the noncompact symplectic group.

**Kleinman’s Algorithm**

**Step 1.** Pick a symplectic matrix of the form

$$S_0 = \begin{bmatrix} I_n & 0 \\ K_0 & I_n \end{bmatrix},$$
with $K_0 = K_0^\top$ and set $j = 0$.

**Step 2.** Compute
\[
\begin{bmatrix}
  A_{11} & A_{12} \\
  A_{21} & -A_{11}^\top
\end{bmatrix}
= S_j^{-1} A S_j.
\]

**Step 3.** Solve the Lyapunov equation $A_{11}^\top Z_j + Z_j A_{11} = A_{21}$ for the symmetric matrix $Z_j \in \text{Sym}_n$.

**Step 4.** Compute
\[
S_{j+1} = S_j \begin{bmatrix} I_n & 0 \\ Z_j & I_n \end{bmatrix}.
\]

**Step 5.** Set $j = j + 1$ and go to Step 2.

We note, that a major difference of our algorithm to Kleinman’s lies in the orthogonal update step 4, which seems advantageous from a numerical stability point of view. Moreover, we believe, that these two algorithms are different on $\text{LG}_n$. Evidence for this distinction arises from the fact that Kleinman’s algorithm is just the Newton algorithm
\[
K_{t+1} = K_t - D R(K_t)^{-1} R(K_t)
\]
for the standard Euclidean Newton vector field of the Riccati vector field $R(K)$. It is thus the Riemannian Newton algorithm on $\text{Sym}_n$ with respect to the standard inner product $\langle \xi, \eta \rangle = \text{tr}(\xi \eta)$. In contrast, the Riemannian Newton algorithm on $\text{LG}_n$ is expressed with respect to the normal Riemannian metric. It can be seen, that this metric pulls back to the Riemannian metric on symmetric matrices
\[
\langle \xi, \eta \rangle_K = 2 \text{tr} (\xi (I + K^2)^{-1}\eta (I + K^2)^{-1})
\]
Thus one would expect that the Riemannian Newton with respect to this metric differs from the standard Newton method.

## 6 A CG-algorithm for vector fields

Conjugate gradient (CG) algorithms are commonly used for minimizing smooth functions. In order to intrinsically fulfill certain constraints, the concept of CG-methods has been lifted to complete Riemannian manifolds, cf. [11]. For minimizing a real valued function $f$ on a complete Riemannian manifold $(M, \langle \cdot, \cdot \rangle)$, these CG-methods consist of iterating sweeps, which schematically are as follows.

**CG-Sweep.** Let $p_0 \in M$ be given.
Set $H_0 := \nabla f(p_0)$ (the Riemannian gradient). Then for $i = 0, ..., n-1$ ($n := \dim M$)

- (Line-Search) Compute $\lambda_i \in \mathbb{R}$ and set $p_{i+1} = \exp_{p_i}(\lambda_i H_i)$;
- (Direction) Compute $H_{i+1}$ tangent to $p_{i+1}$ according to a Riemannian adaptation of the Hestenes-Stiefel-Formula or another formula known from the Euclidian case, cf. [8].
For optimization problems, the line search consists of minimizing, or at least approximately minimizing the restriction of $f$ to $\exp_{p_i}(\mathbb{R}H_i)$. Having in mind a vector field adaptation, the following line search turns out to be appropriate. Let $p_i \in M$ and $H(p_i): T_p M \times T_p M \rightarrow \mathbb{R}$ be the Riemannian Hessian at $p_i$. For $H_i \in T_p M$ we define the step-size in direction $H_i$ as

$$
\lambda_i := -\frac{d}{dt} \bigg|_{t=0} f \circ \exp_{p_i}(tH_i) = -\frac{\langle \nabla f(p_i), H_i \rangle}{H(p_i)(H_i, H_i)},
$$

(29)

Hence $\lambda_i$ can be interpreted as a one-dimensional Newton step along $f \circ \exp_{p_i}(\mathbb{R}H_i)$. Since we consider our optimization task to be a local one, we can expect that a one-dimensional Newton-step locally yields acceptable convergence results. Note, that in the Euclidian space, for a strictly convex quadratic function and for $\exp_{p_i}(tH_i) = p_i + tH_i$, this corresponds to an exact line search.

We abbreviate $\nabla f(p_i) = G_i$. The direction update is given by

$$
H_{i+1} = -G_{i+1} + \gamma_i \tau H_i,
$$

where $\tau$ denotes the Riemannian parallel transport of some $H \in T_p M$ to $\tau H \in T_{p_{i+1}} M$ along the geodesic curve $\exp_{p_i}(tH_i)$. Manifold adaptions for the two most common update formulas are

$$
\gamma_i^{HS} = \frac{\langle G_{i+1}, G_{i+1} - \tau G_i \rangle}{\langle \tau H_i, G_{i+1} - \tau G_i \rangle} \text{ (Hestenes-Stiefel)} \tag{30}
$$

$$
\gamma_i^{PR} = \frac{\langle G_{i+1}, G_{i+1} - \tau G_i \rangle}{\|G_i\|^2} \text{ (Polak-Ribi`ere)} \tag{31}
$$

We refer to [8] for a motivation of these formulas in the Euclidian case.

We now propose an extension of the above ideas to find zeros of a vector field on a Riemannian manifold $M$.

**CG-Sweep for vector fields.** Let $F: M \rightarrow TM$ be a smooth vector field. Let $p_0 \in M$ be given and set $H_0 := F(p_0)$. Then for $i = 0, ..., n - 1 \ (n := \dim M)$

- **(Line-Search)** Set $p_{i+1} = \exp_{p_i}(\lambda_i H_i)$ with

$$
\lambda_i := -\frac{\langle F(p_i), H_i \rangle}{\langle \nabla_{H_i} F(p_i), H_i \rangle},
$$

(32)

where $\nabla_{H_i} F(p_i)$ denotes the covariant derivative of $F$ at $p_i$ in direction $H_i$.

- **(Direction)** Set

$$
H_{i+1} = -F(p_{i+1}) + \gamma_i \tau H_i,
$$

(33)

where $\gamma_i$ is chosen for example via the Polak-Ribi`ere formula as

$$
\gamma_i^{PR} = \frac{\langle F(p_{i+1}), F(p_{i+1}) - \tau F(p_i) \rangle}{\|F(p_i)\|^2}.
$$

(34)
7 Computing Lagrangian invariant subspaces

The ideas of the previous section are now used to compute a Lagrangian invariant subspace by applying the CG-algorithm to the generalized double bracket vector field

\[ F_A : \mathcal{L}G_n \to T\mathcal{L}G_n, \]
\[ F_A(P) = [\Omega, P] + [P, [P, S]]. \]  

Note the invariance property \( F_A(\theta P \theta^T) = \theta F_A(P) \theta^T \) for all \( P \in \mathcal{L}G_n, \ \theta \in \text{OSp}_{2n}, \) and that at

\[ P_I := \begin{bmatrix} I_n & 0 \\ 0 & 0 \end{bmatrix} \]

the vector field yields

\[ F_A(P_I) = \begin{bmatrix} 0 & A_{21} \\ A_{21} & 0 \end{bmatrix}. \]  

Lemma. Let \( H = \begin{bmatrix} 0 & Z \\ Z & 0 \end{bmatrix} \in T_{P_I} \mathcal{L}G. \) Then \( \langle \nabla_H F(P_I), H \rangle = -4\text{tr}A_{11}Z^2. \)

Proof. Let \( \Omega =: \begin{bmatrix} \Omega_{11} & \Omega_{12} \\ -\Omega_{12} & \Omega_{11} \end{bmatrix} \) and \( S =: \begin{bmatrix} S_{11} & S_{12} \\ S_{12} & -S_{11} \end{bmatrix}. \) By Proposition 4,

\[ \nabla_H F(P_I) = [P_I, [P_I, \Omega, H]] - [P_I, [S, H]] \]
\[ = \begin{bmatrix} 0 & \Omega_{11}Z - Z\Omega_{11} \\ \Omega_{11}Z - Z\Omega_{11} & 0 \end{bmatrix} - \begin{bmatrix} 0 & S_{11}Z + ZS_{11} \\ S_{11}Z + ZS_{11} & 0 \end{bmatrix}. \]

Since \( S_{11} + \Omega_{11} = A_{11} \) and \( S_{11} - \Omega_{11} = A_{11}^\top, \) the covariant derivative is

\[ \nabla_H F(P_I) = \begin{bmatrix} 0 & -ZA_{11} - A_{11}^\top Z \\ -ZA_{11} - A_{11}^\top Z & 0 \end{bmatrix} \]
and the claim follows. \( \square \)

In order to achieve an efficient implementation of the CG algorithm, we choose the QR-coordinates \( \nu_{P}^{QR} \) to approximate the Riemannian parallel transport of \( \xi \) along geodesics

\[ \tau^{\exp} = \exp(t[P, H])\xi \exp(-t[P, H]) \]  

by the map

\[ \tau : T_P \mathcal{L}G_n \to T_{\nu_{P}^{QR}(tH)} \mathcal{L}G_n, \]
\[ \xi \mapsto (I_{2n} + t[P, H])_Q \xi (I_{2n} + t[P, H])_Q^\top. \]

The following routine is one sweep in the algorithm for computing a Lagrangian invariant subspace of a Hamiltonian matrix. The input is a Hamiltonian matrix \( A \) and a symplectic and orthogonal \( \theta. \)

The algorithm for computing a Lagrangian invariant subspace of a Hamiltonian matrix now consists of iterating sweeps.
CG Lagrangian invariant subspace algorithm - Sweep

**Step 1.** Set $Z_0 = A_{21}$ and set $i = 0$. For $i = 0, \ldots, \frac{1}{2}n(n+1) - 1$ iterate the following steps:

**Step 2.** Set

$$
\lambda_i = -\frac{\text{tr}(A_{21}^{(i)}Z_i)}{2\text{tr}(A_{11}^{(i)}(Z_i)^2)}
$$

and compute

$$
\tilde{\theta} = \begin{bmatrix}
    I_n & -\lambda_i Z_i \\
    \lambda_i Z_i & I_n
\end{bmatrix}_Q.
$$

**Step 3.** Set $A^{(i+1)} := \tilde{\theta}^T A^{(i)} \tilde{\theta}$ and $\theta^{(i+1)} := \theta^{(i)} \tilde{\theta}$.

**Step 4.** Set

$$
\gamma_i^{(PR)} = -\frac{\text{tr}(A_{21}^{(i+1)}(A_{21}^{(i+1)} - A_{21}^{(i)}))}{\text{tr}((A_{21}^{(i)})^2)}
$$

and set $Z_{i+1} = -A_{21}^{(i+1)} + \gamma_i^{(PR)} Z_i$.

Alternatively, an implementation of the Hestenes-Stiefel formula

$$
\gamma_i^{(HS)} := -\frac{\text{tr}(A_{21}^{(i+1)}(A_{21}^{(i+1)} - A_{21}^{(i)}))}{\text{tr}(Z^{(i)}(A_{21}^{(i+1)} - A_{21}^{(i)}))}
$$

instead of $\gamma_i^{(PR)}$ is possible.

## 8 Numerical Simulations and Discussion

We implemented the CG-algorithm from the last section in Mathematica 5.2. For a symmetric Hamiltonian matrix $A$, the algorithm performs quite well and in that case, the simulations support $N$-step local quadratic convergence, cf. Table 1. This is not too surprising, since in the case of symmetric matrices the used vector field is a gradient vector field indeed and the CG-algorithm is a conjugate gradient algorithm for the function $P \rightarrow \text{tr}(PA)$.

### Table 1.

As a measure of convergence, the norm $\|A_{21}\|$ is returned after each sweep. Here, $A \in \mathbb{R}^{80 \times 80}$.

<table>
<thead>
<tr>
<th>sweep</th>
<th>$\gamma_i^{HS}$</th>
<th>$\gamma_i^{PR}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>19.7888</td>
<td>19.7888</td>
</tr>
<tr>
<td>1</td>
<td>1.9 \cdot 10^{-10}</td>
<td>8.2 \cdot 10^{-7}</td>
</tr>
<tr>
<td>2</td>
<td>7.3 \cdot 10^{-17}</td>
<td>8.3 \cdot 10^{-16}</td>
</tr>
</tbody>
</table>

In the unsymmetric case, convergence could only be observed for small $n$. Figure 1 illustrates the convergence for the unsymmetric case where $n = 10$. In both examples, the entries of the Hamiltonian matrix $A$ were uniformly randomly chosen from $[0,1]$. Further investigation of the unsymmetric case is needed to result in an improved and reliable algorithm.
Figure 1. Convergence behavior of the CG algorithm applied to an unsymmetric Hamiltonian $20 \times 20$ matrix.

Bibliography


