PROBLEMS OF COMBINATORIAL OPTIMIZATION, STATISTICAL SUMS AND REPRESENTATIONS OF THE FULL LINEAR GROUP

A. I. Barvinok

1. Statement of the Problem and Notation. In this paper we consider the complexity of the computation of relative invariants of the full linear group in connection with some problems of combinatorial optimization. The necessary prerequisites from representation theory can be found in [1, 2]; introductions to the combinatorial problems in question can be found in [3, 4].

Let \( V \) be an \( n \)-dimensional vector space over the field \( k \) (the field \( k \) is \( \mathbb{R}, \mathbb{C} \) or \( \mathbb{Q} \)) with scalar product \( \langle , \rangle \) and orthonormal basis \( e_1, e_2, \ldots, e_n \). The space \( V^\otimes k \), the \( k \)-th tensor power of \( V \), can in the usual way be equipped with a scalar product (which will also be denoted by \( \langle , \rangle \)), in which

\[
\langle u_1 \otimes \cdots \otimes u_k, v_1 \otimes \cdots \otimes v_k \rangle = \prod_{i=1}^{k} \langle u_i, v_i \rangle;
\]

\( u_i, v_i \in V \).

The group \( GL(V) \) of nondegenerate linear transformations acts on the space \( V^\otimes k \) by the \( k \)-th tensor power of its natural action on \( V \), and the symmetric group \( S_k \) acts by permutation of the components. For decomposable tensors these actions are expressed as follows:

\[
G(v_1 \otimes \cdots \otimes v_k) = G v_1 \otimes \cdots \otimes G v_k; \quad G \in GL(V);
\]

\[
g(v_1 \otimes \cdots \otimes v_k) = v_{g^{-1}(1)} \otimes \cdots \otimes v_{g^{-1}(k)}; \quad g \in S_k.
\]

These actions can be extended by \( k \)-linearity to the group algebras \( kS_k, kGL \). For more detail see [1, 2].

We fix a number \( l \in \mathbb{N} \) and assume that \( n = lm, m \in \mathbb{N} \). Let \( \varepsilon = \sum_{g \in S_n} \text{sgn} \, g \cdot g \equiv k[S_n] \), where \( \text{sgn} \, g \) is the sign of the permutation \( g \); \( \text{sgn} \, g = 1 \), if \( g \) is an even permutation and \( \text{sgn} \, g = -1 \), if \( g \) is an odd permutation. Let \( B \in V^\otimes l; B = \| B (i_1, \ldots, i_l) \|, 1 \leq i_j \leq n \) (for convenience we write the subscripts here and in future in parentheses). It is well known that the expression

\[
P(B) = \langle B^\otimes m, e(e_1 \otimes \cdots \otimes e_n) \rangle = \sum_{g \in S_n} \text{sgn} \, g \prod_{i=1}^{m} B(g(l(i-1)+1), g(l(i-1)+2), \ldots, g(l(i)))
\]

is a relative invariant of the group \( GL(V) \), i.e., \( P(G(B)) = \det G \cdot P(B) \) (cf. [2]). If the number \( l \) is odd, it is easy to see that \( P(B) = 0 \, \forall B \in V^\otimes l \), and therefore we shall assume from now on that, unless otherwise stated, the number \( l \) is even; \( l = 2p \). The letters \( n, l, m, p \) will denote the values introduced above for the rest of the paper. For \( l = 2 \) the corresponding invariant is well-known and is called the Pfaffian (notation: Pf \( (B) \); often one introduces a normalization factor \( 1/2^m m! \)). It is known that the computation of the Pfaffian can be achieved using a number of arithmetic operations which is polynomial in \( n \) (cf. [1, pp. 229-232; 4, pp. 318-329]).

In this paper we consider the complexity of the computation of the relative invariant \( P(B) \) for fixed \( l > 2 \). The complexity of an algorithm, working with real numbers, here denotes the number of arithmetical operations which it uses (additions, subtractions, multiplications, divisions) and of tests for equality to zero (necessary for admissibility of divisions). The precise definition of this so-called "real complexity" can be found in [5].
We may also assume that the computational model of [6, pp. 34-40] applies. The results obtained here will also apply in the traditional setting of complexity theory under the assumption that all the numbers appearing in the process are rationals (cf. [3]).

The polynomial \( P(B) \) is not the only relative invariant of the group \( GL(V) \). We fix a number \( d \in \mathbb{N} \) and let \( \mu \in k [S_{nd}] \) be an element of the group algebra contained in a two-sided simple ideal \( I_d \), on which an irreducible representation of the group \( S_{nd} \) acts with Young diagram \((d, \ldots, d) \) \((n \text{ terms})\). Let \( B \in \mathcal{V}^\otimes l \) and \( l | nd \). The polynomial \( P_{\mu}(B) = \langle B^\otimes nd, \mu ((e_1 \otimes \ldots \otimes e_n)^{\otimes d}) \rangle \) is the general form of a relative invariant of the group \( GL(V) \): \( P_{\mu}(G(B)) = \det G \cdot P_{\mu}(B) \) (cf. [1; 2, p. 327]). The paper also discusses the complexity of the computation of the relative invariant \( P_{\mu} \). Note that for \( d = 1 \mu = c \cdot c, c \in k, P_{\mu}(B) = c \cdot P(B) \). For \( d > 1 \) and \( t \) odd we have in general \( P_{\mu}(B) \neq 0 \).

To conclude this section we introduce the tensor \( U_{\nu, l} \in \mathcal{V}^\otimes l \) \((\text{often denoted simply by } U)\) which will be used in the sequel: \( U(1, 2, \ldots, l) = U(l + 1, \ldots, 2l) = \ldots = U((m - 1)l + 1, \ldots, ml) = 1, U(i_1, \ldots, i_l) = 0 \), if there does not exist a number \( 0 \leq t \leq ml - 1 \) such that \( i_j = tl + j, 1 \leq j \leq l \). It is easy to verify that \( P(U_{\nu, l}) = ml (\nu \text{ even}) \).

2. Main Results. Definition. The rank of a tensor \( B \in \mathcal{V}^\otimes l \) is the smallest number \( r = \text{rank } B \) such that the tensor \( B \) can be represented in the form

\[
B = \sum_{i=1} u_i^1 \otimes u_i^2 \otimes \ldots \otimes u_i^l, \quad \text{where } u_i^j \in V
\]  
(1)

(cf. [7]).

The 2-rank of a tensor \( B \in \mathcal{V}^\otimes l \) is the smallest number \( r = \text{rank}_s B \) such that the tensor \( B \) can be represented in the form

\[
B = \sum_{i=1} A_i^1 \otimes A_i^2 \otimes \ldots \otimes A_i^l, \quad A_i^j \in V \otimes V.
\]  
(2)

THEOREM 1. 1) Let \( k \in \mathbb{N} \) be a fixed number and assume that at least one of the following conditions holds:

a) \( \text{rank } B = n/l + k \) and the tensor \( B \) is given in the form (1);

b) \( \text{rank}_s B = k \) and the tensor \( B \) is given in the form (2).

Then there exists a polynomial algorithm for the computation of \( P(B) \).

2) The problem of computing the relative invariant \( P \) for an arbitrary \( l \)-tensor is polynomial with respect to the problem of computing \( P(B) \) if we assume that:

a') \( \text{rank } B \leq 2n/l \) and the tensor \( B \) is given in the form (1).

In points 1) and 2) the number \( l \) is assumed to be fixed and the problem extends over \( n \in \mathbb{N}, B \in \mathcal{V}^\otimes l \).

Proof. 1) Below we set out the required algorithms which are readily seen to be polynomial in the number \( n \) of arithmetic operations. The degree of the polynomial depends linearly on \( k \).

a) We extend the set of vectors \( e_1, \ldots, e_n \) to an orthonormal basis of the space \( V' \supseteq V \) by means of the vectors \( e_{n+1}, \ldots, e_{n+k} \); \( \dim V' = rl = n' \). We construct an operator \( G \in \text{End } (V') \) \((\text{End denotes the ring of endomorphisms of the space}) \) such that \( G(e_i) = u_i^1 \), where \( (i - 1)k + j, 1 \leq j \leq l \). Let \( U = U_{\nu, l} \in \mathcal{V}^\otimes l \). We have \( G(U) \) and

\[
P(B) = \langle B^\otimes m, \varepsilon (e_1 \otimes \ldots \otimes e_n) \rangle = \langle G(U)^\otimes m, \varepsilon (e_1 \otimes \ldots \otimes \varepsilon e_n) \rangle.
\]

The space \( L \subset \mathcal{V}^\otimes n \) generated by all vectors of the form \( \varepsilon (e_{i_1} \otimes \ldots \otimes e_{i_n}) \) \((1 \leq i_1 < i_2 < \ldots < i_n \leq n') \) is invariant under the action of the group \( GL(V') \) and of the ring \( \text{End } (V') \). In particular, the subspace \( L \subset \mathcal{V}^\otimes n \) is an isotypic component of the irreducible representation
of the group $S_n$ corresponding to the Young diagram $(1, 1, \ldots, 1)$, and therefore an isotypic component of the corresponding irreducible representation of the group $GL(V')$ (cf. [1, 2]).

Here $\dim L = \binom{n + km - 1}{n} \leq n^k$. For convenience we will abbreviate the vector $e(e_1 \otimes \ldots \otimes e_n)$ by $f_T$, where $T$ is the ordered set of indices $i_1, \ldots, i_n$, and the vector $e(e_1 \otimes \ldots \otimes e_n)$ will be denoted by $f$.

It is clear that we have $\langle U'^m, f_T \rangle = \pm m!$ if the sequence of numbers in $T$ can be recorded in such a way that we have $m$ non-intersecting intervals of length $i$; the sign is given as the sign of corresponding permutation. $\langle U'^m, f_T \rangle = 0$ in all other cases. For the computation of $P(B)$ it suffices to construct a decomposition of the vector $G^k f$ in the basis $\{f_T\}$ of the space $L$. The ring $End(V')$ is generated under multiplication by the generators $E_{ij}(a) = I + a e_{ij}$, where $I$ is the identity operator, $e_{ij}(k, s) = \delta_{ik}\delta_{js}$ a matrix unit. These generators operate as follows on the basis vectors of the space $L$: for $i \neq j$ $E_{ij}(a)f_T = f_T$ if $j \not\in T$ or $i \in T$; $E_{ii}(a)f_T = f_T \pm a f_T \cup i$ otherwise, where the sign is determined by the permutation which rearranges the sequence $\{i_1, \ldots, i_n\} \setminus j \cup i$; $E_{ii}(a)f_T = (1 + a)f_T$ if $i \in T$ and $E_{ii}(a)f_T = f_T$ otherwise (cf. [2, pp. 328–330]). By decomposing the operator $G^k \in End(V')$ into a product of no more than $(n')^3$ generators $E_{ij}(a)$ (a standard problem of linear algebra) and then applying the above formulas we obtain the required decomposition of the vector $G^k f$ in terms of the basis of the space $L$.

b) Let $\{C_i\}, i = 1, \ldots, p$ be tensors, $C_i \subseteq V \otimes V$. It is easily seen that $\langle C_i \otimes \ldots \otimes C_p, f \rangle$ does not depend on the order of the indices $i_1, \ldots, i_p$. Therefore

$$P(B) = \sum_{q_1 + \ldots + q_k = m} \langle \otimes_{i,j} (C_i \otimes q_i f),$$

where the sum is taken over all $(q_1, \ldots, q_k) \subseteq N^k$, $q_1 + \ldots + q_k = m$; the index $i$ ranges over the set $[1: k]$, the index $j$ over the set $[1: p]$. Each of the $\binom{m + k - 1}{m}$ terms is easily computed if we apply the following method. Let $C_1, \ldots, C_k \subseteq V \otimes V$ and $t = (t_1, \ldots, t_s), t_i \subseteq k$.

Put $C_t = t_1 C_1 + \ldots + t_s C_s$.

We have

$$Pf(C_t) = \sum_{q_1 + \ldots + q_s = n/k} \langle C_t \otimes q_s, f \rangle.$$
We have
\[ P(B) = \langle B^\otimes m, f \rangle = \frac{1}{m!} \langle B^\otimes m \otimes U^\otimes m, f \rangle. \]

The computation of the last expression follows from the computation of the values of \( P(B + tU, \lambda) \) for \( m \) distinct values of the parameter \( t \leq k \) similarly to what was done in part 1), b). Here \( B_i = B + tU, \lambda \subseteq V^\otimes t \) and \( \operatorname{rank} B_i \leq 2m' \). The theorem is established.

In order to raise the question of the complexity of the computation of a general invariant \( P_\mu(B) \) as a function of \( n, \mu, B \), it is necessary to determine exactly how the element \( \mu \in S_{md} \) is given since for each fixed \( d > 1 \) the dimension of the ideal \( I_d \) grows exponentially with \( n \). We will therefore assume that the number \( n \) and the element \( \mu \) are fixed and consider the problem over \( B \subseteq V^\otimes t \), \( \dim V = n \). The number \( l \) is not necessarily even now.

**THEOREM 2.** For all numbers \( d, k, l \in \mathbb{N} \) there exists a polynomial \( t(x) \), such that \( \forall n, n = \operatorname{Im} \) and \( \mu \in I_d \) in the computation of the expression \( P_\mu(B) \) on the set \( \{ B \subseteq V^\otimes t \}, \dim V = n \) and the tensor \( B \) satisfies condition 1) a) of theorem 1 \} there exists an algorithm whose complexity does not exceed \( t(n) \).

**Proof.** We repeat the considerations of 1) a) in Theorem 1, except that now the space \( L \) and its basis have to be constructed for each \( n, \mu \) individually. As before we extend the space \( V \) to \( V' \) and construct an element \( G \in \text{End}(V') \). We denote by \( f \) the element \( \mu(e_1 \otimes \ldots \otimes e_n)^d \subseteq V^\otimes md \).

We have
\[ P_\mu(B) = \langle B^\otimes md, f \rangle = \langle U^\otimes md, G^\ast f \rangle. \]

Put \( L = \text{Lin} \{ Gf, G \in \text{End}(V') \} \). The space \( L \) is contained in an isotypic component of the representation \( \pi \) of \( GL(V') \) corresponding to the Young diagram \( (d, \ldots, d) \). Since the space \( L \) is generated as \( GL(V') \)-module by a unique element we have \( \dim L \leq \dim \pi \) and for fixed \( d, k \) it follows from the formula for \( \dim \pi \) (cf [2, p. 326]) that \( \dim L \leq q(n) \), where \( q \) is some polynomial. We choose a basis \( \{ f_\tau \} \) in the space \( L \) (for example as union of standard bases of Weyl modules [2, pp. 320-327]) on which the generators \( E_{ij}(\alpha) \) act as follows:
\[ E_{ij}(\alpha)f_\tau = \sum_{\tau'} p_{ij}^{\tau\tau'}(\alpha)f_\tau', \]
where \( \{ p_{ij}^{\tau\tau'} \} \) are polynomials and \( \deg p_{ij}^{\tau\tau'} \leq d \). For the required algorithm it suffices to compute the values \( \gamma_\tau = \langle U^\otimes md, f_\tau \rangle \).

The algorithm consists of the following computations, starting from formula (4) for the decomposition of the vector \( G^\ast f \) in terms of the basis \( \{ f_\tau \} \) followed by the computation of the values \( P_\mu(B) \) using the coefficients \( \{ \gamma_\tau \} \). The theorem is established.

**Remark.** A similar result can be established for the computation of invariants \( P_\mu(B) \) for tensors of fixed 2-rank. As in the proof of 1) b) of Theorem 1 this requires the ability to compute the corresponding invariant for \( t = 2 \). The corresponding algorithm is based on the reduction of bilinear forms to canonical form [8, pp. 411-440]. This question will be addressed in more detail in a further paper by the author.

3. Problems Concerning Weighted Decompositions, Covers and the Computation of Relative Invariants. Many problems in combinatorial optimization can in general form be expressed as maximum problems:

\[ \max \{ f(g), g \subseteq S_n \} \]
for some function \( f: S_n \to \mathbb{R} \).

We introduce the following statistical sum:

\[
S_f (f; t) = \sum_{g \in S_n} \exp (t f (g)) \mu (g),
\]

where \( \mu: S_n \to \mathbb{R} \) is some weight function on the group \( S_n \); \( t \in \mathbb{R} \).

**Proposition 1.** Assume that the following implication holds: \( f (g) = f (h) \Rightarrow \mu (g) = \mu (h) \).

Then

\[
\lim_{t \to -\infty} t^{-1} \log | S_f (f; t) | = \max \{ f (g): g \in S_n, \mu (g) \neq 0 \}.
\]

Various modifications of this obvious result have been applied in optimization, e.g., in the "annealing" method derived from statistical physics (cf. [9]), although it appears that they have never before been formalized in this way. Below we give examples of problems for which the computation of the sums \( S_f (f; t) \) reduces to the computation of relative invariants.

I. The matching problem (cf. [4, pp. 307-318]). Let \( \| a (i, f) \| (1 \leq i, j \leq n) \) be a real matrix and \( f (g) = \sum_{i=1}^n a (i, g (i)) \). Put \( A_t (i, j) = \exp \{ ta (i, j) \} \) and \( \epsilon (g) = \text{sgn} \, g \). Then

\[
S (f; t) = \det \| A_t (i, j) \|.
\]

II. The problem of weighted decompositions. Given a real tensor \( b = \| b (i_1, \ldots, i_l) \| \in V^\otimes l \), \( \dim V = n = lm \) and \( f (g) = \sum_{i=1}^n b (g (l (i - 1) + 1), g (l (i - 1) + 2), \ldots, g (l i)) \). \( f (g) \) is a summation weight of the decomposition of the set \( \{1, 2, \ldots, n\} \) into \( m \) ordered \( \ell \)-vectors \( \{g (1), \ldots, g (l)\}, \{g (\ell + 1), \ldots, g (2\ell)\}, \ldots, \{g (\ell (m - 1) + 1), \ldots, g (\ell m)\} \), where the weight of the ordered vector \( \{i_1, \ldots, i_\ell\} \) equals the number \( b (i_1, \ldots, i_\ell) \). If the element \( g \) ranges over the entire group \( S_n \), then \( f (g) \) ranges over the values of the weights of all decompositions of the set \( \{1, 2, \ldots, n\} \) into ordered \( \ell \)-vectors. For even \( \ell \) the set of tensors \( b \in V^\otimes \ell \) for which the function \( f \) and the weight function \( \epsilon \) satisfy the condition of Proposition 1 form an open and dense set. Put \( B_t \subseteq V^\otimes \ell \), \( B_t (i_1, \ldots, i_\ell) = \exp \{ tb (i_1, \ldots, i_\ell) \} \). Then it is clear that \( S (f; t) = P (B_t) \). For \( \ell = 2 \) the result \( S (f; t) = Pf (B_t) \) is known in a different form [4, pp. 318-329]. The NP-hard problem of the existence of a decomposition into \( \ell \)-vectors from a given set \( \Sigma \) reduces polynomially in the sense of a probabilistic Turing machine [10] to the computation of the relative invariant \( P \). We put \( B (I) = 0 \), if \( I \not\subseteq \Sigma \) and \( B (I) = x_\ell \subseteq R \) if \( I \subseteq \Sigma \); \( I = \{i_1, \ldots, i_\ell\} \). The polynomial \( P (B) = P (x_\ell, I \subseteq \Sigma) \) is identically zero if and only if there does not exist a decomposition of the set \( \{1, 2, \ldots, n\} \) into \( \ell \)-vectors from \( \Sigma \) (\( \ell \) even). The test of the condition \( P (x_\ell, I \subseteq \Sigma) \neq 0 \) can be accomplished in polynomial time with arbitrarily small error probability if the values \( P (B) \) can be computed polynomially (cf. [10]). This is a strong indication that the problem of computing the invariant \( P \) is of high complexity.

In view of the results of section 2 it is natural to ask: for which tensors \( b \in V^\otimes l \) does the tensor \( B_t \subseteq R \) satisfy the condition \( B_t \leq r \) or \( \text{rank}_2 B_t \leq r \), and the value \( P (B) \) can be computed polynomially? We can list several types of conditions of combinatorial character.

1) Tensors with block structure. Assume that the index set is partitioned into \( r \) parallellepipeds \( \Pi = I_1 \times \ldots \times I_l \), \( I_l \subseteq \{1: n\} \), where for each of the parallelepipeds \( \Pi \):

a) there exist vectors \( u_1, \ldots, u_l \subseteq V \) (depending on \( \Pi \)) such that \( \forall I \subseteq \Pi, I = \{i_1, \ldots, i_l\} \)

\[
b (I) = u_1 (i_1) + u_2 (i_2) + \ldots + u_l (i_l) ;
\]

b) there exist matrices \( a_1, \ldots, a_p \subseteq V \otimes V \) (depending on \( \Pi \)) such that \( \forall I \subseteq \Pi, I = \{i_1, \ldots, i_l\} \)

\[
b (I) = a_1 (i_1, i_2) + a_2 (i_3, i_4) + \ldots + a_p (i_{l-1}, i_l) ;
\]
It is clear that in case a) \( \text{rank} B_t \leq r \) and in case b) \( \text{rank}_2 B_t \leq r \).

2) Tensors of fixed rank with given number of distinct coordinates. Let \( D \subseteq \mathbb{R} \) be the set of the coordinates \( b(i_1, \ldots, i_t) \) of the tensor \( b \). \( |D| = d \) (the smallest combinatorial case is evidently \( D = \{0,1\}, d = 2 \)) and \( \text{rank} b = c \). Then \( \forall t \subseteq \mathbb{R} \) there exists a polynomial \( q(z) \), \( \deg q \leq d - 1 \) such that \( B_t(I) = q(b(I)) \), \( I = \{i_1, \ldots, i_t\} \) (a corollary of Lagrange's interpolation theorem). Then one finds by direct calculation that \( \text{rank}_t B_t \leq \sum_{i=1}^{d-1} \binom{t+c-1}{i} \). If a representation of the tensor \( b \) in the form (1) is known then one can also find a representation of the tensor \( B_t \) in the form (1). An analogous formula holds also for the 2-rank of the tensors \( b \) and \( B_t \).

In the space of tensors 1) b) the tensors for which the function \( f \) and the weight function \( \varepsilon \) satisfy the condition of Proposition 1 form an open and dense set. The function \( f \) for tensors 1) a) does obviously not satisfy the condition of the theorem, but nonetheless, the limit \( \alpha = \lim_{t \to \infty} r^t \log |S_t(f; t)| \) has an obvious combinatorial meaning: in general, \( \alpha \) is the maximum of the values of \( f \) on those decompositions for which no \( 2t \)-vectors belong to one index of the parallelepiped \( \Pi \).

III. Problems with weighted coverings. Assume first that \( b \subseteq V^* \), \( \dim V = n \) and fix a number \( d \in \mathbb{N} \). We denote by \( \bar{a} \) the remainder \( a \mod (n) \), assuming that \( 1 \leq \bar{a} \leq n \). We define the function \( f : S_{nd} \to \mathbb{R} \) by

\[
f(g) = \sum_{i=0}^{nd/t} b(g(i+1), \ldots, g(i+t)).
\]

The value \( f(g) \) is a summation weight of a covering of the set \( \{1, 2, \ldots, n\} \) by \( nd/t \) ordered \( t \)-vectors (possibly with treated elements) where each element of the set belongs to exactly \( d \) vectors (allowing for multiplicities). If the element \( g \) ranges over the full group \( S_{nd} \), then the value \( f(g) \) ranges over the weights of all such coverings. Let \( \mu \subseteq R[S_{nd}] \) and \( \mu \subseteq I_d \). Then, if we put \( B_t(i_1, \ldots, i_t) = \exp \{tb(i_1, \ldots, i_t)\} \) we have \( S_{\mu}(f; t) = P_{\mu*}(B_t) \), \( \mu*(g) = \mu(g^t) \), \( \mu* \subseteq I_d \). As in example II there are some special cases where there exist algorithms for the computation of the sums \( S_{\mu}(f; t) \) for which the number of arithmetic operations and exponentiations is polynomial in \( n \).

**Remark.** If the function \( f \) is integer valued, then the sums \( S_{\mu}(f; t) \) also allow the computation of the multiplicity \( \text{mult}_{\mu}(f; c) = \sum_{(g)} \mu(g) \) of the values \( c \in \text{Im} f \) of the function \( f \) relative to the weight function \( \mu \) according to the obvious formula

\[
S_{\mu}(f; t) = \sum_{c \in \mathbb{Z}} \exp(tc) \text{mult}_{\mu}(f; c).
\]

In the particular examples above (example II, 1, 2) the resulting algorithms are pseudo-polynomial.

The author is grateful to A. M. Vershik for advice with this paper.

**LITERATURE CITED**