Vlastimil Pták
Linear fractional transformations and companion matrices

Commentationes Mathematicae Universitatis Carolinae, Vol. 29 (1988), No. 2, 279--284

COMMENTATIONES MATHEMATICAE UNIVERSITATIS CAROLINAE
29,2 (1988)

LINEAR FRACTIONAL TRANSFORMATIONS AND COMPANION MATRICES

Vlastimil PTÁK

Abstract: With each linear fractional transformation and each positive integer n we consider a set of n polynomials of degree $\leq n-1$. Simple relations for these polynomials are established which provide a geometrical explanation of the recent results of B.A. Shane and S. Barnett on companion matrices.

The mapping which assigns to each linear fractional transformation the matrix of the coefficients of these polynomials is a representation of GL(2).

Key words: Companion matrix, matrix representation.

Classification: Primary 15A03, 15A04, 15A24, 93B25, 94C05
Secondary 14L35, 46N05, 47A05, 47A67

Introduction: In inertia theory and in investigations concerning root location for polynomials the Cayley transformation is frequently used to map the unit disc of the complex plane onto a half-plane; this provides, in particular, a connection between the theory of automata working in discrete time and that of the continuous case.

More generally, it is possible to consider the general bilinear transformation

$$\varphi : A \longrightarrow (cA+d)^{-1}\ (aA+b);$$

if A is a matrix for which cA+d is invertible then the spectrum of the transformed matrix is obtained as the image under φ of the spectrum of A. In an interesting paper [1] B.A. Shane and S. Barnett investigate the effect of φ when applied to a companion matrix; the image under φ of a companion matrix will not be a companion any more in the general case. Nevertheless, Shane and Barnett show that there exists, for each n, a matrix $M_n(\varphi)$ depending only on φ such that, for any companion matrix A for which $\varphi(A)$ exists

$$M(\varphi)\,\varphi(A)M(\varphi)^{-1}$$

is again a companion matrix. The original proof of the authors is based, however, on a highly technical computation of the individual entries. In a subsequent paper [2] N.J. Young succeeded in giving a simpler and more direct proof of the result; his proof uses an interesting characterization of companion matrices in terms of tensor products. Young also shows that the matrices $M_n(\varphi)$ may be used to construct a representation of GL(2).

It is the purpose of the present note to point out a natural interpretation of the result which provides a short and transparent proof. At the same time this proof explains the geometric meaning of the matrix $M(\varphi)$.

2. Preliminaries: Given a polynomial f of degree n written in the form

$$f(x)=x^n-(a_0+a_1x+\ldots+a_{n-1}x^{n-1})$$

we denote by C(f) its companion matrix

$$C(f)=\begin{pmatrix} 0 & 0 & \ldots & 0 & a_0 \\ 1 & 0 & \ldots & 0 & a_1 \\ 0 & 1 & \ldots & 0 & a_2 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 1 & a_{n-1} \end{pmatrix}$$

It will be convenient, for the purposes of this note, to define C(g) for a not necessarily monic polynomial g as the companion of the corresponding monic multiple of g. In the rest of this note f will be a fixed polynomial of degree n written as above. If F stands for the algebra of all polynomials with complex coefficients and if H(f) is the ideal of all multiples of f we denote by X the quotient algebra X=F/H(f). Let S be the operator of multiplication by x on F. Then H(f) is invariant with respect to S and we shall use the same letter S to denote the corresponding operator on X; clearly the polynomials

$$1, x,\ldots,x^{n-1}$$

(more precisely, the corresponding classes modulo f) form a basis of the linear space X. Since $x^n=a_0+a_1x+\ldots+a_{n-1}x^{n-1}$ modulo f the matrix C(f) appears as the matrix of S taken in this basis.

As for changes of bases, we shall use the following convention. If W is the matrix of the operator T in the basis $e_0,\ldots,e_{n-1}$, in other words $Te_j=$ $=\sum_s w_{sj}e_s$, and if a new basis is given by $p_j=\sum_r m_{rj}e_r$ then the matrix of T in the new basis is $M^{-1}WM$.

3. Fractional linear transformations: Let A be a fixed two by two matrix

$$A=\begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

We shall restrict ourselves to the case $\delta=\det A\neq 0$. Define the four linear forms u, v, U, V by the formulae

$$u(x)=ax+b \qquad v(x)=cx+d$$
$$U(x)=dx-b \qquad V(x)=-cx+a$$

and observe that

$$d\,u(x)-b\,v(x)=\delta x$$
$$-c\,u(x)+a\,v(x)=\delta.$$

For x such that $v(x)\neq 0$ define $\varphi(x)=\frac{u(x)}{v(x)}$. It follows from the two formulae above that, for $v(x)\neq 0$

$$(*)\qquad \begin{aligned} U(\varphi(x))&=\frac{\delta x}{v(x)}\\ V(\varphi(x))&=\frac{\delta}{v(x)}\end{aligned}$$

For y such that $V(y)\neq 0$ define $\psi(y)=\frac{U(y)}{V(y)}$.

We have then

1^0 If $v(x)\neq 0$ then $V(\varphi(x))=\frac{\delta}{v(x)}\neq 0$ and $\psi(\varphi(x))=x$. Similarly,

2^0 If $\delta\neq 0$ then, for every λ such that $V(\lambda)\neq 0$, the value $z=\frac{U(\lambda)}{V(\lambda)}$ satisfies $v(z)=\frac{\delta}{V(\lambda)}\neq 0$ and $\varphi(z)=\frac{u(z)}{v(z)}=\lambda$.

Define the polynomials $p_0,\dots,p_{n-1}$ of degree n-1 by the formulae

$$p_j(x)=u(x)^j\,v(x)^{n-1-j}$$

for $j=0,1,\dots,n-1$ and let us show that they are linearly independent. Suppose that

$$\xi_0p_0+\dots+\xi_{n-1}p_{n-1}=0.$$

Consider an arbitrary λ for which $V(\lambda)\neq 0$ and let $z=\frac{U(\lambda)}{V(\lambda)}$; recall that $v(z)\neq 0$ and $\varphi(z)=\lambda$. We have then

$$\Big(\sum_{j=0}^{n-1}\xi_j\lambda^j\Big)v^{n-1}(z)=\sum_{j=0}^{n-1}\xi_j\varphi(z)^jv^{n-1}(z)=\sum_{j=0}^{n-1}\xi_jp_j(z)=0.$$

Since $v(z)\neq 0$ it follows that

$$\sum_{j=0}^{n-1}\xi_j\lambda^j=0.$$

Writing down this relation for n different values of λ we conclude that $\xi_0=\dots=\xi_{n-1}=0$ so that the p_j are linearly independent.

Now consider a polynomial f of degree n

$$f(x)=x^n-(a_0+a_1x+\dots+a_{n-1}x^{n-1}).$$

It suffices to consider the case $n\geq 2$.

Suppose that $(v,f)=1$. We intend to show that there exists a polynomial h of degree $\leq n-1$ and a constant k such that

$$v(x)h(x)+f(x)k=u(x).$$

It is possible to write down h and k explicitly.

If $c=0$ then $d\neq 0$ and it suffices to set $h(x)=\frac{1}{d}\,u(x)$, $k=0$.

If $c\neq 0$ denote by y the point where $v(y)=0$. Since v and f are relatively prime, we have $f(y)\neq 0$. Set $k=\frac{u(y)}{f(y)}$ and observe that the polynomial $u(x)-kf(x)$ will be zero at y; it follows that $\frac{1}{v(x)}\,(u(x)-k\,f(x))$ is a polynomial. Denote it by h so that

$$h(x)=\frac{1}{v(x)f(y)}\,(u(x)f(y)-u(y)f(x)).$$

Since $-cu(y)=-cu(y)+a\ v(y)=\delta$ we have $u(y)=-\frac{\delta}{c}$ whence $k=-\frac{\delta}{c\ f(y)}$.

Now consider $T=C(f)$. We have $f(T)=0$ so that

$$v(T)\ h(T)=u(T).$$

Furthermore, it is easy to see that $v(T)$ is invertible. Indeed, if $0\in\sigma(v(T))$ we have $0\in\sigma(v(T))=v(\sigma(T))$ so that $v(y)=0$ for some $y\in\sigma(T)$ but this is impossible since $(v,f)=1$. It follows that $h(T)=\varphi(T)$.

Given a polynomial f of the form

$$f(x)=x^n-(a_0+a_1x+\dots+a_{n-1}x^{n-1})$$

we write g for the polynomial defined by the formula

$$g(x)=U(x)^n-\sum_{j=0}^{n-1}a_j\ U(x)^j\ V(x)^{n-j}.$$

We intend to show that g has degree n if $(v,f)=1$. Indeed, the coefficient of x^n in the polynomial g equals

$$d^n-\sum_{j=0}^{n-1}a_j\ d^j(-c)^{n-j}.$$

If $c=0$ this expression equals d^n and d is different form zero since $ad=\delta\neq 0$. If $c\neq 0$ we have $f(y)\neq 0$ since $(v,f)=1$ and $v(y)=0$. We have then

$$d^n-\sum a_jd^j(-c)^{n-j}=(-c)^n\left(\frac{d^n}{(-c)^n}-\sum a_j\frac{d^j}{(-c)^j}\right)=(-c)^n\ f(y)\neq 0.$$

We shall show now that $g(h)$ is divisible by f. The following two relations will turn out to be useful to this end. Using the relation $u=hv+kf$ we obtain

$$vU(h)=dvh-bv=d(u-kf)-bv=\delta x-dkf$$
$$vV(h)=-cvh+av=-c(u-kf)+av=\delta+ckf.$$

Using these relations for the products $vU(h)$ and $vV(h)$ we obtain, modulo f,

$$v^n g(h)=(vU(h))^n-\sum a_j(vU(h))^j(vV(h))^{n-j}=\delta^n x^n-\sum a_j\delta^j x^j\delta^{n-j}=\delta^n f(x)$$

and, since $(v,f)=1$, it follows that $g(h)=0$ modulo f.

The next step consists in showing that the operator of multiplication by h, taken in the basis $p_0,\dots,p_{n-1}$, has matrix $C(g)$.

First of all, we show that, for $0\leq j\leq n-2$ the product hp_j equals p_{j+1} modulo f. For $j\leq n-2$ we have $p_j=u^jv^{n-1-j}$ whence

$$hp_j=u^jv^{n-2-j}hv=u^jv^{n-2-j}(u-kf)=u^{j+1}v^{n-2-j}-u^jv^{n-2-j}kf$$

so that

$$hp_j=p_{j+1}\bmod f.$$

To compute hp_{n-1} we need some notation. If G is the leading coefficient of the polynomial g we can write g in the form

$$g(x)=G(x^n-\sum g_jx^j).$$

Now we intend to show that

$$hp_{n-1}=\sum g_jp_j \bmod f.$$

We argue as follows

$$hp_{n-1}-\sum g_jp_j=hu^{n-1}-\sum g_ju^jv^{n-1-j}=h.h^{n-1}v^{n-1}-\sum g_jh^jv^jv^{n-1-j}=$$
$$=v^{n-1}(h^n-\sum g_jh^j)=v^{n-1}g(h)=0.$$

In this manner we have shown that the matrix of $h(S)=\varphi(S)$ in the basis $p_0,p_1,\dots,p_{n-1}$ is exactly $C(g)$. If $M(\varphi)$ stands for the transformation matrix $p_j(x)=\sum M(\varphi)_{sj}x^s$ the result just proved may be formulated as follows.

There exists a matrix $M(\varphi)$ with the following property:

If C is the companion matrix of a polynomial f with $(v,f)=1$ then

$$C'=M(\varphi)^{-1}\varphi(C)M(\varphi)$$

is again a companion matrix.

More precisely, $C'=C(g)$ where

$$g=U^n-\sum_{j=0}^{n-1}a_jU^j\,V^{n-j}.$$

To conclude let us briefly comment on the behaviour of the matrices $M(\varphi)$ with respect to composition of the transformations φ.

Let n be a given positive integer. For each two by two matrix

$$A=\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

we define an n by n matrix $M(A)$ as follows. We begin by constructing a sequence of polynomials of degree $\le n-1$.

$$p_0(A,x),\ p_1(A,x),\dots,p_{n-1}(A,x)$$

and define $M(A)$ by the requirement that

$$M(A)_{jr}x^r=p_j(A,x).$$

The polynomials p_j are defined as follows

$$p_j(A,x)=u(x)^j\ v(x)^{n-1-j}$$

where $u(x)=ax+b$, $v(x)=cx+d$. Let us show now that the mapping $A\to M(A)$ is a representation of $GL(2)$. To see that consider two matrices A and B. It is easy to verify the relation

$$p_j(BA)=\sum M(B)_{js}p_s(A).$$

Indeed, we observe first that

$$u_{BA}(x)=v_A(x)\ u_B(\varphi)$$
$$v_{BA}(x)=v_A(x)\ v_B(\varphi)$$

where $\varphi=u_A(x)\ (v_A(x))^{-1}$. It follows that

$$p(BA,x)=u_{BA}(x)^j\ v_{BA}(x)^{n-1-j}=v_A(x)^{n-1}\ u_B(\varphi)^j\ v_B(\varphi)^{n-1-j}=$$
$$=v_A(x)^{n-1}\ p_j(B,\varphi)=v_A(x)^{n-1}\sum M(B)_{jr}\varphi^r=\sum M(B)_{jr}\ p_r(A,x).$$

Comparing coefficients of x^r on both sides of this relation we find that

$$M(BA)_{jv}=\sum_s M(B)_{js}\ M(A)_{sv}$$

whence

$$M(BA)=M(B)\ M(A).$$

References

[1] B.A. SHANE and S. BARNETT: On the bilinear transformation of companion matrices, Lin. Alg. Appl. 9(1974), 175-184.

[2] N.J. YOUNG: Linear fractional transforms of companion matrices, Glasgow Math. J. 20(1979), 129-132.

Institute of Mathematics, Czechoslovak Academy of Sciences, Žitná 25, 115 67 Praha 1, Czechoslovakia

(Oblatum 17.10. 1987)