Cambridge University Press
978-1-107-12286-4 - Theory and Synthesis of Linear Passive Time-Invariant Networks
Dante C. Youla
Excerpt
More information

1 Auxiliary results

1.1 Introduction

Much of the special necessary mathematical background for this book is developed, as needed, throughout the body of the text. Nonetheless, to avoid loss of clarity and continuity, justification for several supporting matrix arguments and points of analysis has been assigned to this opening chapter. Since the material is completely standard, what follows is in the nature of a detailed, but concise explanatory review.

As regards notation, $\bar{A}$, A', $A^*(=\bar{A}')$, A^{-1}, $\det A$, 1_n, $O_{m,n}$, O_n, and $\mathbf{0}_n$ denote, in the same order, the complex conjugate of matrix A, its transpose, adjoint, inverse and determinant, the n×n identity, the m×n and n×n zero matrices and the zero vector of dimension n. To exhibit entries it is often typographically convenient to write $\mathbf{a} = (a_1, \ldots, a_n)'$ for column vectors and $A = (a_{ij})$ for matrices. In addition, $A \dot{+} B$ is the "direct sum" of matrices A and B and $A = diag[d_1, \ldots, d_n]$ is a diagonal matrix with diagonal elements $d_1, \ldots, d_n$.

An m×n matrix A possesses m rows, each of dimension n, and n columns, each of dimension m, which span associated linear subspaces row(A) and column (A) contained in their respective linear vector spaces E_n and E_m. Let the largest size nonzero minor of A be of order r and abbreviate "dimension" by "dim." The equality

$$\dim \operatorname{row}(A) = r = \dim \operatorname{column}(A) \tag{1.1}$$

is fundamental [1]. By definition, $r \triangleq R(A)$ is the rank of A.

1.2 The rank of a matrix product

THEOREM 1.1 *Let A be m×n and B n×p. Then AB is m×p and the vectors $B\mathbf{x}$, subject to the constraint $AB\mathbf{x} = \mathbf{0}_m$, constitute a linear vector space (LVS) of dimension $R(B) - R(AB)$.*

Proof. Let $q = p - \operatorname{rank} B$ and $r = p - \operatorname{rank} AB$. In the list of r linearly independent solutions $\mathbf{x}_1, \ldots, \mathbf{x}_q, \ldots, \mathbf{x}_r$ of the equation $AB\mathbf{x} = \mathbf{0}_m$, the first q may be chosen as the set of linearly independent solutions of $B\mathbf{x} = \mathbf{0}_n$. To complete the proof it suffices to show that $B\mathbf{x}_{q+1}, \ldots, B\mathbf{x}_r$ are linearly independent. If not,

$$\mathbf{0}_n = \sum_{i=1}^{r-q} c_i B \mathbf{x}_{q+i} = B \sum_{i=1}^{r-q} c_i \mathbf{x}_{q+i} \tag{1.2}$$

for some choice of nontrivial scalars c_i. Such implies that the rightmost sum in (1.2) is a linear combination of $\mathbf{x}_1, \ldots, \mathbf{x}_q$, which is impossible because the vectors $\mathbf{x}_i$, $(i = 1 \rightarrow r)$, are linearly independent. Consequently, the vectors $B\mathbf{x}_{q+1}, \ldots, B\mathbf{x}_r$ form a linear manifold of dimension

$$r - q = R(B) - R(AB), \tag{1.3}$$

Q.E.D.

COROLLARY 1 (Sylvester's law of nullity) *If ABC exists,*

$$R(ABC) \geq R(AB) + R(BC) - R(B). \tag{1.4}$$

Proof. Denote by $\mathtt{S}_1$ the vector space of all vectors $BC\mathbf{x}$ such that $ABC\mathbf{x} = \mathbf{0}$ and by $\mathtt{S}_2$ the space of all vectors $B\mathbf{x}$ such that $AB\mathbf{x} = \mathbf{0}$. Clearly, $\mathtt{S}_1 \subset \mathtt{S}_2$ implies $\dim \mathtt{S}_1 \leq \dim \mathtt{S}_2$. However, owing to Theorem 1.1,

$$\dim \mathtt{S}_1 = R(BC) - R(ABC), \ \dim \mathtt{S}_2 = R(B) - R(AB) \tag{1.5}$$

and (1.4) is immediate, Q.E.D.[1]

COROLLARY 2 *Let A be $m \times n$ and C $n \times p$. Then*

$$R(AC) \geq R(A) + R(C) - n. \tag{1.6}$$

Proof. Choose $B = 1_n$ in (1.4), Q.E.D.

1.3 Compound matrices and Jacobi's theorem

THEOREM 1.2 (Cauchy–Binet) [1] *Let A be $m \times n$, $m \leq n$, and suppose B is $n \times m$. Then AB is square of size $m \times m$ and its determinant, det(AB), is equal to the sum[2] of all ${}^nC_m$ products which can be formed by multiplying a minor of order m from A by the corresponding minor of order m in B.*[3]

With the help of this theorem it is possible to introduce the notion of a compound matrix and to explain its significance by means of three corollaries.

Let a matrix $A^{(k)}$ be defined whose elements are all the $k \times k$ minors of an $m \times n$ matrix A.[4] *Place all such minors which come from the same group of k rows*

[1] The inequalities $R(A+B) \leq R(A) + R(B)$, $R(A-B) \geq |R(A) - R(B)|$, and $R(AB) \leq \min\{R(A), R(B)\}$ are worth remembering.

[2] The number of combination of n things taken m at a time equals ${}^nC_m = \frac{n!}{m!(n-m)!}$.

[3] A minor of order m in A built on columns numbered $i_1, \ldots, i_m$ corresponds to the minor of order m in B built on rows numbered $i_1, \ldots, i_m$.

[4] Naturally, $k \leq \min\{m, n\}$ and $A^{(k)}$ is the zero matrix for $k > R(A)$.

(or columns) of A in the same row (or column) of $A^{(k)}$. Let the position of these elements in either the rows or columns of $A^{(k)}$ be decided by dictionary (or lexical) ordering. For example, minors chosen from rows 1, 2, 4 of A will appear in an earlier row than those from 1, 2, 5 or 1, 3, 4 or 2, 3, 4; and similarly for columns. When so constructed, $A^{(k)}$, the kth *compound of A, possesses ${}^{m}C_k$ rows and ${}^{n}C_k$ columns. Note the following:*

1. *The* kth *compound of an identity is an identity. The* kth *compound of a diagonal matrix D is a diagonal matrix with the k-ary products of the diagonal element d_{ii} arranged in lexical order as diagonal elements of $D^{(k)}$.*
2. *The* kth *compound of $\overline{A}$ is $\overline{(A^{(k)})}$.*
3. *The* kth *compound of A' is $(A^{(k)})'$.*
4. *The* kth *compound of A^* is $(A^{(k)})^*$.*
5. *The* kth *compound of A^{-1} is $(A^{(k)})^{-1}$.*

Hence A symmetric, skew-symmetric, hermitian or skew-hermitian entail, respectively, $A^{(k)}$ symmetric, skew-symmetric, hermitian or skew-hermitian.

COROLLARY 1 *Let $C = AB$. Then*

$$C^{(k)} = (AB)^{(k)} = A^{(k)}B^{(k)}. \tag{1.7}$$

Proof. Assume A to be m×n, B n×p and consider the three compounds $A^{(k)}$, $B^{(k)}$, and $(AB)^{(k)}$. According to the rules of matrix multiplication, any k×k submatrix C_1 of $C = AB$ is a product of a certain k×n submatrix A_1 of A and a certain n×k submatrix B_1 of B. In fact, if C_1 is built on rows $i_1, \ldots, i_k$ and columns $j_1, \ldots, j_k$ of C, A_1 uses rows $i_1, \ldots, i_k$ of A and B_1 uses columns $j_1, \ldots, j_k$ of B. Owing to Cauchy–Binet, $\det C_1$, which is an element of $C^{(k)}$, is a sum of all products of a minor of order k drawn from A_1 and a corresponding minor of order k drawn from B_1. Thus $\det C_1$ is an element of $A^{(k)}B^{(k)}$ obtained by ordinary row-into-column multiplication. Moreover, lexical ordering of $A^{(k)}$ and $B^{(k)}$ assures lexical ordering of C, so that

$$(AB)^{(k)} = A^{(k)}B^{(k)}, \tag{1.8}$$

Q.E.D.

As a consequence, for square A, $AA^* = A^*A$ and $AA^* = 1$ in turn imply $A^{(k)}(A^{(k)})^* = (A^{(k)})^*A^{(k)}$ and $(A^{(k)})^*A^{(k)} = 1$. Apparently, compounding preserves both normality and unitarity, two properties that play an important role in Chapter 7.

Let A be n×n and denote its cofactor matrix by A_c and its adjugate A_c' by adj A. The familiar identity

$$A \cdot \operatorname{adj} A = (\det A)1_n \tag{1.9}$$

is basic.

The correct generalization of adj A to compounds is suggested by the Laplace expansion of a determinant in terms of any set of its k rows (or columns) [1]: replace every element of $A^{(k)}$ by its algebraic (signed) complementary minor in A. Let $\operatorname{adj}^{(k)}A$ denote the transpose of this new equi-size matrix.[5] Then [2]

$$A^{(k)} \cdot \operatorname{adj}^{(k)} A = (\det A) 1_{\mu}, \tag{1.10}$$

where $\mu = {}^{n}C_{k}$.

COROLLARY 2 (Jacobi's theorem) *Any minor of* adj A *of order k is equal to the complementary signed minor in A' multiplied by* $(\det A)^{k-1}$.

Proof. From (1.9) and (1.10),

$$A^{(k)} (\operatorname{adj} A)^{(k)} = (\det A)^{k} 1_{\mu} = (\det A)^{k-1} A^{(k)} \operatorname{adj}^{(k)} A. \tag{1.11}$$

Consequently, if $\det A \neq 0$, then $A^{(k)}$ is nonsingular and its cancellation in (1.11) gives

$$(\operatorname{adj} A)^{(k)} = (\det A)^{k-1} \operatorname{adj}^{(k)} A. \tag{1.12}$$

By continuity, (1.12) is valid also when A is singular, Q.E.D.

COROLLARY 3 *Any minor of A^{-1} of order k equals its complementary signed minor in A' multiplied by* $(\det A)^{-1}$.

Proof. A comparison of the identities $A^{(k)} (A^{-1})^{(k)} = 1$ and (1.10) gives

$$(A^{-1})^{(k)} = (\det A)^{-1} \operatorname{adj}^{(k)} A, \tag{1.13}$$

Q.E.D.

1.4 Singular value decomposition

Let the hermitian n×n matrix H possess the eigenvalues $\lambda_1, \ldots, \lambda_n$. As is well known [1], all λ_i are real and there exists an n×n unitary matrix U, such that

$$H = U \Sigma U^{*}, \ \Sigma = \operatorname{diag}[\lambda_1, \ldots, \lambda_n]. \tag{1.14}$$

Furthermore, if H is nonnegative-definite (n.n.d.), i.e., if $H \geq O_n$, all λ_i are ≥ 0. If also nonsingular, all eigenvalues are >0 and H is positive-definite (p.d.), i.e., $H > O_n$.

THEOREM 1.3 [3] *Let A be $m \times n$ of rank r. There exist $m \times m$ and $n \times n$ unitaries V and U, such that*

$$A = V \left[\begin{array}{c|c} \Sigma & O \\ \hline O & O \end{array} \right] U^{*}, \tag{1.15}$$

where $\Sigma = \operatorname{diag}[\sigma_1, \ldots, \sigma_r]$ and $\sigma_1 \geq \cdots \geq \sigma_r > 0$. The numbers $\sigma_1, \ldots, \sigma_r$ are the nonnegative square roots of the nonzero eigenvalues of $A^{*}A$ arranged in descending order. They constitute the positive singular values of A. Equation (1.15) is the singular value decomposition (SVD) of A.

[5] Remember that a minor of A of order k equals the determinant of a k×k submatrix of A.

COROLLARY *Define*

$$A^+ = U \left[\begin{array}{c|c} \Sigma^{-1} & O \\ \hline O & O \end{array} \right] V^* \tag{1.16}$$

to be the pseudo-inverse of A. Although many vectors $\mathbf{x}$ *minimizing*

$$\rho(\mathbf{x}) = (\mathbf{b} - A\mathbf{x})^*(\mathbf{b} - A\mathbf{x}) \tag{1.17}$$

may exist, the one for which $\|x\| = (\mathbf{x}^*\mathbf{x})^{1/2}$ *is a minimum is unique and given by* $\mathbf{x} = A^+\mathbf{b}$. *In particular, if the equation* $A\mathbf{x} = \mathbf{b}$ *is consistent,* $\mathbf{x} = A^+\mathbf{b}$ *is the solution of smallest norm. Lastly, only* A^+*, as defined in (1.15), satisfies the Penrose requirements*

$$1.\ A^+AA^+ = A^+, \tag{1.18}$$

$$2.\ AA^+A = A, \tag{1.19}$$

$$3.\ (AA^+)^* = AA^+, \tag{1.20}$$

$$4.\ (A^+A)^* = A^+A. \tag{1.21}$$

1.5 Joint diagonalization of hermitian matrices

Let A and B be a pair of n×n hermitian matrices. As is well known and easy to show [1], if A is p.d. there exists an n×n nonsingular matrix T, such that $T^*AT = 1_n$ and $T^*BT = D$, D diagonal. Specifically, since

$$D = 1_n \cdot D = (T^*AT)^{-1}(T^*BT) = T^{-1}(A^{-1}B)T, \tag{1.22}$$

the diagonal elements of D coincide with the eigenvalues of $A^{-1}B$. Although this possibility for joint diagonalization is lost when $n > 1$ and A is only n.n.d., it is still true also if B is $\geq O_n$.

Proof[6]. Obviously, $C = A + B$ is either $> O_n$ or $\geq O_n$. In the former case, C and A, and therefore C, A, and B are simultaneously reducible to diagonal form. However, if C is singular, there exists $\mathbf{x}_1 \neq \mathbf{0}_n$ such that $\mathbf{x}_1^*\mathbf{x}_1 = 1$ and $(A + B)\mathbf{x}_1 = \mathbf{0}_n$. From $A \geq O_n$ and $B \geq O_n$ follows $A\mathbf{x}_1 = B\mathbf{x}_1 = \mathbf{0}_n$.

Incorporate $\mathbf{x}_1$ into the first column of an n×n unitary matrix X_1 and observe that

$$X_1^*AX_1 = (O_1 \dot{+} A_1),\ X_1^*BX_1 = (O_1 \dot{+} B_1). \tag{1.23}$$

Clearly, A_1 and B_1 are hermitian p.d. or n.n.d matrices of order $n - 1$. By the induction hypothesis, $D_1 = X_2^*A_1X_2$ and $D_2 = X_2^*B_1X_2$ are diagonal for some choice of $(n-1)\times(n-1)$ nonsingular matrix X_2. Consequently, $T = X_1(1_1 \dot{+} X_2)$ is nonsingular and diagonalizes both A and B, Q.E.D.

[6] Proceeds by induction on n and differs significantly from the one given by Newcomb in [4].

1.6 Sylvester's law of inertia

THEOREM 1.4 *Let*

$$H = \mathrm{diag}[h_1, \ldots, h_r, 0, \ldots, 0],\ G = \mathrm{diag}[g_1, \ldots, g_r, 0, \ldots, 0] \tag{1.24}$$

be two real $n \times n$ diagonal matrices of rank r, and suppose that $G = T^*HT$ for some choice of $n \times n$ nonsingular matrix T. Then the number of positive h equals the number of positive g.

Proof. Choose an indeterminate n-vector $\mathbf{x} = (x_1, \ldots, x_n)'$ and let $\mathbf{y} = T\mathbf{x}$. Clearly, $\mathbf{x}^*G\mathbf{x} = (T\mathbf{x})^*H(T\mathbf{x}) = \mathbf{y}^*H\mathbf{y}$ may be rewritten as

$$\sum_{i=1}^{r} g_i|x_i|^2 = \sum_{i=1}^{r} h_i|y_i|^2. \tag{1.25}$$

Now assume $g_1 > 0, \ldots, g_k > 0$, $g_{k+1} < 0, \ldots, g_r < 0$, $h_1 > 0, \ldots, h_m > 0$, $h_{m+1} < 0, \ldots, h_r < 0$, $m < k$. Introduce the partition

$$T = \begin{array}{c} \begin{array}{cc} \quad k & \quad n-k \end{array} \\ \left[\begin{array}{c|c} T_{11} & T_{12} \\ \hline T_{21} & T_{22} \end{array}\right] \end{array} \begin{array}{l} \\ m \\ n-m \end{array}. \tag{1.26}$$

Since $m < k$, the linear homogeneous equation $T_{11}\mathbf{x}_1 = \mathbf{0}_m$ admits a nontrivial solution $\mathbf{x}_1 = (x_1, \ldots, x_k)'$. Thus

$$T\mathbf{x} = T\begin{bmatrix} \mathbf{x}_1 \\ \mathbf{0}_{n-k} \end{bmatrix} = \begin{bmatrix} \mathbf{0}_m \\ \mathbf{y}_2 \end{bmatrix}, \tag{1.27}$$

where $\mathbf{y}_2 = (y_{m+1}, \ldots, y_n)'$. With $\mathbf{x}$ so chosen, (1.25) reduces to

$$\sum_{i=1}^{k} g_i|x_i|^2 = \sum_{i=m+1}^{r} h_i|y_i|^2 \leq 0, \tag{1.28}$$

an impossibility because all g_i are > 0 and $\mathbf{x}_1 \neq \mathbf{0}_k$. Accordingly, $m \geq k$ and by symmetry $k \geq m$, so that $k = m$, Q.E.D.

COROLLARY *Let $D = L^*AL$, A $n \times n$ hermitian, L nonsingular and D diagonal. The number of positive and negative diagonal elements in D equal, respectively, the number of positive and negative eigenvalues of A.*

Proof. Write $U^*AU = \Sigma = \mathrm{diag}[\lambda_1, \ldots, \lambda_n]$, where U is unitary and the λ_i are the eigenvalues of A. Since $D = T^*\Sigma T$ and $T = U^*L$ is nonsingular, Theorem 1.4 is applicable, Q.E.D.

1.7 Row and column-reduced polynomial matrices

All elements, and therefore all minors of an m×n polynomial matrix $A(s)$ are polynomials in the variable s. Its normal rank, written $NR(A)$, equals the largest order of a non-identically singular minor. $A(s)$ also possesses m row degrees $\mu_1, \ldots, \mu_m$ and n column degrees $\nu_1, \ldots, \nu_n$. By definition, μ_i is the maximum degree of an element in the ith row, while ν_j is the maximum degree of an element in the jth column.

Let $\Lambda_a(s) = \text{diag}[s^{\mu_1}, \ldots, s^{\mu_m}]$ and $\Lambda_b(s) = \text{diag}[s^{\nu_1}, \ldots, s^{\nu_n}]$. Evidently,

$$C_a = \lim_{s\to\infty} \Lambda_a^{-1}(s)A(s), \; C_b = \lim_{s\to\infty} A(s)\Lambda_b^{-1}(s) \tag{1.29}$$

exist as finite m×n matrices. We say that $A(s)$ is row-reduced if $R(C_a) = m$ and column-reduced if $R(C_b) = n$. Observe that if $A(s)$ is reduced (either row or column), the condition $r = NR(A) = \min\{m, n\}$ is necessary. In general, however, it is not sufficient and our next theorem is decisive.

THEOREM 1.5 [5] *Let the m×n polynomial matrix $A(s)$ satisfy the requirement $NR(A) = r = \min\{m, n\}$. Then*[7]

1. *if $r = m$ there exists an elementary m×m polynomial matrix $E_a(s)$, such that $E_a(s)A(s)$ is row-reduced;*
2. *if $r = n$ there exists an elementary n×n polynomial matrix $E_b(s)$, such that $A(s)E_b(s)$ is column-reduced;*
3. *the row degrees of $E_a(s)A(s)$ and column degrees of $A(s)E_b(s)$ are unique up to a permutation.*

COROLLARY 1[8] *Suppose $A(s)$ is n×n and $\det A(s) \not\equiv 0$. There exists an n×n elementary polynomial matrix $E(s)$ and nonnegative integers $\nu_1, \ldots, \nu_n$, such that*

$$B(s) = A(s)E(s)\text{diag}[s^{-\nu_1}, \ldots, s^{-\nu_n}] \tag{1.30}$$

and its inverse $B^{-1}(s)$ are finite at $s = \infty$.

Proof. Clearly, $r = NR(A) = n$ and the matrix $A(s)E(s)$ can be made column-reduced by an appropriate choice of n×n elementary polynomial matrix $E(s)$. Consequently, if the ν_i are the associated column degrees of $A(s)E(s)$, $C = \text{limit}\, B(s)$ as $s \to \infty$ is nonsingular. But then $C^{-1} = \text{limit}\, B^{-1}(s)$ as $s \to \infty$ is also finite, Q.E.D.

COROLLARY 2 *The degree of the determinant of an n×n column(row)-reduced polynomial matrix $A(s)$ equals the sum of its n column(row) degrees.*

Proof. If, say, $A(s)$ is column-reduced and ν_i is the degree of its ith column, then by definition the rational matrix

$$B(s) = A(s)\text{diag}[s^{-\nu_1}, \ldots, s^{-\nu_n}] \tag{1.31}$$

is nonsingular at $s = \infty$. Since $\det B(s) = \det A(s)/s^{\nu}$ where $\nu = \nu_1 + \cdots + \nu_n$, then $\det B(\infty) \neq 0$ iff degree $\det A(s) = \nu$, Q.E.D.

[7] A polynomial matrix is elementary [1] if it is square and its determinant is a nonzero constant.
[8] The common and most direct application of Theorem 1.5.

1.8 Some key results in the theory of equations

THEOREM 1.6 *Let the real function $f(s)$ be continuous in the closed interval $[a, b] \triangleq \{s : a \leq s \leq b\}$ and suppose $f(a)f(b) < 0$. There exists ξ, $a < \xi < b$, such that $f(\xi) = 0$.*

Proof[9]. To be definite, assume $f(a) > 0, f(b) < 0$ and let ξ denote the l.u.b of all numbers s in $[a, b]$ for which $f(s) > 0$. Owing to the continuity of $f(s), a < \xi < b$. Moreover, this same continuity implies ξ too small if $f(\xi) > 0$ and ξ too large if $f(\xi) < 0$. Hence $f(\xi) = 0$, Q.E.D.

COROLLARY *A real function $f(s)$ continuous in the closed interval $[a, b]$ assumes every value μ between $f(a)$ and $f(b)$.*

Proof. Trivial if $f(a) = f(b)$. Otherwise, let

$$g(s) \triangleq \begin{cases} \mu - f(s), & f(a) < \mu < f(b), \\ f(s) - \mu, & f(a) > \mu > f(b). \end{cases} \tag{1.32}$$

The function $g(s)$ is continuous in $[a, b]$, positive for $s = a$ and negative for $s = b$. Thus, for some choice of ξ in $a < \xi < b$, $g(\xi) = 0 \Rightarrow f(\xi) = \mu$, Q.E.D.

THEOREM 1.7 *A real function $f(s)$ assumes both its inf m, its sup M and every value in between in any closed interval $[a, b]$ in which it is continuous.*

Proof. Firstly, let μ denote either m or M. By definition, there exists a sequence $\{s_r\}$ contained in $[a, b]$, such that

$$\lim_{r \to \infty} f(s_r) = \mu. \tag{1.33}$$

Extract from this sequence a subsequence $\{s_{r'}\}$ that converges to some $\xi \in [a, b]$. Since $f(s)$ is continuous,

$$\mu = \lim_{r' \to \infty} f(s_{r'}) = f(\xi). \tag{1.34}$$

Secondly, if $m = f(\xi_1)$, $M = f(\xi_2)$, and $m < M$, we complete the proof by an obvious application of the corollary to Theorem 1.6, Q.E.D.

COROLLARY 1 (Rolle's Theorem) *Let the real function $f(s)$ be continuous in the closed interval $[a, b]$ and differentiable in the open interval $(a, b) \triangleq \{s : a < s < b\}$. If $f(a) = f(b)$ there exists ξ, $a < \xi < b$, such that*[10] *$f'(\xi) = 0$.*

Proof. If the inf m and sup M of $f(s)$ are equal, $f(s)$ is constant and the result is trivial. Otherwise, $m < M$ and one of the two, say m, is assumed at some point ξ in (a, b).

[9] The reader is expected to know that every bounded set of real numbers possesses a least upper bound (l.u.b. or supremum abbreviated sup) and a greatest lower bound (g.l.b. or infimum abbreviated inf)[6].

[10] In other places, to avoid possible confusion, we sometimes write $f^{(1)}(s), f^{(2)}(s), \ldots,$ instead of $f'(s), f''(s), \ldots,$ to indicate successive derivatives of $f(s)$ of orders one, two, etc.

Clearly, if $f'(\xi) \neq 0$, there exist s in (a, b) for which $f(s) < m$, a contradiction. Hence $f'(\xi) = 0$ and $a < \xi < b$, Q.E.D.

COROLLARY 2 (The Mean Value Theorem) *Let the real function $f(s)$ be continuous in $[a, b]$ and differentiable in (a, b). There exists ξ, $a < \xi < b$, such that*

$$f(b) - f(a) = (b - a)f'(\xi). \tag{1.35}$$

Proof. Consider the function

$$g(s) = f(b) - f(s) - (b - s)P, \tag{1.36}$$

where the constant

$$P = \frac{f(b) - f(a)}{b - a} \tag{1.37}$$

is chosen to guarantee $g(a) = 0$. Clearly, $g(s)$ is continuous in $[a, b]$, differentiable in (a, b) and satisfies $g(a) = g(b) = 0$. According to Rolle's theorem there exists ξ, $a < \xi < b$, such that

$$0 = g'(\xi) = P - f'(\xi) \Rightarrow P = f'(\xi), \tag{1.38}$$

and (1.35) follows, Q.E.D.

THEOREM 1.8 (Darboux)[7] *Let the real function $f(s)$ be continuous and also possess a derivative $f'(s)$ in the closed interval $[a, b]$. Suppose $f'(a) = A$ and $f'(b) = B$. Then all values strictly between A and B are assumed by $f'(s)$ for $s \in (a, b)$.*[11]

Proof. No generality is lost by imposing the constraint $A < B$. Given C, $A < C < B$, it remains to show that there exists $\xi \in (a, b)$, such that $f'(\xi) = C$. Let $g(s) = f(s) - Cs$ and note that

$$g'(s) = f'(s) - C, g'(a) = A - C < 0, g'(b) = B - C > 0. \tag{1.39}$$

Thus $g(s)$ is a continuous differentiable function in $[a, b]$ whose derivative is negative for $s = a$ and positive for $s = b$. Consequently, there exists a point s in (a, b) to the right of a for which $g(s) < g(a)$ and a point s in (a, b) to the left of b for which $g(s) < g(b)$. This means that inf $g(s)$ in $[a, b]$ is actually attained at some point $\xi \in (a, b)$. Accordingly, because ξ is interior to $[a, b]$,

$$g'(\xi) = 0 = f'(\xi) - C \Rightarrow f'(\xi) = C, \tag{1.40}$$

Q.E.D.[12]

[11] The point of Darboux's theorem is that $f'(s)$ need not be continuous! (The reader is reminded that existence of $f'(s)$ at a point $\xi \in (a, b)$ implies that the left- and right-side derivatives of $f(s)$ at $s = \xi$ exist and are equal.)

[12] As a corollary, $f'(s) \neq 0$ for $s \in [a, b]$ iff $f'(s) > 0$ for all s in $[a, b]$ or $f'(s) < 0$ for all s in $[a, b]$. In other words, even when discontinuous, $f'(s)$ can change sign only by passing through a zero.

LEMMA 1.1 *Let $f(s)$ be a real polynomial, i.e., one with only real coefficients. Suppose that $a < b$ and $f(a)f(b) \neq 0$. Then*

$$\operatorname{sign} f(b) = (-1)^r \operatorname{sign} f(a), \tag{1.41}$$

where r denotes the number of zeros of $f(s)$ in (a, b).

Proof. Let $s_1 \leq s_2 \leq \cdots \leq s_r$ denote the zeros of $f(s)$ in (a, b). Clearly, $a < s_1$, $s_r < b$ and

$$f(s) = g(s) \prod_{i=1}^{r} (s - s_i), \tag{1.42}$$

where the real polynomial $g(s)$ is free of zeros in $[a, b]$. As such (Theorem 1.6), sign $g(s)$ is constant in $[a, b]$. Hence, in veiw of (1.42), sign $f(b)$ = sign $g(b)$ = sign $g(a)$ = $(-1)^r \operatorname{sign} f(a)$ and (1.41) follows, Q.E.D.

THEOREM 1.9[13] *Strictly between two successive zeros of a real nontrivial polynomial $f(s)$ lie an* **odd** *number of zeros of its derivative $f'(s)$.*

Proof. Assume $a < b$, let l and m denote the respective multiplicities of a and b as zeros of $f(s)$ and write

$$f(s) = (s - a)^l (s - b)^m g(s). \tag{1.43}$$

Of course, a and b are not zeros of the real polynomial $g(s)$.

Differentiation of (1.43) yields

$$\frac{f'(s)}{f(s)} = \frac{l}{s-a} + \frac{m}{s-b} + \frac{g'(s)}{g(s)}. \tag{1.44}$$

Since l and m are both ≥ 1, examination of (1.44) reveals that for sufficiently small positive $\epsilon < |b - a|$,

$$\frac{f'(a+\epsilon)}{f(a+\epsilon)} > 0, \ \frac{f'(b-\epsilon)}{f(b-\epsilon)} < 0, \tag{1.45}$$

and $(b - \epsilon) - (a + \epsilon) = b - a - 2\epsilon > 0$. In particular, (1.45) implies that $f'(s)$ is free of zeros in the two half-open intervals $a < s \leq a + \epsilon$ and $b - \epsilon \leq s < b$.

By hypothesis, a and b are successive zeros of $f(s)$. Consequently, owing to Lemma 1.1, *sign* $f(a + \epsilon)$ = *sign* $f(b - \epsilon)$ which is consistent with (1.45) iff

$$f'(a+\epsilon) f'(b-\epsilon) < 0. \tag{1.46}$$

With the aid of Lemma 1.1 we now conclude that $f'(s)$ possesses an odd number of zeros in $a + \epsilon < s < b - \epsilon$, Q.E.D.

[13] A strengthening of Rolle's theorem.