ALGEBRAIC METHODS FOR THE SOLUTION OF SOME LINEAR MATRIX EQUATIONS*+

by

T. E. Djaferis and S.K. Mitter
Department of Electrical Engineering and Computer Science
and
Laboratory for Information and Decision Systems
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

*The research of the first author has been supported by ERDA under Grant ERDA-E(49-18)-2087. The research of the second author has been supported by ERDA under Grant ERDA-E(49-18)-2087 and by NASA-NGL-22-009-124. The computational work was done using the computer systems MACSYMA developed by the Math Lab group at M.I.T. The Math Lab group was supported by the Defence Advanced Research Projects Agency, work order 2095, under Office of Naval Research Contract No. N00014-75-C-0661.
+This paper is a revised version of a paper presented at the 1978 IFAC International Congress, held at Helsinki, Finland.

1. Introduction

Linear matrix equations play a very important role in system theory. In this paper we undertake the study of linear matrix equations which take the form

$$\sum_{i=0}^{s} \sum_{j=0}^{t} g_{ij} B^i P A^j = Q \qquad (1.1)$$

where B(mxm), A(nxn) and Q(mxn) are given matrices over some field F and g_{ij} are elements of F, using methods of modern algebra. The emphasis is on the solution of such equations using finite algebraic procedures which are easily implemented on a digital computer.

Particular attention is given to equations PA + BP = Q and P - BPA = Q, and their special subcases PA + A'P = Q (the Lyapunov equation) and P - A'PA = Q the (discrete Lyapunov equation). The Lyapunov equation appears in several areas of control theory such as stability theory, optimal control (evaluation of quadratic integrals), stochastic control (evaluation of covariance matrices) and in the solution of the Algebraic Riccati Equation using Newton's Method.

This paper has been inspired by an important paper by Kalman [2]. Kalman's concern was the characterization of polynomials whose zeros lie in certain algebraic domains (and the unification of the ideas of Hermite and Lyapunov). In this paper we show that the same ideas lead to finite algorithms for the solution of linear matrix equations of the form given above. The Analysis in terms of a module structure on matrices presented here is believed to be new. In a subsequent paper we shall investigate the implications of these ideas on stability theory.

The paper is divided into five sections. In section 2 we define the action f_{BA} over an arbitrary commutative ring with identify and prove a Basic

Lemma. In section 3 we consider equation (1.1) over a field F in great generality and prove the Main Theorem. In section 4 we deal with the equation PA + BP = Q and the Lyapunov Equation PA + A'P = Q for which we give algorithms for obtaining its solution and comment on the arithmetic complexity. We also provide numerical examples and prove a stability theorem. In section 5 we deal with the equation P - BPA = Q as well as with the Discrete Lyapunov Equation P - A'PA = Q. In section 6 we look at equation (1.1) over an integral domain.

2. The Action f_{BA}

Let A be an nxn matrix and B an mxm matrix both over E, a commutative ring with identity. Let E[x,y] be the ring of polynomials in two indeterminates x and y over E. Let $\Psi = (\phi_2(x), \psi_2(y))$ be the ideal in E[x,y] generated by $\phi_2(x)$ the characteristic polynomial of A, and $\psi_2(y)$ the characteristic polynomial of B. Elements of the quotient ring $E[x,y]/\Psi$ are cosets (equivalence classes) denoted by $\Psi + a(x,y)$. The Cayley Hamilton Theorem holds [4] therefore $\phi_2(A) = 0, \psi_2(B) = 0$.

Since $\phi_2(x)$ and $\psi_2(y)$ are monic polynomicals division is possible and as a consequence we can state;

Lemma 1: Let $g(x,y) \in E[x,y]$. Then g(x,y) can be written uniquely as:

$$g(x,y) = t(x,y)\,\phi_2(x)\,\psi_2(y) + p(x,y)\phi_2(x) + q(x,y)\psi_2(y) + r(x,y)$$

where:

the degree of p(x,y) in y is less than m (it may be a polynomial in x) or p(x,y) is zero,

the degree of $q(x,y)$ in x is less than n (it may be a polynomial in y) or $q(x,y)$ is zero, (2.1)

the degree of $r(x,y)$ in y is less than m, in x less than n or $r(x,y)$ is zero.

Proof:

Division in x by $\phi_2(x)$ is possible therefore

$$g(x,y) = a(x,y)\phi_2(x) + b(x,y)$$

where

degree of $b(x,y)$ in x is less than n ($b(x,y)$ may be a polynomial in y) or $b(x,y)$ is zero.

Division in y by $\psi_2(y)$ is possible therefore

$$a(x,y) = t(x,y)\psi_2(y) + p(x,y)$$

where degree of $p(x,y)$ in y is less than m ($p(x,y)$ may be a polynomial in x) or $p(x,y)$ is zero.

Also

$$b(x,y) = q(x,y)\psi_2(y) + r(x,y)$$

where degree of $r(x,y)$ in y is less than m and degree of $r(x,y)$ in x is less than n or $r(x,y)$ is zero.

Now then

$$g(x,y) = t(x,y)\phi_2(x)\psi_2(y) + p(x,y)\phi_2(x) + q(x,y)\psi_2(y) + r(x,y)$$

This representation is unique since suppose

$$g(x,y) = t_1(x,y)\phi_2(x)\psi_2(y) + p_1(x,y)\phi_2(x) + q_1(x,y)\psi_2(y) + r_1(x,y)$$

$$= t_2(x,y)\phi_2(x)\psi_2(y) + p_2(x,y)\phi_2(x) + q_2(x,y)\psi_2(y) + r_2(x,y)$$

with p_1, p_2, q_1, q_2, r_1, r_2 satisfying requirements (2.1).

$$r_1(x,y) - r_2(x,y) = \underset{\alpha}{[t_1-t_2]\phi_2(x)\psi_2(y)} + \underset{\beta}{[p_1-p_2]\phi_2(x)} + \underset{\gamma}{[q_1-q_2]\psi_2(y)}$$

Suppose that $\alpha \neq 0$. Then there exists a term on the r.h.s. say $a\, x^i\, y^j$ $i \geq n$, $j \geq m$. This term cannot be cancelled by either β or γ. Therefore, $\alpha = 0$. Suppose that $\beta \neq 0$. Then there exists a term on the r.h.s. say $b\, x^i\, y^j$ $i \geq n$. This term cannot be cancelled by any term from γ. Therefore $\beta = 0$. But then $\delta = 0$ as well and $r_1(x,y) = r_2(x,y)$. ■

Corollary 1: Let $g_1 = t_1\phi_2(x)\psi_2(y) + p_1\phi_2(x) + q_1\psi_2(y) + r_1$ and $g_2 = t_2\phi_2(x)\psi_2(y) + p_2\phi_2(x) + q_2\psi_2(y) + r_2$ be in the same coset $\Psi + a(x,y)$. Then $r_1 = r_2$. (If $g = t\phi_2\psi_2 + p\phi_2 + q\psi_2 + r$ denote r by $g(x,y) \bmod \Psi$.) ▨

Let MN be the set of mxn matrices over E. Define the action f_{BA}: $E[x,y] \times MN \to MN$ in the following manner:

$$f_{BA}(h(x,y), M) = \sum_{jk} h_{jk}\, B^j\, M\, A^k$$

where $h(x,y) = \sum_{jk} h_{jk}\, y^j\, x^k$ is an element in $E[x,y]$ and M an element in MN.

It can be shown that f_{BA} has the following properties.

i) $f_{BA}(u,M) = uM$ where $u \in E$.

ii) $f_{BA}(g(x,y)+h(x,y), M) = f_{BA}(g(x,y),M) + f_{BA}(h(x,y),M)$

iii) $f_{BA}(g(x,y)h(x,y), M) = f_{BA}(g(x,y), f_{BA}(h(x,y), M))$

$= f_{BA}(h(x,y), f_{BA}(g(x,y), M))$

iv) $f_{BA}(g(x,y), M) = f_{BA}(g(x,y) \bmod \Psi, M)$

v) $f_{BA}(g(x,y), M+N) = f_{BA}(g(x,y), M) + f_{BA}(g(x,y), N)$

Properties i), ii), iii) and v) follow directly from the definition of f_{BA} [1]. Property v) is arrived at by using Lemma 1 and the Cayley-Hamilton theorem.

The definition of f_{BA} allows for the interpretation of MN as an $E[x,y]/\Psi$-module.

Basic Lemma: The set MN of mxn matrices with elements in E is a module over the quotient ring $E[x,y]/\Psi$.

Proof: The set of mxn matrices under addition is an abelian group. Define multiplication (*) of cosets $\Psi + h(x,y)$ and mxn matrices M in the following manner:

$$(\Psi+h(x,y)) * M = f_{BA}(h(x,y) \bmod \Psi, M)$$

The multiplication is well defined and satisfies the properties:

1) $(\Psi+h(x,y)) *(M+N) = (\Psi+h(x,y)) *M + (\Psi+h(x,y)) *N$

2) $(\Psi+h(x,y)) *[(\Psi+g(x,y) *M] = [(\Psi+h(x,y)) \cdot (\Psi+g(x,y))] *M$

3) $[(\Psi+h(x,y)) + (\Psi+g(x,y))] *M = (\Psi+h(x,y)) *M + (\Psi+g(x,y)) *M$

4) $(\Psi+1) *M = M$

for all M,N in MN and all $\Psi+h(x,y)$, $\Psi+g(x,y)$ in $E[x,y]/\Psi$ with $\Psi+1$ being the multiplicative identity in $E[x,y]/\Psi$.

3. The General Equation

Suppose that we restrict E to be some field F and let K be an algebraically closed extension of F. If $f(x,y)$ is an element of $F[x,y]$ we denote by V_f the variety of $f(x,y)$ in A_2^K [7].(1) Let $\lambda_1\ \lambda_2,\ldots,n$ be the eigenvalues of A and $\mu_1, \mu_2,\ldots,\mu_m$ the eigenvalues of B. Suppose that $g(x,y)$ is a polynomial in $F[x,y]$. If $g(x,y)=\sum_{jk} g_{jk}\, y^j x^k$ then we define G_g the mnxmn matrix.

$$G_g = \sum_{jk} g_{jk}\, B^j \otimes A'^k \tag{3.1}$$

where $\otimes$ denotes tensor product, $(A \otimes B = (a_{ij}\, B))$ and A' denotes transpose. The significance of the matrix G_g comes from the following.

Let $\underline{p}$ be the mnxl column vector made up of the entries of matrix $P = (p_{ij})$ written as

$$\underline{p} = [p_{11}\ p_{12}\ p_{13} \cdots p_{1n}\ p_{21}\ p_{22} \cdots p_{2n} \cdots\cdots p_{m1}\ p_{m2} \cdots p_{mn}]'.$$

Let $\underline{q}$ be the mnxl column vector made up of the entries of Q. Then equation (1.1) can simply be written as

$$G_g\ \underline{p} = \underline{q} \tag{3.2}$$

We now state the

Main Theorem: The following statements are equivalent.

1) Equation (1.1) has a unique solution for all Q.

2) G_g is invertible.

3) $g(\lambda_i,\mu_j) \neq 0 \quad \forall \mu_j,\lambda_i \quad 1 \leq i \leq n \qquad 1 \leq j \leq m.$

---

(1) $A_2^K = \{(t_1,t_2) \mid t_1,t_2 \in K\}.$

4) $V_{g(x,y)} \cap V_{\phi_2(x)} \cap V_{\psi_2(y)} = \phi$

5) The coset $\Psi+g(x,y)$ is a unit in $F[x,y]/\Psi$.

Proof:

We will show the equivalences in the order 1) $\Rightarrow$ 2) $\Rightarrow$ 3) $\Rightarrow$ 4) $\Rightarrow$ 5) $\Rightarrow$ 1)

1) $\Rightarrow$ 2

Suppose then that equation (1.1) does have a unique solution for all $\underline{Q}$. Well since equation (1.1) can equivalently be written as $G_g \underline{p} = \underline{q}$ then G_g is invertible.

2) $\Rightarrow$ 3)

From [5] theorem 43.8 we have that the λ_n characteristic values of G_g are $g(\lambda_i,\mu_j)$. Since $\det G_g = \prod_{ij} g(\lambda_i,\mu_j)$ and since $\det G_g \neq 0$ we have that $g(\lambda_i,\mu_j) \neq 0$ for all λ_i,μ_j $1 \leq i \leq n$ $1 \leq j \leq m$.

3) $\Rightarrow$ 4)

If we look at what 3) says it is the following; that the polynomials $g(x,y)$, $\phi_2(x)$ and $\psi_2(y)$ have no common roots in A_2^K. But this is statement 4).

4) $\Rightarrow$ 5)

Now $\Psi+g(x,y)$ is a unit iff there exists a $\Psi+f(x,y)$ such that

$$(\Psi+f(x,y))\cdot(\Psi+g(x,y)) = \Psi + 1 .$$

Now

$\Psi+g(x,y)$ is a unit $\iff \exists\ f(x,y), \ni \Psi + f(x,y)g(x,y) = \Psi + 1$

$\iff \exists\ f(x,y), a_1(x,y), a_2(x,y) \in F[x,y]$ such that

$$f(x,y)g(x,y) + a_1(x,y)\phi_2(x) + a_2(x,y)\psi_2(y) = 1$$

Assume now that 4) holds (i.e., the polynomial h=1 vanishes at every common zero of $g(x,y), \phi_2(x), \psi_2(y)$.) By the Hilbert-Nullstellensatz [7] there exist polynomials $f(x,y)$, $a_1(x,y)$ $a_2(x,y)$ such that

$$f(x,y)\, g(x,y) + a_1(x,y)\phi_2(x) + a_2(x,y)\psi_2(y) = 1$$

this means that $\Psi+g(x,y)$ is a unit in $F[x,y]/\Psi$.

5) $\Rightarrow$1)

Suppose that $\Psi+g(x,y)$ is a unit in $F[x,y]/\Psi$ i.e. $\exists$ $\Psi+f(x,y)$ such that $(\Psi+f(x,y))(\Psi+g(x,y)) = \Psi + 1$. Let $P = f_{BA}(f(x,y) \bmod \Psi, Q) = f_{BA}(f(x,y), Q)$. Show that this is a solution to (1.1).

$$\sum_{i=0}^{s} \sum_{j=0}^{t} g_{ij}\, B^i\, P\, A^j = f_{BA}(g(x,y), P)$$

$$= f_{BA}(g(x,y)f(x,y), Q)$$

$$= f_{BA}(1,Q) = Q$$

The P so defined is the unique solution to (1.1).

Let $P_1 \neq P_2$ be two distinct solutions to (1.1)

$$\Rightarrow f_{BA}(g(x,y), P_1) = f_{BA}(g(x,y), P_2) = Q$$

$$\Rightarrow f_{BA}(f(x,y), f_{BA}(g(x,y), P_1)) = f_{BA}(f(x,y), f_{BA}(g(x,y), P_2))$$

$$\Rightarrow \qquad P_1 = P_2$$

which is a contradiction.

Therefore equation (1.1) has a unique solution for all Q. This completes the proof of the Main Theorem. ▨

Remark 1 In the above proof we have an explicit expression for the solution of equation (1.1). A general method for constructing such an $f(x,y)$ is through a constructive proof of the Hilbert-Nullstellensatz or using Resultant Theory [6]. As will be seen in later pages of this paper for several important equations this generality is unnecessary and easier methods exist.

Remark 2 In our entire construction we have been using the ideal $\Psi = (\phi_2(x), \psi_2(y))$. Other ideals can be used. As an example the ideal $(\bar{\phi}_2(x), \bar{\psi}_2(y))$ where $\bar{\phi}_2(x)$ and $\bar{\psi}_2(y)$ are the minimal polynomicals of A and B respectively. Since $\phi_2(x) = k(x)\bar{\phi}_2(x)$ and $\psi_2(y) = \ell(y)\bar{\psi}_2(y)$ we will be dealing with polynomials of smaller degree. This may have as an effect the reduction in the number of computations performed.

Remark 3 In the special case in which A is missing from equation (1.1) (i.e., suppose it is of the form $\sum_{i=0}^{s} g_i B^i P = Q$) then it would seem that the analysis can take place in some quotient ring $F[y]/\Phi$. This actually is the case. Let Φ be the ideal in $F[y]$ generated by $\psi_2(y)$. Then $\Psi+g(y)$ is a unit in $F[x,y]/\Psi$ if and only if $\Phi+g(y)$ is a unit in $F[y]/\Phi$. This follows from the fact that if $\exists$ $f(y)$, $a_2(y)$ elements of $F[y]$ such that $f(y)g(y) + a_2(y)\psi_2(y) = 1$ then clearly there exist elements $f(x,y)(=f(y))$, $a_1(x,y)(=0)$ and $a_2(x,y)(=a_2(y)$ in $F[x,y]$ such that $f(x,y)g(y) + a_1(x,y)\phi_2(x) + a_2(x,y)\psi_2(y) = 1$. On the other hand if there exist $f(x,y)$, $a_1(x,y)$, $a_2(x,y)$ elements of $F[x,y]$ such that $f(x,y)g(y) + a_1(x,y)\,x^n + a_2(x,y)\psi_2(y) = 1$ then evaluating at $x=0$ we get $f(0,y)g(y) + a_2(0,y)\psi_2(y) = 1$ which means that $\Phi+g(y)$ is a unit in $F[y]/\Phi$. ($\phi_2(x) = x^n$ since $A = 0_n$). The action f_B : $F[y]x\,MN \to MN$ can similarly be

defined as $f_B(h(y), M) = \sum_j h_j B^j P$ and MN becomes an $F[y]/\phi$-module. The solution to $\sum_{i=0}^{s} g_i B^i P = Q$ is then given by $P = f_B(f(x) \bmod \phi, Q)$. We act similarly if B is missing.

Remark 4 Let us look at the very special case when we are dealing with equation $Bp = q$ where p and q are mx1 vectors. In this case $g(x,y) = y$. What we want to do is find $f(y)$ such that $f(y)y + a(y)\psi_2(y) = 1$. If $\psi_2(y) = y^m + k_{m-1}y^{m-1} + \ldots + k_0$ obvious choices are $f(y) = -\frac{1}{k_0} y^{m-1} - \frac{k_{m-1}}{k_0} y^{m-2} - \ldots - \frac{k_1}{k_0}$, $a(y) = \frac{1}{k_0}$ since

$$(-\frac{1}{k_0} y^{m-1} - \frac{k_{m-1}}{k_0} y^{m-2} - \ldots - \frac{k_1}{k_0})y + \frac{1}{k_0} y^m + \frac{k_{m-1}}{k_0} y^{m-1} + \ldots + \frac{k_1}{k_0} y + 1 = 1.$$

Now $k_0 \neq 0$ since for a solution to exist $\det B = k_0 \neq 0$. The solution p is given by:

$$p = f_B(f(y), q)$$

Analyzing this further we get

$$p = -\frac{1}{k_0} \sum_{j=0}^{m-1} B^j q.$$

$$= (-\frac{1}{k_0} \sum_{j=0}^{m-1} B^j) q$$

As would be expected $B^{-1} = -\frac{1}{k_0} \sum_{j=0}^{m-1} B^j$.

We will now close this section by proving two propositions which make clear the method of solution we have adopted.

Let MN be the vector space of mxn matrices over the field F. Let M_n be the vector space of mnx1 vectors over F. Then we have the obvious

vector space isomorphism f: $MN \to M_n$ defined as:

$$f: \begin{bmatrix} p_{11} \\ p_{12} \\ \vdots \\ p_{1n} \\ p_{21} \\ p_{22} \\ \vdots \\ p_{2n} \\ \vdots \\ p_{m1} \\ \vdots \\ p_{mn} \end{bmatrix} \longmapsto \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ & \vdots & & \\ p_{m1} & p_{m2} & \cdots & p_{mn} \end{bmatrix}.$$

Let G_g be as in (3.1). Let the polynomial $\pi(u)$ in $F[u]$ be the characteristic polynomial of G_g, $\pi(u) = \det(u\, I_{mn} - G_g)$. Let $\Pi = (\pi(u))$ be the principal ideal in $F[u]$. Then $F[u]/\Pi$ is a ring.

Define the function h: $F[u]/\Pi \longrightarrow F[x,y]/\Psi$ in the following manner:

$$h: \Pi + a(u) \mapsto \Psi + a(g(x,y))$$

Proposition 3.1. The function h is a ring homomorphism.

Proof:

We first show that h is well-defined. Let $\Pi + a(u) = \Pi + b(u)$ (i.e., $a(u)-b(u) = k(u)\ \pi(u)$). Show that $\Psi + a(g(x,y)) = \Psi + b(g(x,y))$ i.e., show that $a(g(x,y)) - b(g(x,y)) = c_1(x,y)\phi_2(x) + c_2(x,y)\psi_2(y)$. I claim that $\pi(g(x,y)) = k(x,y)\phi_2(x)\psi_2(y) + k_1(x,y)\phi_2(x) + k_2(x,y)\psi_2(y)$.

Suppose that we are working over $K[x,y]$. We have $\pi(u) = (u-v_{11})(u-v_{12})\cdots(u-v_{mn})$ where $v_{11}, v_{12}, \ldots, v_{mn}$ are the mn eignevalues of G_g. Let

$v_{ij} = g(\lambda_i, \mu_j)$. Then

$$\Pi(g(x,y)) = \prod_{i=1}^{n} \prod_{j=1}^{m} (g(x,y) - g(\lambda_i, \mu_j)) \qquad (3.3)$$

Now we can show that each factor $g(x,y) - g(\lambda_i, \mu_j)$ can be written in the form

$$g(x,y) - g(\lambda_i \mu_j) = k_{ij}(x,y)(x-\lambda_i) + \ell_{ij}(y)(y-\mu_j)$$

This can be seen easily from the fact that if $g(x,y) = g_t x^t + g_{t-1} x^{t-1} + \ldots g_1 x + g_0$ then

$$g(x,y) - g(\lambda_i, \mu_j) = [g_t x^{t-1} + (g_{t-1} + g_t \lambda_i) x^{t-2} + (g_{t-2} + g_{t-1}\lambda_i + g_t \lambda_i^2) x^{t-3} + \ldots$$

$$+ (g_1 + g_2 \lambda_2 + \ldots g_t \lambda_i^{t-1})](x-\lambda_i) + g(\lambda_i, y) - g(\lambda_i, \mu_j)$$

Therefore (3.3) can be written as

$$\Pi(g(x,y)) = \prod_{i=1}^{n} \prod_{j=1}^{m} (k_{ij}(x,y)(x-\lambda_i) + \ell_{ij}(y)(y-\mu_j))$$

In expanding this product we see that every term in the sum will be of either of the two forms, $a(x,y)\phi_2(x)$ or $b(x,y)\psi_2(y)$. Therefore $\pi(g(x,y)) = t_1(x,y)\phi_2(x)\psi_2(y) + p_1(x,y)\phi_2(x) + q_2(x,y)\psi_2(y)$ in form (2.1) over $K[x,y]$. Since $F[x,y] \subset K[x,y]$ and form (2.1) is unique we must have that $t_1(x,y)$, $p_1(x,y)$ $q_1(x,y)$ are actually elements of $F[x,y]$.

Therefore $\pi(g(x,y) \in \Psi$ and h is well defined. Now h is a ring homomorphism since

$$h[\Pi + a(u) + \Pi + b(u)] = h[\Pi + (a(u) + b(u))]$$

$$= h(\Pi + a(u)) + h(\Pi + b(u))$$

and

$$h[(\Pi + a(u))(\Pi + b(u))] = h(\Pi + a(u)b(u))$$

$$= h(\Pi + a(u))h(\Pi + b(u))$$

and $h(\Pi + 1) = \Psi + 1$.

This completes the proof of Proposition 3.1. ▨

Now since MN is an $F[x,y]/\Psi$-module, M_n a $F[u]/\Pi$-module and $h:F[u]/\Pi \to F[x,y]/\Psi$ a ring homomorphism, MN can be made into an $F[u]/\Pi$-module in the natural way.

Define multiplication $(\cdot):F[u]/\Pi \times MN \to MN$ by:

$$(\Pi + a(u))\cdot P = h(\Pi + a(u))*P$$

We now have

Proposition 3.2. The map f is an $F[u]/\Pi$-module isomorphism.

Proof:

We already know that f is a vector space isomorphism. In order to show that f is an $F[u]/\Pi$-module isomorphism we just need to show that

$$f((\Pi + a(u))*\underline{p}) = (\Pi + a(u))\cdot f(\underline{p}) = (\Pi + a(u))\cdot P = h(\Pi + a(u))*P$$

Let us show that

$$f((\Pi + u)*\underline{p}) = h(\Pi + u)*P.$$

Well

$$f((\Pi + u)*p) = f(G_g \cdot \underline{p})$$
$$= \sum_{jk} g_{jk} B^j \cdot P \cdot A^k$$
$$= (\Psi + g(x,y))*P$$
$$= h(\Pi + u)*P$$

Now it is clear by induction that

$$f((\Pi + u^i)*\underline{p}) = h((\Pi + u^i))*P$$

Therefore

$$f((\Pi + a(u))*\underline{p}) = h(\Pi + a(u))*P$$

This completes the proof of Proposition 3.2. ▨

The equation we are considering takes the two forms given in (3.2) and (1.1). We know that in order to obtain the unique solution to (3.2) we have to invert the matrix G_g or equivalently find the inverse of $\Pi + u$ in $F[u]/\Pi$. The above Propositions show this is the same as obtaining the inverse of $\Psi + g(x,y)$ in $F[x,y]/\Psi$ while working with form (1.1) of the equation.

In the following two sections we will be concerned with the problem of constructing the solution to several special cases of the general equation. It is of course assumed that a unique solution does exist. We also prove a stability theorem associated with the Lyapunov equation.

4. The Equation PA + BP = Q

As shown when proving the Main Theorem the solution to equation PA + BP = Q is given by

$$P = f_{BA}(f(x,y) \bmod \Psi, Q)$$

where $f(x,y) \in F[x,y]$, $(\Psi + f(x,y))(\Psi + (x + y)) = \Psi + 1$.

It has also been mentioned that such an $f(x,y)$ can be found by using Resultant Theory [6] or from a constructive proof of the Hilbert-Nullstellensatz. But in simple cases like this we need not resort to such general theory.

In carrying out computations, it may be advantageous instead of finding $f(x,y)$ such that $f(x,y)g(x,y) = k_1\phi_2(x) + k_2\psi_2(y) + 1$ to find $f_u(x,y)$ such that $f_u(x,y)g(x,y) = k_1\phi_2(x) + k_2\psi_2(y) + u$ where u is any non-zero element in F. The solution P is then given by $P = (1/u) \cdot f_{BA}(f_u(x,y) \bmod \Psi, Q)$.

We construct $f_u(x,y)$ in this manner.

We do have that

$$x + y \mid \phi_2(x)\psi_2(y) - \phi_1(y)\psi_1(x)$$

where

$$\phi_1(x) = \phi_2(-x), \quad \psi_1(x) = \psi_2(-x).$$

Let

$$p(x,y) = \frac{\phi_2(x)\psi_2(y) - \phi_1(y)\psi_1(x)}{x+y} \tag{4.1}$$

Since $\phi_2(x), \psi_1(x)$ are coprime ($\lambda_i + \mu_j \neq 0$ for all i,j) we have $\lambda_e(x)$, $\mu_e(x)$, $\lambda'_e(x)$, $\mu'_e(x)$

$$\begin{aligned} \lambda_e(x)\psi_1(x) + \mu_e(x)\phi_2(x) &= e \\ \lambda'_e(x)\psi_2(x) + \mu'_e(x)\phi_1(x) &= e \end{aligned} \tag{4.2}$$

Let $f_u(x,y) = \lambda_e(x)\mu_e'(y)p(x,y)$, since

$$f_u(x,y)\ (x+y) = \lambda_e(x)\mu_e'(y)p(x,y)(x+y)$$

$$= \lambda_e(x)\mu_e'(y)\phi_2(x)\psi_2(y) + e\lambda_e'(y)\psi_2(y)$$

$$+ e\mu_e(x)\phi_2(x) - \mu_e(x)\lambda_e'(y)\phi_2(x)\psi_2(y) - e^2$$

With $u = -e^2$ we have $(\Psi + f_u(x,y))(\Psi + (x+y)) = \Psi + u$.

A different method for obtaining an $\bar{f}(x,y)$ such that $\bar{f}(x,y)(x+y) = k_1\phi_2(x) + k_2\psi_2(y) + 1$ is the following:

Divide $\phi_2(x)$ by $x+y$ in x.

$$\phi_2(x) = h(x,y)(x+y) + h(y).$$

For $x = -y$ we have $\phi_2(-y) = \phi_1(y) = h(y)$. Now since $\phi_1(y), \psi_2(y)$ are relatively prime there exist $\lambda(y), \mu(y)$ such that

$$\lambda(y)\phi_1(y) + \mu(y)\psi_2(y) = 1$$

$$\Longrightarrow \quad \lambda(y)[\phi_2(x) - h(x,y)(x+y)] + \mu(y)\psi_2(y) = 1$$

$$\Longrightarrow \quad -\lambda(y)h(x,y)(x+y) + \mu(y)\phi_2(x) + \mu(y)\psi_2(y) = 1$$

Let $\bar{f}(x,y) = -\mu(y)h(x,y)$.

<u>The Lyapunov Equation PA + A'P = Q</u>

The Lyapunov equation is a special case of PA + BP = Q, B = A', A is stable. With the appropriate modifications to the first procedure for constructing the solution we have:

Let $\tau_e(x), \mu_e(x)$ be such that

$$\tau_e(x)\phi_1(x) + \mu_e(x)\phi_2(x) = e \qquad e \neq 0$$

$$p(x,y) = \frac{\phi_2(x)\phi_2(y) - \phi_1(x)\phi_1(y)}{x+y} \tag{4.3}$$

$$f_u(x,y) = \tau_e(x)\tau_e(y)p(x,y)$$

Algorithm for solving the Lyapunov equation $A'P + PA = Q$.

R_1) Obtain $\phi_2(x)$ the characteristic polynomial of A.

R_2) Set $p(x,y) = \dfrac{\phi_2(x)\phi_2(y) - \phi_1(y)\phi_1(x)}{x+y}$.

R_3) Using the Extended Euclidean algorithm or an equivalent method obtain $\tau_e(x)$ and e.

R_4) Form $f_u(x,y) = \tau_e(x)\tau_e(y)P(x,y)$.

R_5) Find $P_u = f_{BA}(f_u(x,y) \bmod \Psi, Q)$.

R_6) Set $P = \frac{1}{u} P_u$, $u = -e^2$.

Computer Implementation

Since we are interested in an exact computer solution we restrict the field of interest to that of the rational numbers R. The algorithm is fully implementable, using the remarkable facilities provided by the computer programming system MACSYMA available at M.I.T. MACSYMA is a large computer programming system used for performing symbolic as well as numerical computations.

Three versions of the algorithm have been constructed and programmed on MACSYMA. They are the Rational Algorithm, the Integer Algorithm, and the Modular Algorithm having names indicative of the mode in which arithmetic operations are carried out.

The Rational Algorithm

It consists of carrying out steps R_1 through R_6 in rational arithmetic.

The Integer Algorithm

Suppose that matrices A and Q contained integer cntries. The polynomials $\phi_2(x)$, $p(x,y)$ then have integer coefficients.

Let $\phi_2(x) = a_n x^n + a_{n-1} x^{n-1} = \ldots + a_0$

$\phi_1(x) = d_n x^n + d_{n-1} x^{n-1} + \ldots + d_0$

Define the 2nx2n matrix S

$$S = \begin{bmatrix} a_n & 0 & & 0 & d_n & 0 & & 0 \\ a_{n-1} & a_n & & 0 & d_{n-1} & d_{n-1} & & 0 \\ \vdots & a_{n-1} & & \vdots & & & & \\ & & \cdots & a_n & d_1 & d_2 & & d_n \\ a_0 & a_1 & & a_{n-1} & d_0 & d_1 & \cdots & d_{n-1} \\ 0 & a_0 & & a_{n-2} & 0 & d_0 & & \\ 0 & 0 & & \vdots & 0 & 0 & & \vdots \\ \vdots & \vdots & & & \vdots & \vdots & & \\ 0 & 0 & & a_0 & 0 & 0 & & d_0 \end{bmatrix}$$

We know that detS $\neq$ 0 since it is the resultant of $\phi_1(x)$ and $\phi_2(x)$ which are coprime. If we let e = detS the linear system

$$S \begin{bmatrix} \lambda_{n-1} \\ \lambda_{n-2} \\ \\ \lambda_0 \\ \tau_{n-1} \\ \tau_{n-2} \\ \\ \tau_0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \\ \\ \\ \\ \\ e \end{bmatrix}$$

has an integer solution and we have integer polynomials

$\tau_e(x) = \tau_{n-1}x^{n-1} + \ldots \tau_0$, $\lambda_e(x) = \lambda_{n-1}x^{n-1} + \ldots + \lambda_0$ which satisfy

$$\tau_e(x)\phi_1(x) + \lambda_e(x)\phi_2(x) = e$$

This means that $f_u(x,y)$ in (4.3) has integer coefficients and so does $f_u(x,y) \bmod \Psi$, which implies that $P_u = f_{BA}(f_u(x,y) \bmod \Psi, Q)$ has integer entries.

The algorithm proceeds as follows.

I_1) Find $\phi_2(x)$ the characteristic polynomial of A.

I_2) Set $p(x,y) = \dfrac{\phi_2(x)\phi_2(y) - \phi_1(x)\phi_1(y)}{x+y}$.

I_3) Find $\tau_e(x)$ and e.

I_4) Form $f_u(x,y) = \tau_e(x)\tau_e(y)p(x,y)$.

I_5) Find $P_u = f_{BA}(f_u(x,y) \bmod \Psi, Q)$.

I_6) Set $P = \frac{1}{u} \cdot P_u$, $u = -e^2$.

The Modular Algorithm

The integer algorithm paves the way for a modular approach to the solution.

Suppose p is a prime that does not divide $e = \det S$. If $A = (a_{ij})$ and $Q = (q_{ij})$ are matrices with integer entries let ${}_pQ = (a_{ij} \bmod p)$ and ${}_pA = (a_{ij} \bmod p)$ be considered as matrices over Z_p. A left subscript p on a polynomial b(x,y) written as ${}_pb(x,y)$ denotes coefficient reduction modulo p. Suppose that coefficient arithmetic is done modulo p. We then have

$$ {}_p\phi_2(x) = \det(Ix - {}_pA) $$

$$ {}_pp(x,y) = \frac{{}_p\phi_2(x)\,{}_p\phi_2(y) - {}_p\phi_1(x)\,{}_p\phi_1(y)}{x+y} $$

$$ {}_p\tau_e(x)\,{}_p\phi_1(x) + {}_p\lambda_e(x)\,{}_p\phi_2(x) = {}_pe $$

$$ {}_pf_u(x,y) = {}_p\tau_e(x)\,{}_p\tau_e(y)\,{}_pp(x,y). $$

Let ${}_pP_u = f_{BA}(f_u(x,y) \bmod_p \Psi, {}_pQ)$ where all arithmetic is done modulo p and ${}_p\Psi = ({}_p\phi_2(x), {}_p\phi_2(y))$ in $Z_p[x,y]$. If ${}_pP_u$ and ${}_pu$ are obtained for a sufficient number of primes, the Chinese Remainder Theorem can be used to find P_u and u making it possible to obtain the solution $P = \frac{1}{u} \cdot P_u$.

The algorithm is as follows:

M_1) Obtain ${}_pA$, ${}_pQ$.

M_2) Obtain ${}_p\phi_2(x) = \det(Ix - {}_pA)$.

M_3) Set ${}_pP(x,y) = \dfrac{{}_p\phi_2(x)\,{}_p\phi_2(y) - {}_p\phi_1(x)\,{}_p\phi_1(y)}{x+y}$.

M_4) Obtain ${}_p\tau_e(x)$, ${}_pe$.

M_5) Set ${}_pf_u(x,y) = {}_p\tau_e(x)\,{}_p\tau_e(y)\,{}_pP(x,y)$.

M_6) Obtain ${}_pP_u = f_{BA}({}_pf_u(x,y) \bmod {}_p\Psi, {}_pQ)$.

M_7) Repeat steps 1-6 for a sufficient number of primes and using the Chinese Remainder Theorem find P_u and $u = -e^2$.

M_8) Set $P = \frac{1}{u} \cdot P_u$.

Since considerable coefficient growth takes place in intermediate computations of the Integer Algorithm, a lot of storage is being used up. In such cases it is advantageous to use the Modular Algorithm.

Arithmetic Complexity of the Integer Algorithm

We are concerned with the number of integer operations (addition, subtraction, multiplication, division) involved in running the Integer Algorithm when A and Q are nxn matrices, using classical operations.

Step I_1: There are several methods for obtaining the Characteristic polynomial $\phi_2(x)$ of a stable matrix. Evaluating $\phi_2(x)$ at n distinct

points and then solving for the coefficients requires $0(n^4)$ operations. If n is small (say $n \leq 20$), evaluating $\phi_2(x)$ at $x = 1$ where $\phi_2(1) = \Lambda$, $\lambda = [\log_{10}\Lambda]$ and then at $x = 10^{\lambda}$ allows one to "read off" the coefficients of $\phi_2(x)$ from a large integer. This procedure requires only $0(n^3)$ operations.

Step I_2: This step can be done in $0(n^2)$ operations.

Step I_3: Solving a linear set of 2n equations simultaneously is and $0(n^3)$ operation.

Step I_4: Performing the multiplication as $\tau_e(x)[\tau_e(y) \cdot P(x,y)]$ requires $0(n^3)$ operations.

Step I_5: Obtaining $f_u(x,y) \bmod \Psi$ involves two polynomial divisions can be done in $0(n^3)$ operations. To form $f_{BA}(f_u(x,y) \bmod \Psi, Q)$ we use $0(n^4)$ operations. In the event that the matrix Q is a product of vectors $Q = c \cdot c'$ this calculation can be done in $0(n^3)$ operations.

Step I_6: It can be done in $0(n^2)$ operations.

It can therefore be seen that the overall calculation requires $0(n^4)$ operations in general and $0(n^3)$ operations in the special cases mentioned.

Storage requirements are much harder to determine since the implementation is on a variable length word computer.

Numerical Examples

We now continue this section by giving two numerical examples.

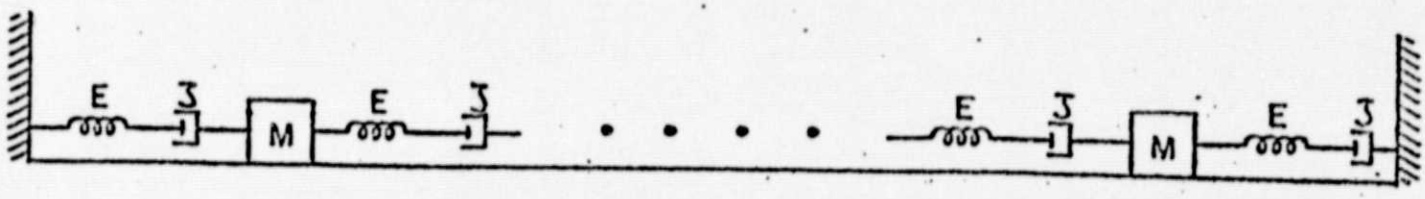
We wish to compute

$$G = \int_0^\infty x'(t) \cdot Q \cdot x(t)dt$$

where $x(t)$ is a solution to

$$\dot{x}(t) = Ax(t) \qquad x(0) = c \qquad (*)$$

The system modelled by (*) is of the form



where the number of blocks is finite.

Example 1: The number of blocks is 5 with ζ=1, K=1, M=10000. Listed are the corresponding A matrix, the Q matrix and the solution P to the equation $PA + A'P = Q$.

Example 2: The number of blocks is 2. Listed are the corresponding A matrix in parametric form, the Q matrix and the parametric solution P of the equation $PA + A'P = Q$. The parametric solution P is valid only for appropriate values of E, M, Z, ($Z=\zeta$).

Example 1:

$$A=\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -\frac{1}{5000} & -\frac{1}{5000} & \frac{1}{10000} & \frac{1}{10000} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{10000} & \frac{1}{10000} & -\frac{1}{5000} & -\frac{1}{5000} & \frac{1}{10000} & \frac{1}{10000} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{10000} & \frac{1}{10000} & -\frac{1}{5000} & -\frac{1}{5000} & \frac{1}{10000} & \frac{1}{10000} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{10000} & \frac{1}{10000} & -\frac{1}{5000} & -\frac{1}{5000} & \frac{1}{10000} & \frac{1}{10000} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{10000} & \frac{1}{10000} & -\frac{1}{5000} & -\frac{1}{5000} \end{bmatrix}$$

$$Q=\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$P = \begin{bmatrix} -\frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\frac{12500}{3} & 0 & -\frac{10000}{3} & 0 & -2500 & 0 & -\frac{5000}{3} & 0 & -\frac{2500}{3} \\ 0 & 0 & -\frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\frac{10000}{3} & 0 & -\frac{20000}{3} & 0 & -5000 & 0 & -\frac{10000}{3} & 0 & -\frac{5000}{3} \\ 0 & 0 & 0 & 0 & -\frac{1}{2} & 0 & 0 & 0 & 0 & 0 \\ 0 & -2500 & 0 & -5000 & 0 & -7500 & 0 & -5000 & 0 & -2500 \\ 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{2} & 0 & 0 & 0 \\ 0 & -\frac{5000}{3} & 0 & -\frac{10000}{3} & 0 & -5000 & 0 & -\frac{20000}{3} & 0 & -\frac{10000}{3} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{2} & 0 \\ 0 & -\frac{2500}{3} & 0 & -\frac{5000}{3} & 0 & -2500 & 0 & -\frac{10000}{3} & 0 & -\frac{12500}{3} \end{bmatrix}$$

Example 2.

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -\frac{2E}{M} & -\frac{2Z}{M} & \frac{E}{M} & \frac{Z}{M} \\ 0 & 0 & 0 & 1 \\ \frac{E}{M} & \frac{Z}{M} & -\frac{2E}{M} & -\frac{2Z}{M} \end{bmatrix}$$

$$Q = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$P = \begin{bmatrix} -\frac{E}{2Z} & 0 & 0 & 0 \\ 0 & -\frac{M}{3Z} & 0 & -\frac{M}{6Z} \\ 0 & 0 & -\frac{E}{2Z} & 0 \\ 0 & -\frac{M}{6Z} & 0 & -\frac{M}{3Z} \end{bmatrix}$$

In closing this section it is interesting to observe how the ideas presented here can be used to prove stability theorems constructively. In particular we prove

Theorem 4.1. Let A be an nxn matrix over the reals R. Let C be a pxn matrix. If A is a stability matrix and (A,C) an observable pair then the equation $PA + A'P = -C'C$ has a unique symmetric positive definite solution P.

Before proceeding with the proof we introduce the notion of a positive polynomial in $R[x,y]$ [2]. If $p(x,y)$ is in $R[x,y]$ we can write it as

$$p(x,y) = \ell'(y)C(p)\ell(x)$$

where $\ell(z)$ is the column vector $1,z,\ldots z^{n-1}$, n with n being one plus the largest power of $p(x,y)$ in either x or y and $C(p)$ an nxn matrix over R. This introduces a bijection between $R[x,y]$ and the set of all square matrices. We then call a polynomial $p(x,y)$ positive if $C(p)$ is i) symmetric and ii) positive definite. One can then prove [2] that $p(x,y)$ in $R[x,y]$ is positive if and only if there exist polynomials $\pi_1, \pi_2, \ldots, \pi_m$ (m the size of $C(p)$) such that

$$p(x,y) = \sum_{i=1}^{m} \pi_i(x)\pi_i(y)$$

where $\{\pi_i(x)\}$ are a basis for the vector space (over R) of polynomials of degree less than m. One can also prove [2] that the $f_u(x,y) \bmod \Psi$ given in (4.3) is positive.

Proof of Theorem 4.1

Since A is a stability matrix $\lambda_i + \lambda_j \neq 0$ for all i,j therefore a unique solution P to the equation $PA + A'P = -C'C$ exists. We also have that $f_u(x,y) \bmod \Psi$ is positive. We can therefore write

$$f_u(x,y) \bmod \Psi = \sum_{i=1}^{n} \pi_i(x)\pi_i(y)$$

We know that the unique solution is given by

$$P = \frac{1}{u} f_{BA}(f_u(x,y) \bmod \Psi, -C'C) \qquad (u = -e^2)$$

$$= \frac{1}{e^2} f_{BA}(\sum_{i=1}^{n} \pi_i(x)\pi_i(y),\ C'C)$$

$$= \frac{1}{e^2} \sum_{i=1}^{n} \pi_i(A')C'C\ \pi_i(A)$$

Since $C'C \geq 0$ we have $P \geq 0$.

Suppose now that there exists $z \neq 0$ such that $z'Pz = 0$.

$\Rightarrow \quad ||C\pi_i(A)z|| = 0 \quad$ for $\quad 1 \leq i \leq n$

$\Rightarrow \quad C\pi_i(A)z = 0 \quad$ for $\quad 1 \leq i \leq n$

Since $\{\pi_i\}$ are a basis there exists a matrix T such that

$$T \begin{bmatrix} \pi_1(x) \\ \pi_2(x) \\ \vdots \\ \pi_n(x) \end{bmatrix} = \begin{bmatrix} 1 \\ x \\ \cdot \\ \cdot \\ x^{n-1} \end{bmatrix}$$

$\Rightarrow f_{BA}(t_{i1}\pi_1(x) + \ldots + t_{in}\cdot\pi_n(x),\ C) = CA^{i-1}$ $1\leq i\leq n$ where $(t_{i1},\ldots,t_{in})$ the i^{th} row of T

$$\Rightarrow \sum_{j=1}^{n} t_{ij} C\pi(A) = CA^{i-1}$$

Define the operator $H = R^n \to R^{np}$ by:

$$H(w) = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} w$$

Since (A,C) observable pair the null space of H is {0}, and since $C\pi_i(A)z = 0$ this implies that $\sum_{j=1}^{n} t_{ij}\pi_j(A)z = 0$ for all $1\leq i\leq \eta$

$\Rightarrow$ $H(z) = 0$.

This is a contradiction since $z \neq 0$.

5. The Equation P-BPA = Q

We again wish to construct f(x,y) such that $(\Psi + f(x,y))(\Psi + (1-xy)) = \Psi + 1$. Let

$$\phi_2(x) = \det(Ix-A) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$$

$$\psi_2(x) = \det(Ix-B) = b_m x^m + b_{m-1}x^{m-1} + \ldots b_1 x + b_0$$

$$\phi_3(x) = a_0 x^n + a_1 x^{n-1} + \ldots + a_n$$

$$\psi_3(x) = b_0 x^m + b_1 x^{m-1} + \ldots + b_m$$

From the above definition we can see that the roots of $\phi_3(x)$ are the values $\frac{1}{\lambda_i}$ where $\lambda_i \neq 0$. Since we assume that a unique solution exists

we must have $1 - \lambda_i \mu_j \neq 0$ for all i,j. Then we must have that $\phi_2(x)$, $\psi_3(x)$ are coprime. Because if they have a non-trivial factor k(x) they must also have at least one common root (i.e., $\lambda_i = \frac{1}{\mu_j}$ for at least some (i,j)).

On the other hand we also have that

a) if $n \geq m$ then $1 - xy \mid y^{n-m}\phi_2(x)\psi_2(y) - \phi_3(y)\psi_3(x)$

b) if $n < m$ then $1 - xy \mid x^{m-n}\phi_2(y)\psi_2(y) - \phi_3(y)\psi_3(x)$

We are now ready to construct f(x,y).

Since $\phi_3(x)\psi_2(x)$ are relatively prime we have $\lambda(x)$ $\mu(x)$ $\lambda'(x)$ $\lambda'(x)$ such that

$$\lambda(x)\psi_3(x) + \mu(x)\phi_2(x) = 1$$

$$\lambda'(x)\psi_2(x) + \mu'(x)\phi_3(x) = 1$$

If $n \geq m$ let

$$p(x,y) = \frac{y^{n-m}\phi_2(x)\psi_2(y) - \phi_3(y)\psi_3(x)}{1 - xy}$$

if $n < m$ let

$$p(x,y) = \frac{x^{m-n}\phi_2(x)\psi_2(y) - \phi_3(y)\psi_3(x)}{1 - xy}$$

Then $f(x,y) = \mu(x)\mu'(y)\ p(x,y)$.

The discrete Lyapunov equation $P - A'PA = Q$ is a special case.

6. Over Integral Domains

Suppose now that we are investigating equation (1.1) over E some integral domain. The next proposition gives a necessary and sufficient

condition for the existence of a unique solution to (1.1) for all Q.

Proposition 6.1. Equation (1.1) has a unique solution over E, for each Q, iff $\Psi + g(x,y)$ is a unit in $E[x,y]/\Psi$.

Proof:

Let $P = f_{BA}(f(x,y) \bmod \Psi, Q)$ where $f(x,y)g(x,y) = k_1(x,y)\phi_2(x) + k_2(x,y)\psi_2(y) + 1$

$$\sum_{i=0}^{s}\sum_{j=0}^{t} g_{ij} B^i P A^j = f_{BA}(g(x,y), P)$$

$$= f_{BA}(g(x,y), f_{BA}(f(x,y), Q))$$

$$= f_{BA}(1, Q) = Q$$

The solution P is unique. This follows in the same manner as in the proof of the Main Theorem.

Suppose that equation (1.1) does have a unique solution for all Q. This means that G_g in (3.2) is invertible. We have that $\pi(u) = \det(Iu - G_g)$. From the Cayley-Hamilton theorem if $\pi(u) = \pi_t u^t + \pi_{t-1} u^{t-1} + \ldots + \pi_0$

$$\pi(G_g) = \pi_t G_g^t + \ldots + \pi_0 I = 0$$

Let $f(u) = -\frac{\pi_t}{\pi_0} u^{t-1} - \frac{\pi_{t-1}}{\pi_0} u^{t-2} - \ldots - \frac{\pi_1}{\pi_0}$. Then $f(u) \cdot u + \frac{1}{\pi_0} \cdot \pi(u) = 1$. Therefore $\Pi + u$ is a unit in $E[u]/\Pi$. The proof of Proposition 3.1 remains valid. Therefore

$$(\Pi + f(u))\ (\Pi + u) = \Pi + 1$$

$$\Rightarrow \quad h(\pi + f(u)) \cdot h(\Pi + u) = \Psi + 1$$

$$\Rightarrow \quad (\Psi + f(g(x,y)))\ (\Psi + g(x,y)) = \Psi + 1$$

which means that $\Psi + g(x,y)$ is a unit in $E[x,y]/\Psi$.

ACKNOWLEDGEMENT

The authors would like to acknowledge interesting dicussions with Professor Thomas Kailath (Stanford University) and Dr. Bernard Levy (formerly at Stanford and now at M.I.T.). In particular, some of the developments in Section 3 arose from discussions with Dr. Bernard Levy.

REFERENCES

[1] T.E. Djaferis, Exact Solution to Lyapunov's Equation Using Algebraic Methods, M.S. Thesis, January 1977.

[2] R.E. Kalman, Algebraic Characterization of Polynomials Whose Zeroes Lie in Certain Algebraic Domains, Proc. N.A.S. (Mathematics), Vol. 64, No. 3, Nov. 1969.

[3] D.E. Knuth, Seminumerical Algorithms, Addison-Wesley, Reading, Mass., 1969.

[4] S. Lang, Algebra, Addison Wesley, Reading, Mass., 1965.

[5] C.C. MacDuffee, The Theory of Matrices, Chelsea, New York, 1946.

[6] B.L. Van Der Waerden, Modern Algebra, Vol. II, Frederick Ungar, 1966.

[7] O. Zariski, P. Samuel, Commutative Algebra, Vol. I, Van Nostrand, Princeton, 1958.