Schwarzian derivatives and uniformization

Takeshi SASAKI and Masaaki YOSHIDA

In every textbook on elementary function theory you can find a definition of Schwarzian derivative

$$\{z;x\} = \frac{1}{2}\left(\frac{z''}{z'}\right)' - \frac{1}{4}\left(\frac{z''}{z'}\right)^2,$$

where $' = d/dx$, of a non-constant function $z = z(x)$ with respect to x. You would also find an exercise to show

(0.1) [PGL(2)-invariance] If a, b, c, and d are constants satisfying $ad - bc \neq 0$, then

$$\left\{\frac{az+b}{cz+d}; x\right\} = \{z;x\}.$$

(0.2) $\{z;x\} = 0$ if and only if $z(x) = (ax+b)/(cx+d)$ for some constants a, b, c, and d satisfying $ad - bc \neq 0$.

(0.3) [change of variable] If y is a non-constant function of x, then

$$\{z;y\} = \{z;x\}\left(\frac{dx}{dy}\right)^2 + \{x;y\}.$$

(0.4) [local behavior] If $z = x^\alpha u$ $(\alpha \neq 0)$, where u is a holomorphic function of x non-vanishing at 0, then

$$\{z;x\} = \frac{1-\alpha^2}{4x^2} + \frac{\text{a function holomorphic at 0}}{x};$$

if $z = \log(xu)$, where u is as above, then

$$\{z;x\} = \frac{1}{4x^2} + \frac{\text{a function holomorphic at 0}}{x}.$$

In this paper we discuss various generalizations of the Schwarzian derivative. We start from recalling how it was found.

Department of Mathematics, Kobe University, Kobe 657 Japan.
Department of Mathematics, Kyushu University, Fukuoka 812 Japan.

1. A paper by H. A. Schwarz

Hermann Amandus Schwarz (1843–1921) is well known through Schwarz inequality, Schwarz's lemma in elementary function theory, reflection principle, etc. In his paper [**Sch**], he treated the hypergeometric differential equation

$$\frac{d^2u}{dx^2}+\frac{\gamma-(\alpha+\beta+1)x}{x(1-x)}\frac{du}{dx}-\frac{\alpha\beta}{x(1-x)}u=0,$$

and studied conditions for the parameters (α,β,γ) that every solution is algebraic, and found explicit expressions of such solutions. Let us leave aside why he got interested in such a problem, and follow his line.

Consider in general a 2nd order linear differential equation

$$\frac{d^2u}{dx^2}+p\frac{du}{dx}+qu=0. \tag{1}$$

Let u_1 and u_2 be two linearly independent solutions, and put $z=u_2/u_1$. One of his main discoveries is the relation

$$\{z;x\}=q-\frac{1}{4}p^2-\frac{1}{2}\frac{dp}{dx} \tag{2}$$

between z and the coefficients of the equation. He proved this by a staightforward computation: We can assume $u_2=zu_1$. Substitute $u_2'=zu_1'+z'u_1$ and $u_2''=zu_1''+2z'u_1'+z''u_1$ into $u_2''+pu_2'+qu_2=0$, and use the identity $u_1''+pu_1'+qu_1=0$. We are led to $2z'u_1'+(z''+pz')u_1=0$, that is,

$$p+\frac{z''}{z'}=-2\frac{u_1'}{u_1}.$$

Differentiate both sides and we have

$$\begin{aligned} p'+\left(\frac{z''}{z'}\right)' &= -2\frac{u_1u_1''-(u_1')^2}{u_1^2}=-2\frac{u_1(-pu_1'-qu_1)-(u_1')^2}{u_1^2} \\ &= 2q+2\left(p\frac{u_1'}{u_1}+\left(\frac{u_1'}{u_1}\right)^2\right). \end{aligned}$$

Since u_1'/u_1 is already expressed in terms of p and z, substituting this expression, we obtain the above relation.

For the hypergeometric equation, the relation above appears to be

$$\{z;x\}=\frac{1-\lambda^2}{4x^2}+\frac{1-\mu^2}{4(1-x)^2}+\frac{1+\nu^2-\lambda^2-\mu^2}{4x(1-x)}, \tag{3}$$

where

$$\lambda=1-\gamma,\quad \mu=\gamma-\alpha-\beta,\quad \nu=\alpha-\beta.$$

On the other hand, the Wronskian $u_2du_1/dx-u_1du_2/dx$ of the solutions u_1 and u_2 is equal to a constant times $\exp(-\int pdx)$, which, in the hypergeometric case, turns out to be a constant multiple of $x^{-\gamma}(1-x)^{\gamma-\alpha-\beta-1}$. Thus if both u_1 and u_2 are algebraic, then γ and $\alpha+\beta$ must be rational numbers. Applying this argument after the change of variable $x\to 1/x$, we know that all of α, β, and γ must be rational numbers.

He next considers a map

$$s: x\mapsto z(x)=u_2(x)/u_1(x),$$

which is nowadays called Schwarz's s-map, and the s-image of the upper x-plane, which turns out to be a triangle bounded by three arcs, is called Schwarz's triangle. The local behavior at $x = 0$ and $x = 1$ can be known by (3) and the property (0.4) of the Schwarzian derivative (the change of variable $x \to 1/x$ leads also the behavior at $x = \infty$); this tells you the three angles of the triangle. Let us sketch his strategy: Possible analytic continuation of s can be illustrated by repeated reflection images along the arcs of the triangle. Since u_1 and u_2 are algebraic, so is s; thus these reflection images cannot be chaotic. This would determine possible three angles of the triangle and so the values of the three parameters.

There are two essential points in his paper: the formula (3), which is the origin of the name Schwarzian derivative, and the map s. In the following we study these by considering their generalizations. A survey of Shwarzian derivatives can be found also in [**MSY2**].

2. Normal forms of differential equations

Let us consider a system of linear differential equations in n variables x with m unknowns u with rank (dimension of the solution space at a generic point) r. Two such systems are said to be *strongly equivalent* if a change of unknowns $u_j \to \sum_i k^i_j u_i$ takes one to the other, where $\det k^i_j$ is a non-zero function of x. Two such systems are said to be *weekly equivalent* if a change of unknowns like above and a change of variables x take one to the other.

Problem is *to find a set of invariants for these equivalence relations.* Invariants should be functions of the coefficients of the equations.

Let us consider the simplest non-trivial case: $n = m = 1$ and $r = 2$, *i.e.*, equations in the form

$$u'' + pu' + qu = 0. \tag{1}$$

If we change the unknown as $u \to ku$ (this means 'substitute $u = kv$ into the equation, derive an equation with unknown v, and then write u instead of v'), we obtain

$$u'' + \left(p + 2\frac{k'}{k}\right) u' + \left(q + p\frac{k'}{k} + \frac{k''}{k}\right) u = 0.$$

If we choose k as a non-zero solution of (1), then the coefficient of u vanishes, and moreover if we change the variable x into a non-trivial solution of this new equation, then eventually the equation changes into $u'' = 0$. Therefore, any equation (1) is weakly equivalent to $u'' = 0$; so there is no invariant in this case. Thus we consider the strong equivalence class represented by the equation (1). If we choose k as above, the survived coefficient $p + 2k'/k$ would serve as an invariant; but unfortunately this quantity is not determined uniquely by the coefficient of the equation (1). If we choose k so that the coefficient of u' vanishes, then the survived coefficient can be luckily expressed in terms of p, q, and their derivatives; indeed, using $k'/k = -p/2$, we have

$$q + p\frac{k'}{k} + \frac{k''}{k} = q - \frac{1}{4}p^2 - \frac{1}{2}p',$$

which is exactly the Schwarzian derivative $\{z; x\}$ according to (2). This can be paraphrased as follows: In the (strong) equivalence class of the equation (1), there is a unique equation without the term u', the coefficient of u is given by (2), the

Schwarzian derivative. In other words, for a given function $q(x)$, the non-linear equation

$$\{z; x\} = q(x)$$

can be solved as $z = u_2/u_1$, where u_1 and u_2 are linearly independent solution of the linear equation

$$u'' + qu = 0.$$

We present several lucky generalizations and a semi-lucky case in this note.

3. Geometrical treatment

Schwarz's s-map suggests a geometrical counterpart of a system of linear differential equations in n variables x with m unknowns u with rank r: the map

$$s : x \longmapsto u_1(x) : \cdots : u_r(x) \in \mathrm{Gr}(m, r),$$

where $u_1, \ldots, u_r$ are column m-vectors giving linearly independent solutions, and $\mathrm{Gr}(m, r)$ denotes the Grassmannian variety

$$\mathrm{GL}(m)\backslash\{m \times r\text{-matrix of rank } m\};$$

in particular $\mathrm{Gr}(1, r) = \mathbf{P}^{r-1}$, the $(r-1)$-dimensional projective space. Note that the group $\mathrm{GL}(r)$ acts naturally on $\mathrm{Gr}(m, r)$ (from the right); in other words, the group $\mathrm{GL}(r)$ is the group of motions of $\mathrm{Gr}(m, r)$.

Note that two strongly equivalent systems define the same s-maps, up to the group of motions, and that two weekly equivalent systems give rise to the same images under the s-maps, up to the group of motions.

The problem in §2 can be geometrized as follows: For two maps from the x-space to $\mathrm{Gr}(m, r)$, decide when they are the same up to the group of motions. The weak version of this problem is given by replacing 'they' by 'their images'.

The solution in case $m = n = 1$ and $r = 2$ is given as follows: For given maps s_1 and s_2 both from x-space to $\mathbf{P}^1$, the maps s_1 and s_2 are related linear fractionally if and only if $\{s_1; x\} = \{s_2; x\}$.

In general the study concerning the problem above is called *geometry of submanifolds* or *projective differential geometry.*

We only have quite restricted results, some of which we present in this note.

4. $m = n = 1$: Projective curves

We consider maps of 1-dimensional source space to the target $\mathbf{P}^{r-1}$. Since such a map is called a curve, the problem is to find invariants for curves in the projective space up to the projective motion group $\mathrm{PGL}(r)$.

An analogous problem is often taught and solved in an undergraduate course: A set of invariants of curves, curvature and torsion, in the euclidean 3-space up to the group of rigid motions appear as coefficients of a normalized equation (called Frenet-Serret equation). Our problem is an variant of this problem; we utilize projective motions in place of rigid motions.

This is solved by Halphen, Laguerre and Forsyth; a modern treatment can be found in [**Sea**]. Let us summarize their results when $r = 3$, that is, when the target is the plane $\mathbf{P}^2$. A curve in the plane can be expressed by a system of homogeneous coordinates as $u(x) = u_0(x) : u_1(x) : u_2(x)$. The differential equation for the unknown u

$$\begin{vmatrix} u''' & u'' & u' & u \\ u_0''' & u_0'' & u_0' & u_0 \\ u_1''' & u_1'' & u_1' & u_1 \\ u_2''' & u_2'' & u_2' & u_2 \end{vmatrix} = 0$$

admits u_0, u_1, and u_2 as solutions. If

$$\begin{vmatrix} u_0'' & u_0' & u_0 \\ u_1'' & u_1' & u_1 \\ u_2'' & u_2' & u_2 \end{vmatrix} \neq 0,$$

that is, if the curve is nondegenerate, then the equation has the form

$$u''' + p_1 u'' + p_2 u' + p_3 u = 0.$$

Conversely, for a 3rd order linear ordinary differential equation, three linearly independent solutions give rise to a plane curve; other choice of solutions give another curve which is projectively equivalent to the old one. The equation above for a given curve u is not unique either, indeed, though ρu where ρ is a non-zero function gives the same curve as u, the corresponding differential equation changes. It is easy to see that a suitable choice of ρ makes the coefficient of u'' zero. Let us write the resulting equation as

$$u''' + P_2 u' + P_3 u = 0,$$

where the coefficients can be expressed as rational functions in $p_1, \ldots, p_3$ and their derivatives (parallel computation as in §2). Now take a solution $f(x)$ of the equation

$$\{f; x\} = \frac{1}{4} P_2 \tag{4}$$

and change the variable x into $y = f(x)$, and take a new unknown $w = f'u$, then the equation with unknown w and variable y is of the form

$$\frac{d^3 w}{dy^3} + Rw = 0, \tag{5}$$

actually $R = (P_3 - P_2'/2)/(f')^3$. In this way, a plane curve determines uniquely an equation of the form (5).

Moreover, since any solution of (4) is expressed linear fractionally by f, the parametrization of the curve determined by (5) has exactly a freedom of PGL(2). In this sense, a plane curve carries a natural projective structure.

Anyway in this case the quantity R (expressible as a rational function of $p_1, \ldots, p_3$ and their derivatives) serves as a complete invariant of curves for the weak equivalence relation. Precisely speaking we should take Rdy^3 rather than R itself. An immediate but important consequence is that R vanishes identically if and only if the curve is a conic. A generalization of this statement will be given in §7.

For general $r \geq 3$, a similar argument is known as Laguerre-Forsyth's theory.

5. $m = 1$ and $r = n + 1$: Projective structures

In this section, we find invariants for non-degenerate maps

$$x = (x^1, \ldots, x^n) \mapsto u(x) = u^0(x) : \cdots : u^n(x) \in \mathbf{P}^n$$

under the group PGL$(n + 1)$ of motions of $\mathbf{P}^n$.

¿From the middle of the 19th century, study of systems of linear differential equations admitting algebraic solutions became popular. Algebraic functions are mult-valued, not easy to handle. So they wanted to derive, from the coefficients of the equations, rational functions in the unknown, which are single-valued after substituting (multi-valued) algebraic solutions. Around 1870, several people such as J. Liouville, P. Pepin, L. Fuchs, F. Klein, F. Brioschi, C. Jordan, and E. Goursat tried to make a general theory. Schwarz's paper introduced in §1 solved this problem for the hypergeometric equation. P. Painlevé studied this problem for equations for which $m = 1$, $n = 2$, and $r = 3$. (cf. Notes by Painlevé on 1887.5.31, and by E. Goursat on 1887.5.16 in Comptes Rendus.) Historical background can be found in [**Bou**]. Let us follow their line. Consider a (non-degenerate) map

$$(x, y) \mapsto (z, w),$$

which is multi-valued in a linear fractional way:

$$(z, w) \mapsto (Z, W) = \left(\frac{az + bw + c}{a''z + b''w + c''}, \frac{a'z + b'w + c'}{a''z + b''w + c''} \right).$$

Find *simple* relations in z, w, Z, W, and their derivatives with respect to (x, y), which are independent of the coefficients a, b, c, a', ... Their results are as follows: Put

$$I(z, w) = \frac{z_{xx}w_x - w_{xx}z_x}{z_x w_y - w_x z_y}, \qquad J(z, w) = \frac{w_{yy}z_y - z_{yy}w_y}{z_x w_y - w_x z_y},$$

$$M(z, w) = \frac{z_{xx}w_y - w_{xx}z_y + 2(z_{xy}w_x - w_{xy}z_x)}{3(z_x w_y - w_x z_y)},$$

$$N(z, w) = \frac{w_{yy}z_x - z_{yy}w_x + 2(w_{xy}z_x - z_{xy}w_x)}{3(z_x w_y - w_x z_y)}$$

then

$$I(z, w) = I(Z, W), \quad J(z, w) = J(Z, W),$$
$$M(z, w) = M(Z, W), \quad N(z, w) = N(Z, W).$$

They called these expression the fundamental differential invariants. Note that the (original) Schwarzian derivative involves 3rd derivative, but these invariants involve only derivatives up to 2nd order.

Next, for the map $(x, y) \mapsto (z, w)$, define functions u^0, u^1, and u^2 as

$$u^0 = \rho = (z_x w_y - z_y w_x)^{-1/3}, \qquad u^1 = z\rho, \qquad u^2 = w\rho.$$

Note that the map can be regarded as the map

$$(x, y) \longmapsto 1 : z : w = u^0 : u^1 : u^2 \in \mathbf{P}^2$$

to the target space $\mathbf{P}^2$. It is very important to realize that the factor ρ is so chosen that any linear fractional transformation of (z, w) causes (u^0, u^1, u^2) a linear transformation. Let us derive differential equations, with unknown u, satisfied by u^0, u^1, and u^2:

$$\begin{vmatrix} u_{xx} & u_x & u_y & u \\ u^0_{xx} & u^0_x & u^0_y & u^0 \\ u^1_{xx} & u^1_x & u^1_y & u^1 \\ u^2_{xx} & u^2_x & u^2_y & u^2 \end{vmatrix} = 0, \qquad \begin{vmatrix} u_{xy} & u_x & u_y & u \\ u^0_{xy} & u^0_x & u^0_y & u^0 \\ u^1_{xy} & u^1_x & u^1_y & u^1 \\ u^2_{xy} & u^2_x & u^2_y & u^2 \end{vmatrix} = 0,$$

$$\begin{vmatrix} u_{yy} & u_x & u_y & u \\ u^0_{yy} & u^0_x & u^0_y & u^0 \\ u^1_{yy} & u^1_x & u^1_y & u^1 \\ u^2_{yy} & u^2_x & u^2_y & u^2 \end{vmatrix} = 0.$$

The coefficients of u_{xx}, u_{xy}, and u_{yy} are

$$\begin{vmatrix} u^0_x & u^0_y & u^0 \\ u^1_x & u^1_y & u^1 \\ u^2_x & u^2_y & u^2 \end{vmatrix} = \begin{vmatrix} \rho_x & \rho_y & \rho \\ u_x\rho & u_y\rho & 0 \\ v_x\rho & v_y\rho & 0 \end{vmatrix} = (\rho)^3(u_xv_y - u_yv_x) = 1.$$

The other coefficients turn out to be as

$$\begin{aligned} u_{xx} &= Mu_x - Iu_y + Au, \\ u_{xy} &= -Nu_x - Mu_y + Bu, \\ u_{yy} &= -Ju_x + Nu_y + Cu, \end{aligned}$$

where

$$A = 2(M^2 + IN) - M_x + I_y, \quad B = IJ - MN + M_y + N_x,$$
$$C = 2(N^2 + JM) - N_y + J_x.$$

The computation above can be summarized as follows: For a given map

$$s : (x, y) \mapsto v^0 : v^1 : v^2 \in \mathbf{P}^2,$$

one can find a non-zero function ρ such that $u^0 = \rho v^0$, $u^1 = \rho v^1$, and $u^2 = \rho v^2$ solve a system

$$\begin{aligned} u_{xx} &= p_1u_x + q_1u_y + r_1u, \\ u_{xy} &= p_2u_x + q_2u_y + r_2u, \\ u_{yy} &= p_3u_x + q_3u_y + r_3u, \end{aligned}$$

satisfying

$$p_1 + q_2 = 0, \qquad q_3 + p_2 = 0;$$

the coefficients p_i, q_i, and r_i are uniquely determined by s, and can be expressed in terms of $z = v^1/v^0$ and $w = v^2/v^0$; these give the *relation between the solutions and the coefficients.*

Let us formulate this argument in general for $n \geq 2$; formulae will be shorter. We consider a map

$$s : x = (x^1, \dots, x^n) \mapsto z = (z^1, \dots, z^n).$$

Denote the Jacobi matrix by

$$j(z, x) = (j^k_i), \quad j^k_i = \partial z^k/\partial x^i,$$

and set $J^k_i(z, x) = \partial x^k/\partial z^i$. We assume that s is non-degenerate, that is, $\det j(z, x) \neq 0$. Putting

$$\sigma(z, x) = \frac{1}{n+1} \log \det j(z, x), \qquad \sigma_i(z, x) = \frac{\partial\sigma}{\partial x^i},$$
$$\gamma^k_{ij}(z, x) = \sum_\ell \frac{\partial^2 z^\ell}{\partial x^i \partial x^j} J^k_\ell(z, x),$$

we define the Schwarzian derivatives of z as

$$S^k_{ij}(z; x) = \gamma^k_{ij}(z, x) - \delta^k_i\sigma_j(z, x) - \delta^k_j\sigma_i(z, x).$$

Note that if $n = 1$ this expression is just zero, and that if $n = 2$ they coincide the fundamental differential invariants above. Easy to check that

$$S^k_{ij}(z;x) = S^k_{ji}(z;x), \qquad \sum_k S^k_{ik} = 0.$$

Let us compare the properties of these derivatives with those of the original Schwarzian derivative (cf. (0.1),...,(0.4)). We have (cf. [**Y1**])

(6.1) [PGL$(n+1)$-invariance]

$$S^k_{ij}(Az;x) = S^k_{ij}(z;x), \quad A \in \mathrm{PGL}_{n+1}.$$

(6.2) $S^k_{ij}(z;x) = 0 \longleftrightarrow z = Ax$.

(6.3) [change of variables] If y is another set of variables,

$$S^k_{ij}(z;y) - \sum_{p,q,r} S^r_{pq}(z;x) j^p_i(x;y) j^q_j(x;y) J^k_r(x,y) = S^k_{ij}(x;y).$$

(6.4) [local behavior] If a map $x = (x^1, \dots, x^n) \mapsto z = (z^1, \dots, z^n),\ n \geq 2$ ramifies along $x^1 = 0$ as

$$z^1(x) = (x^1)^\alpha v^1, \quad z^2(x) = v^2, \dots, z^n(x) = v^n, \quad \left| \frac{\partial z}{\partial x} \right| = (x^1)^{\alpha-1} u,$$

where $v^j (1 \leq j \leq n)$ and u are holomorphic not divisible by x^1. Then for $2 \leq i, j, k \leq n$,

$$S^k_{ij}\{z;x\}, \quad S^k_{1j}\{z;x\} + \delta^k_j \frac{1}{n+1} \frac{\alpha - 1}{x^1}, \quad \frac{1}{x^1} S^1_{ij}\{z;x\},$$
$$S^1_{1j}\{z;x\}, \quad x^1 S^k_{11}\{z;x\}, \quad S^1_{11}\{z;x\} - \frac{n-1}{n+1} \frac{\alpha-1}{x^1}$$

are holomorphic. Logarithmic ramification implies $\alpha = 0$.

Let us state the conclusion obtained:

Conclusion: Two maps z_1 and z_2 from the $x = (x^1, \dots, x^n)$-space to $\mathbf{P}^n$ are projectively equivalent if and only if

$$S^k_{ij}(z_1;x) = S^k_{ij}(z_2;x), \quad i,j,k = 1, \dots, n.$$

If we put $\rho = (\det j(z,x))^{-1/(n+1)}$, then $u^0 = \rho$, $u^1 = \rho z^1$, ..., $u^n = \rho z^n$ solve the system

$$u_{ij} = \sum_k S^k_{ij} u_k + S^0_{ij} u, \tag{7}$$

where

$$S^0_{ij} = \frac{1}{n-1} \left(\sum_{\ell,k} S^\ell_{ik} S^k_{\ell j} - \sum_k \frac{\partial}{\partial x^k} S^k_{ij} \right);$$

this is the *relation between the solutions and the coefficients.*

6. The uniformizing equation of the moduli space of the marked cubic surfaces

As an application of the relation between the solutions and the coefficients given in the previous section, we present in this section a system of differential equations which gives a uniformization of the moduli space M of the marked cubic surfaces.

Let us recall just an essence of cubic surfaces. Any non-singular cubic surface in $\mathbf{P}^3$ can be obtained from $\mathbf{P}^2$ by blowing up six points, such that no three points are collinear and no conic passes through the six points. The marking means, in this setup, just the numbering of these six points. Let us represent these six points by a 3×6-matrix, where the j-th column gives homogeneous coordinates of the j-th point. Since no three points are collinear, we can assume that the first four points have the coordinates $1 : 0 : 0$, $0 : 1 : 0$, $0 : 0 : 1$, and $1 : 1 : 1$; so the matrix is of the form

$$\begin{pmatrix} 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & x^1 & x^2 \\ 0 & 0 & 1 & 1 & x^3 & x^4 \end{pmatrix}. \tag{8}$$

The assumption on the six points can be now stated as the non-vanishing of

$$\begin{aligned} D(x): \quad = & \prod_{j=1}^{4} x^j(x^j-1)\cdot(x^1-x^2)(x^1-x^3)(x^2-x^4)(x^3-x^4) \\ & \times (x^1x^4-x^2x^3)\{(x^1-1)(x^4-1)-(x^2-1)(x^3-1)\} \\ & \times \{x^1(x^2-1)(x^3-1)x^4-(x^1-1)x^2x^3(x^4-1)\}. \end{aligned}$$

The moduli space M of the marked (non-singular) cubic surfaces can be identified with

$$\{x = (x^1, \dots, x^4) \in \mathbf{C}^4 \mid D(x) \neq 0\}.$$

In [**ACT**], for each cubic surface $S(x)$ determined by $x \in M$, they considers five *periods* $u^0, \dots, u^4$ of $S(x)$ (precisely speaking, periods of the triple cyclic cover of $\mathbf{P}^3 - S(x)$), and shows that the (multi-valued) map

$$s : M \ni x \longmapsto u^0(x) : \cdots : u^4(x) \in \mathbf{P}^4$$

has its image in the 4-ball

$$B_4 = \{1 : z^1 : \cdots : z^4 \in \mathbf{P}^4 \mid |z^1|^2 + \cdots + |z^4|^2 < 1\},$$

and that the inverse map is defined on B_4 and is single-valued. This can be paraphrased as 'the moduli space M is isomorphic to the quotient of B_4 under a discontinuous group acting on it' or simply 'the moduli space admits a complex hyperbolic structure'.

Now we are ready to apply our Schwarzian derivatives. The period map s above should be a Schwarz map of a system, say E, of differential equations defined on M. Analogous stories such as 'elliptic curves and the hypergeometric equation', 'certain curves admitting a cyclic automorphism group and Appell-Lauricella's hypergeometric equation' (see [**Y2**]), and 'a 4-parameter family of K3 surfaces and the hypergeometric equation of type $(3, 6)$' (see §8) have been known and loved, and these fed many mathematicians. So if we can explicitly know the system E, that would be a lot of fun.

Let us denote the Schwarzian derivatives $S_{ij}^k(s;x)$ of the period map with respect to $x = (x^1, x^2, x^3, x^4)$ by S_{ij}^k, then the system E should be expressed as

$$u_{ij} = \sum_k S_{ij}^k u_k + S_{ij}^0 u,$$

where the coefficients S_{ij}^0 can be determined by S_{ij}^k as is explained in §5. Important is that since the Schwarzian derivatives are invariant under PGL(4), the coefficients S_{ij}^k must be rational functions in x with poles only along the divisor $D(x) = 0$. Moreover the poles along this divisor are fairly restricted by the local properties (6.4), since it is known that the period map s ramifies along this divisor with index 3.

This information is sufficient to determine E, however it needs a complicated computation. If we recall a classical fact that the moduli space M admits a regular action of the Weyl group of type E_6, and require E to be invariant under this group action, then the computation becomes much simpler. We do not dare tabulate all the coefficients of E, but we show just one of them:

$$\begin{aligned} S_{23}^1 \quad &= -\frac{1}{3}\prod_{j=1}^4 x^j(x^j-1)\cdot(x^1-x^2)(x^1-x^3)(x^2-x^4)(x^3-x^4) \\ &\quad\times x^1(x^1-1)\cdot x^4(x^4-1)(x^1-x^3)(x^1-x^2)/D(x). \end{aligned}$$

On the other hand, as is suggested in [**MT**], this system E has a relation with Appell-Lauricella's hypergeometric differential equation. Indeed, E can be obtained from Appell-Lauricella's system of type D in nine variables with special parameters by restricting this system on a 4-dimensional subvariety, extracting from this restricted one a subsystem of rank 5, and finally performing an algebraic coordinate change on this subsystem; for more detail, see [**SY3**] and [**SY4**].

7. $m = 1$ and $r = n + 2$: Conformal structure

We consider systems in n variables $x = (x^1, \dots, x^n)$ with an unknown u of rank $n + 2$ (cf [**SY2**]). When $n = 1$, these are equations for plane curves studied in §4. Let us regard one variable say, x^n, special and write the system as

$$\frac{\partial^2 u}{\partial x^i \partial x^j} = g_{ij}\frac{\partial^2 u}{\partial x^1 \partial x^n} + \sum_{k=1}^n A_{ij}^k \frac{\partial u}{\partial x^k} + A_{ij}^0 u, \quad (1 \le i, j \le n) \tag{9}$$

where

$$g_{ij} = g_{ji},\ g_{1n} = 1,\ A_{ij}^k = A_{ji}^k,\ A_{ij}^0 = A_{ji}^0,\ A_{1n}^k = A_{1n}^0 = 0,$$

and assume $\det g_{ij} \ne 0$. The *normalization factor* is defined as

$$e^\theta = \det\left(\mathbf{u}, \frac{\partial \mathbf{u}}{\partial x^1}, \dots, \frac{\partial \mathbf{u}}{\partial x^n}, \frac{\partial^2 \mathbf{u}}{\partial x^1 \partial x^n}\right),$$

where $\mathbf{u}$ is a column vector consisting of $(n + 2)$ linearly independent solutions; this quantity is independent of the choice of $\mathbf{u}$ up to multiplicative constants. By multiplying a suitable function to the unknown u, we can assume that the condition

$$\det(e^\theta g_{ij}) = 1$$

holds.

This system is said to satisfy *quadric condition* if the image of the Schwarz map

$$x \mapsto \mathbf{u}(x) \in \mathbf{P}^{n+1}$$

is contained in a certain quadratic hypersurface.

We assume $n \geq 3$; we studied already the case $n = 1$ and, when $n = 2$, the situation is a bit different (see [**SY1**]). Then the main result is as follows: If the system satisfies the quadric condition, then the coefficients are expressed as rational functions in g_{ij} and their derivatives:

$$A^j_{ik} = \Gamma^j_{ik} - g_{ik}\Gamma^j_{1n}, \quad A^0_{ik} = -S_{ik} + g_{ik}S_{1n},$$

where Γ^j_{ik} and S_{ik} are the Christoffel symbols and the Schouten tensor of $e^\theta g_{ij} =: h_{ij}$. They are defined as follows:

$$\Gamma^j_{ik} = \frac{1}{2}\sum_l h^{jl}(h_{il,k} + h_{kl,i} - h_{ik,l}), \quad dh_{il} = \sum_k h_{il,k}dx^k.$$

Let ${R^j}_{ikl}$ be the Riemannian curvature tensor:

$$d\pi^j_i - \sum_k \pi^k_i \wedge \pi^j_k = \frac{1}{2}\sum_{k,l} {R^j}_{ikl} dx^k \wedge dx^l, \quad \pi^j_i = \sum_k \Gamma^j_{ik}dx^k.$$

The Ricci and the scalar curvatures are defined by

$$R_{ij} = \sum_l {R^l}_{ilj}, \quad R = \sum_{i,j} h^{ij}R_{ij},$$

respectively, and finally the Schouten tensor (relative to h_{ij}) is defined as

$$S_{ik} = \frac{1}{n-2}\left(R_{ik} - \frac{R}{2(n-1)}h_{ik}\right).$$

8. The uniformizing equation of the moduli space of a 4-dimensional family of K3 surfaces

We present a system of differential equations which gives a uniformization of the moduli space X of the K3 surfaces obtained as double covers of $\mathbf{P}^2$ branching along six lines, that no three lines meet (cf. [**MSY1**], [**Y2**]). Let us represent six lines in the plane by a 3×6-matrix, where the j-th column gives coefficients of the linear equation of the j-th line. Since we can assume that the first four lines are defined by $t^1 = 0, t^2 = 0, t^3 = 0, t^1 + t^2 + t^3 = 0$, these matrices are of the form (8). The assumption on the six lines can now be stated as the non-vanishing of

$$\begin{aligned} D_1(x) : \; &= \prod_{j=1}^{4} x^j(x^j - 1) \cdot (x^1 - x^2)(x^1 - x^3)(x^2 - x^4)(x^3 - x^4) \\ &\times (x^1x^4 - x^2x^3)\{(x^1 - 1)(x^4 - 1) - (x^2 - 1)(x^3 - 1)\}, \end{aligned}$$

and the moduli space X can be identified with

$$\{x = (x^1, x^2, x^3, x^4) \in \mathbf{C}^4 \mid D_1(x) \neq 0\}.$$

Let $K3(x)$ be the surface obtained by desingularizing the double cover of the plane branching along the six lines parameterized by $x \in X$. On each surface $K3(x)$, there are exactly six linearly independent 2-cycles, say $\gamma_0(x), \dots, \gamma_5(x)$, and a unique holomorphic 2-form, say $\phi(x)$ up to multiplicative constant. The period map

$$x \longmapsto u^0(x) : \cdots : u^5(x) \in \mathbf{P}^5, \quad \text{where} \quad u^j(x) = \int_{\gamma_j(x)} \phi(x)$$

happens to have its image in a hyperquadric. The system of differential equations is of the form (9) in the previous section; so all the coefficients can be expressed in terms of the coefficients g_{ij} of $\partial^2 u/\partial x^1 \partial x^4$. They are given as

$$\begin{aligned}
g_{12} &= \frac{x^4-x^3}{x^1-x^2}, \quad g_{13} = \frac{x^4-x^3}{x^1-x^2}, \quad g_{14} = 1, \\
g_{23} &= 1, \quad g_{23} = \frac{x^3-x^1}{x^2-x^4}, \quad g_{34} = \frac{x^2-x^1}{x^3-x^4}, \\
g_{11} &= \frac{x^2x^3-x^4}{x^1(1-x^1)} - \frac{x^3(x^4-x^2)}{x^1(x^1-x^3)} - \frac{x^2(x^4-x^3)}{x^1(x^1-x^2)}, \\
g_{22} &= \text{make exchanges } x^1 \leftrightarrow x^2 \text{ and } x^3 \leftrightarrow x^4 \text{ in } g_{11}, \\
g_{33} &= \text{make an exchange } x^2 \leftrightarrow x^3 \text{ in } g_{22}, \\
g_{44} &= \text{make an exchange } x^1 \leftrightarrow x^4 \text{ in } g_{11}.
\end{aligned}$$

On the other hand, it is known that this system is equivalent to the hypergeometric system $E(3,6;\alpha_1,\dots,\alpha_6)$ of type $(3,6)$ with parameters $\alpha_j = 1/3$.

9. $n = 1$ and $r = md$: Systems of ordinary differential equations of order d

Let us consider a system of ordinary differential equations with m unknowns $\mathbf{u} = {}^t(u_1,\dots,u_m)$ of order d

$$\mathbf{u}^{(d)} = \sum_{k=0}^{d-1} P_k \mathbf{u}^{(k)}. \tag{10}$$

Under a transformation of unknowns

$$\mathbf{v} = K\mathbf{u}, \quad K = (k_i^j(x))_{1\le i,j\le m}, \ \det K \neq 0, \tag{11}$$

the system (10) is transformed into

$$\mathbf{v}^{(d)} = (KP_1 + dK')K^{-1}\mathbf{v}^{(d-1)} + \cdots .$$

Choosing K so that $KP_1 + dK' = 0$, we can assume

$$P_1 = 0. \tag{12}$$

Change of variable $y = f(x)$ together with the transformation (11), then the system (10) changes into

$$\begin{aligned}
&(f')^d\mathbf{v}^{(d)} + a_d(f')^{d-2}f''\mathbf{v}^{(n-1)} + (b_d(f')^{d-3}f''' + c_d(f')^{d-4}(f'')^2)\mathbf{v}^{(d-2)} + \cdots \\
&\quad = K\mathbf{u}^{(d)} + dK'\mathbf{u}^{(d-1)} + \frac{d(d-1)}{2}K''\mathbf{u}^{(n-2)} + \cdots ,
\end{aligned}$$

where

$$a_d = \frac{d(d-1)}{2}, \ b_d = \frac{d(d-1)(d-2)}{6}, \ c_d = \frac{d(d-1)(d-2)(d-3)}{8}.$$

¿From this, we get

$$(f')^d\mathbf{v}^{(d)} = (d(f')^{d-1}K'K^{-1} - a_d(f')^{d-2}f'')\mathbf{v}^{(d-1)} + (f')^{d-2}\tilde{P}_2\mathbf{v}^{(d-2)} + \cdots ,$$

where

$$\tilde{P}_2 = KP_2K^{-1} - b_d\frac{f''}{f'} - c_d\left(\frac{f''}{f'}\right)^2 + d\,a_{d-1}\frac{f''}{f'}K'K^{-1} \tag{13}$$

$$-d(d-1)K'K^{-1}K'K^{-1} + \frac{d(d-1)}{2}K''K^{-1}.$$

Hence, to preserve the condition (12), K must be of the form

$$K = (f')^{(n-1)/2}B, \qquad B : \text{a constant matrix.}$$

On the other hand, from (13), we have

$$\frac{\mathrm{tr}\tilde{P}_2}{m} = \frac{d(d-1)(d+1)}{6}\{f;x\} + \frac{\mathrm{tr}P_2}{m}.$$

Thus, by choosing f satisfying $\mathrm{tr}\tilde{P}_2 = 0$, we get the following conclusion:

1. The system (10) can be normalized as

$$P_1 = 0 \qquad \text{and} \qquad \mathrm{tr}\,P_2 = 0.$$

2. A transformation of unknown $\mathbf{u}$ and that of variable preserving this condition has the form

$$y = \frac{\alpha x + \beta}{\gamma x + \delta},$$

$$\mathbf{v} = (\gamma x + \delta)^{1-d}B\mathbf{u}, \quad (B : \text{a constant matrix}).$$

3. Under this transformation, the matrix-valued quadratic form P_2dx^2 changes only by a conjugate action of B.

We can moreover see that the i-differential form

$$R_i = \sum_{j=0}^{i-2} a_{i,j}\left(\frac{d}{dx}\right)^j P_{i-j}(dx)^i \qquad (i \geq 2),$$

where

$$a_{i,j} = (-1)^j\frac{(2i-j-2)!(n-i+j)!}{j!(i-j-1)!}$$

are also invariants up to conjugation of B. ([**W**],[**Sea**]). Note that when $m = 1$, $\mathrm{tr}P_2 = 0$ implies $P_2 = 0$, of course. When $m = 1$ and $d = 3$, the conclusion above reduces to that in §4, where projective curves are studied.

When $m = 2$ and $d = 2$, our argument can be translated into the geometry of ruled surfaces in $\mathbf{P}^3$. In fact, four($= r = md$) linearly independent solutions

$$\begin{pmatrix} u_1^0 \\ u_2^0 \end{pmatrix}, \ldots, \begin{pmatrix} u_1^3 \\ u_2^3 \end{pmatrix},$$

viewed as two curves

$$u_1 = (u_1^0, \ldots, u_1^3) \quad \text{and} \quad u_2 = (u_2^0, \ldots, u_2^3)$$

define a ruled surface swept by lines joining $u_1(x)$ and $u_2(x)$; these lines can be thought of points in $\mathrm{Gr}(2,4)$.

10. $n = m^2$ and $r = 2m$: Systems modelled after $\mathrm{Gr}(m, m^2)$

The reader might think that for any m, n, and r, there would be some analogue of the Schwarzian derivative. Unfortunately things are not so simple. In this section, we consider a seemingly simple case: systems of linear differential equations in m^2 variables $x^{ij}(1 \le i, j \le m)$ of rank $2m$ with m unknowns $u^k(1 \le k \le m)$ and the changes K of unknowns

$$(u^k) \to (\sum_l K^k_l u^l), \quad \det(K^k_l) \not\equiv 0;$$

two systems related under such changes are said to be equivalent ([**SY5**], cf. [**SYY**]). As the ratio of two linearly independent solutions of 2nd order ordinary differential equation defines a map (the Schwarz map) from the x-space to the projective line, any $2m$ linearly independent solutions of our system defines a map from the $x = (x^{ij})$-space to the $(m, 2m)$-Grassmannian variety $\mathrm{Gr}(m, 2m)$; two equivalent systems define the same Schwarz map. We assume that this map is non-degenerate. We shall show that we can follow the classical argument to some extent but not to the quite same.

The system in question is an innocent analogue of the classical model case $(n = 1, m = 1, r = 2m)$, we just replace scalar variable and scalar unknown by $m \times m$-matrices. Or, this can be seen as a generalization of the case treated in §9 of order $d = 2$ to a high dimensional source space. When $m = 2$, since $\mathrm{Gr}(2, 4)$ can be embedded in $\mathbf{P}^5$ as a quadratic hypersurface, the uniformizing equation of the K3 surfaces presented in §8 can be transformed into a system modelled after $\mathrm{Gr}(2, 4)$.

Let us regard one variable, say x^{11}, special and write down our system as

$$E = E_m(a, b, \alpha, \beta) \begin{cases} u^k_{:11:11} &= \sum_l \alpha^k_l u^l_{:11} + \sum_l \beta^k_l u^l, \\ u^k_{:ij} &= \sum_l a^k_{ijl} u^l_{:11} + \sum_l b^k_{ijl} u^l, \end{cases}$$

$1 \le k, l, i, j \le m$, where $f_{:ij}$ stands for $\partial f/\partial x^{ij}$, and

$$a^k_{11l} = \delta^k_l, \quad b^k_{11l} = 0.$$

The Schwarz map of the system E is non-degenerate if and only if $m^2 \times m^2$-determinant

$$W = \det(a^k_{ijl})_{(i,j),(k,l)}$$

does not vanish identically.

The transformation

$$u^k \to \sum_l K^k_l u^l, \quad \det K^k_l \ne 0$$

changes the coefficients a as

$$a^k_{ijl} \to \sum K^k_p a^p_{ijq} (K^{-1})^q_l,$$

in other words,

$$\mathbf{A} = (a^k_l) \to K\mathbf{A}K^{-1}, \quad a^k_l = \sum a^k_{ijl} dx^{ij},$$

and α as

$$\mathcal{A} = (\alpha^k_l) \to (2K_{:11} + K\mathcal{A})K^{-1}.$$

Thus we can normalize the system as $\mathcal{A} = 0$, but still remains a freedom of transformations K satisfying $K_{:11} = 0$; this implies that the survived coefficients of this normalized system can not be expressed by the coefficients of the given system E. Nevertheless we can control the system: we can extract a set of essential coefficients as follows.

Under the assumptions $W \neq 0$ and $\mathbf{A} \wedge \mathbf{A} \neq 0$, where $\mathbf{A} = (a^k_l)$, the coefficients a determine the other coefficients b and β up to adding an exact 1-form $dk(x)$, where $k(x)$ is independent of x^{11}, to b^i_i $(i = 1, \ldots, m)$. This ambiguity is caused by the scalar transformation $K = k(x)I_n$.

Hence we have the following conclusion: Two systems $E_m(a, b, \alpha, \beta)$ and $E_m(\bar{a}, \bar{b}, \bar{\alpha}, \bar{\beta})$ are equivalent if only if there is K such that

$$\bar{\mathbf{A}} = K^{-1}\mathbf{A}K, \quad K = (K^k_l),\ \mathbf{A} = (a^k_l),\ a^k_l = \sum a^k_{ijl}\, dx^{ij},$$

provided that $W \neq 0$ and $\mathbf{A} \wedge \mathbf{A} \neq 0$.

11. Generalizations in different context

The Schwarzian derivative is used in the theory of univalent functions and in the study of the Teichmüller spaces. A fundamental fact known as Nehari's theorem can be stated as

> If a function z holomorphic in $|x| < 1$ is univalent, then
> $$|\{z; x\}| \leq 6(1 - |z|^2)^{-2}.$$
> Conversely if
> $$|\{z; x\}| \leq 2(1 - |z|^2)^{-2},$$
> then z is univalent in $|x| < 1$.

This leads to the boundedness of the Bers embedding of the Teichmüller space. Kobayashi-Wada [**KY**] defined a variant of the Schwarzian derivative for non-degenerate maps between two Riemann manifolds, and established a generalization of Nehari's theorem.

On the other hand, Sato [**Sat**] studied obstructions for 2nd order ordinary differential equations $y'' = f(x, y, y')$ to be reduced to the equation $y'' = 0$ under transformations of (x, y)-space, and found that the obstructions can be expressed in terms of the fundamental differential invariants I, J, M, and N appeared in §5. He moreover derived obstructions for 3rd order ordinary differential equations $y''' = f(x, y, y', y'')$ to be reduced to the equation $y''' = 0$ under contact transformations of (x, y, y')-space, and called them contact Schwarzian derivatives.

References

[ACT] D. Allcock, J. Carlson, and D. Toledo, A complex hyperbolic structure for moduli of cubic surfaces, Comptes Rendus Acad. Sci. **326**(1998), 49–54.

[Bou] A. Boulanger, Contribution à l'étude des équations différentielles linéaires homogènes intégrable algébriquement, J. de l'Ecole Polytechnique **4**(1898), 1–122.

[KY] O. Kobayashi and M. Wada, Circular geometry and the Schwarzian, to appear in FarEast J. Math.

[MSY1] K. Matsumoto, T. Sasaki and M. Yoshida, The monodromy of the period map of a 4-parameter family of K3 surfaces and the Aomoto-Gel'fand hypergeometric function of type (3,6), Internat. J. of Math. **3** (1992), 1–164.

[MSY2] —, Recent progress of Gauss-Schwarz theory and related geometric structures, Kyushu J. Math. **47**(1993), 283–381.

[MT] K. Matsumoto and T. Terasoma, Theta constants associated to cubic threefolds, math.AG/0008024.

[Sat] H. Sato, Schwarzian derivatives of contact diffeomorphisms, Lobachevskii J. Math. **4**(1999), 89–98.

[Sas] T. Sasaki, *Projective Differential Geometry and Linear Homogeneous Differential Equations*, Rokko Lectures in Math **5**(1999), Dept. of Math., Kobe Univ.

[Sch] H. A. Schwarz, Ueber diejenigen Fälle, inwelchen die Gaussische hypergeometrische Reihe eine algebraische Funktion ihres vierten Elementes darstellt, J. reine und angew. Math. **75**(1872), 292–335.

[Sea] Y. Se-ashi, A geometric construction of Laguerre-Forsyth's canonical forms of linear ordinary differential equations, Advanced Studies in Pure Math. **22**(1993), 265–297.

[SY1] T. Sasaki and M. Yoshida, Linear differential equations in two variables of rank four I, Math. Ann. **282**(1988), 69–93.

[SY2] —, Linear differential equations modeled after hyperquadrics, Tôhoku Math. J. **41**(1989), 321–348.

[SY3] —, The uniformizing differential equation of the complex hyperbolic structure on the moduli space of marked cubic surfaces, Proc. Japan Acad. **75**(1999), 129–133.

[SY4] —, The uniformizing differential equation of the complex hyperbolic structure on the moduli space of marked cubic surface II, Journal of Physics A: Math. and Gen. **34**-11(2001).

[SY5] —, Invariant theory for linear differential systems modelled after the Grassmannian $Gr(n, 2n)$, Kyushu Univ. preprint 2001-1.

[SYY] T. Sasaki, K. Yamaguchi, and M. Yoshida, On the rigidity of differential systems modelled on hermitian symmetric spaces and disproofs of a conjecture concerning modular interpretations of configuration spaces, Advanced Studies in Pure Math. **25**(1997), 315–354.

[W] E. J. Wilczynski, *Projective Differential Geometry and Curves and Ruled Surfaces*, Teubner 1906; reprinted by Chelsea Publ. Co.

[Y1] M. Yoshida, *Fuchsian Differential Equations*, Aspects of Mathematics **E11**, Vieweg, Wiesbaden 1987.

[Y2] M. Yoshida, *Hypergeometric Functions, My Love*, Vieweg, Wiesbaden 1997.