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 $v$ incident with $L$ and collinear with $x$.
A3. Each line (respectively, point) is incident with $1 + t$ points (respectively, $1 + s$ 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 $S$, $m > 2$ and $n > 2$. Let $k_i$ be the number of $x_i$'s with which $y_i$ is collinear, $1 < i < n$, $0 < k_i < m$. Our main results consist of the following two theorems.

Theorem I.1.

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

Theorem I.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: $(m - 1) \mid s$.

There are two corollaries that deserve mention.

Corollary I.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.
Proof. The inequality \( t < s^2 \) 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 1.2. □

Corollary 1.4. Let \( x \) and \( y \) be noncollinear points of \( S \) with \( s > 1 \) and \(|\text{sp}(x, y)| = 1 + p\). Then \( pt < s^2 \). 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 \( \vec{x} \neq \vec{0} \) define the Rayleigh quotient \( R(\vec{x}) \) for \( A \) by

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

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

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

Perhaps not so well known is the following.

II.1. Let \( \vec{x} \) be a nonzero vector in \( R^n \) 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 \( R^n \) normalized so that \( \vec{x} \cdot \vec{x} = 1 \). Then \( R(\vec{x}) = \vec{x} \cdot A \vec{x} = \sum c_i \vec{x}_i \) with \( \sum c_i^2 = 1 \). Hence \( \mu_1 = \mu_1 \cdot \sum c_i^2 < \sum c_i^2 \mu_i = \vec{x} \cdot A \vec{x} = R(\vec{x}) \), with equality holding if and only if \( \mu_i = \mu_i \) whenever \( c_i \neq 0 \). It follows that \( R(\vec{x}) = \mu_i \) if and only if \( \vec{x} \) belongs to the eigenspace associated with \( \mu_i \). 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 \( \Gamma = \Gamma_1 + \cdots + \Gamma_r \) 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

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

$$A^A = \begin{pmatrix} \delta_j \\ \delta_i \end{pmatrix}_{1 \leq i, j \leq r}; \quad A^\Gamma = \begin{pmatrix} \gamma_{ij} \\ \gamma_{ji} \end{pmatrix}_{1 \leq i, j \leq s}.$$

$$A_\Delta = \text{diag}\left(\sqrt{\delta_1}, \ldots, \sqrt{\delta_r}\right); \quad A_\Gamma = \text{diag}\left(\sqrt{\gamma_1}, \ldots, \sqrt{\gamma_s}\right).$$

$$\hat{A}_\Delta = A_\Delta A^A (A_\Delta)^{-1} = \begin{pmatrix} \delta_j \\ \sqrt{\delta_i \delta_j} \end{pmatrix}_{1 \leq i, j \leq r}.$$

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

Hence $\hat{A}_\Delta$ and $\hat{A}_\Gamma$ are real symmetric matrices with real characteristic values equal to those of $A^A$ 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 $(\bar{x} \cdot \hat{A}_\Gamma \bar{x})/(\bar{x} \cdot \bar{x})$ and $(\bar{y} \cdot \hat{A}_\Delta \bar{y})/(\bar{y} \cdot \bar{y})$, $\bar{y} \neq \bar{x} \in R^r$, $\bar{0} \neq \bar{y} \in R^r$.

Let $0 \neq \bar{y} = (y_1, \ldots, y_r)^T \in R^r$. Then put $\bar{x} = (\ldots, x_a, \ldots)^T$, where $x_a = \gamma_{ai}/\delta_i$ whenever $a \in I$, $1 \leq a \leq s$. Then

$$\sum_{a=1}^{s} x_a^2 = \sum_{i=1}^{r} \left( \sum_{a \in I} \left( \gamma_{ai}/\gamma_{ii} \right)^2 \right) = \sum_{i=1}^{r} \frac{y_i^2}{\delta_i} \left( \sum_{a \in I} \gamma_{ai} \right) = \sum_{i=1}^{r} y_i^2,$$

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

$$\bar{x} \cdot \hat{A}_\Gamma \bar{x} = \sum_{a, \beta = 1}^{s} x_a \frac{\gamma_{a \beta}}{\sqrt{\gamma_a \gamma_\beta}} x_\beta$$

$$= \sum_{i, j = 1}^{r} \left( \sum_{a \in I} \frac{\gamma_{a \beta}}{\sqrt{\gamma_a \gamma_\beta}} \frac{y_i \gamma_{ai}}{\sqrt{\delta_i}} \frac{y_j \gamma_{aj}}{\sqrt{\delta_j}} \right)$$

$$= \sum_{i, j = 1}^{r} \frac{y_i}{\delta_i} \left( \sum_{a \in I} \frac{\gamma_{a \beta}}{\sqrt{\delta_a \delta_\beta}} \right) y_j$$

$$= \sum_{i, j = 1}^{r} y_i \left( \frac{\delta_j}{\sqrt{\delta_i \delta_j}} \right) y_j = \bar{y} \cdot \hat{A}_\Delta \bar{y}.$$

This implies that any value of $(\bar{y} \cdot \hat{A}_\Delta \bar{y})/(\bar{y} \cdot \bar{y})$ is also a value of $(\bar{x} \cdot \hat{A}_\Gamma \bar{x})/(\bar{x} \cdot \bar{x})$. Hence the following is a corollary of (2) and II.1.

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

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use
Proof. The first part of the result is evident. So let $\overline{0} \neq \overline{y} = (y_1, \ldots, y_r)^T$ satisfy $A^\Delta \overline{y} = \lambda_1 \overline{y} = \mu_1 \overline{y}$. Then $A^\Delta \overline{y} = (y_1\sqrt{\delta_1}, \ldots, y_r\sqrt{\delta_r})^T$ is a characteristic vector of $A^\Delta$ belonging to $\lambda_1 = \mu_1$. Hence $\overline{z} = (\ldots, z_a, \ldots)^T$, $z_a = y_i\sqrt{\gamma_a}$ for $a \subseteq i$, is a characteristic vector of $A^\Gamma$ belonging to $\lambda_1$ (by the proof of II.1). It follows that $\overline{z}$ 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 $S = (\mathcal{P}, \mathcal{E}, I)$ 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 $\delta_1 = m, \delta_2 = n, \delta_3 = v - (m + n)$, $\delta_{i_1} = n(n - 1), \delta_{i_2} = \delta_{i_2} = \Sigma_{j=1}^n(m - k_j)mn = \Sigma$, where $\Sigma = \Sigma_{j=1}^n k_j$. Since $\Sigma_{j=1}^n(\delta_j / \delta_i) = ts^2$, we also have $\delta_{i_3} = \delta_1 ts^2 - \delta_{i_2} - \delta_{i_1} = ts^2 m - (mn - \Sigma) - m(m - 1)$. Similarly, $\delta_{i_3} = ts^2 n - (mn - \Sigma) - n(n - 1)$. Using these results it is now routine to complete the calculation of $A^\Delta$.

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

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 $\Gamma = \Gamma_1 + \cdots + \Gamma_e$ 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

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

It is easy to calculate $\text{tr}(A^\Delta)$ from (3) and then to write $b$ as follows.
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.

$$\det(A^\Delta) = \left| \begin{array}{ccc} m - 1 & n - \Sigma/m & 1 \\ 1 - \Sigma/n & \Sigma/m - 1 & 0 \\ m[s^2 + 1 - m - n] + \Sigma & n[s^2 + 1 - m - n] + \Sigma & 1 \\ v - m - n \\ v - m - n \end{array} \right|. \quad (6)$$

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

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

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

$$0 < (s - l)(m + n + s^2 - 1)mn + 2mn^2 - (1 + s)^2. \quad (8)$$

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.

$$\Sigma_2 = \frac{mn + \sqrt{m^2n^2 + (s^2 - 1)(m + n)mn + (s^2 - 1)^2 mn}}{1 + s}. \quad (9)$$

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 $A^\Delta + sI$.

$$A^\Delta + sI = \begin{bmatrix} m - 1 + s & 0 & ts^2 + 1 - m \\ 0 & n - 1 + s & ts^2 + 1 - n \\ * & * & * \end{bmatrix}, \quad (10)$$

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

$$y_1 = \frac{m - 1 - ts^2}{s + m - 1}; \quad y_2 = \frac{n - 1 - ts^2}{s + n - 1}. \quad (11)$$

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.
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 \( x \), which must equal \(-s\), is actually \( t_1 y_1 + t_2 y_2 + ts^2 - t_1 - t_2 = s \). 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 \( O \) 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 \( O \) 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 \( 1 + (n - 1)/s > 3 \) 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

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use