AN ABSTRACT OF THE THESIS OF

DAVID WARREN AKERS for the MASTER OF SCIENCE
(Name) (Degree)

in MATHEMATICS presented on April 24, 1969
(Major) (Date)

Title: FIRST ORDER DIFFERENTIAL CORRECTIONS IN THE EIGENVALUE PROBLEM

Abstract approved: Redacted for Privacy
Harry Goheen

Let A be an $n \times n$ real, symmetric matrix with distinct characteristic values $\lambda_1, \lambda_2, \ldots, \lambda_n$. Then there exists an orthogonal matrix P such that $PAP^T = \Lambda = (\lambda_i)$. Given a small symmetric change, ΔA, in the matrix A, we can calculate the resulting changes, ΔP, and $\Delta\Lambda$, in P and Λ respectively. We next assume that the change in A is dependent on time t. In particular, we assume that $A(t)$ is a differentiable function of t. Then, if $A_0 = A(0)$ has distinct characteristic values, we show that for sufficiently small t, P and Λ are differentiable functions of t, and that $A(t)$ also has distinct characteristic values.

First Order Differential Corrections in the Eigenvalue Problem

by

David Warren Akers

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Master of Science

June 1969

APPROVED:

Redacted for Privacy

Professor of Mathematics

in charge of major

Redacted for Privacy

Chairman of Department of Mathematics

Redacted for Privacy

Dean of Graduate School

Date thesis is presented April 24, 1969

Typed by Clover Redfern for David Warren Akers

TABLE OF CONTENTS

FIRST ORDER DIFFERENTIAL CORRECTIONS IN THE EIGENVALUE PROBLEM

I. INTRODUCTION

From an historical standpoint, the theory of matrices and their characteristic values has been of significant interest and importance to many areas in the physical sciences. These characteristic values, or eigenvalues as they are also called, of a square $n \times n$ matrix A, are determined by solving the characteristic equation $\det(A - \lambda I) = 0$. This equation is a polynomial of degree n, and by the Fundamental Theorem of Algebra has exactly n (not necessarily distinct) roots, each of which is called a characteristic value of A. If λ is a characteristic value of A, then X is said to be a characteristic vector belonging to λ if $XA = \lambda X$.

As we have said, in many fields the characteristic values and their corresponding vectors of a given matrix A may have a special physical significance. For example, consider the theory of molecular vibrations (2). A certain symmetric matrix arises in connection with a particular kind of motion of a given molecule. By calculating the characteristic values and vectors of that matrix, we determine the frequency, amplitude, and phase of the vibration of each atom of the molecule about its equilibrium position. Now suppose that an atom of that molecule is replaced by an isotopic atom of the same element. The result is a change in the frequencies of vibration because the original matrix, and consequently its characteristic values, have been changed due to the change in mass of the atom.

We can consider this problem entirely out of its physical context and in a purely mathematical one. The problem can be stated as follows. Suppose we have a given symmetric matrix A with distinct characteristic values, and suppose we change A by a small symmetric amount ΔA. What will be the change in the characteristic values and in the diagonalizing matrix P, that is, the orthogonal matrix P such that $PAP^T = \Lambda$, where Λ is the diagonal matrix of characteristic values of A?

A reasonable extension of the problem is to next assume that the change in A is dependent on time t. If, in fact, $A(t)$ is a differentiable function of t, and $A_0 = A(0)$ has distinct characteristic values, we are able to show that for sufficiently small t, both P and Λ are differentiable functions of t, and $A(t)$ also has distinct characteristic values.

It is very important that we assume that A_0 has distinct characteristic values, which, in the physical context, means that we do not have any degenerate frequencies. If we do not make this assumption, the procedures and calculations used here will clearly not hold. For a treatment of the degeneracy problem in the setting of quantum mechanics, see Kemble (1), who refers to a method first used by Van Vleck.

II. CHANGES IN THE CHARACTERISTIC VALUES AND DIAGONALIZING MATRIX

Let A be a real, symmetric, $n \times n$ matrix with n distinct characteristic values $\lambda_1, \lambda_2, \ldots, \lambda_n$. Then by the Principal Axes Theorem for real, symmetric matrices, we know that A is simultaneously similar to and congruent to the diagonal matrix $\Lambda = (\lambda_i)$. Thus there exists an orthogonal matrix P such that $PAP^{-1} = PAP^T = \Lambda$. Since P is orthogonal, $PP^T = P^TP = I$.

Suppose now that we change the matrix A by some small amount ΔA such that $A + \Delta A$ is still symmetric. Then Λ will be changed by some small amount, $\Delta\Lambda$, and $\Lambda + \Delta\Lambda$ will still be diagonal. The diagonalizing matrix P will then be changed and $P + \Delta P$ will also be orthogonal. Thus we have the following two equations.

$$(2.1) \qquad (P+\Delta P)(A+\Delta A)(P+\Delta P)^T = \Lambda + \Delta\Lambda$$

$$(2.2) \qquad (P+\Delta P)(P+\Delta P)^T = (P+\Delta P)^T(P+\Delta P) = I.$$

We wish to determine ΔP and $\Delta\Lambda$ given a certain small change, ΔA, in the matrix A. For notational convenience in the following development, we will denote $(\Delta P)^T$ by ΔP^T.

Carrying out the multiplication on the left hand side of Equation (2.1) we obtain

$$(2.3) \qquad (P+\Delta P)(A+\Delta A)(P+\Delta P)^T = (P+\Delta P)(A+\Delta A)(P^T+\Delta P^T)$$

$$= PAP^T + \Delta PAP^T + P\Delta AP^T + \Delta P\Delta AP^T$$

$$+ PA\Delta P^T + \Delta PA\Delta P^T + P\Delta A\Delta P^T + \Delta P\Delta A\Delta P^T$$

$$= \Lambda + \Delta\Lambda.$$

Now, if ΔA is sufficiently small, then so is ΔP. Thus any term in the above expression containing at least two "Δ" terms can be neglected. (This fact is readily seen when considering A, P, and Λ as functions of time and the proof will be left until then). Since $PAP^T = \Lambda$ we now have

$$(2.4) \qquad \Delta PAP^T + P\Delta AP^T + PA\Delta P^T = \Delta\Lambda.$$

Similarly, from Equation (2.2) we obtain

$$(2.5) \qquad \Delta PP^T + P\Delta P^T = 0$$

and this implies

$$(2.6) \qquad \Delta P = -P\Delta P^T P.$$

Now $PAP^T = \Lambda$ implies that $PA = \Lambda P$. Substituting this expression and Equation (2.6) into Equation (2.4) we have

$$(-P\Delta P^T P)AP^T + P\Delta AP^T + \Lambda P\Delta P^T = \Delta\Lambda.$$

Thus

$$(-P\Delta P^T)(PAP^T) + P\Delta AP^T + \Lambda P\Delta P^T = \Delta\Lambda.$$

and so

$$(2.7) \qquad (-P\Delta P^T)\Lambda + \Lambda\,(P\Delta P^T) = \Delta\Lambda - P\Delta AP^T.$$

Recalling that we wish to solve for ΔP and $\Delta\Lambda$ given ΔA, consider the (i, j)th element of each side of Equation (2.7). Since (2.7) is an equation of two matrices, these (i, j)th elements must be equal. First, the (i, j)th element of $\Lambda(P\Delta P^T)$ is $\lambda_i(P\Delta P^T)_{ij}$ where $(P\Delta P^T)_{ij}$ is the (i, j)th element of $P\Delta P^T$. Also, the (i, j)th element of $(-P\Delta P^T)\Lambda$ is $(-P\Delta P^T)_{ij}\lambda_j = -\lambda_j(P\Delta P^T)_{ij}$. Hence the (i, j)th element of the left hand side of Equation (2.7) is $(\lambda_i-\lambda_j)(P\Delta P^T)_{ij}$. Next, the (i, j)th element of $\Delta\Lambda$ is $\Delta\lambda_i\delta_{ij}$, where δ_{ij} is the Kronecker delta; that is, $\delta_{ij} = 0$ if $i \neq j$ and $\delta_{ij} = 1$ if $i = j$. Thus the (i, j)th element of the right hand side of (2.7) is $\Delta\lambda_i\delta_{ij} - (P\Delta AP^T)_{ij}$. Since these (i, j)th elements are equal, it follows that

$$(2.8) \qquad (\lambda_i-\lambda_j)(P\Delta P^T)_{ij} = \Delta\lambda_i\delta_{ij} - (P\Delta AP^T)_{ij}\,.$$

Suppose now that $i = j$. Then the left side of Equation (2.8) is zero, $\delta_{ii} = 1$, and we have

$$(2.9) \qquad \Delta\lambda_i = (P\Delta AP^T)_{ii} \quad \text{for} \quad i = 1, 2, \ldots, n.$$

Since $\Delta\Lambda = (\Delta\lambda_i)$ is a diagonal matrix, it is completely determined by Equation (2. 9).

Next suppose that $i \neq j$. Then we have

$$(\lambda_i - \lambda_j)(P\Delta P^T)_{ij} = -(P\Delta A P^T)_{ij}$$

since $\delta_{ij} = 0$. The characteristic values of A are distinct and so $\lambda_i - \lambda_j \neq 0$. Thus

$$(P\Delta P^T)_{ij} = \frac{1}{\lambda_j - \lambda_i}(P\Delta A P^T)_{ij} . \tag{2.10}$$

For convenience of notation, let $B = P\Delta A P^T = (b_{ij})$ and let $C = (c_{ij})$ where $c_{ii} = 0$ and

$$c_{ij} = \frac{1}{\lambda_j - \lambda_i} b_{ij}.$$

Then,

$$P\Delta P^T = D + C \tag{2.11}$$

where D is some diagonal matrix. Clearly B is symmetric, since ΔA is, and hence C is skew-symmetric, that is, $C^T = -C$. Consider now $(P\Delta P^T)^T$. We see that

$$\Delta P P^T = (P\Delta P^T)^T = (D+C)^T = D^T + C^T = D - C .$$

By Equation (2. 5), $P\Delta P^T + \Delta P P^T = 0$ and so

$$P\Delta P^T + \Delta P P^T = (D+C) + (D-C) = 0.$$

This implies that $2D = 0$ and hence $D = 0$. From this and Equation (2.11) it follows that

$$P \Delta P^T = C. \tag{2.12}$$

Thus $\Delta P^T = P^T C$ and so

$$\Delta P = (P^T C)^T = C^T P = -CP. \tag{2.13}$$

III. PRINCIPAL THEOREM

Equation (2. 9) of Chapter II shows how to determine a small change in the characteristic values of a symmetric matrix A, given a small symmetric change in that matrix. Similarly, Equation (2.13) gives the change in the diagonalizing matrix P. Suppose now that the change in A is dependent on time t. Let us further suppose that A is a differentiable function of t. In particular, suppose that $A(t) = A_0 + t(A_1 - A_0)$ for $0 \leq t \leq 1$, where A_0 and A_1 are real, symmetric, $n \times n$ matrices, and A_0 has n distinct characteristic values $\lambda_1^0, \lambda_2^0, \ldots, \lambda_n^0$. Let $P_0 = (p_{ij}^0)$ be the orthogonal diagonalizing matrix for A_0. Then

$$P_0 A_0 P_0^T = \Lambda_0 = (\lambda_i^0) \quad \text{and} \quad P_0 P_0^T = P_0^T P_0 = I .$$

It is clear that at time $t = 0$, $A(t)$ takes on the value A_0. Note also that A is symmetric for any t. Thus, for any t, our two basic equations hold; that is, there exists an orthogonal matrix P such that

$$PAP^T = \Lambda \tag{3. 1}$$

$$PP^T = P^T P = I. \tag{3. 2}$$

Since A is a function of time, it follows that P and Λ are also. As in Chapter II, we obtain the following equation from (3. 1).

$$\Delta PAP^T + P\Delta AP^T + \Delta P\Delta AP^T + PA\Delta P^T + \Delta PA\Delta P^T + P\Delta A\Delta P^T + \Delta P\Delta A\Delta P^T = \Delta\Lambda. \tag{3.3}$$

Let us next assume that P and Λ are differentiable functions of t. We will show that this assumption is a valid one when we prove the existence of a set of continuous solutions to a certain system of differential equations.

Now divide both sides of Equation (3.3) by Δt. This gives

$$\frac{\Delta P}{\Delta t} AP^T + P\frac{\Delta A}{\Delta t} P^T + \frac{\Delta P}{\Delta t}\Delta AP^T + PA\frac{\Delta P^T}{\Delta t} + \frac{\Delta P}{\Delta t} A\Delta P^T + P\frac{\Delta A}{\Delta t}\Delta P^T + \frac{\Delta P}{\Delta t}\Delta A\Delta P^T = \frac{\Delta\Lambda}{\Delta t}. \tag{3.4}$$

Since A, P, and Λ are differentiable functions of t,

$$\lim_{\Delta t \to 0} \Delta A = \lim_{\Delta t \to 0} \Delta P = \lim_{\Delta t \to 0} \Delta P^T = 0.$$

Also

$$\lim_{\Delta t \to 0} \frac{\Delta A}{\Delta t} = \frac{dA}{dt}, \quad \lim_{\Delta t \to 0} \frac{\Delta P}{\Delta t} = \frac{dP}{dt},$$

$$\lim_{\Delta t \to 0} \frac{\Delta P^T}{\Delta t} = \frac{dP^T}{dt}, \quad \text{and} \quad \lim_{\Delta t \to 0} \frac{\Delta\Lambda}{\Delta t} = \frac{d\Lambda}{dt}.$$

Thus, taking the limit of Equation (3.4) as Δt tends to zero gives the following equation.

$$(3.5) \qquad \frac{dP}{dt} AP^T + P \frac{dA}{dt} P^T + PA \frac{dP^T}{dt} = \frac{d\Lambda}{dt}.$$

We see now that the terms with two or more "Δ" terms as factors become zero in the limit. This is why we were able, in Chapter II, to neglect those terms, providing we chose ΔA small enough.

Similarly, from Equation (3.1) we obtain

$$(3.6) \qquad \frac{dP}{dt} P^T + P \frac{dP^T}{dt} = 0.$$

Now suppose that we can choose $\Delta t = dt$ small enough so that we do not go through a degenerate case. That is, suppose that we can choose t_0 small enough such that $A(t)$ will have distinct characteristic values for every t such that $0 \leq t \leq t_0$. (Recall that we have already assumed that $A(0) = A_0$ has distinct characteristic values. This is essential.) Once again, by showing the existence of a unique set of continuous solutions to a certain system of differential equations, we will show that this assumption is correct. This is, in fact, the main result we are seeking.

We can now proceed with Equations (3.5) and (3.6) in an analogous way to the method used in Chapter II. Doing so we obtain the following equations.

$$(3.7) \qquad \frac{d\lambda_i}{dt} = \left(P \frac{dA}{dt} P^T\right)_{ii} \quad \text{for} \quad i = 1, 2, \ldots, n$$

(3.8) $$\frac{dP}{dt} = -CP$$

where $C = (c_{ij})$ is defined by $c_{ii} = 0$,

$$c_{ij} = \frac{1}{\lambda_j - \lambda_i} b_{ij} \quad \text{with} \quad b_{ij} = (P \frac{dA}{dt} P^T)_{ij} .$$

We can also write Equation (3.8) in a more convenient fashion.

(3.9) $$\frac{dp_{ij}}{dt} = -(CP)_{ij} \quad \text{for} \quad i, j = 1, 2, \ldots, n.$$

Since $A(t) = A_0 + t(A_1 - A_0)$, it is clear that $\frac{dA}{dt} = A_1 - A_0$. Let $A_1 - A_0 = M = (m_{ij})$ and let P_i denote the ith row of P. Then Equation (3.7) becomes

(3.10) $$\frac{d\lambda_i}{dt} = P_i M P_i^T \quad \text{for} \quad i = 1, 2, \ldots, n.$$

We can write out Equations (3.9) and (3.10) explicitly by carrying out the indicated matrix multiplication.

(3.11)
$$\frac{d\lambda_i}{dt} = P_i M P_i^T = m_{11} p_{i1}^2 + m_{22} p_{i2}^2 + \ldots + m_{nn} p_{in}^2 + 2 \sum_{k<j} m_{kj} p_{ik} p_{ij}$$

$$\text{for} \quad i = 1, 2, \ldots, n.$$

$$(3.12) \qquad \frac{dp_{ij}}{dt} = -(CP)_{ij} = -\sum_{k=1}^{n} c_{ik}p_{kj}$$

$$= -\left(\sum_{k=1}^{i-1} \frac{1}{\lambda_k - \lambda_i} (PMP^T)_{ik}p_{kj} + \sum_{k=i+1}^{n} \frac{1}{\lambda_k - \lambda_i} (PMP^T)_{ik}p_{kj}\right)$$

$$\text{for} \quad i, j = 1, 2, \ldots, n.$$

We notice now that Equations (3.11) and (3.12) define a system of $n^2 + n$ differential equations. This is the system of differential equations that we will show has a unique set of continuous solutions. Furthermore, these solutions will assume the values $\lambda_1^0, \lambda_2^0, \ldots, \lambda_n^0, p_{11}^0, \ldots, p_{nn}^0,$ at time $t = 0$.

Before we can continue, we need several definitions and a generalized form of Taylor's Theorem.

Definition. Let $f(x)$ be a real-valued function defined on an open interval (a, b), and let x_0 be an interior point of (a, b). Suppose there exists a neighborhood $N(x_0)$ and a positive number L such that $x \in N'(x_0)$ implies $|f(x) - f(x_0)| < L|x - x_0|$. Then f is said to satisfy a Lipschitz condition at the point x_0.

This definition can be extended easily to a function of several variables as follows.

Definition. Let $(x_1^0, x_2^0, \ldots, x_n^0)$ be a point in R^n. Let

$f(x_1, x_2, \ldots, x_n)$ be a real-valued function defined on the domain D, where D is defined by $|x_1 - x_1^0| \leq a_1$, $|x_2 - x_2^0| \leq a_2, \ldots,$ $|x_n - x_n^0| \leq a_n$. Then f is said to satisfy a Lipschitz condition at the point $(x_1^0, x_2^0, \ldots, x_n^0)$ if there exists constants $L_1, L_2, \ldots, L_n$ such that, for all interior points $(x_1, x_2, \ldots, x_n)$ of the domain D, we have that

$$|f(x_1, x_2, \ldots, x_n) - f(x_1^0, x_2^0, \ldots, x_n^0)|$$

$$< L_1|x_1 - x_1^0| + L_2|x_2 - x_2^0| + \ldots + L_n|x_n - x_n^0| .$$

In order to state a generalized Taylor's Theorem for a function of several variables in a form analogous to the theorem for a function of one variable, we first need to define the concept of higher order differentials.

<u>Definition.</u> Let f be a real-valued function defined on a subset of R^n. The (first order) differential of f, denoted df, is a function of $2n$ variables, defined for those points $\underline{x} \in R^n$ where f has all its partial derivatives, and for every $\underline{t} \in R^n$, by the equation

$$df(\underline{x}\,;\underline{t}) = \sum_{i=1}^{n} D_i f(x) t_i$$

where $\underline{x} = (x_1, x_2, \ldots, x_n)$, $\underline{t} = (t_1, t_2, \ldots, t_n)$ and

$$D_i f(\underline{x}) = \frac{\partial f}{\partial x_i}(\underline{x}).$$

The second order differential of f, denoted $d^2 f$, is defined by the equation

$$d^2 f(\underline{x}; \underline{t}) = \sum_{i=1}^{n} \sum_{j=1}^{n} D_{i,j} f(\underline{x}) t_j t_i$$

where

$$D_{i,j} f(\underline{x}) = \frac{\partial f}{\partial x_i \partial x_j}(\underline{x}).$$

The third order differential of f, denoted $d^3 f$, is defined by the equation

$$d^3 f(\underline{x}; \underline{t}) = \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} D_{i,j,k} f(\underline{x}) t_k t_j t_i .$$

The mth order differential, denoted $d^m f$, is defined similarly.

We can now state the following

<u>Theorem (Taylor)</u>. Let f have continuous partial derivatives of order m at each point of an open set S in R^n. If $\underline{a} \in S$, $\underline{b} \in S$, $\underline{a} \neq \underline{b}$, and if the line segment, $L(\underline{a}, \underline{b})$, joining $\underline{a}$ and $\underline{b}$

lies entirely in S, then there exists a point $\underline{z}$ on $L(\underline{a}, \underline{b})$ such that

$$f(\underline{b}) = f(\underline{a}) + \sum_{k=1}^{m-1} \frac{1}{k!} d^k f(\underline{a}\,;\underline{b-a}) + \frac{1}{m!} d^m f(\underline{z}\,;\underline{b-a}).$$

The proof of this theorem can be found in many analysis texts and will be omitted here.

We can again consider our system of $n^2 + n$ differential equations. Instead of writing the actual expressions, we will use the following function notation for convenience.

(3.13)

$$\frac{d\lambda_1}{dt} = f_1(\lambda_1, \ldots, \lambda_n, p_{11}, \ldots, p_{nn})$$

$$\frac{d\lambda_2}{dt} = f_2(\lambda_1, \ldots, \lambda_n, p_{11}, \ldots, p_{nn})$$

$$\vdots$$

$$\frac{d\lambda_n}{dt} = f_n(\lambda_1, \ldots, \lambda_n, p_{11}, \ldots, p_{nn})$$

$$\frac{dp_{11}}{dt} = f_{n+1}(\lambda_1, \ldots, \lambda_n, p_{11}, \ldots, p_{nn})$$

$$\vdots$$

$$\frac{dp_{nn}}{dt} = f_{n^2+n}(\lambda_1, \ldots, \lambda_n, p_{11}, \ldots, p_{nn}).$$

It is clear from Equations (3.11) and (3.12) that each of the above functions $f_1, \ldots, f_{n^2+n}$, is continuous and satisfies the conditions of Taylor's Theorem for the open set S in R^{n^2+n} that is defined by t_0. Now let $\underline{a} = (\lambda_1^0, \ldots, \lambda_n^0, p_{11}^0, \ldots, p_{nn}^0)$ and let $\underline{b} = (\lambda_1, \ldots, \lambda_n, p_{11}, \ldots, p_{nn}^0)$. Then for $i = 1, \ldots, n^2+n$, we can expand f_i into the following Taylor's series.

$$f_i(\underline{b}) = f_i(\underline{a}) + \sum_{k=1}^{m-1} \frac{1}{k!} d^k f_i(\underline{a}\,; \underline{b-a}) + R_i$$

where $R_i = \frac{1}{m!} d^m f_i(\underline{z_i}\,; \underline{b-a})$, m any fixed integer, $z_i \in L(\underline{a}, \underline{b})$. Hence, for $i = 1, \ldots, n^2+n$,

$$f_i(\underline{b}) = f_i(\underline{a}) + df_i(\underline{a}\,; \underline{b-a}) + \frac{1}{2} d^2 f_i(\underline{a}\,; \underline{b-a})$$
$$+ \ldots + \frac{1}{(m-1)!} d^{m-1} f_i(\underline{a}\,; \underline{b-a}) + R_i\,.$$

Now if we choose t small enough, then $\underline{b}$ is not too far from $\underline{a}$. Thus the higher order differentials can be neglected since they can be made as small as we wish. Hence $f_i(\underline{b}) - f_i(\underline{a})$ will be approximately $df_i(\underline{a}\,; \underline{b-a})$. We now have the following system.

$$\frac{d\lambda_1}{dt} = f_1(\underline{b}) = f_1(\underline{a}) + \frac{\partial f_1}{\partial \lambda_1}(\underline{a})(\lambda_1 - \lambda_1^0) + \ldots + \frac{\partial f_1}{\partial p_{nn}}(\underline{a})(p_{nn} - p_{nn}^0) \tag{3.14}$$

$$\vdots$$

$$\frac{d\lambda_n}{dt} = f_n(\underline{b}) = f_n(\underline{a}) + \frac{\partial f_n}{\partial \lambda_1}(\underline{a})(\lambda_1 - \lambda_1^0) + \ldots + \frac{\partial f_n}{\partial p_{nn}}(\underline{a})(p_{nn} - p_{nn}^0)$$

$$\frac{dp_{11}}{dt} = f_{n+1}(\underline{b}) = f_{n+1}(\underline{a}) + \frac{\partial f_{n+1}}{\partial \lambda_1}(\underline{a})(\lambda_1 - \lambda_1^0) + \ldots + \frac{\partial f_{n+1}}{\partial p_{nn}}(\underline{a})(p_{nn} - p_{nn}^0)$$

$$\vdots$$

$$\frac{dp_{nn}}{dt} = f_{n^2+n}(\underline{b}) = f_{n^2+n}(\underline{a}) + \frac{\partial f_{n^2+n}}{\partial \lambda_1}(\underline{a})(\lambda_1 - \lambda_1^0) + \ldots + \frac{\partial f_{n^2+n}}{\partial p_{nn}}(\underline{a})(p_{nn} - p_{nn}^0)$$

Now let

$$L_{11} = \left|\frac{\partial f_1}{\partial \lambda_1}(\underline{a})\right|, \quad L_{12} = \left|\frac{\partial f_1}{\partial \lambda_2}(\underline{a})\right|, \ldots, L_{n^2+n,\, n^2+n} = \left|\frac{\partial f_{n^2+n}}{\partial p_{nn}}(\underline{a})\right|.$$

Then, by the triangle inequality, we see that, for $i = 1, \ldots, n^2+n$, f_i satisfies the following Lipschitz condition.

$$|f_i(\underline{b}) - f_i(\underline{a})| \leq L_{i1}|\lambda_1 - \lambda_1^0| + \ldots + L_{i,\, n^2+n}|p_{nn} - p_{nn}^0|.$$

Since each f_i is continuous and also satisfies the above Lipschitz condition at $(\lambda_1^0, \ldots, \lambda_n^0, p_{11}^0, \ldots, p_{nn}^0)$, we know from the theory of differential equations that for small enough t, the system (3. 13) has a unique set of continuous solutions $\lambda_1(t), \lambda_2(t), \ldots, \lambda_n(t)$, $p_{11}(t), \ldots, p_{nn}(t)$. Furthermore, these solutions assume the values $\lambda_1^0, \ldots, \lambda_n^0, p_{11}^0, \ldots, p_{nn}^0$ respectively at time $t = 0$.

We now state this result formally as a theorem.

Theorem. Let A_0 and A_1 be real, symmetric, $n \times n$ matrices, and suppose that A_0 has n distinct characteristic values. Let $A(t)$ be a differentiable function of t. In particular, let $A(t) = A_0 + t(A_1 - A_0)$ for $0 \leq t \leq 1$. Then, providing t is chosen small enough, $A(t)$ will also have distinct characteristic values. Furthermore, $P(t)$ and $\Lambda(t)$ will also be differentiable functions with respect to t, where $P(t)$ is the orthogonal diagonalizing matrix for $A(t)$, and $\Lambda(t)$ is the diagonal matrix of characteristic values of $A(t)$.

It should be noted here that we know that the system (3.13) has a set of solutions even if we only know that each function f_i is continuous, regardless of whether or not we know that each satisfies a Lipschitz condition. However, we want to be sure that the solutions are unique, since we want a uniquely determined diagonalizing matrix and unique characteristic values for a given matrix $A(t)$. To guarantee this, it is sufficient (but not necessary) to show, as we have done, that each of the functions satisfies a Lipschitz condition.

IV. NUMERICAL EXAMPLE

Let us now consider an example to illustrate some of the methods and results of the previous chapters.

Let $A_0 = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$. Then A_0 is clearly symmetric, and to determine the characteristic values, we solve the characteristic equation $\det(A_0 - \lambda I) = 0$. Thus we solve

$$\det(A_0 - \lambda I) = \det\begin{pmatrix} 1-\lambda & 2 \\ 2 & 1-\lambda \end{pmatrix} = 0.$$

Then

$$(1-\lambda)^2 - 4 = 0$$

$$\lambda^2 - 2\lambda + 1 - 4 = 0$$

$$\lambda^2 - 2\lambda - 3 = 0$$

$$(\lambda - 3)(\lambda + 1) = 0 .$$

Hence the characteristic values of A_0 are $\lambda_1^0 = 3$ and $\lambda_2^0 = -1$.

We next find the diagonalizing matrix P_0. It is easy to prove that the characteristic vectors belonging to distinct characteristic values of a real, symmetric matrix are orthogonal. Thus, to determine P_0, we need only find the characteristic vectors X_1 and X_2 belonging to $\lambda_1^0 = 3$ and $\lambda_2^0 = -1$ respectively, and normalize them. Then P_0 will be the matrix whose rows are X_1 and X_2.

Let $X_1 = (x_1, y_1)$ and $X_2 = (x_2, y_2)$. Then $X_1 A_0 = \lambda_1^0 X_1$

and $X_2 A_0 = \lambda_2^0 X_2$. We first solve for X_1.

$$(x_1, y_1)\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} = 3(x_1, y_1) = (3x_1, 3y_1)$$

$$x_1 + 2y_1 = 3x_1$$

$$2x_1 + y_1 = 3y_1$$

Solving this system of linear equations, we see that x_1 must equal y_1. So let $x_1 = \frac{1}{\sqrt{2}} = y_1$. Then $X_1 = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$ and X_1 is already normalized, that is, has length one. Similarly, we can show that $X_2 = (\frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}})$ and so

$$P_0 = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}.$$

A check shows that $P_0 A_0 P_0^T = \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix} = \Lambda_0$ and that $P_0 P_0^T = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I$.

Now let $\Delta A_0 = \begin{pmatrix} .03 & .02 \\ .02 & .04 \end{pmatrix}$. Then by Equation (2.9) we have

$$\Delta\lambda_i = (P_0 A_0 P_0^T)_{ii} \quad \text{for} \quad i = 1, 2.$$

Now

$$P_0 \Delta A_0 P_0^T = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} .03 & .02 \\ .02 & .04 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$$

$$= \begin{pmatrix} .055 & -.005 \\ -.005 & .015 \end{pmatrix}.$$

Hence $\Delta\lambda_1 = .055$ and $\Delta\lambda_2 = .015$. We see then that the characteristic values of the matrix $A_0 + \Delta A_0 = \begin{pmatrix} 1.03 & 2.02 \\ 2.02 & 1.04 \end{pmatrix}$ should be $\lambda_1 = 3.055$ and $\lambda_2 = -.985$. Calculating the characteristic values of $A_0 + \Delta A_0$ directly from the characteristic equation we obtain $\lambda_1 = 3.04$ and $\lambda_2 = -.98$. (If we had chosen ΔA_0 even smaller, and had carried out the arithmetic to more decimal places, we could have made the results even closer.)

From Equation (2.13) we see that $\Delta p_{ij} = -(CP_0)_{ij}$. Now

$$CP_0 = \begin{pmatrix} 0 & c_{12} \\ c_{21} & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} \frac{c_{12}}{\sqrt{2}} & -\frac{c_{12}}{\sqrt{2}} \\ \frac{c_{21}}{\sqrt{2}} & \frac{c_{21}}{\sqrt{2}} \end{pmatrix}$$

where $c_{12} = \frac{1}{\lambda_2^0 - \lambda_1^0} (P_0 \Delta A_0 P_0^T)_{12}$ and $c_{21} = -c_{12}$. Thus

$$c_{12} = -\frac{1}{4}(-.005) = .00125$$

and

$$c_{21} = -.00125.$$

Hence

$$\Delta p_{11} = \frac{-1}{\sqrt{2}} \times (.00125), \quad \Delta p_{12} = \frac{1}{\sqrt{2}} \times (.00125),$$

$$\Delta p_{21} = \frac{1}{\sqrt{2}} \times (.00125), \quad \text{and} \quad \Delta p_{22} = \frac{1}{\sqrt{2}} \times (.00125)$$

It is obvious from our example so far that the calculations for even a

2×2 case are rather tedious, and for larger cases, calculation by hand becomes quite impractical.

Again let $A_0 = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ and let $A_1 = \begin{pmatrix} 2 & -1 \\ -1 & 3 \end{pmatrix}$. Define $A(t)$ by the equation $A(t) = A_0 + t(A_1 - A_0)$ for $0 \leq t \leq 1$. Let $A_1 - A_0 = M$. Then

$$M = \begin{pmatrix} 1 & -3 \\ -3 & 2 \end{pmatrix} = \frac{dA}{dt} .$$

By Equation (3. 11) we have that

$$\frac{d\lambda_i}{dt} = m_{11}p_{i1}^2 + m_{22}p_{i2}^2 + 2 \sum_{k<j} m_{kj}p_{ik}p_{ij} \quad \text{for} \quad i = 1, 2$$

By Equation (3. 12) we have

$$\frac{dp_{ij}}{dt} = -\left(\sum_{k=1}^{i-1} \frac{1}{\lambda_k - \lambda_i} (PMP^T)_{ik} p_{kj} + \sum_{k=i+1}^{2} \frac{1}{\lambda_k - \lambda_i} (PMP^T)_{ik} p_{kj} \right)$$

$$\text{for} \quad i, j = 1, 2$$

We thus obtain the following system of six differential equations.

$$\frac{d\lambda_1}{dt} = p_{11}^2 + 2p_{12}^2 - 6p_{11}p_{12}$$

$$\frac{d\lambda_2}{dt} = p_{21}^2 + 2p_{22}^2 - 6p_{21}p_{22}$$

$$\frac{dp_{11}}{dt} = \frac{1}{\lambda_1-\lambda_2}(p_{11}p_{21}^2-3p_{11}p_{21}p_{22}-3p_{12}p_{21}^2+2p_{12}p_{21}p_{22})$$

$$\frac{dp_{12}}{dt} = \frac{1}{\lambda_1-\lambda_2}(p_{11}p_{21}p_{22}-3p_{11}p_{22}^2-3p_{12}p_{21}p_{22}+2p_{12}p_{22}^2)$$

$$\frac{dp_{21}}{dt} = \frac{1}{\lambda_2-\lambda_1}(p_{11}^2p_{21}-3p_{11}^2p_{22}-3p_{11}p_{12}p_{21}+2p_{11}p_{12}p_{22})$$

$$\frac{dp_{22}}{dt} = \frac{1}{\lambda_2-\lambda_1}(p_{11}p_{12}p_{21}-3p_{11}p_{12}p_{22}-3p_{12}^2p_{21}+2p_{12}^2p_{22}).$$

As we have shown in Chapter III, each of these functions satisfies a Lipschitz condition at the point

$$(\lambda_1^0, \lambda_2^0, p_{11}^0, p_{12}^0, p_{21}^0, p_{22}^0) = (3, -1, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}})$$

providing we choose t small enough. To demonstrate the Lipschitz conditions explicitly for the system, it is necessary to calculate 36 Lipschitz constants. These are determined by taking the partial derivative of each function with respect to each of λ_1, λ_2, p_{11}, p_{12}, p_{21} and p_{22} successively, and evaluating that partial derivative at the point $(3, -1, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}})$.

Since each function is continuous and satisfies a Lipschitz condition, for sufficiently small t, the system has a unique set of continuous solutions $\lambda_1(t)$, $\lambda_2(t)$, $p_{11}(t)$, $p_{12}(t)$, $p_{21}(t)$ and $p_{22}(t)$

that assume the values 3, -1, $\frac{1}{\sqrt{2}}$, $\frac{1}{\sqrt{2}}$, $\frac{1}{\sqrt{2}}$, and $-\frac{1}{\sqrt{2}}$, respectively, at time $t = 0$.

BIBLIOGRAPHY

1. Kemble, Edwin C. The fundamental principles of quantum mechanics. New York, Dover, 1958. 611 p.

2. Wilson, E. Bright, Jr., J. C. Decius and Paul C. Cross. Molecular vibrations: the theory of infrared and Raman vibrational spectra. New York, McGraw-Hill, 1955. 388 p.