*The Mycielski-Paszkowski Diffusion Problem**

RAYMOND M. REDHEFFER

Communicated by R. BELLMAN

0. Summary of the mathematical content. This paper is devoted, for the most part, to the proof of a specific mathematical theorem. Let a $3n$ by $3n$ matrix S be defined by

$$S(x, y) = \begin{bmatrix} p & Q & 0 \\ q & P & 0 \\ r & R & 1 \end{bmatrix}$$

where the entries are real n by n matrix functions of the two real variables x and y . Suppose that $S(y, y) = 1$, that $S_v(y, y)$ is piecewise continuous, that pP is nonsingular, and that the entries of S satisfy the six nonlinear functional equations (10), (11). Then the elements s_{ii} of S satisfy

$$s_{ii} \geq 0, \quad \sum_i s_{ii} \leq 1, \quad \text{for } y > x,$$

if and only if the matrix $S_v(y, y)$ is such that its off-diagonal entries are nonnegative and the sum of the elements in each column is nonpositive, for all y under consideration.

It is found that existence of $S_v(y, y)$, together with the functional equations, implies existence of $S_v(x, y)$. Hence the entries of S satisfy the system of ordinary differential equations which one would get by formal differentiation. The theorem just stated is equivalent to a similar theorem about this system (Sec. 7).

The immediate interest of the result is that it solves a problem that has been outstanding for several years. In an investigation summarized below, MYCIELSKI & PASZKOWSKI were led to the differential system forming the subject of our theorem, except that the coefficients were constant. Even with this simplification they were unable to establish the desired inequalities in

* The preparation of this paper was sponsored in part by the Office of Naval Research. Reproduction in whole or in part is permitted for any purpose of the United States Government.

the solution. The present result settles the question not only for constant coefficients but for arbitrary functions. In the course of the analysis we also generalize the probability model, and give a new derivation of the functional equations.

Although our primary purpose is to answer the question raised by MYCIELSKI & PASZKOWSKI, the analysis fits into a broader context. Functional and differential equations similar to these arise in a variety of significant investigations, ranging from transmission-line theory [3] to neutron diffusion [8]. The inequalities assert the dissipative character of the process; *e.g.*, that the matrix corresponds to a passive transducer in the transmission-line case, or that it admits a probability model in the case of diffusion. The key step in analysis of these inequalities is to pass from relations in the small to those in the large; and our attack gives a simple, systematic method of taking this step. Thus the semigroup methods that have long been used to study linear equations can be carried over to a broad class of nonlinear problems.

Although this paper is concerned with real coefficients and l_1 norm, the case of complex coefficients and l_2 norm is developed with equal ease [6]. The methods also carry over from matrices to operators on a Hilbert space $\mathcal{H}$. The essential desideratum is that the functional equations for operators on $\mathcal{H}$ be suitably related to a certain "scattering process" on $\mathcal{H} \times \mathcal{H} \times \cdots \times \mathcal{H}$.

1. Motivation: The probability model. For $n = 1$ the probability problem to which we have referred was first introduced by C. RYLL-NARDZEWSKI (*Soc. Polonaise de Math.*, Jan. 2, 1954) and, in more general form, by MYCIELSKI & PASZKOWSKI [1]. The latter's problem is the following: A material endowed with mass is distributed on a straight line Π , and a molecule moving on this line is reflected, transmitted, or absorbed when it encounters a given segment I of Π . These phenomena are random and independent with respect to non-overlapping segments. Indeed, the probabilities associated with a given segment depend only on the direction of motion of the molecule and on the mass of material in the segment. The probabilities of transmission, reflection, and absorption for a segment of mass x are designated by $p(x)$, $q(x)$, $r(x)$ for a molecule moving from left to right, and by $P(x)$, $Q(x)$, $R(x)$ for a molecule moving from right to left. These functions are thought, at first, to be defined without reference to the rest of the line segment. That is, the segment of mass x is withdrawn from the line, and a molecule is incident upon it. In this situation the foregoing expressions give the probabilities of the various events that can happen.

Next the effect of the rest of the line is taken into account. By considering two adjacent segments of masses x and y , one is led to the nonlinear functional equations

$$(1) \quad \begin{aligned} p(x+y) &= p(x)p(y)[1 - Q(x)q(y)]^{-1}, \\ q(x+y) &= q(x) + p(x)P(x)q(y)[1 - Q(x)q(y)]^{-1}. \end{aligned}$$

(The procedure yielding (1) is conceptually similar to that in the derivation of the Poisson law. There the probability $p_n(x)$ of n points in an interval of length x is determined by the functional equation one gets from two adjacent segments x and y .) MYCIELSKI & PASZKOWSKI determine the unknown probabilities as functions of x using (1), the equations which result therefrom by symmetry, and the fact that the variables actually represent probabilities. The last means that the process is *dissipative* in the sense of this paper.

At about the same time the author had encountered equations similar to (1) in a problem on transmission-line theory [3]. This fact suggested that the *scattering matrix* of circuit analysis could profitably be introduced into the probability problem, and such proved to be the case [4]. The new method allowed the probabilities associated with a given interval I to be arbitrary functions of its endpoints (x, y) , and a complete characterization of p, q, r, P, Q, R was given under that hypothesis. The case $p(x, y) = p(y - x), \dots$ corresponds to the case [1].

In [2] the problem is generalized by PASZKOWSKI as follows. Let $\Pi_1, \Pi_2, \dots, \Pi_n$ be parallel lines each endowed with mass. On this system a molecule moves in such a way that it can change the direction of its progress, jump from one line to another (but only in a direction perpendicular to the lines), or be stopped. In the latter case it does not again start to move. Let $I_1, I_2, \dots, I_n$ with $I_i \subset \Pi_i$ be segments which are orthogonal projections one on the other. Under the hypothesis that the molecule has just entered the i^{th} segment I_i from the left, PASZKOWSKI lets p_{ij} denote the conditional probability that it emerges from the j^{th} segment at the right; q_{ji} , the probability that it emerges from the j^{th} segment at the left; and r_{ji} , the probability that it stops on the j^{th} segment. Thus p_{ij} , q_{ji} , and r_{ji} can be thought to be probabilities of transmission, reflection, and absorption, respectively, when the molecule is incident from the left. Similarly, P_{ij} , Q_{ji} , and R_{ji} denote the probabilities of transmission, reflection, and absorption when the molecule is incident from the right. Functional equations similar to (1) are obtained in [2] for this problem, differential equations are found from those, and certain inequalities for the coefficients are derived from the fact that the variables are probabilities.

Although the author solves the problem completely when the segments I_i are identical, he remarks that he does not know whether his inequalities are sufficient to ensure that $p, q, \dots$ are nonnegative in the general case. One purpose of the present note is to answer this question in the affirmative, and thus to show that the solution in [2] is a complete solution. Another purpose is to generalize the assumptions as outlined previously for $n = 1$. That is, instead of having an additive measure as in [2], we shall allow $p_{ij}, q_{ji}, \dots$ to be arbitrary functions of the end-points (x, y) of the intervals I_k used in the definition. This increase in generality corresponds to having arbitrary functions for the coefficients of the associated differential equations, whereas in [2] the coefficients are constant.

2. The scattering matrix. To introduce a scattering matrix we need something to be scattered. As in [4] this requirement can be met by presenting the molecule to the system $\{I_i\}$ in a random manner. Specifically, let the column vector V_2 be such that its i^{th} component V_2^i measures the probability that we toss the molecule at the segment I_i from the right. Similarly V_3^i represents the probability that we toss the molecule at I_i from the left, whereas A^i represents the probability that we place the molecule on the segment I_i . The vectors V_2 , V_3 , and A have nonnegative components and the sum of all $3n$ components is 1.

Under the foregoing hypothesis the molecule can emerge from the segment I_i at the right in $2n$ mutually exclusive ways, corresponding to the $2n$ components of V_2 and V_3 . Hence if the corresponding probability vector is V_1 , we get

$$(2) \quad V_1^i = \sum_{i=1}^n Q_{ii} V_2^i + \sum_{i=1}^n p_{ii} V_3^i.$$

Here the j^{th} component V_1^i represents the probability of emerging from I_i at the right.

Similarly, if V_4^i represents the probability of emerging from I_i at the left, under the hypothesis (V_2, V_3, A) , then

$$(3) \quad V_4^i = \sum_{i=1}^n P_{ii} V_2^i + \sum_{i=1}^n q_{ii} V_3^i.$$

Finally, the probability a^i of coming to rest on I_i is

$$(4) \quad a^i = \sum_{i=1}^n R_{ii} V_2^i + \sum_{i=1}^n r_{ii} V_3^i + A^i.$$

We define an $n \times n$ matrix p by $p' = (p_{ii})$, where the prime denotes the transpose, and, similarly, q, r, P, Q, R are defined in terms of $q_{ii}, r_{ii}, \dots$. In this notation (2)–(4) assume the form

$$(5) \quad \begin{bmatrix} V_1 \\ V_4 \\ a \end{bmatrix} = \begin{bmatrix} p & Q & 0 \\ q & P & 0 \\ r & R & 1 \end{bmatrix} \begin{bmatrix} V_3 \\ V_2 \\ A \end{bmatrix}.$$

Since (5) gives

$$(6) \quad \begin{bmatrix} V_1 \\ V_4 \end{bmatrix} = \begin{bmatrix} p & Q \\ q & P \end{bmatrix} \begin{bmatrix} V_3 \\ V_2 \end{bmatrix},$$

the study of Q, P, q, p can be carried out independently of R and r . In this paper the $3n$ by $3n$ matrix in (5) is called the *scattering matrix* associated with the probability problem; the vectors (V_3, V_2, A) and (V_1, V_4, a) are the *incident* and *scattered* vectors, respectively.

3. The functional equations. As in [1] and [2], we consider two adjacent sets of intervals $\{I_i\}$ and $\{J_i\}$ with I_i at the left of J_i . The end-points of the intervals J_i lie on a common orthogonal trajectory of Π_i , and so do the ends of I_i . We denote the systems by I and J respectively, and the union by $I \cup J$. It is assumed that the events associated with I and J are independent.

Let a molecule be presented to $I \cup J$ in such a way that V_2^i gives the probability that it enters J_i from the right, V_5^i the probability that it enters I_i from the left, and A^i the probability that it is placed on the segment $I_i \cup J_i$. The latter probability vector we shall suppose decomposed in the form

$$A = A_1 + A_2$$

where A_1^i is the probability that the molecule is placed on I_i and A_2^i the probability that it is placed on J_i .

Under this hypothesis let V_3 give the probability that the molecule enters J from the left, and V_4 the probability that it enters I from the right. Since I and J are adjacent V_3 gives also the probability of leaving I from the right, and V_4 the probability of leaving J from the left. The vectors V_3 and V_4 are supposed to take full account of the interaction between I and J . Although we shall not require expressions for V_3 and V_4 , they can be thought to be defined by infinite series analogous to the series in [2].

If V_1 represents the probability of emerging from $I \cup J$ at the right and V_6 at the left, then (5) gives

$$(7) \quad \begin{aligned} \begin{pmatrix} V_3 \\ V_6 \\ a_1 \end{pmatrix} &= \begin{pmatrix} p_1 & Q_1 & 0 \\ q_1 & P_1 & 0 \\ r_1 & R_1 & 1 \end{pmatrix} \begin{pmatrix} V_5 \\ V_4 \\ A_1 \end{pmatrix}, \\ \begin{pmatrix} V_1 \\ V_4 \\ a_2 \end{pmatrix} &= \begin{pmatrix} p_2 & Q_2 & 0 \\ q_2 & P_2 & 0 \\ r_2 & R_2 & 1 \end{pmatrix} \begin{pmatrix} V_3 \\ V_2 \\ A_2 \end{pmatrix} \end{aligned}$$

where the subscript 1 refers to I and 2 to J . Here a_1^i gives the probability that the molecule comes to rest on I_i and a_2^i that it comes to rest on J_i . By the assumed independence the vector

$$a = a_1 + a_2$$

gives the probability of coming to rest on the various segments of $I \cup J$.

Upon eliminating V_3 and V_4 from the system (7) and adding the equations involving A_1 and A_2 , we find that

$$(8) \quad \begin{pmatrix} V_1 \\ V_6 \\ a \end{pmatrix} = \begin{pmatrix} p_3 & Q_3 & 0 \\ q_3 & P_3 & 0 \\ r_3 & R_3 & 1 \end{pmatrix} \begin{pmatrix} V_5 \\ V_2 \\ A \end{pmatrix}$$

where the latter matrix has the entries

$$\begin{aligned}
 p_3 &= p_2(1 - Q_1 q_2)^{-1} p_1, \\
 (9) \quad q_3 &= q_1 + P_1(1 - q_2 Q_1)^{-1} q_2 p_1, \\
 r_3 &= R_1(1 - q_2 Q_1)^{-1} q_2 p_1 + r_2(1 - Q_1 q_2)^{-1} p_1 + r_1.
 \end{aligned}$$

Interchanging small and capital letters and the subscripts 1 and 2 yields P_3 , Q_3 , and R_3 . Because of the probability interpretation it is natural to assume (and we do assume) that $1 - q_2 Q_1$ and $1 - Q_1 q_2$ are nonsingular. Later we shall require that p_i and P_i be nonsingular, a condition which, in the presence of the other inequalities, implies the condition just stated.

The arbitrary character of V_5 , V_2 , and A shows that the matrix in (8) must be the scattering matrix associated with $I \cup J$. Hence if the end-points of I and J are (x, y) and (y, z) , comparison of (5) and (8) gives the functional equations

$$\begin{aligned}
 (10a) \quad p(x, z) &= p(y, z)[1 - Q(x, y)q(y, z)]^{-1}p(x, y), \\
 (10b) \quad q(x, z) &= q(x, y) + P(x, y)[1 - q(y, z)Q(x, y)]^{-1}q(y, z)p(x, y), \\
 (10c) \quad r(x, z) &= R(x, y)[1 - q(y, z)Q(x, y)]^{-1}q(y, z)p(x, y) \\
 &\quad + r(y, z)[1 - Q(x, y)q(y, z)]^{-1}p(x, y) + r(x, y), \\
 (11a) \quad P(x, z) &= P(x, y)[1 - q(y, z)Q(x, y)]^{-1}P(y, z), \\
 (11b) \quad Q(x, z) &= Q(x, y) + p(y, z)[1 - Q(x, y)q(y, z)]^{-1}Q(x, y)P(y, z), \\
 (11c) \quad R(x, z) &= r(y, z)[1 - Q(x, y)q(y, z)]^{-1}Q(x, y)P(y, z) \\
 &\quad + R(x, y)[1 - q(y, z)Q(x, y)]^{-1}P(y, z) + R(y, z).
 \end{aligned}$$

When equations (97), (99), (100), (102), (103), and (104) of [2] are written in matrix notation they are found to agree with the equations just given. However, for our purposes the important fact is not that we have a new derivation of the functional equations, but that *the functional equations are consistent with (7)–(9)*.

When $Q = q = 0$ then (10a) and (11a) are the Chapman-Kolmogoroff equations for this problem, and (10c) and (11c) are equally evident. A model for this case can be obtained by imagining a turbulent but steady-state flow. If a particle is introduced at a point i on the cross-section at x of the stream, $p_{i,j}(x, y)$ gives the probability that the particle emerges at j on the cross-section y .

It might appear that our results are less general than those of [2], in that the argument x of $p_{i,j}(x)$, $q_{i,j}(x)$, $\dots$ in [2] is a vector. That is, in [2]

$$x = [\mu_1(I_1), \mu_2(I_2), \dots, \mu_n(I_n)]$$

where μ_i is the measure defined on I_i . However, this situation is really encompassed by ours, since we allow $p(x, y)$, $\dots$ to be an *arbitrary* matrix function

of the ends (x, y) of the family $\{I_i\}$. To get PASZKOWSKI's form we only need to take

$$p(x, y) = p \left[\int_x^y d\mu_1, \int_x^y d\mu_2, \dots, \int_x^y d\mu_n \right]$$

where p on the right is a function of n arguments; and similarly for $q, r, \dots$. We shall keep the arbitrary functional form throughout the analysis, though PASZKOWSKI assumes

$$\mu_1 = \mu_2 = \dots = \mu_n$$

in his differential equations and the results depending thereon.

4. The differential equations. The functional equations satisfied by p, q, P, Q are the same as those considered by REID [5]. To apply REID's results we assume from now on that p and P are nonsingular and that the derivatives

$$(12) \quad p_\nu(x, y), \quad q_\nu(x, y), \quad P_\nu(x, y), \quad Q_\nu(x, y)$$

exist on the line $y = x$. The physical interpretation gives the initial conditions [2]

$$(13) \quad p(x, x) = P(x, x) = 1, \quad q(x, x) = Q(x, x) = 0$$

where 1 denotes the identity matrix and 0 the zero matrix. These conditions are explicitly postulated.

Theorem 3.2 of [5] shows, then, that p, q, P , and Q satisfy the differential equations

$$(14) \quad \begin{cases} Q_\nu = \Lambda + kQ + QK + Q\Lambda Q, \\ P_\nu = P(K + \lambda Q), \\ p_\nu = (k + Q\lambda)p, \\ q_\nu = P\lambda p, \end{cases}$$

$$(15) \quad \begin{cases} -Q_x = p\Lambda P, \\ -P_x = (K + q\Lambda)P, \\ -p_x = p(k + \Lambda q), \\ -q_x = \lambda + Kq + qk + q\Lambda q. \end{cases}$$

The argument of the functions p, q, P, Q is (x, y) in each case, whereas the argument of λ, k, Λ, K is y in the first four equations and x in the last four. These functions are given in terms of (12) by

$$\Lambda(x) = Q_\nu(x, x), \quad K(x) = P_\nu(x, x), \quad k(x) = p_\nu(x, x), \quad \lambda(x) = q_\nu(x, x),$$

as is clear when we set $y = x$ in (14).

The equations agree with (112), (113), (114), and (115) of [2], and with the other four equations obtained from these by symmetry. The reason we get a minus sign in (15) is that our x specifies the left-hand end-point of the intervals I , whereas x in the differential equations of [2] represents the masses, supposed all equal, of I_i . Thus the x in [2] decreases when our x increases. The functions λ , k , Δ , K are constant in [2] because the assumption of an additive measure, the same on all intervals, gives $p(x, y) = p(y - x)$, $q(x, y) = q(y - x)$, $\dots$. In this case $k = p'(0)$, $\lambda = q'(0)$, $\dots$.

Although the equations for r and R were not considered by REED, they are readily obtained by a procedure similar to his. (Cf. also [2] and [7].) It is supposed that

$$r(x, x) = R(x, x) = 0$$

and that

$$\mu(x) = r_v(x, x), \quad M(x) = R_v(x, x)$$

exist. If we subtract $r(x, y)$ from the functional equation (10c), divide by $z - y$, and then let $z \rightarrow y$, we get

$$r_v = R\lambda p + \mu p.$$

Similarly

$$-R_x = r\Delta P + MP.$$

To get r_x we set $z = x$ in (10c) and in (10a). The identity

$$(1 - qQ)^{-1}q = q(1 - Qq)^{-1}$$

then yields

$$(16) \quad -r(x, y) = R(x, y)q(y, x)p^{-1}(y, x) + r(y, x)p^{-1}(y, x).$$

Upon differentiating with respect to x and using the relations already obtained we get

$$\begin{aligned} -r_x(x, y) = & \mu + Mq(x, y) + r(x, y)k + r(x, y)\Delta q(x, y) \\ & + [R(x, y)P(y, x) + r(x, y)Q(y, x) + R(y, x)]\lambda \end{aligned}$$

where the argument of μ , M , $\dots$ is x . In this derivation we needed (16) and the identity

$$P(x, y)q(y, x)p^{-1}(y, x) = -q(x, y)$$

which follows by setting $z = x$ in (10a) and (10b). Upon setting $z = x$ in (11a) and (11c) we find that the coefficient of λ is 0 in the equation for r_x . Since the result for R_v can be got by symmetry, we have shown that

$$(17a) \quad \begin{cases} R_v = M + \mu Q + RK + R\lambda Q, \\ r_v = (R\lambda + \mu)p, \end{cases}$$

$$(17b) \quad \begin{cases} -R_x = (r\Lambda + M)P, \\ -r_x = \mu + Mq + rk + r\Lambda q. \end{cases}$$

These results agree with (116) and (119) of [2].

5. The constraints. The conditions for physical realizability can be presented from two viewpoints. On the one hand, we have, as in [2],

$$(18) \quad \begin{cases} p_{ii} \geq 0, & q_{ii} \geq 0, & r_{ii} \geq 0, & P_{ii} \geq 0, & Q_{ii} \geq 0, & R_{ii} \geq 0, \\ \sum_{i=1}^n (p_{ii} + q_{ii} + r_{ii}) = 1, & \sum_{i=1}^n (P_{ii} + Q_{ii} + R_{ii}) = 1. \end{cases}$$

On the other hand, if the $3n$ incidence vector (V_s, V_2, A) of (5) represents an exhaustive set of probabilities, then the scattered vector (V_1, V_4, a) must have the same property. That is, if the first $3n$ -vector has nonnegative components whose sum is 1, the same must be true of the second.

The reader can easily show that the two conditions are equivalent. When they hold, the associated $3n$ by $3n$ matrix is termed a *probability matrix*. Our objective is to find necessary and sufficient conditions on the coefficients of the differential system for the solutions to form a probability matrix in this sense. As a preliminary, we give a condition for (9) to be well-defined.

Let $A = (a_{ii})$ be an $m \times m$ matrix with nonnegative elements. We define

$$\|A\| = \sup \sum \sum a_{ii} V^i$$

where $V = V^i$ ranges over all vectors with components V^i satisfying

$$V^i \geq 0, \quad \sum V^i = 1.$$

Then $\|AB\| \leq \|A\| \|B\|$, $\|A + B\| \leq \|A\| + \|B\|$, and each element of A is at most equal to $\|A\|$.

If the matrix (5) is a probability matrix then the $n \times n$ matrices p, P, q, Q, r, R all have norm at most 1; and actually $\|Q\| < 1$ provided P is nonsingular. Thus when the matrices (7) are probability matrices and P_1 or p_2 is nonsingular, we have $\|q_2 Q_1\| < 1$. It follows that

$$\sum (q_2 Q_1)^n = (1 - q_2 Q_1)^{-1}$$

converges, so that $1 - q_2 Q_1$ is nonsingular (see also [2], Lemma 1). Similarly for $1 - Q_1 q_2$. The series representation shows, incidentally, that the inverses have nonnegative entries.

The foregoing discussion of constraints fits into the pattern set by [1] and [2]. For our own purposes, however, a slightly more general formulation is preferable. We say that a matrix S is *dissipative* if it has nonnegative entries and $\|S\| \leq 1$ (cf. [4]). When applied to our $3n \times 3n$ matrices this notion does not require that the probabilities be exhaustive, but allows some other event to happen.

For example, there might be a finite probability that the particle disappears, or that it lands on an extra line Π_{n+1} .

Any matrix $S = (s_{ij})$ is dissipative if and only if

$$s_{ii} \geq 0, \quad \sum_i s_{ii} \leq 1.$$

Hence a necessary and sufficient condition for our $3n \times 3n$ matrix to be dissipative is given by (18) with the two equality signs replaced by " $\leq$ ".

6. The star product. The matrix (8), with entries defined by (9), is termed the *star product*

$$(19) \quad \begin{bmatrix} p_1 & Q_1 & 0 \\ q_1 & P_1 & 0 \\ r_1 & R_1 & 1 \end{bmatrix} * \begin{bmatrix} p_2 & Q_2 & 0 \\ q_2 & P_2 & 0 \\ r_2 & R_2 & 1 \end{bmatrix}$$

of the matrices (7). We shall form such products only for dissipative matrices with nonsingular pP , so that the result is well-defined. The relevance of the star product rests partly on the following:

Lemma 1. *If the matrices (19) are dissipative and have nonsingular p 's and P 's the same is true of the star product.*

It is clear that the star product has nonnegative entries, and that its transmission entries p and P are nonsingular. Moreover, if the vectors appearing in (7) have nonnegative components, the dissipative character of the matrices gives

$$\begin{aligned} \sum (V_3^i + V_6^i + a_i^i) &\leq \sum (V_5^i + V_4^i + A_i^i), \\ \sum (V_1^i + V_4^i + a_i^i) &\leq \sum (V_3^i + V_2^i + A_i^i). \end{aligned}$$

Upon adding these inequalities we get

$$\sum (V_1^i + V_6^i + a_i^i) \leq \sum (V_5^i + V_2^i + A_i^i)$$

with $A = A_1 + A_2$ and $a = a_1 + a_2$. The latter inequality shows that the matrix (8) is dissipative, since the vectors V_2 , V_5 , and A can be prescribed at will. This completes the proof.

Lemma 2. *Star multiplication of the matrices considered in Lemma 1 is associative.*

Denoting the three dissipative $3n \times 3n$ matrices by S_i , consider the three systems

$$\begin{bmatrix} V_{1+2i} \\ V_{4+2i} \\ a_i \end{bmatrix} = S_i \begin{bmatrix} V_{3+2i} \\ V_{2+2i} \\ A_i \end{bmatrix}$$

obtained when $i = 0, 1, 2$. By the definition (7)–(9) of star multiplication

$$\begin{bmatrix} V_1 \\ V_6 \\ a_0 + a_1 \end{bmatrix} = (S_1 * S_0) \begin{bmatrix} V_5 \\ V_2 \\ A_0 + A_1 \end{bmatrix}, \quad \begin{bmatrix} V_3 \\ V_8 \\ a_1 + a_2 \end{bmatrix} = (S_2 * S_1) \begin{bmatrix} V_7 \\ V_4 \\ A_1 + A_2 \end{bmatrix}.$$

Applying the definition again yields

$$(19) \quad \begin{bmatrix} V_1 \\ V_8 \\ a \end{bmatrix} = S \begin{bmatrix} V_7 \\ V_2 \\ A \end{bmatrix}$$

where $a = a_0 + a_1 + a_2$, $A = A_0 + A_1 + A_2$, and

$$S = S_2 * (S_1 * S_0) \quad \text{or} \quad S = (S_2 * S_1) * S_0,$$

depending upon the procedure. Since V_1 , V_2 , and A are arbitrary, S is uniquely specified by (19); and hence the associative law is verified. The conditions of Lemma 1 are needed only to ensure that the star products exist.

Lemma 3. *Let p, P, q, Q, r, R be dissipative $n \times n$ matrix functions of (x, y) , with pP nonsingular, and suppose*

$$S(x, y) = \begin{bmatrix} p & Q & 0 \\ q & P & 0 \\ r & R & 1 \end{bmatrix}, \quad S(x, x) = 1.$$

If the entries of S satisfy the differential equations (14) and (17a), then

$$S(x, z) = S(x, y) * S(y, z).$$

In other words, the six functional equations (10), (11) hold. It was shown in [5] and [6] that (10a), (10b), (11a), and (11b) follow from the differential system with the stated initial conditions, and hence we need establish (10c) and (11c) only.

By (10a) we can write (10c) as

$$r(x, z) = [R(x, y)q(y, z) + r(y, z)]p^{-1}(y, z)p(x, z) + r(x, y).$$

Because $S(x, x) = 1$ the desired result follows if it is shown that the right side of the foregoing equation is independent of y . The derivative with respect to y is computed by use of the differential system. The resulting expression involves, at first, $K, \Lambda, M, k, \lambda$, and μ . However, the coefficients of the first four functions reduce identically to 0, and the coefficients of λ and μ are 0 by virtue of (10a). Thus (10c) is verified, and (11c) follows by symmetry, since (15) and (17b) hold [5]. This completes the proof. We remark that the result gives, in particular,

$$S(x, y) * S(y, x) = 1.$$

7. The main theorem. The foregoing considerations lead to the following:

Theorem 1. *In a given interval $y \in I$, $(x, y) \in I \times I$ let the entries of the matrix*

$$S(x, y) = \begin{bmatrix} p & Q & 0 \\ q & P & 0 \\ r & R & 1 \end{bmatrix}$$

be $n \times n$ matrix functions of (x, y) which satisfy the differential system

$$\begin{aligned} P_y &= P(K + \lambda Q), & p_y &= (k + Q\lambda)p, \\ Q_y &= \Lambda + kQ + QK + Q\lambda Q, & q_y &= P\lambda p, \\ R_y &= M + \mu Q + RK + R\lambda Q, & r_y &= (R\lambda + \mu)p, \end{aligned}$$

together with the initial conditions $S(x, x) = 1$. Suppose that the $n \times n$ matrix coefficients

$$\lambda(y), \quad k(y), \quad \mu(y), \quad \Lambda(y), \quad K(y), \quad M(y)$$

are piecewise continuous. Then the matrix $S(x, y)$ is dissipative for $x \in I$, $y \in I$, $y > x$ if and only if the matrix

$$C(y) = \begin{bmatrix} k & \Lambda & 0 \\ \lambda & K & 0 \\ \mu & M & 0 \end{bmatrix}$$

is such that its off-diagonal elements are nonnegative and the sum of the elements in each column is nonpositive, for $y \in I$.

The differential equation and initial conditions give

$$S(x, y) = 1 + C(x)\Delta y + o(\Delta y)$$

where $\Delta y = y - x$. The stated condition on C follows readily from the dissipative character of $S(x, y)$, since

$$C(x) = \lim_{y \rightarrow x} \frac{S(x, y) - S(x, x)}{y - x}.$$

To obtain the converse consider a closed interval J of continuity of $C(y)$. We define

$$C_\epsilon(y) = G_\epsilon(y) + 2\epsilon D - (3n + 1)\epsilon E, \quad \epsilon > 0,$$

where D is the $3n \times 3n$ matrix whose elements are all 1 and E is the $3n \times 3n$ identity matrix (denoted elsewhere by 1 in this paper). The matrix $G_\epsilon(y)$ has continuously differentiable elements which approximate those of $C(y)$ within ϵ . Then C_ϵ satisfies the conditions stated for C and the inequalities are strict.

Indeed, the off-diagonal entries are $\geq \epsilon$ and the sum of the entries in any column is $\leq -\epsilon$.

If $S_\epsilon(x, y)$ is the scattering matrix which arises when the coefficient matrix C is replaced by C_ϵ , the relation

$$S_\epsilon(x, y) = 1 + C_\epsilon(x)\Delta y + o(\Delta y)$$

shows that S_ϵ is dissipative for Δy small and positive. Also the initial conditions show that $p_\epsilon(x, y)$ and $P_\epsilon(x, y)$ are nonsingular, at each fixed x , provided $|y - x|$ is sufficiently small. In the closed subinterval $J = [a, b] \subset I$ the uniform boundedness of $\partial^2 S_\epsilon(x, y)/\partial y^2$ in $J \times J$ enables us to assign a positive lower bound to the smallness of Δy necessary to achieve the aforesaid results. Thus if $a \leq x < y \leq b$, the interval $[x, y]$ can be divided by a finite number of points

$$(21) \quad x = x_1 < x_2 < x_3 \cdots < x_m = y$$

such that $S_\epsilon(x_k, x_{k+1})$ is dissipative and $p_\epsilon(x_k, x_{k+1}) P_\epsilon(x_k, x_{k+1})$ is nonsingular. Since $S_\epsilon(x, y)$ can be computed by repeated use of Lemma 3, we deduce by Lemma 1 that $S_\epsilon(x, y)$ is dissipative. A similar argument enables us to pass from the subintervals of I in which $C_\epsilon(y)$ is continuous to the whole interval I . Finally, since

$$S(x, y) = \lim_{\epsilon \rightarrow 0} S_\epsilon(x, y)$$

we get the theorem.

The reason for introducing G_ϵ is to have uniform boundedness of the remainder in the Taylor's series for $S_\epsilon(x, y)$. The Heine-Borel theorem does not apply directly, because we only know that each x is the left-hand end-point of a set of half-open covering intervals. And the exceptional set that arises when we apply Sierpinski's lemma causes trouble in the proof.

Nevertheless, we can take $G_\epsilon = C$ if the argument is modified as follows. Each x is the center of an open interval in which the sign of the off-diagonal entries and of the column sums for

$$S_\epsilon(x, y) - 1$$

agree with those for $C_\epsilon(x)\Delta y$. Also, $P_\epsilon p_\epsilon$ is nonsingular in some such interval. By the Heine-Borel theorem we select a finite set $\{x_k\}$ whose intervals give a covering. Let ξ_k , with $x_{k-1} < \xi_k < x_k$, be in the common part of the intervals for x_{k-1} and for x_k . By the functional equations

$$S_\epsilon(\xi_k, x_k) * S_\epsilon(x_k, \xi_k) = 1.$$

Thus we see that $S_\epsilon(\xi_k, x_k)$ is dissipative and has nonsingular pP . Since

$$S_\epsilon(x_{k-1}, x_k) = S_\epsilon(x_{k-1}, \xi_k) * S_\epsilon(\xi_k, x_k),$$

we can complete the argument as before.

Corollary 1. *Under the hypothesis of the theorem the matrices $p(x, y)$ and $P(x, y)$ are nonsingular.*

For proof, note that each x is interior to some open interval in which Pp is nonsingular. By the Heine-Borel theorem we can select a finite covering. Since the matrices are known to be dissipative, the proof is completed by use of Lemmas 1 and 3.

Corollary 2. *Under the hypothesis of the theorem $S(x, y)$ is a probability matrix for $x \in I, y \in I, y > x$ if and only if the off-diagonal elements of $C(y)$ are non-negative and the sum of the elements in each column is zero.*

Since the theorem shows that the entries in $S(x, y)$ are nonnegative, all we have to establish is that the equalities of (18) hold. It was stated in [2] that this follows from the differential system, but for completeness we give the proof here. With

$$w = p + q + r, \quad W = P + Q + R,$$

the equations yield

$$\begin{aligned} w_y &= (W\lambda + \mu + k)p, \\ W_y &= WK + \Lambda + M + (W\lambda + k + \mu)Q. \end{aligned}$$

The elements of the transposed matrices are denoted by subscripts as elsewhere in this paper; thus

$$w' = (w_{ij}), \quad \lambda' = (\lambda_{ij}), \quad \mu' = (\mu_{ij}), \quad k' = (k_{ij}),$$

and similarly for capital letters. Using the summation convention we display the matrix equations in full:

$$\begin{aligned} \frac{\partial w_{ij}}{\partial y} &= p_{ik}(\lambda_{km}W_{mi} + \mu_{ki} + k_{ki}), \\ \frac{\partial W_{ij}}{\partial y} &= K_{ik}W_{ki} + \Lambda_{ij} + M_{ij} + Q_{ik}(\lambda_{km}W_{mi} + k_{ki} + \mu_{ki}). \end{aligned}$$

The conditions we have on C agree with [2], equations (120)–(121):

$$\begin{aligned} k_{ii} &\geq 0, \quad K_{ii} \geq 0 \quad (i \neq j), \\ \lambda_{ij} &\geq 0, \quad \mu_{ij} \geq 0, \quad \Lambda_{ij} \geq 0, \quad M_{ij} \geq 0, \\ \sum_i (k_{ij} + \lambda_{ij} + \mu_{ij}) &= 0, \quad \sum_i (K_{ij} + \Lambda_{ij} + M_{ij}) = 0. \end{aligned}$$

Hence summation of the differential equations on j yields

$$\begin{aligned} \frac{\partial w_i}{\partial y} &= p_{ik}(\lambda_{km}W_m - \lambda_{km}\delta_m), \\ \frac{\partial W_i}{\partial y} &= K_{ik}W_k - K_{ik}\delta_k + Q_{ik}(\lambda_{km}W_m - \lambda_{km}\delta_m) \end{aligned}$$

where

$$w_i = \sum_j w_{ij}, \quad W_i = \sum_j W_{ij}, \quad \delta_i = 1.$$

Now $w_i = W_i = 1$ at $y = x$, and $w_i = W_i \equiv 1$ satisfies the differential equations. Since the coefficients are bounded, the solution is unique; and this completes the proof.

BIBLIOGRAPHY

- [1] J. MYCIELSKI & S. PASZKOWSKI, Sur un problème du calcul de probabilité (I), *Studia Math.* **15** (1956), 188–200.
- [2] S. PASZKOWSKI, Sur un problème du calcul de probabilité (II), *Studia Math.* **15** (1956), 273–298.
- [3] R. M. REDHEFFER, Novel uses of functional equations, *J. Rat. Mech. and Anal.* **3** (1954), 271–279.
- [4] R. M. REDHEFFER, The Riccati Equation: Initial values and inequalities, *Math. Annalen* **133** (1957), 235–250.
- [5] W. T. REID, Solutions of a Riccati matrix differential equation as functions of initial values, *J. Math. Mech.* **8** (1959), 221–230.
- [6] R. M. REDHEFFER, Inequalities for a matrix Riccati equation, *J. Math. Mech.* **8** (1959), 349–367.
- [7] R. M. REDHEFFER, On solutions of Riccati's equation as functions of the initial values, *J. Rat. Mech. and Anal.* **5** (1956), 835–848.
- [8] R. BELLMAN & R. KALABA, Invariant Imbedding, Random Walk and Scattering, *J. Math. Mech.* **8** (1959).

University of California
Los Angeles, California