AN INVERSE PROBLEM FOR HAMILTONIAN SYSTEMS

IAN W. KNOWLES

Dedicated to Professor Michael Eastham on the occasion of his 65th birthday.

Abstract. In a linear Hamiltonian system for which the Dirichlet principle is valid, solutions to boundary value problems can be identified as the unique minimizers of the quadratic functional associated with the system. The inverse problem, in which coefficient functions in the differential equations are identified as unique minimizers of a related functional, is discussed, together with conditions under which recovery can occur.

1. Introduction

We consider linear Hamiltonian systems of the form
\[-Jx' = H(t)x, \quad \alpha \leq t \leq \beta,\] (1.1)
where \(J\) is a constant invertible skew-symmetric matrix of order \(2n\) defined by
\[J = \begin{pmatrix} 0 & -I_n \\ I_n & 0 \end{pmatrix},\] (1.2)
and
\[x = \begin{pmatrix} y \\ z \end{pmatrix}, \quad H(t) = \begin{pmatrix} A(t) & B^T(t) \\ B(t) & C(t) \end{pmatrix},\] (1.3)
are corresponding partitions of the vector \(x\) and the symmetric matrix \(H\); here, \(A\) and \(C\) are symmetric \(n \times n\) matrices, and (1.1) may be rewritten in the explicit system form
\[y' = -B(t)y - C(t)z,\] (1.4)
\[z' = A(t)y + B^T(t)z.\] (1.5)

Here, and in the sequel, all entries in the Hamiltonian matrices are assumed to be \(L^1[\alpha, \beta]\) functions. The basic theory of such systems may be found in [1, Ch. 2] and [2, p. 384]. Standard examples include selfadjoint linear homogeneous differential equations of order \(2n\), commonly written in the form
\[L_{2n}(u) = \sum_{k=0}^{n} (-1)^k [p_k(t)u^{(k)}]^{(k)} = 0, \quad \alpha \leq t \leq \beta,\] (1.6)
with \(p_n > 0\); if we set \(y = (y_k)\) and \(z = (z_k), 1 \leq k \leq n,\) where \(y_k = u^{(k-1)}\) and
\[z_k(t) = \sum_{j=k}^{n} (-1)^{j-k} [p_j(t)u^{(j)}]^{(j-k)} ,\]
then the Hamiltonian matrix is given by

\[ A = \text{diag} \{ p_0, p_1, \ldots, p_{n-1} \}, \quad C = -p_1^{-1} \text{diag} \{ 0, 0, \ldots, 1 \}, \]

\[ B = -\begin{pmatrix}
0 & 1 & 0 & \ldots & 0 \\
0 & 0 & 1 & \ldots & 0 \\
0 & 0 & 0 & \ldots & 1 \\
0 & 0 & 0 & \ldots & 0 \\
\end{pmatrix}. \]

(1.7)  (1.8)

Another special case is the matrix Sturm-Liouville equation,

\[-(P(t)y'(t))' + Q(t)y(t) = 0, \quad \alpha \leq t \leq \beta, \]

where \( y \) is an \( n \)-vector and \( P \) and \( Q \) are symmetric \( n \times n \) matrices with \( P \) positive definite; here, \( z = Py' \), and

\[ H(t) = \begin{pmatrix} Q(t) & 0 \\ 0 & -P^{-1}(t) \end{pmatrix}. \]

(1.9)

It is known [2, Ch. XI, §6] that, if the function \( Q \) is in \( L^1[\alpha, \beta] \), and the Sturm-Liouville equation

\[-u'' + Q(x)u = 0, \]

is disconjugate on \( [\alpha, \beta] \), then not only can Dirichlet boundary value problems involving this equation be solved uniquely, but also the Dirichlet principle holds, so that these solutions can be obtained as unique minimizers of the Dirichlet functional.

In [3] an associated inverse problem, involving the unique recovery of \( Q \) from \( u \) by minimization, was considered. In this context, as \( u'' = Qu \), we have a new approach to the old problem of numerical differentiation. To be more specific as to the method, assume that a positive solution \( u \) of (1.10) is given, and let \( E \) denote the set of functions \( q \) in \( L^1[\alpha, \beta] \) for which the equation

\[ A_qv = -v'' + q(x)v = 0, \]

is disconjugate on \( [\alpha, \beta] \), i.e. every solution of (1.11) has no more than one zero on this interval. For \( q \) in \( E \), define the functional

\[ G(q) = \int_\alpha^\beta (u' - u_q')^2 + q(x)(u - u_q)^2 \, dx \]

(1.10)  (1.11)  (1.12)

where \( v = u_q \) is the solution of the boundary value problem consisting of equation (1.11) together with the boundary conditions

\[ v(\alpha) = u(\alpha), \quad v(\beta) = u(\beta). \]

(1.13)

In particular, \( u_Q = u \). It is shown in [3] that, for all \( q \) in \( E \)

\[ G(q) \geq G(Q) = 0, \]

and that the second Fréchet differential of \( G \) is given by

\[ G''(q)[h, h] = 2 \left( A_q^{-1}(u_q h), u_q h \right) \]

(1.14)

for each \( h \) in \( L^1[\alpha, \beta] \), where \( A_q^{-1} \) denotes the inverse of the operator \( A_q \) associated with equ. (1.11) and homogeneous Dirichlet boundary conditions. Further, it follows from the positivity of \( A_q \) that \( G''(q)[\cdot, \cdot] \) is a positive definite quadratic form for each \( q \) in \( E \), so that \( G \) is a strictly convex functional on \( E \). This means that \( Q \) may be recovered as the unique global minimum of the functional \( G \).
In contemplating extensions of this work to more general Hamiltonian systems, one must be mindful at the outset of an intrinsic obstruction that is already present in second order equations. Consider the possibility of recovering \( P \) and \( Q \) from a knowledge of the solutions \( y \) of the equation
\[
-(P(x)y')' + Q(x)y = 0, \quad 0 \leq x \leq 1. \tag{1.15}
\]
If we change the independent variable to \( t \) defined by
\[
t = \int_0^x \frac{ds}{P(s)} \quad ds, \quad z(t) = y(x(t)),
\]
we obtain the equation
\[
-z''(t) + P(x(t))Q(x(t))z(t) = 0, \quad 0 \leq t \leq \gamma,
\]
where \( \gamma = \int_0^1 \frac{ds}{P(s)} \). It is clear that, even with a knowledge of the solutions \( z \), one should expect to recover at best only the product \( PQ \), and not the functions \( P \) and \( Q \) individually.

We are interested in the problem of recovering the Hamiltonian matrix \( H \) from a knowledge of the partitions \( y \), or \( z \), of the solutions \( x = (y, z)^T \) of (1.1). Given the above observation, it would seem that the recovery of all of \( H \) in this manner could be a little too ambitious; in particular, the simultaneous recovery of \( A \) and \( C \) could lead to difficulties. It is their separate identification that we pursue here.

Consider first the recovery of the symmetric matrix of functions \( A(t) \). To this end, let vectors \( y_i, 1 \leq i \leq n \), be given satisfying the boundary conditions
\[
y_i(\alpha) = 0, \quad y_i(\beta) = \eta_i, \tag{1.16}
\]
where the set \( \{\eta_i : 1 \leq i \leq n\} \) is linearly independent in \( \mathbb{R}^n \), and the vectors \( x_i = (y_i, z_i)^T, 1 \leq i \leq n \), are solutions of (1.1). Let \( D \) denote the set of all \( n \times n \) symmetric matrices \( a(t) \) such that the linear Hamiltonian system
\[
-Jx' = \begin{pmatrix} a(t) & B(t) \\ B^T(t) & C(t) \end{pmatrix} x, \quad \alpha \leq t \leq \beta, \tag{1.17}
\]
is disconjugate on \([\alpha, \beta]\), i.e. for no solution \( x = (y, z)^T \) of (1.17) does the vector \( y(t) \) vanish more than once on this interval; we note in passing that, as a consequence, the system (1.17) also satisfies the condition \([C]\) of [1], namely that for no non-trivial solution \( x = (y, z)^T \) of (1.17) does the vector \( y \) vanish on any subinterval of \([\alpha, \beta]\). For \( a \in D \) define
\[
G_i(a) = \int_\alpha^\beta (y_i - y_{i,a})^T a(t)(y_i - y_{i,a}) - (z_i - z_{i,a})^T C(t)(z_i - z_{i,a}) \, dt, \tag{1.18}
\]
where \( x = x_{i,a} = (y_{i,a}, z_{i,a})^T \) is the unique solution of (1.17) satisfying the boundary conditions
\[
y_{i,a}(\alpha) = y_i(\alpha), \quad y_{i,a}(\beta) = y_i(\beta). \tag{1.19}
\]

Then we have the following theorem:

**Theorem 1.1.** Assume that the matrix \( C(t) \) is non-positive definite for \( \alpha \leq t \leq \beta \), and let vectors \( y_i, 1 \leq i \leq n \), be given so that the boundary conditions (1.16) are satisfied and \( x_i = (y_i, z_i)^T, 1 \leq i \leq n \), are solutions of (1.1). Then the functional
\[
G(a) = \sum_{i=1}^n G_i(a) \tag{1.20}
\]
is non-negative and strictly convex on \( D \), with a unique global minimum at \( a = A \).

Theorem 1.1 is but one variation on this theme. One could, as a second example, consider the recovery of the matrix of functions \( C \) from a knowledge of solutions of (1.1). Again, let vectors \( z_i, 1 \leq i \leq n \), be given satisfying the boundary conditions

\[
z(\alpha) = 0, \quad z(\beta) = \eta_i, \tag{1.21}
\]

where the set \( \{ \eta_i : 1 \leq i \leq n \} \) is linearly independent in \( \mathbb{R}^n \), and the vectors \( x_i = (y_i, z_i)^T, 1 \leq i \leq n \), are solutions of (1.1). Let \( \tilde{D} \) denote the set of all \( n \times n \) symmetric matrices \( c(t) \) such that the linear Hamiltonian system

\[
-Jx' = \begin{pmatrix} A(t) & B(t) \\ B(t) & c(t) \end{pmatrix} x, \quad \alpha \leq t \leq \beta,
\]

(1.22)
is skew-disconjugate on \([\alpha, \beta] \), i.e. the system obtained from (1.22) by interchanging \( y \) and \( z \) is disconjugate in the usual sense. Skew disconjugate systems can be obtained, for example, from equations of the form (1.6) in the case \( p_0 = p_1 = \ldots = p_{n-2} = 0, p_{n-1} > 0 \), by interchanging \( p_n \) and \( p_{n-1} \) and requiring that the new system be disconjugate according to the usual definition. For \( c \in \tilde{D} \) define

\[
F_i(c) = \int_{\alpha}^{\beta} (y_i - y_{i,c})^T A(t)(y_i - y_{i,c}) - (z_i - z_{i,c})^T c(t)(z_i - z_{i,c}) dt, \quad \tag{1.23}
\]

where \( x = x_{i,c} = (y_{i,c}, z_{i,c})^T \) is the solution of (1.22) satisfying the boundary conditions

\[
z_{i,c}(\alpha) = z_i(\alpha), \quad z_{i,c}(\beta) = z_i(\beta). \quad \tag{1.24}
\]

Then we arrive at

**Theorem 1.2.** Assume that the matrix \( A(t) \) is non-negative definite for \( \alpha \leq t \leq \beta \), and let vectors \( z_i, 1 \leq i \leq n \), be given so that the boundary conditions (1.21) are satisfied and \( x_i = (y_i, z_i)^T, 1 \leq i \leq n \), are solutions of (1.1). Then the functional

\[
F(c) = \sum_{i=1}^{n} F_i(c) \tag{1.25}
\]
is non-negative and strictly convex on \( \tilde{D} \), with a unique global minimum at \( c = C \).

One can also consider Hamiltonian matrices of the form

\[
\hat{H}_\lambda(t) = \begin{pmatrix} A_1(t) + \lambda A_2(t) & B(t) \\ B(t) & C(t) \end{pmatrix}, \tag{1.26}
\]

and recover two symmetric matrix functions, \( A_1 \) and \( A_2 \) from a knowledge of solutions of the equation

\[
-Jx' \hat{=} \hat{H}_\lambda(t)x, \quad \alpha \leq t \leq \beta. \tag{1.27}
\]

Let vectors \( y_{i,\lambda}, 1 \leq i \leq n \), be given satisfying the boundary conditions (1.16) and such that \( x_{i,\lambda} = (y_{i,\lambda}, z_{i,\lambda})^T, 1 \leq i \leq n \), are solutions of (1.27) for \( \lambda = \lambda_1 \) and \( \lambda = \lambda_2 \neq \lambda_1 \). Let \( \hat{D} \) denote the set of pairs of \( n \times n \) symmetric matrices \((a_1(t), a_2(t))\) such that the linear Hamiltonian system

\[
-Jx' = \begin{pmatrix} a_1(t) + \lambda a_2(t) & B(t) \\ B(t) & C(t) \end{pmatrix} x, \quad \alpha \leq t \leq \beta, \tag{1.28}
\]
is disconjugate on \([\alpha, \beta]\). For \((a_1, a_2) \in \hat{D}\) define

\[
G_{ij}(a_1, a_2) = \int_{\alpha}^{\beta} \left( (y_{i,\lambda_j} - y_{i,\lambda_j,a})^T (a_1(t) + \lambda_j a_2(t))(y_{i,\lambda_j} - y_{i,\lambda_j,a}) - (z_{i,\lambda_j} - z_{i,\lambda_j,a})^T C(t)(z_{i,\lambda_j} - z_{i,\lambda_j,a}) \right) dt, \tag{1.29}
\]

where \(x = x_{i,\lambda_j,a} = (y_{i,\lambda_j,a}, z_{i,\lambda_j,a})^T\) is the solution of (1.17) satisfying the boundary condition

\[
y_{i,\lambda_j,a}(\alpha) = y_{i,\lambda_j}(\alpha), \quad y_{i,\lambda_j,a}(\beta) = y_{i,\lambda_j}(\beta). \tag{1.30}
\]

Then we have

**Theorem 1.3.** Assume that the matrix \(C(t)\) is non-positive definite for \(\alpha \leq t \leq \beta\), and let vectors \(y_{i,\lambda_j}, 1 \leq i \leq n\), be given satisfying the boundary conditions (1.16) and such that \(x_{i,\lambda_j,a} = (y_{i,\lambda_j,a}, z_{i,\lambda_j,a})^T\), \(1 \leq i \leq n\), are solutions of (1.27) for \(\lambda = \lambda_1\) and \(\lambda = \lambda_2 \neq \lambda_1\). Then

\[
G(a_1, a_2) = \sum_{i=1}^{n} \sum_{j=1}^{2} G_{ij}(a_1, a_2) \tag{1.31}
\]

is a non-negative and strictly convex functional on \(\hat{D}\), having a unique global minimum at \(a_1 = A_1\) and \(a_2 = A_2\).

Finally, one can focus attention on special classes of the matrices \(A\) and \(C\). For example, one can consider the situation in which the matrix \(A(t)\) takes the form

\[
A(t) = \begin{pmatrix}
Q(t) & 0 & \ldots & 0 \\
0 & 0 & \ldots & 0 \\
o & 0 & \ldots & 0 \\
o & 0 & \ldots & 0
\end{pmatrix}. \tag{1.32}
\]

As there are fewer functions to recover in \(A\), we require correspondingly fewer known partitions of solutions of (1.1). In this case, let \(y\) be a given vector such that \(x = (y, z)^T\) is a non-trivial solution of (1.1), where \(A\) has the form (1.32). Let \(D_1\) denote the set of all \(n \times n\) symmetric matrices \(a(t)\) of the form

\[
a(t) = \begin{pmatrix}
q(t) & 0 & \ldots & 0 \\
0 & 0 & \ldots & 0 \\
o & 0 & \ldots & 0 \\
o & 0 & \ldots & 0
\end{pmatrix}. \tag{1.33}
\]

such that the linear Hamiltonian system

\[-Jx' = \begin{pmatrix}
a(t) & B^T(t) \\
B(t) & C(t)
\end{pmatrix} x, \quad \alpha \leq t \leq \beta, \tag{1.34}\]

is disconjugate on \([\alpha, \beta]\). For \(a \in D_1\) define

\[
G(a) = \int_{\alpha}^{\beta} \left( (y - y_a)^T a(t)(y - y_a) - (z - z_a)^T C(t)(z - z_a) \right) dt, \tag{1.35}
\]

where \(x = x_a = (y_a, z_a)^T\) is the unique solution of (1.34) satisfying the boundary condition

\[
y_a(\alpha) = y(\alpha), \quad y_a(\beta) = y(\beta). \tag{1.36}
\]

Then, as a final result, we have
Theorem 1.4. Assume that the matrix $C(t)$ is non-positive definite for $\alpha \leq t \leq \beta$, and a vector $y$ is given such that $x = (y, z)^T$ is a solution of (1.1). Then the functional $G$ defined by (1.35) is non-negative and strictly convex on $D_1$, with a unique global minimum when $q = Q$.

2. Basic Theory

For ease of reference, we list here some standard material, as well as some other ancillary results. The first theorem is essentially [1, Proposition 2] and provides for us the comforting fact that, as long as we restrict attention to disconjugate equations, boundary value problems may always be solved uniquely.

Theorem 2.1. The linear Hamiltonian system (1.17) is disconjugate on $I = [\alpha, \beta]$ if and only if for every two distinct points $t_1$ and $t_2$ in $I$ and arbitrary $n$-vectors $\eta_1$ and $\eta_2$ there is a unique solution $x = (y, z)^T$ such that $y(t_1) = \eta_1$, $y(t_2) = \eta_2$.

If we define the distance between two equations of the form (1.17) with Hamiltonian matrices

$$H_1 = \begin{pmatrix} a_1 & B^T \\ B & C \end{pmatrix}, \quad H_2 = \begin{pmatrix} a_2 & B^T \\ B & C \end{pmatrix},$$

to be

$$\int_{\alpha}^{\beta} |a_1(t) - a_2(t)| \, dt$$

and we further assume that the $n \times n$ matrix $C$ is negative definite, then, by [1, Proposition 14], the set $D$ defined above, is both open and convex, as are the sets $\tilde{D}$, $\hat{D}$, and $D_1$ by an analogous argument.

It is advantageous to consider also the associated matrix equations

$$Y' = -B(t)Y - C(t)Z$$

$$Z' = A(t)Y + B^T(t)Z,$$

where $Y(t)$ and $Z(t)$ are $n \times n$ matrix functions of $t$. A solution $(Y, Z)$ of this system is called isotropic if

$$Y^T(t)Z(t) - Z^T(t)Y(t) = 0. \quad (2.2)$$

We then have [1, Theorem 1, p. 36]

Theorem 2.2. Suppose that $C(t) \geq 0$ for all $\alpha \leq t \leq \beta$. If the equation (1.1) is disconjugate on $[\alpha, \beta]$, then there exists an isotropic solution $(Y, Z)$ of (2.1) such that $Y(t)$ is invertible for all $\alpha \leq t \leq \beta$.

Also of interest is the following

Lemma 2.3. Let $(Y_0, Z_0)$ be an isotropic solution of (2.1) such that $Y_0(t)$ is invertible for all $\alpha \leq t \leq \beta$. Then all solutions $(Y, Z)$ of (2.1) are given by

$$Y(t) = Y_0(t)[M + S_0(t)N], \quad \quad \quad (2.3)$$

$$Z(t) = Z_0(t)[M + S_0(t)N] + (Y_0^T)^{-1}(t)N, \quad \quad \quad (2.4)$$

where $M$ and $N$ are arbitrary constant matrices and

$$S_0(t) = \int_{\alpha}^{t} Y_0^{-1}(s)C(s)(Y_0^T)^{-1}(s) \, ds. \quad \quad \quad (2.5)$$
This is [1, Proposition 1, p. 35]. We also need [1, Proposition 2, p. 38].

Lemma 2.4. Let \((Y_0, Z_0)\) be an isotropic solution of (2.1) such that \(Y_0(t)\) is invertible for all \(\alpha \leq t \leq \beta\) and assume that \(C(t) \geq 0\) for all \(\alpha \leq t \leq \beta\), and that for every non-trivial solution \((y(t), z(t))\) of (1.1), \(y(t)\) does not vanish identically in any subinterval of \([\alpha, \beta]\) (equivalently, condition \([C]\) in [1, p. 36] holds). Then the symmetric matrix

\[
S_0(t) = \int_\alpha^t Y_0^{-1}(s)C(s)(Y_0^T)^{-1}(s)\, ds
\]

is a strictly increasing function of \(t\).

The quadratic form associated with the Hamiltonian system (1.17) is defined by

\[
Q(x) = \int_\alpha^\beta y^T(t)a(t)y(t) - z^T(t)C(t)z(t)\, dt; \quad (2.6)
\]

a vector \(x = (y, z)^T\) is called admissible with respect to this form if \(y(\alpha) = y(\beta) = 0\).

We have, from [1, Theorem 12],

Theorem 2.5. If \(C(t)\) is non-positive definite over \([\alpha, \beta]\), and \(a \in \mathcal{D}\), then \(Q\) is a positive definite quadratic form with respect to all admissible vectors \(x = (y, z)^T\), and, if we set \(Lx = -Jx' - Hx\), then

\[
Q(x) = \int_\alpha^\beta x^T Lx \, dt. \quad (2.7)
\]

3. Uniqueness

In an inverse problem, perhaps the first question that arises naturally is that of uniqueness, i.e. when does the given data uniquely specify the quantity whose recovery is under consideration. With respect to the recovery of the matrix \(A\) we have

Theorem 3.1. If the matrices \(A_1\) and \(A_2\) both give rise to the vector solutions \(x_i = (y_i, z_i)^T, 1 \leq i \leq n\), of the Hamiltonian system (1.1), where the given set of vectors \(\{y_i(t) : 1 \leq i \leq n\}\) spans \(\mathbb{R}^n\) for all \(\alpha \leq t \leq \beta\), then \(A_1 = A_2\).

Proof. Notice that, for \(1 \leq i \leq n\),

\[
z'_i = A_1(t)y_i + B^T(t)z_i
\]

\[
z'_i = A_2(t)y_i + B^T(t)z_i,
\]

so that, after subtraction,

\[
(A_1(t) - A_2(t))y_{i,\lambda} = 0,
\]

and the result follows. \(\square\)

In similar fashion, unique recovery of \(C\) is guaranteed, provided that the corresponding set \(\{z_i(t) : 1 \leq i \leq n\}\) has the same spanning property. For the third situation considered above, we have
Theorem 3.2. If the matrix pairs \((A_1, A_2)\) and \((\tilde{A}_1, \tilde{A}_2)\) both give rise to the \(2n\) solutions \(x_{i,\lambda_j} = (y_{i,\lambda_j}, z_{i,\lambda_j})^T\), \(1 \leq i \leq n, 1 \leq j \leq 2\), of the Hamiltonian system (1.27), where each of the sets \(S_j = \{y_{i,\lambda_j}(t) : 1 \leq i \leq n\}, 1 \leq j \leq 2\), spans \(\mathbb{R}^n\) for all \(\alpha \leq t \leq \beta\), then \(A_1 = \tilde{A}_1\) and \(A_2 = \tilde{A}_2\).

Proof. For \(1 \leq i \leq n\) and \(1 \leq j \leq 2\) we have

\[
z_{i,\lambda_j}' = (A_1(t) + \lambda_j A_2(t)) y_{i,\lambda_j} + B^T(t) z_{i,\lambda_j} + B(t) z_{i,\lambda_j},
\]

so that for \(1 \leq i \leq n\) and \(1 \leq j \leq 2\) we have

\[
[(A_1 - \tilde{A}_1) + \lambda_j(A_2 - \tilde{A}_2)] y_{i,\lambda_j}.
\]

It follows that \((A_1 - \tilde{A}_1) + \lambda_j(A_2 - \tilde{A}_2) = 0\) for \(j = 1, 2\), and hence that \(A_1 = \tilde{A}_1\) and \(A_2 = \tilde{A}_2\).

\[\Box\]

4. The Functionals \(G_i\)

We first gather some of the more useful properties of the functionals \(G_i\) defined in (1.18):

Theorem 4.1. (i) For all \(a\) in \(\mathcal{D}\), \(G_i(a) \geq G_i(A) = 0\) and

\[
G_i(a) = Q(x_i) - Q(x_{i,a}). \tag{4.1}
\]

(ii) For all \(n \times n\) symmetric matrices \(v\) of \(\mathcal{L}^1[\alpha, \beta]\) functions,

\[
G_i'(a)[v] = \int_\alpha^\beta y_i^T(v(t)y_i - y_{i,a}v(t)y_{i,a}) dt. \tag{4.2}
\]

(iii) For all \(n \times n\) symmetric matrices of \(\mathcal{L}^1[\alpha, \beta]\) functions \(v\) and \(w\)

\[
G_i''(a)[v,w] = -2 \int_\alpha^\beta y_i^T(y_{i,a}v(t)m_{i,a,w}) dt, \tag{4.3}
\]

where \(r_{i,a,w} = (m_{i,a,w}, n_{i,a,w})^T\) satisfies the inhomogeneous linear Hamiltonian system

\[
\begin{align*}
m'_{i,a,w} &= -B(t)m_{i,a,w} - C(t)n_{i,a,w} \\
n'_{i,a,w} &= a(t)m_{i,a,w} + B(t)n_{i,a,w} + w y_{i,a},
\end{align*} \tag{4.4}
\]

and the boundary condition

\[
m_{i,a,w}(\alpha) = m_{i,a,w}(\beta) = 0. \tag{4.5}
\]

Proof. The first part of (i) follows directly from the definition of \(G\). To verify the second part, first note the identity

\[
G_i(a) = \int_\alpha^\beta (y_i^T a(t)y_i - z_i^T C(t) z_i) - (y_{i,a}^T a(t)y_{i,a} - z_{i,a}^T C(t) z_{i,a}) dt
\]

\[
+ 2 \int_\alpha^\beta y_{i,a}^T a(t)(y_{i,a} - y_i) - z_{i,a}^T C(t)(z_{i,a} - z_i) dt. \tag{4.6}
\]
It is enough to show that the second integral vanishes. We have

\[
\int_{\alpha}^{\beta} y_{i,a}^T a(t)(y_{i,a} - y_{i}) - z_{i,a}^T C(t)(z_{i,a} - z_{i}) \, dt
\]

\[
= \int_{\alpha}^{\beta} [y_{i,a}^T a(t) + z_{i,a}^T B(t)](y_{i,a} - y_{i})
\]

\[
- z_{i,a}^T [C(t)(z_{i,a} - z_{i}) + B(t)(y_{i,a} - y_{i})] \, dt
\]

\[
= \int_{\alpha}^{\beta} z_{i,a}^T (y_{i,a} - y_{i}) - z_{i,a}^T [C(t)(z_{i,a} - z_{i}) + B(t)(y_{i,a} - y_{i})] \, dt
\]

\[
= - \int_{\alpha}^{\beta} z_{i,a}^T [y_{i,a}' - y_{i}' + C(t)(z_{i,a} - z_{i}) + B(t)(y_{i,a} - y_{i})] \, dt
\]

\[
= 0,
\]

as required.

To prove part (ii) observe first that, for \( v \) a symmetric matrix of \( L^1[\alpha, \beta] \) functions and \( \epsilon \) a real number, and using (4.1),

\[
G_i(a + \epsilon v) - G_i(a
\]

\[
= \epsilon \int_{\alpha}^{\beta} (y_{i,a}^T v(t)y_{i} - y_{i,a + \epsilon v}^T v(t)y_{i,a + \epsilon v}) \, dt
\]

\[
+ \int_{\alpha}^{\beta} z_{i,a + \epsilon v}^T C(t)z_{i,a + \epsilon v} - z_{i,a}^T C(t)z_{i,a}
\]

\[
- y_{i,a + \epsilon v}^T a(t)y_{i,a + \epsilon v} + y_{i,a}^T a(t)y_{i,a} \, dt,
\]

\[
= \epsilon \int_{\alpha}^{\beta} (y_{i,a}^T v(t)y_{i} - y_{i,a + \epsilon v}^T v(t)y_{i,a + \epsilon v}) \, dt
\]

\[
+ \int_{\alpha}^{\beta} z_{i,a + \epsilon v}^T C(t)(z_{i,a + \epsilon v} - z_{i,a}) - y_{i,a + \epsilon v}^T a(t)(y_{i,a + \epsilon v} - y_{i})
\]

\[
+ (z_{i,a + \epsilon v} - z_{i,a})^T C(t)z_{i,a} - (y_{i,a + \epsilon v} - y_{i,a})^T a(t)y_{i,a} \, dt.
\]

(4.7)

Now

\[
\int_{\alpha}^{\beta} z_{i,a + \epsilon v}^T C(t)(z_{i,a + \epsilon v} - z_{i,a}) - y_{i,a + \epsilon v}^T a(t)(y_{i,a + \epsilon v} - y_{i}) \, dt
\]

\[
= \int_{\alpha}^{\beta} z_{i,a + \epsilon v}^T [C(t)(z_{i,a + \epsilon v} - z_{i,a}) + B(t)(y_{i,a + \epsilon v} - y_{i})]
\]

\[
- [z_{i,a + \epsilon v}^T B(t) + y_{i,a + \epsilon v}^T a(t)](y_{i,a + \epsilon v} - y_{i,a}) \, dt
\]

\[
= - \int_{\alpha}^{\beta} z_{i,a + \epsilon v}^T (y_{i,a + \epsilon v} - y_{i,a})' + z_{i,a + \epsilon v}^T (y_{i,a + \epsilon v} - y_{i,a})
\]

\[
+ c y_{i,a + \epsilon v}^T v(t)(y_{i,a + \epsilon v} - y_{i,a}) \, dt
\]

\[
= -\epsilon \int_{\alpha}^{\beta} y_{i,a + \epsilon v}^T v(t)(y_{i,a + \epsilon v} - y_{i,a}) \, dt,
\]
and

\[
\int_\alpha^\beta (z_{i,a+\epsilon v} - z_{i,a})^T C(t) z_{i,a} - (y_{i,a+\epsilon v} - y_{i,a})^T a(t) y_{i,a} dt \\
= \int_\alpha^\beta [(z_{i,a+\epsilon v} - z_{i,a})^T C(t) + (y_{i,a+\epsilon v} - y_{i,a})^T B^T(t)] z_{i,a} \\
- (y_{i,a+\epsilon v} - y_{i,a})^T [a(t) y_{i,a} + B^T(t) z_{i,a}] dt \\
= - \int_\alpha^\beta (y'_{i,a+\epsilon v} - y'_{i,a})^T z_{i,a} + (y_{i,a+\epsilon v} - y_{i,a})^T z'_{i,a} dt \\
= 0.
\]

Consequently,

\[
G_i(a + \epsilon v) - G_i(a) = \epsilon \int_\alpha^\beta y_i^T v(t) y_i - y_{i,a+\epsilon v}^T v(t)y_{i,a+\epsilon v} \\
- y_{i,a+\epsilon v}^T v(t)(y_{i,a+\epsilon v} - y_{i,a}) dt, \quad (4.8)
\]

and (ii) follows easily.

Using (ii) we have, for \(v\) and \(w\) symmetric matrices,

\[
G'_i(a + \epsilon w)[v] - G'_i(a)[v] \\
= \int_\alpha^\beta y_i^T v(t) y_i - y_{i,a+\epsilon w}^T v(t)y_{i,a+\epsilon w} dt \\
= - \int_\alpha^\beta y_{i,a+\epsilon w}^T v(t)(y_{i,a+\epsilon w} - y_{i,a}) + (y_{i,a+\epsilon w} - y_{i,a})^T v(t)y_{i,a} dt.
\]

Note that

\[
y'_{i,a+\epsilon w} - y'_{i,a} = -B(t)(y_{i,a+\epsilon w} - y_{i,a}) - C(t)(z_{i,a+\epsilon w} - z_{i,a}) \\
z'_{i,a+\epsilon w} - z'_{i,a} = a(t)(y_{i,a+\epsilon w} - y_{i,a}) + B^T(t)(z_{i,a+\epsilon w} - z_{i,a}) + \epsilon v(t)y_{i,a+\epsilon w}.
\]

If we divide these equations by \(\epsilon\) and let \(\epsilon\) tend to zero, we find that

\[
\lim_{\epsilon \to 0} \frac{y_{i,a+\epsilon w} - y_{i,a}}{\epsilon} = m_{i,a,w} \quad (4.9)
\]

and

\[
\lim_{\epsilon \to 0} \frac{z_{i,a+\epsilon w} - z_{i,a}}{\epsilon} = n_{i,a,w}, \quad (4.10)
\]

where \(r_{i,a,w} = (m_{i,a,w}, n_{i,a,w})^T\) satisfies the equations (4.4) and the boundary condition (4.5); noting the symmetry of \(v\), part (iii) now follows. \(\square\)

The functionals \(F_i\), \(G_{ij}\), and \(G\) defined in (1.23), (1.29), and (1.35) each have analogous properties, but we omit the details.
5. Proof of Theorem 1.1

We begin the proof of Theorem 1.1 by showing that the functionals $G_i$ are convex. From (4.3), as $v$ is symmetric,

\[
G_i''(a)[v, v] = -2 \int_{\alpha}^{\beta} y_{i,a}^T v(t) m_{i,a,v} dt = -2 \int_{\alpha}^{\beta} m_{i,a,v}^T y_{i,a} v(t) dt
\]

\[
= -2 \int_{\alpha}^{\beta} m_{i,a,v}^T \left[ n_{i,a,v}^t - a(t) m_{i,a,v} - B^T(t) n_{i,a,v} \right] dt
\]

\[
= 2 \int_{\alpha}^{\beta} m_{i,a,v}^T \left[ m_{i,a,v} a(t) m_{i,a,v} + B^T(t) n_{i,a,v} \right] dt
\]

\[
- [m_{i,a,v} B^T(t) + n_{i,a,v}^T C(t)] n_{i,a,v} dt
\]

\[
= 2 \int_{\alpha}^{\beta} m_{i,a,v}^T a(t) m_{i,a,v} - n_{i,a,v}^T C(t) n_{i,a,v} dt
\]

\[
= 2Q(r_{i,a,v}).
\]

The convexity of $G_i$ now follows from Theorem 2.5. From the same theorem we also have that if $G''(a)[v, v] = 0$ then $G_i''(a)[v, v] = 0$, and hence $r_{i,a,v} = 0$, for $1 \leq i \leq n$, so that from (4.4), $v y_{i,a} = 0$ for all $i$. Set $Y_{i,a} = (y_{1,a}, \ldots, y_{n,a})$. Then

\[
v Y_{i,a} = 0,
\]

(5.1)

where

\[
Y_{i,a}(\alpha) = 0, \quad Y_{i,a}(\beta) = (\eta_1, \ldots, \eta_n),
\]

(5.2)

and

\[
Y_{i,a}' = -B(t) Y_{i,a} - C(t) Z_{i,a},
\]

\[
Z_{i,a}' = a(t) Y_{i,a} + B^T(t) Z_{i,a},
\]

(5.3)

Let $(Y_0, Z_0)$ be an isotropic solution of (5.3) for which $Y_0(t)$ is invertible for all $\alpha \leq t \leq \beta$; the existence such a solution is guaranteed by Theorem 2.2. By Lemma 2.3

\[
Y_{i,a}(t) = Y_0(t)[M + S_0(t)N],
\]

where $M$ and $N$ are constant matrices and $S_0(t)$ is defined by (2.5). From (5.2) it follows that $M = 0$. From Lemma 2.4, noting that condition [C] of [1] holds as a consequence of disconjugacy, we have that $S_0(\beta)$ is invertible and, as $Y_{i,a}(\beta)$ is also invertible, it follows that $N$ is invertible. In consequence, $Y_{i,a}(t)$ is invertible for all $\alpha < t \leq \beta$ and by (5.1) we see that the matrix function $v(t)$ is zero almost everywhere on $[\alpha, \beta]$ and $G$ is strictly convex.

Theorems 1.2 and 1.3 may be proven in similar fashion; we omit the details. The proof of strict convexity in Theorem 1.4 is simpler in that the matrix $v(t)$ has the form (1.33), so only one vector $y_a$ is needed. Indeed, we need only the fact that, for the first component of $y_a$, the zeros cannot cluster, so that it is in consequence non-zero almost everywhere in $[\alpha, \beta]$. 
6. Applications

**PDE Parameter Identification.** Hamiltonian systems involving a parameter \( \lambda \) arise naturally from certain partial differential equation parameter identification problems. Consider, by way of a relatively simple example, the problem of recovering the coefficient functions \( \theta(t) > 0, p(t) > 0 \) and \( q(t) > 0 \) from a knowledge of two solutions \( w_i(t, s), i = 1, 2, \) of the equation

\[
\theta(t) \frac{\partial w}{\partial s} + \frac{\partial^4 w}{\partial t^4} - \frac{\partial}{\partial t} (p(t) \frac{\partial w}{\partial t}) + q(t) w(t, s) = 0, \quad 0 \leq s \leq 1, \quad 0 \leq t \leq 1, \quad (6.1)
\]

defined by the boundary conditions

\[
w_i(t, 0) = w_i(t, 1) = 0, \quad 0 \leq t \leq 1, \quad i = 1, 2,
\]

\[
w_i(0, s) = \frac{\partial}{\partial t} w_i(0, s) = 0, \quad 0 \leq s \leq 1, \quad i = 1, 2,
\]

\[
w_1(1, s) = 1, \quad \frac{\partial}{\partial t} w_1(1, s) = 0, \quad 0 \leq s \leq 1,
\]

\[
w_2(1, s) = 0, \quad \frac{\partial}{\partial t} w_2(1, s) = 1, \quad 0 \leq s \leq 1.
\]

Let

\[
u(t, \lambda) = \int_0^1 e^{-\lambda s} w(t, s) \, ds.
\]

For \( i = 1, 2 \) one can transform the \( w_i(t, s) \) data to data for the solutions \( u = u_i(t, \lambda) \) of the equation

\[
\frac{d^4 u}{dt^4} - \frac{d}{dt} (p(t) \frac{du}{dt}) + [-\lambda \theta(t) + q(t)]u = 0,
\]

where, if we set \( y_{i,\lambda}(t) = (u_i(t, \lambda), u_i'(t, \lambda))^T \), then

\[
y_{i,\lambda}(0) = 0, \quad i = 1, 2, \quad y_{1,\lambda}(1) = \frac{1 - e^{-\lambda}}{\lambda} (1, 0)^T, \quad y_{2,\lambda}(1) = \frac{1 - e^{-\lambda}}{\lambda} (0, 1)^T.
\]

The equation may be written in Hamiltonian form (1.26) where

\[
A_1(t) = \begin{pmatrix} q(t) & 0 \\ 0 & p(t) \end{pmatrix}, \quad A_2(t) = \begin{pmatrix} -\theta(t) & 0 \\ 0 & 0 \end{pmatrix},
\]

and

\[
B(t) = \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix}, \quad C(t) = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}.
\]

and the matrix \( C \) is non-positive definite.

In similar fashion, one can also incorporate higher order derivatives in \( s \) and \( t \) in equation (6.1) into this process.

**Numerical Differentiation.** The problem of effectively computing the derivative functions \( u^{(m)}, m \geq 1, \) given (possibly noisy) data for \( u(t), 0 \leq t \leq 1, \) is generally known as numerical differentiation. We can assume that \( u(t) > 0 \) for \( 0 \leq t \leq 1, \) and we consider the linear equation of order \( 2n \)

\[
(-1)^n \frac{d^{2n} u}{dt^{2n}}(t) + Q_n(t) u(t) = 0, \quad 0 \leq t \leq 1. \quad (6.2)
\]

We make use of Theorem 1.4. Setting \( n = 1, \) so that in (1.34)

\[
A = (Q_1), \quad B = (-1), \quad C = (-1), \quad y = (u), \quad z = (u'), \quad x = (u, u')^T,
\]
and \( u \) is the function being differentiated. Note that, by Theorem 1.4(ii), a knowledge of \( u \) alone allows the minimization to proceed. We minimize \( G(a) \) to obtain \( Q_1 \), with \( u' \) appearing as a by-product. Next, set \( n = 2 \); then
\[
A = \begin{pmatrix} Q_2 & 0 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}, \quad C = \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix},
\]
and \( y(t) = (u(t), u'(t))^T \) is now known, and \( x = (y, z)^T \), where \( z = (-u^{(3)}, u^{(2)})^T \), is a solution of the version of (1.1) obtained by using (6.2). Again \( G(a) \) is minimized to obtain \( Q_2 \), with \( u^{(2)} \) and \( u^{(3)} \) as by-products. This process may be continued for the higher derivatives.

References