The Sturm sequence property of $\{f_k(\lambda)\}$ The sequences $\{f_k(a)\}$ and $\{f_k(b)\}$ can be used to determine the number of roots of $f_n(\lambda)$ which are contained in [a,b]. To do this, introduce the following integer valued function $s(\lambda)$. Define $s(\lambda)$ to be the number of agreements in sign of consecutive members of the sequence $\{f_k(\lambda)\}$; and if the value of some member $f_j(\lambda)=0$, let its sign be chosen opposite to that of $f_{j-1}(\lambda)$. We will show later that $f_j(\lambda)=0$ implies $f_{j-1}(\lambda)\neq 0$.

Example Consider the sequence $f_0(\lambda),\ldots,f_6(\lambda)$ given in (9.4.5) of the last example. For $\lambda=3$,

$$(f_0(\lambda),\ldots,f_6(\lambda)) = (1,-1,0,-1,0,1).$$

The corresponding sequence of signs is

$$(+,-,+,+,-,+,+),$$

and $s(3)=2$.

We now state the basic result used in computing the roots of $f_n(\lambda)$ and thus the eigenvalues of T.

Theorem 9.5 Let T be a real symmetric tridiagonal matrix of order n, as given in (9.3.22). Let the sequence $\{f_k(\lambda) \mid 0\leq k\leq n\}$ be defined as in (9.4.2), and assume all $\beta_\ell\neq 0$, $\ell=1,\ldots,n-1$. Then the number of roots of $f_n(\lambda)$ which are greater than $\lambda=a$ is given by $s(a)$, which is defined in the preceding paragraph. For $a<b$, the number of roots in the interval $a<\lambda\leq b$ is given by $s(a)-s(b)$.

Proof The proof of this result is fairly lengthy; and those who wish to skip it may go directly to the example following the proof. This theorem is characteristic of all Sturm sequences, but we will consider only $\{f_k(\lambda)\}$. The proof is divided into several parts, beginning with a derivation of properties of $\{f_k(\lambda)\}$.

(1) No two consecutive polynomials have a common zero. For a proof by contradiction, suppose

$$f_j(\lambda) = f_{j-1}(\lambda) = 0$$

for some $j \geq 2$. Then using (9.4.3) with $k=j$ and solving for $f_{j-2}(\lambda)$,

$$f_{j-2}(\lambda) = \frac{1}{\beta_{j-2}^2}[(\alpha_j-\lambda)f_{j-1}(\lambda)-f_j(\lambda)]$$

$$= 0.$$

Continue the argument inductively using (9.4.3) to prove

$$f_k(\lambda) = 0, \qquad \text{all } k \leq j.$$

But this will contradict the definition $f_0(\lambda) \equiv 1$.

(2) The zeroes of $f_{k-1}(\lambda)$ interlace those of $f_k(\lambda)$, for $k=2,3,\ldots,n$. For the simplest case, consider the zeroes of $f_1(\lambda)$ amd $f_2(\lambda)$. The single zero of $f_1(\lambda)$ is $\lambda=\alpha_1$. Since $f_2(\pm\infty)=+\infty$ and $f_2(\alpha_1)=-\beta_1^2<0$, there must be zeroes of $f_2(\lambda)$ to the left and right of $\lambda=\alpha_1$.

Assume the result is true for the zeroes of $f_1(\lambda),\ldots,f_{k-1}(\lambda)$. We will prove it is true for the zeroes of $f_k(\lambda)$. Let the roots of $f_{k-1}(\lambda)$ and $f_{k-2}(\lambda)$ be $\{\lambda_1,\ldots,\lambda_{k-1}\}$ and $\{\mu_1,\ldots,\mu_{k-2}\}$, respectively. By assumption,

$$\lambda_{k-1}<\mu_{k-2}<\lambda_{k-2}<\cdots<\lambda_2<\mu_1<\lambda_1. \qquad (9.4.6)$$

Note that since degree $f_{k-1}(\lambda)=k-1$ and since there are exactly $k-1$ roots $\lambda_1,\ldots,\lambda_{k-1}$, they must all be simple roots. The same is true of the roots $\mu_1,\ldots,\mu_{k-2}$ for $f_{k-2}(\lambda)$.

Examine the sign of $f_k(\lambda)$ at $\lambda=\lambda_{j-1}$ and $\lambda=\lambda_j$. From (9.4.3),

$$f_k(\lambda_j) = (\alpha_k - \lambda_j) f_{k-1}(\lambda_j) - \beta_{k-1}^2 f_{k-2}(\lambda_j)$$

$$= -\beta_{k-1}^2 f_{k-2}(\lambda_j),$$

$$f_k(\lambda_{j-1}) = -\beta_{k-1}^2 f_{k-2}(\lambda_{j-1}). \tag{9.4.7}$$

Because $f_{k-2}(\lambda)$ has a single simple root μ_{j-1} between λ_{j-1} and λ_j, we have

$$\text{sign } f_{k-2}(\lambda_j) = -\text{sign } f_{k-2}(\lambda_{j-1}).$$

But from (9.4.7) the same must be true of $f_k(\lambda_j)$ and $f_k(\lambda_{j-1})$; and therefore, $f_k(\lambda)$ must have a root of odd multiplicity between λ_{j-1} and λ_j. This is true for each $k=2,3,\ldots,k-1$.

Since

$$f_\ell(\lambda) = (-1)^\ell \lambda^\ell + \text{lower degree terms}, \quad \text{all } \ell \geq 1,$$

we have

$$\text{sign } f_k(-\infty) = \text{sign } f_{k-2}(-\infty),$$

$$\text{sign } f_k(+\infty) = \text{sign } f_{k-2}(+\infty). \tag{9.4.8}$$

Also from (9.4.7),

$$\text{sign } f_k(\lambda_1) = -\text{sign } f_{k-2}(\lambda_1) = -\text{sign } f_{k-2}(+\infty).$$

The last equality follows since f_{k-2} has no zeroes to the right of μ_1. Using (9.4.8), $f_k(\lambda)$ must have a zero to the right of λ_1. A similar argument shows $f_k(\lambda)$ has a root to the left of λ_{k-1}. Counting roots of $f_k(\lambda)$, there are k so far. But degree $f_k(\lambda)=k$ implies these must be all the roots, and they must all be simple. This completes the proof that the zeroes of $f_{k-1}(\lambda)$ interlace those of $f_k(\lambda)$.

(3) We now prove the main conclusion of the theorem, that the number of roots of $f_n(\lambda)$ which are greater than a is equal to s(a). The proof will be an induction on n.

For the sequence with n=1,

$$\{f_0(a), f_1(a)\} = \{1, \alpha_1 - a\},$$

it is straightforward that the number of agreements in sign equal the number of roots of $f_1(\lambda)$ greater than a.

Assume the result is true for the sequence

$$f_0(a), \ldots, f_{k-1}(a),$$

and let the number of roots of $f_{k-1}(\lambda)$ which are greater than a be denoted by m. Denote the roots of $f_{k-1}(\lambda)$ by $\{\lambda_i\}$ and the roots of $f_k(\lambda)$ by $\{\nu_i\}$. Then the induction hypothesis and part (2) imply

$$\lambda_1 > \lambda_2 > \ldots > \lambda_m > a \geq \lambda_{m+1} > \ldots > \lambda_{k-1}, \tag{9.4.9}$$

$$\nu_1 > \lambda_1 > \nu_2 > \ldots > \nu_m > \lambda_m > \nu_{m+1} > \lambda_{m+1} > \ldots > \lambda_{k-1} > \nu_k. \tag{9.4.10}$$

We wish to prove that the number of roots of $f_k(\lambda)$ which are greater than a equals the number of agreements in sign in the sequence

$$\{f_0(a), \ldots, f_k(a)\}. \tag{9.4.11}$$

From (9.4.9) and (9.4.10), the number of roots of $f_k(\lambda)$ which are greater than a must be either m or m+1. The proof of our desired result breaks into three cases.

<u>case (1)</u> $a \neq \lambda_{m+1}, \nu_{m+1}$. Write

$$\begin{aligned} f_{k-1}(\lambda) &= (\lambda_1 - \lambda) \ldots (\lambda_{k-1} - \lambda), \\ f_k(\lambda) &= (\nu_1 - \lambda) \ldots (\nu_k - \lambda). \end{aligned} \tag{9.4.12}$$

There are two subcases. First, if $\lambda_{m+1} < a < \nu_{m+1}$, then the number of roots of $f_k(\lambda)$ greater than a is m+1. Also from (9.4.12),

$$\text{sign } f_{k-1}(a) = (-1)^{(k-1)-m}, \quad \text{sign } f_k(a) = (-1)^{k-(m+1)},$$

which are equal. Since these agree, the number of agreements in sign in the sequence (9.4.11) is m+1, using the original number m from the sequence

$$\{f_0(a), \ldots, f_{k-1}(a)\}.$$

This is the desired result. For the second subcase, $\nu_{m+1} < a < \lambda_m$, the number of roots of $f_k(\lambda)$ which are greater than a is just m. Also,

$$\text{sign } f_{k-1}(a) = (-1)^{(k-1)-m}, \quad \text{sign } f_k(a) = (-1)^{k-m},$$

which are different; and the desired result follows as before.

<u>case (2)</u> $a=\nu_{m+1}$. Thus $f_k(a)=0$; and by agreement, the sign of $f_k(a)$ is opposite to that of $f_{k-1}(a)$. Thus the number of agreements in sign in (9.4.11) is still m, the same as the number of zeroes of $f_k(\lambda)$ which are greater than $a=\nu_{m+1}$.

<u>case (3)</u> $a=\lambda_{m+1}$. Then $f_{k-1}(a)=0$ and there are m+1 zeroes of $f_k(\lambda)$ which are greater than a. Using (9.4.3),

$$f_k(a) = -\beta_{k-1}^2 f_{k-2}(a);$$

and by construction of the signs in (9.4.11), due to $f_{k-1}(a)=0$,

$$\text{sign } f_{k-1}(a) = -\text{sign } f_{k-2}(a).$$

Combining these gives

$$\text{sign } f_k(a) = \text{sign } f_{k-1}(a);$$

and thus the number of agreements in sign in (9.4.11) is m+1, the desired result.

This completes the proof.

■

<u>Calculation of the eigenvalues</u> The past theorem will be the basic tool in locating and separating the roots of $f_n(\lambda)$. To begin, calculate an interval which contains the roots. Using the Gerschgorin circle Theorem 9.1, all eigenvalues are contained in the interval [a,b], with

$$a = \underset{1 \le i \le n}{\text{Minimum}} \{\alpha_i - |\beta_i| - |\beta_{i-1}|\},$$

$$b = \underset{1 \le i \le n}{\text{Maximum}} \{\alpha_i + |\beta_i| + |\beta_{i-1}|\},$$

where $\beta_0 = \beta_n = 0$.

We use the bisection method on [a,b] to divide it into smaller subintervals. Theorem 9.5 is used to determine how many roots are contained in a subinterval, and we seek to obtain subintervals that will each contain one root. If some eigenvalues are nearly equal, then we continue subdividing until the root is found with sufficient accuracy. Once a subinterval is known to contain a single root, we can switch to a more rapidly convergent method.