Solutions to Riccati's Problem as Functions of Initial Values

G. S. McCARTY, JR.

Communicated by E. HILLE

Introduction. R. M. REDHEFFER has constructed [1] a set of functions defined by the solution of a RICCATI problem. Considering the solutions of his defining equations as functions of both the independent variable y and an initial point x at which the functions vanish, he shows that if

$$(1) \quad \begin{aligned} u_v &= a(y) + 2b(y)u + c(y)u^2, & u(x, x) &= 0; \\ v_v &= b(y) + c(y)u, & v(x, x) &= 0; \\ w_v &= c(y) \exp(2v), & w(x, x) &= 0 \end{aligned}$$

in an open convex set symmetric about $y = x$; and if $v(x, y)$ is continuous in x , then the partial derivatives with respect to x exist there and satisfy

$$(2) \quad \begin{aligned} -w_x &= c(x) + 2b(x)w + a(x)w^2; \\ -v_x &= b(x) + a(x)w; \\ -u_x &= a(x) \exp(2v). \end{aligned}$$

Furthermore, REDHEFFER shows the following functional equations are valid whenever u, v and w satisfy (1) and the points (x, y) , (x, z) and (y, z) are in the set:

$$(3) \quad \begin{aligned} \frac{u(x, z) - u(y, z)}{u(x, y)} &= \exp [-v(x, y) + v(x, z) + v(y, z)], \\ 1 - u(x, y)w(y, z) &= \exp [v(x, y) - v(x, z) + v(y, z)], \\ \frac{w(x, z) - w(x, y)}{w(y, z)} &= \exp [v(x, y) + v(x, z) - v(y, z)]. \end{aligned}$$

In Theorem III of [1] the first of equations (3) is considered in a form which holds when $u(x, x)$ and $v(x, x)$ differ from zero, namely

$$\frac{u(x, z) - u(y, z)}{u(x, y) - u(y, y)} = \frac{\exp [v(x, z) + v(y, z)]}{\exp [v(x, y) + v(y, y)]}.$$

W. T. REID [3] has extended these results to matrix differential equations of the RICCATI type. Some of his methods and results are used below to enlarge the class of admissible coefficients.

The present note continues these studies by generalizing systems (1), (2) and (3) to the case of arbitrary initial values. Since $v(x, y)$ can be altered by an additive constant, there is no loss of generality in setting $v(x, x) = 0$. However, setting $u(x, x) = z$ for arbitrary complex z leads to new results. Some of these specialize to the previous results when $z = 0$, but differentiation with respect to this initial value yields relations which have no analog in the case $z = 0$. The principal results are stated in Theorem I.

Definition of the functions. Following [3], let S denote an open connected set of triples (x, y, z) where x and y are real and z is complex. Let S have the property that $(x_0, y_0, z_0) \in S$ only if the closed line segment with end-points $(x_0, y_0, z_0), (x_0, x_0, z_0)$ belongs to S . Let U, Λ, V, W be complex-valued functions having argument (x, y, z) , defined over S and having everywhere there the properties:

(P₀) $1 - zU \neq 0$ and $1 - z^2 \Lambda V \neq 0$;

(P₁) the partial derivatives U_y, Λ_y, V_y, W_y exist;

(P₂) $\Lambda \neq 0$ and $V \neq 0$;

(P₃) there exist complex-valued functions a, b, c with argument y such that

$$(4) \quad \begin{aligned} a) \quad U_y &= a + 2bU + cU^2, & U(x, x, z) &= z; \\ b) \quad \Lambda_y &= (b + cU)\Lambda, & \Lambda(x, x, z) &= 1; \\ c) \quad V_y &= \frac{b + cU + (a + bU)z}{1 - zU} V - \frac{\Lambda(c + 2bz + az^2)z}{1 - zU} V^2, \\ & & V(x, x, z) &= 1; \\ d) \quad W_y &= (c + 2bz + az^2)V^2, & W(x, x, z) &= z. \end{aligned}$$

The following statements of functional dependence will be needed.

Lemma I. If U, Λ, V, W satisfy the hypotheses P₀, P₁, P₂, P₃ over a set S , then

$$V(1 - zU) = \Lambda(1 - zW).$$

Proof. This may be seen by defining a new function, f , over S :

$$f = \frac{V(1 - zU) - \Lambda(1 - zW)}{\Lambda}.$$

The hypothesis shows f has a derivative over S and satisfies $f_y(x, y, z) = 0$, $f(x, x, z) = 0$. Thus $f(x, y, z) = 0$ along the line joining (x, y, z) and (x, x, z) , and so f is identically zero in S .

An immediate consequence is that $1 - zW(x, y, z) \neq 0$.

Lemma II. For fixed z_0 , if U, Λ, V, W satisfy P_0, P_1, P_2, P_3 at some point $(x_0, y_0, z_0) \in S$, then there exists a set U, V, W defined in a neighborhood $(x, y, 0)$ of $(x_0, y_0, 0)$ and satisfying (4) at $(x_0, y_0, 0)$.

Proof. For clarity, use the symbols of (1) to denote a solution of (4) at $z = 0$ by u, e^*, w . Computation shows that if U, Λ, V, W have argument (x_0, y_0, z_0) then (4) is satisfied by

$$\begin{aligned} u(x_0, y_0) &= \frac{U - z_0 \Lambda V}{1 - z_0^2 \Lambda V}, \\ e^{*(x_0, y_0)} &= \frac{V(1 - z_0 U)}{1 - z_0^2 \Lambda V} = \frac{\Lambda(1 - z_0 W)}{1 - z_0^2 \Lambda V}, \\ w(x_0, y_0) &= \frac{W - z_0 \Lambda V}{1 - z_0^2 \Lambda V}, \end{aligned}$$

where the argument z_0 is held fast. Since the uniqueness of solutions to (4) in the plane $z = 0$ is shown in Lemma 2.1 of [3], different choices of z_0 lead to the same functions u, e^*, w .

Theorem I. If U, Λ, V, W satisfy the hypotheses P_0, P_1, P_2, P_3 over a set S , then these functions are continuous in (x, y, z) over S and have partial derivatives defined there with respect to x and z which satisfy

$$\begin{aligned} \text{a)} \quad & -W_x = c + 2bW + aW^2, \\ \text{b)} \quad & -V_x = (b + aW)V, \\ (5) \quad & \text{c)} \quad -\Lambda_x = \frac{b + aW + (c + bW)z}{1 - zW} \Lambda - \frac{V(a + 2bz + cz^2)z}{1 - zW} \Lambda^2, \\ & \text{d)} \quad -U_x = (a + 2bz + cz^2)\Lambda^2, \end{aligned}$$

where a, b and c have argument x . Moreover,

$$\begin{aligned} \text{a)} \quad & U_x = \Lambda^2, & \text{b)} \quad \Lambda_x &= \frac{(W - z\Lambda V)\Lambda}{1 - zW}, \\ (6) \quad & \text{c)} \quad V_x &= \frac{(U - z\Lambda V)V}{1 - zU}, & \text{d)} \quad W_x &= V^2. \end{aligned}$$

Proof. The proof of Theorem I proceeds by first establishing that (4) admits a unique solution over S , and then exhibiting a set U, Λ, V, W which satisfies (4), (5) and (6).

Uniqueness is established by the method of [3]. It is necessary to introduce a matrix $H(x, y, z)$ which serves the same purpose as REID's matrix $H(x, y)$. Define the matrices G and H by

$$G(y) = \begin{bmatrix} -b(y) & -c(y) \\ a(y) & b(y) \end{bmatrix},$$

$$H(x, y, z) = \frac{1}{V(1-zU)} \begin{bmatrix} 1 - z^2\Lambda V & -W + z\Lambda V \\ U - z\Lambda V & \Lambda V - UW \end{bmatrix},$$

where for brevity the functions U, Λ, V, W are understood to have argument (x, y, z) . It may be noted that $H(x, y, z)$ at $z = 0$ reduces to REID's $H(x, y)$.

By (P_0) and (P_2) , H is defined over S ; by (P_1) , H has there a partial derivative with respect to y . Clearly $\det H(x, y, z) = 1$ and $H(x, x, z) = I$. An elementary but lengthy calculation shows that $H_y(x, y, z) = G(y)H(x, y, z)$. It follows that $(H^{-1})_y = -H^{-1}G$. Moreover, if $H_0 = H_0(x, y, z)$ is any solution of $H_{0y} = GH_0$ over S then $(H^{-1}H_0)_y = 0$ there. Particularly, if $H_0(x, x, z) = I$ then $H^{-1}H_0 = I$ so that $H_0 = H$. Thus H is unique. Choosing $H_0 = H(x, y, 0)$ shows that $H(x, y, z)$ is independent of z ; this yields (6).

Now denote this unique matrix H by

$$H = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}.$$

If $z = 0$ then inspection shows the functions U, Λ, V, W are determined by $\alpha, \beta, \gamma, \delta$. Considering points where $z \neq 0$, and remembering that $V \neq 0$ by (P_2) and $1 - zU \neq 0$ by (P_0) , one may verify the relations

$$(\alpha - z\gamma)V = 1,$$

$$\Lambda = \frac{U\alpha - \gamma}{z},$$

$$W = \frac{U\alpha - \beta - \gamma + Uz\beta}{\alpha - z\gamma},$$

$$(z\alpha + z^2\beta)U^2 - (\alpha + z\beta + z\gamma + z^2\delta)U + \gamma + z\delta = 0.$$

If the coefficient of U^2 in this quadratic, namely $z\alpha + z^2\beta$, should ever vanish, then $U = 1/z$, a contradiction of (P_0) . Therefore the quadratic formula may be used to find that the two solutions are:

$$U_1 = \frac{\gamma + z\delta}{\alpha + z\beta}, \quad U_2 = 1/z.$$

Since the second solution is in contradiction to (P_0) , U_1 is seen to be the unique function U defined by H . But then U uniquely defines the functions W and Λ , so that the quadruple U, Λ, V, W is uniquely defined by H over S . It follows that the solution of (4) is unique.

Actually it is easily verified that, in terms of the functions of (1),

$$U(x, y, z) = u(x, y) + \frac{z \exp 2v(x, y)}{1 - zw(x, y)},$$

$$(7) \quad \Lambda(x, y, z) = \frac{\exp v(x, y)}{1 - zw(x, y)},$$

$$V(x, y, z) = \frac{\exp v(x, y)}{1 - zu(x, y)},$$

$$W(x, y, z) = w(x, y) + \frac{z \exp 2v(x, y)}{1 - zu(x, y)}$$

is a solution of (4), and therefore the unique solution over S .

By Theorem 2.1 of [3] applied in the scalar case, the functions of (1) are continuous in (x, y) . It follows that U , Λ , V and W are continuous in (x, y, z) over S . Note here that the points for which either $z = w^{-1}(x, y)$ or $z = u^{-1}(x, y)$ have been excluded from S by (P_1) together with (7).

Again using the results of [1] and [3] as given in (2), one may verify (5). Thus the functions defined in (7) satisfy (4), (5) and (6), and the proof of the theorem is complete.

It may be remarked that a simple computation shows that (P_0) is in reality just one restriction, not two as it appears to be. The use of BERNOULLI equations in (4c) and (5c) is chosen to conform more closely to the notation of [1], [2] and [3]; linear expressions offer little simplification. The definition of the initial value $\Lambda(x, x, z)$ may be generalized to an arbitrary nonzero complex constant; such a constant will then appear in the definitions (4) in an obvious way. On the other hand, the author was unable satisfactorily to generalize (4) to the case where $W(x, x, z)$ differs from z and $V(x, x, z)$ is not 1.

Some properties of the functions.

Theorem II. *If the hypothesis of Theorem I is satisfied, the points concerned are all in S and the expressions below are defined, then*

$$(8) \quad \begin{aligned} \text{a) } W(x, y, z) &= \frac{1}{U(y, x, 1/z)}, \\ \text{b) } zU(y, x, 1/z) &= \frac{\Lambda(y, x, 1/z)}{V(x, y, z)} = \frac{z}{W(x, y, z)}. \end{aligned}$$

Proof. It is easy to verify that, under the hypotheses, $1/U(y, x, 1/z)$ is a solution of (5a). Further, expression of U and W in terms of the functions of (1) reduces (8a) to an identity. Differentiation of (8a) with respect to z gives (8b).

Theorem III. *If U , Λ , V , W satisfy (4) and the points concerned are in S , then for arbitrary p*

$$(9) \quad \begin{aligned} \text{a) } U(x, y, z) &= U[p, y, U(x, p, z)], \\ \text{b) } \Lambda(x, y, z) &= \Lambda(x, p, z)\Lambda[p, y, U(x, p, z)], \\ \text{c) } V(x, y, z) &= V(p, y, z)V[x, p, W(p, y, z)], \\ \text{d) } W(x, y, z) &= W[x, p, W(p, y, z)]. \end{aligned}$$

Since (9) at the plane $z = 0$ yields (3), these relations offer a natural generalization of the functional equations of [1].

Corollary.

$$\Lambda(x, y, z) = \frac{1}{\Lambda[y, x, U(x, y, z)]},$$

$$V(x, y, z) = \frac{1}{V[y, x, W(x, y, z)]}.$$

Proof of Theorem III. Identity (9a) is a restatement of Theorem II of [2], where it is proved in greater generality. It may also be seen from (5d) and (4a), whence, using subscripts to denote differentiation by 1st, 2nd, or 3rd variable,

$$-U_1[y, z, U(x, y, p)] = U_3[y, z, U(x, y, p)]U_2(x, y, p)$$

which implies

$$\frac{\partial}{\partial y} U[y, z, U(x, y, p)] = 0.$$

Setting $y = z$ and renaming variables gives (9a).

In entirely symmetric fashion, (9d) may be shown from (4d) and (5a). It is also a consequence of (8) and (9a). Identities (9b) and (9c) can be verified most easily by using (3) and (7). Alternatively it may be mentioned that (9b) and (9c) are consequences of (4), (5) and (6) and the technique used in the case of (9a); to verify this, first notice the following differential relations:

$$V_z = (b + az)V + (c + 2bz + az^2)VV_z,$$

$$-\Lambda_z = (b + cz)\Lambda + (a + 2bz + cz^2)\Lambda\Lambda_z.$$

Further properties of the matrix H.

Theorem IV. *If the hypotheses of Theorem I are satisfied, and H is the matrix defined there, then*

$$\text{a) } H_z = -HG(x),$$

$$(10) \quad \text{b) } H^{-1} = \frac{1}{\Lambda(1 - zW)} \begin{pmatrix} \Lambda V - UW & W - z\Lambda V \\ -U + z\Lambda V & 1 - z^2\Lambda V \end{pmatrix},$$

$$\text{c) } H(x_1, x_3, z) = H(x_2, x_3, z)H(x_1, x_2, z).$$

Proof. The proof of (10c) rests on Theorem VII of [1], where the matrix $M(x, y)$ defined there is seen to be equivalent to the transpose of $H(x, y, z)$. Equation (10c) yields a set of (complicated) functional equations analogous to (3) but containing the arbitrary initial value z .

Setting $x_1 = x_3$ in (10c) and differentiating with respect to x_2 gives (10a). This relation is equivalent to (5).

BIBLIOGRAPHY

- [1] REDHEFFER, R. M., "On Solutions of Riccati's Equation as Functions of the Initial Values," *J. M. R. A.*, **5** (1956), 835-848.
- [2] REDHEFFER, R. M., "The Riccati Equation: Initial Values and Inequalities," *Math. Annalen*, **133** (1957), 235-250.
- [3] REID, W. T., "Solutions of a Riccati Matrix Differential Equation as Functions of Initial Values," *J. Math. and Mech.*, **8** (1959), 221-230.

University of California
at Los Angeles