An Extension of Sylvester's Law of Inertia

R. RADO
The University
Reading, England

Communicated by L. Mirsky

---

1. Sylvester's law of inertia states that if a real quadratic form Q in independent variables $t_1, \ldots, t_n$ satisfies equations of the type

$$Q = a_1 x_1^2 + \cdots + a_n x_n^2 = b_1 y_1^2 + \cdots + b_n y_n^2,$$

where the a_ν and b_ν are real numbers and the x_ν and y_ν real linear forms in $t_1, \ldots, t_n$ such that the x_ν are linearly independent and the y_ν are linearly independent, then the number of indices ν such that $a_\nu > 0$ equals the number of ν such that $b_\nu > 0$, and similarly for $a_\nu < 0$ and for $a_\nu = 0$. In this note we prove an extension of this theorem in which, in analogy with the fact that the indices ν fall into three classes according to the sign of a_ν, we consider not two but three representations of Q. The method of proof closely resembles that used to prove the classical version of the theorem.

2. Let n be a positive integer and $t_1, \ldots, t_n$ independent indeterminates over the field of real numbers. The letters a, b, c denote real numbers and the letters x, y, z real linear forms in the t_ν. The vector space of all such linear forms is denoted by V_n. Independence means linear independence. For any given numbers $a_1, \ldots, a_n$ we define three sets

$$A^+ = \{\nu : a_\nu > 0\}, \quad A^- = \{\nu : a_\nu < 0\}, \quad A^0 = \{\nu : a_\nu = 0\},$$

and similarly B^+, B^-, B^0 in terms of $b_1, \ldots, b_n$, etc. For finite sets S the symbol $|S|$ denotes the number of elements of S. Thus

$$A^+ \cup A^- \cup A^0 = \{1, \ldots, n\}, \qquad |A^+| + |A^-| + |A^0| = n. \tag{1}$$

Linear Algebra and Its Applications **1, 29–31 (1968)**

THEOREM. *Suppose that* $x_1, \ldots, x_n$ *are independent, that* $y_1, \ldots, y_n$ *are independent, and that* $z_1, \ldots, z_n$ *are independent. Let*

$$\sum_1^n a_\nu x_\nu^2 = \sum_1^n b_\nu y_\nu^2 = \sum_1^n c_\nu z_\nu^2 = Q. \tag{2}$$

Then (i) *the forms* $x_\nu(\nu \in A^+)$, $y_\nu(\nu \in B^-)$, $z_\nu(\nu \in C^0)$ *constitute a basis of* V_n, *and* (ii) *the subspace of* V_n *generated by the forms* $x_\nu(\nu \in A^+ \cup A^-)$ *depends on* Q *only.*

Remark. Sylvester's law follows from the theorem. For, by (i),

$$|A^+| + |B^-| + |C^0| = |B^+| + |A^-| + |C^0| = n,$$

and by (ii), $|A^+| + |A^-| = |B^+| + |B^-|$. These relations imply that $|A^+| = |B^+|$ and $|A^-| = |B^-|$.

3. We need two simple lemmas which, for convenience, will be proved. In these we assume the hypotheses of the theorem.

LEMMA 1. *Suppose that the form* $\sum (1 \leqslant r, s \leqslant n) a_{rs} x_r x_s$, *where* $a_{rs} = a_{sr}$, *takes the value zero whenever the* t_ν *are replaced by arbitrary real numbers. Then* $a_{rs} = 0$ *for all* r, s.

Proof. By choosing the t_ν in such a way that $x_1 = 1$ and $x_\nu = 0$ for $\nu \geqslant 2$ we find $a_{11} = 0$, and, in case $n \geqslant 2$, by making $x_1 = x_2 = 1$ and $x_\nu = 0$ for $\nu \geqslant 3$ we find $2a_{12} = 0$. The lemma follows by symmetry.

LEMMA 2. *Let* $0 \leqslant k, l \leqslant n$; $a_r \neq 0$ *for* $1 \leqslant r \leqslant k$ *and* $b_s \neq 0$ *for* $1 \leqslant s \leqslant l$. *Let* $\sum (1 \leqslant r \leqslant k) a_r x_r^2 = \sum (1 \leqslant s \leqslant l) b_s y_s^2$. *Then* $k = l$, *and* $x_1, \ldots, x_k$ *generate the same space as* $y_1, \ldots, y_l$.

Proof. Let $y_s = \sum (1 \leqslant r \leqslant n) c_{sr} x_r$ $(1 \leqslant s \leqslant l)$. Then

$$\sum (1 \leqslant r \leqslant k) a_r x_r^2 = \sum (1 \leqslant s \leqslant l) b_s \sum (1 \leqslant \alpha, \beta \leqslant n) c_{s\alpha} c_{s\beta} x_\alpha x_\beta.$$

Let $k < r_0 \leqslant n$. Then, by Lemma 1,

$$\sum (1 \leqslant s \leqslant l) b_s c_{sr_0} c_{s\beta} = 0 \qquad \text{for} \quad 1 \leqslant \beta \leqslant n.$$

Hence

$$\sum (1 \leqslant s \leqslant l) b_s c_{sr_0} \sum (1 \leqslant \beta \leqslant n) c_{s\beta} x_\beta = 0,$$

i.e., $\sum (1 \leqslant s \leqslant l) b_s c_{sr_0} y_s = 0$. This implies that $c_{sr_0} = 0$ for $1 \leqslant s \leqslant l$, so that $y_s = \sum (1 \leqslant r \leqslant k) c_{sr} x_r$ for $1 \leqslant s \leqslant l$. Hence $l \leqslant k$, and the lemma follows by symmetry.

4. We now prove the theorem. Part (ii) follows from Lemma 2. Let the t_ν be replaced by any real numbers such that

$$x_\nu = 0 (\nu \in A^+), \qquad y_\nu = 0 (\nu \in B^-), \qquad z_\nu = 0 (\nu \in C^0). \tag{3}$$

Then, by (2), $Q \leqslant 0$ and $Q \geqslant 0$, so that $Q = 0$ and therefore $x_\nu = 0$ for $\nu \in A^-$. Now we conclude, by (ii), that $z_\nu = 0$ for $\nu \in C^+ \cup C^-$, so that $z_1 = \cdots = z_n = 0$. In view of the fact that the t_ν have arbitrary real values subject only to (3), it follows that the $x_\nu (\nu \in A^+)$ together with the $y_\nu (\nu \in B^-)$ and the $z_\nu (\nu \in C^0)$ generate V_n. Hence

$$|A^+| + |B^-| + |C^0| \geqslant n. \tag{4}$$

We now interchange in all six possible ways the roles played by the three pairs (a_ν, x_ν), (b_ν, y_ν), (c_ν, z_ν) and add the six inequalities corresponding to (4). In view of the symmetry of the procedure we obtain

$$2(|A^+| + |A^-| + |A^0| + |B^+| + |B^-| + |B^0| + |C^+| + |C^-| + |C^0|) \geqslant 6n. \tag{5}$$

However, by (1) there is equality in (5) and therefore also in (4). This proves (i) and completes the proof of the theorem.

5. We conclude with remarks on abstract extensions. Both lemmas remain true if the real field is replaced by any field whose characteristic is different from 2. They are false for every field of characteristic 2, as is shown by the equations $t_1^2 + t_2^2 + t_3^2 = (t_1 + t_2)^2 + t_3^2 = t_1^2 + (t_2 + t_3)^2$ which hold in such fields. The theorem holds in every ordered field.

Received April 2, 1967.