103. On the General Schwarzian Lemma.

By Masao SUGAWARA.

Tokyo Bunrika Daigaku, Tokyo.

(Comm. by T. TAKAGI, M.I.A., Dec. 12, 1941.)

We aim to generalize the Schwarzian lemma in the theory of functions of one variable to the case of the higher dimension and apply it to the characterization of the displacements of the general Poincaré-space.

Definition 1. Let $W^{(m,n)}=(w_{ij})$ and $Z^{(m,n)}=(z_{ij})$ be two matrices of (m, n) type. We call W *an analytic function* $f(Z)$ of Z, if the elements w_{ij} of W are analytic functions of the elements z_{ij} of Z, it is called *regular at a point* Z, if w_{ij} are regular functions of the elements z_{ij} in a neighbourhood of Z, and it is called *regular in a domain* D of Z, if it is regular at every point of D.

Definition 2. Let $A^{(m,n)}$ be a matrix and $\mathfrak{x}$ be a n-dimensional vector with the length 1. Put $|A|=\underset{|\mathfrak{x}|=1}{\mathrm{l.\,u.\,b.}}|A\mathfrak{x}|$.

We call $|A|$ *the absolute value* of A or *the norm* of A.

Definition 3. A function W of Z is called *partially constant*, if there exist two constant unitary matrices U, V of the dimension m and n resp. such that $UWV=(a_{ij})$ in which a_{ii}, $i=1, 2, \dots, r$ are all constant and all the elements $a_{ij}=0$, where $i\neq j$ and $i\leqq r$ or $j\leqq r$.

Theorem 1. When $f(Z)$ is regular in a closed domain D of Z, the maximum absolute value of $f(Z)$ is taken on the boundary of D. It is taken also at inner points of D, when and only when W is partially constant.

Definition 4. A matrix is said to be of a *D-form*, if it is of the form $(a_{ij}\delta_{ij})$ in which δ_{ij} means the Kronecker's symbol; namely $\delta_{ij}=1$, if $i=j$, and $=0$ in other cases.

Definition 5. A matrix B of a D-form is said to be a *N-form* of a matrix A, if there exist two constant unitary matrices U and V such that $A=UBV$.

Proof of the theorem 1. Let $Z_0=(z_{ij}^0)$ be an inner point of D at which $f(Z)$ takes its maximum absolute value. We can evidently assume that $f(Z_0)\neq 0$. Take two constant unitary matrices U and V such that $Uf(Z_0)V$ is a D-form $(z_i\delta_{ij})$ and $|f(Z_0)|=|z_1|$.

Put $f_1(Z)=Uf(Z)V=(x_{ij})$, then $f_1(Z)$ is also a regular analytic function of Z in D such that $|f_1(Z)|=|f(Z)|$ and $|x_{11}^0|=|z_1|$, x_{11}^0 being the value of x_{11} at the point Z_0. As Z_0 is an inner point of D we can find a domain $F=(Z; |z_{ij}-z_{ij}^0|\leqq\varepsilon)$ in D, if we take ε sufficiently small. As x_{11} takes its maximum absolute value on the boundary of F and as $|f(Z_0)|=|x_{11}^0|\geqq|f_1(Z)|\geqq|x_{11}|$, x_{11} is a constant. Hence $x_{i1}=0$ $(i=2, \dots, m)$, as $|f(Z)|^2\geqq\sum_{k=1}|x_{k1}|^2$ and $x_{1j}=0$ $(j=2, \dots, n)$, because $|f(Z)|^4\geqq|x_{11}|^2(|x_{11}|^2+|x_{12}|^2+\cdots+|x_{1n}|^2)$; the right-hand side of the last inequality being the square of the length of the 1st column-vector of $\bar{W}'W$. By the same reason we have $x_{ij}=x_{ji}=0$ $i\neq j$ for all values of j and for all values of i for which $|z_i|=|z_1|$ hold. Hence W is partially constant, if it takes its maximum absolute value at an inner point of D

Theorem 2. (*The general Schwarzian lemma*)

Let $f(Z)$ be such a regular analytic function of Z in the domain $R=(Z; |Z|\leqq 1)$, that $f(0)=0$ and $|f(Z)|\leqq 1$, when $|Z|=1$, then

1°. *$|f(Z)|\leqq|Z|$ at every point of the domain R;*

2°. *If $|f(Z)|=|Z|$ at every point of a neighbourhood S of one inner point Z_0, it has the form*

$$f(Z)=UZV \quad \text{or} \quad UZ'V,$$

where U and V are two constant unitary matrices and the last case takes place only when Z is a square matrix, i. e. $m=n$.

Proof of 1°[1]. Let Z_0 be a point of the domain R and put $Z=tZ_0$, t being a complex variable, then the elements x_{ij} of the function $f_1(Z)$ introduced in the proof of the theorem 1 are regular analytic functions of t. Moreover $x_{11}=0$ when $t=0$ and $|x_{11}|\leqq|f(Z)|\leqq 1$, when $|tZ_0|=1$. By the Schwarzian lemma in the case of one variable we have $|x_{11}|\leqq|Z_0t|$, when $|Z_0t|\leqq 1$. Hence we get $|x_{11}^0|\leqq|Z_0|$, if we put $t=1$.

2°[2]. Let Z_0 be an inner point such that $|f(Z)|=|Z|$ at every point of a neighbourhood S of Z_0. Let Z_1 be a N-form of Z_0, namely $Z_0=U_0Z_1V_0$, where U_0 and V_0 are constant unitary matrices.

Put $f^*(Z)=f(U_0ZV_0)$, then $f^*(Z)$ is a regular analytic function of Z in R. Let $Z_1^*=(z_i^*\delta_{ij})$ be a point near to Z_1 such that $1>|z_1^*|>|z_2^*|>\cdots>|z_n^*|>0$. Make a function $f_1^*(Z)=(x_{ij})$ from $f^*(Z)$ at Z_1^* as we did $f_1(Z)$ from $f(Z)$ at Z.

We take a neighbourhood S' of Z_1^* such that $U_0S'V_0\subset S$ and consider the behavior of the function $f_1^*(Z)$ in S'. Let $Z=(z_i^*t\delta_{ij})$, t being a complex variable, then $|x_{11}|$ is equal to $|Z_1^*|=|z_1^*|$ when $t=1$ owing to the assumption, and $|x_{11}|\leqq 1$ when $|z_1^*t|=1$. $x_{11}=0$ when $t=0$ as in the proof of 1°, so we have $x_{11}=e_1z_1^*t$ by the Schwarzian lemma in the case of one variable, e_1 being a constant of the absolute value 1. Take a fixed value of t such that $Z_1^*t\in S'$. Then $|Z|=|z_1^*t|=|f_1^*(Z)|=|x_{11}|$, so we have $x_{1j}=x_{i1}=0$; $i,j\geqq 2$ as in the proof of the theorem 1. Now we vary Z in the domain $G=(Z\colon z_{11}=z_1^*t,\ z_{i1}=z_{1j}=0$ $i,j\geqq 2,\ |z_{ij}-tz_i^*\delta_{ij}|\leqq\eta$ $i,j\geqq 2)$, where η is so small that $|Z|=|z_1^*t|$ holds. As x_{11} is regular analytic, its maximum absolute value is taken on the boundary of G. On the other hand we have $|f_1^*(Z)|=|z_1^*t|$. Hence x_{11} is constant and equals to $e_1z_1^*t$ and $x_{i1}=x_{1j}=0$ $i,j\geqq 2$. Hence we have more generally $x_{11}=e_2z_{11}$, e_2 being a constant of the absolute value 1 and $x_{i1}=x_{1j}=0$ $i,j\geqq 2$ at every point Z of R at which $z_{i1}=z_{1j}=0$ $i,j\geqq 2$ hold.

---

1) See K. Morita Analytical characterization of the displacements of in a general Poincaré's space.

2) Here we assume that $m\geqq n$. If otherwise, we take transposed matrices.

Next we consider the function $f_1^*(Z)$ in a neighbourhood of a point $Z_2=(z_i'\delta_{ij})$ in which $|z_1'|$ is suitably small, while $z_i'=z_i^*$, $i\geqq 2$. Let $(x_i'\delta_{ij})$ be a N-form of the matrix $g(Z)=(x_{ij})$ $(i=2,\dots,m,\ j=2,\dots,n)$ at Z_2, namely $U_2 g(Z_2)V_2=(x_i'\delta_{ij})$, where U_2, V_2 are constant unitary matrices of the dimention $m-1$ and $n-1$ resp. and $|x_2'|\geqq|x_3'|\geqq\cdots\geqq|x_n'|$.

Multiplying the matrix $\begin{pmatrix}1 & 0\\ 0 & U_2\end{pmatrix}$ to $f_1^*(Z)$ by left and $\begin{pmatrix}1 & 0\\ 0 & V_2\end{pmatrix}$ by right, we have a regular function $f_2^*(Z)=(y_{ij})$ in R such that $y_{i1}=y_{1j}=0$, $i,j\geqq 2$ and $y_{11}=e_2 z_{11}$, if $z_{i1}=z_{1j}=0$ $i,j\geqq 2$. We consider the function $f_2^*(Z)$ in the domain $(Z\colon z_{22}=\text{constant } z_2^*,\ z_{i2}=z_{2j}=0\ i,j\neq 2,\ |z_{ij}-\delta_{ij}z_i'|\leqq\eta,\ i,j\neq 2)$, η being so small taken that $|Z|=|z_2^*|$. Then $|f_2^*(Z)|=|z_2^*|$. If we take y_{22} and z_2^* instead of x_{11} and z_1^* resp., we see that $y_{i2}=y_{2j}=0$ $i,j\neq 2$ $y_{22}=z_{22}e_3$, e_3 being a constant of the absolute value 1, if $z_{i2}=z_{2j}=0$, $i,j\neq 2$; and so on.

So we have

Lemma 1. There exists a regular analytic function $\hat{f}(Z)$ *in* R *such that* $\hat{f}(Z)=U^*f(U_0ZV_0)V^*=(v_{ij})$, *where* U^*, V^*, U_0, V_0 *are constant unitary matrices,* $v_{ii}=\varepsilon_i z_{ii}$, ε_i *being a constant of the absolute value 1, and* $v_{ik}=v_{ji}=0$ $(k,j\neq i\ \ i=1,2,\dots,n)$, *if* $z_{ik}=z_{ji}=0$, $k,j\neq i$.

Moreover

Lemma 2 We can find two unitary matrices U and V such that $f(Z)=UZV$ *at every point* Z_0 *of* R.

$$\text{Proof.}\quad f(Z_0)=f^*(Z_1)=U^{*-1}\hat{f}(Z_1)V^{*-1}=U^{*-1}U^{**}Z_1V^{**}V^{*-1}$$
$$=U^{*-1}U^{**}U_0^{-1}Z_0V_0^{-1}V^{**}V^{*-1}=UZ_0V,$$

where U and V are unitary.

From the lemmas 1 and 2 we get

Lemma 3. The function $f(Z)=(w_{ij})$ *is linear homogeneous in* Z, *namely* w_{ij} *are linear homogeneous functions of the elements* z_{ij} *of* Z.

Proof. As $f(Z)$ is regular analytic, it is sufficient to prove that $f(tZ)=tf(Z)$ holds, t being any complex number, at every point Z_0 of R. We have indeed

$$f(tZ_0)=f^*(tZ_1)=U^{*-1}\hat{f}(tZ_1)V^{*-1}=tU^{*-1}\hat{f}(Z_1)V^{*-1}$$
$$=tf^*(Z_1)=tf(Z_0)\,.$$

From the lemmas 1 and 3, we have at once

Lemma 4. v_{kk} *is a linear homogeneous function of* z_{ik} and z_{kj} $(i=1,\dots,m,\ j=1,\dots,n)$ *and the absolute value of the coefficient of* z_{kk} *is 1, while* v_{kj}, v_{ik}, $i,j\neq k$ *are linear homogenous functions of the variables* z_{kj} *and* z_{ik}, $i,j\neq k$, *only.*

We proceed to determine these linear functions.

To study the 1st column or row of $\hat{f}(z)$ we may put $z_{ij}=0$, $i,j\geqq 2$. Put $Z_1^{(0)}=z_{11}e_{11}$, $Z_1^{(1)}=\sum_{i=1}^{m}z_{i1}e_{i1}$, $Z_1^{(2)}=\sum_{j=1}^{n}z_{1j}e_{1j}$[1], then $\hat{f}(Z_1^{(0)})=\varepsilon_1 z_{11}e_{11}$. For

1) e_{ij} means a matrix whose (i,j)-component is 1 and the other components are all zero.

$\hat{f}(Z_1^{(1)})$ two cases are possible, namely

$$(1)\quad \hat{f}(Z_1^{(1)})=\sum_{i=1}^{m} v_{i1}(Z_1^{(1)})e_{i1} \qquad \text{or} \qquad (2)\quad \hat{f}(Z_1^{(1)})=\sum_{j=1}^{n} v_{1j}(Z_1^{(1)})e_{1j}\,,$$

because the rank of $Z_1^{(1)}$ is 1, thus the rank of $\hat{f}(Z_1^{(1)})$ is also 1 by the lemma 2 and z_{11} is contained only in v_{11}.

Here $v_{k1}(Z_1^{(1)})$ (or $v_{1k}(Z_1^{(1)})$) is a constant multiple of z_{k1} in (1) (or in (2)) $k\geqq 2$, with the condition

$$(3)\quad \sum_{i=1}^{m} |z_{i1}|^2=\sum_{i=1}^{m} |v_{i1}(Z_1^{(1)})|^2 \text{ in (1), (or } \sum_{j=1}^{n} |v_{1j}(Z_1^{(1)})|^2 \text{ in (2))}\,.$$

As we assume that $n\leqq m$, (2) can occur only when $m=n$. In the case (1)

$$(4)\quad \begin{pmatrix} v_{11}(Z_1^{(1)}) \\ v_{21}(Z_1^{(1)}) \\ \vdots \\ v_{m1}(Z_1^{(1)}) \end{pmatrix} = \begin{pmatrix} c_1 a_1 a_2 \cdots a_{m-1} \\ 0\; c_2\; 0 \qquad 0 \\ \cdots\cdots\cdots\cdots \\ 0\;0\;0 \qquad c_m \end{pmatrix} \begin{pmatrix} z_{11} \\ z_{21} \\ \vdots \\ z_{m1} \end{pmatrix} = U_1 \begin{pmatrix} z_{11} \\ z_{21} \\ \vdots \\ z_{m1} \end{pmatrix}.$$

Such a transformation U_1 is constant and unitary by (3). It follows that $a_i=0$, $|c_i|=1$, namely U_1 is a diagonal unitary form $(c_i\delta_{ij})$. If we consider transposed matrices $Z_1^{(2)\prime}$ and $\hat{f}(Z_1^{(2)})'$ instead of $Z_1^{(1)}$ and $\hat{f}(Z_1^{(1)})$, we see that $\hat{f}(Z_1^{(2)})$ has the form $\sum_{j=1}^{n} v_{1j}(Z_1^{(2)})e_{1j}$, because if it take the form $\sum_{i=1}^{m} v_{i1}(Z_1^{(2)})e_{i1}$, the rank of $\hat{f}(Z_1^{(1)}+Z_1^{(2)}-Z_1^{(0)})=\hat{f}(Z_1^{(1)})+\hat{f}(Z_1^{(2)})-\hat{f}(Z_1^{(0)})$ is 1, while that of $Z_1^{(1)}+Z_1^{(2)}-Z_1^{(0)}$ is 2. Moreover (5) $\big(v_{11}(Z_1^{(2)}), v_{12}(Z_1^{(2)}), \dots, v_{1n}(Z_1^{(2)})\big)=(z_{11}, z_{12}, \dots, z_{1n})V_1$, where V_1 means a constant unitary diagonal matrix $(d_i\delta_{ij})$ by the same reason as before. As $Z_1^{(1)}+Z_1^{(2)}-Z_1^{(0)}=\begin{pmatrix} z_{11}z_{12}\cdots z_{1n} \\ z_{21} \\ \vdots \qquad 0 \\ z_{m1} \end{pmatrix}$ and as $v_{11}=\varepsilon_1 z_{11}$, we can omit $Z_1^{(1)}$ and $Z_1^{(2)}$ in the formula (4) and (5), because $\hat{f}(Z_1^{(1)}+Z_1^{(2)}-Z_1^{(0)})=\hat{f}(Z_1^{(1)})+\hat{f}(Z_1^{(2)})-\hat{f}(Z_1^{(0)})=\begin{pmatrix} v_{11}v_{12}\cdots v_{1n} \\ v_{21} \\ \vdots \qquad 0 \\ v_{m1} \end{pmatrix}$. The same circumstances also hold about other columns and rows. In these cases, the case (1) is only possible if it hold already about first column (or row), because v_{1i} really contains z_{1i}.

Therefore if we put $Z_i^{(1)}=\sum_{k=1}^{m} z_{ki}e_{ki}$, $Z_i^{(2)}=\sum_{k=1}^{n} z_{ik}e_{ik}$, we have

$$\begin{pmatrix} v_{1i} \\ v_{2i} \\ \vdots \\ v_{mi} \end{pmatrix} = U_i \begin{pmatrix} z_{1i} \\ z_{2i} \\ \vdots \\ z_{mi} \end{pmatrix},\quad (v_{i1}, v_{i2}, \dots, v_{in})=(z_{i1}, z_{i2}, \dots, z_{in})V_i\,,$$

where U_i and V_i are constant unitary diagonal matrices.

We assume that $\hat{f}(Z)$ is so taken that $v_{11}=z_{11}$, which can be obtained by multiplication ε_1^{-1} to the former $\hat{f}$, then $c_1=1$ in U_1 and $d_1=1$ in V_1. Thus we get

$$U_1^{-1}\hat{f}(Z)V_1^{-1}=\begin{pmatrix} z_{11}\, z_{12} & \cdots & z_{1n} \\ z_{21}\, p_{22} & \cdots & p_{2n} \\ \cdot \quad \cdot & & \cdot \\ z_{m1} p_{m2} & \cdots & p_{mn} \end{pmatrix}$$

in which $p_{ij}=\varepsilon_{ij}z_{ij}$, ε_{ij} being constant of the absolute value 1. Now put $Z=z_{11}e_{11}+z_{i1}e_{i1}+z_{1j}e_{1j}+z_{ij}e_{ij}$, then $U_1^{-1}\hat{f}(Z)V_1^{-1}=z_{11}e_{11}+z_{i1}e_{i1}+z_{1j}e_{1j}+\varepsilon_{ij}z_{ij}e_{ij}$. As the rank of Z is equal to that of $U_1^{-1}\hat{f}(Z)V_1^{-1}$, det. $\begin{vmatrix} z_{11} & z_{1j} \\ z_{i1} & \varepsilon_{ij}z_{ij} \end{vmatrix}=0$, whenever det. $\begin{vmatrix} z_{11} & z_{1j} \\ z_{i1} & z_{ij} \end{vmatrix}=0$. This is possible only when $\varepsilon_{ij}=1$. Thus we get $\hat{f}(Z)=U_1ZV_1$. Hence $f(Z)=U^{*-1}\hat{f}(U_0^{-1}ZV_0^{-1})V^{*-1}=U^{*-1}U_1U_0^{-1}ZV_0^{-1}V_1V^{*-1}=UZV$, where U and V are constant and unitary.

When (2) takes place, we consider Z and $\hat{f}(Z)'$ instead of Z and $\hat{f}(Z)$ resp. and it leads to the result.

$$f(Z)=UZ'V.$$

Now we restrict ourselves to the case of symmetrical matrices, that is to say when Z and W are all symmetric, the circumstances remain almost the same as the above case and the slight modification will show that

$$f(Z)=UZU'$$

The main difference of the proof is as follows. Let $z_{ij}=0$, $i,j\geqq 2$, then $v_{ij}=0$ $i\neq j$, $i,j\geqq 2$, because v_{ij} is a linear combination of z_{1i} and z_{i1} on one side and that of z_{1j} and z_{j1} on the other side.

If $v_{kk}\neq 0$ for some $k\geqq 2$, the other columns (or rows), are linear combinations of the 1st and the kth column (or row), because the rank of Z is 2 in this case, thus the rank of $\hat{f}(Z)$ also 2 by the lemma 2. But they do not contain z_{11}, so they are constant multiples of the kth column (or row). Hence $\hat{f}(Z)$ has the form in which $v_{kk}\neq 0$ and the other elements v_{ij} are all zero except v_{11}, v_{1k}, v_{k1}. If we take z_{1k}, $\neq 0$ and we chose z_{11} such that the minor det. $\begin{vmatrix} v_{11} & v_{1k} \\ v_{k1} & v_{kk} \end{vmatrix}=0$, the rank of $\hat{f}(Z)$ will become 1, while the rank of Z is 2. This is a contradiction. Hence $v_{kk}=0$ for all values $k\geqq 2$.

Hence we have

$$\hat{f}(Z)=\begin{pmatrix} v_{11}v_{12} \cdots v_{1m} \\ v_{21} \\ \vdots \qquad 0 \\ v_{m1} \end{pmatrix}, \quad \text{if} \quad Z=\begin{pmatrix} z_{11}z_{12} \cdots z_{1m} \\ z_{21} \\ \vdots \qquad 0 \\ z_{m1} \end{pmatrix}.$$

where v_{1k}, $k\geqq 2$, is a constant multiple of z_{1k} with the condition

$$|v_{11}|^2+2\sum_{k=1}^{m}|v_{1k}|^2=|z_{11}|^2+2\sum_{k=1}^{m}|z_{1k}|^2,$$

From here we get the result as we did in the case of non-symmetry, taking $V=U'$.

As an application of the general Schwarzian lemma, we note here a theorem due to K. Morita[1], namely.

The one to one analytical mapping of the space R onto itself is the displacement or transposition or their combination;

$$f(Z)=(U_1Z+U_2)(U_3Z+U_4)^{-1}, \qquad f(Z)=Z',$$

where $$U=\begin{pmatrix}U_1U_2\\U_3U_4\end{pmatrix}, \quad U'S\bar{U}=S, \quad S=\begin{pmatrix}E_m & 0\\ 0 & -E_n\end{pmatrix}.$$

In the case of symmetrical matrices we must add a condition of symmetry

$$U'JU=J, \qquad J=\begin{pmatrix}0 & E_n\\ -E_n & 0\end{pmatrix}.$$

Indeed such a one to one analytical transformation $f(Z)$ that $f(0)=0$ is a regular analytic function of Z with the inverse regular function, so that from the theorem 2, we get at once $|W|=|Z|$ for all values of Z of R. Thus it is a function of our class.

1) K. Morita, loc. cit. Theorem 3.