UNIVALENT FUNCTIONS WHICH MAP ONTO REGIONS
OF GIVEN TRANSFINITE DIAMETER

P. L. DUREN AND M. M. SCHIFFER

ABSTRACT. By a variational method, the sharp upper bound is obtained for the
second coefficients of normalized univalent functions which map the unit disk
onto regions of prescribed transfinite diameter, or logarithmic capacity.

0. INTRODUCTION

Let S be the usual class of functions $f(z) = z + a_2 z^2 + \cdots$ analytic and
univalent in the unit disk $\mathbb{D}$. Many years ago, Pick [7] used Bieberbach's in-
equality $|a_2| \le 2$ to obtain the sharp bound $|a_2| \le 2(1 - 1/M)$ for functions
$f \in S$ with $|f(z)| < M$ for some number $M > 1$. The extremal functions
map the unit disk onto the disk $|w| < M$ minus a radial slit of suitable length.

Our purpose is to solve a related problem: to find the maximum of $|a_2|$
among all functions $f \in S$ whose range has a given transfinite diameter R,
$1 < R < \infty$. This problem is considerably harder than Pick's, but we shall see
that for "large" R the solution is somewhat similar. The sharp inequality is

$$|a_2| \le 2 - e^2/4R$$

and the extremal function maps the disk onto a certain Jordan region minus a
radial spire. For small R the situation is rather different. The sharp bound is
then determined implicitly through elliptic integrals, and the extremal function
maps the disk onto a certain Jordan region with analytic boundary. The division
occurs at the transfinite diameter

$$R = e^2 \pi^2/64 = 1.139 \ldots .$$

This unexpected phenomenon, the dual nature of the solution, gives the problem
a particular interest.

A more detailed summary of our results appears in the theorem stated at
the end of the paper. The proof uses a special variational method devised to

preserve both the class S and the transfinite diameter of the range. It combines the standard methods of interior variation and boundary variation, and it leads ultimately to a quadratic differential for the boundary arcs in the extremal situation. The trajectory structure then allows three possible cases, each of which is eventually seen to occur for suitable values of R.

A few preliminary remarks about transfinite diameter (or logarithmic capacity) are in order. We begin by recalling the original definition. For a bounded set E in the complex plane, the nth diameter is

$$d_n(E) = \left\{ \sup_{z_1, \ldots, z_n \in E} \prod_{j < k} |z_j - z_k| \right\}^{2/n(n-1)}.$$

It can be shown that $d_{n+1}(E) \leq d_n(E)$. The transfinite diameter of E is defined by $d(E) = \lim_{n \to \infty} d_n(E)$.

The transfinite diameter of a circular disk is its radius. According to a result of Pólya ([8]; see [4, Chapter VII, §2]), the outer area of a set E is bounded above by πR^2, where $R = d(E)$. Thus for functions $f \in S$ with $f(\mathbb{D})$ of transfinite diameter R, Pólya's theorem gives the inequality

$$1 + \sum_{n=2}^{\infty} n|a_n|^2 \leq R^2.$$

It follows that $R > 1$ unless f is the identity mapping. Incidentally, Pólya's inequality also gives the crude estimate

$$|a_2| \leq 2^{-1/2}(R^2 - 1)^{1/2}.$$

It is important to recall a result of Szegö [13] which equates transfinite diameter with logarithmic capacity. Given a compact set E of positive transfinite diameter in the plane, let

$$g(w) = g(w, \infty) = \log|w| + \gamma + O(1/w)$$

be Green's function of the unbounded component of the complement of E. The number γ is known as the Robin constant of E, and $e^{-\gamma}$ is its logarithmic capacity. Szegö's theorem asserts that $R = e^{-\gamma}$, or $\gamma = -\log R$, where $R = d(E)$. Because of the conformal invariance of Green's function, this basic relation accounts for the natural occurrence of transfinite diameter in function-theoretic problems.

Szegö's theorem also allows the interpretation of the transfinite diameter as a conformal radius. Suppose for simplicity that the compact set E is simply connected. Let

$$\zeta = h(w) = w + c_0 + c_1/w + \cdots$$

map the complementary domain $\widetilde{E}$ conformally onto a circular region $|\zeta| > \rho$. Then Green's function of $\widetilde{E}$ is

$$g(w, \infty) = \log|h(w)| - \log \rho,$$

which shows that $\rho = R$. An equivalent statement is that a function of the form

$$\tilde{f}(z) = Rz + b_0 + b_1/z + \cdots$$

maps $|z| > 1$ conformally onto $\tilde{E}$.

For basic facts about transfinite diameter and related topics, the reader may consult Goluzin [4, Chapter VII] or Hille [5, Chapter 16]. Further information is in Pólya and Szegö [9, Part IV, Chapter 2]. See also Tsuji [14].

1. THE VARIATIONAL METHOD

Let S_R be the class of functions $f \in S$ whose range has transfinite diameter $R < \infty$. Although S_R is a normal family, it is not compact. Indeed, a "pinching" construction shows that a sequence of functions $f_n \in S_R$ may converge uniformly on compact sets to a function $f \in S_Q$ for some $Q < R$. We claim, however, that $\widehat{S}_R = \bigcup_{Q \le R} S_Q$ is a compact normal family. To prove this, let the functions $f_n \in \widehat{S}_R$ map $\mathbb{D}$ onto a sequence of regions Δ_n with transfinite diameters $R_n \le R$. Let $E_n = \overline{\Delta}_n$, and let

$$\tilde{f}_n(z) = R_n z + b_{0n} + b_{1n}/z + \cdots$$

map $|z| > 1$ conformally onto $\tilde{E}_n$. We may choose a subsequence $\{n_k\}$ such that $f_{n_k}(z) \to f(z)$ and $\tilde{f}_{n_k}(z) \to \tilde{f}(z)$, uniformly on compact sets. Then $f \in S$ and $\tilde{f}$ maps $|z| > 1$ onto the complement of a compact set E with transfinite diameter $Q \le R$. Since $f(\mathbb{D}) \subset E$, it follows that $f \in S_P$ for some $P \le Q$. This proves the compactness of $\widehat{S}_R$.

The compactness of $\widehat{S}_R$ ensures that $\operatorname{Re}\{a_2\}$ attains a maximum within the family. We may safely assume that the maximum occurs for a function $f \in S_R$, because the extremal value of $\operatorname{Re}\{a_2\}$ will be seen to increase monotonically with the transfinite diameter of $f(\mathbb{D})$.

The extremal functions will be found by applying a variation which preserves the class S_R. Such a variation is constructed as follows. Given $f \in S_R$, fix three distinct points w_1, w_2, $w_3 \notin \overline{f(\mathbb{D})}$ and consider the perturbation

$$(1) \qquad w^* = w + \sum_{j=1}^{3} \varepsilon_j \frac{w}{w - w_j},$$

where the ε_j are small complex parameters. If $\varepsilon = \max |\varepsilon_j|$ is sufficiently small, then clearly

$$(2) \qquad f^*(z) = f(z) + \sum_{j=1}^{3} \varepsilon_j \frac{f(z)}{f(z) - w_j} = \sum_{n=1}^{\infty} a_n^* z^n$$

is univalent in $\mathbb{D}$. Observe that

$$(3) \qquad a_1^* = 1 - \sum_{j=1}^{3} \varepsilon_j w_j^{-1},$$

$$(4) \qquad a_2^* = a_2 - \sum_{j=1}^{3} \varepsilon_j (a_2 w_j^{-1} + w_j^{-2}).$$

In view of (3), the condition

$$(5) \qquad \sum_{j=1}^{3} \varepsilon_j w_j^{-1} = 0$$

will ensure that $f^* \in S$.

The transfinite diameter R^* of the region $f^*(\mathbb{D})$ can be calculated by the method of interior variation, which gives a formula for Green's function $g^*(w, \infty)$ of the complement of $\overline{f^*(\mathbb{D})}$. Let $g(w, \infty)$ be Green's function of the complement of $\overline{f(\mathbb{D})}$ with pole at infinity, and let $p(w) = p(w, \infty)$ be its (multiple-valued) analytic completion. Then we have the variational formula (cf. [11])

$$R^* = R \left(1 - \operatorname{Re} \left\{ \sum_{j=1}^{3} \varepsilon_j w_j p'(w_j)^2 \right\} \right) + O(\varepsilon^2).$$

In order to have $R^* = R$, we must first require

$$(6) \qquad \sum_{j=1}^{3} \varepsilon_j w_j p'(w_j)^2 = 0.$$

Then under conditions (5) and (6), the variation (2) may be modified by adding a term of order ε^2 to achieve $R^* = R$ while retaining $a_1^* = 1$. The modified variation preserves the class S_R and changes formula (4) for a_2^* only by the addition of a term of order ε^2. (This last step is based on the implicit function theorem and is a standard technique for constrained variation. See [6 or 12] for a more detailed discussion.)

Now hold the ε_j fixed at values (not all zeros) satisfying (5) and (6), and let ε_j be replaced by $\rho e^{i\theta} \varepsilon_j$, where $e^{i\theta}$ is an arbitrary constant of unit modulus and ρ is a small positive parameter. Note that conditions (5) and (6) are unaffected. After modifying the variation (2) as indicated above, we obtain $f^* \in S_R$ and

$$(7) \qquad a_2^* = a_2 - \rho e^{i\theta} \sum_{j=1}^{3} \varepsilon_j (a_2 w_j^{-1} + w_j^{-2}) + O(\rho^2).$$

Suppose now that f maximizes $\operatorname{Re}\{a_2\}$ in S_R. Then $\operatorname{Re}\{a_2^*\} \leq \operatorname{Re}\{a_2\}$, and (7) gives

$$\sum_{j=1}^{3} \varepsilon_j (a_2 w_j^{-1} + w_j^{-2}) = 0.$$

In view of (5), this reduces to

$$(8) \qquad \sum_{j=1}^{3} \varepsilon_j w_j^{-2} = 0.$$

Conditions (5), (6), and (8) constitute a linear homogeneous system of equations with a nontrivial solution $(\varepsilon_1, \varepsilon_2, \varepsilon_3)$. Therefore, the determinant of coefficients vanishes:

$$(9) \qquad \begin{vmatrix} w_1^{-1} & w_2^{-1} & w_3^{-1} \\ w_1^{-2} & w_2^{-2} & w_3^{-2} \\ w_1 p'(w_1)^2 & w_2 p'(w_2)^2 & w_3 p'(w_3)^2 \end{vmatrix} = 0.$$

This is a necessary condition for the extremal function f, valid for each choice of distinct points w_j outside the closure of the range of f.

In order to exploit condition (9), we hold w_2 and w_3 fixed $(w_2 \neq w_3)$ and regard $w_1 = w$ as variable. Expansion of determinant (9) along the first column gives

$$\lambda_1 w^{-1} + \lambda_2 w^{-2} + \lambda_3 w p'(w)^2 = 0,$$

where $\lambda_j = \lambda_j(w_2, w_3)$ and $\lambda_3 \neq 0$ since $w_2 \neq w_3$. Thus we have

$$(10) \qquad p'(w)^2 = aw^{-2} + bw^{-3}$$

for some constants a and b.

Recall now that $p(w)$ is the analytic completion of Green's function. It has the form

$$p(w) = \log w + c_0 + c_1 w^{-1} + c_2 w^{-2} + \cdots$$

near infinity, and so

$$(11) \qquad p'(w)^2 = w^{-2} - 2c_1 w^{-3} + \cdots.$$

A comparison of (10) and (11) shows that $a = 1$. Thus

$$(12) \qquad w^2 p'(w)^2 = 1 + b/w, \qquad b \in \mathbb{C}.$$

On the other hand, $p(w)$ is purely imaginary on the boundary of $\overline{f(\mathbb{D})}$, and so we may conclude informally that the boundary is composed of arcs $w = w(t)$ which satisfy the differential equation

$$[p'(w(t))w'(t)]^2 < 0.$$

In view of (12), this suggests that the outer boundary of $f(\mathbb{D})$ lies on trajectories of the quadratic differential

$$(13) \qquad -\left(1 + \frac{b}{w}\right) \frac{dw^2}{w^2} > 0.$$

In fact, the entire boundary of $f(\mathbb{D})$ satisfies (13). For a proof which will apply also to internal spires (if any), we now develop a special kind of boundary

variation which preserves the family S_R. This is an auxiliary variation which will make use of expression (12) for $p'(w)$.

Let Γ be the boundary of $f(\mathbb{D})$. Choosing an arbitrary point $w_0 \in \Gamma$, we begin with the boundary variation (see [10] or [3, Chapter 10])

$$(14) \qquad w^* = w + \frac{a\rho^2 w}{w_0(w - w_0)} + O(\rho^3),$$

which is analytic and univalent outside a small subcontinuum of Γ containing w_0, and vanishes at the origin. Next we choose distinct points w_1, $w_2 \notin \overline{f(\mathbb{D})}$ and modify the variation (14) by adding two terms $\varepsilon_j w(w - w_j)^{-1}$, where the ε_j are of order ρ^2 and will be specified later. This produces a variation

$$w^* = V_\rho(w) = w + \frac{a\rho^2 w}{w_0(w - w_0)} + \frac{\varepsilon_1 w}{w - w_1} + \frac{\varepsilon_2 w}{w - w_2} + O(\rho^3)$$

which is analytic and univalent outside a small subcontinuum of Γ near w_0 except for small disks centered at w_1 and w_2. Then $f^* = V_\rho \circ f$ is analytic and univalent in $\mathbb{D}$, and $f^*(0) = 0$. In order to make $f^{*\prime}(0) = 1$, we must require that $V_\rho'(0) = 1$, or

$$(15) \qquad a\rho^2 w_0^{-2} + \varepsilon_1 w_1^{-1} + \varepsilon_2 w_2^{-1} = 0.$$

Then $f^* \in S$.

The transfinite diameter of $f^*(\mathbb{D})$ is

$$R^* = R[1 - \text{Re}\{\varepsilon_1 w_1 p'(w_1)^2 + \varepsilon_2 w_2 p'(w_2)^2\}] + O(\rho^3).$$

In order to make $R^* = R$, we put a second demand

$$(16) \qquad \varepsilon_1 w_1 p'(w_1)^2 + \varepsilon_2 w_2 p'(w_2)^2 = 0$$

on the variation. We then adjust the error term $O(\rho^3)$ to have $R^* = R$ and $a_1^* = 1$. Thus $f^* \in S_R$.

In view of (12), condition (16) takes the form

$$\frac{\varepsilon_1}{w_1}\left(1 + \frac{b}{w_1}\right) + \frac{\varepsilon_2}{w_2}\left(1 + \frac{b}{w_2}\right) = 0.$$

Combining this with (15), we obtain

$$(17) \qquad a\rho^2 w_0^{-2} = b(\varepsilon_1 w_1^{-2} + \varepsilon_2 w_2^{-2}).$$

A calculation now gives

$$
\begin{aligned}
a_2^* = a_2 &- \frac{a\rho^2}{w_0^2}\left(a_2 + \frac{1}{w_0}\right) - \frac{\varepsilon_1}{w_1}\left(a_2 + \frac{1}{w_1}\right) \\
(18) \qquad &- \frac{\varepsilon_2}{w_2}\left(a_2 + \frac{1}{w_2}\right) + O(\rho^3) \\
= a_2 &- \frac{a\rho^2}{bw_0^2}\left(1 + \frac{b}{w_0}\right) + O(\rho^3),
\end{aligned}
$$

where (15) and (17) have been used. The assumption that $b \neq 0$ is justified by (12). Indeed, if $w^2 p'(w)^2 = 1$, then Green's function takes the form $g(w) = \log|w| - \log R$, and the circle $|w| = R$ is the outer boundary of $f(\mathbb{D})$. This returns us to Pick's problem (see §0), so the sharp bound would be $|a_2| \leq 2(1 - 1/R)$. Later, however, it will be apparent that $|a_2|$ can always be larger for functions in the class S_R.

Since $f^* \in S_R$ and f is extremal, we have $\text{Re}\{a_2^*\} \leq \text{Re}\{a_2\}$. Thus (18) gives

$$\text{Re}\left\{ \frac{a\rho^2}{bw_0^2} \left(1 + \frac{b}{w_0} \right) + O(\rho^3) \right\} \geq 0$$

for each $w_0 \in \Gamma$. It now follows from the fundamental lemma of the method of boundary variation ([10]; see also [3, p. 297]) that Γ consists of analytic arcs which satisfy

$$(19) \qquad -\frac{1}{b} \left(1 + \frac{b}{w} \right) \frac{dw^2}{w^2} > 0.$$

It will be seen presently that $b > 0$, so that (19) and (13) are equivalent.

2. ANALYSIS OF THE QUADRATIC DIFFERENTIAL

The quadratic differential (19) has a simple zero at $-b$, a triple pole at the origin, and a double pole at infinity. Parametrizing Γ by $w = f(e^{it})$, we conclude from (19) that

$$(20) \qquad F(z) = \frac{z^2 f'(z)^2}{bf(z)^3}[f(z) + b] \geq 0$$

on $|z| = 1$. Since $f \in S$, it is clear that F is analytic in $\mathbb{D}$ except for a simple pole at the origin with residue 1. Taking into account all possible singularities of Γ at singular points of the quadratic differential, one shows that F is analytic and nonvanishing on the unit circle, except perhaps for one double zero. By the Schwarz reflection principle, F has a meromorphic continuation to the Riemann sphere which satisfies $F(1/\bar{z}) = \overline{F(z)}$. Thus F has the form

$$F(z) = z + 2c + 1/z,$$

where $c \geq 1$. But a short calculation gives $2c = a_2 + 1/b$. Since $0 < a_2 < 2$, we conclude that $b > 0$. In particular, conditions (19) and (13) are equivalent.

Because $b > 0$, the trajectories of (13) are symmetric with respect to the real axis, and the real segment $-b < w < 0$ is a trajectory. The structure of the trajectories is shown in Figure 1.

There are now three cases to consider.

Case I. Γ is an analytic Jordan curve surrounding the point $-b$. Then F is analytic on the unit circle and $F(z) \neq 0$ in $\overline{\mathbb{D}}$ except for a simple zero at the point $-\eta$ $(0 < \eta < 1)$ where $f(-\eta) = -b$.
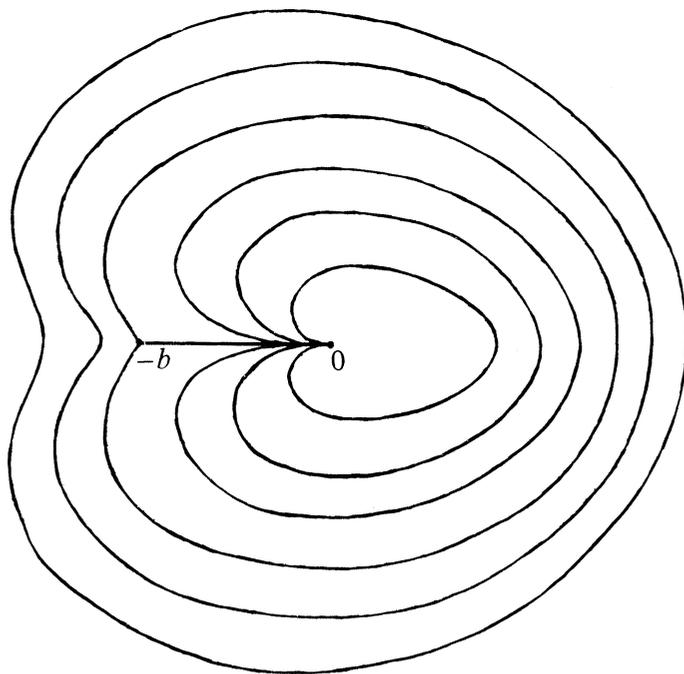
FIGURE 1. Trajectories of the quadratic differential

Case II. Γ is a Jordan curve passing through $-b$. Then F is analytic and nonvanishing on the unit circle except at the point -1 where $f(-1) = -b$. A calculation shows that F has a double zero at -1. Except for the pole at the origin, F is analytic and nonvanishing in $\mathbb{D}$.

Case III. Γ consists of a Jordan curve through $-b$ plus a segment $[-b, \tau]$ of the real axis, where $-b < \tau < -\frac{1}{4}$. Then we find that F is analytic and nonvanishing in $\overline{\mathbb{D}}\backslash\{0\}$ except for a double zero at the point -1 where $f(-1) = \tau$.

In Cases II and III, the double zero on the boundary implies that $c = 1$ and so

(21) $$F(z) = z(1 + 1/z)^2$$

and $b = (2 - a_2)^{-1} > \frac{1}{2}$. Because of the perfect square, integration of the differential equation is elementary in Cases II and III. This will be carried out in §3, while Case I will be deferred to §§4 and 5.

3. CASES II AND III: CORNERS AND INTERNAL SPIRES

Suppose now that either Case II or III prevails, so that F is given by (21) and $b = (2 - a_2)^{-1}$. The differential equation for $w = f(z)$ may then be expressed in the form

$$\left(1 + \frac{b}{w}\right)^{1/2} \frac{dw}{w} = \sqrt{b}\left(1 + \frac{1}{z}\right) \frac{dz}{\sqrt{z}}.$$

Integration of the left-hand side gives

$$G(w) = \int_{-b}^{w} \left(1 + \frac{b}{w}\right)^{1/2} \frac{dw}{w}$$

$$(22) \qquad = -2\left(1 + \frac{b}{w}\right)^{1/2} + \log \frac{1 + (1 + b/w)^{1/2}}{1 - (1 + b/w)^{1/2}}$$

$$= \log w + \log(-4/b) - 2 + O(1/w).$$

Recalling that Γ is a trajectory of (13), we see that $\mathrm{Re}\{G(w)\} = 0$ on Γ. Furthermore, G is analytic outside $\overline{f(\mathbb{D})}$ apart from the logarithmic singularity at infinity. It follows that $\mathrm{Re}\{G(w)\}$ is Green's function

$$g(w) = \log|w| - \log R + O(1/w)$$

of the complement of $\overline{f(\mathbb{D})}$, and that

$$(23) \qquad\qquad \log R = \log b - \log 4 + 2,$$

or $R = be^2/4$. But $b = (2 - a_2)^{-1}$, so $a_2 = 2 - e^2/4R$. Because f maximizes $\mathrm{Re}\{a_2\}$, this gives the sharp bound

$$(24) \qquad\qquad |a_2| \le 2 - e^2/4R$$

whenever Case II or III occurs.

It remains to determine the values of R for which Case II or III can occur as an extremal configuration. These will correspond to the values of $b = 4R/e^2$ for which the Jordan curve of Case II bounds a region Ω_b of inner radius $r_b \ge 1$. If $r_b > 1$, then Ω_b can be modified by the insertion of a radial slit (as in Case III) to reduce its inner radius to 1. If $r_b < 1$, however, no admissible modification can change Ω_b to the range of a function in S, and Case I must prevail.

It is important, therefore, to calculate the inner radius r_b of the region Ω_b. Let

$$w = \varphi(z) = \alpha z + \beta z^2 + \cdots, \qquad \alpha = r_b > 0,$$

map $\mathbb{D}$ conformally onto Ω_b. Parametrizing the boundary of Ω_b by $w = \varphi(e^{it})$ and invoking (13), we obtain as in §2

$$(25) \qquad \left(1 + \frac{b}{w}\right)\frac{dw^2}{w^2} = \frac{b}{\alpha z}\left(1 + \frac{1}{z}\right)^2 dz^2.$$

For small $\varepsilon > 0$, integration of (25) gives

$$G(f(-\varepsilon)) = \int_{-b}^{f(-\varepsilon)} \left(1 + \frac{b}{w}\right)^{1/2} \frac{dw}{w} = \left(\frac{b}{\alpha}\right)^{1/2} \int_{-1}^{-\varepsilon} (1 + z) z^{-3/2}\, dz$$

$$= 2\left(\frac{b}{\alpha}\right)^{1/2} (\varepsilon^{1/2} - \varepsilon^{-1/2} - 2)i.$$

But by (22),

$$(26) \qquad G(f(-\varepsilon)) = -2i\left(\frac{b}{\alpha\varepsilon}\right)^{1/2} - \pi i + O(\sqrt{\varepsilon}).$$

Comparing the two expressions and letting ε tend to 0, we obtain

$$r_b = \frac{16b}{\pi^2} = \frac{64R}{e^2\pi^2}.$$

Thus $r_b < 1$ precisely when

$$R < e^2\pi^2/64 = 1.139\ldots.$$

This proves that the extremal function cannot have the form of Case II or III if $1 < R < e^2\pi^2/64$. In other words, *each extremal function must have the form of Case I if* $1 < R < e^2\pi^2/64$.

So far it is unclear whether Case I or III prevails if $R > e^2\pi^2/64$, or whether Case I or II prevails if $R = e^2\pi^2/64$. In the next two sections we shall show that if an extremal function has the form of Case I, then $1 < R < e^2\pi^2/64$. Since we have just proved the converse, this will imply that Case I prevails if $R < e^2\pi^2/64$, Case II if $R = e^2\pi^2/64$, and Case III if $R > e^2\pi^2/64$.

4. Case I: analytic Jordan curves

Suppose now that Case I describes an extremal function $f \in S_R$ for some R. Then f maps $\mathbb{D}$ onto the interior of an analytic Jordan curve Γ which is a trajectory of the quadratic differential (13) for some $b > 0$. Furthermore, f satisfies the differential equation

(27) $$F(z) = z + 2c + 1/z$$

for some $c = \frac{1}{2}(a_2 + 1/b) > 1$, where F is defined by (20). The point $-\eta$ where $f(-\eta) = -b$ is a zero of F, so the quadratic equation gives $\eta = c - (c^2 - 1)^{1/2}$.

Now consider the function

$$H(w) = \int_{w_0}^{w} \left\{ 1 + \frac{b}{w} \right\}^{1/2} \frac{dw}{w} = G(w) - G(w_0),$$

where $w_0 = f(-1)$ is the point at which Γ meets the negative real axis. Because Γ satisfies (13), it follows as before that $\mathrm{Re}\{H(z)\}$ is Green's function of the region outside Γ. Comparison with (22) gives

$$\log R = \log b - \log 4 + 2 + \mathrm{Re}\{G(w_0)\}.$$

But by (27)

$$G(w_0) = \int_{-b}^{w_0} \left\{ 1 + \frac{b}{w} \right\}^{1/2} \frac{dw}{w} = \sqrt{b} \int_{-\eta}^{-1} \left\{ z + 2c + \frac{1}{z} \right\}^{1/2} \frac{dz}{z}.$$

Hence

(28) $$\log R = 2 + \log(b/4) + \sqrt{b}\,I,$$

where

(29) $$I = \int_{\eta}^{1} \left\{ 2c - \left(t + \frac{1}{t} \right) \right\}^{1/2} \frac{dt}{t}.$$

A further relation is obtained from the fact that Γ must enclose a region of inner radius 1. For small $\varepsilon > 0$ we have by (27)

$$b^{-1/2}G(f(-\varepsilon)) = \int_{-\eta}^{-\varepsilon} \left(t + 2c + \frac{1}{t}\right)^{1/2} \frac{dt}{t}$$

$$= -i\int_{\varepsilon}^{\eta}\{(1 - 2ct + t^2)^{1/2} - 1\}t^{-3/2}\,dt + 2i(\eta^{-1/2} - \varepsilon^{-1/2}).$$

Referring to (26) with $\alpha = 1$, we conclude that

$$(30) \qquad\qquad -b^{-1/2}\pi = 2\eta^{-1/2} - J,$$

where

$$(31) \qquad\qquad J = \int_0^\eta \{(1 - 2ct + t^2)^{1/2} - 1\}t^{-3/2}\,dt.$$

The integral (31) is easily simplified. Integration by parts gives

$$J = 2\eta^{-1/2} + 2\int_0^\eta (1 - 2ct + t^2)^{-1/2}(t - c)t^{-1/2}\,dt.$$

Observe now that the integrand has an analytic extension to the closed unit disk with a branch cut along the segment from 0 to η. By Cauchy's integral theorem, the contour may be deformed to the unit circle $\mathbb{T}$. Thus (30) becomes

$$\pi b^{-1/2} = \int_{\mathbb{T}} \left(z - 2c + \frac{1}{z}\right)^{-1/2}(z - c)\frac{dz}{z},$$

or

$$(32) \qquad\qquad \int_0^{2\pi} (c + \cos\theta)^{1/2}\,d\theta = \frac{\pi\sqrt{2}}{\sqrt{b}},$$

where

$$(33) \qquad\qquad 2c = a_2 + 1/b > 2.$$

Equations (28), (32), and (33) allow the computation of a_2 and R, with c as a parameter. The computation is facilitated by expressing the integrals in terms of standard elliptic integrals. This will be done in the next section.

5. ELLIPTIC INTEGRALS

The integral I defined by (29) may be transformed as follows. Let $k = (1 - \eta)/(1 + \eta)$ and make the substitution $x = (1 - t)/(1 + t)$. Then

$$2I = \int_\eta^{1/\eta} \left\{2c - \left(t + \frac{1}{t}\right)\right\}^{1/2} \frac{dt}{t}$$

$$= 2\sqrt{2}\int_{-k}^k \left\{\frac{1 + k^2}{1 - k^2} - \frac{1 + x^2}{1 - x^2}\right\}^{1/2} \frac{dx}{1 - x^2},$$

since $2c = \eta + 1/\eta$. Making the further substitution $x = k \sin \theta$, we then obtain

$$I = 4k^2 (1 - k^2)^{-1/2} \int_0^{\pi/2} \cos^2 \theta (1 - k^2 \sin^2 \theta)^{-3/2} \, d\theta,$$

where $0 < k < 1$.

The last integral can be expressed in terms of the standard complete elliptic integrals

$$K = K(k) = \int_0^{\pi/2} (1 - k^2 \sin^2 \theta)^{-1/2} \, d\theta$$

and

$$E = E(k) = \int_0^{\pi/2} (1 - k^2 \sin^2 \theta)^{1/2} \, d\theta$$

of the first and second kinds, respectively (see [1] or [2]). It may be verified that

$$(34) \qquad\qquad I = (4/k')(K - E),$$

where $k' = (1 - k^2)^{1/2}$ is the complementary modulus. Substituting (34) into (28), we have

$$(35) \qquad\qquad \log R = 2 + \log(b/4) + 4\sqrt{b}(K - E)/k'.$$

The integral in (32) can also be reduced to standard form. The half-angle formula gives

$$\int_0^{2\pi} (c + \cos \theta)^{1/2} \, d\theta = \frac{4\sqrt{2}}{k'} E',$$

where

$$K' = K'(k) = K(k'); \qquad E' = E'(k) = E(k').$$

Thus (32) reduces to

$$(36) \qquad\qquad 4E' = \pi k'/\sqrt{b},$$

while (33) takes the form

$$(37) \qquad\qquad a_2 = 2\frac{1 + k^2}{1 - k^2} - \frac{16E'^2}{\pi^2 k'^2}.$$

Combining (35) and (36) to eliminate b, we obtain the basic relation

$$(38) \qquad \log R = \log(e^2 \pi^2/64) + 2\log k' - 2\log E' + \pi(K - E)/E'.$$

Observe first that for $k = 0$ (a limiting value for Case I), $K = E = \pi/2$ and $E' = 1$, so that (38) gives $R = e^2 \pi^2/64$. For $k = 1$ we have $E = 1$ and $E' = \pi/2$, while

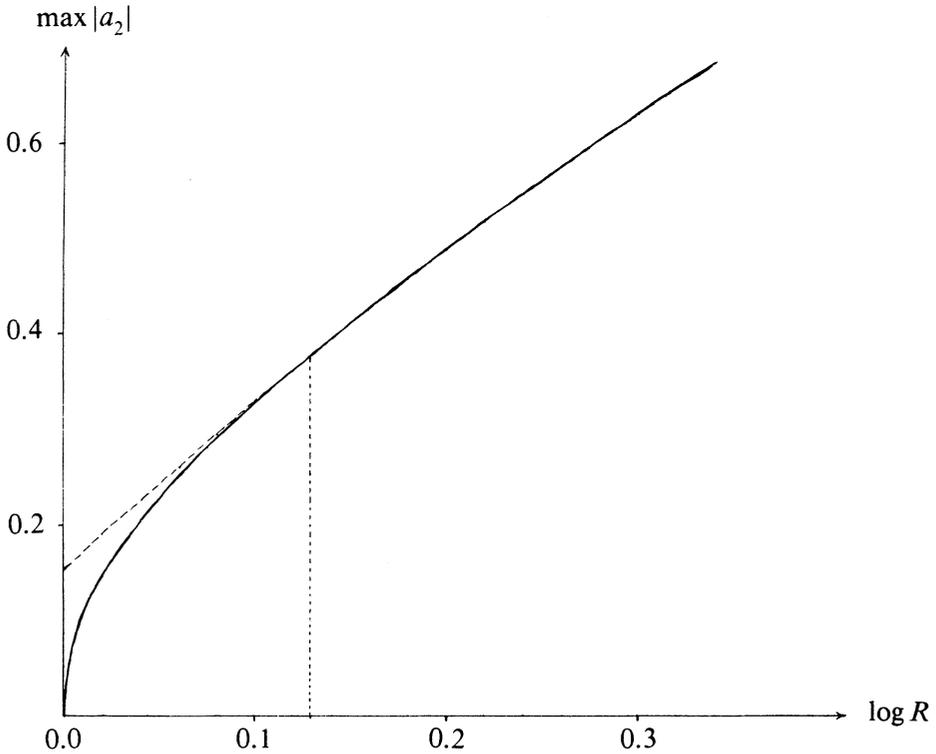$$(39) \qquad\qquad \lim_{k \to 1} \{K(k) - \log(4/k')\} = 0.$$

FIGURE 2. The solution for small R

(See [2, #112.01].) Thus a short calculation based on (38) shows that $R \to 1$ as $k \to 1$.

We shall now show on the basis of (38) that R is a strictly decreasing function of k, $0 < k < 1$. This will establish our assertion that $1 < R < e^2\pi^2/64$ if the extremal function has the form of Case I. It will also show that R uniquely determines k, b, c, and a_2 via equations (38), (36), and (37).

In using (38) to calculate the derivative of $\log R$ with respect to k, we must use the formulas [2, #710.00 and #710.02]

$$(40) \qquad \frac{dK}{dk} = \frac{E - k'^2 K}{kk'^2}; \qquad \frac{dE}{dk} = \frac{E - K}{k}.$$

Using also the simple formula $dk'/dk = -k/k'$, we find after some manipulation

$$(41) \qquad \begin{aligned} k'^2 E'^2 \frac{d(\log R)}{dk} &= -2kE'K' + \pi k(E'K + EK' - KK') \\ &= k(\pi^2/2 - 2E'K'), \end{aligned}$$

where Legendre's relation [2, #110.10] was used in the last step. But the Schwarz

inequality shows immediately that $E'K' > \pi^2/4$. This completes the proof that

$$\frac{d(\log R)}{dk} < 0, \qquad 0 < k < 1.$$

Consequently, we have shown that $1 < R < e^2\pi^2/64$ *whenever Case I prevails.* Hence in view of the results in §3, we have proved the following theorem.

Theorem. *For all functions* $f \in S$ *with range of transfinite diameter* R, *the sharp inequality* $|a_2| \leq 2 - e^2/4R$ *holds if* $R \geq e^2\pi^2/64$. *If* $1 < R < e^2\pi^2/64$, *the sharp inequality involves elliptic integrals and is implicitly determined by* (37) *and* (38), *with* k ($0 < k < 1$) *acting as a parameter. In all cases the extremal function is unique up to a rotation, and the function maximizing* $\text{Re}\{a_2\}$ *maps the disk onto a region bounded by trajectories of the quadratic differential* (13), *as shown in Figure* 1, *where* $b = 4e^{-2}R$ *if* $R \geq e^2\pi^2/64$ *and* $b = \pi^2(1-k^2)E'(k)^{-2}/16$ *if* $1 < R < e^2\pi^2/64$. *If* $R > e^2\pi^2/64$, *the range of the extremal function is bounded by a Jordan curve which is analytic except for a corner at the point* $-b$, *plus a slit along the negative real axis from* $-b$ *toward the origin. The length of this linear slit is determined by the requirement that the curves bound a region of inner radius* 1. *If* $R = e^2\pi^2/64$, *the slit is not present. If* $R < e^2\pi^2/64$, *the extremal range is bounded by an analytic Jordan curve surrounding the point* $-b$. *This trajectory of* (13) *is determined by the requirement that its interior have inner radius* 1.

The relations (37) and (38), together with a table of complete elliptic integrals, can be used to compute numerical values for the sharp bound on $|a_2|$ in S_R for $1 < R < e^2\pi^2/64$. A sample is shown in Table 1, and a graph is plotted in Figure 2. In Figure 2 the vertical dotted line marks the junction of the two curves at $R = e^2\pi^2/64$, and the dashes to the left indicate the continuation of the curve $a_2 = 2 - e^2/4R$. The graph suggests that the two curves have a common tangent line at the junction, and a calculation confirms this. One may compute da_2/dk from (37), using (40). Equation (41) gives dR/dk. Taking the ratio and letting k tend to 0, one finds from (39) that

$$\frac{da_2}{dR} = \frac{1024}{e^2\pi^4} = 1.4222\ldots$$

at $R = e^2\pi^2/64$. The curve $a_2 = 2 - e^2/4R$ is easily seen to have the same derivative there.

Note added in proof. Professor J. A. Jenkins has pointed out to us that the theorem of this paper is implicit in Theorem 7.3 of his book *Univalent functions and conformal mapping* (Springer-Verlag, 1958). His derivation relies on the General Coefficient Theorem. Our variational approach has other applications; for instance, R. Cunningham has applied it to find the maximum of $|f(\zeta)|$ for fixed $\zeta \in \mathbb{D}$ as f ranges over the class S_R.

TABLE 1. Maximum of $|a_2|$ for $1 \leq R \leq e^2\pi^2/64$

k^2	R	max $|a_2|$
1.00	1.0000	0.0000
.90	1.0002	0.0131
.80	1.0008	0.0278
.70	1.0020	0.0444
.60	1.0040	0.0633
.50	1.0073	0.0853
.40	1.0124	0.1115
.30	1.0206	0.1437
.20	1.0343	0.1856
.15	1.0448	0.2125
.10	1.0599	0.2459
.08	1.0680	0.2622
.06	1.0779	0.2809
.04	1.0905	0.3031
.02	1.1077	0.3315
.01	1.1198	0.3501
.00	1.1395	0.3789

REFERENCES

1. Bateman Manuscript Project (A. Erdélyi, ed), *Higher transcendental functions*, Vol. II, McGraw-Hill, New York, 1953.

2. P. F. Byrd and M. D. Friedman, *Handbook of elliptic integrals for engineers and physicists*, Springer-Verlag, Berlin and New York, 1954.

3. P. L. Duren, *Univalent functions*, Springer-Verlag, Berlin and New York, 1983.

4. G. M. Goluzin, *Geometric theory of functions of a complex variable*, Moscow, 1952; German. transl., Deutscher Verlag, Berlin, 1957; 2nd ed., Moscow, 1966; English transl., Amer. Math. Soc., Providence, R. I., 1969.

5. E. Hille, *Analytic function theory*, Vol. II, Ginn, Boston, Mass., 1962; 2nd ed., Chelsea, New York, 1987.

6. J. A. Hummel, *Lagrange multipliers in variational methods for univalent functions*, J. Analyse Math. **32** (1977), 222–234.

7. G. Pick, *Uber die konforme Abbildung eines Kreises auf ein schlichtes und zugleich beschränktes Gebiet*, S.-B. Kaiserl. Akad. Wiss. Wien **126** (1917), 247–263.

8. G. Pólya, *Beitrag zur Verallgemeinerung des Verzerrungssatzes auf mehrfach zusammenhän-gende Gebiete.* I, II, S.-B. Preuss. Akad. Wiss. (1928), 228–232; 280–282.

9. G. Pólya and G. Szegö, *Aufgaben und Lehrsätze aus der Analysis,* Zweiter Band, Springer-Verlag, Berlin, 1925; English transl. *Problems and theorems in analysis,* vol. II, Springer-Verlag, New York, 1976.

10. M. Schiffer, *A method of variation within the family of simple functions,* Proc. London. Math. Soc. **44** (1938), 432–449.

11. _____, *Hadamard's formula and variation of domain-functions,* Amer. J. Math. **68** (1946), 417–448.

12. M. Schiffer and D. C. Spencer, *Functionals of finite Riemann surfaces,* Princeton Univ. Press, Princeton, N. J., 1954.

13. G. Szegö, *Bermerkungen zu einer Arbeit von Herrn M. Fekete: Über die Verteilung der Wurzeln bei gewissen algebraischen Gleichungen mit ganzzahligen Koeffizienten,* Math. Z. **21** (1924), 203–208.

14. M. Tsuji, *Potential theory in modern function theory,* Maruzen, Tokyo, 1959; 2nd ed., Chelsea, New York, 1975.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF MICHIGAN, ANN ARBOR, MICHIGAN 48109

DEPARTMENT OF MATHEMATICS, STANFORD UNIVERSITY, STANFORD, CALIFORNIA 94305