AN INEQUALITY FOR GENERALIZED QUADRANGLES

STANLEY E. PAYNE

Abstract. Let $S$ be a generalized quadrangle of order $(s, t)$. Let $X$ and $Y$ be disjoint sets of pairwise noncollinear points of $S$ such that each point of $X$ is collinear with each point of $Y$. If $m = |X|$ and $n = |Y|$, then $(m - 1)(n - 1) < s^2$. When equality holds, severe restrictions are placed on $m$, $n$, $s$, and $t$.

I. Prolegomena. A generalized quadrangle of order $(s, t)$, $s > 1$, $t > 1$, is a point-line incidence geometry $S = (\mathcal{P}, \mathcal{L}, I)$ with point set $\mathcal{P}$, line set $\mathcal{L}$, and symmetric point-line incidence relation $I$ satisfying the following axioms:

A1. No two points are incident with two lines in common.

A2. If $x$ is a point not incident with a line $L$, then there is a unique point $y$ incident with $L$ and collinear with $x$.

A3. Each line (respectively, point) is incident with $1 + s$ points (respectively, $1 + t$ lines).

Throughout this note $S = (\mathcal{P}, \mathcal{L}, I)$ will denote a generalized quadrangle (GQ) of order $(s, t)$, $s > 1$, $t > 1$. Let $X = \{x_1, \ldots, x_m\}$ and $Y = \{y_1, \ldots, y_n\}$ be disjoint sets of pairwise noncollinear points of $S$, $m \geq 2$ and $n \geq 2$. Let $k_i$ be the number of $x_i$'s with which $y_i$ is collinear, $1 \leq i \leq n$, $0 \leq k_i \leq m$. Our main results consist of the following two theorems.

Theorem 1.1.

$$(1 + s) \sum_{i=1}^{n} k_i < mn + \sqrt{m^2n^2 + (s^2 - 1)(m + n)mn + (s^2 - 1)^2 mn}.$$ 

Theorem 1.2. Let $k_i = m$ for all $i$, i.e. each $y_i$ is collinear with each $x_j$. Then $(m - 1)(n - 1) < s^2$. If equality holds, then one of the following must occur.

(i) $m = n = 1 + s$, and each point of $Z = \mathcal{P} \setminus (X \cup Y)$ is collinear with precisely two points of $X \cup Y$.

(ii) $m \neq n$. If $m < n$, then $s \mid t$, $s < t$, $n = 1 + t$, $m = 1 + s^2 / t$, and each point of $S$ is collinear with either $1$ or $1 + t/s$ points of $Y$ according as it is or is not collinear with some point of $X$. Note: $(m - 1) | s$.

There are two corollaries that deserve mention.

Corollary 1.3. If there is a GQ $S$ with order $(s, t)$, $s > 1$, then $t < s^2$. If $t = s^2$, then each triad of points has exactly $1 + s$ centers.

Received by the editors February 14, 1977.


Key words and phrases. Generalized quadrangles, Rayleigh quotient, incidence matrix.

© American Mathematical Society 1978
Proof. The inequality \( t \leq s^2 \) is due to D. G. Higman ([3], [4]). Alternate treatments appear in Bose [1] and Cameron [2]. In the present setting a proof is obtained by putting \( X = \{x_1, x_2\} \) where \( x_1 \) and \( x_2 \) are not collinear, \( Y = \text{tr}(X) \) = the set of 1 + \( t \) points collinear with both \( x_1 \) and \( x_2 \), and then applying Theorem I.2. □

**Corollary I.4.** Let \( x \) and \( y \) be noncollinear points of \( \mathcal{S} \) with \( s > 1 \) and \(|\text{sp}(x, y)| = 1 + p\). Then \( pt \leq s^2 \). If \( pt = s^2 \) and \( p < t \), then each point \( z \) collinear with no point of \( \text{sp}(x, y) \) must be collinear with exactly \( 1 + t/s \) points of \( \text{tr}(x, y) \).

**Proof.** For the original proof and an explanation of the notation see Thas [7]. In the present setting put \( X = \text{sp}(x, y) \), \( Y = \text{tr}(x, y) \). □

The proofs depend on a general matrix theoretic approach due to Sims. As the treatment in [5] does not include the “case of equality,” we first give an exposition of this method.

**II. A matrix-theoretic technique.** If \( \vec{x} = (x_1, \ldots, x_n)^T \) and \( \vec{y} = (y_1, \ldots, y_n)^T \) are column vectors of real numbers, then \( \vec{x} \cdot \vec{y} = \sum x_i y_i \) denotes their usual dot product. If \( A \) is a real, symmetric, \( n \times n \) matrix, then for each \( \vec{x} \neq 0 \) define the Rayleigh quotient \( R(\vec{x}) \) for \( A \) by

\[
R(\vec{x}) = \frac{\vec{x} \cdot A \vec{x}}{\vec{x} \cdot \vec{x}}. \tag{1}
\]

It is well known that \( A \) has real characteristic roots, say \( \mu_1 \leq \cdots \leq \mu_n \), and that

\[
\mu_1 = \min_{\vec{x}: \vec{x} \neq 0} R(\vec{x}) \leq \max_{\vec{x}: \vec{x} \neq 0} R(\vec{x}) = \mu_n. \tag{2}
\]

Perhaps not so well known is the following.

**II.1.** Let \( \vec{x} \) be a nonzero vector in \( \mathbb{R}^n \) for which \( R(\vec{x}) = \mu_i \) for either \( i = 1 \) or \( i = n \). Then \( \vec{x} \) is a characteristic vector of \( A \) belonging to the characteristic value \( \mu_i \).

**Proof.** Let \( \vec{x}_1, \ldots, \vec{x}_n \) be an orthonormal basis of characteristic vectors of \( A \) ordered so that \( A \vec{x}_i = \mu_i \vec{x}_i \). Let \( \vec{x} \) be an arbitrary nonzero vector of \( \mathbb{R}^n \) normalized so that \( \vec{x} \cdot \vec{x} = 1 \). Then \( R(\vec{x}) = \vec{x} \cdot A \vec{x} \) and \( \vec{x} = \sum c_i \vec{x}_i \) with \( \sum c_i^2 = 1 \). Hence \( \mu_1 = \mu_1 \cdot \sum c_i^2 \leq \sum c_i^2 \mu_i = \vec{x} \cdot A \vec{x} = R(\vec{x}) \), with equality holding if and only if \( \mu_i = \mu_i \) whenever \( c_i \neq 0 \). It follows that \( R(\vec{x}) = \mu_1 \) if and only if \( \vec{x} \) belongs to the eigenspace associated with \( \mu_1 \). The argument for \( \mu_n \) is similar. □

We continue to let \( A = (a_{ij}) \) denote an \( n \times n \) real symmetric matrix. Let \( \Delta = \Delta_1 + \cdots + \Delta_s \) and \( \Gamma = \Gamma_1 + \cdots + \Gamma_s \) be partitions of \( \{1, \ldots, n\} \). Suppose that \( \Gamma \) is a refinement of \( \Delta \), and write \( i \subseteq j \) whenever \( \Gamma_i \subseteq \Delta_j \), \( 1 \leq i \leq s \), \( 1 \leq j \leq r \). Put \( \delta_i = |\Delta_i|, \gamma_i = |\Gamma_i| \). Let

\[
\delta_{ij} = \sum_{\mu \in \Delta_i} a_{\mu \nu}, \quad \gamma_{ij} = \sum_{\mu \in \Gamma_i \wedge \nu \in \Gamma_j} a_{\mu \nu}.
\]
So $\delta_j = \delta_i$ and $\gamma_{ij} = \gamma_{ij}$ by the symmetry of $A$. Define the following matrices:

$$A^\Delta = \begin{pmatrix} \delta_{ij} \\ \delta_i \end{pmatrix}_{1 < i, j < r}, \quad A^\Gamma = \begin{pmatrix} \gamma_{ij} \\ \gamma_i \end{pmatrix}_{1 < i, j < s}.$$ 

$$\tilde{A}_\Delta = \text{diag}(\sqrt{\delta_1}, \ldots, \sqrt{\delta_r}); \quad \tilde{A}_\Gamma = \text{diag}(\sqrt{\gamma_1}, \ldots, \sqrt{\gamma_s}).$$

$$\hat{A}_\Delta = \tilde{A}_\Delta A^\Delta (\tilde{A}_\Delta)^{-1} = \begin{pmatrix} \delta_{ij} \\ \delta_i \delta_j \end{pmatrix}_{1 < i, j < r},$$

$$\hat{A}_\Gamma = \tilde{A}_\Gamma A^\Gamma (\tilde{A}_\Gamma)^{-1} = \begin{pmatrix} \gamma_{ij} \\ \gamma_i \gamma_j \end{pmatrix}_{1 < i, j < s}.$$  

Hence $\hat{A}_\Delta$ and $\hat{A}_\Gamma$ are real symmetric matrices with real characteristic values equal to those of $A^\Delta$ and $A^\Gamma$, respectively. The smallest and largest characteristic roots of $\hat{A}_\Gamma$ and $\hat{A}_\Delta$ are the minimum and maximum, respectively, of $(\tilde{x} \cdot \hat{A}_\Gamma \tilde{x})/(\tilde{x} \cdot \tilde{x})$ and $(\tilde{y} \cdot \hat{A}_\Delta \tilde{y})/(\tilde{y} \cdot \tilde{y})$, $\tilde{x} \neq \tilde{x} \in R^r$, $\tilde{y} \neq \tilde{y} \in R^s$.

Let $\tilde{y} = (y_1, \ldots, y_r)^T \in R^r$. Then put $\tilde{x} = (\ldots, x_a, \ldots)^T$, where $x_a = y_{i_a} \sqrt{\gamma_{i_a}/\delta_i}$ whenever $a \in i$, $1 \leq a \leq s$. Then

$$\sum_{a=1}^s x_a^2 = \sum_{i=1}^r \left( \sum_{a \in i} \left( y_{i_a} \sqrt{\gamma_{i_a}/\gamma_i} \right)^2 \right) = \sum_{i=1}^r \frac{y_i^2}{\delta_i} \left( \sum_{a \in i} \gamma_{i_a} \right) = \sum_{i=1}^r y_i^2,$$

implying $\tilde{x} \cdot \tilde{x} = \tilde{y} \cdot \tilde{y}$. And

$$\tilde{x} \cdot \hat{A}_\Gamma \tilde{x} = \sum_{a, \beta = 1}^s x_a \frac{\gamma_{a\beta}}{\sqrt{\gamma_a \gamma_\beta}} x_\beta = \sum_{i, j = 1}^r \left[ \sum_{a \in i, \beta \in j} \frac{\gamma_{a\beta}}{\sqrt{\gamma_a \gamma_\beta}} \cdot \frac{y_i \sqrt{\gamma_a}}{\sqrt{\delta_i}} \cdot \frac{y_j \sqrt{\gamma_\beta}}{\sqrt{\delta_j}} \right],$$

$$= \sum_{i, j = 1}^r y_i \left[ \sum_{\beta \in j} \frac{\gamma_{a\beta}}{\sqrt{\delta_i \delta_j}} \right] y_j = \tilde{y} \cdot \hat{A}_\Delta \tilde{y}.$$  

This implies that any value of $(\tilde{y} \cdot \hat{A}_\Delta \tilde{y})/(\tilde{y} \cdot \tilde{y})$ is also a value of $(\tilde{x} \cdot \hat{A}_\Gamma \tilde{x})/(\tilde{x} \cdot \tilde{x})$. Hence the following is a corollary of (2) and I.I.

II.2. If $\mu_1 \leq \cdots \leq \mu_r$ are the characteristic roots of $A^\Delta$ and $\lambda_1 \leq \cdots \leq \lambda_s$ are the characteristic roots of $A^\Gamma$, then $\lambda_1 < \mu_1 < \mu_s < \lambda_s$. If $\tilde{y} = (y_1, \ldots, y_r)^T$ satisfies $A^\Delta \tilde{y} = \lambda_1 \tilde{y}$ (so $\lambda_1 = \mu_1$), then $A^\Gamma \tilde{x} = \lambda_1 \tilde{x}$, where $\tilde{x} = (\ldots, x_a, \ldots)^T$ is defined by $x_a = y_{i_a}$ whenever $a \in i$. A similar result holds in case $\lambda_n = \mu_n$. 

License or copyright restrictions may apply to redistribution; see http://www.ams.org/journal-terms-of-use
PROOF. The first part of the result is evident. So let \( \tilde{0} \neq \tilde{y} = (y_1,\ldots,y_r)^T \) satisfy \( A_{\Delta} \tilde{y} = \lambda_1 \tilde{y} = \mu_1 \tilde{y} \). Then \( A_{\Delta} \tilde{y} = (y_1\sqrt{\delta_1},\ldots,y_r\sqrt{\delta_r})^T \) is a characteristic vector of \( \hat{A}_{\Delta} \) belonging to \( \lambda_1 = \mu_1 \). Hence \( \tilde{z} = (\ldots,z_a,\ldots)^T \), \( z_a = y_1\sqrt{\gamma_a} \) for \( a \subseteq i \), is a characteristic vector of \( \hat{A}_r \) belonging to \( \lambda_1 \) (by the proof of II.1). It follows that \( \tilde{z} \) as given in the statement of II.2 is a characteristic vector of \( A^\Gamma \) associated with \( \lambda_1 \). A similar proof holds in case \( \lambda_n = \mu_n \). □

III. Applications to generalized quadrangles. Let \( S = (\mathcal{P}, \mathcal{L}, I) \) be a GQ of order \((s, t)\). Let \( X \) and \( Y \) be as in the hypothesis of Theorem 1.1, and put \( Z = \mathcal{P} \setminus (X \cup Y) \), so \( |Z| = r = v - (m + n) \), where \( v = (1 + s)(1 + st) = |\mathcal{P}| \). For some ordering of \( \mathcal{P} \) let \( A \) be the \((0, 1)\)-matrix \( A = (a_{ij}) \) defined by \( a_{ij} = 1 \) if the \( i \)th and \( j \)th points of \( \mathcal{P} \) are not collinear in \( S \); \( a_{ij} = 0 \) otherwise. It follows that \( A \) is symmetric with minimum polynomial given by \( f(x) = (x + s)(x - t)(x - ts^2) \). Let \( \Delta = \Delta_1 + \Delta_2 + \Delta_3 \) be the partition of \( \{1, \ldots, v\} \) determined by \( X, Y, \) and \( Z \); i.e. points of \( X, Y, Z \), respectively, are indexed by \( \Delta_1, \Delta_2, \Delta_3 \), respectively. As \( \delta_i = |\Delta_i| \), we have \( \delta_1 = m, \delta_2 = n, \delta_3 = v - (m + n), \delta_{11} = n(n - 1), \delta_{12} = \delta_{21} = \sum_{k=1}^n (m - k)mn = \Sigma, \) where \( \Sigma = \sum_{k=1}^n k \). Since \( \sum_{j=1}^n (\delta_j/\delta_i) = ts^2 \), we also have \( \delta_{13} = \delta_{1}ts^2 - \delta_{12} - \delta_{11} = ts^2m - (mn - \Sigma) - m(m - 1) \). Similarly, \( \delta_{23} = ts^2n - (mn - \Sigma) - n(n - 1) \). Using these results it is now routine to complete the calculation of \( A^\Delta \).

\[
A^\Delta = \begin{bmatrix}
m - 1 & n - \Sigma/m & ts^2 + 1 - m - n + \Sigma/m \\
m - \Sigma/n & n - 1 & ts^2 + 1 - m - n + \Sigma/n \\
A_1 & A_2 & A_3
\end{bmatrix}
\]

where

\[
A_1 = \frac{m[ts^2 + 1 - m - n] + \Sigma}{v - m - n}, \quad A_2 = \frac{n[ts^2 + 1 - m - n] + \Sigma}{v - m - n}
\]

and

\[
A_3 = ts^2 - \frac{(m + n)[ts^2 + 1 - m - n] + 2\Sigma}{v - m - n}.
\]

Let \( (x - ts^2)(x - r_1)(x - r_2) \) be the characteristic polynomial of \( A^\Delta \) with the roots ordered so that \( r_1 < r_2 < ts^2 \). Let \( \Gamma = \Gamma_1 + \cdots + \Gamma_v \) be the identity partition of \( \{1, \ldots, v\} \), so \( \Gamma \) is a refinement of \( \Delta \). Then \( A^\Gamma = A \) has numerical range \([-s, ts^2]\) which must then contain all characteristic roots of \( A^\Delta \). Indeed, the proof of Theorem 1.1 amounts to calculating \( r_1 \) and using the inequality \(-s < r_1 \). We now proceed to do this.

Put \( (x - r_1)(x - r_2) = x^2 - bx + c \), so that \( 2r_1 = b - \sqrt{b^2 - 4c} \). Hence \(-s < r_1 \) simplifies to

\[
0 < s^2 + bs + c, \quad b = r_1 + r_2 = \text{tr}(A^\Delta) - ts^2, \quad c = \text{det}(A^\Delta)/ts^2. \tag{4}
\]

It is easy to calculate \( \text{tr}(A^\Delta) \) from (3) and then to write \( b \) as follows.
\[ b = \frac{(m + n)(s + st + 2) - 2v - 2\Sigma}{v - m - n}. \]  \hspace{1cm} (5)

To calculate \( \det(A^\Delta) \), add the first and second columns of \( A^\Delta \) to the third column and then subtract the first row from the second. At this point \( \det(A^\Delta) \) appears as follows.

\[
\begin{vmatrix}
  m - 1 & 1 - \Sigma/n & 0 \\
  m[ts^2 + 1 - m - n] + \Sigma & n[ts^2 + 1 - m - n] + \Sigma & 1 \\
  v - m - n & v - m - n & 1
\end{vmatrix}
\]  \hspace{1cm} (6)

Expanding by the third column and simplifying, one may calculate \( c \) to be as follows.

\[
c = \frac{\det(A^\Delta)}{ts^2} = \frac{(1 + s + st)(2\Sigma - m - n) + v - v\Sigma^2/mn}{v - m - n}. \]  \hspace{1cm} (7)

Using the values for \( b \) and \( c \) given in (5) and (7), (4) may be rewritten as follows.

\[ 0 < (s - 1)(m + n + s^2 - \lambda)mn + 2mn^2 - (1 + \lambda)^22. \]  \hspace{1cm} (8)

Equality in (8) gives two roots \( \Sigma_1 \) and \( \Sigma_2 \) for which (8) says \( \Sigma_1 < \Sigma < \Sigma_2 \), if \( \Sigma_1 < \Sigma_2 \). But \( \Sigma_2 \) is easily evaluated.

\[
\Sigma_2 = \frac{mn + \sqrt{m^2n^2 + (s^2 - 1)(m + n)mn + (s^2 - 1)^2mn}}{1 + s}. \]  \hspace{1cm} (9)

Clearly \( \Sigma < \Sigma_2 \) is just the inequality in Theorem 1.1. If each \( k_i = m \), then \( \Sigma = mn \), and the inequality of Theorem 1.1 reduces to \( (m - 1)(m - 1) < s^2 \), the inequality of Theorem 1.2.

We now use II.2 to investigate the case of equality in Theorem 1.2. Suppose that \( k_i = m \) for all \( i \), so \( \Sigma = mn \), and suppose that \( (m - 1)(n - 1) = s^2 \), so \( -s \) is a characteristic root of \( A^\Delta \). Hence a nonzero characteristic vector of \( A^\Delta \) belonging to \( -s \) must span the null space of \( A^\Delta + sI \).

\[
A^\Delta + sI = \begin{pmatrix}
  m - 1 + s & 0 & ts^2 + 1 - m \\
  0 & n - 1 + s & ts^2 + 1 - n \\
  * & * & *
\end{pmatrix}
\]  \hspace{1cm} (10)

where we need not bother to calculate the third row, since the rank must equal 2. Clearly \( \vec{y} = (y_1, y_2, 1)^T \) spans the null space of \( A^\Delta + sI \), where

\[
y_1 = \frac{m - 1 - ts^2}{s + m - 1}; \hspace{1cm} y_2 = \frac{n - 1 - ts^2}{s + n - 1}. \]  \hspace{1cm} (11)

Let us assume that the points of \( \mathcal{P} \) are ordered (for the construction of \( A \)) so that the first \( m \) points are those of \( X \), the next \( n \) points are those of \( Y \), and the last \( v - m - n \) points are those of \( Z \). Then by II.2, \( \vec{x} \) must be a characteristic vector of \( A^\Gamma = A \) belonging to \( \lambda_1 = -s \), where \( \vec{x} \) is as follows.
\[ \bar{x} = \begin{pmatrix} y_1, \ldots, y_1, & y_2, \ldots, y_2, & 1, \ldots, 1 \\ m \times & n \times & (v - m - n \times) \end{pmatrix}^T. \tag{12} \]

For the first \( m + n \) rows of \( A \) this yields no new information. But let \( z \in Z \) be the \( i \)th point, \( i > m + n \). Suppose \( z \) is not collinear with \( t_1 \) points of \( X \), is not collinear with \( t_2 \) points of \( Y \), and hence is not collinear with \( ts^2 - t_1 - t_2 \) points of \( Z \). Then the product of the \( i \)th row of \( A \) with \( \bar{x} \), which must equal \(-s\), is actually \( t_1y_1 + t_2y_2 + ts^2 - t_1 - t_2 = s \). After a little simplification this becomes

\[ \frac{t_1}{s + m - 1} + \frac{t_2}{s + n - 1} = 1. \tag{13} \]

If \( z \) lies on a line joining a point of \( X \) and a point of \( Y \), then \( t_1 = m - 1 \) and \( t_2 = n - 1 \), i.e., since \( S \) has no triangles, \( z \) is collinear with a unique point of \( X \) and with a unique point of \( Y \). On the other hand, if \( z \) is not on such a line either \( t_1 = m \) or \( t_2 = n \). Suppose \( t_1 = m \), so \( z \) is collinear with no point \( X \). Using (13) we find that the number of points of \( Y \) collinear with \( z \) is

\[ n - t_2 = 1 + (n - 1)/s. \tag{14} \]

Similarly, any point of \( Y \) collinear with no point of \( Y \) must be collinear with \( 1 + (m - 1)/s \) points of \( X \). If \( m = n = s + 1 \), this says each point not on a line joining a point of \( X \) with a point of \( Y \) must be collinear with two points of \( X \) and none of \( Y \) or with two of \( Y \) and none of \( X \). If \( 1 < m < s + 1 \), so \( 1 + (m - 1)/s \) is not an integer, then each point of \( S \) is collinear with some point of \( Y \). This implies that each point \( z \) of \( Z \) is either on a line joining points of \( X \) and \( Y \) or is collinear with \( 1 + (n - 1)/s \) points of \( Y \). Clearly \( n < 1 + t \). Suppose \( n < 1 + t \) and let \( x_1 \in X \). Then there is some line \( L \) through \( x_1 \) not incident with any point of \( Y \). But then any point \( z \) on \( L \), \( z \neq x_1 \), cannot be collinear with any point of \( Y \), a contradiction. Hence it must be that \( n = 1 + t \), from which it follows that \( m = 1 + s^2/t \). This essentially completes the proof of Theorem 1.2.

A similar treatment is available for the restriction on the parameters of a subquadrangle, a combinatorial proof of which is found in [6].

**REFERENCES**

1. R. C. Bose, *Geometric and pseudo-geometric graphs \( (q^2 + 1, q + 1, 1) \)*, J. Geometry 2 (1972), 75–93.
7. , *On generalized quadrangles with parameters \( s = q^2 \) and \( t = q^3 \)*, Geometria Dedicata 5 (1976), 485–496.

**DEPARTMENT OF MATHEMATICS, MIAMI UNIVERSITY, OXFORD, OHIO 45056**