An Oscillation Theorem for a Sturm–Liouville Eigenvalue Problem

By Martin Bohner of Ulm

(Received October 7, 1994)

Abstract. We consider a certain Sturm–Liouville eigenvalue problem with self-adjoint and non-separated boundary conditions. We derive an explicit formula for the oscillation number of any given eigenfunction.

1. Introduction

Let $a, b, \alpha_0, \beta_0, \alpha_1, \beta_1 \in \mathbb{R}$, $a < b$, let $q : [a, b] \times \mathbb{R} \to \mathbb{R}$ be a function, and let the conditions

\[
\begin{align*}
(A_1) & \quad q \in C([a, b] \times \mathbb{R}), \\
& \quad q(t; \lambda_1) < q(t; \lambda_2) \quad \text{for all } t \in (a, b), \quad \lambda_1 < \lambda_2, \\
& \quad \lim_{\lambda \to -\infty} q(t; \lambda) = -\infty \quad \text{for all } t \in (a, b),
\end{align*}
\]

\[
(A_2) \quad \alpha_0 \beta_1 - \alpha_1 \beta_0 = 1
\]

be satisfied (which will be denoted by $(A)$). We consider the Sturm–Liouville eigenvalue problem $(E)$, consisting of a linear differential equation of second order

\[
(E_1) \quad \ddot{x}(t) + q(t; \lambda)x(t) = 0 \quad (t \in [a, b])
\]

subject to the two (under $(A_2)$) linearly independent and self-adjoint boundary conditions

\[
(E_2) \quad \begin{cases} 
\alpha_0 \dot{x}(a) + \beta_0 x(a) = x(b), \\
\alpha_1 \dot{x}(a) + \beta_1 x(a) = -\dot{x}(b).
\end{cases}
\]
(Note that the boundary conditions are non-separated and that any two linearly independent, self-adjoint, and non-separated boundary conditions may be equivalently transformed into \((E_2)\) under \((A_2)\), as is shown in [BA].) Assuming \((A)\), there exists, as is well-known, (see e.g. [BI]) an at most countably infinite number of real and isolated eigenvalues \(\lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \cdots\) of \((E)\). This number is infinite e.g. whenever \(\lim_{\lambda \to -\infty} q(\cdot; \lambda) = \infty\) holds on some nonempty subinterval of \([a, b]\), and it is finite e.g. whenever \(\lim_{\lambda \to -\infty} q(\cdot; \lambda) \leq \kappa \in \mathbb{R}\) holds on \([a, b]\) (similar statements are true when assuming \(q \in C([a, b] \times I)\) for some real and open interval \(I\); see e.g. [IN, p. 231]). The main result of this paper is an oscillation formula ((OF) in Theorem 3.1 below) for the problem \((E)\) under \((A)\), which calculates the oscillation number \(Z_p\), i.e., the exact number of zeros of a given eigenfunction \(x_p\) corresponding to an eigenvalue \(\lambda_p\) on the interval \((a, b]\) for \(p \in \mathbb{N}\).

In 1909, G. D. Birkhoff ([BI]) already discussed this problem, but his results were not correct, as was shown by G. Baur ([BA]) in 1988. The main difference between those works and the present paper is that here we make use of an oscillation theorem for general Hamiltonian problems which has been developed by G. Baur, W. Kratz, and A. Peyermhoff in [KR-PE, BA-KR, KR]. With this new method one can see that one of the 13 cases in [BA] actually never occurs; and the remaining 12 cases are unified by our oscillation formula

\((\text{OF})\)

\[ Z_p = p - \mu_p - \nu_p, \]

where \(\mu_p, \nu_p \in \mathbb{R}\) are constants, and where \(\text{ind} x = 1\) if \(x < 0\) and 0 otherwise. Of course, the new method also yields the classical oscillation theorem of Sturm, where the boundary conditions need to be separated.

2. Notation and auxiliary results

Throughout we denote by Ker, Im, rank, def, ind, resp. the kernel, image, rank, defect (i.e., the dimension of the kernel), negative index (i.e., the number of negative eigenvalues) resp. of a matrix. Our main tool will be a general oscillation theorem for self-adjoint differential systems due to W. Kratz ([KR, Theorem 1]). For convenience we restate this theorem (Theorem 2.1 below) for our special case; its proof is slightly different but essentially the same as in [KR]. To begin with, we provide the necessary notation:

Let \(\lambda_p \in \mathbb{R}\) be an eigenvalue of \((E)\), let \(x_p\) be a corresponding eigenfunction with \(x_p(b) \neq 0\), and let \(x_p^* = -\frac{x_p}{x_p(b)}\). We use \(x(t, \lambda)\) to denote the solution of the initial value problem

\[
\begin{align*}
\ddot{x}(t) + q(t, \lambda)x(t) & = 0 \\
x(a) & = \alpha_0 \\
\dot{x}(a) & = -\beta_0
\end{align*}
\]

(Note that \(x_p^*(t) \dot{x}(t, \lambda_p) - \dot{x}_p^*(t)x(t, \lambda_p) \equiv 1\) holds on \([a, b]\).) Furthermore, we define
the following matrices

\[
\Lambda_p = \begin{pmatrix}
1 & 0 & 0 \\
-1 & 0 & -x(b, \lambda_p) \\
-x_p^*(b) & 0 & 1 + x(b, \lambda_p)
\end{pmatrix}, \quad R = \begin{pmatrix}
1 & 0 & 0 \\
0 & \alpha_0 & 0 \\
0 & \alpha_1 & 1
\end{pmatrix},
\]

and

\[
X_p = \begin{pmatrix}
0 & 0 & -1 \\
-x_p^*(a) & -x_p^*(a) & -\alpha_0 \\
0 & -1 & x(b, \lambda_p)
\end{pmatrix}.
\]

Now, the orthogonal decomposition of \(\text{Im } R^T\) into \(\text{Im } X_p\) and its orthogonal complement leads to a unique matrix \(S_p^*\) and another matrix \(S_p\) such that \(R^T = X_p S_p + S_p^*\) and \(X_p^T S_p^* = 0\) hold. Then we consider a matrix \(T_p\) with \(\text{Im } T_p = \text{Ker } S_p^*\) and finally we define \(Q_p = T_p^T \Lambda_p S_p T_p\). With this setting, Theorem 1 from [KR] reads as follows:

**Theorem 2.1.** Assume (A). Let \(\lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \cdots\) be the eigenvalues of (E) and let \(x_p\) be an eigenfunction with \(x_p(b) \neq 0\) corresponding to the eigenvalue \(\lambda_p\). Then we have

\[
z_p + \text{ind } Q_p = n_p + \text{rank } T_p - 2,
\]

where \(z_p\) denotes the number of zeros of \(x_p\) in \((a, b)\), and where \(n_p\) denotes the number of eigenvalues (counting multiplicities) of (E) that are less than \(\lambda_p\).

3. The oscillation formula

We shall now construct the matrices \(Q_p\) and \(T_p\), compute \(\text{ind } Q_p\) and \(\text{rank } T_p\), and apply Theorem 2.1 to prove the following theorem which is the main result of this paper.

**Theorem 3.1.** Assume (A). Let \(\lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \cdots\) be the eigenvalues (counting multiplicities) of (E). An eigenfunction \(x_p\) corresponding to an eigenvalue \(\lambda_p, p \in \mathbb{N}\), vanishes exactly

\[
\text{(OF)} \quad Z_p = p - \text{ind } \mu_p - \text{ind } \nu_p
\]

times on \((a, b]\), where the constants \(\mu_p\) and \(\nu_p\) are defined by

\[
\begin{align*}
\mu_p &= -\alpha_0 x_p(a) x_p(b) - (1 - \text{def } \alpha_0) \text{def } x_p(a) \\
\nu_p &= (-1)^{p-1} (\alpha_0 - \beta_0 \text{def } \alpha_0).
\end{align*}
\]

The remainder of this section is devoted to the proof of our main result.

**Lemma 3.2.** With the assumptions and notation of Theorem 2.1 we have

\[
\text{(OF*)} \quad z_p = 1 + n_p - \text{ind } \mu_p - \text{ind } x(b, \lambda_p).
\]
Proof. For convenience we use the abbreviations

\[ N_p = \text{def } x_p(a) = \begin{cases} 1 & \text{if } x_p(a) = 0, \\ 0 & \text{if } x_p(a) \neq 0, \end{cases} \]

and

\[ N_p^* = \begin{cases} 0 & \text{if } x_p(a) = 0, \\ \frac{1}{x_p(a)} & \text{if } x_p(a) \neq 0. \end{cases} \]

Using the matrices

\[
S_p = \frac{1}{1 + N_p \alpha_0^2} \begin{pmatrix} N_p^* \alpha_0 + (1 - N_p) x(b, \lambda_p) & -N_p^* \alpha_0 & -N_p^* \alpha_0 \hat{x}_p^*(b) \\ -x(b, \lambda_p) & -N_p \alpha_0^2 x(b, \lambda_p) & N_p \alpha_0^2 \hat{x}(b, \lambda_p) - 1 \\ -1 & -N_p \alpha_0^2 & -N_p \alpha_0^2 \hat{x}_p^*(b) \end{pmatrix},
\]

\[
S_p^* = \frac{1}{1 + N_p \alpha_0^2} \begin{pmatrix} N_p \alpha_0^2 & -N_p \alpha_0^2 & -N_p \alpha_0^2 \hat{x}_p^*(b) \\ -N_p \alpha_0 & N_p \alpha_0 & N_p \alpha_0 \hat{x}_p^*(b) \\ 0 & 0 & 0 \end{pmatrix},
\]

\[
T_p = \begin{pmatrix} 1 - N_p & N_p & N_p \hat{x}_p^*(b) \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},
\]

one can easily verify the relations \( X_p S_p + S_p^* = R^T \), \( X_p^T S_p^* = 0 \), and \( \text{Im } T_p = \text{Ker } S_p^* \) (use the formulae \( N_p^* x_p^*(a) = 1 - N_p \), \( N_p^* x_p^*(a) = 0 \), \( 1 + \hat{x}(b, \lambda_p) = -\hat{x}_p^*(b) x(b, \lambda_p) \), and \( \alpha_1 = x_p^*(a) + \alpha_0 \hat{x}_p^*(b) \)). Now, we have \( \text{rank } T_p = 3 - N_p \) and

\[
Q_p = \frac{1}{1 + N_p \alpha_0^2} \begin{pmatrix} N_p^* \alpha_0 + (1 - N_p) x(b, \lambda_p) & -N_p^* \alpha_0 \\ -N_p^* \alpha_0 & N_p^* \alpha_0 + N_p x(b, \lambda_p) (1 + \alpha_0^2) \\ -N_p \alpha_0 \hat{x}_p^*(b) & \hat{x}_p^*(b) \{ N_p^* \alpha_0 + N_p x(b, \lambda_p) (1 + \alpha_0^2) \} \\ 0 & 0 & \{ \hat{x}_p^*(b) \}^2 \{ N_p^* \alpha_0 + N_p x(b, \lambda_p) (1 + \alpha_0^2) \} \end{pmatrix},
\]

\[
\text{ind } Q_p = \text{ind } \left\{ \begin{pmatrix} N_p^* + N_p & N_p^* \\ 0 & 1 \end{pmatrix} Q_p \begin{pmatrix} N_p^* + N_p & 0 \\ 0 & \hat{x}_p^*(b) \end{pmatrix} \right\} = \text{ind } \begin{pmatrix} \{ N_p^* \}^2 x(b, \lambda_p) & 0 & 0 \\ 0 & N_p^* \alpha_0 + N_p x(b, \lambda_p) (1 + \alpha_0^2) & 0 \\ 0 & 0 & 0 \end{pmatrix}.
\]
Bohner, An Oscillation Theorem

\[ \text{ind} \left\{ N_p \right\} z(b, \lambda_p) + \text{ind} \left\{ N_p \alpha_0 - N_p z(b, \lambda_p) \right\} \]

\[ \text{ind} x(b, \lambda_p) + \text{ind} \mu_p - N_p. \]

An application of Theorem 2.1 yields (OF*).

It is now possible to transform (OF*) into the "nicer" formula (OF). To do so, we need to provide some facts about the function \( x(\cdot, \lambda_p) \) that occurs in (OF*).

**Lemma 3.3.** Assume (A). Let \( \lambda_1^* < \lambda_2^* < \lambda_3^* < \cdots \) denote the eigenvalues of the problem

\[
\begin{cases}
\ddot{x}(t) + q(t, \lambda)x(t) = 0 \\
\alpha_0 \dot{x}(a) + \beta_0 x(a) = 0 \\
x(b) = 0.
\end{cases}
\]

Then we have

(i) \( \lambda^* \) is an eigenvalue of \((E^*)\) iff \( x(b, \lambda^*) = 0 \);

(ii) if \( \lambda \in (\lambda_{p-1}^*, \lambda_p^*) \), then \( x(\cdot, \lambda) \) has exactly \( p - 1 \) zeros on \((a, b)\);

(iii) putting \( \lambda_0^* = -\infty \), the comparison result \( \lambda_{p-1}^* \leq \lambda_p \leq \lambda_p^* \) holds for all \( p \in \mathbb{N} \) (where \( \lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \cdots \) are the eigenvalues of problem \((E)\));

(iv) if \( x(\cdot, \lambda) \) has exactly \( q \) zeros on \((a, b)\), and if \( \nu_q = (-1)^{q-1}(\alpha_0 - \beta_0 \text{def} \alpha_0) \), then

\[ \text{sgn} x(b, \lambda) = \text{sgn} \nu_{q+1} \quad \text{if} \quad x(b, \lambda) \neq 0, \quad \text{and} \]

\[ \text{sgn} \dot{x}(b, \lambda) = \text{sgn} \nu_q \quad \text{if} \quad x(b, \lambda) = 0. \]

**Proof.** (i), (ii), and (iii) are well-known facts (see e.g. [BA, KR-PE]), and (iv) follows from the simplicity of the zeros of \( x(\cdot, \lambda) \).

**Proof of Theorem 3.1.** Put \( \lambda_0 = \infty \) and let \( p \in \mathbb{N} \).

Suppose \( x_p(b) \neq 0 \). If \( \lambda_p \in (\lambda_{p-1}^*, \lambda_p^*) \), then \( x(b, \lambda_p) \neq 0 \) and \( x(\cdot, \lambda_p) \) has exactly \( p - 1 \) zeros on \((a, b)\). Hence (note \( \lambda_{p-1} \leq \lambda_{p-1}^* < \lambda_p \) \( Z_p = 1 + (p - 1) - \text{ind} \mu_p - \text{ind} \nu_p \)). Otherwise, we have \( \lambda_p \in \{ \lambda_{p-1}^*, \lambda_p^* \} \). Now, \( x(b, \lambda_p) = 0 \) and \( \dot{x}(b, \lambda_p) = -1 \) (since \( x_p(b) \dot{x}(b, \lambda_p) - \dot{x}_p(b) x(b, \lambda_p) = 1 \)). Also, \( \lambda_p \) is a double eigenvalue since both \( x_p(\cdot) \) and \( x(\cdot, \lambda_p) \) solve \((E)\). We have that \( n_p = p - 1 = p - 1 - (1 - \text{ind} \nu_{p-1}) = p - 1 - \text{ind} \nu_p \) if \( \lambda_p = \lambda_p^* > \lambda_{p-1}^* \leq \lambda_{p-1} \) (note that \( x(\cdot, \lambda_p) \) vanishes \( p - 1 \) times on \((a, b)\)) and \( n_p = p - 2 = p - 1 - \text{ind} \nu_{p-2} = p - 1 - \text{ind} \nu_p \) if \( \lambda_{p-1} \leq \lambda_p = \lambda_{p-1}^* < \lambda_p^* \leq \lambda_{p+1} \), \( p \neq 1 \) (observe that \( x(\cdot, \lambda_p) \) has now \( p - 2 \) zeros in \((a, b)\)).

Now, suppose that \( x_p(b) = 0 \). Then \( x(\cdot, \lambda_p) \) and \( x_p(\cdot) \) are linearly dependent and we have \( \mu_p = 0 \) as well as \( \dot{x}(b, \lambda_p) = -1 \). (For the proof of the latter statement pick the solution \( y \) of \((E_1)\) where \( \lambda = \lambda_p \) with \( y(a) = -\alpha_1 \) and \( \dot{y}(a) = \beta_1 \). Now we have \( y(b) - \dot{x}(b, \lambda_p) = 2 \) (see e.g. [BL]) and \( \dot{x}(b, \lambda_p) y(b) = -1 \) which yields \( y(b) = 1 \) and \( \dot{x}(b, \lambda_p) = -1 \).) If \( \lambda_p = \lambda_p^* \), then \( x(\cdot, \lambda_p) \) has \( p - 1 \) zeros on \((a, b)\), i.e., \( Z_p = p \) and \( \text{ind} \nu_p = 1 - \text{ind} \nu_{p-1} = 0 \). If \( \lambda_p = \lambda_{p-1}^* \), \( p \neq 1 \), then \( x(\cdot, \lambda_p) \) vanishes \( p - 2 \) times on \((a, b)\), i.e., \( Z_p = p - 1 \) and \( \text{ind} \nu_p = \text{ind} \nu_{p-2} = 1 \).
This proves that our oscillation formula (OF) is valid in any case. □

**Remark 3.4.** Observe that the case $\alpha_0 < 0$, $\mu_1 < 0$ can never occur (see e.g. [BA, p.192]), since an application of (OF) would yield the impossible number of $Z_1 = 1 - 1 - 1 = -1$ zeros of an eigenfunction $x_1$ corresponding to the eigenvalue $\lambda_1$ of (E) on $(a, b]$.

**References**


*Universität Ulm*
*Abteilung Mathematik V*
*Helmholtzstr. 18*
*D-88069 Ulm*
*Germany*
*e-mail:*
*bohner@mathematik.uni-ulm.de*