On a generalization of Hadamard's theorem on determinants*

Otto Szász

Monatshefte für Mathematik und Physik 28.1 (1917) pp. 253-257

Let

$$D = \sum \pm a_{1,1}\, a_{2,2} \cdots a_{n,n}$$

be a determinant of order n, whose elements are any complex numbers; let us denote the conjugate value of the number $a_{i,k}$ by $\bar{a}_{i,k}$. Then,

$$D^2 = D\bar{D} = \sum \pm c_{1,1}\, c_{2,2} \cdots c_{n,n},$$

in which

$$c_{i,k} = a_{i,1}\bar{a}_{k,1} + \cdots + a_{i,n}\bar{a}_{k,n}.$$

J. Hadamard[1] derived an upper bound for the determinant by proving that

$$\sum \pm c_{1,1}\, c_{2,2} \cdots c_{n,n} \leq c_{1,1}\, c_{2,2} \cdots c_{n,n}. \quad \text{(I)}$$

Hadamard proved the following to derive this theorem

$$D\bar{D} \leq \left(\sum \pm c_{1,1}\, c_{2,2} \cdots c_{n-1,n-1}\right) c_{n,n}. \quad \text{(II)}$$

Since Hadamard's theorem plays a role in the theory of integral equations, it may not be superfluous to provide a proof of the theorem that is based only on the most elementary properties of the determinants; this can be found in §1. In §2 I derive a generalization of the theorem (II) given by E. Fischer, in a shorter way, using a theorem of Weierstrass related to continuous functions[2].

*This work is a translation from the Hungarian of Az Hadamard-féle determinanstétel egy elemi bebizonyitása, Mathematikai és Physikai Lapok, XIX, 1910, p. 221-227; however 3 which is added here appeared in the "Math. és Phys. Lapok" under the title: "Egy determinimstételrol".

[1]Bulletin des sciences mathématiques (2). XVII (1893). Résolution d'une question relative aux déterminants.

[2]For other references, see Wirtinger, Monatshefte f. Math. u. Phys. (1907). E. Fischer, Archiv der Math. u. Phys., Bd. 13 (1908)

I recently found that E.J. Nanson [Note from the translator: Edward John Nanson (1850-1936)] also provided evidence of the result (I)[3] He obtained the same result by proving the following theorem.

Let

$$\Delta = [b_{i,k}]_1^n$$

be a symmetric determinant, of which all principal minors are positive, then the following inequality applies

$$\Delta \leq b_{1,1} \cdots b_{n,n}.$$

This result only refers to determinants with real elements. However, it is easy to extend Nanson's view to Hermite determinants. This theorem is no more general than theorem (I), because the square matrix $(b_{i,k})$ can (under the given conditions) be factored as a product of a matrix $(u_{i,s})$ with the complex conjugate $(\bar{u}_{i,s})$, such that

$$b_{i,k} = \sum_{s=1}^{N} u_{i,s}\bar{u}_{k,s}, \quad (i,k = 1,2,\ldots,n),(N \leq n).$$

See Fischer, loc. cit., p. 34, point 2.

In §3 I give a simple proof of Fischer's theorem (relation III of Fischer's note) by generalizing Nanson's reasoning.

1.[4]

I transform the given determinant - without changing its value - to the form:

$$D = \sum \pm\alpha_{1,1} \cdots \alpha_{n,n},$$

so that the following relationships should exist between its elements,

$$\sum_{j=1}^{n} \alpha_{i,j}\bar{\alpha}_{k,j} = 0,$$

[3]Messenger of Mathematics 31 (1901), A determinant inequality. Nanson noted that Lord Kelvin proved this result in 1886 and communicated it to T. Muir, and later this result appeared in "The Educational Times". From a kind message from Mr. Nanson I learnt that Muir published the result in 1901 in "The Educational Times" as "Question 14792".

[4]Recently, T. Boggio (Bulletin des Sciences Mathématiques, 1911) has published a proof of the theorem that differs little from my proof. Mr. T. Boggio was so friendly, writing his regret to say that he had no knowledge of my note, which was published in Hungarian. His note contains further references. [Note from the translator: Tommaso Boggio (1877-1963) was an Italian mathematician.]

or in short terms

$$\gamma_{i,k} = 0,$$

for

$$\begin{pmatrix} i = 1, 2, \ldots, n \\ k = 1, 2, \ldots, n \\ i \neq k \end{pmatrix}$$

which is expressed so that the rows are orthogonal to each other.[5]

Then,

$$D\bar{D} = \gamma_{1,1} \cdots \gamma_{n,n}. \qquad (\gamma)$$

The transformation is achieved in the following steps:

I leave the first row of the determinant unchanged; so it gives

$$\alpha_{1,j} = a_{1,j}, \; j = 1, \ldots, n.$$

The second row becomes orthogonal by placing the following in its place,

$$\alpha_{2,j} = a_{2,j} + x_{1,1}\,\alpha_{1,j}, \; j = 1, \ldots, n.$$

where the still undetermined number $x_{1,1}$ is found by the following equation,

$$\sum_{j=1}^{n} (a_{2,j} + x_{1,1}\,\alpha_{1,j})\,\bar{\alpha}_{1,j} = 0.$$

Now the third line becomes orthogonal to the previous ones by placing,

$$\alpha_{3,j} = a_{3,j} + x_{1,2}\,\alpha_{1,j} + x_{2,2}\,\alpha_{2,j}, \; j = 1, \ldots, n,$$

where $x_{1,2}$ and $x_{2,2}$ are determined through the equations,

$$\sum_{j=1}^{n} (a_{3,j} + x_{1,2}\,\alpha_{1,j})\,\bar{\alpha}_{1,j} = 0,$$

$$\sum_{j=1}^{n} (a_{3,j} + x_{2,2}\,\alpha_{2,j})\,\bar{\alpha}_{2,j} = 0.$$

I continue this until the last row becomes orthogonal to the previous ones.

This transformation is clear because the x's can be successively calculated from linear equations, unless the coefficient of any x vanishes, but this would mean that for a value of the index i,

$$\alpha_{i,1} = 0, \; \alpha_{i,2} = 0, \ldots, \alpha_{i,n} = 0.$$

[5]If $\alpha_{i,1}, \ldots, \alpha_{i,n}$ are real and the coordinates of a vector in n-dimensional space, then the equations mean that the vectors are orthogonal to each other.

We can exclude this case from our considerations, because it gives $D = 0$, and the theorem is trivial.

This leads to Hadamard's theorem since the individual steps do not increase the product $c_{1,1} \cdots c_{n,n}$, but at most leave it unchanged. Because the first row remained unchanged,

$$c_{1,1} = \gamma_{1,1};$$

changing the second row only changes the factor $c_{2,2}$; namely $x_{1,1}$ is determined such that the expression

$$\phi(x) = \sum_{j=1}^{n} (a_{2,j} + x\,\alpha_{1,j})\,(\bar{a}_{2,j} + \bar{x}\,\bar{\alpha}_{1,j})$$

for $x = x_{1,1}$ becomes minimum; $\phi(x)$ namely differs from the product

$$\begin{aligned}\psi(x) \quad = \quad & \left[x(\sum_{j=1}^{n} \alpha_{1,j}\bar{\alpha}_{1,j})^{\frac{1}{2}} + \frac{\sum_{j=1}^{n} \bar{\alpha}_{1,j} a_{2,j}}{(\sum_{j=1}^{n} \alpha_{1,j}\bar{\alpha}_{1,j})^{\frac{1}{2}}} \right] \cdot \\ & \left[\bar{x}(\sum_{j=1}^{n} \alpha_{1,j}\bar{\alpha}_{1,j})^{\frac{1}{2}} + \frac{\sum_{j=1}^{n} \alpha_{1,j} \bar{a}_{2,j}}{(\sum_{j=1}^{n} \alpha_{1,j}\bar{\alpha}_{1,j})^{\frac{1}{2}}} \right]\end{aligned}$$

only by an additive constant, so instead of the minimum of $\phi(x)$ we can consider that of $\psi(x)$; $\psi(x)$ never becomes negative, so if it takes the value 0, then it is the minimum, but this is obtained for $x = x_{1,1}$. Since $c_{2,2} = \phi(0)$ and $\gamma_{2,2} = \phi(x_{1,1})$ it is clear that

$$c_{2,2} \geq \gamma_{2,2}.$$

A similar remark applies to the change in the third row; we have seen that for the change from the row $a_{2,i} + x\alpha_{1,i}\,(i = 1, \ldots, n)$ to the row $\alpha_{1,i}\,(i = 1, \ldots, n)$ to obtain orthogonality, the value of x has to be such that the expression $\sum_i (a_{2,i} + x\alpha_{1,i})(\bar{a}_{2,i} + \bar{x}\bar{\alpha}_{1,i})$ becomes the minimum – mutatis mutandis– one now uses the value of x_1 for which the row

$$a_{3,i} + x_1\alpha_{1,i} + x_2\alpha_{2,i} \;\; (i = 1, \ldots, n)$$

becomes orthogonal to the row $\alpha_{1,i}\,(i = 1, \ldots, n)$, at the same time the expression $\sum_i (a_{3,i} + x_1\alpha_{1,i} + x_2\alpha_{2,i})(\bar{a}_{3,i} + \bar{x}_1\bar{\alpha}_{1,i} + \bar{x}_2\bar{\alpha}_{2,i})$ is a minimum (x_2 being for the moment an undefined parameter). This is done for $x_1 = x_{1,2}$, and further for the value of x_2 for which the row $a_{3,i} + x_{1,2}\alpha_{1,i} + x_2\alpha_{2,i}\,(i = 1, \ldots, n)$ becomes orthogonal to the row $\alpha_{2,i}\,(i = 1, \ldots, n)$, at the same time the expression

$$\sum_i (a_{3,i} + x_{1,2}\alpha_{1,i} + x_2\alpha_{2,i})\,(\bar{a}_{3,i} + \bar{x}_{1,2}\bar{\alpha}_{1,i} + \bar{x}_2\bar{\alpha}_{2,i})$$

is a minimum.

It is thus proven that

$$c_{3,3} \geq \gamma_{3,3}.$$

It is clear that a similar remark applies to the remaining steps, so further

$$c_{4,4} \geq \gamma_{4,4}, \ \ldots, \ c_{n,n} \geq \gamma_{n,n}.$$

From these relations, combined with the equation (γ), it follows that

$$c_{1,1} \cdots c_{n,n} \geq D\bar{D},$$

is proven.

It is clear that in (I) the equal sign applies if any of the $c_{i,i}$ disappears (i.e. if $a_{i,1} = 0, \ldots, a_{i,n} = 0$), or if $c_{i,k} = 0$ for each i and k ($i \neq k$); but only in these cases, because if none of these conditions is met, I can proceed to another determinant D in the way specified, so that the upper limit decreases, so it could not already be the minimum.

Using the same method of proof, the theorem can be derived for any matrix.

2.

The generalization of Theorem (II) given by Fischer is as follows; let

$$M = \begin{Vmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{Vmatrix}, \quad m \leq n$$

be any matrix, then

$$M\bar{M} \leq \left(\sum \pm c_{1,1} \cdots c_{\rho,\rho}\right)\left(\sum \pm c_{\rho+1,\rho+1} \cdots c_{m,m}\right) \quad (M)$$

or, in short,

$$M\bar{M} \leq \Gamma(a_{i,k}),$$

ρ being one of the numbers $1, 2, \ldots, m-1$. Here

$$M\bar{M} = \left(\sum \pm c_{1,1} \cdots c_{m,m}\right).$$

I distinguish three cases:

α) All ρ-th order minors, formed from the ρ first columns, vanish; in this case the result is trivial ($M\bar{M} = 0$).

β) All minors of ρ-th order formed from the first ρ columns vanish, with the exception of the minor occurring in Γ; in this case the equal sign applies in relation (M).

γ) In addition to the minor $\sum \pm c_{1,1} \cdots c_{\rho,\rho}$, there is at least one minor formed from the first ρ columns that does not vanish[6].

In this case, I can assume, without restricting generality, that for a k that is larger than ρ,

$$\sum \pm c_{k,1} \cdots c_{\rho,\rho} \neq 0.$$

Now the elements $a_{i,k}$ of the matrix are variable, so that

$$\sum \pm c_{1,1} \cdots c_{m,m} = K \quad (F)$$

is constant and different from zero, which means that the case α) is excluded.

So I consider a $mn - 1$ dimensional set of matrices in which the matrix M is contained. Since Γ is a continuous function of $a_{i,k}$ that never becomes negative, it has a minimum in condition (F). Let this minimum be achieved with the elements of the matrix

$$M_u = \left\| \begin{matrix} u_{1,1} & \cdots & u_{1,n} \\ \vdots & & \vdots \\ u_{m,1} & \cdots & u_{m,n} \end{matrix} \right\|,$$

if I can prove that the case β) occurs for this matrix, our result has already been reached, because then it is

$$\Gamma_{\min} = \Gamma(u_{i,k}) = M_u \bar{M}_u = K,$$

so

$$\Gamma(a_{i,k}) \geq K,$$

or

$$\Gamma(a_{i,k}) \leq M\bar{M}.$$

I prove that M_u satisfies β) like this: I assume the opposite, then, according to the case γ)

$$\sum \pm z_{1,1} \cdots z_{\rho,\rho} \neq 0$$

where

$$z_{i,k} = \sum_{j=1}^{n} u_{i,j} \bar{u}_{k,j},$$

and I show that I can determine x so that

$$M'_u = \left\| \begin{matrix} u_{1,1} + x\, u_{k,1} & \cdots & u_{1,n} + x\, u_{k,n} \\ u_{2,1} & \cdots & u_{2,n} \\ \vdots & & \vdots \\ u_{m,1} & \cdots & u_{m,n} \end{matrix} \right\|$$

[6] A fourth case is impossible.

belonging to $\Gamma(M'_u)$ becomes smaller than $\Gamma(M_u)$; however, this is a contradiction, because $\Gamma(M_u)$ is the minimum according to the requirement[7].

$\Gamma(M'_u)$ and $\Gamma(M_u)$ differ only in their first factor; this is in $\Gamma(M_u)$:

$$\sum \pm z_{1,1}\cdots z_{\rho,\rho} = \begin{vmatrix} z_{1,1} & \cdots & z_{1,\rho} \\ \vdots & & \vdots \\ z_{\rho,1} & \cdots & z_{\rho,\rho} \end{vmatrix} = \left\| \begin{matrix} u_{1,1} & \cdots & u_{1,n} \\ \vdots & & \vdots \\ u_{\rho,1} & \cdots & u_{\rho,n} \end{matrix} \right\| \cdot \left\| \begin{matrix} \bar{u}_{1,1} & \cdots & \bar{u}_{1,n} \\ \vdots & & \vdots \\ \bar{u}_{\rho,1} & \cdots & \bar{u}_{\rho,n} \end{matrix} \right\|$$

and the corresponding factor in $\Gamma(M'_u)$ is:

$$\left\| \begin{matrix} u_{1,1}+xu_{k,1} & \cdots & u_{1,n}+xu_{k,n} \\ u_{2,1} & \cdots & u_{2,n} \\ \vdots & & \vdots \\ u_{\rho,1} & \cdots & u_{\rho,n} \end{matrix} \right\| \cdot \left\| \begin{matrix} \bar{u}_{1,1}+\bar{x}\bar{u}_{k,1} & \cdots & \bar{u}_{1,n}+\bar{x}\bar{u}_{k,n} \\ \bar{u}_{2,1} & \cdots & \bar{u}_{2,n} \\ \vdots & & \vdots \\ \bar{u}_{\rho,1} & \cdots & \bar{u}_{\rho,n} \end{matrix} \right\|,$$

after the multiplication is carried out, this product becomes

$$\begin{aligned} \sum \pm z_{1,1}\cdots z_{\rho,\rho} \quad &+ \quad x\sum \pm z_{k,1}\, z_{2,2}\cdots z_{\rho,\rho} + \bar{x}\sum \pm \bar{z}_{k,1}\, \bar{z}_{2,2}\cdots \bar{z}_{\rho,\rho} \\ &+ \quad x\bar{x}\sum \pm z_{k,k}\, z_{2,2}\cdots z_{\rho,\rho} \end{aligned}$$

and this can obviously be made smaller than $\sum \pm z_{1,1}\cdots z_{\rho,\rho}$ by selecting the x appropriately, because the coefficient of x is - according to the requirement - different from zero. Theorem (M) is thus fully proven.

It is clear that in cases α) and β) in (M) the equality applies; but only in these cases, because in the case γ) as we saw, the value of Γ is not the smallest[8]

By repeatedly applying the result to the individual factors of Γ, I get that

$$\begin{aligned} M\bar{M} \quad &\leq \quad (\sum \pm c_{1,1}\cdots c_{\rho_1,\rho_1})\,(\sum \pm c_{\rho_1+1,\rho_1+1}\cdots c_{\rho_2,\rho_2})\cdots \\ &\cdots \quad (\sum \pm c_{\rho_k+1,\rho_k+1}\cdots c_{m,m}) \end{aligned}$$

$$(\rho_1 < \rho_2 < \cdots < \rho_k < m)$$

This relation contains all preceding results as special cases.

[7] It is clear that $M_u\bar{M}_u = K$.

[8] I can only briefly note afterwards that the case β) only occurs if $[c_{i,k}]_1^\rho \neq 0$ and $c_{i,k} = 0$ for $i = \rho+1,\ldots,m$; $k = 1,\ldots,\rho$ (also specified by Fischer); this results from the fact that the determinant of the system of equations

$$\sum_{j=1}^{\rho} c_{i,j}D_{k,j} = 0, \quad (k = 1,\ldots,\rho,\ i > \rho),$$

does not vanish. Here $D_{i,k}$ means the minor corresponding to $c_{i,k}$ in $[c_{i,k}]_1^\rho$.

3.

We prove the following theorem:

If every principal minor of an Hermite's determinant $\Delta = [b_{i,k}]_1^n$ is positive, then

$$\sum \pm b_{1,1}\cdots b_{n,n} \le (\sum \pm b_{1,1}\cdots b_{\rho-1,\rho-1})\,(\sum \pm b_{\rho,\rho}\cdots b_{n,n}).$$

The theorem is obviously true for $n = 1, 2$. I assume it is true for the orders that are less than n and prove that it then also applies to the order n.

Apparently the adjoint determinant of Δ

$$\sum \pm B_{1,1}\cdots B_{n,n}$$

is also an Hermite determinant; according to a theorem of Jacobi,

$$\sum \pm B_{1,1}\cdots B_{\rho,\rho} = \Delta^{\rho-1}\sum \pm b_{\rho+1,\rho+1}\cdots b_{n,n},$$

so every principal minor of the adjoint determinant is positive. Now the hypothesis is

$$\sum \pm B_{1,1}\cdots B_{\rho,\rho} \le B_{\rho,\rho}\sum \pm B_{1,1}\cdots B_{\rho-1,\rho-1} \quad (2)$$

$$B_{\rho,\rho} \le (\sum \pm b_{1,1}\cdots b_{\rho-1,\rho-1})\,(\sum \pm b_{\rho+1,\rho+1}\cdots b_{n,n}) \quad (3)$$

From relations (1), (2) and (3) it now follows that

$$\Delta^{\rho-1} \le (\sum \pm b_{1,1}\cdots b_{\rho-1,\rho-1})\,(\sum \pm B_{1,1}\cdots B_{\rho-1,\rho-1}).$$

Since

$$\sum \pm B_{1,1}\cdots B_{\rho-1,\rho-1} = \Delta^{\rho-2}(\sum \pm b_{\rho,\rho}\cdots b_{n,n})$$

one has[9]

$$\Delta \le (\sum \pm b_{1,1}\cdots b_{\rho-1,\rho-1})\,(\sum \pm b_{\rho,\rho}\cdots b_{n,n}) \quad (4)$$

The result can be somewhat generalized because of the following result:

If all the principal minors of the determinant Δ are positive and if there are numbers $b_1,\ldots,b_n$ such that

$$\left[b_{i,k}\,\frac{b_i}{b_k}\right]_1^n$$

becomes an Hermite determinant, then

$$\sum \pm b_{1,1}\cdots b_{n,n} \le (\sum \pm b_{1,1}\cdots b_{\rho-1,\rho-1})\,(\sum \pm b_{\rho,\rho}\cdots b_{n,n}).$$

[9]Here, too, it is easy to show that equality applies to non-vanishing Δ in (4) if and only if $b_{i,k} = 0$ for $i = \rho,\ldots,n$. This is proved by induction but I cannot show it because of lack of space.

The theorem follows from the fact that $b_{i,k}$ and $\left[b_{i,k} \begin{matrix} b_i \\ b_k \end{matrix}\right]_1^n$ match in all of their principal minors.

Remark: Let Δ be an orthogonal determinant, then

$$\Delta = 1 = \prod_{i=1}^{n} \left(\sum_{j=1}^{n} b_{i,j} \bar{b}_{i,j} \right)^{\frac{1}{2}}$$

and on the other hand

$$1 = \prod_{i=1}^{n} b_{i,i},$$

therefore,

$$b_{i,i} = 1, \quad b_{i,k} = 0, \; i \neq k.$$

In other words: The only positively definite Hermite form of n variables whose discriminant is an orthogonal determinant is $x_1 \bar{x}_1 + \cdots + x_n \bar{x}_n$.