The Annals of Probability
1983, Vol. 11, No. 4, 833–846

BROWNIAN MOTION, GEOMETRY, AND GENERALIZATIONS OF PICARD'S LITTLE THEOREM

BY S. I. GOLDBERG[1] AND C. MUELLER[2]

University of Illinois and University of Texas

Brownian motion is introduced as a tool in Riemannian geometry to show how useful it is in the function theory of manifolds, as well as the study of maps between manifolds. As applications, a generalization of Picard's little theorem, and a version of it for Riemann surfaces of large genus are given.

Introduction. The purpose of this paper is to introduce Brownian motion as a tool in Riemannian geometry. In other branches of analysis, notably potential theory, and more recently complex analysis, probabilistic methods have already achieved important results. From one point of view this is not surprising, since Brownian motion is intimately connected with harmonic functions, the Laplacian, and other fundamental objects in analysis. The purpose here is to show that Brownian motion is also a useful and fundamental tool in the function theory of manifolds, and in the study of maps between manifolds. Indeed, Brownian motion on manifolds has been studied for some time by probabilists, and there are many fascinating differences between Brownian motion on manifolds and on Euclidean space. Pinsky [12] presents a useful survey of this work. On the other hand, whereas nonpositive curvature plays an essential role in geometric function theory, this is not the case in our considerations.

In choosing a geometric problem to attack, the authors were guided by the analogy with Brownian motion on Euclidean space. One of the most intriguing results in that setting has been Burgess Davis's probabilistic proof of the little Picard theorem [2], based on the winding properties of Brownian motion. It is felt that many less obvious applications of Brownian motion to geometry are possible, and the adventurous reader is invited to explore these possibilities.

Our main problem arises from the generalized Picard theorem for harmonic maps between manifolds, proved by Goldberg and Har'El [6], and Davis's technique is used to extend this result. The theorem states that, given certain assumptions on the sectional curvatures of the Riemannian manifolds M and N, any harmonic map of bounded dilatation between M and N must be constant. Using probability, Kendall [9] has already proved a restricted version of the Goldberg-Har'El theorem which requires an additional condition on the curvature of N. Our result still requires some restrictions, but includes many cases not covered by the original theorem. In particular, while the Goldberg-Har'El theorem assumes that the sectional curvatures of N are bounded above by a negative constant, our theorem allows the curvature to tail off to 0, and even allows regions of positive curvature. The main probabilistic tool used is the tail σ-field of Brownian motion, which is trivial for some manifolds but not for others.

Our second theorem is also motivated by Davis' work, and deals with Picard's theorem for Riemann surfaces. It is our contention that the connection between Brownian motion and complex analysis could well be exploited to study complex manifolds. Chern [1] has extended Picard's theorem to Riemann surfaces of low genus. Using the winding properties of Brownian motion on manifolds, a version of Picard's theorem which holds for Riemann surfaces of large genus is established.

The paper is divided into two parts, each dealing with one of the theorems. The proof

Received July 1982; revised January 1983.

[1] Queen's Quest Professor, Queen's University, Kingston, Canada.

[2] NSF Postdoctoral Fellow, University of Illinois, Urbana, 1979–81.

AMS 1980 *subject classifications*. 32H25, 53C21, 60J65.

Key words and phrases. Brownian motion, harmonic and quasiconformal mappings, sectional curvature, tail σ-field.

of the first theorem is given in Part I in the form of a sequence of lemmas. Some of the techniques are similar to those of Kendall [9], but the actual tactics require considerable modification.

Whereas the manifolds considered in Part I are noncompact and of arbitrary dimension, only compact Riemann surfaces are considered in Part II. Let M and N be compact Riemann surfaces with a finite number of points deleted. Any such surface N is homeomorphic to a sphere with $t(N)$ tori attached and $p(N)$ points deleted. Let $n(N) = 2t(N) + p(N)$. If $t(M) > 0, p(M) > 0, p(N) > 1$ and $n(N) > n(M)$, then the second theorem says that any holomorphic map from M to N must be a constant map.

Part I

1. Definitions and statement of the theorem. Let M and N be complete Riemannian manifolds with dimensions m, n, metrics ${}^M g_{ij}$, ${}^N g_{\alpha\beta}$, and Christoffel symbols ${}^M\Gamma^k_{ij}$, ${}^N\Gamma^\gamma_{\alpha\beta}$, respectively. Assume that $F:M \to N$ is a C^2 map. F is said to be harmonic [3] if its second fundamental form

$$(\nabla(dF))^\gamma_{ij} = \frac{\partial^2 F^\gamma}{\partial x^i \partial x^j} - {}^M\Gamma^k_{ij}\frac{\partial F^\gamma}{\partial x^k} + {}^N\Gamma^\gamma_{\alpha\beta}\frac{\partial F^\alpha}{\partial x^i}\frac{\partial F^\beta}{\partial x^j}$$

has trace 0. Define the tensor $\xi^{\alpha\beta}(x)$ on M by

$$\xi^{\alpha\beta}(x) = {}^M g_{ij}\left[\frac{\partial F^\alpha}{\partial x^i}\frac{\partial F^\beta}{\partial x^j}\right](x), \quad x \in M.$$

Note that $(\xi^{\alpha\beta}(x))$ is a symmetric matrix, so the eigenvalues are nonnegative, and we may order them as follows: $\lambda_1(x) \geq \lambda_2(x) \geq \cdots \geq \lambda_n(x) \geq 0$. F is said to be *K-quasiconformal* [7] if for all $x \in M$, $\lambda_1(x) \leq K^2\lambda_n(x)$. It is of *K-bounded dilatation* if for all $x \in M$, $\lambda_1(x) \leq K^2\lambda_2(x)$. In [6], Goldberg and Har'El proved the following:

THEOREM. *Let M be a complete connected Riemannian manifold with nonnegative Ricci curvature, and let N be a Riemannian manifold with negative sectional curvature bounded away from zero. Then a harmonic mapping $f:M \to N$ of bounded dilatation is a constant mapping.*

Our theorem relaxes the requirement that the curvature be bounded away from zero, but imposes some additional conditions. First, we wish to set up polar coordinates (r, θ) on N. Choose a point $x_0 \in N$, and if θ is in the unit sphere of the tangent space N_{x_0} and $r > 0$, associate a point $p \in N$ with (r, θ) through the exponential map: $p = \exp_{x_0}(r\theta)$. Now we can state our conditions on N:

(i) *The sectional curvatures of N are bounded below by $-L^2 < 0$.*
(ii) *Each of the sectional curvatures at $(r, \theta) \in N$ determined by dr and some other tangent vector, is bounded above by $K(r)$, where $K(r)$ satisfies:*
  (a) *For some $\varepsilon > 0$, $-K(r) \sim r^{2\varepsilon-2}$;*
  (b) *There exists a C^∞ solution $u(r)$ of the equation*

$$u''(r) = K(r)u(r), \quad u(0) = 0, \quad u'(0) = 1,$$

and $u'(r)$ is always positive.

Note that we can always find such a solution if $K(r)$ is smooth and everywhere negative. Condition (ii, a) allows the sectional curvature to tail off towards 0, a case not covered by the Goldberg-Har'El theorem, but (i) is not required for their theorem. Condition (ii, b) ensures that the map $(r, \theta) \to N$ is one-to-one. Indeed, let the manifold N' be defined by the metric

$$ds^2 = dr^2 + u(r)^2 \textstyle\sum_{i=1}^{n-1} d\theta_i^2$$

where $(dr, d\theta_i)$ is an orthonormal frame. An easy computation shows that the sectional curvature with respect to dr and any vector perpendicular to it is $-u''(r)/u(r) = K(r)$. Since $u'(r) > 0$, we know that there are no points conjugate to the origin ($r = 0$) in N'. By the Rauch comparison theorem [4] and condition (ii), there are no points conjugate to x_0 in N, so polar coordinates are unique.

We need to consider Brownian motion X_t on M. For a definition and many properties, see [12]. The tail σ-field of Brownian motion is the σ-field $\bigcap_{T>0}\sigma\{X_t : t > T\}$. Brownian motion on manifolds such as $\mathbb{R}^n$ has trivial tail σ-field, but this is false in hyperbolic space. We can now state the main theorem.

THEOREM 1. *Suppose that M is a complete connected Riemannian manifold with N satisfying* (i) *and* (ii). *Assume that Brownian motion on M has trivial tail σ-field. Then every K-quasiconformal harmonic map $F: M \to N$ is constant.*

Geometrically, we can take for M a manifold of nonnegative Ricci curvature by virtue of a theorem of Yau [16] regarding bounded harmonic functions on such manifolds.

2. Proof of Theorem 1. Assume F is not constant. The following results will be essential to the proof. Some of our techniques are similar to those of Kendall [9]. Let $x_N \in N\backslash\{x_0\}$ be a point in the range of F.

LEMMA 1. *Let X_t be Brownian motion on M, with X_0 chosen so that $F(X_0) = x_N$. There is a random time change $\sigma(t)$ and a constant $C > 0$, such that if $\rho_t = r \circ F(X_{\sigma(t)})$, then $d\rho_t = a(X_{\sigma(t)})\, dB_t + b(X_{\sigma(t)})\, dt$, when $F(X_t) \neq x_0$, where $B(t)$ is some Brownian motion, $1/K \leq |a(X_{\sigma(t)})| \leq 1$, and $b(X_{\sigma(t)}) \geq C\rho_t^{\varepsilon/2-1}$ if ρ_t is larger than some constant R.*

The statement uses the stochastic calculus, which is discussed in [10].

Let τ be a stopping time for $X_{\sigma(t)}$.

LEMMA 2. *If $\rho_t^{(\tau)}$ is the distance of $F(X_{\sigma(t)})$ from $F(X_{\sigma(\tau)})$, then for $t > \tau$,*

$$d\rho_t^{(\tau)} = a^{(\tau)}(X_{\sigma(t)})\, dB_t^{(\tau)} + b^{(\tau)}(X_{\sigma(t)})\, dt,$$

where

$$\frac{1}{K} \leq |a^{(\tau)}(X_{\sigma(t)})| \leq 1$$

and

$$b^{(\tau)}(X_{\sigma(t)}) \leq \frac{n}{2} L \coth(L\rho_t^{(\tau)}).$$

PROOFS OF LEMMAS 1 AND 2. Let $\sigma(t) = \int_0^t (ds/\lambda_1(x_s))$. Recall that $\lambda_1(x) \geq \cdots \geq \lambda_n(x)$ are the eigenvalues of $(\xi^{\alpha\beta}(x))$. By Ito's lemma,

$$\begin{aligned} d\rho_t &= \sqrt{\xi^{\alpha\beta}\frac{\partial r}{\partial y^\alpha}\frac{\partial r}{\partial y^\beta}}\, dX_{\sigma(t)} + \frac{1}{2}\xi^{\alpha\beta}\left[\frac{\partial^2 r}{\partial y^\alpha \partial y^\beta} - {}^N\Gamma^\gamma_{\alpha\beta}\frac{\partial r}{\partial y^\gamma}\right] d\sigma(t) \\ &= \sqrt{\frac{\xi^{\alpha\beta}}{\lambda_1}\frac{\partial r}{\partial y^\alpha}\frac{\partial r}{\partial y^\beta}}\, dB_t + \frac{1}{2\lambda_1}\xi^{\alpha\beta}\left[\frac{\partial^2 r}{\partial y^\alpha \partial y^\beta} - {}^N\Gamma^\gamma_{\alpha\beta}\frac{\partial r}{\partial y^\gamma}\right] dt, \end{aligned}$$

where we have suppressed the dependence on $X_{\sigma(t)}$, and where B_t is a new Brownian motion. Now, since λ_1 is the maximum eigenvalue of $(\xi^{\alpha\beta})$, and since $\partial r/\partial y^\alpha$ is a vector of length 1, we have $|a(X_{\sigma(t)})| \leq 1$. The lower bound follows from the K-quasiconformal condition. Repeating the same argument in the context of Lemma 2, we find that

$$\frac{1}{K} \leq |a^{(\tau)}(X_{\sigma(t)})| \leq 1.$$

Fix a point $x \in M$, and let (z^γ) be a set of normal geodesic coordinates with origin $F(x)$, such that (z^γ) corresponds to the eigenvectors of $(\xi^{\alpha\beta})$, under the map F. For these coordinates, ${}^N\Gamma^\gamma_{\alpha\beta}(F(x)) = 0$, and we may write

$$\xi^{\alpha\beta}(x)\left[\frac{\partial^2 r}{\partial y^\alpha \partial y^\beta} - {}^N\Gamma^\gamma_{\alpha\beta}\frac{\partial r}{\partial y^\gamma}\right](F(x)) = \sum_{\alpha=1}^n \lambda_\alpha(x)\frac{d^2 r(z^\alpha(s))}{ds^2}\bigg|_{s=0}.$$

Hence,

$$b(X_{\sigma(t)}) = \frac{1}{2}\sum_{\alpha=1}^n \frac{\lambda_\alpha}{\lambda_1}(X_{\sigma(t)})\frac{d^2}{ds^2}r(z^\alpha(s))\bigg|_{s=0}.$$

In the case of Lemma 2, let H be the hyperbolic space of dimension n and constant curvature $-L^2$. A standard computation shows, that if (r, θ) are polar coordinates in H, and if $z(s)$ is a geodesic, then $(d^2/ds^2)r(z(s))\,|_{s=0} \le L\coth(Lr(z(0)))$. Then, the Hessian comparison theorem of Greene and Wu [5], together with condition (i), implies that

$$\frac{d^2}{ds^2}r(z^\alpha(s))\bigg|_{s=0} \le L\coth(Lr(X_{\sigma(t)})).$$

This inequality, together with the formula for $d\rho_t$, implies the bound on $b^{(\tau)}$ stated in Lemma 2. The proof of Lemma 2 is thus complete.

As for Lemma 1, we again use the Hessian comparison theorem, where P is the manifold with the following metric

$$ds^2 = dr^2 + g(r)\sum_{\alpha=1}^{n-1} d\theta_\alpha^2,$$

where $(d\theta_\alpha)$ is a set of normal tangents perpendicular to dr, and $g(r)$ is defined below. First, note that we can find a C^∞ function $\hat{K}(r) > K(r)$ such that for some $R > 0$, $r > R$ implies

$$\hat{K}(r) = \frac{\varepsilon}{2}\left(\frac{\varepsilon}{2} - 1\right)r^{\varepsilon/2-2} = \frac{d^2}{dr^2}r^{\varepsilon/2}.$$

In addition, we may require that a C^∞ solution $g(r)$ of $\hat{u}'' = \hat{K}\hat{u}$, $g(0) = 0$, $g'(0) = 1$ exists for all $r > 0$ and has strictly positive first derivative. Now, choose $0 < c_1 < c_2$ such that

$$c_1 e^{R^{\varepsilon/2}} \le g(R) \le c_2 e^{R^{\varepsilon/2}}, \quad c_1\frac{\partial e^{x^{\varepsilon/2}}}{\partial x}\bigg|_{x=R} \le \frac{\partial g(x)}{\partial x}\bigg|_{x=R} \le c_2\frac{\partial e^{x^{\varepsilon/2}}}{\partial x}\bigg|_{x=R}.$$

It is easy to see from the equation for g that $c_1 e^{r^{\varepsilon/2}}$ and $c_2 e^{r^{\varepsilon/2}}$ are lower and upper bounds, respectively, for $g(r)$, $r > R$, and that $g'(r)$ is likewise bounded below and above by $(c_1 e^{r^{\varepsilon/2}})'$ and $(c_2 e^{r^{\varepsilon/2}})'$, $r > R$. A computation shows that, for z a geodesic perpendicular to dr,

$$\frac{d^2 r(z(s))}{ds^2}\bigg|_{s=0} = \frac{g'(r)}{g(r)} \ge \frac{c_1}{c_2}\frac{\varepsilon}{2}r^{\varepsilon/2-1}, \quad r > R.$$

Thus, by the Hessian comparison theorem, we have for the manifold N,

$$\frac{d^2 r(z^\alpha(s))}{ds^2}\bigg|_{s=0} \ge C' r^{\varepsilon/2-1}, \quad r > R.$$

This, together with the formula for $d\rho_t$, completes the proof of Lemma 1.

Now we will find the speed at which ρ_t goes to ∞.

LEMMA 3. $\liminf_{t\to\infty} \frac{\rho_t}{t^{2/(4-\varepsilon)}} > c_3 > 0$.

PROOF. The event

$$\left\{\liminf_{t\to\infty}\frac{\rho_t}{t^{2/(4-\varepsilon)}} > c_3\right\}$$

belongs to the tail σ-field of Brownian motion on M, which is trivial by assumption. Therefore this event has probability 0 or 1, and it suffices to show that it holds with positive probability.

Consider the process M_t defined by

$$dM_t = a(X_{\sigma(t)})\,dB_t, \quad M_0 = \rho_0,$$

and note that M_t is ρ_t with drift subtracted. By the law of the iterated logarithm, there exists a random time $T < \infty$ such that $t > T$ implies

$$|M_t| \le 2\sqrt{t \log\log t} \le t^{2/(4-\varepsilon)}.$$

Let $D_t = \rho_t - M_t$, so that $dD_t = b(X_{\sigma(t)})\,dt$. D_t is the drift of ρ_t.

We wish to compare D_t with $D_t^*(t \ge T)$, where D_t^* is defined by

$$dD_t^* = C[D_t^* + (t-T)^{2/(4-\varepsilon)}]^{\varepsilon/2-1}\,dt, \quad D_T^* = 0$$

for $t \ge T$. Note that, for the appropriate constant c_1,

$$D_t^* = c_1(t-T)^{2/(4-\varepsilon)}.$$

For $t \ge T$ and $\rho_t > R$, we have by Lemma 1 that

$$b(X_{\sigma(t)}) \ge C\rho_t^{\varepsilon/2-1} \ge C[D_t + (t-T)^{2/(4-\varepsilon)}]^{\varepsilon/2-1}.$$

Let $T^* \ge T$ be a function of T, and let A_{T^*} be the set of paths for which $T < T^*$ and for which $\rho_t > R$ provided $T \le t \le T^*$. Clearly, $P(A_{T^*}) > 0$. From the above inequality on $b(X_{\sigma(t)})$, we see that on A_{T^*}, $D_t^* \le D_t$ for $T \le t < T^*$. Therefore, if $\rho_s > R$ for $T \le s \le t$, then

$$\rho_t = D_t + M_t \ge D_t^* + M_t \ge c_1(t-T)^{2/4-\varepsilon} - 2\sqrt{t\log\log t}.$$

Now let T^* be the supremum of the times t for which the above expression is $\le R$. Then, on the set A_{T^*},

$$\rho_t \ge c_1 t^{2/(4-\varepsilon)} - 2\sqrt{t\log\log t} > c_3 t^{2/(4-\varepsilon)}$$

for an appropriate constant c_3 and all sufficiently large t. Since $P(A_{T^*}) > 0$, this establishes Lemma 3.

LEMMA 4. *Let τ_Q be the first time that $\rho_t \ge Q$. Then, for h sufficiently large, there is a $q \in (0, 1)$ and a constant C_1 such that*

$$P\{\rho_{t+\tau_Q} > C_1(t+h)^{2/(4-\varepsilon)} \text{ for } t > 0 \,|\, X_{\sigma(\tau_Q)}\} > q$$

uniformly in $X_{\sigma(\tau_Q)}$.

PROOF. Since $\limsup_{t\to\infty}(|B(t)|/2(t\log\log t)^{1/2}) = 1$ with probability 1, it easily follows that given of < 1, we can find h so large that

$$P\{|B(t)| \le 2\sqrt{(t+h)\log\log(t+h)}, \text{all } t > 0\} \ge q.$$

Let $dM_t = a(X_{\sigma(\tau_Q+t)})\,dB_{\tau_Q+t}$, $M_0 = Q$. Recall that $|a(X_{\sigma(t)})| \le 1$, so that if A is the event that $|M(t)| \le 2\sqrt{(t+h)\log\log(t+h)}$, all $t > 0$, then $P(A) \ge q$. Choose Q so large that for $\rho_t > Q/2$, $b(X_{\sigma(t)}) \ge C\rho_t^{\varepsilon/2-1}$. Let $D_t = \rho_{t+\tau_Q} - M_t$, so $D_0 = 0$. Then, for $\rho_{\tau_Q+t} > Q$ on the set A,

$$b(X_{\sigma(\tau_Q+t)}) \ge C[D_t + (t+2h)^{2/(4-\varepsilon)}]^{\varepsilon/2-1}.$$

Now, we repeat the reasoning of Lemma 3. Let $d\tilde{D}_t = C[\tilde{D}_t + (t+2h)^{2/(4-\varepsilon)}]^{\varepsilon/2-1}dt$. Then, one solution is $\tilde{D}_t = C_1(t+h)^{2/(4-\varepsilon)}$. Let Q be so large that $2(C_1+4)h^{2/(4-\varepsilon)} \le Q$. Then, by the formula for $d\tilde{D}_t$ it follows that $\tilde{D}_t \le D_t$. Thus, on the set A

$$\begin{aligned}\rho_{t+\tau_Q} &= D_t + M_t \ge \tilde{D}_t + M_t \\ &\ge C_1(t+2h)^{2/(4-\varepsilon)} - 2\sqrt{(t+h)\log\log(t+h)} \ge \tfrac{1}{2}C_1(t+h)^{2/(4-\varepsilon)}\end{aligned}$$

if h is large enough. Since $P(A) \ge q$, this proves the lemma.

LEMMA 5. *Let* $r = \inf(r(x), r(y))$ *for* $x, y \in N$. *If* $|\Theta(x, y)|$ *denotes the angular distance between* $\theta(x)$ *and* $\theta(y)$, *then for some constants* C *and* R, $r > R$

$$|\Theta(x, y)| \leq C\frac{{}^N d(x, y)}{\exp(r^{\varepsilon/2})},$$

where ${}^N d$ *denotes the distance in the manifold* N.

PROOF. We use the Rauch comparison theorem, for the same manifold P as used in the proof of Lemma 1. An easy computation shows that the formula holds for P, so it must hold for N.

LEMMA 6. *There exists a random time* T *such that for all* $m > T$,

$$\sup_{\sigma(m) \leq t \leq \sigma(m+1)} {}^N d(F(X_t), F(X_{\sigma(m)})) \leq m.$$

PROOF. Using the language of Lemma 2, we must show, for $m > T$, that

$$\sup_{m \leq t \leq m+1} \rho_t^{(m)} \leq m.$$

The argument is similar to an idea of Prat [13], and is also used by Kendall [9] and Pinsky [12]. We note by Lemma 2, for $|\rho_t| > 1$, say, that the drift of $\rho_t^{(m)}$ is bounded above by a constant K. Also, in order that $\rho_t^{(m)}$ hit m, it must first hit $2m/3$. Let τ be the first time after m that it does so. Then,

$$P\{\sup_{m \leq t \leq m+1} \rho_t^{(m)} \geq m\} \leq P\{\sup_{\tau \leq t \leq m+1} \rho_t^{(m)} \geq m\}$$
$$\leq P\left\{\sup_{\tau \leq t \leq m+1} \rho_t^{(m)} \geq m,\ \inf_{\tau \leq t \leq m+1} \rho_t^{(m)} \leq \frac{m}{3}\right\}.$$

Note that the total drift $\int_m^t b^{(m)}(X_{\sigma(t)})\, dt$ of $\rho_t^{(m)}$ is bounded by a constant C if $\tau \leq t \leq m + 1$ and if $\rho_t^{(m)} \geq m/3$. Thus, by Lemma 2, $|a^{(m)}(X_{\sigma(t)})| \leq 1$, and the above probability is bounded by

$$P\left\{\sup_{0 \leq t \leq 1} |B_t| \geq \frac{m}{3}\right\} \leq 4P\left\{B(1) \geq \frac{m}{3} - C_0\right\} \leq C \exp\left(-\frac{1}{2}\left(\frac{m}{3} - C_0\right)^2\right)$$

for some constants, C, C_0. This probability sums over m, so by the Borel-Cantelli lemma, $\sup_{m \leq t \leq m+1} \rho_t^{(m)} \geq m$ occurs only finitely often.

LEMMA 7. *Fix* $\delta > 0$. *If* τ *is a stopping time, we may choose an integer* J *so large that*

$$P\{\sup_{\tau \leq t-m \leq \tau+1} \rho_t^{(\tau+m)} \leq m + J, \text{ for all } m \geq 1\} \geq 1 - \delta.$$

PROOF. Let A_m be the event that $\sup_{\tau \leq t-m \leq \tau+1} \rho_t^{(\tau+m)} \geq m + J$. By the strong Markov property, $\{A_m\}$ are independent events, and so using the same estimate as in Lemma 6,

$$P\{A_m\} \leq C \exp\left(-\frac{1}{2}\left(\frac{m+J}{3} - K\right)^2\right).$$

If J is large enough, $\sum P(A_m) \leq \delta$. Then, by Kroenecker's lemma,

$$P\{\text{no } A_m \text{ occurs}\} = \prod (1 - P(A_m)) \geq 1 - \delta.$$

We now show that

$$\theta(\omega) = \lim_{t \to \infty} \theta(F(X_t))$$

exists. First, we show that $\lim_{k\to\infty}\theta(F(X_{\sigma(k)}))$ exists. Lemmas 2, 3 and 5 show that for k large enough

$$|\Theta(F(X_{\sigma(k)}), F(X_{\sigma(k+1)}))| \leq C\frac{k}{\exp((c'k)^{\varepsilon/(4-\varepsilon)})}.$$

This sums over k, so the limit must exist. By Lemmas 6 and 3,

$$\sup_{k\leq t\leq k+1}|\Theta(F(X_{\sigma(k)}), F(X_{\sigma(t)}))| \leq C\frac{k}{\exp((c'k)^{\varepsilon/(4-\varepsilon)})} \to 0,$$

for sufficiently large k. Thus, $\theta(\omega)$ exists. Now, we need to show that $\theta(\omega)$ is not trivial. This argument is due to Kendall. By a theorem of Stroock and Varadhan [15, Theorem 3.1], $\theta(F(X_{\sigma(\tau_Q)}))$ has positive probability of hitting any open set in the parameter space of θ. Choose q, δ in Lemmas 4 and 7 such that $q-\delta>0$. Fix Δ, and choose h so large that

$$\sum_{m=1}^{\infty} C\frac{m+J}{\exp((c_0(m+h)^{2/(4-\varepsilon)})^{\varepsilon})} < \Delta.$$

Lemmas 4, 5, and 7 show that, conditioned on $X_{\sigma(\tau_Q)}$, with probability at least $q-\delta$, all of the following inequalities hold:

$$\sup_{\tau_Q+m<t\leq\tau_Q+m+1}|\Theta(F(X_{\sigma(\tau_Q+m)}), F(X_{\sigma(t)}))| \leq C\frac{m+J}{\exp((c_0(t+h)^{2/(4-\varepsilon)})^{\varepsilon/2})},$$

$m=1, 2, \cdots$. Thus, conditioned on $X_{\sigma(\tau_Q)}$, with probability at least $q-\delta$, $|\Theta(F(X_{\sigma(\tau_Q)}), F(X_t))| < \Delta$ for all $t>\sigma(\tau_Q)$. Now, fix θ_0, and let $p=P\{|\Theta(F(X_{\sigma(\tau_Q)}), \theta_0)|<\Delta\}$. From what we have said, $p>0$ and $P\{|\theta_0-\theta(\omega)|<2\Delta\}\geq p(q-\delta)>0$. Since θ_0 and Δ are arbitrary, it follows that $\theta(\omega)$ is nontrivial. However, $\theta(\omega)$ is the tail σ-field of X_t, which was assumed to be trivial. This is a contradiction, so F must be constant.

Part II

3. Definitions and statement of the theorem. Let M and N be compact Riemann surfaces with a finite number of points deleted. It is a standard fact that any such surface N is homeomorphic to a sphere with $t(N)$ tori attached and $p(N)$ points deleted. Let $n(N) = 2t(N)+p(N)$.

THEOREM 2. *Let $F:M\to N$ be a holomorphic map. If $t(M)>0, p(M)>0, p(N)>1$, and $n(N)-n(M)>0$, then F is constant.*

4. Proof of Theorem 2. The proof uses Brownian motion on Riemann surfaces, which can be defined modulo time changes as follows. Let $\{m_i\}$ be coordinate patches exhausting M, and for each i, let $c_i: m_i\to\mathbb{C}$, where $\mathbb{C}$ is the complex plane, be a holomorphic map. For $p\in M$, choose a patch $m_{i(1)}$ containing p. Let $W(t)$ be Brownian motion in $\mathbb{C}$, and let σ_1 be the first time that $p+W(t)$ hits the boundary of $c_{i(1)}(m_{i(1)})$. For $0\leq t\leq\sigma_1$, let $B(t)=c_{i(1)}^{-1}(p+W(t))$. Next, let $m_{i(2)}$ be a patch containing $B(\sigma_1)$, and let σ_2 be the first time that $c_{i(2)}(B(\sigma_1))+W(t)-W(\sigma_1)$ hits the boundary of $m_{i(2)}$. For $\sigma_1\leq t\leq\sigma_2$, let

$$B(t)=c_{i(2)}^{-1}(c_{i(2)}(B(\sigma_1))+W(t)-W(\sigma_1)),$$

and note that we have defined $B(t)$ to be continuous at σ_1. For $k>2$, define σ_k and $B(t)$, $\sigma_k\leq t\leq\sigma_{k+1}$, analogously.

Note that $B(t)$ is defined only modulo time changes; different coordinate patches lead to different time scales.

We shall require the following theorem by Lévy [2].

THEOREM 3. *If f is a holomorphic function with domain Ω, then, until $W(t)$ leaves*

Ω, the process $f(W(t))$ has the same distribution as $W(p(t)) + f(0)$, where

$$p(t) = \int_0^t |f'(W(s))|^2\, ds$$

is a new time scale.

COROLLARY. *If h is a holomorphic function on M, then the process $h(B(t))$ has the same distribution as $W(p(t)) + h(B(0))$, where $p(t)$ is a new time scale.*

PROOF OF THE COROLLARY. Using the definition of $B(t)$, it follows that on each time interval $\sigma_k \le t \le \sigma_{k+1}$, $h(B(t))$ is a holomorphic function of $W(t)$. This allows us to apply Lévy's theorem.

We wish to study the winding properties of Brownian motion on M and N. To this end, we compute the fundamental groups of M and N.

LEMMA 8. *Let $p(N) > 0$. Then the fundamental group of N is the free group with $n(N) - 1$ generators. The generators may be taken to be the cycles around single points (except one) and the canonical generators of the tori.*

PROOF. Easy topology shows that a torus with one point deleted is homotopic to a sphere with 3 points deleted, and the generators of the torus correspond, under the homotopy, to cycles around 2 of the points. Therefore, a sphere N with $t(N)$ tori attached and $p(N) > 0$ points deleted is homotopic to a sphere N' with $t(N) - 1$ tori attached and $p(N) + 2$ points deleted. If the original generators $\{\alpha_i\}$ were the canonical generators of the tori plus cycles around every point except one, then $\{\alpha_i\}$ corresponds, under the homotopy, to a similar set of generators $\{\alpha_i'\}$ on N'. Thus, by induction, N is homotopic to a sphere N'' with $2t(N) + p(N)$ points deleted, and $\{\alpha_i\}$ corresponds to a set $\{\alpha_i''\}$ consisting of cycles around every point except one. In this case, it is standard that the fundamental group of N'' is the free group with $2t(N) + p(N) - 1$ generators $\{\alpha_i''\}$. This proves the lemma.

Now assume the conditions of Theorem 2, and suppose that F is not constant. Let $G(M)$ be the fundamental group of M with respect to the base point p, and let $G(N)$ be the fundamental group of N with respect to the base point $F(p)$. Let $\{\alpha_i^M\}$, $\{\alpha_i^N\}$ be the generators of $G(M)$, $G(N)$, respectively, as described in Lemma 8.

Choose 2 generators α_1^M, α_2^M of $G(M)$ which are generators of a torus, and let $\alpha_3^M, \cdots, \alpha_{n(M)-1}^M$ be the remaining generators. Let $H(M)$ be the smallest normal subgroup of $G(M)$ containing $\alpha_3^M, \cdots, \alpha_{n(M)-1}^M$ and the commutator $[G(M), G(M)]$. Let $\hat{G}(M) = G(M)/H(M)$, and note that $\hat{G}(M)$ is isomorphic to $\mathbb{Z} \times \mathbb{Z}$.

We will now use Brownian motion on M to construct a recurrent process X_i taking values in $\hat{G}(M)$. After constructing a homomorphic image $\hat{G}(N)$ of the fundamental group of N, we will define a transient process Y_i taking values in $\hat{G}(N)$. We will show, however, that the converse of Theorem 2 implies that Y_i is a function of X_i. This leads to a contradiction which establishes the theorem.

Recall that p is contained in the patch $m_{i(1)}$. Let $\mathcal{O}$, $\mathcal{O}_2$, $\mathcal{O}_3$ be neighborhoods of p, contained in $m_{i(1)}$, whose images under $c_{i(1)}$ are discs of radii ε, 2ε, 3ε, respectively. We need the following theorem of Harris [8] which deals with Markov chains x_n with state space X.

THEOREM 4. *Suppose the following conditions hold:*

(1) *There exists a σ-finite measure m on X such that if $m(E) > 0$, then for all $x_0 \in X$, $P\{x_n \in E \text{ infinitely often} \mid x_0\} = 1$;*

(2) *We can find $A \subset X$ such that $0 < m(A) < \infty$, and a constant $C > 0$ for which*

$$P\{x_1 \in dy \mid x_0 = x\} \ge Cm(dy)$$

for all $x, y \in A$. Assume also, that if ρ is the recurrence time to A, that $E(\rho \mid x_0 = x)$ is bounded independently of x.

Then, x_n has a unique invariant probability measure.

Now, let $B^Q(t)$ be the Brownian motion on M whose initial distribution $B^Q(0)$ is the measure Q, concentrated on the boundry of $\mathcal{O}$. Let $\tau_0 = 0$; given τ_i, let τ_{i+1} be the first time after τ_i for which $B(\tau_{i+1}) \in \partial\mathcal{O}$, and $\{B(t): \tau_i \le t \le \tau_{i+1}\}$ corresponds to a nontrivial element of $\hat{G}(M)$. Note that since M is compact except for deleted points, all τ_i are finite.

Next, let $x_n = B^Q(\tau_n)$, and note that x_n is a Markov process with respect to the fields $\mathcal{F}_n = \sigma\{B(t): t \le \tau_n\}$. Q is said to be an invariant measure for x_n if, for all $E \in \partial\mathcal{O}$,

$$Q(E) = \int_{x\in\partial\mathcal{O}} Q(dx)P\{x_1 \in E \mid x_0 = x\}.$$

Let m be the measure on $\partial\mathcal{O}$ induced by Lebesgue measure on the circle $c_{i(1)}(\partial\mathcal{O})$.

We will show that $x_n = B^Q(\tau_n)$ satisfies the conditions of Theorem 4 for all Q. Let $\tau_{i+1}^{(2)}$ be the first time after τ_i that $B^Q(t)$ hits $\partial\mathcal{O}_2$, and $\{B^Q(t): \tau_i \le t \le \tau_{i+1}^{(2)}\}$ corresponds to a nontrivial element of $\hat{G}(M)$; let $\tau_{i+1}^{(3)}$ be the first time after $\tau_{i+1}^{(2)}$ that $B^Q(t)$ hits $\partial\mathcal{O}$ or $\partial\mathcal{O}_3$. Then, for any $x_0 \in \partial\mathcal{O}$,

$$\begin{aligned} P\{B^Q(\tau_n) \in E \mid B^Q(0) = x_0\} &\ge P\{B^Q(\tau_n^{(3)}) \in E \mid B^Q(0) = x_0\} \\ &\ge \inf_{x\in\partial\mathcal{O}_2} P\{B^Q(\tau_n^{(3)}) \in E \mid B(\tau_n^{(2)}) = x\} \\ &= \inf_{x\in c_{i(1)}(\partial\mathcal{O}_2)} P\{W(\tau_n^{(3)}) \in c_{i(1)}(E) \mid W(\tau_n^{(2)}) = x\}. \end{aligned}$$

Recalling that $c_{i(1)}(\mathcal{O}_j)$ are discs with radii in the ratios 1:2:3, it follows from standard probability theory that the above infimum is bounded below by $\delta m(E)$, say. Thus, by the strong Markov property,

$$\begin{aligned} P\{B^Q(\tau_n) \notin E \text{ for } n \ge K\} &\le \textstyle\prod_{n=K+1}^{\infty} \sup_{x\in\partial\mathcal{O}} P\{B^Q(\tau_n) \notin E \mid B^Q(\tau_{n-1}) = x\} \\ &\le \textstyle\prod_{n=K+1}^{\infty} (1 - \delta m(E)) = 0. \end{aligned}$$

Thus, condition (1) holds.

As for condition 2, let $A = \partial\mathcal{O}$, and note that by the previous reasoning

$$\begin{aligned} P\{B^Q(\tau_1) \in dy \mid B_Q(0) = x\} &\ge \inf_{x\in\partial\mathcal{O}_2} P\{B^Q(\tau_1^{(3)}) \in dy \mid B^Q(\tau_n^{(2)}) = x\} \\ &\ge \delta\, m(dy). \end{aligned}$$

Now, by definition, the recurrence time to A is 1, and so condition (2) is satisfied. Therefore, x_n has an invariant measure, so we may set Q equal to this measure.

Let $\hat{X}_i$ be the element of $\hat{G}(M)$ corresponding to $\{B^Q(t): 0 \le t \le \tau_i\}$. Since $\hat{G}(M) \simeq \mathbb{Z} \times \mathbb{Z}$, we may regard $\hat{X}_i$ as a process on $\mathbb{Z} \times \mathbb{Z}$.

LEMMA 9. *Let $\Delta_i = \hat{X}_i - \hat{X}_{i-1}$. Then $\{\Delta_i\}$ is a stationary process with $E|\Delta_i|^4 < \infty$. Furthermore, the process is symmetric, so that Δ_i and $-\Delta_i$ have the same distribution.*

PROOF. We first prove symmetry. Let $\bar{\tau}_i$ be the last time between τ_i and τ_{i+1}, properly, that $B^Q(t) \in \partial\mathcal{O}$. By a standard time reversal argument, $\{B^Q(\bar{\tau}_{i+1} - t): 0 \le t \le \bar{\tau}_{i+1} - \bar{\tau}_i\}$ has the same distribution $\{B^Q(\tau_i + t): 0 \le t \le \tau_{i+1} - \tau_i\}$. However, these two paths correspond to elements of $\hat{G}(M)$ which are inverses of each other. This verifies symmetry.

We now prove that $E\Delta_i^4 < \infty$ by using the fact that the generators of $\hat{G}(M)$ derive from a torus. Let l_1, l_2 be smooth paths representing α_1^M, α_2^M, respectively. Note that, since $\{B^Q(t): \tau_i \le t \le \tau_{i+1}\}$ is not homotopic to 0, it must cross l_1 or l_2 at least once, transversally. Define a sequence of stopping times $\hat{\tau}_i^{(n)}$ such that $\hat{\tau}_i^{(0)} = \tau_i$, and $\hat{\tau}_i^{(n+1)}$ is the first time after $\hat{\tau}_i^{(n)}$ that $\alpha = \{B(t): \hat{\tau}_i^{(n)} \le t \le \hat{\tau}_i^{(n+1)}\}$ intersects either l_1 or l_2 twice, transversally. As can be easily seen from the fact that l_1, l_2 correspond to the generators of a torus, $P\{\hat{\tau}_i^{(n+1)} < \tau_{i+1} \mid \hat{\tau}_i^{(n)} < \tau_{i+1}\}$ is bounded above by a constant, say $1 - \delta$, which is less than 1. Therefore, by the strong Markov property, $P\{\hat{\tau}_i^{(n)} < \tau_{i+1}\} \le (1 - \delta)^n$. Let n_i be the first integer such that $\hat{\tau}_i^{(n)} > \tau_{i+1}$, and note that $|\hat{X}_{i+1} - \hat{X}_i| < 2n_i$. By the above inequality, $P\{n_i > K\} \le (1 - \delta)^k$, so $E|\Delta_i|^4 < \infty$.

It remains to show that $\{\Delta_i\}$ is stationary. But this follows from the strong Markov property of Brownian motion, and the fact that $B^Q(\tau_i)$ has the same distribution Q for all i.

We say that $\hat{X}_i$ is *recurrent* if it returns to 0 infinitely often, with probability 1.

LEMMA 10. *$\hat{X}_i$ is recurrent.*

PROOF. If the variables Δ_i were independent, it would easily follow that $\hat{X}_i$ was recurrent. In our case, we need a central limit theorem for stationary processes, such as Theorem 9 of Phillip [11]. Although his theorem is not stated in terms of vector-valued processes, inspection shows that his arguments carry through without change in this case, yielding the following theorem.

THEOREM 5. *Let $\mathscr{M}_{k,l}$, $k, l \in Z \cup \{\pm\infty\}$ be the σ-field generated by $\{\Delta_i : k < i < l\}$. Suppose that $\{\Delta_i\}$ is a stationary process with $E\Delta_i = 0$ and $E|\Delta_i|^2 < \infty$. Assume that the following mixing condition holds*:

$$\sup_k \sup_{A \in \mathscr{M}_{k+l,\infty}} |P\{A \mid \mathscr{M}_{0,k}\} - P\{A\}| \le \varphi(l)$$

and $\sum_l \varphi^{1/5}(l) < \infty$. Then

$$\sigma_{ij}^2 = E(\Delta_1^i \Delta_1^j) + \sum_{\nu=1}^{\infty} E(\Delta_1^i \Delta_{\nu+1}^j + \Delta_1^j \Delta_{\nu+1}^i)$$

exists, where the superscripts on Δ refer to vector components. Moreover, if σ^2 is positive definite, then

$$\mathscr{L}\left(\frac{1}{N^{1/2}} \sum_{k=1}^N \Delta_k\right) \to N(0, \sigma^2).$$

By Lemma 9, $E|\Delta_i|^4 < \infty$ and the variables Δ_i are symmetric, so it also follows that $E\Delta_i = 0$.

Let us verify the mixing condition. We will use the notation of Theorem 4, so that $\mathscr{O} \subset \mathscr{O}_2$ are circular neighborhoods of p (with respect to some coordinate patch). Let $\mathscr{O}_2$ be fixed and choose $\mathscr{O}$ to be small, as required in what follows. Let σ be the first time that B hits $\mathscr{O}_2$. Recall that the τ_i are the stopping times used to define $\hat{X}_i$, and note that $\sigma < \tau_i$, almost surely. If S is the initial distribution of B on $\partial\mathscr{O}$, let $\tilde{S}$ be the distribution of $B(\sigma)$ on $\partial\mathscr{O}_2$. Thus, $B^Q(\sigma)$ has distribution $\tilde{Q}$. Standard estimates using the Poisson kernel show that if $\varepsilon > 0$, we may choose $\mathscr{O}$ so small that for all initial distributions S on $\partial\mathscr{O}$,

$$\tilde{S}(dx) \ge (1 - \varepsilon)\tilde{Q}(dx).$$

Now let $S^{(n)}$ be the distribution of $B(\tau_n)$, given that $B(0)$ has distribution S. Since $\tilde{Q}$ leads to the stationary distribution Q, we have shown that, for all S,

$$S^{(1)}(dx) \ge (1 - \varepsilon)Q(dx)$$

and, therefore, by the strong Markov property,

$$S^{(n)}(dx) \ge (1 - \varepsilon^n)Q(dx).$$

Also, if $A \in \mathscr{M}_{k+l,\infty}$, then by the strong Markov property, $P(A \mid \mathscr{M}_{0,k})$ depends only on the distribution of $B(\tau_{k+l})$ given $B(\tau_{k-1})$. By the above, this distribution is greater than or equal to $(1 - \varepsilon^l)Q$. Thus, the mixing condition holds with $\varphi(l) = \varepsilon^l$.

Next, we must show that

$$\sigma_{ij}^2 = E(\Delta_1^i \Delta_1^j) + \sum_{\nu=1}^{\infty} E(\Delta_1^i \Delta_{\nu+1}^j + \Delta_1^j \Delta_{\nu+1}^i)$$

is positive definite. Let $Q_{x_1}^{(k)}$ be the distribution of $B(\tau_k)$ conditioned on X_1. We claim that for each $\varepsilon > 0$, $\mathscr{O}$ can be chosen so small that

$$(*) \qquad Q_{x_1}^{(k)} = (1 - \varepsilon^k)Q + \varepsilon^k \hat{Q}_{x_1}^{(k)}$$

where $\hat{Q}_{x_1}^{(k)}$ is a probability distribution depending on X_1. This would follow as in the previous argument, if we could establish the equality for $k = 1$. Let $\mathcal{O} \subset \mathcal{O}_2 \subset \mathcal{O}_3$ as before, with $\mathcal{O}_2$, $\mathcal{O}_3$ fixed and $\mathcal{O}$ to be chosen sufficiently small. Let $\sigma_0^{(2)} = \sigma_0^{(3)} = 0$, let $\sigma_{i+1}^{(2)}$ be the first time after $\sigma_i^{(3)}$ that $B(t)$ hits $\partial \mathcal{O}_2$. Let $\sigma_i^{(3)}$ be the first time after $\sigma_i^{(2)}$ that $B(t)$ hits $\partial \mathcal{O}_3$ or $\partial \mathcal{O}$. By standard probability theory, if $\mathcal{O}$ is a sufficiently small disc, and if $S_x^{(k)}$ is the distribution of $B(\tau_k^{(3)})$ given that $B(\tau_k^{(2)}) = x$ and $B(\tau_k^{(3)}) \in \partial \mathcal{O}$, then

$$S_x^{(k)} = \left(1 - \frac{\varepsilon}{2}\right) U + \frac{\varepsilon}{2} \hat{S}_x^{(k)}$$

where U is the uniform distribution on $\partial \mathcal{O}$ and $\hat{S}_x^{(k)}$ is a probability distribution depending on x. Since both Q and $\hat{Q}_{x_1}^{(k)}$ are mixtures of the distributions $S_x^{(k)}$, equation $(*)$ follows for $k = 1$, and thus for all k.

Using $(*)$, we may define events A_k such that, conditioned on X_1, the distribution of $B(\tau_k)1(A_k)$ is $(1 - \varepsilon^k)Q$ and the distribution of $B(\tau_k)1(A_k^c)$ is $\varepsilon^k \hat{Q}_{x_1}^{(k)}$. The choice of A_k may require the use of auxiliary random variables independent of $\{B(t)\}$. We also require that A_k be independent of $\{\Delta_i : i > k\}$. This insures that, conditioned on Δ_1 and $A_k (A_k^c$, respectively), Δ_{k+1} has the same distribution as if $B(\tau_k)$ had the distribution $Q (Q_{x_1}^{(k)}$, respectively). Now let

$$Y_{k+1} = \Delta_{k+1} 1(A_k), \; Z_{k+1} = \Delta_{k+1} 1(A_k^c).$$

By the above, $E[Y_{k+1} \mid \Delta_1] = (E\Delta_{k+1})P(A_k) = 0$. We also wish to show that $EZ_{k+1}^2 < C\varepsilon^k$ for some constant C. Since $P(A_k^c) = \varepsilon^k$, this would follow if we could show, in particular, that

$$E[\Delta_{k+1}^4 \mid B(\tau_k) = x] < C$$

uniformly in x. But in Lemma 9, the proof that $E\Delta_1^4 < \infty$ did not depend on the starting point $B(0)$, so the inequality holds. Now $|E\Delta_1^i \Delta_{k+1}^j| \le E|\Delta_1^i Y_{k+1}^j| + E|\Delta_1^i Z_{k+1}^j|$ where the superscripts refer to vector components. The first term on the right is 0, so applying the Cauchy-Schwartz inequality to the second term we get

$$E|\Delta_1^i Z_{k+1}^j| \le (E|\Delta_1^i|^2 E|Y_{k+1}^j|^2)^{1/2} \le C\varepsilon^{k/2}.$$

Thus, in the equation

$$\sigma_{ij}^2 = E\Delta_1^i \Delta_1^j + \textstyle\sum_{\nu=1}^{\infty} E(\Delta_1^i \Delta_{\nu+1}^j + \Delta_1^j \Delta_{\nu+1}^i),$$

the second term is bounded in absolute value by a constant times $\varepsilon^{1/2}$.

It is easy to see that Δ_1 is not restricted to a subspace of $\mathbb{R}^2$, and thus $E\Delta_1^i \Delta_1^j$ is a positive definite matrix. A simple but tedious calculation of the characteristic equation then shows that if ε is small enough, σ_{ij}^2 must be positive definite.

Next, let $u_n = P\{\Delta_1 + \cdots + \Delta_k = 0\}$. We intend to show that $u_{2k} > Ck^{-1}$, so that the u_k sum to infinity. First, we need the following estimate. Let $Q_x^{(k)}$ be the distribution of $B(\tau_k)$ given that $\hat{X}_k = x$. Then by previous arguments, $Q_x^{(k)} > (1 - \varepsilon)Q$ for all x, so by the strong Markov property,

$$\begin{aligned} P\{\hat{X}_{2k} = 0 \mid \hat{X}_k = x\} &= P\{\hat{X}_k = -x \mid B(0) \text{ has distribution } Q_x^{(k)}\} \\ &\ge (1 - \varepsilon)P(\hat{X}_k = -x) = (1 - \varepsilon)P(\hat{X}_k = x) \end{aligned}$$

since $\hat{X}_k$ is symmetric. Thus,

$$u_{2k} = \textstyle\sum_{x \in \mathbb{Z}^2} P\{\hat{X}_{2k} = 0 \mid \hat{X}_k = x\} P\{\hat{X}_k = x\} \ge (1 - \varepsilon) \sum_{x \in \mathbb{Z}^2} P\{\hat{X}_k = x\}^2.$$

Now by Theorem 5, we know that for some constant C, and for k large enough,

$$(**) \qquad P\{|\hat{X}_k| < k^{1/2}\} \ge C.$$

By a well-known argument, the sum of the squares $\sum P\{\hat{X}_k = x\}^2$ is minimized subject to the constraint $(**)$ when each point $\{x \in \mathbb{Z}^2 : |x| < k^{1/2}\}$ has equal probability, proportional to $1/k$. Therefore, for some constant C,

$$u_{2k} \geq C \sum_{x \in \mathbb{Z}^2, |x| < k^{1/2}} \frac{1}{k^2} \geq \frac{1}{k},$$

so the u_k sum to infinity.

Let $f_{l-k|k}$ denote the probability, given that $\hat{X}_k = 0$, that $Y_i = \hat{X}_{k+i}$ returns to 0 for the first time at $i = l - k$. Thus, $f_{l-k|k} = P\{\hat{X}_{k+1} \neq 0, \cdots, \hat{X}_{l-1} \neq 0, \hat{X}_l = 0 \mid \hat{X}_k = 0\}$. Since $\hat{X}_0 = 0$, there must be a last time $k < l$ such that $\hat{X}_k = 0$, and so

$$u_l = \sum_{k=0}^{l-1} u_k f_{l-k|k} \quad \text{for} \quad l \geq 1, \quad u_0 = 1.$$

Adding these equations, we obtain

$$(***) \qquad \sum_{l=0}^{L} u_l = \sum_{l=1}^{L} \sum_{k=0}^{l-1} u_k f_{l-k|k} + 1 = \sum_{k=0}^{L} u_k \sum_{l=k+1}^{L} f_{l-k|k} + 1.$$

Suppose that $\hat{X}_i$ were not recurrent, so that if A is the event that $\hat{X}_i$ ever returns to 0, $P\{A\} < 1 - \delta$ for some $\delta > 0$. We claim that $P\{A \mid B(0)\} < (1 - \varepsilon)(1 - \delta) + \varepsilon < 1$. Indeed, if σ is the time (with respect to $B(t)$) that $\hat{X}_i$ first returns to 0, familiar arguments show that the distribution of $B(\sigma)$ is greater than $(1 - \varepsilon)Q$. The assertion then follows, and from the strong Markov property, we conclude that for all k

$$\sum_{l=k+1}^{\infty} f_{l-k|k} < (1 - \varepsilon)(1 - \delta) + \varepsilon < 1.$$

But since $\sum_{k=0}^{\infty} u_k = \infty$, this contradicts equation $(***)$. Thus $\hat{X}_i$ is recurrent, and Lemma 10 is proven.

Unfortunately, $\hat{X}_i$ is not suited to our purposes since it is induced by B^Q and not B. Let X_i be the process induced on $\hat{G}(M)$ by $B(t)$.

LEMMA 11. *X_i is recurrent.*

PROOF. Let $A, \hat{A}$ be the event of recurrence of $X, \hat{X}$, respectively. By Lemma 3.1 of [2], given $\varepsilon > 0$, we can choose $\mathcal{O}$ so small that

$$|P(A) - P(\hat{A})| \leq \varepsilon P(\hat{A}),$$

and so $P(A) \geq 1 - \varepsilon$. Since ε was arbitrary, $P(A) = 0$, and X_i is recurrent.

Recall that $H(M)$ is the smallest normal subgroup of $G(M)$ containing $\alpha_3^M, \cdots, \alpha_{n(M)-1}^M$ and $[G(M), G(M)]$. Now, F induces a map F_H from $H/[H, H]$ to $G(N)/[G(N), G(N)]$. Since these are commutative groups with difference in dimension at least 3, it follows that there must be at least 3 generators $\alpha_1^N, \alpha_2^N, \alpha_3^N$ of $G(N)$ which generate a subgroup of $G(N)/[G(N), G(N)]$ modulo $F_H(H/[H, H])$ isomorphic to $\mathbb{Z} \times \mathbb{Z} \times \mathbb{Z}$.

Let $H(N)$ be the smallest normal subgroup of $G(N)$ containing $\alpha_4^N, \cdots, \alpha_{n(N)-1}^N$ and $[G(N), G(N)]$, and let $\hat{G}(N) = G(N)/H(N)$. From what we have said, F induces a map $\hat{F}$ from $\hat{G}(M)$ to $\hat{G}(N)$.

Following the steps of the previous case, we could now construct a process $\hat{Y}_i$ from $F(B(t))$, and obtain an invariant measure Q on $F(\mathcal{O})$. Note that by Lévy's theorem, $F(B(t))$ is Brownian motion on N with a new time scale.

LEMMA 12. *$\hat{Y}_i$ is transient. That is, $\hat{Y}_i = 0$ only finitely often with probability* 1.

PROOF. Let $\Delta_i = \hat{Y}_i - \hat{Y}_{i-1}$. Now consider a path of length k. We will define events A_i, for $i = 0, \cdots, k$ by induction. First, let A_0 be the entire sample space. Now suppose that we have defined $A_0, \cdots, A_{k-1}$. As before, it is easily shown that if $\mathcal{O}$ is sufficiently small, then, conditioned on $A_0, \cdots, A_{k-1}$ and $\hat{Y}_1, \cdots, \hat{Y}_k$, the distribution of $B(\tau_k)$ is always greater than $(1 - \varepsilon)Q$. Expanding the probability space if necessary, we can choose A_k to be independent of $A_0, \cdots, A_{k-1}$ and $\hat{Y}_1, \cdots, \hat{Y}_k$ such that

(i) $P(A_k) = 1 - \varepsilon$,

(ii) The distribution of $F(B(\tau_k))$ given $A_0, \cdots, A_k$ and $\hat{Y}_1, \cdots, \hat{Y}_k$ is Q.

We define further random variables as follows. Our strategy is to single out those Δ_{2j} for which both A_{2j-1} and A_{2j} occur. By construction, these Δ_{2j} will then be independent of all other steps, given $A_{2j-1} \cap A_{2j}$. Let

$$S_k = \sum_{i \le k, i \text{ odd}} \Delta_i + \sum_{i \le k, i \text{ even}} \Delta_i 1((A_{i-1} \cap A_i)^c),$$

$$T_k = \sum_{i \le k, i \text{ even}} \Delta_i 1(A_{i-1} \cap A_i), \quad \text{and} \quad N_k = \sum_{i \le k, i \text{ even}} 1(A_{i-1} \cap A_i).$$

Thus, N_k is the number of events $A_{i-1} \cap A_i$ which occur. We claim that conditioned on N_k, S_k and T_k are independent. Indeed, conditioned on N_k and S_k, the distribution of T_k is the N_k-fold convolution of the distribution of Δ_1 given $A_0 \cap A_1$. Since this distribution does not depend on S_k, independence follows. It is easy to see that this distribution is not restricted to a 2-dimensional subspace of $\mathbb{Z}^3$. Therefore, by a theorem in Spitzer [14], page 72, $\sup_{x \in \mathbb{Z}^3} P\{T_k = x \mid N_k\} \le C\, N_k^{-3/2}$ for some constant C independent of N_k. Since S_k and T_k are independent given N_k, and since $\hat{Y}_k = S_k + T_k$, it follows that

$$P\{\hat{Y}_k = 0 \mid N_k\} \le C\, N_k^{-3/2}.$$

We show that $\sum_{k=0}^{\infty} P\{\hat{Y}_k = 0\} < \infty$. Note that $P\{N_k = i\} = \binom{k}{i}(1-\varepsilon)^i \varepsilon^{k-i}$, so that

$$\begin{aligned} P\{\hat{Y}_k = 0, N_k = i\} &= P\{Y_k = 0 \mid N_k = i\} P\{N_k = i\} \\ &\le C\, i^{-3/2} \binom{k}{i} (1-\varepsilon)^i \varepsilon^{k-i} \\ &\le C' i^{1/2} \frac{k!}{(i+2)!\,(k-i)!} (1-\varepsilon)^i \varepsilon^{k-i} \\ &\le C'' k^{-3/2} \binom{k+2}{i+2} (1-\varepsilon)^{i+2} \varepsilon^{k-i}. \end{aligned}$$

Summing over $i = 1, \cdots, k$, we obtain, after relabeling C,

$$P\{\hat{Y}_k = 0\} \le C k^{-3/2}((1-\varepsilon) + \varepsilon)^{k+2} = C k^{-3/2}.$$

Thus, $\sum_{k=0}^{\infty} P\{\hat{Y}_k = 0\} < \infty$. By the Borel-Cantelli lemma, $\hat{Y}_k$ is transient. This verifies Lemma 12.

Davis's agrument, as in Lemma 11, then shows that if Y_i is the process induced by $F(B(t))$ with $B(0) = p$, then Y_i is also transient. But then $\hat{F}(X_i) = Y_i$, and X_i is recurrent. This contradiction shows that F must be constant.

Added in proof. T. J. Lyons has communicated to us a short classical proof of Theorem 2.

REFERENCES

[1] Chern, S. S. (1960). Complex analytic mappings of Riemann surfaces. *I Amer. J. Math.* **82** 323–337.

[2] Davis, B. (1975). Picard's theorem and Brownian motion. *Trans. Amer. Math. Soc.* **213** 353–362.

[3] Eells, J. and Lemaire, L. (1978). A report on harmonic maps. *Bull. London Math. Soc.* **10** 1–68.

[4] Cheeger, J. and Ebin, D. G. (1975). *Comparison Theorems in Riemannian Geometry*. North Holland, Amsterdam.

[5] Greene, R. E. and Wu, H. (1979). Function theory on manifolds which possess a pole. *Lecture Notes in Math.* **699**. Springer Verlag, Berlin.

[6] Goldberg, S. I. and Har'El, Z. (1977). A general Schwarz lemma for Riemannian manifolds. *Bull. Greek Math. Soc.* **18** 141–148.

[7] Goldberg, S. I. and Ishihara, T. (1976). Harmonic quasiconformal mappings of Riemannian manifolds. *Amer. J. Math.* **98** 225–240.

[8] Harris, T. E. (1956). The existence of stationary measures for certain Markov processes. *Proc. of the 3rd Berkeley Symp. on Math. Statist. and Probab. Vol. II* 113–124.

[9] Kendall, W. S. (1983). Brownian motion and a generalized little Picard's theorem. *Trans.*

Amer. Math. Soc., **275** 751–760.

[10] McKEAN, H. P. JR. (1969). *Stochastic Integrals.* Academic, New York.

[11] PHILLIP, W. (1969). The central limit theorem for mixing sequences of random variables. *Z. Wahrsch. verw. Gebiete.* **12** 155–171.

[12] PINSKY, M. (1978). Stochastic Riemannian geometry. *Probabilistic Analysis and Related Topics.* Vol. I, A. T. Barucha-Reid (ed). Academic, New York.

[13] PRAT, J.-J. (1975). Etude asymptotique et convergence angulaire du mouvement brownien sur une variété à courbure negative. *C. R. Acad. Sci. Paris Sér. A-B* **280** 1539–1542.

[14] SPITZER, F. (1976). *Principles of Random Walk.* Springer Verlag, Berlin.

[15] STROOCK, D. W. and VARADHAN, S. R. S. (1970). On the support of diffusion processes with applications to the strong maximum principle. *Proc. of the 6th Berkeley Symp. on Math. Statist. and Probab., Vol. III*, 333–360.

[16] YAU, S.-T. (1975). Harmonic functions on complete Riemannian manifolds. *Comm. Pure Appl. Math.* **28** 201–228.

DEPARTMENT OF MATHEMATICS
1409 WEST GREEN STREET
UNIVERSITY OF ILLINOIS
URBANA, ILLINOIS 61801

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF TEXAS
AUSTIN, TEXAS 78712