Weak Schwarzians, Bounded Hyperbolic Distortion, and Smooth Quasisymmetric Functions

M. Chuaqui
University of Pennsylvania

B. Osgood
Stanford University and IHES

1 Introduction

In this paper we wish to show how some techniques based on a Sturm comparison theorem for the differential equation associated with the Schwarzian derivative can be used to study two problems. First, to estimate the quasisymmetry quotient of a function in terms of bounds on its Schwarzian. Here, the bounds on the Schwarzian are much like those one finds in the theory of univalent functions, and the result is a sufficient condition for a function to be quasisymmetric. This is discussed in Section 3. Second, to study how much mappings of an interval distort distances in the hyperbolic metric. These results are Schwarz-Pick type lemmas and are discussed in Section 4. Apart from the differential equations arguments there are interesting issues having to do with smoothness. In Section 5 we combine the estimates for hyperbolic distances with those for quasisymmetry quotients to obtain a result expressing a quasisymmetric function of the type we have been considering as a composition of functions whose quasisymmetry quotients are arbitrarily close to 1. Finally, in Section 6 we construct some examples to show that there is no obvious necessary condition for a function to be quasisymmetric corresponding to the sufficient conditions in Section 3.

We work with real valued functions of a real variable. Let $f : I \to \mathbf{R}$ be an increasing homeomorphism, where I is an open interval that may be the whole real line. The *quasisymmetry quotient* of f is

$$kf(x,h) = \frac{f(x+h) - f(x)}{f(x) - f(x-h)} \tag{1.1}$$

for $x, x+h, x-h \in I$. The function is called *quasisymmetric* if $kf(x,h)$ is bounded below away from zero and above away from ∞. Because of $kf(x,-h) = kf(x,h)^{-1}$ we may assume that $h > 0$ for this definition. One says that f is $k-$*quasisymmetric*, $k \geq 1$, if

$$\frac{1}{k} \leq kf(x,h) \leq k.$$

A similarity is 1-quasisymmetric, and the functions f and $g = af + b$, $a, b \in \mathbf{R}$, have $kf = kg$.

When f is monotonic and three times differentiable its Schwarzian derivative is

$$Sf = \left(\frac{f''}{f'}\right)' - \frac{1}{2}\left(\frac{f''}{f'}\right)^2 = \frac{f'''}{f'} - \frac{3}{2}\left(\frac{f''}{f'}\right)^2. \tag{1.2}$$

We remind the reader of the chain rule for the Schwarzian,

$$S(f \circ g) = (S(f) \circ g)(g')^2 + Sg, \tag{1.3}$$

and the fact that Sf is identically zero if and only if f is a Möbius transformation, $f(x) = (ax+b)/(cx+d)$.

If u is the solution to the initial value problem

$$u'' + \frac{1}{2}pu = 0, \quad u(0) = 1,\, u'(0) = 0, \tag{1.4}$$

on an interval containing the origin, and

$$f(x) = \int_0^x u^{-2}(t)\, dt\,,$$

then $Sf = p$ and $f(0) = 0, f'(0) = 1, f''(0) = 0$. For brevity we say that a function f with these values at the origin is *normalized*. This normalization, and the question of when a function can or cannot be normalized, is important in our work. To explain a little more, we can first achieve $f(0) = 0$ and $f'(0) = 1$ using only affine transformations, which affect neither the Schwarzian nor the quasisymmetry quotient of f. The further (parabolic) Möbius transformation

$$f^\dagger = f/(1 + a_2 f),\; a_2 = (1/2)f''(0)$$

will then obtain $(f^\dagger)''(0) = 0$. We will be normalizing in this way very frequently, so we will often use the dagger notation for the normalized function. However this last transformation, which still does not affect the Schwarzian though it does change the quasisymmetry quotient, will allow $f^\dagger$ to become unbounded if there is a point x_0 where $f(x)$ tends to the value $-1/a_2$ as $x \to x_0$. We treat this question in Section 2.

For applications to quasisymmetric functions, where in general no smoothness is required beyond continuity, we would naturally like to relax the conditions needed on a function to define its Schwarzian. For instance, by the Rademacher-Stepanov theorem, if f is of class $C^{2,1}_{\text{loc}}$ then Sf is defined a.e. by the formula (1.2). While one would like to define a still 'weaker' Schwarzian, the class $C^{2,1}_{\text{loc}}$ does comes up in our work in two ways. First, for the initial value problem (1.4) one can easily prove existence, uniqueness and the relevant comparison theorem when the coefficient p is in L^∞_{loc}, in which case the solution u will be $C^{1,1}_{\text{loc}}$ and f will be $C^{2,1}_{\text{loc}}$. This is done in Section 2; probably one can do better here. Second, and to us more surprising, is that $C^{2,1}_{\text{loc}}$ is the degree of smoothness that is *implied* by controlling the amount of distance distortion in the hyperbolic metric. This regularity result in turn implies a compactness theorem in $C^{2,1}_{\text{loc}}$ for the space of functions whose Schwarzians are bounded in the hyperbolic metric.

Here briefly is a summary statement of our main results, with more complete definitions, statements, and additional results in the later sections. We let $d_J(x, y)$ be the hyperbolic distance between x and y in an interval J. For constants $0 < r \le 1$, $1 \le s < \infty$, F_r and G_s are the normalized functions on $I = (-1, 1)$ with

$$SF_r(x) = \frac{2(1 - r^2)}{(1 - x^2)^2}, \quad SG_s(x) = \frac{-2(s^2 - 1)}{(1 - x^2)^2}.$$

These functions respectively decrease and increase the hyperbolic distance on I by the constant factors of r and s. They are also extremals for bounding the Schwarzian and the quasisymmetry quotient.

Theorem *Let $f : I \to \mathbf{R}$ be a non-constant, increasing function. Suppose there are numbers $0 < r \leq 1$ and $1 \leq s < \infty$ such that for every open subinterval $J \subseteq I$*

$$d_{F_r(J)}(F_r(x), F_r(y)) \leq d_{f(J)}(f(x), f(y)) \leq d_{G_s(J)}(G_s(x), G_s(y)) \tag{1.5}$$

for all $x, y \in J$. Then $f \in C^{2,1}_{\mathrm{loc}}(I)$ and

$$SG_s \leq Sf \leq SF_r \ \ a.e.. \tag{1.6}$$

Conversely, if $f \in C^{2,1}_{\mathrm{loc}}(I)$ and (1.6) holds then so does (1.5). The set of normalized functions satisfying either of these conditions is a compact family in $C^{2,1}_{\mathrm{loc}}(I)$ of $k(r,s)-$quasisymmetric functions, with

$$k(r,s) = \frac{s}{r} \max\left\{ \frac{2^{1-r}}{2^r - 1}, \frac{2^s - 1}{2^{1-s}} \right\}.$$

The compactness here is in the topology of local uniform convergence of a sequence of functions together with the sequences of first and second derivatives, and weak* convergence of the third derivatives as elements of L^∞_{loc}.

We remark that smooth quasisymmetric functions have not been of primary interest in the subject, certainly as far as their relations to quasiconformal mappings and Teichmüller theory go, where totally singular functions are the rule. However, the Schwarzian, expanding and contracting maps, quasisymmetry, and questions of smoothness have also all played a role in one-dimensional dynamics. See the important papers [4], [12] by de Melo-van Strien and by Sullivan, to cite some recent work. For example, in [12] and [10] it is proved that a map of an interval is locally bi-Lipschitz in the hyperbolic metric if and only if it is of class $C^{1+\mathrm{Zygmund}}$. Also, in their paper [6] Gardiner and Sullivan study some cases when their symmetric quasisymmetric functions are C^1. Our interest in the Schwarzian has come from univalent functions, and the present paper follows up on work in [2], [3]. Though the amount of smoothness we require here may still not be satisfactory, the arguments seemed to be worth developing. We feel this is so partly because the differential equations arguments work so naturally and involve the explicit and interesting extremal functions, and partly because we do not use quasiconformal extensions. For both of these reasons the estimates are elementary and fairly precise.

We refer to the book by O. Lehto [9] for an excellent account of just about all of the background material that is needed, as well as to the paper [3] for some of the results in Section 2.

2 Weak Schwarzians and Bounds on f from Sf

For the following discussion we suppose that functions are defined on the interval $(-1,1)$. As mentioned above, if $f \in C^{2,1}_{\text{loc}}(-1,1)$ then Sf is defined a.e. and can be regarded as an element of $L^\infty_{\text{loc}}(-1,1)$. A technical remark may be in order here. A function whose derivative exists a.e. is not necessarily absolutely continuous of course, but a function with a 'weak derivative', as in the theory of distributions, is. For functions in $C^{2,1}$, f'' is absolutely continuous and so f''' is its weak derivative in whatever setting one is working. Thus there is some justification for calling Sf a 'weak Schwarzian' when we start with f in $C^{2,1}_{\text{loc}}$. This will come up in Section 4 where in one instance we are able to define Sf almost everywhere for a C^1 function with a log convex derivative.

To bring in the differential equation, we now have:

Theorem 1 *Let $p \in L^\infty_{\text{loc}}(-1,1)$. There is a unique solution $u \in C^{1,1}_{\text{loc}}(-1,1)$ of*

$$u'' + \frac{1}{2}pu = 0 \ a.e., \quad u(0) = 1,\, u'(0) = 0. \tag{2.1}$$

If

$$f(x) = \int_0^x u^{-2}(t)\,dt\,,$$

then $f \in C^{2,1}_{\text{loc}}(-1,1)$, as long as $u \neq 0$, and $Sf = p$ a.e..

This may be standard, but for completeness we sketch a proof based on a fixed point method. We piece together solutions on small intervals, so we need to solve the equation in a neighborhood of any point x_0 with any initial conditions $u(x_0) = a$, $u'(x_0) = b$. Let $x_0 \in (-1,1)$ and for $0 < \epsilon < 1$, small, let J be the centered interval $(x_0 - \epsilon, x_0 + \epsilon)$, which we assume is compactly contained in $(-1,1)$. Let

$$X = \{\phi \in C^{1,1}(J) \,:\, \phi(x_0) = 0,\, \phi'(x_0) = b\},$$

and let $T : X \to X$ be defined by $T\phi = \psi$ where

$$\psi'(x) = b - \frac{1}{2}\int_{x_0}^x (p\phi)(s)\,ds - \frac{a}{2}\int_{x_0}^x p(s)\,ds\,. \tag{2.2}$$

X is a complete metric space with the usual norm on $C^{1,1}(J)$, and we claim that T is a contraction provided ϵ is sufficiently small.

First, if $T\phi_1 = \psi_1$ and $T\phi_2 = \psi_2$ then $||{\psi_1}' - {\psi_2}'||_\infty \le (\epsilon/2)||p||_{\infty,\bar{J}}||\phi_1 - \phi_2||_\infty$. Next, since $\psi_1(x_0) = \psi_2(x_0) = 0$, this implies that $||\psi_1 - \psi_2||_\infty \le (\epsilon^2/2)||p||_{\infty,\bar{J}}||\phi_1 - \phi_2||_\infty < (\epsilon/2)||p||_{\infty,\bar{J}}||\phi_1 - \phi_2||_\infty$. Finally, it also follows from the definition of T that the Lipschitz constant for ${\psi_1}' - {\psi_2}'$ is at most $(\epsilon/2)||p||_{\infty,\bar{J}}||\phi_1 - \phi_2||_\infty$. Hence

$$||\psi_1 - \psi_2||_{1,1} \le (3\epsilon/2)||p||_{\infty,\bar{J}}||\phi_1 - \phi_2||_\infty \le (3\epsilon/2)||p||_{\infty,\bar{J}}||\phi_1 - \phi_2||_{1,1}\,.$$

Therefore T has a unique fixed point if $\epsilon < 2/(3||p||_{\infty,\bar{J}})$. If ϕ is that fixed point then $u = a + \phi$ is the solution to

$$u'' + \frac{1}{2}pu = 0, \quad u(x_0) = a,\, u'(x_0) = b$$

on J.

This solves the equation on small intervals with initial data at the center points. Starting with an interval around zero with the initial conditions $u(0) = 1$, $u'(0) = 0$ we can piece together a $C^{1,1}_{\text{loc}}$ solution on $(-1, 1)$ by covering any given compact set with enough such intervals so that adjacent overlapping intervals contain each others centers. (The estimates for the size of the intervals, the $\epsilon's$, can be made uniform in terms of $||p||_\infty$ on a slightly larger compact set.) The remaining assertions in the statement of the Theorem are immediate.

The version of the Sturm comparison theorem that we need may be stated as follows. We consider the two initial value problems

$$u'' + pu = 0,\, a.e. \quad u(0) = 1,\, u'(0) = 0,\, p \in L^\infty_{\text{loc}}[0, 1),$$

and

$$v'' + qv = 0,\, a.e. \quad v(0) = 1,\, v'(0) = 0,\, q \in L^\infty_{\text{loc}}[0, 1).$$

Suppose $q \geq p$. By this we mean that the integral of $q - p$ against smooth, non-negative functions of compact support in $(-1, 1)$ is non-negative. Then $u \geq v$ until the first zero of v. The proof follows the classical case almost word for word. Consider $w = uv' - vu'$, which is Lipschitz (locally), and then use that $w' = (p - q)uv$ is ≤ 0 as an element of $L^\infty_{\text{loc}}[0, 1)$ provided $uv > 0$ to get w decreasing.

After these generalities we now want to discuss the specific types of bounds and comparisons we will be using. The goal is to obtain bounds on a function from bounds on its Schwarzian. All of our subsequent work is based on this. The results stated below are from [3], where the notation was different and the setting was analytic functions in the disk. Nothing is required here beyond the comparison theorem as stated above. In fact, working with functions which are not analytic allows much more flexibility in constructing examples based upon the differential equation, as we shall see in Section 6.

Let r and s be two constants with $0 < r \leq 1$ and $1 \leq s < \infty$. Define

$$F_r(x) = \frac{1}{r}\frac{(1 + x)^r - (1 - x)^r}{(1 + x)^r + (1 - x)^r}, \quad G_s(x) = \frac{1}{s}\frac{(1 + x)^s - (1 - x)^s}{(1 + x)^s + (1 - x)^s}.$$

As a function of x the behavior is different when the parameter is less than or greater than one, so we prefer to use two names for the function. F_r is concave up on $(0, 1)$ and concave down on $(-1, 0)$ while the reverse is true for G_s. F_r and G_s are odd and are normalized at the origin. They fit together alternately on $(-1, 0]$ and $[0, 1)$ to give C^2 functions on $(-1, 1)$ that do not change concavity. To maintain the distinction between the two functions we write the Schwarzians as

$$SF_r(x) = \frac{2(1 - r^2)}{(1 - x^2)^2}, \quad SG_s(x) = \frac{-2(s^2 - 1)}{(1 - x^2)^2}.$$

Notice that with the F_r we have functions whose Schwarzians are positive but are always less than $2/(1-x^2)^2$, while with the G_s the Schwarzians can be as negative as we please.

These functions are 'extremal' for many of the problems we shall study. Here we stress only the basic estimates that depend on the Schwarzian, and the question of normalizing. Some of their other properties will be elaborated in Sections 3 and 4.

Suppose that f is a $C^{2,1}_{\mathrm{loc}}$, increasing, normalized function on $(-1,1)$ whose Schwarzian satisfies the bounds

$$SG_s \le Sf \le SF_r \tag{2.3}$$

on $(-1,1)$. Again, when we write such an inequality between L^∞_{loc} functions, which we will be doing frequently, we mean for it to hold in the distributional sense, as we explained in connection with Sturm comparison theorem. That theorem gives upper and lower bounds for $u=(f')^{-1/2}$ on $[0,1)$ in terms of $v=(F_r')^{-1/2}$ and $w=(G_s')^{-1/2}$ which lead to

$$G_s'(x) \le f'(x) \le F_r'(x), \tag{2.4}$$
$$G_s(x) \le f(x) \le F_r(x), \tag{2.5}$$

for $x \in [0,1)$. For inequalities on $(-1,0]$ we define $g(x)=-f(-x)$, which has $Sg(x) = Sf(-x)$, and apply the above bounds to g for $x \ge 0$. Since F_r and G_s are odd one then finds that (2.5) is replaced by

$$F_r(x) \le f(x) \le G_s(x), \tag{2.6}$$

for $x \in (-1,0]$, while (2.4) continues to hold, as is, on $(-1,0]$. Two consequences of (2.5) and (2.6) that we will often use are

$$1/s \le f(1) \le 1/r\,, \tag{2.7}$$

$$-1/r \le f(-1) \le -1/s\,. \tag{2.8}$$

As for cases of equality we only need a fairly weak statement, that if f agrees with one of the extremals F_r, G_s at an endpoint ± 1 then it must agree with the corresponding function on the half interval $[0,\pm 1]$. This follows easily from integrating the inequalities on the derivatives (F_r' and G_s' are integrable).

An interesting issue associated with these estimates is that of the normalization; when is it possible to normalize and what happens if it is not possible? First, a normalized function satisfying even just the upper bound

$$Sf \le SF_r \tag{2.9}$$

will be subject to

$$|f(x)| \le |F_r(x)| \tag{2.10}$$

on all of $(-1,1)$, from (2.5) and (2.6) above. It will therefore be bounded on $[-1,1]$ by $\pm 1/r$. Suppose f is not normalized but satisfies (2.9). As explained in the Introduction, we may assume that $f(0)=0, f'(0)=1$ and then normalize to get the second derivative to vanish at the origin by defining

$$f^\dagger = f/(1+a_2 f),\; a_2 = (1/2)f''(0),$$

at the expense of possibly introducing a singularity if there is a point $x_0 \in [-1,1]$ where $f(x)$ tends to the value $-1/a_2$ as $x \to x_0$. Now, $f^\dagger$ also satisfies (2.9) and so it will satisfy (2.10) on $(-|x_0|,|x_0|)$. But since $F_r(x)$ is bounded on $[-1,1]$, $f^\dagger$ cannot become unbounded as $x \to \pm x_0$. We conclude that any f satisfying (2.9) can be normalized and will then be subject to the bounds (2.10) on $[-1,1]$. In fact, the argument shows that $-1/a_2$ lies in the complement of the closure of the range of f.

Finally, if f is normalized and satisfies the lower bound

$$SG_s \le Sf, \tag{2.11}$$

then f will satisfy

$$|G_s(x)| \le |f(x)| \tag{2.12}$$

on $(-1,1)$. If f is not normalized but satisfies (2.11) then it may not be possible to normalize further by defining $f^\dagger$ without introducing a singularity. $f^\dagger$ will satisfy (2.12) as far as it is regular. In any case, a function with $f(0)=0$, $f'(0)=1$, whose Schwarzian has this lower bound will be subject to the coefficient inequality

$$|a_2| \le s. \tag{2.13}$$

This follows because even if $f^\dagger$ is not regular on all of $(-1,1)$, its range will always cover the interval $(-1/s,1/s)$, and $f = f^\dagger/(1-a_2 f^\dagger)$ is regular.

Remark 1 For later applications we also need versions of these estimates for functions on an interval $(-R,R)$, with the same normalization at the origin. The new extremals are

$$\tilde{F}_r(x) = RF_r(\frac{x}{R}), \quad \tilde{G}_s(x) = RG_s(\frac{x}{R}),$$

with

$$S\tilde{F}_r(x) = \frac{2R^2(1-r^2)}{(R^2-x^2)^2}, \quad S\tilde{G}_s(x) = \frac{-2R^2(s^2-1)}{(R^2-x^2)^2}.$$

If we replace (2.3) by

$$S\tilde{G}_s \le Sf \le S\tilde{F}_r \tag{2.14}$$

for a normalized function f on $(-R,R)$, then all the discussion above goes through with $\tilde{F}_r$ and $\tilde{G}_s$ replacing F_r and G_s, respectively. For example, the inequalities (2.7), (2.8) are replaced by

$$R/s \le f(R) \le R/r\,, \tag{2.15}$$

$$-R/r \le f(-R) \le -R/s\,. \tag{2.16}$$

The coefficient bound (2.13) is replaced by $|a_2| \le s/R$.

We could also translate the origin and formulate the results for intervals centered at a point x_0; this amounts to a trivial change. It is not so easy to give clean statements when the normalization is not at the center of the interval. This point comes up in Section 5.

3 Sufficient Conditions for Quasisymmetry

The differential equation (1.4) for the Schwarzian and the estimates that come from it are well suited to studying the quasisymmetry quotient. Consider the identity

$$kf(x,h) = \frac{f(x+h)-f(x)}{f(x)-f(x-h)} = \frac{\displaystyle\int_x^{x+h} \exp\left(\int_x^y \frac{f''}{f'}(t)\,dt\right)\,dy}{\displaystyle\int_{x-h}^{x} \exp\left(\int_x^y \frac{f''}{f'}(t)\,dt\right)\,dy}\,. \tag{3.1}$$

Observe that the variable of integration y is greater than x on the top and less than x on the bottom. Thus an upper bound for f''/f' will simultaneously bound the numerator from above and the denominator from below. A lower bound for f''/f' will do the reverse. This is the basis for the following Lemma.

Lemma 1 *Suppose the $C^{2,1}_{\mathrm{loc}}$ function f is normalized and satisfies $SG_s \le Sf \le SF_r$ on $(-1,1)$.*
(i) If $0 \le x-h < x$ then

$$kG_s(x,h) \le kf(x,h) \le kF_r(x,h). \tag{3.2}$$

(ii) If $x < x+h \le 0$ then

$$kF_r(x,h) \le kf(x,h) \le kG_s(x,h). \tag{3.3}$$

(iii) If $x-h < 0 \le x$ then

$$\frac{G_s(x+h)-G_s(x)}{G_s(x)-F_r(x-h)} \le kf(x,h) \le \frac{F_r(x+h)-F_r(x)}{F_r(x)-G_s(x-h)}. \tag{3.4}$$

(iv) If $x \le 0 < x+h$ then

$$\frac{G_s(x+h)-F_r(x)}{F_r(x)-F_r(x-h)} \le kf(x,h) \le \frac{F_r(x+h)-G_s(x)}{G_s(x)-G_s(x-h)}. \tag{3.5}$$

Proof. As in Section 2, let u and v be the solutions of the initial value problems

$$u'' + \frac{1}{2}(Sf)u = 0,\ u(0)=1,\ u'(0)=0, \tag{3.6}$$

$$v'' + \frac{1}{2}(SF_r)v = 0,\ v(0)=1,\ v'(0)=0, \tag{3.7}$$

where the first equation is meant to hold a.e.. Note first that

$$\frac{f''}{f'} = -2\frac{u'}{u}\,, \quad \frac{F_r''}{F_r'} = -2\frac{v'}{v}\,. \tag{3.8}$$

Using the differential equations and the right hand inequality in (3.2) we have

$$(u'v - uv')' = \frac{1}{2}(SF_r - Sf)uv \geq 0,$$

and hence $u'v - uv' \geq 0$ because of the initial conditions. In other words

$$-2\frac{u'}{u} \leq -2\frac{v'}{v}, \quad \text{on } [0,1), \tag{3.9}$$

since u and v are positive. As promised, we can now conclude that

$$\begin{aligned} kf(x,h) &= \frac{\displaystyle\int_x^{x+h} \exp\left(-2\int_x^y \frac{u'}{u}(t)\,dt\right)\,dy}{\displaystyle\int_{x-h}^{x} \exp\left(-2\int_x^y \frac{u'}{u}(t)\,dt\right)\,dy} \\ &\leq \frac{\displaystyle\int_x^{x+h} \exp\left(-2\int_x^y \frac{v'}{v}(t)\,dt\right)\,dy}{\displaystyle\int_{x-h}^{x} \exp\left(-2\int_x^y \frac{v'}{v}(t)\,dt\right)\,dy} = kF_r(x,h), \end{aligned} \tag{3.10}$$

as long as $0 \leq x - h$. This proves the right hand inequality in (3.2). For the left hand inequality we let w be the solution to

$$w'' + \frac{1}{2}(SG_s)w = 0,\ w(0) = 1,\ w'(0) = 0, \tag{3.11}$$

and we find, in the same manner as above, that

$$-2\frac{u'}{u} \geq -2\frac{w'}{w}, \quad \text{on } [0,1). \tag{3.12}$$

This leads to the lower bound $kf \geq kG_s$ in (3.5) and completes the proof of Part (i).

The inequalities in Part (ii) follow from those in (3.2) of Part (i). For $y \in [0,1)$ let $g(y) = -f(-y)$. Then g is increasing, normalized and $Sg(y) = Sf(-y)$. Hence $kG_s(y,h) \leq kg(y,h) \leq kF_r(y,h)$ when $y - h \geq 0$. But now, it is easy to check that $kg(y,h) = kf(-y,h)^{-1}$, and also that $kF_r(y,h) = kF_r(-y,h)^{-1}$ and $kG_s(y,h) = kG_s(-y,h)^{-1}$, the latter two identities holding because F_r and G_s are odd. Flipping the inequalities, and the hypotheses, and writing x for $-y$, we obtain (3.3).

For the proof of the inequalities (3.4) in Part (iii) we have to mix the estimates for u'/u on either side of 0. We treat the numerator and denominator of (3.1) separately. To do this we have available in addition to (3.9) and (3.12), and for the same reasons, the two bounds

$$-2\frac{u'}{u} \leq -2\frac{w'}{w}, \quad \text{on } (-1,0], \tag{3.13}$$

and

$$-2\frac{u'}{u} \geq -2\frac{v'}{v}, \quad \text{on } (-1,0]. \tag{3.14}$$

Suppose that $x - h < 0 \le x$. Since $x + h > x \ge 0$ we find from (3.9) that

$$\begin{aligned} f(x+h) - f(x) &= f'(x) \int_x^{x+h} \exp\left(-2 \int_x^y \frac{u'}{u}(t)\, dt\right) dy \\ &\le f'(x) \int_x^{x+h} \exp\left(-2 \int_x^y \frac{v'}{v}(t)\, dt\right) dy \\ &= f'(x) v(x)^2 \left\{F_r(x+h) - F_r(x)\right\}. \end{aligned} \tag{3.15}$$

Next, to estimate $f(x) - f(x-h)$ from below we write

$$f(x) - f(x-h) = \\ f'(x) \left\{ \int_{x-h}^0 \exp\left(-2 \int_x^0 \frac{u'}{u}(t)\, dt - 2 \int_0^y \frac{u'}{u}(t)\, dt\right) dy + \int_0^x \exp\left(-2 \int_x^y \frac{u'}{u}(t)\, dt\right) dy \right\}.$$

In the first exponentiated integral we use (3.9), in the second we use (3.13), and in the third we use (3.9). This, with $v(0) = w(0) = 1$ and $F_r(0) = G_s(0) = 0$, gives

$$\begin{aligned} f(x-h) - f(x) &\ge f'(x) \left\{ v(x)^2 \int_{x-h}^0 w^{-2}(y)\, dy + v(x)^2 \int_0^x v^{-2}(y)\, dy \right\} \\ &= f'(x) v(x)^2 \left\{F_r(x) - G_s(x-h)\right\}. \end{aligned} \tag{3.16}$$

Combining (3.15) with (3.16) we obtain

$$kf(x,h) \le \frac{F_r(x+h) - F_r(x)}{F_r(x) - G_s(x-h)}.$$

The proof of the lower bound in (3.4) follows along the same lines. First, using (3.12) we obtain the lower bound

$$f(x+h) - f(x) \ge f'(x) w(x)^2 \{G_s(x+h) - G_s(x)\}.$$

Next, splitting the integrals again in a way that makes it possible to apply (3.12), (3.14) and (3.12) yields the upper bound

$$f(x) - f(x-h) \le f'(x) w(x)^2 \{G_s(x) - F_r(x-h)\}.$$

Combining these gives the lower bound for $kf(x,h)$ in (3.4).

This proves Part (iii) of the Lemma. Fortunately, we can deduce the inequalities (3.5) in Part (iv), in the case when $x \le 0$ but $x + h > 0$, via the same trick we used in the proof of Part (ii). That is, we can apply the inequalities in Part (iii) to the function $g(y) = -f(-y)$. Using as before the identity $kg(y,h) = kf(-y,h)^{-1}$ and the fact that F_r and G_s are odd leads quickly from (3.4) to (3.5). This completes the proof of Lemma 1.

There is an aspect of the proof of this Lemma which we will use in Section 4. Namely, the comparison theorem gives locally uniform bounds for $|f|$ and for f', with the latter being bounded below away from zero. Hence from the bounds on f''/f' we obtain bounds for $|f''|$, and then the bounds on the Schwarzian entail bounds in L^∞_{loc} for f'''.

It follows from the results in [2] on quasiconformal extensions that F_r and G_s are quasisymmetric on $(-1,1)$. This is not obvious because $F_r'(x) = O((1-\ x^2)^{r-1})$, $G_s'(x) = O((1-x^2)^{s-1})$ as $x \to \pm 1$. With some effort, using primarily the concavity, one can show directly that F_r and G_s are k-quasisymmetric with

$$k = \frac{2^{1-r}}{2^r - 1} \text{ for } F_r, \tag{3.17}$$

and

$$k = \frac{2^s - 1}{2^{1-s}} \text{ for } G_s. \tag{3.18}$$

We will not give the details of the calculations. One may thus view Parts (i) and (ii) of Lemma 1 as stating that a normalized map having the given bounds on its Schwarzian is $k-$quasisymmetric on $(-1,0)$ and on $(0,1)$ with

$$k = \max\left\{\frac{2^{1-r}}{2^r-1}, \frac{2^s-1}{2^{1-s}}\right\}$$

for both intervals.

Having obtained the estimates for kF_r and kG_s in (3.17) and (3.18), it turns out to be easier to write the inequalities (3.4), (3.5) in the second half of the Lemma in the form

$$\begin{aligned} \frac{G_s(x) - G_s(x-h)}{G_s(x) - F_r(x-h)} kG_s(x,h) &\le kf(x,h) \qquad (3.19)\\ &\le kF_r(x,h)\frac{F_r(x) - F_r(x-h)}{F_r(x) - G_s(x-h)}, \qquad (3.20) \end{aligned}$$

for $x - h < 0 \le x$, and

$$\begin{aligned} \frac{G_s(x+h) - F_r(x)}{F_r(x+h) - F_r(x)} kF_r(x,h) &\le kf(x,h) \qquad (3.21)\\ &\le kG_s(x,h)\frac{F_r(x+h) - G_s(x)}{G_s(x+h) - G_s(x)} \qquad (3.22) \end{aligned}$$

for $x \le 0 < x + h$, and to estimate the *new* quantities which appear. One can show that on the right hand sides the new quantities tend to a maximum of s/r as $(x,h) \to (0,1)$, and that on the left they tend to a minimum of r/s as (x,h) tends to the same point. Again this uses strongly the concavity of the individual functions and also the fact that F_r and G_s can be pieced together to give smooth functions which do not change concavity at the origin. Again, we omit the details.

We collect all these estimates together with Lemma 1 as a Theorem:

Theorem 2 *Suppose the $C^{2,1}_{\text{loc}}$ function f is normalized and satisfies $SG_s \le Sf \le SF_r$ on $(-1,1)$. Then f is $k(r,s)-$quasisymmetric with*

$$k(r,s) = \frac{s}{r}\max\left\{\frac{2^{1-r}}{2^r-1}, \frac{2^s-1}{2^{1-s}}\right\}.$$

The particular value for k is not so important, but it is important that it tends to 1 as $r, s \to 1$. This fact also follows from the estimates in [2].

What happens if a function satisfies $SG_s \le Sf \le SF_r$ but is not normalized? If $f(1)$ or $f(-1)$ are infinite, which could happen, then kf can tend to zero or to infinity. We study this problem in the following way. The quasisymmetry quotient is unaffected by affine transformations of the function, so we may continue to assume at the outset that $f(0) = 0$ and $f'(0) = 1$. As explained in Section 2, the upper bound $Sf \le SF_r$ allows us to normalize further by defining $f^\dagger = f/(1 + a_2 f)$, $a_2 = (1/2)f''(0)$. The quasisymmetry quotients of f and $f^\dagger$ are related by

$$kf^\dagger(x, h) = \frac{1 + a_2 f(x-h)}{1 + a_2 f(x+h)} kf(x, h) = qf(x, h)kf(x, h). \tag{3.23}$$

Theorem 2 provides estimates for $kf^\dagger$ and we want to estimate qf from above and below. We give two ways of doing this. One is by making the assumption that $f(-1)$ and $f(1)$ are bounded, and the other is by restricting the size of a_2. The latter actually has that f is bounded as a consequence. We recall from (2.13) in Section 2 that we *always* have the coefficient estimate $|a_2| \le s$ when f satisfies the lower bound $SG_s \le Sf$.

Lemma 2 *Let $f \in C^{2,1}_{\text{loc}}(-1, 1)$ satisfy $SG_s \le Sf \le Sf_r$ with $f(0) = 0$, $f'(0) = 1$.*
(i) Suppose $-\infty < a = f(-1) < 0 < f(1) = b < \infty$, and let $m = \max\{b/|a|, |a|/b\}$. Then

$$\frac{1}{m}\frac{r}{s} \le qf \le \frac{s}{r}m. \tag{3.24}$$

(ii) If $|a_2| < r$, then $-\infty < a = f(-1) < 0 < f(1) = b < \infty$, and

$$\frac{r - |a_2|}{r + |a_2|} \le qf \le \frac{r + |a_2|}{r - |a_2|}. \tag{3.25}$$

Note that the bounds in Part (ii) tend to 1 as $a_2 \to 0$.

Proof. We prove Part (i) first, and we may suppose that $a_2 \ne 0$. Since $f^\dagger(0) = f(0) = 0$ and $(f^\dagger)'(0) = f'(0) = 1$ it follows that f and $f^\dagger$ have the same sign. Hence $1 + a_2 f(x) > 0$ for all x. Write

$$qf = \frac{f(x-h) + \dfrac{1}{a_2}}{f(x+h) + \dfrac{1}{a_2}}.$$

From

$$a + \frac{1}{a_2} \le f(x) + \frac{1}{a_2} \le b + \frac{1}{a_2}$$

and

$$\begin{aligned} a + \frac{1}{a_2} &= \frac{1}{a_2}(1 + a_2 f(-1)) > 0, \quad \text{if}\, a_2 > 0, \\ b + \frac{1}{a_2} &= \frac{1}{a_2}(1 + a_2 f(1)) < 0, \quad \text{if}\, a_2 < 0, \end{aligned}$$

we have

$$\frac{a+\dfrac{1}{a_2}}{b+\dfrac{1}{a_2}} \leq qf \leq \frac{b+\dfrac{1}{a_2}}{a+\dfrac{1}{a_2}}\,, \tag{3.26}$$

while if $a_2 < 0$ then

$$\frac{b+\dfrac{1}{a_2}}{a+\dfrac{1}{a_2}} \leq qf \leq \frac{a+\dfrac{1}{a_2}}{b+\dfrac{1}{a_2}}. \tag{3.27}$$

Now write $f = f^\dagger/(1-a_2 f^\dagger)$ and

$$\frac{b+\dfrac{1}{a_2}}{a+\dfrac{1}{a_2}} = \frac{1-a_2 f^\dagger(-1)}{1-a_2 f^\dagger(1)}.$$

Then using the estimates for normalized functions (2.7), (2.8), that is, $f^\dagger(-1) \geq -1/r$ and $f^\dagger(1) \geq 1/s$, we have

$$\frac{1}{1-a_2 f^\dagger(1)} = \frac{b}{f^\dagger(1)} \leq sb,$$
$$1-a_2 f^\dagger(-1) = \frac{f^\dagger(-1)}{a} \leq \frac{-1}{ar}.$$

Hence for $a_2 > 0$

$$\frac{r}{s}\frac{|a|}{b} \leq qf \leq \frac{s}{r}\frac{b}{|a|},$$

while for $a_2 < 0$ this is rearranged to

$$\frac{r}{s}\frac{b}{|a|} \leq qf \leq \frac{s}{r}\frac{|a|}{b}.$$

Combining these yields (3.24).

The proof of (3.25) in Part (ii) relies on differential equations. Suppose $0 < |a_2| < r$. Then the value $1/a_2$ is not attained by the function F_r, so that $H = F_r/(1-a_2F_r)$ is regular on $(-1,1)$. In fact, if v is the solution of the initial value problem

$$v'' + \frac{(1-r^2)}{(1-x^2)^2}v = 0, \quad v(0) = 1,\, v'(0) = -a_2,$$

then $v = (H')^{-1/2}$. We can apply the comparison theorem to conclude that if f satisfies $f(0) = 0, f'(0) = 1, f''(0) = 2a_2$, and $Sf(x) \leq 2(1-r^2)(1-x^2)^{-2}$ then

$$|f(x)| \leq |H(x)| = \left|\frac{F_r(x)}{1-a_2F_r(x)}\right|$$

on $(-1,1)$. Hence

$$b = f(1) \le \frac{F_r(1)}{1 - a_2 F_r(1)} = \frac{1}{r - a_2},$$
$$a = f(-1) \ge \frac{F_r(-1)}{1 - a_2 F_r(-1)} = \frac{-1}{r + a_2}.$$

If again we treat separately the cases $a_2 > 0$ and $a_2 < 0$ then these last inequalities combined with (3.26) and (3.27) lead to the inequalities in (3.25), and so completes the proof of the Lemma.

We now see that we can drop the normalization hypothesis in Theorem 2 provided we replace it by either assumption in Lemma 2, and modify the quasisymmetry constant accordingly.

Finally, readers familiar with the role of the Schwarzian in the theory of univalent functions may wonder if there is a sufficient condition for quasisymmetry in terms of f''/f'. We raised this question for different reasons in [2], and we treat it here only briefly. There are similarities to the situation with the Schwarzian, but there is also an interesting difference.

Let $0 \le t < 1$ and let L_t and M_t be solutions of

$$\frac{L_t''}{L_t'}(x) = \frac{2t}{1-x^2} \quad \text{and} \quad \frac{M_t''}{M_t'}(x) = \frac{-2t}{1-x^2}$$

on $(-1,1)$, with $L_t(0) = M_t(0) = 0$ and $L_t'(0) = M_t'(0) = 1$. Then

$$L_t(x) = \int_0^x \left(\frac{1+y}{1-y}\right)^t dy\,, \quad M_t(x) = \int_0^x \left(\frac{1-y}{1+y}\right)^t dy\,. \tag{3.28}$$

These are the extremals for bounding f''/f' corresponding to F_r and G_s for the Schwarzian. They are hypergeometric functions, so we lose the elementary nature of some of the estimates. More importantly, for $t \ge 1$, $L_t(1) = +\infty$, $M_t(-1) = -\infty$, and *both* L_t and M_t fail to be quasisymmetric, whereas this did not happen with the lower extremal G_s for the Schwarzian. (However, the failure of quasisymmetry here is a little more subtle. See Remark 3 at the end of this Section.) Thus with f''/f', for all intents and purposes it makes sense to consider only *symmetric* upper and lower bounds.

Theorem 3 *If f satisfies*

$$\left|\frac{f''}{f'}(x)\right| \le \frac{2t}{1-x^2} \tag{3.29}$$

on $(-1,1)$ for some $0 \le t < 1$, then

$$kM_t(x,h) \le kf(x,h) \le kL_t(x,h)$$

for $x, x-h, x+h \in (-1,1)$, $h > 0$. For $0 \le t < 1$, the function f is $\alpha(t)-$quasisymmetric, where $\alpha(t) = -L_t(1)/L_t(-1)$.

Actually, the proof will show that we can give the quasisymmetry bounds for kf in terms of either extremal function. We did not make an issue of the smoothness of f, but in the spirit of the earlier results $C^{1,1}_{\mathrm{loc}}$ would suffice.

Proof. We may assume that $f(0) = 0, f'(0) = 1$. First, it follows easily that $f'(x) = O((1-x)^{-t})$ as $x \to 1$, hence $f(1) < \infty$. Similarly $f(-1) > -\infty$. Now write

$$\frac{f''}{f'}(x) = \frac{2tg(x)}{1-x^2}\,,$$

where $|g(x)| \leq 1$. Next, we write the quasisymmetry quotient as

$$kf(x,h) = \frac{\displaystyle\int_x^{x+h} \exp\left(2t\int_x^y \frac{g(\tau)}{1-\tau^2}\,d\tau\right)\,dy}{\displaystyle\int_{x-h}^{x} \exp\left(2t\int_x^y \frac{g(\tau)}{1-\tau^2}\,d\tau\right)\,dy}\,. \tag{3.30}$$

The choice $g = 1$ both maximizes the numerator and minimizes the denominator. Hence, using (3.28),

$$kf(x,h) \leq \frac{\displaystyle\int_x^{x+h} \exp\left(2t\int_x^y \frac{1}{1-\tau^2}\,d\tau\right)\,dy}{\displaystyle\int_{x-h}^{x} \exp\left(2t\int_x^y \frac{1}{1-\tau^2}\,d\tau\right)\,dy} = \frac{\displaystyle\int_x^{x+h} \left(\frac{1+y}{1-y}\right)^t\,dy}{\displaystyle\int_{x-h}^{x} \left(\frac{1+y}{1-y}\right)^t\,dy} = kL_t(x,h)\,.$$

Likewise, if we choose $g = -1$ in (3.30) then we obtain

$$kM_t(x,h) \leq kf(x,h).$$

This proves the first part of the Theorem.

We now want to estimate $kM_t(x,h)$ and $kL_t(x,h)$. Suppose first that $x \geq 0$. Because L_t is concave up, $kL_t(x,h) \leq kL_t(x,1-x)$, and with some work we find that

$$\max_{0\leq x\leq 1} kL_t(x,1-x) = \frac{L_t(1)-L_t(0)}{L_t(0)-L_t(-1)} = -\frac{L_t(1)}{L_t(-1)} = \alpha(t) < \infty$$

for all $0 \leq t < 1$. Similarly, because M_t is concave down, $kM_t(x,h) \geq kM_t(x,1-x)$ for $0 \leq x \leq 1$, and here the result is

$$\min_{0\leq x\leq 1} kM_t(x,1-x) = \frac{M_t(1)-M_t(0)}{M_t(0)-M_t(-1)} = -\frac{M_t(1)}{M_t(-1)} = \beta(t) > 0\,.$$

Thus for $0 \leq x < 1$

$$\beta(t) \leq kf(x,h) \leq \alpha(t).$$

To get bounds for $kf(x,h)$ when $x < 0$ we employ the familiar trick of considering the function $g(x) = -f(-x)$ and applying what has already been proved. This leads to

$$\frac{1}{\alpha(t)} \leq kf(x,h) \leq \frac{1}{\beta(t)}$$

for $-1 < x < 0$. Therefore, for all x, h

$$\min\{\beta(t), \frac{1}{\alpha(t)}\} \leq kf(x,h) \leq \max\{\alpha(t), \frac{1}{\beta(t)}\}.$$

But from (3.28), $\alpha(t) = 1/\beta(t)$, whence

$$\frac{1}{\alpha(t)} \leq kf(x,h) \leq \alpha(t)$$

as desired. Note that $\alpha(t) = \infty$ for $t \geq 1$ so, as we remarked earlier, both L_t and M_t fail to be quasisymmetric in this range.

Remark 2 We also need versions of results in this Section for functions on the interval $(-R, R)$. We also recall Remark 1 in Section 2. In Lemma 1 we need only replace F_r and G_s by $\tilde{F}_r$ and $\tilde{G}_s$ for the statement and the proof to remain otherwise unchanged. More importantly, the bounds for $k\tilde{F}_r$ and $k\tilde{G}_s$ are the same as for kF_r and kG_s in (3.17) and (3.18). The same is true for the estimates of the mixed quantities in (3.19)–(3.22). That is, the bound for the quasisymmetry quotient in Theorem 2 will be the same for normalized maps on $(-R, R)$ satisfying $S\tilde{G}_s \leq Sf \leq S\tilde{F}_r$.

The situation in Lemma 2 is a little different. Again we replace F_r and G_s by $\tilde{F}_r$ and $\tilde{G}_s$. With $a = f(-R) < 0 < f(R) = b$ the statement in (3.24) is unchanged. For the second part of the Lemma we make the assumption that $|a_2| \leq r/R$ (a corresponding strengthening of the estimate $|a_2| \leq s/R$). Then (3.25) is replaced by

$$\frac{r - |a_2|R}{r + |a_2|R} \leq qf \leq \frac{r + |a_2|R}{r - |a_2|R}. \tag{3.31}$$

Remark 3 It is possible to refine the differential equations arguments we have used to obtain the following result, whose proof we will not give here.

Theorem *Suppose that $f \in C^{2,1}_{\mathrm{loc}}(-1,1)$ satisfies*

$$-\infty < \liminf_{|x|\to 1}(1-x^2)^2 Sf(x) \quad \textit{and} \quad \limsup_{|x|\to 1}(1-x^2)^2 Sf(x) < 2.$$

Then either $f(1) = -f(-1) = \infty$ or else some Möbius transformation of f is quasisymmetric on $(-1,1)$.

(The Schwarzian is in L^∞_{loc} so the hypotheses have to be interpreted in the distributional sense. For instance, to say that the lim sup as $x \to 1$ is < 2 means that there exist x_0 and $b < 2$ such that $(1-x^2)^2 Sf(x) \leq b$ on $[x_0, 1)$ in the distributional sense.)

This is analogous to a theorem of Gehring and Pommerenke [7] on univalent functions with a quasiconformal extension. It has an interesting consequence for the extremal functions L_t and M_t used to bound f''/f'. Taking L_t, for example, we compute that

$$SL_t(x) = \frac{4tx - 2t^2}{(1-x^2)^2}.$$

Hence the limits of $(1-x^2)^2 SL_t(x)$ as $x \to \pm 1$ are $-4t-2t^2$ at -1, and $4t-2t^2$ at 1, thus $> -\infty$ in either case. The limit at -1 is ≤ 0 and the limit at 1 is < 2 if $t > 1$. Since L_t maps only the endpoint $+1$ to infinity we see from the Theorem above that when $t > 1$ some Möbius transformation of L_t will be quasisymmetric. A similar discussion holds for M_t, getting the same limits but at the opposite endpoints. So for $t > 1$ the failure of quasisymmetry of the extremals L_t, M_t can be eliminated via a Möbius transformation. The catch is that, unlike the Schwarzian, the expression f''/f' is *not* invariant under general Möbius transformations. On the other hand, for $t = 1$, $L_1(x) = \log(1-x)^{-2} + x$, $M_1(x) = \log(1+x)^2 - x$, and no Möbius transformation will make these functions quasisymmetric on $(-1, 1)$.

Remark 4 The quasisymmetry quotient determines a function up to a similarity. For suppose $kf = kg$. We may first apply similarity transformations to obtain $f(0) = g(0) = 0$ and $f(1) = g(1) = 1$, and then it is easy to show that $f = g$ at dyadic points. This implies $f = g$ under only the assumption of continuity. If we allow for some differentiability the same uniqueness statement follows quite differently from

$$(\frac{\partial}{\partial h} kf)(x, 0) = \frac{f''}{f'}(x). \tag{3.32}$$

This equation also allows one to represent f directly in terms of its quasisymmetry quotient, though not in a particularly interesting way.

A problem which we do not address, but which has been in the background of much of our work, is the corresponding question of existence. To what extent can one prescribe the quasisymmetric distortion, not just the bounds, but the positive, bounded function that measures the distortion at each point and at each scale? Allowing again for some differentiability, there are several other interesting identities which the quasisymmetry quotient must satisfy, and which might cast some shadow as necessary and sufficient conditions for an existence theorem for continuous quasisymmetric functions. For example, one also has

$$(\Delta kf)(x, 0) = \left(\frac{f''}{f'}(x)\right)^2.$$

A trivial consequence (for smooth maps, at least) is that kf is harmonic if and only if f is a similarity.

One can also get the Schwarzian derivative out of the quasisymmetry quotient. If we change coordinates to $u = x + h$ and $v = x - h$ then

$$(\frac{\partial^2}{\partial v^2} kf)(u, u) = -\frac{1}{2} Sf(u).$$

Unfortunately we have not been able to make much use of these and other similar identities.

4 Schwarz-Pick Lemmas and $C^{2,1}$ Smoothness

Let J be an interval (a, b). By analogy with a two dimensional disk we define

$$\lambda_J(t)dt = \frac{(b-a)dt}{2(b-t)(t-a)} \tag{4.1}$$

to be the Poincaré metric for J, and

$$d_J(x,y) = \left|\int_x^y \lambda_J(t)\,dt\right| = \left|\log \frac{(y-a)(b-x)}{(x-a)(b-y)}\right| \tag{4.2}$$

to be the corresponding hyperbolic distance. For the Poincaré metric, $\lambda_J(t)$ is the arithmetic mean of the reciprocals of the distances from t to the endpoints, while it is often helpful to view the distance $d_J(x,y)$ as the logarithm of a cross-ratio. For instance the invariance of the hyperbolic distance under Möbius transformations is a visible consequence of the latter. We discuss this briefly at the end of this Section.

If J is the centered interval (x_0-h, x_0+h) then the hyperbolic metric takes the form

$$\lambda_J(x)dx = \frac{h\,dx}{h^2-(x-x_0)^2}\,. \tag{4.3}$$

Now let f be an increasing function. We compare the Poincaré metrics on (x_0-h, x_0+h) and $(f(x_0-h), f(x_0+h))$ and find that whenever $Sf(x_0)$ exists we can write

$$\frac{\lambda_{f(J)}(f(x_0))f'(x_0)}{\lambda_J(x_0)} = 1 - \frac{1}{6}Sf(x_0)h^2 + o(h^2). \tag{4.4}$$

For C^4 functions the next non-zero term would be $O(h^4)$ because the left hand side is actually even in h. Infinitesimally, a function with a negative Schwarzian therefore increases hyperbolic distances, while a function with a positive Schwarzian decreases hyperbolic distances. This phenomenon on a global scale, much discussed in dynamics, is the subject of this Section, and the extremal functions F_r and G_s are the models.

Let I be the interval $(-1,1)$. The Möbius transformation $P(x) = (1+x)/(1-x)$ is an isometry of $(I, \lambda_I dx)$ and the positive half-line $\mathbf{R}^+$ with its Poincaré metric $\lambda_{\mathbf{R}^+}(x)dx = dx/2x$. For any $\alpha > 0$ the map $y = x^\alpha$ is a smooth, incereasing map of $\mathbf{R}^+$ to itself with $dy/2y = \alpha(dx/2x)$. It decreases or increases hyperbolic distances on $\mathbf{R}^+$ when $\alpha < 1$ or $\alpha > 1$, respectively. Now set $\phi(x) = (1/\alpha)(P^{-1}(P(x)^\alpha))$. Then $F_r(x)$ and $G_s(x)$ are $\phi(x)$ for $\alpha = r$ and s, respectively. Furthermore, the extremals have the stronger property of being distance decreasing, or increasing, on all subintervals, though not by a constant amount as they do for the whole interval. That is, if $J \subseteq (-1,1)$ is any open subinterval then

$$d_{F_r(J)}(F_r(x), F_r(y)) \le d_J(x,y), \quad d_J(x,y) \le d_{G_s(J)}(G_s(x), G_s(y)).$$

This can be checked directly, but it also has to do precisely with the Schwarzians being of one sign.

We will have to talk about functions which are increasing or decreasing in the ordinary sense along with functions which increase or decrease hyperbolic distances. To keep this straight with as few words as possible, we will refer to the latter properties as *expanding* or *contracting*. Observe that if f is contracting then f^{-1} is expanding on the range of f, and vice-versa.

We start with some Schwarz-Pick type inequalities, in an infinitesimal form, under the assumption that the function is $C^{2,1}_{\mathrm{loc}}$.

Lemma 3 *Let $I = (-1, 1)$ and let $f : I \to \mathbf{R}$ be an increasing $C^{2,1}_{\text{loc}}$ function. Let $J \subseteq I$ be an open interval in I.*

(a) If $SG_s \leq S \leq SF_r$ then

$$r\lambda_I(x) \leq \lambda_{f(I)}(f(x))f'(x) \leq s\lambda_I(x), \ x \in I. \tag{4.5}$$

Equality at a single point in either inequality in (4.5) implies that f is a Möbius conjugation of the corresponding extremal F_r or G_s.

(b) $Sf \leq 0$ on I if and only if $\lambda_J(x) \leq \lambda_{f(J)}(f(x))f'(x)$, $x \in J$, for all J. If equality holds at a single point in the latter inequality then f is a Möbius transformation on J.

(c) $Sf \geq 0$ on I if and only if $\lambda_{f(J)}(f(x))f'(x) \leq \lambda_J(x)$, $x \in J$, for all J. If equality holds at a single point in the latter inequality then f is a Möbius transformation on J.

The fact that functions with a positive (negative) Schwarzian are contracting (expanding) in the hyperbolic metric is due to de Melo and van Strien [4] using cross-ratio. We thought it was worthwhile to give a different proof in the present context, especially because the differential equations argument we use also gives the case of equality.

Proof. For Part (a) we first show that

$$r \leq \lambda_{f(I)}(f(0))f'(0) \leq s. \tag{4.6}$$

As always, we may assume that $f(0) = 0, f'(0) = 1$ without changing (4.6) and we may further normalize to the function $f^\dagger = f/(1+a_2 f)$, with $a_2 = (1/2)f''(0)$, without introducing any singularities. Since Möbius transformations are hyperbolic isometries it then suffices to show that

$$r \leq \lambda_{f^\dagger(I)}(0) \leq s \tag{4.7}$$

in order to deduce (4.6).

Let $f^\dagger(-1) = a < 0 < f^\dagger(1) = b$. Then from (4.3) we get

$$\lambda_{f^\dagger(I)}(0) = \frac{b-a}{-2ab} = \frac{1}{2}\left(\frac{1}{b} - \frac{1}{a}\right). \tag{4.8}$$

The inequalities (2.7) and (2.8) now give

$$-\frac{1}{r} \leq a \leq -\frac{1}{s}, \quad \frac{1}{s} \leq b \leq \frac{1}{r},$$

from which we obtain (4.7). If equality holds in (4.7) in either inequality this forces *both* a and b to have the corresponding extreme value. This implies that $f^\dagger$ is the same extremal function on each interval $(-1, 0]$, $[0, 1)$.

The general result (4.5) at a point $x_0 \in (-1, 1)$ follows by considering $y = (x + x_0)/(1 + x_0 x)$ and $h(x) = f(y)$. Then $(1-x^2)^2 Sh(x) = (1-y^2)^2 Sf(y)$. Therefore $r \leq \lambda_{h(I)}(0)g'(0) \leq s$, while also

$$\lambda_{h(I)}(h(0))h'(0) = \lambda_{h(I)}(f(x_0))f'(x_0)(1 - {x_0}^2) = \lambda_{f(I)}(f(x_0))f'(x_0)(1 - {x_0}^2).$$

The case of equality stated in Part (a) also follows, since if equality holds in (4.5) at some point x_0 we can precompose f with a Möbius transformation of the interval to itself to assume that the point is 0, and then apply the previous argument.

To prove Part (b) we first show that if $Sf \leq 0$ on I then

$$\lambda_I(x) \leq \lambda_{f(I)}(f(x))f'(x), \ x \in I.$$

Because the Schwarzian is negative we can define $f^\dagger$, as above, without introducing a singularity. As in Part (a) it is then enough to show that $1 \leq \lambda_{f^\dagger(I)}(0)$, where $\lambda_{f^\dagger(I)}(0)$ is given by (4.8). This follows from (2.10) with $r = 1$, according to which $a \geq -1$ and $b \leq 1$. Furthermore $\lambda_{f^\dagger(I)}(0) = 1$ if and only if $b = -a = 1$, which can only happen if $f^\dagger$ is the identity, hence f is Möbius. We get the expanding property on a subinterval J, and also the case of equality, by considering $f \circ \varphi$, where φ is an affine map of I to J. This proves the sufficiency in Part (b). The necessity follows from (4.4).

Part (c) follows easily from Part (b) by considering f^{-1}. It is also possible to give a direct proof along the lines of Part (b), with the complication that when Sf is only bounded below one cannot normalize without possibly introducing a singularity. Thus it is necessary to distinguish a number of cases, and we will not give this version of the proof.

We have one further comment about Parts (b) and (c). Though we have not been able to formulate a general statement, it seems that the property of a function being contracting or expanding has to do with the Schwarzian being 'mostly' of one sign on an interval. For example, the function $f(x) = x^3 + x$ has $Sf(x) = 6(1 - 6x^2)/(1 + 3x^2)^2$, hence $Sf(x) \geq 0$ if and only if $x^2 \leq 1/6$. So for certain f is contracting as a map from J to $f(J)$ where J is any subinterval of $(-1/\sqrt{6}, 1/\sqrt{6})$. However, one can check that f is still contracting as a map from $(-1/2, 1/2)$ to $(-5/8, 5/8)$ $\quad (f(1/2) = 5/8)$, though it will be *expanding* on small intervals contained in $(-1/2, 1/2)$ near the endpoints since the Schwarzian will be negative between $\pm 1/\sqrt{6}$ and $\pm 1/2$.

Corollary 1 *Let $f : I \to \mathbf{R}$ be an increasing $C^{2,1}_{\text{loc}}$ function. Then $SG_s \leq Sf \leq SF_r$ on I if and only if*

$$\lambda_{F_r(J)}(F_r(x)){F_r}'(x) \leq \lambda_{f(J)}(f(x))f'(x) \leq \lambda_{G_s(J)}(G_s(x)){G_s}'(x), \tag{4.9}$$

$x \in J$, for all open subintervals $J \subseteq I$. If equality holds at a single point in either inequality in (4.9) then f is the corresponding extremal function on J up to a Möbius transformation of J.

This follows from Parts (b) and (c) of the Lemma via the chain rule for the Schwarzian (1.3) applied to the compositions $f{F_r}^{-1}$ and $f{G_s}^{-1}$.

We next have two Lemmas *implying* degrees of smoothness of functions on $(-1, 1)$ when the change in hyperbolic distance is controlled. In the first we ask that the function be contracting in the same strong sense as the extremal F_r, i.e., that it be contracting on all subintervals. In the second, it is the version of (4.9) for hyperbolic distances, not the infinitessimal statement in terms of the metric, that is the key to proving $C^{2,1}$ smoothness.

Lemma 4 *Let $f : I \to \mathbf{R}$ be an increasing function. Suppose that for every open subinterval $J \subseteq I$*

$$d_{f(J)}(f(x), f(y)) \leq d_J(x, y) \tag{4.10}$$

for all $x, y \in J$. Then f is C^1 on I. If $f'(x) = 0$ for any x then f is constant, otherwise f' is never zero and $\log f'$ is a convex function.

Proof. It is easy to see that f is continuous on I. We show that it is differentiable there. Let $-1 < a < x < b < 1$ and let $J = (a, b)$. From (4.2), given $\epsilon > 0$ there is a $\delta > 0$ such that

$$d_J(x, y) \leq (1+\epsilon)\lambda_J(x)|x - y|,$$

and

$$d_{f(J)}(f(x), f(y)) \geq (1-\epsilon)\lambda_{f(J)}(f(x))|f(x) - f(y)|,$$

provided $|x - y| < \delta$. This yields the following upper bound for the difference quotient at x:

$$\left|\frac{f(x) - f(y)}{x - y}\right| \leq \frac{1+\epsilon}{1-\epsilon} \frac{b-a}{f(b) - f(x)} \frac{f(b) - f(x)}{b - x} \frac{f(x) - f(a)}{x - a}\,.$$

Taking the lim sup as $y \to x$ and then letting $\epsilon \to 0$ we conclude that

$$D^+ f(x) \leq \frac{b-a}{f(b) - f(a)} \frac{f(b) - f(x)}{b - x} \frac{f(x) - f(a)}{x - a}\,. \tag{4.11}$$

Now fix a and x and let $b \to x$ from the right along any sequence. Then (4.11) implies together with the continuity of f that

$$D^+ f(x) \leq \liminf_{b \to x^+} \frac{f(b) - f(x)}{b - x}\,.$$

Similarly,

$$D^+ f(x) \leq \liminf_{a \to x^-} \frac{f(x) - f(a)}{x - a}\,.$$

It follows that all the limits are the same and hence that $f'(x)$ exists. Thus (4.11) holds with $f'(x)$ in place of $D^+ f(x)$.

Next, writing down the hyperbolic distances from (4.2), the condition (4.10) with $a < x < y < b$ is

$$\frac{f(b) - f(x)}{f(b) - f(y)} \frac{f(y) - f(a)}{f(x) - f(a)} \leq \frac{b - x}{b - y} \frac{y - a}{x - a}\,, \tag{4.12}$$

which we rewrite as

$$\frac{f(b) - f(y)}{b - y} \frac{f(x) - f(a)}{x - a} \geq \frac{f(b) - f(x)}{b - x} \frac{f(y) - f(a)}{y - a}\,.$$

Knowing that f is differentiable we can let $x \to a, y \to b$ to obtain

$$f'(b) f'(a) \geq \left(\frac{f(b) - f(a)}{b - a}\right)^2\,. \tag{4.13}$$

The inequalities (4.13) and (4.11) for f' together show that f is C^1, lower semicontinuity following from the former and upper semicontinuity from the latter. Equality in (4.13) for all a, b characterizes Möbius tranformations. Also from (4.13), if $f'(x_0) = 0$ at any point x_0 then f is constant, so we now assume that $f' > 0$.

We next show that $\log f'$ is convex. Taking the logarithm in (4.13) we get

$$\frac{1}{2}(\log f'(b) + \log f'(a)) \geq \log \frac{f(b) - f(a)}{b - a},$$

and so we need to verify that

$$\log \frac{f(b) - f(a)}{b - a} \geq \log f'\left(\frac{a+b}{2}\right).$$

But from (4.11)

$$f'\left(\frac{a+b}{2}\right) \leq \frac{b-a}{f(b)-f(a)} \frac{f(b) - f\left(\frac{a+b}{2}\right)}{\frac{b-a}{2}} \frac{f\left(\frac{a+b}{2}\right) - f(a)}{\frac{b-a}{2}},$$

and we bound the right hand side from above by exactly $(f(b) - f(a))/(b - a)$ on applying the inequality between the arithmetic and geometric means to the numerators of the last two factors. This completes the proof.

Corollary 2 *If f satisfies the hypotheses of Lemma 4 and is not constant, then Sf exists as a locally L^1 function and $Sf(x) \geq 0$ wherever it is defined.*

Proof. According to Lemma 4 $\log f'$ is convex. From general facts on convex functions (see, for example [5]) one knows that $(\log f')' = f''/f'$ will be an increasing function with at most a countable number of jump discontinuities. By Lebesgue's theorem $(f''/f')'$ exists a.e. in I and is measurable, and so the same goes for Sf. It follows from the distance decreasing property and (4.4) that $Sf(x_0) \geq 0$ at a point x_0 where it exists. Sf is locally integrable because

$$\int_x^y \left(\frac{f''}{f'}\right)'(t)\, dt \leq \frac{f''}{f'}(y) - \frac{f''}{f'}(x). \tag{4.14}$$

This proves the Corollary, but we have a few additional comments. See Remark 7 at the end of this Section.

Next, we find that the smoothness improves if along with a function being contracting we ask that the amount by which it contracts be regulated from below by the extremal function F_r.

Lemma 5 *Let $f : I \to \mathbf{R}$ be a non-constant, increasing function. Suppose there is a number $0 < r \leq 1$ such that for every open subinterval $J \subseteq I$*

$$d_{F_r(J)}(F_r(x), F_r(y)) \leq d_{f(J)}(f(x), f(y)) \leq d_J(x, y) \tag{4.15}$$

for all $x, y \in J$. Then $f \in C^{2,1}_{\text{loc}}(I)$ and $Sf \geq 0$.

The heart of the proof is to manipulate f by composing it with the extremals or their inverses. We can do this most easily, without worrying about domains, if we rescale the extremals to rF_r and sG_s so that they map $(-1,1)$ onto itself. This affects neither the Schwarzian nor any hyperbolic distances. There are some other advantages to this, for if we write

$$\Phi_\alpha(x) = P^{-1}(P(x)^\alpha) \quad \text{with} \quad Px = \frac{1+x}{1-x} \tag{4.16}$$

as we did to see the expanding and contracting properties of F_r and G_s, then we easily find that

$$\Phi_\alpha \Phi_\beta = \Phi_{\alpha\beta}. \tag{4.17}$$

Thus for $\alpha > 0$ the $\{\Phi_\alpha\}$ form a one-parameter group of mappings of I onto itself.

Proof. By Lemma 4 $\log f'$ is a convex function. We reiterate that f''/f' is therefore a continuous, increasing function in the complement of a countable set in I, and at points in this countable set $\log f'$ has a left hand and right hand derivative with the former being smaller than the latter. At a jump of f''/f' the left-hand limit is equal to the left-hand derivative of $\log f'$ and similarly from the right side.

Let x_0 be a point where f''/f' has a presumed jump discontinuity. Without loss of generality we may assume that $x_0 = 0$. Furthermore, since the hypotheses are unaffected by a compositon Tf with a Möbius transformation we may 'partially normalize' f and assume that $f(0) = 0, f'(0) = 1$ and that the left hand derivative $D(\log f')(0^-) = 0$. Then the right hand derivative $D(\log f')(0^+)$ will be non-negative. This may introduce a singularity somewhere in the interval, but we will be working only in a neighborhood of the origin. To show that a jump in the second derivative cannot occur we bound the change in f''/f' at points on either side of the origin.

We now bring in the rescaled extremals (4.16), $\Phi_r = rF_r, \Phi_s = sG_s$. The hypothesis is the same with Φ_r in place of F_r. We form $f{\Phi_r}^{-1} = f\Phi_s$, $s = 1/r$, which is a hyperbolically expanding function. Consider also $\Phi_2(x) = 2x/(1+x^2)$. One easily checks that ${\Phi_2}''/{\Phi_2}'$ is decresing on I, and because ${\Phi_2}''(0) = 0$ we have

$$\frac{{\Phi_2}''}{{\Phi_2}'}(y) \le 0 \le \frac{{\Phi_2}''}{{\Phi_2}'}(x) \tag{4.18}$$

when $x \le 0 \le y$. The function $\tilde{f} = f\Phi_{2s}$ is more expanding than Φ_s, and we want to show that the property (4.18) of ${\Phi_2}''/{\Phi_2}'$ is also shared by $\tilde{f}''/\tilde{f}'$, wherever it exists, at least near the origin.

For this, write $\Phi_2 = h\tilde{f}$, where $h = {\Phi_s}^{-1}f^{-1} = \Phi_r f^{-1}$. Then h is contracting and

$$\log {\Phi_2}' = (\log h') \circ \tilde{f} + \log \tilde{f}'. \tag{4.19}$$

Since the extremals have zero second derivative at the origin, it follows from (4.19) that the left hand derivative of $\log h'$ at $x = 0$ is 0, and because h is contracting, we conclude again from Lemma 3 that, wherever it exists,

$$\frac{h''}{h'}(x) \le 0 \le \frac{h''}{h'}(y), \tag{4.20}$$

for $x \leq 0 \leq y$ near zero, opposite to (4.18). But (4.18), (4.19) and (4.20), along with the fact that $\tilde{f}$ is increasing and $\tilde{f}(0) = 0$, are exactly what we need to conclude that, whenever it exists,

$$\frac{\tilde{f}''}{\tilde{f}'}(y) \leq 0 \leq \frac{\tilde{f}''}{\tilde{f}'}(x), \tag{4.21}$$

when $x \leq 0 \leq y$ are near zero.

We get bounds for the change in f''/f' on either side of zero directly from this. First,

$$\frac{\tilde{f}''}{\tilde{f}'} = \left(\frac{f''}{f'} \circ \Phi_{2s}\right) \Phi'_{2s} + \frac{\Phi''_{2s}}{\Phi'_{2s}}$$

wherever f''/f' exists. Thus (4.21) implies

$$0 \leq \frac{f''}{f'}(\Phi_{2s}(y))\Phi'_{2s}(y) - \frac{f''}{f'}(\Phi_{2s}(x))\Phi'_{2s}(x) \leq \frac{\Phi''_{2s}}{\Phi'_{2s}}(x) - \frac{\Phi''_{2s}}{\Phi'_{2s}}(y), \tag{4.22}$$

for $x \leq 0 \leq y$ near zero, wherever f''/f' exists. The intermediate map Φ_{2s} is smooth, so (4.22) shows that in fact no jump can occur in f''/f' at 0. Thus f''/f' exists and is continuous at 0. Furthermore, (4.22) also shows that the difference quotient of f''/f'' at 0 is bounded by $(\Phi''_{2s}/\Phi'_{2s})'(0)$.

Finally, an increasing function with bounded difference quotient at each point of an interval must be Lipschitz on compact subsets. Hence $f \in C^{2,1}_{\text{loc}}(I)$. We now know that $Sf(x)$ exists a.e. in I. The fact that $Sf \geq 0$ again follows from (4.4) using the fact that f is contracting on every subinterval. This completes the proof.

Corollary 3 *Let $f : I \to \mathbf{R}$ be a non-constant, increasing function. Suppose there is a number $1 \leq s < \infty$ such that for every open subinterval $J \subseteq I$*

$$d_J(x,y) \leq d_{f(J)}(f(x), f(y)) \leq d_J(G_s(x), G_s(y)) \tag{4.23}$$

for all $x, y \in J$. Then $f \in C^{2,1}_{\text{loc}}(I)$ and $Sf \leq 0$.

The smoothness assertion follows from the preceding Lemma by considering $f{G_s}^{-1}$; if we rescale, then ${G_s}^{-1}$ is an F_r. The Schwarzian is negative this time because f is expanding on each subinterval.

Incidentally, the function $\Phi_2(x) = 2x/(1+x^2)$ in the proof of Lemma 4 is (aside from the factor 2) the Koebe function $x/(1-x)^2$ normalized to have second derivative zero at the origin. Other extremals with negative Schwarzian, $s > 1$, would work to get the property (4.18) near the origin, which was crucial to getting the argument started. For using

$$\left(\frac{{\Phi_s}''}{{\Phi_s}'}\right)' = S\Phi_s + \frac{1}{2}\left(\frac{{\Phi_s}''}{{\Phi_s}'}\right)^2,$$

it follows from the normalization ${\Phi_s}''(0) = 0$ that

$$\left(\frac{{\Phi_s}''}{{\Phi_s}'}\right)'(0) = S\Phi_s(0) = -2(s^2-1) < 0.$$

We used the Koebe function mostly for sentimental reasons. Also, the proof actually gives a more general result, namely that we can change the qualifiers and allow r and s to depend on the subinterval J. We have not been able to make any particular use of the stronger versions.

As a consequence of the preceding work we can now prove:

Theorem 4 *Let $f : I \to \mathbf{R}$ be a non-constant, increasing function. Suppose there are numbers $0 < r \leq 1$ and $1 \leq s < \infty$ such that for every open subinterval $J \subseteq I$*

$$d_{F_r(J)}(F_r(x), F_r(y)) \leq d_{f(J)}(f(x), f(y)) \leq d_{G_s(J)}(G_s(x), G_s(y)) \tag{4.24}$$

for all $x, y \in J$. Then $f \in C^{2,1}_{\mathrm{loc}}(I)$ and

$$SG_s \leq Sf \leq SF_r. \tag{4.25}$$

Conversely, if $f \in C^{2,1}_{\mathrm{loc}}(I)$ and (4.25) holds then so does (4.24).

Note that we are *not* assuming that f is normalized. We can also add that if equality holds in (4.24), in either inequality, for a single pair of points x, y on a given interval J, then f is the corresponding extremal on J up to a Möbius transformation of J. For if, say,

$$d_{F_r(J)}(F_r(x), F_r(y)) = d_{f(J)}(f(x), f(y)),$$

then we find that

$$d_{F_r(J)}(F_r(x), F_r(z)) = d_{f(J)}(f(x), f(z))$$

for all $x \leq z \leq y$. This implies that equality holds at x at the infinitessimal level, and hence $f = F_r$ on J up to a Möbius transformation by Corollary 1.

Proof. Again it is more convenient to work with the rescaled extremals, so we suppose first that f satisfies (4.24) with Φ_r, Φ_s in place of F_r, G_s. Then using (4.17) the map $g = f{\Phi_s}^{-1} = f\Phi_{1/s}$ satisfies

$$d_{\Phi_{r/s}(J)}(\Phi_{r/s}(x), \Phi_{r/s}(y)) \leq d_{g(J)}(g(x), g(y)) \leq d_J(x, y).$$

By Lemma 5 the function $g \in C^{2,1}_{\mathrm{loc}}(I)$ with $Sg \geq 0$. Hence $f \in C^{2,1}_{\mathrm{loc}}(I)$ as well, and by the chain rule for the Schwarzian, (1.3), $Sf \geq S\Phi_s = SG_s$. Similarly, by forming $h = f\Phi_r^{-1}$ and applying Corollary 3 we get that $Sf \leq SF_r$.

The converse, in infinitessimal form, has already appeared as Corollary 1. This completes the proof.

Combining Theorems 2 and 4 we see that a normalized function satisfying the inequalities (4.24) is quasisymmetric with constant provided by Theorem 2. This seems to be difficult to show, with any constant, without going through the Schwarzian.

One would expect a compactness result to go along with the regularity theorem above. Let $\mathcal{S}(r,s)$ be the set of increasing, $C^{2,1}_{\mathrm{loc}}$ functions f on I with $SG_s \leq Sf \leq SF_r$ and let $\mathcal{SN}(r,s)$ be the subset of $\mathcal{S}(r,s)$ of normalized functions. The topology we use on $C^{2,1}_{\mathrm{loc}}(I)$, and on $\mathcal{S}(r,s)$, $\mathcal{SN}(r,s)$ is the metric space topology on $C^2(I)$ and the weak* topology on

$L^{\infty}_{\text{loc}}(I)$. We recall that the former comes from the family of seminorms for an exhaustion of I by compact sets (the sup norms of f, f', f'' on compact sets), and the latter (measuring the local Lipschitz constants by the L^{∞} norm of f''' on compact sets) amounts to saying that $g_n \to g$, weak*, if

$$\int_I g_n \phi \to \int_I g\phi$$

for all $\phi \in L^1(I)$ of compact support. See, e.g. [13].

Theorem 5 *Let $0 < r \leq 1$, $1 \leq s < \infty$. $\mathcal{SN}(r,s)$ is compact in $C^{2,1}_{\text{loc}}(I)$. Let $\{f_n\}$ be a sequence $\mathcal{S}(r,s)$ and suppose that the sequence $\{f_n(0)\}$ converges and that for some point $p \neq 0$ that the sequence $\{f_n(p)\}$ is bounded. Then a subsequence of $\{f_n\}$ converges in the $C^{2,1}_{\text{loc}}$ topology to a function f which is either constant or an element of $\mathcal{S}(r,s)$. Under the stronger assumption that $\{f_n\}$ converges locally uniformly on I to a function f we have that either f is constant or that $f \in \mathcal{S}(r,s)$ and the full sequence $\{f_n\}$ converges in $C^{2,1}_{\text{loc}}$ to f.*

Proof. To save notation, anytime we pass from a sequence to a subsequence, which we shall have to do several times, we will use the same indices for each. Let $\{f_n\}$ be a sequence in $\mathcal{SN}(r,s)$. From the inequalities (2.4)–(2.6) derived from the comparison theorem, we get uniform bounds for $|f_n|$ and uniform bounds above and below for ${f_n}'$ on any compact set. Hence from the Arzela-Ascoli theorem there is a subsequence and a function $f \in C^0(I)$ with $f_n \to f$ locally uniformly on I. Now note that with the bounds for ${f_n}'$ we see as in the proof of Lemma 1, (3.8), (3.9), (3.12), (3.13), (3.14), that we also get local uniform bounds for $|{f_n}''/{f_n}'|$ and therefore for $|{f_n}''|$. Hence for another subsequence ${f_n}' \to f'$. In particular $f(0) = 0$ and $f'(0) = 1$. The limit function f' is subject to the same locally uniform, upper and lower bounds as the ${f_n}'$, and hence is a non-constant, increasing function on I. From Theorem 4, the functions f_n all satisfy (4.24) and therefore so too does the limit function f. The same Theorem then implies that $f \in C^{2,1}_{\text{loc}}(I)$ with $SG_s \leq Sf \leq SF_r$, so already we know that $f \in \mathcal{S}(r,s)$.

We want to get convergence in the full topology and the normalization for the second derivative. For this, with $SG_s \leq Sf_n \leq SF_r$ and the above, we also have local uniform bounds for the L^{∞} norm of ${f_n}'''$. Passing to another subsequence we then get ${f_n}'' \to f''$ locally uniformly, and pruning further still we obtain a weak* convergent subsequence of the third derivatives with ${f_n}''' \to f'''$, weak*, by the Banach-Alaoglu theorem applied on the compact sets in an exhaustion of I. Finally, $f''(0) = 0$ from the convergence of the second derivatives. Hence $\mathcal{SN}(r,s)$ is compact in the $C^{2,1}_{\text{loc}}$ topology. (With a little more involved argument it is actually possible to circumvent the use of Theorem 4 in the proof of compactness, but it is not as natural.)

For the second part of the Theorem, let $\{f_n\}$ be a sequence in $\mathcal{S}(r,s)$ and assume that $f_n(0)$ converges and that $\{f_n(p)\}$ is bounded. By working with $f_n - f_n(0)$ we may also assume that all $f_n(0) = 0$. Let

$${f_n}'(0) = b_n > 0, \quad \frac{1}{2}\frac{{f_n}''}{{f_n}'}(0) = c_n.$$

Note that since $SG_s \leq Sf_n$, it follows from (2.13) that the c_n are all $\leq s$ in absolute value. Form the normalized sequence

$$f_n{}^\dagger = \frac{f_n}{b_n + c_n f_n}\,, \quad \left(f_n = \frac{b_n f_n{}^\dagger}{1 - c_n f_n{}^\dagger} \right), \tag{4.26}$$

in $\mathcal{SN}(r,s)$.

First, we claim that for each compact set K there is a $d > 0$ so that $1 - c_n f_n{}^\dagger(x) \geq d$ for all $x \in K$. Certainly $1 - c_n f_n{}^\dagger(x) > 0$ for all $x \in I$ because it is equal to 1 at $x = 0$ and f_n and $f_n{}^\dagger$ have the same sign. If the claim is false then there is a compact set K, a sequence of points $\{x_n\}$ in K converging to a point $x_0 \neq 0$ in K, and a subsequence (same notation) with $1 - c_n f_n{}^\dagger(x_n) \to 0$. Since the c_n are bounded we may also extract a convergent subsequence, say $c_n \to c$. Now, the $f_n{}^\dagger$ are in $\mathcal{SN}(r,s)$ which is compact, so with one more subsequence we can obtain $1 - c_n f_n{}^\dagger \to 1 - cg$, in $C^{2,1}_{\text{loc}}$, for a $g \in \mathcal{SN}(r,s)$, with $1 - c_n f_n{}^\dagger(x_n) \to 1 - cg(x_0) = 0$. So $c \neq 0$, and since g is increasing we can take an x to the left or right of x_0, depending on the sign of c, to get $1 - cg(x) < -\epsilon < 0$. But then eventually $1 - c_n f_n{}^\dagger(x) < 0$ and this is a contradiction.

Next, suppose that some subsequence $b_n \to \infty$. Working at the point p, for a suitable convergent subsequence of the $f_n{}^\dagger$ and the c_n we would have from (4.26) that $f_n{}^\dagger(p) \to g(p) = 0$, since $f_n(p)$ is bounded. But $p \neq 0$ by assumption and $g \in \mathcal{SN}(r,s)$ vanishes only at zero. This contradiction shows that the sequence of derivatives $b_n = f_n{}'(0)$ must be bounded.

If a subsequence $b_n \to 0$ then, since $1 - c_n f_n{}^\dagger$ is bounded below away from zero on any compact set, it follows that f_n tends locally uniformly to the constant 0, and in fact in $C^{2,1}_{\text{loc}}$ since the $f_n{}^\dagger$ together with their first, second and third (a.e.) derivatives are locally uniformly bounded.

Now, suppose a subsequence of the b_n has a non-zero limit b. Again we may assume that a subsequence of the c_n converges to c, and a further subsequence of the $f_n{}^\dagger$ converges in $C^{2,1}_{\text{loc}}$, to conclude again from (4.26) that a subsequence of the f_n converge in $C^{2,1}_{\text{loc}}$ to a function f. This time f is in $\mathcal{S}(r,s)$ with $f'(0) = b$ and $f''(0) = 2bc$. This settles the first claims of the 'near compactness' of $\mathcal{S}(r,s)$.

Finally, consider the stronger assumption that $\{f_n\}$ converges to a function f locally uniformly on I. We can again assume that all $f_n(0) = 0$ and follow the preceding argument through. But now we can also deduce, first of all, that the full sequence $\{b_n\}$ must have a limit. For different accumulation points lead to mutually exclusive conclusions about the limit function f; either that $f = 0$ or that $f'(0) = b > 0$. When the limit of the b_n is $b > 0$ then any accumulation point c of the c_n gives $f''(0) = 2bc$, so again the c_n must also have a limit.

We now claim that the full sequence $\{f_n\}$ must converge to f in $C^{2,1}_{\text{loc}}$. This is clear if $b_n \to 0$, in which case $f = 0$, for the same reasons as above. If $b_n \to b > 0$ then $c_n \to c$ and if g is any accumulation point of $\{f_n{}^\dagger\}$ in $C^{2,1}_{\text{loc}}$ then

$$f = \frac{bg}{1 - cg}\,,$$

from (4.26). In other words, g is unique, and $g = f^\dagger$. It follows that the full sequence of the ${f_n}^\dagger$ must be converging in $C^{2,1}_{\mathrm{loc}}$ and the same is then true of the f_n. This completes the proof.

Remark 5 One can easily state 'conformally invariant' versions of these results. One case we shall need in the next Section is when f is defined on $I = (-R, R)$. Then $\lambda_I(x) = R/(R^2 - x^2)$ and, referring to Remark 1 in Section 2, Part (a) of Lemma 3 reads

$$\begin{aligned} &S\tilde{G}_s \leq Sf \leq S\tilde{F}_r \ \text{ implies} \\ &\qquad r\lambda_I(x) \leq \lambda_{f(I)}(f(x))f'(x) \leq s\lambda_I(x). \end{aligned} \tag{4.27}$$

We could also translate the center of the interval, and with corresponding translations of $\tilde{F}_r$ and $\tilde{G}_s$ the statement would look the same.

Remark 6 The inequality (4.12) in the proof of Lemma 4 expresses a distortion of the cross-ratio

$$(x_1, x_2, x_3, x_4) = \frac{x_1 - x_3}{x_1 - x_4} \frac{x_2 - x_4}{x_2 - x_3} \,,$$

namely,

$$(f(x), f(y), f(b), f(a)) \leq (x, y, b, a),$$

for $a < x < y < b$. That Lemma dealt with contracting maps (and so with positive Schwarzian). Other authors have obtained and used similar distortions of the cross-ratio when $Sf < 0$, notably Singer in [11], and De Melo and van Strien in [4]. See also the papers of Sullivan [12] and Guckenheimer [8]. The most general form of the relationship (4.4) between the Schwarzian and the distortion of cross-ratio is in Ahlfors [1].

Remark 7 These comments are an addendum to Lemma 4 and Corollary 2 on the C^1 smoothness of contracting functions. Recall that such a function f has a log convex derivative, so that f''/f' is an increasing function with at most a countable number of jump discontinuities. It is certainly possible for jumps to occur, so whatever extra smoothness might still follow from the hypotheses one cannot get up to C^2. For an example of this we piece together the Möbius transformations $f(x) = x$ for $-1 < x \leq 0$ and $f(x) = x/(1-x)$ for $0 \leq x < 1$. Then $f''(x)$ jumps by 2 at $x = 0$. The function is a hyperbolic isometry on subintervals of $(-1, 1)$ not containing the origin, and it is easy to check that it decreases hyperbolic distances on all subintervals containing the origin.

It is likely that this sort of construction can be extended to get more jump discontinuities, but we would like to be able to say more about the properties of contracting functions away from the jumps. First note that f''/f' will be absolutely continuous if and only if it has no jumps and equality holds in (4.14). So, as we remarked at the beginning of Section 2, one should perhaps not refer to Sf as a 'weak Schwarzian' without these latter conditions also holding. But might it be that a contracting function has f''/f' absolutely continuous on the complement of a discrete set of points where it does have positive jumps? We do not know, but in trying to understand this question we were led to construct 'virtual Möbius transformations'. These are C^2 functions f with a third derivative a.e., with $Sf = 0$ wherever it exists, but with f''/f' not absolutely continuous. Briefly, the construction goes as follows.

Let q be a continuous, increasing function on $(-1,1)$ with supremum $\leq 1/2$ and with $q' = 0$ a.e.. It is easy to see that the operator

$$(Tg)(x) = q(x) + \frac{1}{2}\int_0^x g(t)^2\,dt$$

maps the closed unit ball in $C^0(-1,1)$ into itself, and if we restrict functions $g \in C^0(-1,1)$ to $[-c,c]$, $0 < c < 1$ it is contracting. We thus get a fixed point for each such compact subset. The functions agree on their common domains by uniqueness, and thus define a continuous function h on $(-1,1)$ such that

$$h(x) = q(x) + \frac{1}{2}\int_0^x h(t)^2\,dt\,.$$

Then h is a continuous, increasing function with

$$h' = \frac{1}{2}h^2$$

almost everywhere. Now let f be a solution to $f''/f' = h$. Then f is C^2, f''/f' is increasing but not absolutely continuous, and $Sf = 0$ a.e.. This completes the construction. Such a function cannot be a hyperbolic isometry unless it is an honest Möbius transformation (this follows from invariance of cross-ratio), but we do not know whether it can be contracting in the hyperbolic metric.

5 Factoring Quasisymmetric Maps via the Schwarzian

In this Section we want to show how one may apply the results of the previous two Sections to prove that a function satisfying the usual upper and lower bounds on its Schwarzian can be factored as a compositon of maps whose quasisymmetry quotients are arbitrarily close to one. Compare the statements to this end in [9] on pages 36 and 89 for quasisymmetric mappings of the line. That theorem is based on quasiconformal extensions and a decomposition theorem for quasiconformal maps.

We state the result for normalized functions on $(-1,1)$.

Theorem 6 *Let f be a normalized $C^{2,1}_{\mathrm{loc}}$ function with $SG_s \leq Sf \leq SF_r$ on $(-1,1)$. Given any $\epsilon > 0$ there exists a number N depending on ϵ, r and s, Möbius transformations $T_1, \ldots, T_{N-1}$, and $C^{2,1}_{\mathrm{loc}}$ functions $h_1, \ldots, h_N$ such that*
(i) $f = T_1 \cdots T_{N-1} h_N \cdots h_1$.
(ii) All maps in the composition have quasisymmetry quotients bounded between $1-\epsilon$ and $1+\epsilon$ on their domains.
One can take $N = O(\frac{1}{\epsilon}\log\frac{1}{\epsilon})$, where the implied constant depends on r and s.

Proof. The proof is an iterative construction. We describe the plan in general terms first. Let $I_1 = (-1,1)$ and write f_1 for f. By solving a differential equation and appealing to the chain rule for the Schwarzian, we would like to produce a map h_1 defined on I_1, with quasisymmetry quotient close to 1, that will take a fraction of Sf_1 away from f_1. That is, for $f_2 = f_1 {h_1}^{-1}$ the bounds for Sf_2 on $I_2 = h_1(I_1)$ will have improved over those for Sf_1 on I_1, meaning that both the upper and lower bounds will have moved closer to zero. We will have written $f = f_1 = f_2 h_1$ and we can try to repeat the procedure with f_2 in place of f_1, and so on. We cannot do this quite so simply. For f_2 to replace $f_1 = f$ in the argument it must be a normalized function defined on a centered interval (meaning, centered at the origin). We can and will make $I_2 = h_1(I_1)$ centered, but h_1 and f_2 will not be normalized. We can still bound kh_1, but we pay the price of keeping track of a term qh_1, via Lemma 2 in Section 3. Next, it is not $f_2 = f_1 {h_1}^{-1}$ that should replace f_1 in order to iterate the construction, but rather it is the *normalized* function ${f_2}^\dagger$ that we need. The extra Möbius transformation required to renormalize is the source of the $T's$ in the statement of the Theorem; $f = f_1 = f_2 h_1 = T_1 {f_2}^\dagger h_1$. Again, we pay the price of keeping track of a qT_1 along with $k{f_2}^\dagger$. The choice of N depends on several conditions which will come up in the course of the proof. We proceed to the details.

To begin with, for the estimates we have to make it is convenient to write

$$\rho_1 = 1 - r^2, \quad \sigma_1 = s^2 - 1.$$

Thus the main hypothesis is

$$-2\sigma_1 \lambda_{I_1}^2 \leq Sf_1 \leq 2\rho_1 \lambda_{I_1}^2.$$

Let g_1 be the normalized solution to $Sg_1 = (1/n)Sf_1$, where $n > 1$ is a positive number depending on ϵ, r and s to be chosen later. Let $a = g_1(-1) < 0 < g_1(1) = b$. In general, (a, b) will not be centered. If we define h_1 by

$$g_1 = {h_1}^\dagger = \frac{h_1}{1 + a_2(h_1)h_1}, \text{ or } h_1 = \frac{g_1}{1 - a_2(h_1)g_1}, \text{ where } a_2(h_1) = \frac{1}{2}\left(\frac{1}{b} + \frac{1}{a}\right),$$

then $I_2 = h_1(-1, 1) = h_1(I_1)$ will be a bounded, centered interval *provided* that h_1 is regular. To address this we estimate $a_2(h_1)$. Using (2.7), (2.8) we have,

$$\begin{aligned} \frac{1}{\sqrt{1 + \sigma_1/n}} &\leq \; b \; \leq \frac{1}{\sqrt{1 - \rho_1/n}} \\ \frac{-1}{\sqrt{1 - \rho_1/n}} &\leq \; a \; \leq \frac{-1}{\sqrt{1 + \sigma_1/n}}. \end{aligned}$$

It follows that

$$|a_2(h_1)| \leq \frac{1}{2}(\sqrt{1 + \sigma_1/n} - \sqrt{1 - \rho_1/n}\,). \tag{5.1}$$

Next, because the map g_1 satisfies

$$Sg_1(x) \leq \frac{2\rho_1}{n} \frac{1}{(1 - x^2)^2},$$

by (2.10) it will not attain the value $-1/a_2(h_1)$, and hence h_1 will be regular, if

$$|a_2(h_1)| < \sqrt{1 - \rho_1/n}, \text{ i.e., if } \sqrt{1 + \sigma_1/n} < 3\sqrt{1 - \rho_1/n}. \tag{5.2}$$

With a given ρ_1 and σ_1 this last inequality holds if $n > (\sigma_1 + 9\rho_1)/8$. For this (preliminary) choice of n we can conclude that $h_1(I_1)$ is the centered interval $I_2 = (-R_2, R_2)$, where $R_2 = 2ab/(a-b)$. Note that R_2 satisfies

$$\sqrt{1-\rho_1/n} \leq \frac{1}{R_2} \leq \sqrt{1+\sigma_1/n}\,. \tag{5.3}$$

Next, we estimate the quasisymmetry quotient $kh_1 = (qh_1)^{-1}kg_1$ by appealing to Theorem 2 for kg_1 and to Lemma 2 for qh_1. For the latter, it is more convenient to use the second set of inequlities (3.25) since we have already chosen n in (5.2) so that the hypothesis '$|a_2| < r$' holds. This gives

$$\frac{3\sqrt{1-\rho_1/n} - \sqrt{1+\sigma_1/n}}{\sqrt{1+\sigma_1/n} + \sqrt{1-\rho_1/n}} \leq qh_1 \leq \frac{\sqrt{1+\sigma_1/n} + \sqrt{1-\rho_1/n}}{3\sqrt{1-\rho_1/n} - \sqrt{1+\sigma_1/n}}\,. \tag{5.4}$$

In terms of ρ_1 and σ_1 the estimate for kg_1 from Theorem 2 is complicated to write down. Recall, however, that it does tend to 1 as, in this case, $\sqrt{1-\rho_1/n}$ and $\sqrt{1+\sigma_1/n}$ tend to 1, that is, as $n \to \infty$. From this and from (5.4) it is clear that we can make make kh_1 lie between $1 \pm \epsilon$ for n sufficiently large, and from the explicit bounds it is not too hard to show that n should be of the order

$$n = O\Big(\frac{\rho_1 + \sigma_1}{\epsilon}\Big)\,. \tag{5.5}$$

We now examine $f_2 = f_1 {h_1}^{-1}$ on $I_2 = h_1(I_1)$. From the chain rule for the Schwarzian (1.3) and the fact that $Sh_1 = Sg_1 = (1/n)Sf_1$ we compute that

$$Sf_2(y) = \frac{n-1}{n} Sf_1(x) \frac{1}{{h_1}'(x)^2}\,,\ \ y = h_1(x). \tag{5.6}$$

Now, from Part (a) of Lemma 3, (4.5), in the last Section, we have

$$(\sqrt{1-\rho_1/n}\,)\lambda_{I_1}(x) \leq \lambda_{I_2}(h_1(x)){h_1}'(x) \leq (\sqrt{1+\sigma_1/n}\,)\lambda_{I_1}(x)\,.$$

Using this in (5.6) leads to

$$-2\sigma_2\lambda_{I_2}^2 \leq Sf_2 \leq 2\rho_2\lambda_{I_2}^2\,, \tag{5.7}$$

where

$$\rho_2 = \frac{n-1}{n-\rho_1}\rho_1\,, \quad \sigma_2 = \frac{n-1}{n+\sigma_1}\sigma_1\,. \tag{5.8}$$

The bounds on the Schwarzian have improved because

$$\rho_2 < \rho_1 \text{ and } \sigma_2 < \sigma_1. \tag{5.9}$$

Now, f_2 is not normalized, but rather $f_2(0) = 0, {f_2}'(0) = 1$ and ${f_2}''(0) = -{h_1}''(0) = -2a_2(h_1)$. Hence the normalized function is

$${f_2}^{\dagger} = \frac{f_2}{1 - a_2(h_1)f_2} = {T_1}^{-1}f_2\,,$$

where T_1 is the Möbius transformation

$$T_1(x) = \frac{x}{1 + a_2(h_1)x} \,.$$

Since the bounds for Sf_2 have improved, f_2 does not assume the value $1/a_2(h_1)$ and ${f_2}^\dagger$ is therefore regular on I_2. It also satisfies (5.7) because the Schwarzians are the same.

We have now written

$$f = T_1 {f_2}^\dagger h_1$$

where ${f_2}^\dagger$ is a normalized function on the centered interval $I_2 = h_1(I_1) = (-R_2, R_2)$ whose Schwarzian has the bounds given for Sf_2 in (5.7), and where kh_1 is between $1 \pm \epsilon$ on I_1. To complete this step of the construction we have to estimate the quasisymmetry quotient kT_1 on ${f_2}^\dagger(I_2)$. For this we observe that the identity map is a normalization of T_1, that is id$= {T_1}^\dagger = T_1/(1 - a_2(h_1)T_1)$, and hence $1 = (qT_1)(kT_1)$ from (3.23). We cannot estimate qT_1 using Lemma 2 because ${f_2}^\dagger(I_2)$ is not necessarily centered. However, we can work directly with

$$qT_1(x, h) = \frac{1 + a_2(h_1)(x - h)}{1 + a_2(h_1)(x + h)} \,. \tag{5.10}$$

The length of ${f_2}^\dagger(I_2)$ is at most $2R_2/\sqrt{1 - \rho_2}$ from (5.7) and (2.15), (2.16). Also, because $h > 0$ and $x - h, x + h$ lie in the interval we see that $x \pm h$ can contribute up to half this, or $\pm R_2/\sqrt{1 - \rho_2}$, to the numerator and denominator. If n is large, R_2 is close to 1, from (5.3), and $\rho_2 < \rho_1$ from (5.8) (for any n). Finally, $a_2(h_1)$ tends to 0, from (5.1). It follows that for n sufficiently large kT_1 lies between $1 \pm \epsilon$. In fact, one can show that n should again be of the size in (5.5). We make a choice of n, of this order, so all the requirements above are satisfied.

(These last estimates are exactly where we have used the hypothesis that $f = f_1$ is normalized. If not, but still with $f_1(0) = 0, {f_1}'(0) = 1$, then we would have ${f_2}''(0) = {f_1}''(0) - {h_1}''(0)$, so the estimates for kT_1 would depend on $a_2(f_1)$ as well. This is not a major complication, but in presenting the proof we felt it was easiest to deal with it separately, after the normalized case was settled.)

We now iterate this construction. There is one thing that changes from the first to the second step, but not after that. We must apply the versions of the earlier inequalities *et al* which are for a general centered interval $(-R, R)$, in the first instance for the interval $(-R_2, R_2)$. However, the results as formulated in Remark 1 in Section 2, Remark 2 in Section 3, and Remark 5 in Section 4 are such that this requires no essential modification. Most helpfully, one sees that the choice of n in the first step works also in the second step and then in all subsequent steps. This turns out to be so by virtue of the way the bounds improve, as in (5.7), (5.8), and (5.9).

After j steps we will have written

$$f = T_1 \cdots T_j {f^\dagger}_{j+1} h_j \cdots h_1 \,. \tag{5.11}$$

The $h's$ and $T's$ have quasisymmetry quotients bounded by $1 \pm \epsilon$, and ${f^\dagger}_{j+1}$ is a normalized function on the centered interval $I_{j+1} = h_j(I_j)$ with

$$-2\sigma_{j+1}\lambda^2_{I_{j+1}} \le S{f^\dagger}_{j+1} \le 2\rho_{j+1}\lambda^2_{I_{j+1}} \,, \tag{5.12}$$

where

$$\rho_{j+1} = \frac{n-1}{n-\rho_j}\rho_j \quad \sigma_{j+1} = \frac{n-1}{n+\sigma_j}\sigma_j\,. \tag{5.13}$$

The choice of n is fixed in the first step and is of the order (5.5).

Starting with ρ_1, σ_1, and n, the maps

$$\rho \mapsto \frac{n-1}{n-\rho}\rho\,, \quad \sigma \mapsto \frac{n-1}{n+\sigma}\sigma \tag{5.14}$$

iterate to zero. Hence after finitely many steps, say $j+1 = N$ in (5.11), the Schwarzian $S{f_N}^\dagger$ will be so small that ${f_N}^\dagger$ will have quasisymmetry quotient bounded by $1 \pm \epsilon$. We can get rough bounds for N by approximating the maps in (5.14) by linear ones. Using that $n = O((\rho_1 + \sigma_1)/\epsilon)$ one can show that N grows like

$$N = O\left(\frac{1}{\epsilon}\log\frac{1}{\epsilon}\right),$$

in terms of ϵ. We put $h_N = {f_N}^\dagger$ and stop at this point. This completes the proof of the Theorem. Naturally, one can formulate the Theorem for a function on an arbitrary bounded interval.

Recall from Section 3 that a non-normalized function f satisfying the usual bounds on its Schwarzian will not necessarily be quasisymmetric unless we make the additional assumption that either it is bounded or that $a_2 = (1/2)f''(0)$ is small, the latter being a stronger condition. Under either assumption, say that $|f(x)| \leq c$, it is straightforward to give a corresponding factorization into functions with small quasisymmetry constants. First, normalize f as always, writing $f = Vf^\dagger$, $Vx = x/(1-a_2x)$, and factor $f^\dagger$ according to the Theorem. Because kV might be large we break V up into $Vx = \tilde{V}^m x$ with

$$\tilde{V}x = \frac{x}{1-(a_2/m)x},$$

where we have to choose m to make $k\tilde{V}$ small. Let $J = f^\dagger(-1,1)$. Then we have to estimate $k\tilde{V}$, equivalently $q\tilde{V}$, on J and its sucessive images under iterating $\tilde{V}$. As we saw in (5.10) above, we need to know about the lengths of these intervals. But we can get uniform estimates for $q\tilde{V}$ for each of the $\tilde{V}$ factors in V exactly because we know that $f = \tilde{V}^m f^\dagger = Vf^\dagger$ is bounded. We will not go through the calculations, but m can be chosen of the order $O(1/\epsilon)$, where the constant depends on $|a_2|$ and c, to make each $k\tilde{V}$ lie between $1 \pm \epsilon$.

Finally, one can combine combine Theorem 6 in this Section with the results in the last Section on expanding and contracting functions to get a nice geometric picture. Suppose f is a normalized $C^{2,1}_{\text{loc}}$ function satisfying $SG_s \leq Sf \leq SF_r$ on $I = (-1,1)$. Using Theorem 1 we can find a normalized, $C^{2,1}_{\text{loc}}$ solution ϕ to the equation $S\phi = \max\{Sf, 0\}$. Then $S\phi \geq 0$ so ϕ will be contracting, while $\psi = f\phi^{-1}$, having negative Schwarzian, will be expanding. This factors $f = \psi\phi$ as a composition of an expanding and contracting function, and we would now like to apply Theorem 6 to factor ϕ and ψ further into maps of small hyperbolic distortion (and small quasisymmetry quotient). For ϕ we have the upper and lower bounds $0 \leq S\phi \leq SF_r$ so Theorem 6 applies directly. This gives $\phi = A\phi_1 \ldots \phi_N$, where A is a Möbius transformation, which is a hyperbolic isometry, and all the factors ϕ_j are contracting because they all have positive Schwarzians. Furthermore, the Schwarzians are shrinking, so it is clear we can make each ϕ_j as close to a hyperbolic isometry on its domain as we please, possibly by changing the N. Next, by Lemma 3, Part (a) appplied to ϕ, we have

$$r\lambda_I(x) \le \lambda_{\phi(I)}(\phi(x))\phi'(x),$$

so using the chain rule (1.3) for the Schwarzian we find that

$$\frac{-2(s^2-1)}{r^2}\lambda^2_{\phi(I)} \le Sh \le 0.$$

The lower bound for for Sh on $\phi(I)$ is worse by the factor $1/r^2$ than the original lower bound for Sf on I, but it is still the Schwarzian of an extremal on $\phi(I)$. We can then invoke Theorem 6 to factor ψ as $\psi = B\psi_1 \ldots \psi_{N'}$, where B is Möbius and the ψ_j are expanding, but nearly isometries.

6 Constructions

In this Section we prove two theorems which give examples showing some limitations to what one might hope to be true for the relations between the Schwarzian and quasisymmetry. For instance, though a small Schwarzian implies a small quasisymmetry constant, the converse does not hold.

Theorem 7 *There is a smooth, bi-Lipschitz function f on $(-1,1)$ with $Sf \ge 0$ and $\sup(1-x^2)^2 Sf(x) = \infty$.*

Proof. Once again we consider the initial value problem

$$u'' + pu = 0,\ u(0) = 1,\ u'(0) = 0 \tag{6.1}$$

on $(-1,1)$. We will construct a smooth, non-negative function p so that

$$\frac{1}{2} \le u \le 1, \quad \text{and} \tag{6.2}$$

$$\sup(1-x^2)^2 p(x) = \infty. \tag{6.3}$$

Then

$$f(x) = \int_0^x u^{-2}(t)\,dt$$

satisfies $1 \le f' \le 4$, so it is bi-Lipschitz, and $\sup(1-x^2)^2 Sf(x) = \sup(1-x^2)^2 p(x) = \infty$.

First, let p be identically zero on $(-1,0]$. Let (a_n, b_n) be a sequence of disjoint intervals in $(0,1)$ with $a_n < b_n < a_{n+1} < \ldots$ and $b_n \to 1$. Let φ_n be a non-negative, smooth cut-off function with maximum 1 and with compact support in (a_n, b_n). On each interval (a_n, b_n) we set

$$p = \frac{n\varphi_n(x)}{(1-x^2)^2}$$

and let p be zero elsewhere on $(0,1)$. The initial value problem (6.1) then makes sense. The condition (6.3) is satisfied, as is $u \le 1$ because $p \ge 0$ and so u is concave down on $[0,1)$.

We want to see that the intervals (a_n, b_n) can be chosen inductively so that for each n

$$u(b_n) > \frac{1}{2} + \frac{1}{4^n}, \tag{6.4}$$

$$u'(b_n) > -\frac{1}{4^n}\frac{1}{1-b_n}. \tag{6.5}$$

This means the following. Notice that u is affine in between the consecutive intervals (a_n, b_n). These conditions on u at the endpoints b_n provide that the prolongation of any such straight line segment in the graph of u intersects the line $x = 1$ above $1/2$. Then both inequalities in (6.2) will hold and the construction will be complete.

For this, note that

$$\begin{aligned} u'(a_n) \ge u'(b_n) &= u'(a_n) + \int_{a_n}^{b_n} u''(x)\,dx = u'(a_n) - n\int_{a_n}^{b_n} \frac{\varphi_n(x)}{(1-x^2)^2}\,dx \\ &\ge u'(a_n) - n\int_{a_n}^{b_n} \frac{dx}{(1-x^2)^2} \ge u'(a_n) - \frac{n(b_n - a_n)}{(1-a_n)(1-b_n)}\,. \end{aligned}$$

We can choose $b_n - a_n$ to tend to zero so rapidly that the last term tends to zero, and so $u'(a_n)$ $(= u'(b_{n-1}))$ and $u'(b_n)$ are then also so close that we can satisfy (6.4) and (6.5) inductively.

This completes the proof of the theorem. We remark that we chose the lower bound $u \ge 1/2$ and the other numbers only to be definite. The construction can be modified to produce an f with $1 \le f' \le 1 + \epsilon$ and $\sup(1-x^2)^2 Sf(x) = \infty$ for any $\epsilon > 0$.

Finally, experience may indicate that a negative Schwarzian is a good property for quasisymmetry, but the next result shows that one still needs a finite lower bound.

Theorem 8 *There is a smooth function f which is not quasisymmetric on $(-1,1)$, with $Sf \le 0$ and $\inf(1-x^2)^2 Sf(x) = -\infty$.*

Proof. The construction is again based on the initial value problem (6.1). Let (a_n, b_n) be a sequence of disjoint intervals in $(0,1)$ with $a_n < b_n < a_{n+1} < \ldots$ and $b_n \to 1$ and with the additional property that $\delta_n = b_n - a_n < a_n - b_{n-1}$. Again we start by setting $p = 0$ on $(-1,0]$. This time we want to inductively define the function p on $(-1,1)$ so that: (i) $p \le 0$ on $(0,1)$ and is supported in the union of the (a_n, b_n), and, (ii) if

$$f(x) = \int_0^x u^{-2}(x)\,dx\,,$$

then given $f(a_n) - f(a_n - \delta_n)$, p is defined on (a_n, b_n) in such a way that

$$k_n = \frac{f(b_n) - f(a_n)}{f(a_n) - f(a_n - \delta_n)} < \frac{1}{n}. \tag{6.6}$$

Condition (ii) makes sense inductively because u is affine off each $[a_n, b_n]$. To show that this is possible we need a lemma.

Lemma 6 *Let $x_0 \in (0,1)$, let c be a positive constant and let v be a solution of*

$$v'' - \frac{c}{(1-x^2)^2} v = 0,\ v(x_0) > 0, v'(x_0) \geq 0\,.$$

Then given $\epsilon > 0, \delta > 0$ there exists $c_0 > 0$ such that

$$\int_{x_0}^{x_0+\delta} v^{-2}(s)\, ds < \epsilon$$

for all $c \geq c_0$.

Proof of Lemma 6. It is clear that $v(x) \geq v(x_0)$ for $x \geq x_0$. Write

$$\begin{aligned} v(x) &= v(x_0) + \int_{x_0}^{x} v'(s)\, ds = v(x_0) + \int_{x_0}^{x} \left\{ v'(x_0) + \int_{x_0}^{s} v''(t)\, dt \right\} ds \\ &\geq v(x_0) + \int_{x_0}^{x}\int_{x_0}^{s} \frac{cv(t)}{(1-t^2)^2}\, dt\, ds \\ &\geq v(x_0) \left\{ 1 + c \int_{x_0}^{x}\int_{x_0}^{s} \frac{1}{(1-t^2)^2}\, dt\, ds \right\}. \end{aligned}$$

This shows that given $\mu > 0$, $v(x)$ tends uniformly to ∞ as $c \to \infty$ for $x \geq x_0 + \mu$. Hence, given $\epsilon > 0$, $\delta > 0$ choose $\mu > 0$ small enough so that

$$\int_{x_0}^{x_0+\mu} v^{-2}(s)\, ds < \epsilon/2\,,$$

and then c_0 large enough so that for $c \geq c_0$,

$$\int_{x_0+\mu}^{x_0+\delta} v^{-2}(s)\, ds < \epsilon/2\,.$$

This completes the proof of the Lemma.

Returning now to the proof of the Theorem, on the interval (a_n, b_n) we let

$$p(x) = -\frac{c_n \varphi_n(x)}{(1-x^2)^2}\,,$$

where φ_n is a smooth cut-off function on (a_n, b_n) as in the proof of the preceding Theorem. It follows from the Lemma above, more accurately its proof, that given the difference $f(a_n) - f(a_n - \delta_n)$ there is a constant $c_n > 0$ sufficiently large, and a cut-off function φ_n such that

$$f(b_n) - f(a_n) = \int_{a_n}^{b_n} u^{-2}(s)\, ds < \frac{1}{n}(f(a_n) - f(a_n - \delta_n)).$$

This construction defines the function p, hence f, on $(-1,1)$. We have $Sf \leq 0$, $\inf(1-x^2)^2 p(x) = \inf(1-x^2)^2 Sf(x) = -\infty$, and $\inf kf(x,h) = 0$.

References

[1] L. Ahlfors. Cross-ratios and Schwarzian derivatives in $\mathbf{R}^n$. In *Complex Analysis:Articles Dedicated to Albert Pfluger on the Occasion of his 80th Birthday*, pages 1–15, Boston, 1989. Birkhauser.

[2] M. Chuaqui and B. Osgood. The Schwarzian derivative and conformally natural quasiconformal extensions from one to two to three dimensions. *Math. Ann.*, 292:267–280, 1992.

[3] M. Chuaqui and B. Osgood. Sharp distortion theorems associated with the Schwarzian derivative. *Jour. London Math. Soc.(2)*, 48:289–298, 1993.

[4] W. de Melo and S. van Strien. A structure theorem in one dimensional dynamics. *Annals of Math.*, 129:519–546, 1989.

[5] W. Fenchel. Convex cones, sets, and functions. Notes by D.W. Blackett of lectures at Princeton University, 1953.

[6] F. Gardiner and D. Sullivan. Symmetric structures on a closed curve. *Amer. Jour. Math.*, 114:683–736, 1992.

[7] F.W. Gehring and Ch. Pommerenke. On the Nehari univalence criterion and quasicircles. *Comment. Math. Helv.*, 59:226–242, 1984.

[8] J. Guckenheimer. Sensitive dependence to initial conditions for one-dimensional maps. *Comm. Math. Phys.*, 70:133–160, 1979.

[9] O. Lehto. *Univalent functions and Teichmüller spaces.* Springer-Verlag, New York, 1987.

[10] W. Paluba. PhD thesis, CUNY, 1992.

[11] D. Singer. Stable orbits and bifurcations of maps of the interval. *SIAM J. Appl. Math.*, 35:260–267, 1978.

[12] D. Sullivan. Bounds, quadratic differentials, and renormalization conjectures. In *Mathematics into the Twenty-First Century, Volume 2*, Providence, RI, 1991. American Math. Soc.

[13] K. Yosida. *Functional Analysis.* Springer-Verlag, New York, 1965.