Advanced Studies in Pure Mathematics 11, 1987
Commutative Algebra and Combinatorics
pp. 259–275

Linear Diophantine Equations and Invariant Theory of Matrices

Yasuo Teranishi

Introduction

In this paper, we shall study the Poincaré series of the ring of invariants of $n \times n$ matrices under the simultaneous adjoint action of $GL(n)$. This ring of invariants was studied by Procesi [3] and others. If $n=2$, it is well known that the ring of invariants of two generic matrices X and Y is a polynomial ring generated by 5 algebraically independent invariants

$$\operatorname{tr}(X),\ \operatorname{tr}(X^2),\ \operatorname{tr}(Y),\ \operatorname{tr}(Y^2),\ \operatorname{tr}(XY),$$

and hence the Poincaré series is

$$1/(1-s)(1-s^2)(1-t)(1-t^2)(1-st). \quad \text{(See [1]).}$$

However if $n \geq 3$, the ring of invariants is not polynomial ring. The Poincaré series of the ring of invariants for generic $n \times n$ matrices is related with the following generating function $F(t)$ of a linear diophantine equations defined by

$$F(t) = \sum_{r \geq 0} h(r) t^r,$$

where $h(r)$ is the number of $n \times n$ matrices $l = (l_{ij}) \in M(n, N)$ with the property:

$$\sum_{i,j} l_{ij} = r \quad \text{and} \quad \sum_j l_{ij} = \sum_j l_{ji}, \quad 1 \leq i \leq n.$$

General "reciprocity theorems" of the generating function of a linear diophantine equations is established by Stanley ([4], [5], [7]). We shall give simple proofs of some Stanley's results in [5].

By using a combinatorial method, we shall calculate the Poincaré series of the ring of invariants of two 4×4 generic matrices.

Received October 20, 1985.

Notations

N: the set of non-negative integers.
Z: the set of integers.
Q: the set of rational numbers.
C: the set of complex numbers.
$M(r, n, Z)$: the set of $r\times n$ matrices with Z-coefficients.
$M(n, R)$: the set of $n\times n$ matrices over a ring R.
For $a=(a_1, \cdots, a_n)$ and $b=(b_1, \cdots, b_n)$, $a<b$ means that $a_i<b_i$, $1\leq i\leq n$.
For $l=(l_{ij}) \in M(n, R)$, $|l|=\sum_{i,j} l_{ij}$.
For an integer l, $\underline{l}=(l, \cdots, l) \in Z^n$.
For $x=(x_1, \cdots, x_n) \in C^n$, $|x|<1$ means that $|x_i|<1$, $1\leq i\leq n$.

Acknowledgment

The author is grateful to Daniel Montanari for pointing out some mistakes in the original manuscript.

§ 1. Stanley's combinatorial reciprocity theorems

Let us consider a finite system of linear inhomogeneous diophantine equations (=I.D.E. system)

$$E_1(x):\ a_{11}x_1+\cdots+a_{1n}x_n=b_1$$
$$\vdots$$
$$E_r(x):\ a_{r1}x_1+\cdots+a_{rn}x_n=b_r.$$

Let $A=(a_{ij})$, $A \in M(r, n, Z)$, and $b=(b_1, \cdots, b_r) \in Z^r$. We denote by (A, b) for the I.D.E. system above.

For an n tuple $l=(l_1, \cdots, l_n) \in Z^n$, we denote by $E(A, b, l)$ the set of all solutions $m=(m_1, \cdots, m_n) \in Z^n$, $m\geq l$, to the I.D.E. system (A, b). For $l=(l_1, \cdots, l_n) \in Z^n$, let

$$F_l(A, b, x)=\sum_{m\in E(A,b,l)} x^m, \qquad x^m=x_1^{m_1}, \cdots, x_n^{m_n}.$$

Then $F_l(A, b, x)$ is a rational function in n variables $x_1, \cdots, x_n$. Let a_i be the i-th column vector of the matrix A. Let $\varepsilon_1, \cdots, \varepsilon_r$ be coordinate functions on C^r, and write, for $l \in Z^r$,

$$\varepsilon^l=\prod_i \varepsilon_i^{l_i}, \qquad l=(l_1, \cdots, l_r).$$

Let (A, b) be an I.D.E. system. For $l=(l_1, \cdots, l_n) \in Z^n$, let $G_l(X, \varepsilon)$ be the rational function in variables $\varepsilon_1, \cdots, \varepsilon_r, x_1, \cdots, x_n$ given by

$$G_l(x, \varepsilon)=\frac{\prod_i (\varepsilon^{a_i}x_i)^{l_i}\varepsilon^{-b}}{\prod_i (1-\varepsilon^{a_i}x_i)}.$$

If $l=(0, \cdots, 0)$, we write $G(x, \varepsilon)$ for $G_l(x, \varepsilon)$.

Lemma 1.1. *Suppose* $|x_1|<1, \cdots, |x_n|<1$. *Then*

$$F_l(A, b, x)=\left(\frac{1}{2\pi\sqrt{-1}}\right)^r \int\cdots\int G_l(x, \varepsilon)\frac{d\varepsilon_1\cdots d\varepsilon_r}{\varepsilon_1\cdots\varepsilon_r}$$

where the j-th integral from inside is taken over the counterclockwise unit circle in the complex ε_j*-plane.*

Proof. For $m=(m_1, \cdots, m_n)\in Z^n$, consider the integral

$$\int_0^1\cdots\int_0^1 \prod \varepsilon^{a_i m_i}\varepsilon^{-b}d\varphi_1\cdots d\varphi_r, \qquad \varepsilon_i=\exp 2\pi\sqrt{-1}\,\varphi_i.$$

Then this integral equals 1 or 0 according as $Am=b$ or not. Therefore we have;

$$\begin{aligned}F_l(A, b, x)&=\sum_{m\geq l}\int_0^1\cdots\int_0^1 \prod_i (\varepsilon^{a_i}x_i)^{m_i}\varepsilon^{-b}\,d\varphi_1\cdots d\varphi_r\\ &=\int_0^1\cdots\int_0^1 G_l(x, \varepsilon)d\varphi_1\cdots d\varphi_r\\ &=\left(\frac{1}{2\pi\sqrt{-1}}\right)^r\int\cdots\int G_l(x, \varepsilon)\frac{d\varepsilon_1\cdots d\varepsilon_r}{\varepsilon_1\cdots\varepsilon_r}.\end{aligned}$$

Suppose that $|x_1|>1, \cdots, |x_n|>1$. Then changing the variables $\varepsilon_j\to\varepsilon_j^{-1}$,

$$F_{\underline{1}-l}(A, -b, 1/x)=(-1)^{n-1}\left(\frac{1}{2\pi\sqrt{-1}}\right)^r\int\cdots\int G_l(x, \varepsilon)\frac{d\varepsilon_1\cdots d\varepsilon_r}{\varepsilon_1\cdots\varepsilon_r}$$

where $\underline{1}=(1, \cdots, 1)\in Z^n$, $1/x=(1/x_1, \cdots, 1/x_n)$ and the j-th integral is taken over the clockwise circle $|\varepsilon_j|=1$. Therefore we have the following

Theorem 1.1. *Taking the integrals over the counterclockwise (resp. clockwise) unit circles, let* $H_0(x)$ *(resp.* $H_\infty(x)$*) be the rational function in* $x_1, \cdots, x_n$ *defined by, in* $|x_1|<1, \cdots, |x_n|<1$ *(resp.* $|x_1|>1, \cdots, |x_n|>1$*),*

$$\int\cdots\int G_l(x, \varepsilon)\frac{d\varepsilon_1\cdots d\varepsilon_r}{\varepsilon_1\cdots\varepsilon_r}.$$

Then $F_l(A, b, x)$ *and* $F_{\underline{1}-l}(A, -b, 1/x)$ *are related by*

$$F_l(A, b, x)=(-1)^{n-r}F_{\underline{1}-l}(A, -b, 1/x)$$

if and only if $H_0(x)=H_\infty(x)$.

For an I.D.E. system (A, b), we denote by $d(A)$ the number of variables appearing in (A, b) minus the rank of A. If

$$F_{\underline{0}}(A, b, x)=(-1)^{d(A)}F_{\underline{1}}(A, -b, 1/x),$$

we say that (A, b) has the R-poroperty.

Lemma 1.2. *(A, b) has the R-property if and only if, for any $l \in \boldsymbol{Z}^n$,*

$$F_l(A, b+Al, x)=(-1)^{d(A)}F_{\underline{1}-l}(A, -b-Al, 1/x).$$

Proof. It follows from the definition of $F_l(A, b, x)$ that

$$F_l(A, b, x)=x^l F_{\underline{0}}(A, b-Al, x)$$

and

$$F_{\underline{1}-l}(A, -b, 1/x)=x^l F_{\underline{1}}(A, -b+Al, 1/x).$$

Then the proof follows immediately.

For an I.D.E. system, pick an integer k, $1\leq k\leq n$. We consider a new I.D.E. system (A', b'), $A' \in M(r, n, \boldsymbol{Z})$, $b' \in \boldsymbol{Z}^r$, defined as follows: $b'=(b(k)_1, \cdots, b(k)_r)$,

$$E'_1(x)=0=b(k)_1,$$
$$E'_2(x)=a_{2k}E_1(x)-a_{1k}E_2(x)=b(k)_2=a_{2k}b_1-a_{1k}b_2,$$
$$\cdots\cdots$$
$$E'_r(x)=a_{rk}E_1(x)-a_{1k}E_r(x)=b(k)_r=a_{rk}b_1-a_{1k}b_r.$$

We call (A', b') the k-eliminated system of (A, b), and denote by $(A(k), b(k))$ the k-eliminated system (A', b').

For an integer i, let C_i (resp. $-C_i$) denote the counterclockwise (resp. clockwise) unit circle in the complex ε_j-plane. We fix $\varepsilon_2, \cdots, \varepsilon_r$ ($|\varepsilon_i|=1, 2\leq i\leq r$) and consider $G(x, \varepsilon)$ as a function in $x=(x_1, \cdots, x_n)$ and ε_1. The integral

$$\int_{C_1} G(x, \varepsilon)\frac{d\varepsilon_1}{\varepsilon_1}, \qquad |x_1|<1, \quad 1\leq i\leq n,$$

can be computed by the residue theorem of the complex function theory

$$\frac{1}{2\pi\sqrt{-1}}\int_{C_1} G(x, \varepsilon)\frac{d\varepsilon_1}{\varepsilon_1}=\mathrm{Res}_{\varepsilon_1=0}G(x, \varepsilon)/\varepsilon_1+\sum_\lambda \mathrm{Res}_{\varepsilon_1=\lambda}G(x, \varepsilon)/\varepsilon_1$$

where $\sum$ is taken over all poles λ in $|\varepsilon_1|<1$.

If, $|x|>1$, similarly we have:

$$\frac{1}{2\pi\sqrt{-1}}\int_{-C_1} G(x, \varepsilon)\frac{d\varepsilon_1}{\varepsilon_1}=\mathrm{Res}_{\varepsilon_1=\infty}G(x, \varepsilon)/\varepsilon_1+\sum_{\lambda}\mathrm{Res}_{\varepsilon_1=\lambda}G(x, \varepsilon)/\varepsilon_1$$

where $\sum$ is taken over all poles λ of $G(x, \varepsilon)$ in $|\varepsilon_1|>1$.

Theorem 1.2. *Let (A, b) be an I.D.E. system. Let $R_0(x)$ and $R_\infty(x)$ be rational function in n variables x defined by*

$$R_0(x)=\int_{C_2}\cdots\int_{C_r}(\mathrm{Res}_{\varepsilon_1=0}G(x, \varepsilon)/\varepsilon_1)\frac{d\varepsilon_2\cdots d\varepsilon_r}{\varepsilon_2\cdots\varepsilon_r},$$

and

$$R_\infty(x)=\int_{-C_2}\cdots\int_{-C_r}(\mathrm{Res}_{\varepsilon_1=\infty}G(x, \varepsilon)/\varepsilon_1)\frac{d\varepsilon_2\cdots d\varepsilon_r}{\varepsilon_2\cdots\varepsilon_r}.$$

Suppose that the following conditions hold:

(1) $R_0(x)=R_\infty(x)$,

(2) *for any integer k satisfying $a_{1k}<0$, the k-eliminated system $(A(k), b(k))$ has the R-property.*

Then (A, b) has the R-property.

Proof. Let λ be a pole of the function $G(x, \varepsilon)$, $|x|<1$. Then λ is a root of the equation in ε_1:

$$1-\varepsilon_1^{a_{1k}}\cdots\varepsilon_r^{a_{rk}}x_k=0, \text{ for some } k \text{ such that } a_{1k}<0.$$

i.e., $\lambda=(\varepsilon_2^{a_{2k}}\cdots\varepsilon_r^{a_{rk}}x_k)^{-1/a_{1k}}$ for some fixed choice of the $-a_{1k}$-th root.

A direct computation shows that the residue of $G(x, \varepsilon)$ at $\varepsilon_1=\lambda$ is, under the assumption $|x|<1$, given by

$$\mathrm{Res}_{\varepsilon_1=\lambda}G(x, \varepsilon)/\varepsilon_1=-\frac{x_k^{-b_1/a_{1k}}}{a_{1k}}G(y, \varepsilon^{-1/a_{1k}}),$$

where $G(y, \varepsilon)$ denotes the function obtained from by the replacement a_i, $b\to a(k)_i$, $b(k)$, and $y_i=x_k^{-a_{1i}/a_{1k}}x_i$, $1\leq i\leq n$.

To compute $F(A, b, x)$ we can replace ε_j, $2\leq j\leq r$, with $\varepsilon^{-a_{1k}}$ (p. 230, [5]) in the following integral

$$\left(\frac{1}{2\pi\sqrt{-1}}\right)^{n-1}\int_{C_2}\cdots\int_{C_r}(\mathrm{Res}_{\varepsilon_1=\lambda}G(x, \varepsilon)/\varepsilon_1)$$

Then the integral above is, up to the factor $-x_k^{-1/a_{1k}}/a_{1k}$ replaced with

$$F(A(k), b(k), y)=\left(\frac{1}{2\pi\sqrt{-1}}\right)^{n-1}\int_{C_1}\cdots\int_{C_r} G(y, \varepsilon)\frac{d\varepsilon_1\cdots d\varepsilon_r}{\varepsilon_1\cdots\varepsilon_r}.$$

On the other hand, if $|x|>1$, all poles of $G(x, \varepsilon)$ are of the form

$$\lambda=(\varepsilon_2^{a_{2k}}\cdots\varepsilon_r^{a_{rk}}x_k)^{-1/a_{1k}}.$$

Therefore we can apply the same computation. Then our assumptions (1), (2), and Theorem 1.1 imply that (A, b) has the R-property.

We now suppose that the first equation of an I.D.E. system (A, b) has the R-property. Then by Proposition 10.3 in [5],

$$\mathrm{Res}_{\varepsilon_1=0}G(x, \varepsilon)/\varepsilon_1=\mathrm{Res}_{\varepsilon_1=\infty}G(x, \varepsilon)/\varepsilon_1=0.$$

thus $R_0(x)=R_\infty(x)$. Therefore in this case, we have the following

Theorem 1.3. *Let (A, b) be an I.D.E. system. Suppose that the first equation $a_{11}x_1+\cdots+a_{1n}x_n=b_1$ of (A, b) has the R-property as an I.D.E. system with one equation and, for any k satisfying $a_{1k}<0$, the k-eliminated system $(A(k), b(k))$ has the R-property. Then (A, b) has the R-property.*

The next proposition gives a simple criterion to have the R-property.

Proposition 1.1. *Let (A, b) be an I.D.E. system with r equations. Suppose the following inequalities hold:*

$$\sum_{i-} a_{ji}<-b_j+\sum_{i} a_{ji}l_i<\sum_{i+} a_{ji}, \qquad 1\le j\le r,$$

where $\sum_{i-} a_{ji}$ (resp. $\sum_{i+} a_{ji}$) denotes the sum of all (j, i)-entries of A satisfying $a_{ji}<0$ (resp. >0).

Then, for $l\in \boldsymbol{Z}^n$,

$$F_l(A, b, x)=(-1)^{d(A)}F_{\underline{1}-l}(A, b, 1/x).$$

Proof. We may assume that $l=0$ by Lemma 1.2. If $r=1$, the assertion is true by Proposition 10.4 in [5]. We proceed by induction on r. For any k, $1\le k\le n$, it is easy to show that

$$\sum_{i-} a(k)_{ji}<-b(k)_j<\sum_{i+} a(k)_{ji}, \qquad 1\le j\le r.$$

Then by Theorem 1.3, (A, b) has the R-property.

Proposition 1.2. *Let (A, b) be an I.D.E. system. Suppose that (A, b) has a solution $s=(s_1, \cdots, s_n)\in \boldsymbol{Q}^n$, $l_i-1<s_i\le l_i$ and $(A, 0)$ has a positive solution. Then*

$$F_l(A, b, x)=(-1)^{d(A)}F_{\underline{1}-l}(A, -b, 1/x), \qquad l \in Z^n.$$

Proof. We may assume that $l=0$ by Lemma 1.2. Then it follows from the assumption that (A, b) satisfies the condition in Proposition 1.1. Hence (A, b) has the R-property.

Proposition 1.3 (Proposition 8.3. [5]). *Suppose an I.D.E. system* $(A, 0)$ *has a solution*

$$x=(x_1, \cdots, x_n),\ x_1>0, \cdots, x_g>0,\ x_{g+1}<0, \cdots, x_n<0.$$

Then, $$F_l(A, \underline{0}, x)=(-1)^{d(A)}F_{\underline{1}-l}(A, \underline{0}, 1/x),$$

where $l=(\underbrace{0, \cdots, 0}_{g}, \underbrace{1, \cdots, 1}_{n-g})$.

Proof. By the assumption we have:

$$\sum_{1\le i\le g}' a_{ji}-\sum_{g+1\le i\le n}'' a_{ji}\ge 0, \quad -\sum_{1\le i\le g}'' a_{ji}+\sum_{g+1\le i\le n} a_{ji}\ge 0, \quad 1\le j\le r,$$

where $\sum'$ (resp. $\sum''$) denotes the sum of all terms >0 (resp. <0). Hence if $A\neq(0)$, we have: $\sum_{i-} a_{ji}<0<\sum_{i+} a_{ji}$.

Then by Proposition 1.1, $(A, 0)$ has the R-property. If $A=(0)$, it is obvious that $(A, \underline{0})$ has the R-property.

We shall need the following

Proposition 1.4 (13.3 Corollary [7]). *If* $\underline{1} \in E(A, \underline{0}, \underline{0})$, $F_0(A, \underline{0}, x)$ *satisfies the following functional equation*:

$$F_0(A, \underline{0}, 1/x)=(-1)^{d(A)}xF_0(A, \underline{0}, x).$$

§ 2. The ring of invariant of a semisimple group

Let G be a connected semisimple linear algebraic group, V_i, $1\le i\le l$, vector spaces over the complex number field C and ρ_i: $G\to GL(V)$ rational representations of G. Let $C[V]$ denote the polynomial ring over the vector space $V:=\oplus_i V_i$.

We denote by $C[V]_d$, $d=(d_1, \cdots, d_l) \in N^l$, the vector space of polynomials with degree $d_1, \cdots, d_l$ with respect to $V_1, \cdots, V_l$. This gives an N^l-graded structure of $C[V]^G$

$$C[V]=\bigoplus_{d\in N^l} C[V]_d.$$

Let R denote the ring of invariants of $C[V]$. Since $C[V]_d$ is a G-invariant subspace of $C[V]$, R has the structure of an N^l-graded algebra

$$R = \bigoplus_{d \in N^l} R_d, \qquad R_d = R \cap C[V]_d.$$

For $d=(d_1, \cdots, d_l) \in Z^l$, let us write $x^d = x_1^{d_1} \cdots x_l^{d_l}$. The Poincaré series of R is defined by

$$F(R, x) = \sum_{d \in N^l} \dim R_d x^d.$$

As well known, $F(R, x)$ is a rational function in l variables $x=(x_1, \cdots, x_l)$ and R is a Gorenstein ring by a theorem of Hochster-Roberts [2]. By Stanley's theorem [6], this is equivalent to say that $F(R, x)$ satisfies the following functional equation

$$F(R, x) = (-1)^d x^a F(R, 1/x),$$

where $d = \dim R$ and $a \in Z^l$.

Let K be a maximal compact subgroup of G and T a maximal torus of K. Let $\alpha_1, \cdots, \alpha_r$ be roots of K with respect to T and W the Weyl group of K. Considering a root as a function on T, let $D(g)$ be the function on T defined by

$$G(g) = (1-\alpha_1(g)) \cdots (1-\alpha_r(g)).$$

Then by Molien-Weyl formula [9], we have

$$(*) \qquad F(R, x) = \frac{1}{|W|} \int_T \frac{D(g)}{\prod_{i \geq 1} \det(1 - x_i g)} dg, \qquad |x_i| < 1$$

where dg is the Haar-measure on T.

Let us consider a special case.

Theorem 2.1. *Let ρ_i be the adjoint representation of G and $V_i =$ Lie G, $1 \leq i \leq l$. Then, if $l \geq 2$, $F(R, x)$ satisfies the following functional equation*

$$F(R, x) = (-1)^d x^a F(R, 1/x),$$

where $d = \dim V - \dim G$, $a = \dim G$.

Proof. Let $R(x_1, \cdots, x_l)$ be the function defined by

$$R(x_1, \cdots, x_l) = (1-x_1)^{\dim T} F(R, x).$$

Then, by $(*)$,

$$R(1, x_2, \cdots, x_l) = \frac{1}{|W|} \int_T \frac{dg}{\prod_{i \geq 2} \det(1 - x_i g)}$$

By Lemma 1.1, $|W|R(1, x_2, \cdots, x_l)$ is the generating function of solutions for an I.D.E. system $(A, \underline{0})$. Since G is semisimple,

$$\underline{1} \in E(A, \underline{0}, \underline{0}).$$

Therefore by Proposition 1.4, $R(1, x_2, \cdots, x_l)$ satisfies: for some $r \in \boldsymbol{N}$, $R(1, x_2, \cdots, x_l) = (-1)^r (x_2 \cdots x_l)^{-\dim G} R(1, 1/x_2, \cdots, 1/x_l)$, and hence $a = \dim G$. It follows from the following proposition that $d = \dim V - \dim G$.

Proposition 2.1. *If* $l \geq 2$, $\dim R = \dim V - \dim G$.

Proof. For $v \in V$, G_v denotes the isotropy subgroup of G. Then one sees easily that min. dim $G_v = 0$, and we have:

$$\begin{aligned} \dim R &= \dim V - \text{max. dim } G_v \\ &= \dim V - \dim G + \text{min. dim } G_v \\ &= \dim V - \dim G. \end{aligned}$$

Specializing x_i, $1 \leq i \leq l$, with a variable t, we consider R as an $\boldsymbol{N}$-graded algebra:

$$R = \bigoplus_{m \in \boldsymbol{N}} R_m.$$

The Poincaré series in one variable t is defined by

$$F(R, t) = \sum_{m \in \boldsymbol{N}} \dim R_m t^m.$$

Let $f_1, \cdots, f_m$, $m = \dim R$, be a homogeneous system of parameters of R. Since R is a Cohen-Macaulay ring, R is a free module over $\boldsymbol{C}[f_1, \cdots, f_m]$. Let $\varphi_1, \cdots, \varphi_r$ be a homogeneous system of generators of this module:

$$R = \bigoplus_i \varphi_i \boldsymbol{C}[f_1, \cdots, f_m].$$

Then

$$F(R, t) = \frac{\sum t^{\deg \varphi_j}}{\prod (1 - t^{\deg f_j})}.$$

It follows from the functional equation of $F(R, t)$ that

$$\dim G = \sum_j (\deg f_j - 1) - \frac{2}{r} \sum_i \deg \varphi_i.$$

Let us consider the Laurent expansion of $F(R, t)$ at $x = 1$:

$$F(R, t)=\frac{a}{(1-t)^m}+\frac{b}{(1-t)^{m-1}}+\cdots.$$

Then the coefficients a, b are given by

$$a=\frac{r}{\prod \deg f_j}$$

and

$$b=\frac{r\sum(\deg f_j-1)-2\sum \deg \varphi_i}{2\prod \deg f_j}$$

Thus a and b are related by dim $G=2b/a$.

§ 3. The Poincaré series of two generic $n\times n$ matrices

In this section, we shall study the invariant ring in the following situation:

$$G=GL(n, C),\quad V_1=V_2=M(n, C),\quad V=V_1\oplus V_2,$$
$$\rho_i=\text{the adjoint representation of } GL(n, C),\quad 1\leq i\leq 2.$$

Let us denote by X_{ij} (resp. Y_{ij}), $1\leq i, j\leq n$, the coordinate functions on V_1 (resp. V_2) with respect to the canonical basis of $M(n, C)$. Let X and Y be $n\times n$ generic matrices defined by

$$X=(x_{ij}),\qquad Y=(y_{ij}).$$

The Poincaré series of R is, in this case, the formal power series in two variables:

$$F(s, t)=\sum_{d\in N^2} \dim R_d s^{d_1}t^{d_2},\qquad d=(d_1, d_2).$$

By Molien-Weyl formula, $F(s, t)$ ($|s|<1$ and $|t|<1$) equals

$$(**)\qquad \frac{1}{n!}\int_0^1\cdots\int_0^1 \frac{\Delta\bar{\Delta}}{f(s)f(t)}d\varphi_1\cdots d\varphi_n,$$

where $\Delta=\prod_{i,j}(\varepsilon_i-\varepsilon_j)$, $\varepsilon_i=\exp 2\pi\sqrt{-1}\,\varphi_i$, $\bar{\Delta}$ is the complex conjugate of Δ and

$$f(x)=\prod_{i,j}(1-x\varepsilon_i\varepsilon_j^{-1}).$$

The functional equation of $F(s, t)$ is given by

$$F(1/s, 1/t)=(-1)^{n+1}(st)^{n^2}F(s, t).$$

For a finite sequence $\lambda=(\lambda_1, \cdots, \lambda_n)$ of nonnegative integers, the weight of λ is the sum of all terms of λ and is denoted by $|\lambda|$

A partition is a sequence $\lambda=(\lambda_1, \cdots, \lambda_n)$ of nonnegative integers in nonincreasing order $\lambda_1 \geq \cdots \geq \lambda_n$.

We denote by Y_n the set of all partitions with n terms:

$$Y_n := \{\lambda=(\lambda_1, \cdots, \lambda_n) : \lambda \text{ is a partition}\}.$$

For a partition λ, let s_λ denote the Schur function. For partitions λ, μ, ν in Y_n, let $c^\nu_{\lambda\mu}$ denote the nonnegative integer defined by

$$s_\lambda s_\mu = \sum_\nu c^\nu_{\lambda\mu} s_\nu.$$

Proposition 3.1.

$$\dim R_d = \sum_{\substack{|\lambda|=d_1 \\ |\mu|=d_2}} \sum_\nu (c^\nu_{\lambda\mu})^2, \qquad d=(d_1, d_2).$$

Proof. Let Λ be the ring of symmetric polynomials in n independent variables with $\boldsymbol{Z}$-coefficients. Let (,) be a scalar product defined by

$$(f, g) = \frac{1}{n!} \int_0^1 \cdots \int_0^1 f(\varepsilon_1, \cdots, \varepsilon_n) g(\varepsilon_1, \cdots, \varepsilon_n) \Delta \bar{\Delta} d\varphi_1 \cdots d\varphi_n,$$

$$\varepsilon_i = \exp(2\pi\sqrt{-1}\varphi_i), \quad 1 \leq i \leq n.$$

Then the Schur function s_λ form an orthonormal basis of Λ with respect to this scalar product. By (**) and the Cauchy identity

$$\frac{1}{\prod_{i,j}(1-x_i y_j)} = \sum_{\lambda,\mu} s_\lambda(x_1, \cdots, x_n) s_\mu(y_1, \cdots, y_n),$$

It follows that

$$F(s, t) = \sum_{\lambda,\mu} (s_\lambda s_\mu, s_\lambda s_\mu) s^{|\lambda|} t^{|\mu|}$$
$$= \sum_{\lambda,\mu} \sum_\nu (c^\nu_{\lambda\mu})^2 s^{|\lambda|} t^{|\mu|}.$$

Thus we obtain the desired result.

Consider the function $P(s, t) = (1-s)^n F(s, t)$. Then $P(s, t)$ is a rational function holomorphic in $\{(s, t) : |s|<1, |t|<1\}$. We set $F(t) = P(1, t)$.

Proposition 3.2. *Let $E(r)$ be the subset of $M(n, \boldsymbol{C})$ defined by*

$$E(r) = \{l=(l_{ij}) \in M(n, \boldsymbol{N}) : |l|=r \text{ and for all } i,\ 1 \leq i \leq n,\ \sum_j l_{ij} = \sum_j l_{ji}\}.$$

Then we have

$$F(t)=\sum_{r\geq 0} h(r)t^r, \qquad \text{where } h(r)=\sharp E(r).$$

Proof. From the definition of $F(t)$, it follows that

$$F(t)=\int_0^1\cdots\int_0^1 \frac{d\varphi_1\cdots d\varphi_n}{(1-t\varepsilon_i\varepsilon_j^{-1})}$$
$$=\sum_{l_{ij}}\int_0^1\cdots\int_0^1 \prod\left(\frac{\varepsilon_i}{\varepsilon_j}\right)d\varphi_1\cdots d\varphi_n t^{|l|}, \qquad l=(l_{ij}).$$

Since

$$\int_0^1\cdots\int_0^1 \prod\left(\frac{\varepsilon_i}{\varepsilon_j}\right)^{l_{ij}} d\varphi_1\cdots d\varphi_n=\begin{cases}1, & \text{if } l\in E,\\ 0, & \text{otherwise,}\end{cases}$$

we obtain the desired result.

We set $E=\bigcup E(r)$. Let sym (n) be the symmetric group of n letters. For $\sigma\in$ sym (n), let p_σ denote the permutation matrix corresponding to σ and e_σ denote the $n\times n$ matrix obtained by replacing diagonal entries in p_σ with zeros. For i, $1\leq i\leq n$, denote by e_i the $n\times n$ matrix having 1 in (i, i) entry and zeros in the others.

Lemma 3.1. *Any matrix in E can be written as an N-combination of* e_i, $1\leq i\leq n$, *and* e_σ, $\sigma\in$ sym (n).

Proof. Let $\underline{a}$ be a matrix in E. Take some nonnegative integers $m_1,\cdots,m_n$ such that $\underline{a}+\sum m_i e_i$ is an integer stochastic matrix. Then by Berkoff-Von Neumann theorem, $\underline{a}+\sum m_i e_i$ can be written as an N-combination of permutation matrices. So we have

$$\underline{a}+\sum m_i e_i=\sum l_i e_i+\sum l_\sigma e_\sigma,$$

for suitable nonnegative integers l_i, l_σ.

Comparing the diagonal entries in the expression above, we see that $m_i\leq l_i$ for all $1\leq i\leq n$. Hence we have

$$\underline{a}=\sum_i (l_i-m_i)e_i+\sum l_\sigma e_\sigma.$$

This shows that $\underline{a}$ is written as an N-combination of e_i, e_σ.

A matrix $\underline{a}$ in E is called *completely fundamental* if whenever $m\underline{a}=\underline{b}+\underline{c}$, for some positive integer m and $\underline{b}$, $\underline{c}\in E$, then $\underline{b}=r\underline{a}$ for some nonnegative integer r, $r\leq m$. Then the set of completely fundamental elements of E consists of the following matrices:

e_i ($1 \leq i \leq n$), e_σ ($\sigma \in$ {cyclic permutations}$-\{e\}$, e is the unit in sym(n)),

Proposition 3.3.
(1) $F(1/t) = -t^{n^2}F(t)$.
(2) *There is a polynomial* $R(t)$ *with integer coefficients such that*

$$F(t) = \frac{R(t)}{(1-t)^n \prod (1-t^{|\sigma|})}$$

where σ *runs over all cyclic permutations* ($\sigma \neq e$) *in* sym(n).

Proof. (1) follows from the functional equation of $F(s, t)$, and (2) follows from 3.7 Theorem [7].

Example 1. If $n=2$, it follows from Proposition 3.3 that $R(t)=1$, and hence we have

$$F(t) = \frac{1}{(1-t)^2(1-t^2)}.$$

In this case, the Poincaré series is given by

$$F(s, t) = \frac{1}{(1-s)(1-s^2)(1-t)(1-t^2)(1-st)}.$$

In fact, the ring of invariants R is a polynomial ring generated by 5 algebraically independent invariants $\mathrm{tr}(X)$, $\mathrm{tr}(X^2)$, $\mathrm{tr}(Y)$, $\mathrm{tr}(Y^2)$, $\mathrm{tr}(XY)$ where tr denotes trace of a matrix (See [1], [8]).

Example 2. If $n=3$, one sees immediately that $h(0)=1$, $h(1)=3$ and $h(2)=6$. Hence $F(t)$ is of the form

$$F(t) = 1 + 3t + 6t^2 + (\text{higher terms}).$$

Then by Proposition 3.3, $R(t) = 1 - t^6$ and so

$$F(t) = \frac{1+t^3}{(1-t)^3(1-t^2)^3(1-t^3)}.$$

In this case, $\mathrm{tr}(X)$, $\mathrm{tr}(X^2)$, $\mathrm{tr}(X^3)$, $\mathrm{tr}(Y)$, $\mathrm{tr}(Y^2)$, $\mathrm{tr}(Y^3)$, $\mathrm{tr}(XY)$, $\mathrm{tr}((XY)^2)$, $\mathrm{tr}(X^2Y)$, $\mathrm{tr}(XY^2)$ are a homogeneous system of parameters of the ring of invariants R. Denoting by C the subring of R generated by these invariants, we have

$$R = C + \mathrm{tr}(XYX^2Y^2)C.$$

The Poincaré series $F(s, t)$ is given by

$$F(s, t)=\frac{1+s^{3}t^{3}}{Q(s, t)}$$

where

$$Q(s, t)=(1-s)(1-s^{2})(1-s^{3})(1-t)(1-t^{2})(1-t^{3})(1-st)\times(1-s^{2}t)(1-st^{2})(1-s^{2}t^{2}).$$

§ 4. The Poincaré series of the ring of invariants for $n=4$

As an application we shall determine the Poincaré series $F(s, t)$ for $n=4$. We shall need the following proposition in [8].

Proposition 4.1. *Let $f_1, \cdots, f_{17}$ be the invariants of R defined by*

$$f_1=\operatorname{tr}(X),\quad f_2=\operatorname{tr}(X^2),\quad f_3=\operatorname{tr}(X^3),\quad f_4=\operatorname{tr}(X^4),$$
$$f_5=\operatorname{tr}(Y),\quad f_6=\operatorname{tr}(Y^2),\quad f_7=\operatorname{tr}(Y^3),\quad f_8=\operatorname{tr}(Y^4),$$
$$f_9=\operatorname{tr}(XY),\quad f_{10}=\operatorname{tr}(X^2Y^2),\quad f_{11}=\operatorname{tr}(XY^2),$$
$$f_{12}=\operatorname{tr}(X^2Y),\quad f_{13}=\operatorname{tr}(XY^3),\quad f_{14}=\operatorname{tr}(X^3Y),$$
$$f_{15}=\operatorname{tr}(XYXY),\quad f_{16}=\operatorname{tr}(XY^2XY^2),\quad f_{17}=\operatorname{tr}(X^2YX^2Y).$$

Then these invariants $f_1, \cdots, f_{17}$ are homogeneous system of parameters of the ring of invariants R. Let C denote the subring of R generated by these invariants $f_1, \cdots, f_{17}$.

Theorem 4.1. *If $n=4$, the Poincaré series $F(s, t)$ is given by $F(s, t)=R(s, t)/Q(s, t)$, where*

$$\begin{aligned}Q(s, t)=&(1-s)(1-s^{2})(1-s^{3})(1-s^{4})(1-t)(1-t^{2})\\&\times(1-t^{3})(1-t^{4})(1-st)(1-s^{2}t^{2})^{2}(1-st^{2})\\&\times(1-s^{2}t)(1-st^{3})(1-s^{3}t)(1-s^{2}t^{4})(1-s^{4}t^{2}),\\R(s, t)=&1+s^{2}t^{3}+2s^{3}t^{3}+s^{3}t^{4}+s^{4}t^{3}+s^{3}t^{6}+s^{6}t^{3}+2s^{4}t^{4}\\&+s^{3}t^{5}+s^{5}t^{3}+s^{4}t^{5}+s^{5}t^{4}+2s^{5}t^{5}+s^{4}t^{6}+s^{6}t^{4}\\&+2s^{5}t^{6}+2s^{6}t^{5}+2s^{6}t^{6}+s^{6}t^{8}+s^{8}t^{6}+2s^{6}t^{7}+2s^{7}t^{6}\\&+2s^{7}t^{7}+s^{7}t^{8}+s^{8}t^{7}+s^{7}t^{9}+s^{9}t^{7}+2s^{8}t^{8}+s^{8}t^{9}\\&+s^{9}t^{8}+s^{9}t^{6}+s^{6}t^{9}+s^{10}t^{9}+s^{12}t^{12}.\end{aligned}$$

Proof. Let $\varphi_1, \cdots, \varphi_r$ be homogeneous generators of R over the subring C. Let $S(s, t)$ be the polynomial defined by

$$S(s,t)=\sum h_{ij}s^i t^j, \qquad h_{ij}=\#\{\varphi_k\colon \deg\varphi_k=(i,j)\}.$$

We shall prove that $S(s,t)=R(s,t)$.

It follows from the functional equation of $F(t)$ that $S(1,t)$ is a polynomial of degree 12 of the form

$$S(1,t)=\sum_i a_i t^i, \qquad a_i=a_{12-i}, \qquad 0\leq i\leq 12.$$

For a matrix A, we mean by weight of A the summation of all entries of A. Then one can easily obtain;

(1) all matrices with weight 3 which can not be written as a N-combination of matrices with weight lower than 3 are $A_1, \cdots, A_8$, where

$$A_1=\begin{bmatrix}0&0&0&0\\0&0&1&0\\0&0&0&1\\0&1&0&0\end{bmatrix}, \quad A_2=\begin{bmatrix}0&0&1&0\\0&0&0&0\\0&0&0&1\\1&0&0&0\end{bmatrix},$$

$$A_3=\begin{bmatrix}0&1&0&0\\0&0&0&1\\0&0&0&0\\1&0&0&0\end{bmatrix}, \quad A_4=\begin{bmatrix}0&1&0&0\\0&0&1&0\\1&0&0&0\\0&0&0&0\end{bmatrix},$$

$$A_5={}^tA_1, \quad A_6={}^tA_2, \quad A_7={}^tA_3, \quad A_8={}^tA_4.$$

(2) all matrices with weight 4 which can not be written as a N-combination of matrices with weight lower than 4 are $B_1, \cdots, B_6$, where

$$B_1=\begin{bmatrix}0&1&0&0\\0&0&1&0\\0&0&0&1\\1&0&0&0\end{bmatrix}, \quad B_2=\begin{bmatrix}0&1&0&0\\0&0&0&1\\1&0&0&0\\0&0&1&0\end{bmatrix},$$

$$B_3=\begin{bmatrix}0&0&1&0\\0&0&0&1\\0&1&0&0\\1&0&0&0\end{bmatrix},$$

$$B_4={}^tB_1, \quad B_5={}^tB_2, \quad B_6={}^tB_3.$$

Therefore $F(t)$ is of the form

$$F(t)=\frac{1}{(1-t)^4(1-t^2)^6}\quad(1+8t^3+6t^4+\text{higher terms}).$$

Then, by using the functional equation of $F(t)$, we obtain:

(***) $\quad S(1,t)=1+t^2+6t^3+5t^4+6t^5+10t^6+6t^7+5t^8+6t^9+t^{10}+t^{12}.$

We need the following

Lemma 4.1 (Proposition 5.1 [8]). *R is generated by invariants of the form*

$$\mathrm{tr}(X^{a_1}Y^{a_2}X^{a_3}Y^{a_4}),\quad \mathrm{tr}(X^aYX^aY^2X^aY^3),\quad \mathrm{tr}(Y^aXY^aX^2Y^aX^3),$$
$$0\leq a,\ a_1,\ \cdots,\ a_4\leq 3,\quad \text{and}\quad \mathrm{tr}(XYX^2Y^2X^3Y^3).$$

We recall the Cayley-Hamilton theorem for $n\times n$ matrices:

$$X_{\sigma(1)}\cdots X_{\sigma(n)}+\sum_k\sum_u\sum_\sigma q_u\,\mathrm{tr}(X_{\sigma(1)}\cdots X_{\sigma(u_1)})\ \vdots\ X_{\sigma(k+1)}X_{\sigma(k+2)}\cdots X_{\sigma(n)}=0,$$

for suitable $q_u\in\mathcal{Q}$ and j-tuples $u=(u_1,\cdots,u_j)$ such that $1\leq u_1\leq u_2\leq\cdots\leq u_j$ and $u_1+\cdots+u_j=k$. Here σ ranges over all permutations on $\{1,2,\cdots,n\}$.

Lemma 4.2.

(1) $h_{ij}=0,\quad \text{if } i\leq 2,\ j\geq 4,$

(2) $h_{ij}=0,\quad \text{if } i\leq 4,\ j\geq 7,$

(3) $h_{33}\leq 2,\quad h_{34}\leq 1,\quad h_{35}\leq 1,\quad h_{43}\leq 1,$
$h_{44}\leq 2,\quad h_{45}\leq 1,\quad h_{46}\leq 1,\quad h_{55}\leq 2,$
$h_{75}=0,\quad h_{63}\leq 1,\quad h_{65}\leq 2.$

Proof. This follows from the Cayley-Hamilton theorem and Lemma 4.1.

We continue the proof of Theorem 4.1. By (***) and Lemma 4.2, we have equalities in Lemma 4.2 (3) and $h_{66}=2$, $h_{23}=1$. Since $h_{ij}=h_{ji}$ and $h_{ij}=h_{12-i,12-j}$, we obtain $S(s,t)=R(s,t)$.

References

[1] E. Formanek, P. Halpin and W.-C. W. Li, The Poincaré series of the ring of 2×2 generic matrices, J. Algebra, **69** (1981), 105–112.

[2] M. Hochster and L. Roberts, Ring of invariants of reductive groups acting on regular rings are Cohen-Macaulay, Adv. in Math., **13** (1974), 115–175.

[3] C. Procesi, The invariant theory of $n\times n$ matrices, Adv. in Math., **19** (1976), 306–381.

[4] R. Stanley, Linear homogeneous diophantine equations and magic labeling of graphs, Duke Math. J., **40** (1973), 607–632.

[5] ——, Combinatorial reciprocity theorems, Adv. in Math., **14** (1974), 194–253.

[6] ——, Hilbert functions of graded algebras, Adv. in Math., **28** (1978), 57–83.

[7] ——, Combinatorics and commutative algebra, Progress in Math., **41** (1983).
[8] Y. Teranishi, The ring of invariants of matrices, Nagoya Math. J., **104** (1986), 149–161.
[9] H. Weyl, Zur Darstellungtheorie und Invarientenabzahlug der projectiven, der Komplex und Drehungsgruppe, Ges. Abh. Bd III, 1–25.

Department of Mathematics
Faculty of Science
Nagoya University
Chikusa-ku, Nagoya 464
Japan