Can. J. Math. Vol. **47** (1), 1995 pp. 3–27.

ON THE ATTRACTING ORBIT OF A NON-LINEAR TRANSFORMATION ARISING FROM RENORMALIZATION OF HIERARCHICALLY INTERACTING DIFFUSIONS PART I: THE COMPACT CASE

J. B. BAILLON, PH. CLÉMENT, A. GREVEN AND F. DEN HOLLANDER

ABSTRACT. This paper analyzes the n-fold composition of a certain non-linear integral operator acting on a class of functions on $[0,1]$. The attracting orbit is identified and various properties of convergence to this orbit are derived. The results imply that the space-time scaling limit of a certain infinite system of interacting diffusions has universal behavior independent of model parameters.

0. **Introduction and main results.** The present paper studies the iterates of a non-linear transformation F acting on a class of functions $g\colon [0,1] \longrightarrow [0,\infty)$. The problem arises in a probabilistic context, which is explained in Sections 0.2, 0.3 and 0.6. The rest of the paper focuses on the analytic aspects. The main results are formulated in Section 0.5. Sections 1–3 contain the proofs.

0.1. *The transformation.* Let $(\nu_\theta^g)_{\theta\in[0,1]}$ be the family of probability measures on $[0,1]$ given by

$$\nu_\theta^g(dx) = \frac{1}{Z_\theta^g}\frac{1}{g(x)}\exp\left[-\int_\theta^x \frac{y-\theta}{g(y)}\,dy\right]dx \quad \bigl(\theta\in(0,1)\bigr) \tag{0.1}$$

where Z_θ^g is the normalizing constant and g is any function satisfying

$$g(0) = g(1) = 0 \tag{0.2i}$$

$$g(x) > 0 \text{ for } x \in (0,1) \tag{0.2ii}$$

$$g \text{ is Lipschitz continuous on } [0,1]. \tag{0.2iii}$$

At the boundary points $\theta = 0$ and $\theta = 1$ set $\nu_0^g = \delta_0$ resp. $\nu_1^g = \delta_1$ (point measures). Define the transformation F acting on g by

$$(Fg)(\theta) = \int_0^1 g(x)\nu_\theta^g(dx) \quad (\theta\in[0,1]). \tag{0.3}$$

Our goal will be to identify the subclass of (0.2) for which

$$a_nF^ng \longrightarrow g^* \quad (n\longrightarrow\infty), \tag{0.4}$$

Received by the editors May 25, 1993.
AMS subject classification: Primary: 60K35; secondary: 47H99.

either pointwise or in a suitable norm, where a_n are normalizing constants (tending to infinity) *independent* of g, and the limit g^* will turn out to be

$$g^*(\theta) = \theta(1-\theta). \tag{0.5}$$

The probability measure $\nu_\theta^g(dx)$ is the equilibrium of a diffusion on $[0,1]$ with drift towards θ and with local diffusion coefficient $g(x)$. The value $(Fg)(\theta)$ is the average of $g(x)$ under $\nu_\theta^g(dx)$. Since $\nu_\theta^g(dx)$ itself depends on g the transformation F is non-linear.

0.2. *Motivation.* We shall first explain how the question (0.4) arises in a probabilistic context, in particular in connection with attempts to explain *universal behavior* of systems of interacting diffusions. The reader who is interested only in the analytic aspects of (0.4) may skip Sections 0.2, 0.3 and 0.6 (with the exception of (0.14–17)).

In the study of systems of interacting diffusions (typically countably many) one finds that certain properties of the long term behavior are universal in whole classes of diffusions (see [CG], [FG]). In order to understand this phenomenon one introduces *renormalization schemes* and one tries to carry out the following two-step program:

(I) Prove that block averages on successive space-time scales converge to a time-inhomogeneous Markov chain. The state space of this Markov chain is the same as that of the single-component diffusions, the time index labels the scales, and the transition kernels are given in terms of the iterates $\{F^n g : n \geq 0\}$ where g is the diffusion function of the single components and F is a transformation determined by the interaction (see Section 0.3).

(II) Prove that F has a unique attracting orbit. Identify the domain of attraction of this orbit for convergence either pointwise or in a suitable norm. Find the speed of convergence to the attractor (see Sections 0.4 and 0.5).

Those interacting diffusions whose components have a diffusion function g in the domain of attraction of g^* (see (0.4)) display a long term behavior that is dictated by the attractor and that therefore is universal.

Step I has been carried out for a number of systems that arise in population genetics: [DG1–3], [CGS], [DGV]. For more details on applications in this area, see [SF] and [S]. In this paper we embark on Step II by treating the transformation arising from the model in [DG3].

Universality is a theme that plays an important role in many areas. For a broad reference in the context of interacting particle systems, see [L], [D] and [G].

0.3. *Background of (0.1–3).* In order to give the reader some guidance we shall briefly describe the model under consideration and formulate the main result of [DG3] leading up to (0.1–3). Equation (0.7) below defines our system of interacting diffusions in a probabilistic language. At the end of this subsection we shall indicate how the system can be described in terms of generators and semigroups.

For integer $N \geq 1$ let Ω_N be the countable group of sequences

$$\Omega_N = \left\{\xi = (\xi_i)_{i\geq 1} : \xi_i \in \{0,1,\ldots,N-1\}, \xi_i \neq 0 \text{ finitely often}\right\} \tag{0.6}$$

with component-wise addition modulo N. Consider the Markov process $\left(X^N(t)\right)_{t\geq 0}$ with state space $[0,1]^{\Omega_N}$ defined to be the unique strong solution of the following system of stochastic differential equations (a typical element $x \in [0,1]^{\Omega_N}$ is written $x = (x_\xi)_{\xi\in\Omega_N}$):

$$\begin{aligned} dX_\xi^N(t) &= \sum_{\eta\in\Omega_N} q_N(\xi,\eta)[X_\eta^N(t) - X_\xi^N(t)]\,dt + \sqrt{2g\left(X_\xi^N(t)\right)}\,dW_\xi(t) \\ X^N(0) &= X^N. \end{aligned} \tag{0.7}$$

Here $\left\{\left(W_\xi(t)\right)_{t\geq 0} : \xi \in \Omega_N\right\}$ is a collection of independent standard Brownian motions, g is any diffusion function satisfying (0.2), and $q_N(\cdot,\cdot)$ is a homogeneous transition kernel on $\Omega_N \times \Omega_N$ given by

$$\begin{aligned} q_N(\xi,\eta) &= q_N(0,\eta-\xi) \\ &= \sum_{l\geq k}\left(\frac{c_{l-1}}{N^{l-1}}\right)\frac{1}{N^l} \quad \text{for } \xi,\eta\in\Omega_N \text{ such that } d(\xi,\eta)=k \quad (k\geq 1), \end{aligned} \tag{0.8}$$

where $d(\cdot,\cdot)$ is the metric

$$\begin{aligned} d(\xi,\eta) &= d(0,\eta-\xi) \\ &= \inf\{k\geq 0 : \xi_i = \eta_i \text{ for } i>k\} \end{aligned} \tag{0.9}$$

and $(c_k)_{k\geq 0}$ is any sequence of positive numbers satisfying $\sum_{k\geq 0} c_k N^{-k} < \infty$.[1] The form written in (0.8) is convenient as will be apparent from (0.12) below. Condition (0.2) is sufficient (and essentially necessary) for (0.7) to have a unique strong solution (see [S]).

The long term behavior of *hierarchical mean-field systems* like (0.7) can be studied by taking the limit $N \longrightarrow \infty$ and looking at a whole sequence of space-time scales. More precisely, introduce the block averages

$$\hat{X}_{\xi,k}^N(t) = \frac{1}{N^k}\sum_{\eta:d(\xi,\eta)\leq k} X_\eta^N(t) \quad (\xi\in\Omega_N, k\geq 0) \tag{0.10}$$

and consider

$$\left\{\left(\hat{X}_{\xi,k}^N(tN^l)\right)_{t\geq 0} : \xi\in\Omega_N\right\} \quad (k,l\geq 0). \tag{0.11}$$

This is a collection of random fields, indexed by k and l, which are to be viewed as *space-time renormalizations* of our original system (k is the space scale, l is the time scale). The analysis of these renormalized systems is based on the following rewrite of (0.7):

[1] This condition ensures that $q_N(\cdot,\cdot)$ is a transition kernel. Namely, $\sum_\eta q_N(\xi,\eta) = \sum_{k\geq 0} c_k N^{-k}$ for all $\xi\in\Omega_N$, as is easily computed via $|\{\eta : d(\xi,\eta)\leq l\}| = N^l$.

$$d\hat{X}^N_{\xi,0}(t) = \left\{ \sum_{k=1}^{\infty} \left(\frac{c_{k-1}}{N^{k-1}} \right) [\hat{X}^N_{\xi,k}(t) - \hat{X}^N_{\xi,0}(t)] \right\} dt + \sqrt{2g\left(\hat{X}^N_{\xi,0}(t)\right)}\, dW_\xi(t). \tag{0.12}$$

Theorem 0 below says that for large N there are three types of behavior depending on k and l: (a) $k > l$: the components are approximately constant; (b) $k = l$: the components fluctuate according to some diffusion with drift towards the initial density; (c) $k < l$: given $\hat{X}^N_{\xi,l}(sN^l) = \theta$ for some $1 << s << N$, the process

$$\left(\hat{X}^N_{\xi,k}(sN^l + tN^k)\right)_{t \geq 0} \tag{0.13}$$

is some diffusion with a drift towards a random value depending on θ, k and l, whose distribution can be explicitly calculated.

To state Theorem 0 in a precise form we define

$$\nu^{g,c}_\theta(dx) = \frac{1}{Z^{g,c}_\theta} \frac{1}{g(x)} \exp\left[-c \int_\theta^x \frac{y-\theta}{g(y)}\, dy \right] dx \tag{0.14}$$

$$(F_c g)(\theta) = \int_0^1 g(x) \nu^{g,c}_\theta(dx), \tag{0.15}$$

which are modifications of (0.1) and (0.3) allowing for an additional parameter $c > 0$. The probability measure $\nu^{g,c}_\theta$ is the unique equilibrium of the diffusion

$$dY(t) = c[\theta - Y(t)]\, dt + \sqrt{2g\left(Y(t)\right)}\, dW(t). \tag{0.16}$$

Define $\left(Y^{g,c}_\theta(t)\right)_{t\geq 0}$ to be the stationary solution of (0.16). Define the iterates

$$F^{(n)} g = F_{c_{n-1}} \circ \cdots \circ F_{c_0} g \quad (n \geq 0) \tag{0.17}$$

with $F^{(0)} g = g$ and $(c_n)_{n\geq 0}$ the sequence appearing in (0.8).

THEOREM 0 ([DG3]). *Let the initial state X^N have a distribution that is homogeneous, ergodic and satisfies $E(X^N_\xi) = \theta$ for all $\xi \in \Omega_N$. Then as $N \longrightarrow \infty$ the following weak convergence holds on path space (with $s(N) \longrightarrow \infty$ and $s(N) = o(N)$):*

(a) $k > l$: $\left(\hat{X}^N_{0,k}(tN^l)\right)_{t\geq 0} \Rightarrow \theta$

(b) $k = l$: $\left(\hat{X}^N_{0,k}\left(s(N)N^k + tN^k\right)\right)_{t\geq 0} \Rightarrow \left(Y^{F^{(k)}g, c_k}_\theta(t)\right)_{t\geq 0}$

(c) $k < l$: $\left(\hat{X}^N_{0,k}\left(s(N)N^l + tN^k\right)\right)_{t\geq 0} \Rightarrow \left(Y^{F^{(k)}g, c_k}_{\theta^{(l+1)}_k}(t)\right)_{t\geq 0}$,

where $(\theta^{(l+1)}_{l+1-m})_{m=0}^{l+1}$ is the backward time-inhomogeneous Markov chain on $[0, 1]$, *starting from $\theta^{(l+1)}_{l+1} = \theta$ and evolving with transition kernel at time $l + 1 - n$ given by*

$$K_{F^{(n)}g, c_n}(u, dv) = \nu^{F^{(n)}g, c_n}_u(dv). \tag{0.18}$$

Theorem 0 provides a *multiple space-time scale analysis* of (0.7). The Markov chain defined through (0.18) is called the *interaction chain* and describes how the fluctuations propagate through the levels as a result of the interaction.

Parts (a-c) should be interpreted as follows. At time $s(N)N^l$ the $(l+1)$-st block average has not yet begun to fluctuate (because $s(N) = o(N)$) and therefore still has the initial value θ (and the same for all the higher block averages). The l-th block average, however, has already begun to fluctuate, and in fact has reached equilibrium (because $s(N) \longrightarrow \infty$). The equilibrium is that of a diffusion with diffusion function $F^{(l)}g$ (see below) and with drift towards θ (the value of the $(l+1)$-st block average). In other words, the l-th block average equals the random variable $\theta_l^{(l+1)}$, which has distribution $\nu_\theta^{F^{(l)}g, c_l}$. The $(l-1)$-st block average now diffuses with a drift towards $\theta_l^{(l+1)}$, and so on, all the way down to the single-component level. *Each lower level fluctuates faster and equilibrates subject to the value of the block average one level up.*

The fact that the diffusion function at level k is $F^{(k)}g$ comes from (0.12) via martingale arguments (see [DG2] Section 3). The important point to observe here is that the $(k-1)$-st level equilibrates faster than the k-th level fluctuates. As a result, $(F^{(k)}g)(u)$ is the expectation of $(F^{(k-1)}g)(v)$ under the equilibrium distribution $\nu_u^{F^{(k-1)}g, c_{k-1}}(dv)$. This is what gives rise to (0.15) and (0.17).

Finally, let us briefly indicate how to describe our system (0.7) in analytical terms. Denote by $C([0,1]^{\Omega_N})$ the set of continuous functions on $[0,1]^{\Omega_N}$, the latter endowed with the product topology. Denote by $C_0^2([0,1]^{\Omega_N})$ the subset consisting of those functions that depend on only finitely many components and are twice differentiable w.r.t. these components. Define for $f \in C([0,1]^{\Omega_N})$

$$(S_t f)(x) = E\Big(f\big(X^N(t)\big) | X^N(0) = x\Big),$$

where E is expectation under the law of $\big(X^N(t)\big)_{t\geq 0}$. Then $(S_t)_{t\geq 0}$ is a semigroup of contractions on $C([0,1]^{\Omega_N})$, which has the Feller-property and whose generator is the closure of the following operator defined on $C_0^2([0,1]^{\Omega_N})$:

$$(Gf)(x) = \Big\{\sum_{\xi,\eta} q_N(\xi,\eta)(x_\eta - x_\xi)\frac{\partial}{\partial x_\xi} + \sum_\xi g(x_\xi)\frac{\partial^2}{\partial x_\xi^2}\Big\}(f)(x).$$

See [S] for a proof of these facts. The diffusion defined in (0.16) can be represented in a similar fashion.

0.4. *The attracting orbit.* Note that F in (0.3) is the special case of F_c in (0.15) when $c = 1$, so F^n is $F^{(n)}$ in (0.17) when $c_k \equiv 1$. Because of the obvious relation

$$F_c g = cF\Big(\frac{1}{c}g\Big), \tag{0.19}$$

most of the analysis, as we shall see, reduces to understanding the case $c_k \equiv 1$. However, for the behavior and the applications of the model the general case is important.

The key fact about the transformation F_c is that it preserves the form Const$\cdot x(1-x)$. Indeed, one checks from (0.14–15) by explicit calculation that for any $d > 0$ (see Proposition 1 in Section 1)

$$F_c(dg^*) = \frac{d}{1+\frac{d}{c}}g^* \tag{0.20}$$

where we recall that $g^*(x) = x(1-x)$ (see (0.5)). By induction it follows that for $F^{(n)}$ defined in (0.17) one has

$$F^{(n)}(dg^*) = \frac{d}{1 + a_n d} g^* \tag{0.21}$$

where a_n is defined by

$$a_n = \sum_{k=0}^{n-1} c_k^{-1}. \tag{0.22}$$

The explicitly calculated orbit in the r.h.s. of (0.21) will be the attracting orbit.

0.5. *Main theorems.* Let $\mathcal{H}$ denote the class of all functions satisfying (0.2). It is straightforward to check (see [DG2] Lemma 2.2) that $F_c\mathcal{H} \subset \mathcal{H}$ for all $c > 0$ (see also the remark at the end of this section). There are two cases to distinguish:

CASE A. $\sum_{k\geq 0} c_k^{-1} = \infty$

CASE B. $\sum_{k\geq 0} c_k^{-1} < \infty$.

Our basic convergence result reads (recall (0.5), (0.17) and (0.22)):

THEOREM 1. *For all $g \in \mathcal{H}$*

$$\lim_{n\to\infty} a_n F^{(n)} g = g^\infty \quad \textit{pointwise on } [0,1] \textit{ and in } C([0,1]) \tag{0.23}$$

with

(a) Case A: $g^\infty = g^*$

(b) Case B: $g^\infty \neq g^*$.

In Case B, g^∞ depends on g.

Theorems 2 and 3 below sharpen the result in (0.23) for Case A, Theorem 4 for Case B. In Theorem 5 we formulate a general smoothing property.

The convergence in (0.23) does not tell us much about what happens close to the boundaries because $(F^{(n)}g)(0) = (F^{(n)}g)(1) = 0$ for all $n \geq 0$. For Case A we can sharpen the statement by introducing the following norm on functions f: $[0,1] \longrightarrow [0,\infty)$:

$$\|f\| = \sup_{x\in(0,1)} \frac{|f(x)|}{x(1-x)} = \left\| \frac{f}{g^*} \right\|_{C([0,1])}. \tag{0.24}$$

There are now two classes of functions to distinguish:

$$\begin{aligned} \mathcal{H}_1 &= \left\{ g \in \mathcal{H} : \liminf_{x\downarrow 0} x^{-2} g(x) > 0 \text{ and } \liminf_{x\uparrow 1} (1-x)^{-2} g(x) > 0 \right\} \\ \mathcal{H}_2 &= \left\{ g \in \mathcal{H} : \limsup_{x\downarrow 0} x^{-2} g(x) = 0 \text{ or } \limsup_{x\uparrow 1} (1-x)^{-2} g(x) = 0 \right\}. \end{aligned} \tag{0.25}$$

THEOREM 2. *Assume Case A.*

(a) If $g \in \mathcal{H}_1$ then there exist $0 < c_g < C_g < \infty$ such that

$$c_g \leq a_n \| a_n F^{(n)} g - g^* \| \leq C_g \quad \textit{for all } n \textit{ sufficiently large.} \tag{0.26}$$

(b) If $g \in \mathcal{H}_2$ then

$$\|a_n F^{(n)} g - g^*\| \geq 1 \quad \textit{for all } n. \tag{0.27}$$

The bounds in (0.26) not only sharpen (0.23) but also give a *speed of convergence* result. We shall see in Section 1 (Proposition 3 below) that in fact the speed of convergence is order a_n^{-1} uniformly in θ.

The dichotomy between $\mathcal{H}_1$ and $\mathcal{H}_2$ has the following origin. Define the following subclasses of $\mathcal{H}_1$:

$$\tag{0.28} \begin{aligned} \bar{\mathcal{H}} &= \left\{ g \in \mathcal{H}_1 : \int_0^1 \frac{x(1-x)}{g(x)}\, dx < \infty \right\} \\ \mathcal{H}^* &= \left\{ g \in \mathcal{H}_1 : \lim_{x \downarrow 0} x^{-1} g(x) > 0 \text{ and } \lim_{x \uparrow 1} (1-x)^{-1} g(x) > 0 \right\}. \end{aligned}$$

We shall see in Section 2.1 that

$$F_c \bar{\mathcal{H}} \subset \mathcal{H}^* \quad \text{for all } c > 0. \tag{0.29}$$

Since $\mathcal{H}^* \subset \bar{\mathcal{H}}$, (0.29) implies that $F_c \mathcal{H}^* \subset \mathcal{H}^*$ for all $c > 0$. The class $\mathcal{H}^*$ turns out to be attracting for $\mathcal{H}_1$. Define

$$\tag{0.30} \begin{aligned} \bar{n}(g) &= \inf\{n \geq 0 : F^{(n)} g \in \bar{\mathcal{H}}\} \\ n^*(g) &= \inf\{n \geq 0 : F^{(n)} g \in \mathcal{H}^*\}. \end{aligned}$$

THEOREM 3. *Assume Case A.*

(a) If $g \in \mathcal{H}_1 \setminus \bar{\mathcal{H}}$ then $n^*(g) = \bar{n}(g) + 1 < \infty$. If in addition

$$\tag{0.31} \begin{aligned} &\lim_{x \downarrow 0} x^{-2} g(x) = l \\ &\lim_{x \uparrow 1} (1-x)^{-2} g(x) = r \end{aligned}$$

then

$$n^*(g) = 2 + \inf\{n \geq 0 : a_n (l \wedge r) \geq 1\}. \tag{0.32}$$

(b) If $g \in \mathcal{H}_2$ then $n^*(g) = \bar{n}(g) = \infty$.

Theorem 3(a) shows that when $g \in \mathcal{H}_1$ the iterates $F^{(n)} g$ eventually develop a positive slope at the boundaries, which is the same boundary behavior as that of g^*. When $g \in \mathcal{H}_2$, on the other hand, $F^{(n)} g$ has zero slope for all n at one or both of the boundaries, which explains (0.27). The class $\bar{\mathcal{H}}$ through which $F^{(n)} g$ passes to reach $\mathcal{H}^*$ will be interpreted in Section 0.6.

Next we turn to Case B, where the situation is different. Since g^∞ depends on g the main question here is what g^∞ looks like. This is answered in the following analogue of Theorem 3. Define

$$\tag{0.33} \begin{aligned} \mathcal{H}_1(d) &= \left\{ g \in \mathcal{H} : \liminf_{x \downarrow 0} x^{-2} g(x) > d \text{ and } \liminf_{x \uparrow 1} (1-x)^{-2} g(x) > d \right\} \\ \mathcal{H}_2(d) &= \left\{ g \in \mathcal{H} : \limsup_{x \downarrow 0} x^{-2} g(x) \leq d \text{ or } \limsup_{x \uparrow 1} (1-x)^{-2} g(x) \leq d \right\}. \end{aligned}$$

THEOREM 4. *Assume Case B.*

(a) If $g \in \mathcal{H}_1(a_\infty^{-1})$ *then* $n^*(g) = \bar{n}(g) + 1 < \infty$. *Again (0.32) holds subject to (0.31).*

(b) If $g \in \mathcal{H}_2(a_\infty^{-1})$ *then* $n^*(g) = \bar{n}(g) = \infty$.

The norm $\|\cdot\|$ defined in (0.24) is no longer appropriate in Case B because $g^\infty \not\equiv g^*$. Therefore we have no analogue of Theorem 2 for Case B. One can attempt to adapt the norm, but since g^∞ depends on g this is of less interest anyway.

We conclude our description of F with the following smoothing property.

THEOREM 5. *For all* $g \in \mathcal{H}$ *and* $c > 0$ *the function* $\theta \longrightarrow (F_c g)(\theta)$ *is* C^∞ *on* $(0, 1)$.

REMARK. The class where the map F_c is defined can actually be chosen much larger than $\mathcal{H}$ given by (0.2). Namely, let $\mathcal{H}'$ be the class of functions $g\colon [0,1] \longrightarrow [0,\infty)$ satisfying

$$g \text{ measurable} \tag{0.34i}$$

$$g^{-1} \text{ locally integrable on } (0,1) \tag{0.34ii}$$

$$\int_0^{1/2} \frac{dx}{g(x)} = \int_{1/2}^1 \frac{dx}{g(x)} = \infty. \tag{0.34iii}$$

We shall see in Section 2.1 that $F_c\mathcal{H}' \subset \mathcal{H}$ for all $c > 0$, so that after one iteration one falls onto the original class $\mathcal{H}$. The class $\mathcal{H}$ was needed for (0.7) and (0.16) to have a unique strong solution, but (0.14) and (0.15) are well-defined in $\mathcal{H}'$.

0.6. *Interpretation.* In this section we continue with the discussion started in Section 0.3 and interpret Cases A, B and Theorems 1–4 from the probabilistic point of view.

We start by explaining the dichotomy between Cases A and B. Suppose, as we did in Theorem 0, that the system in (0.7) starts in an initial state X^N which has a distribution that is homogeneous, ergodic and satisfies $E(X_\xi^N) = \theta$ for all $\xi \in \Omega_N$. Then, as is shown in [DG3], two types of behavior are possible on the single-component level, namely as $t \longrightarrow \infty$

$$\mathcal{L}\big(X^N(t)\big) \Rightarrow \mu_\theta \quad \text{(Case 1)} \tag{0.35i}$$

$$\mathcal{L}\big(X^N(t)\big) \Rightarrow (1-\theta)\delta_{\{x\equiv 0\}} + \theta\delta_{\{x\equiv 1\}} \quad \text{(Case 2).} \tag{0.35ii}$$

Here μ_θ in Case 1 is some non-degenerate equilibrium state on $[0,1]^{\Omega_N}$ that is again homogeneous, ergodic and with density θ, while in Case 2 the limit is degenerate with point masses at the traps $\{x \equiv 0\}$ and $\{x \equiv 1\}$. Case 1 is called *stable*, Case 2 is called *clustering*. The latter means that the system develops patterns of growing blocks in which the components are either all close to 0 or all close to 1.

Now, glancing at (0.11) we see that the interaction (= drift) term and the fluctuation (= diffusion) term compete: without fluctuation ($g \equiv 0$) the system goes to $\delta_{\{x\equiv\theta\}}$, without interaction ($c_k \equiv 0$) it goes to $(1-\theta)\delta_{\{x\equiv 0\}} + \theta\delta_{\{x\equiv 1\}}$. Therefore one expects to get Case 1 when the interaction is strong and Case 2 when the interaction is weak. Indeed, it is proved in [DG3] that for all $g \in \mathcal{H}$

$$\text{(0.36)} \qquad \begin{aligned} &\text{Case 1:} \quad \sum_{k\geq 0} c_k^{-1} < \infty \quad \text{(Case B)} \\ &\text{Case 2:} \quad \sum_{k\geq 0} c_k^{-1} = \infty \quad \text{(Case A).} \end{aligned}$$

To get some feeling for why this is so we have to return to the interaction chain that was defined through (0.18). According to Theorem 0(c), the distribution of the single components (*i.e.*, space scale $k = 0$) at time $s(N)N^l$ (*i.e.*, time scale l) is the equilibrium of the diffusion in (0.16) with $c = c_0$, diffusion function g, and drift towards $\theta_0^{(l+1)}$. In Section 1 we shall see that what is responsible for Theorem 1 is the following dichotomy as $l \longrightarrow \infty$:

$$\text{(0.37)} \qquad \begin{aligned} &\mathcal{L}(\theta_0^{(l+1)}) \Rightarrow \gamma_\theta, \text{ some law with } \gamma_\theta(0,1) = 1 \quad \text{(Case A)} \\ &\mathcal{L}(\theta_0^{(l+1)}) \Rightarrow (1-\theta)\delta_0 + \theta\delta_1. \quad \text{(Case B)} \end{aligned}$$

In other words, for every θ the interaction chain has a non-trivial *entrance law* only in Case A. This explains (0.35) and (0.36).

We can now interpret Theorems 1 and 2. What Theorem 1 shows is that the dichotomy between stable and clustering is *universal* in g. Moreover, Theorems 1 and 2 show that in the clustering case the diffusion function at level k is close to $a_k^{-1}x(1-x)$ for k large. This in turn can be shown to imply that *the laws governing the formation and growth of the clusters are universal* in g too. Indeed, it is proved in [DG3] that for $c_k \equiv c$ and subject to the property $\lim_{n\to\infty} \|a_n F^{(n)} g - g^*\| = 0$ one has as $l \longrightarrow \infty$

$$\text{(0.38)} \qquad (\theta^{(l)}_{(1-\alpha)l})_{\alpha\in[0,1)} \Rightarrow (Y_{\log(\frac{1}{1-\alpha})})_{\alpha\in[0,1)}$$

(the limit is independent of c). Here $(Y_t)_{t\geq 0}$ is the Fisher-Wright diffusion, *i.e.*, the diffusion on $[0,1]$ generated by $\frac{1}{2}x(1-x)\frac{\partial^2}{\partial x^2}$, starting at $Y_0 = \theta$. If one defines

$$\text{(0.39)} \qquad \tau = \inf\{\alpha \in [0,1) : Y_{\log(\frac{1}{1-\alpha})} = 0 \text{ or } 1\},$$

then (0.38) says that at time scale l (*i.e.*, time N^l) the largest cluster has a hierarchical radius equal to $(1-\tau)l$ (*i.e.*, volume $N^{(1-\tau)l}$) for large l. This means that the clusters grow at a *random linear speed* τ in the hierarchical distance.

The importance of Theorems 3 and 4 is in another direction. To explain why, we first make the following observation. The diffusion generated by $g(x)\frac{\partial^2}{\partial x^2}$ (so (0.16) with $c = 0$) has both 0 and 1 as *accessible boundary points* iff $\int_0^1 \frac{x(1-x)}{g(x)}\,dx < \infty$ (see [B] Proposition 16.43), *i.e.*, the diffusion eventually hits one of the traps at 0 or 1 iff $g \in \bar{\mathcal{H}}$. Therefore Theorem 3(a) says that for every $g \in \mathcal{H}_1 \setminus \bar{\mathcal{H}}$ there exists $\bar{n}(g) < \infty$ such that

$$\text{(0.40)} \qquad \begin{aligned} &n < \bar{n}(g): \quad (F^{(n)}g)(x)\frac{\partial^2}{\partial x^2} \text{ has at least one non-accessible boundary} \\ &n \geq \bar{n}(g): \quad (F^{(n)}g)(x)\frac{\partial^2}{\partial x^2} \text{ has both boundaries accessible.} \end{aligned}$$

This change of character at $n = \bar{n}(g)$ has interesting consequences for *large* but *finite* systems in the mean field limit $N \longrightarrow \infty$. Namely, consider the situation where $c_k > 0$ for $k < m$ and $c_k = 0$ for $k \geq m$. Then (0.7) breaks up into a collection of independent subsystems each of size N^m. (Recall (0.7–9) and note that $d(\xi, \eta), d(\xi, \chi) \leq m$ implies $d(\eta, \chi) \leq m$.) It turns out that as $N \longrightarrow \infty$ one gets the following behavior ([DG1],[CGS]):

$$\mathcal{L}\big(X^N(tN^m)\big) \Rightarrow \int Q_t^{(m)}(\theta, d\theta')\Big\{\bigotimes_{\Omega_\infty} \nu_{\theta'}^{F^{(m)}g,0}(\cdot)\Big\} \quad (t > 0) \tag{0.41}$$

$$\mathcal{L}\big(\hat{X}_{0,m}^N(tN^m)\big) \Rightarrow Q_t^{(m)}(\theta, \cdot) \quad (t > 0).$$

Here $\Omega_\infty = \cup_N \Omega_N$ and $(Q_t^{(m)})_{t \geq 0}$ is the semigroup on $[0,1]$ generated by $F^{(m)}g(x)\frac{\partial^2}{\partial x^2}$. Since for every finite N the process $\big(\hat{X}_{0,m}^N(t)\big)_{t\geq 0}$ is a diffusion controlled by g and *not* by $F^{(m)}g$ (see (0.12)), we have the following remarkable situation: For any $m \geq \bar{n}(g)$ and $g \in \mathcal{H}_1 \setminus \bar{\mathcal{H}}$

$$\big(\hat{X}_{0,m}^N(t)\big)_{t\geq 0} \text{ has at least one non-accessible boundary for every } N \tag{0.42}$$

$$\big(\hat{X}_{0,m}^\infty(t)\big)_{t\geq 0} \text{ has both boundaries accessible,}$$

where $\big(\hat{X}_{0,m}^\infty(t)\big)_{t\geq 0}$ is the diffusion with semigroup $(Q_t^{(m)})_{t\geq 0}$ starting at θ. In other words, the system has *two phases*: (1) $m \geq \bar{n}(g)$: the mean field limit of the system has an accessible boundary where the original system has not; (2) $m < \bar{n}(g)$: both systems have the same boundary behavior. The existence of the first phase is due to the "cooperation" of the components.

The above observation is important for models with $g(x) = d\big(x(1-x)\big)^2$, which are of interest in genetics: so-called Ohta-Kimura diffusions ([OK]). According to (0.31) and (0.32) there is a *phase transition* in the parameter d: the system switches between the two phases when d crosses the value a_∞^{-1}. Thus for the qualitative behavior of large finite collections of Ohta-Kimura diffusions the constant d is in fact crucial.

1. **Proof of Theorems 1 and 2.** The following four relations are the key to Theorems 1 and 2:

PROPOSITION 1. *For all $g \in \mathcal{H}$, $c > 0$ and $\theta \in [0,1]$*

$$\int_0^1 \nu_\theta^{g,c}(dx) = 1 \tag{1.1a}$$

$$\int_0^1 x\nu_\theta^{g,c}(dx) = \theta \tag{1.1b}$$

$$\int_0^1 x^2\nu_\theta^{g,c}(dx) = \theta^2 + \frac{1}{c}(F_c g)(\theta) \tag{1.1c}$$

$$\int_0^1 g(x)\nu_\theta^{g,c}(dx) = (F_c g)(\theta). \tag{1.1d}$$

PROOF. (a), (d) are (0.14), (0.15). It is straightforward to check (b), (c) (see also Lemma 3 in Section 2.2). One way is via Itô's formula using that $\nu_\theta^{g,c}(dx)$ is the equilibrium of (0.16). The derivation along this line also makes it clear that what matters for (b), (c) is not so much the explicit form of $\nu_\theta^{g,c}(dx)$ but rather its equilibrium property. ∎

For $g \in \mathcal{H}$ and $c > 0$ define the probability kernel on $[0,1] \times [0,1]$:

$$K_{g,c}(x, dy) = \nu_x^{g,c}(dy). \tag{1.2}$$

For $g \in \mathcal{H}$ and $(c_k)_{k\geq 0}$ any sequence of positive numbers define the compositions

$$K^{(n)}(x, dy) = K_{F^{(n)}g,c_n} \circ \ldots \circ K_{F^{(0)}g,c_0}(x, dy) \quad (x, y \in [0,1], n \geq 0) \tag{1.3}$$

$$F^{(n)}g = F_{c_{n-1}} \circ \ldots \circ F_{c_0}g \quad (n \geq 0). \tag{1.4}$$

(Note that $F^{(0)}g = g$ and $K^{(0)} = K_{g,c_0}$. See (0.17) and (0.18) for the probabilistic background.)

PROPOSITION 2. *For all $g \in \mathcal{H}$, $\theta \in [0,1]$ and $n \geq 0$*

$$\int_0^1 K^{(n)}(\theta, dy) = 1 \tag{1.5a}$$

$$\int_0^1 yK^{(n)}(\theta, dy) = \theta \tag{1.5b}$$

$$\int_0^1 y^2 K^{(n)}(\theta, dy) = \theta^2 + a_{n+1}(F^{(n+1)}g)(\theta) \tag{1.5c}$$

$$\int_0^1 g(y)K^{(n)}(\theta, dy) = (F^{(n+1)}g)(\theta) \tag{1.5d}$$

where $a_n = \sum_{k=0}^{n-1} c_k^{-1}$ (see (0.22)).

PROOF. (a), (b) are immediate from (1.1a), (1.1b); (d) follows from (1.1d) via (1.4); (c) is obtained by combining (d) with (1.1c). ∎

Subtracting (c) from (d) in (1.5) we get

$$0 \leq \int_0^1 y(1-y)K^{(n)}(\theta, dy) = \theta(1-\theta) - a_{n+1}(F^{(n+1)}g)(\theta). \tag{1.6}$$

We now give the proof of Theorems 1 and 2.

PROOF OF THEOREM 1.

CASE A. Because $a_n \longrightarrow \infty$ it follows from (1.6) that $(F^{(n)}g)(\theta) \longrightarrow 0$ for all $\theta \in [0,1]$. By (1.5d) this implies

$$K^{(n)}(\theta, dy) \Rightarrow (1-\theta)\delta_0 + \theta\delta_1 \quad (n \longrightarrow \infty) \tag{1.7}$$

($\Rightarrow$ means weak convergence of measures). To get (1.7) we use that g is strictly bounded away from zero on any closed interval contained in (0,1) (because of (0.2ii), (0.2iii)) and we note that the weights $1-\theta, \theta$ come from (1.5b). Inserting (1.7) into (1.6) we get

$$\lim_{n\to\infty} a_n(F^{(n)}g)(\theta) = \theta(1-\theta). \tag{1.8}$$

CASE B. Define for $0 \le m \le n$

$$K^{(n,m)}(x,dy) = K_{F^{(n)}g,c_n} \circ \cdots \circ K_{F^{(m)}g,c_m}(x,dy) \tag{1.9}$$
$$F^{(n,m)}g = F_{c_{n-1}} \circ \cdots \circ F_{c_m} g.$$

Then the analogues of (1.5b), (1.5c) read

$$\int_0^1 yK^{(n,m)}(\theta,dy) = \theta \tag{1.10}$$
$$\int_0^1 y^2 K^{(n,m)}(\theta,dy) = \theta^2 + (a_{n+1} - a_{m+1})(F^{(n+1,m)}g)(\theta).$$

Consequently (since $a_\infty < \infty$)

$$\lim_{m\to\infty} \sup_{n\ge m} \int_0^1 (y-\theta)^2 K^{(n,m)}(\theta,dy) = 0. \tag{1.11}$$

Hence, because g is continuous,

$$\lim_{m\to\infty} \sup_{n\ge m} \int_0^1 g(y)K^{(n,m)}(\theta,dy) = g(\theta). \tag{1.12}$$

Since $K^{(n)} = K^{(n,m)} \circ K^{(m-1)}$ it now follows that

$$\int_0^1 g(y)K^{(n)}(\theta,dy) = (F^{(n+1)}g)(\theta) \tag{1.13}$$

is a Cauchy sequence for every θ. Its limit is what we define as $g^\infty(\theta)$.

Thus we have proved pointwise convergence for Case A and Case B. The convergence in $C([0,1])$ will follow from the pointwise convergence if we can show that the sequence $\{a_n F^{(n)}g : n \ge 0\}$ is uniformly equicontinuous on $\mathcal{H}$ (Arzela-Ascoli). For this it suffices to show that $\sup_n L(a_n F^{(n)}g) < \infty$, where $L(f)$ denotes the Lipschitz constant of f. But by (1.5c) and (1.10), the latter in turn is implied by the following lemma:

LEMMA 1. *$L(K_{g,c}f) \le L(f)$ for all $f, g \in \mathcal{H}$ and $c > 0$.*

A proof of Lemma 1 via coupling techniques is given in [DG2] Lemma 2.2. We give an analytic proof in Section 2.7. ■

PROOF OF THEOREM 2. Assume Theorem 3 and the following lemma:

LEMMA 2. *If $g_1 \le g_2$ then $F_c g_1 \le F_c g_2$ for all $c > 0$.*

The proof of Lemma 2 is deferred to Section 2.4. We proceed by showing how Theorem 2 follows. Recall (0.25) and (0.28).

According to Theorem 3(a), if $g \in \mathcal{H}_1$ then there exists $n^*(g) < \infty$ such that

$$F^{(n^*(g))}g \in \mathcal{H}^*. \tag{1.14}$$

The following proposition will give the proof:

PROPOSITION 3. *For every $g \in \mathcal{H}^*$ there exist $0 < c_g < C_g < \infty$ such that for all $n \ge 0$*

$$\frac{c_g}{a_n} \le 1 - \frac{a_n(F^{(n)}g)(\theta)}{\theta(1-\theta)} \le \frac{C_g}{a_n} \quad \textit{uniformly in } \theta. \tag{1.15}$$

PROOF. The upper bound is obtained as follows. Since $g \in \mathcal{H}^*$ we have (see (0.2ii), (0.2iii) and (0.28))

$$g(y) \geq \delta y(1-y) \quad \text{for some } \delta > 0. \tag{1.16}$$

Substitute this inequality into (1.6) to obtain

$$\begin{aligned} \theta(1-\theta) - a_{n+1}(F^{(n+1)}g)(\theta) &\leq \frac{1}{\delta}\int_0^1 g(y)K^{(n)}(\theta, dy) \\ &= \frac{1}{\delta}(F^{(n+1)}g)(\theta) \\ &\leq \frac{1}{\delta a_n}\theta(1-\theta), \end{aligned} \tag{1.17}$$

where the equality is (1.5d) and the last inequality holds because $a_{n+1}(F^{(n+1)}g)(\theta) \leq \theta(1-\theta)$, as is obvious from (1.6).

The lower bound follows from Lemma 2. Indeed, since for every $g \in \mathcal{H}$ also

$$g(y) \leq \Delta y(1-y) \quad \text{for some } \Delta < \infty, \tag{1.18}$$

we can apply Lemma 2 and (0.21) to conclude

$$a_n(F^{(n)}g)(y) \leq a_n \frac{\Delta}{1 + a_n\Delta} y(1-y). \tag{1.19}$$

■

Proposition 3 combined with (1.9) shows that (1.15) holds for all $g \in \mathcal{H}_1$ and $n \geq n^*(g)$, after a shift of the sequence (c_k) over a distance $n^*(g)$. This completes the proof of Theorem 2(a).

It has already been explained in Section 0.5 why Theorem 2(b) is immediate from Theorem 3(b). ■

2. **Boundary behavior of $F_c g$.** This section is devoted to studying the relation between the boundary behavior of g and $F_c g$. The results derived here will be used in Section 3 to prove Theorems 3 and 4. Section 2.1 contains four main propositions. These are proved in Sections 2.2, 2.3 and 2.5. In Sections 2.4 and 2.7 we prove Lemmas 2 resp. 1, which were already used in Section 1. The proof of Theorem 5 is in Section 2.6.

2.1. *Main propositions.* Because of (0.19) it suffices to consider $c = 1$. We formulate our results only for the left boundary at $\theta = 0$, the right boundary at $\theta = 1$ being analogous. We start by assuming that g satisfies (0.34i), (0.34ii) which are the minimal conditions required for Fg to be well-defined.

PROPOSITION 4. *$\lim_{\theta \downarrow 0}(Fg)(\theta) = c$ exists with*

$$\begin{aligned} c &\in (0,\infty) \quad \textit{if } \int_0^{1/2} \frac{dx}{g(x)} < \infty \\ c &= 0 \quad \textit{if } \int_0^{1/2} \frac{dx}{g(x)} = \infty. \end{aligned} \tag{2.1}$$

PROPOSITION 5. *Assume* $\int_0^{1/2} \frac{dx}{g(x)} = \infty$. *Then* $\lim_{\theta\downarrow 0} \theta^{-1}(Fg)(\theta) = c'$ *exists with*

$$\begin{aligned} &c' \in (0,\infty) \quad \textit{if} \int_0^{1/2} \frac{x}{g(x)}\,dx < \infty \\ &c' = 0 \quad \textit{if} \int_0^{1/2} \frac{x}{g(x)}\,dx = \infty. \end{aligned} \tag{2.2}$$

From (2.1) and (2.2) we see that (0.34iii) is the natural condition to add in order to ensure that the iterates $F^n g$ remain zero at the boundaries.[2] We also see from (2.2) that $F\mathcal{H}' \subset \mathcal{H}$, as was claimed in the remark at the end of Section 0.5 (recall (0.2), (0.34) and Lemma 1). Note that the first case in (2.2) corresponds to the class $\bar{\mathcal{H}}$ and shows that $F\bar{\mathcal{H}} \subset \mathcal{H}^*$, as was claimed in (0.29).

To determine the domain of attraction of the class $\mathcal{H}^*$ the following explicitly calculable example is important:

PROPOSITION 6. *Let* $g(x) = dx^2$ $(x \in [0,1], d > 0)$. *Then as* $\theta \downarrow 0$

$$\begin{aligned} &0 < d < 1 \quad (Fg)(\theta) \sim \frac{d}{1-d}\theta^2 \\ &d = 1 \quad (Fg)(\theta) \sim \theta^2 \log\left(\frac{1}{\theta}\right) \\ &d > 1 \quad (Fg)(\theta) \sim c_d \theta^{1+\frac{1}{d}} \ (c_d > 0). \end{aligned} \tag{2.3}$$

The important point to note here is that the curvature at the boundary increases under F and that the case $d > 1$ leads to the boundary behavior as in $\bar{\mathcal{H}}$.[3]

The final statement is the following technical property showing that *the left and the right boundary behavior are decoupled.* Define

$$(G^\eta g)(\theta) = \frac{\int_0^\eta g(x)\nu_\theta^g(dx)}{\int_0^\eta \nu_\theta^g(dx)} \quad (\theta, \eta \in [0,1]). \tag{2.4}$$

This is the conditional expectation of g under $\nu_\theta^g(dx)$ given $x \leq \eta$. Note that $G^1 = F$.

PROPOSITION 7. *For every* $g \in \mathcal{H}'$ *(recall (0.34)) and* $\eta \in (0,1)$

$$(G^\eta g)(\theta) \sim [1 + C_g(\eta)]^{-1}(Fg)(\theta) \quad (\theta \downarrow 0) \tag{2.5}$$

with

$$0 \leq C_g(\eta) \leq \frac{1-\eta}{\eta} \tag{2.6i}$$

$$\textit{If } g(x) \leq x^2 \textit{ for } x \in [0,\eta) \textit{ then } C_g(\eta) = 0. \tag{2.6ii}$$

[2] This condition is what confines our diffusion, defined in (0.7) and (0.16), to the interval [0, 1] in a natural way. See [B] Definitions 16.48–49.

[3] Observe that the g in Proposition 6 has $g(1) \neq 0$ and therefore does not satisfy (0.34iii). However, due to (2.5) it will still be useful, as we shall see later.

2.2. *Proof of Propositions 4 and 5.* We begin by rewriting the definition of F into a form more suitable for manipulations. Namely we put

$$\mu_\theta^g(x) = \frac{1}{g(x)} \exp\left[-\int_\theta^x \frac{y-\theta}{g(y)}\,dy\right] \quad \big(\theta \in (0,1)\big) \tag{2.7}$$

$$(Fg)(\theta) = \frac{\int_0^1 g(x)\mu_\theta^g(x)\,dx}{\int_0^1 \mu_\theta^g(x)\,dx} \quad \big(\theta \in (0,1)\big). \tag{2.8}$$

The integrand in the numerator now has a nice shape property.

LEMMA 3. *For all* $\theta \in (0,1)$

$$\frac{\partial}{\partial x}[g(x)\mu_\theta^g(x)] = (\theta - x)\mu_\theta^g(x) \tag{2.9}$$

$$g(\theta)\mu_\theta^g(\theta) = 1. \tag{2.10}$$

Hence $x \longrightarrow g(x)\mu_\theta^g(x)$ *is increasing on* $(0,\theta)$, *decreasing on* $(\theta, 1)$ *and has a maximum 1 at* θ.

PROOF. Immediate from (2.7). ∎

To complete Lemma 3, define

$$\begin{aligned} l(\theta) &= \lim_{x\downarrow 0} g(x)\mu_\theta^g(x) \\ r(\theta) &= \lim_{x\uparrow 1} g(x)\mu_\theta^g(x). \end{aligned} \tag{2.11}$$

LEMMA 4. $l(\theta) = 0$ *for all* $\theta \in (0,1)$ *iff* $\int_0^{1/2} \frac{dx}{g(x)} = \infty$, *and similarly for* $r(\theta)$.

PROOF. By the monotone convergence theorem we have from (2.7)

$$l(\theta) = \exp\left[-\int_0^\theta \frac{\theta - y}{g(y)}\,dy\right]. \tag{2.12}$$

Now substitute the inequalities

$$\frac{\theta}{2}\int_0^{\theta/2} \frac{dy}{g(y)} \le \int_0^\theta \frac{\theta-y}{g(y)}\,dy \le \theta \int_0^\theta \frac{dy}{g(y)}. \tag{2.13}$$

∎

The proof of the first part of Proposition 4 is easy. Indeed, if $\int_0^{1/2} \frac{dx}{g(x)} < \infty$ then by Lebesgue's dominated convergence theorem applied to (2.7) and (2.8)

$$\lim_{\theta\downarrow 0}(Fg)(\theta) = \frac{\int_0^1 \exp\left[-\int_0^x \frac{y}{g(y)}\,dy\right]dx}{\int_0^1 \frac{1}{g(x)} \exp\left[-\int_0^x \frac{y}{g(y)}\,dy\right]dx} \tag{2.14}$$

with the r.h.s. obviously positive and finite.

Since the second part of Proposition 4 is implied by Proposition 5 we continue with the proof of the latter. This will need some preparatory estimates contained in Lemmas 5 and 6 below. Define for $0 \le a < b \le 1$

$$I([a,b];\theta) = \int_a^b \mu_\theta^g(x)\,dx. \tag{2.15}$$

LEMMA 5. *For every* $\theta \in (0,1)$

$$I([0,\theta];\theta) \geq \frac{1}{\theta}\big(1-l(\theta)\big) \tag{2.16}$$

$$I([\theta,1];\theta) \geq \frac{1}{1-\theta}\big(1-r(\theta)\big).$$

For every $0 < x_1 \leq \theta \leq x_2 < 1$

$$I([0,x_1];\theta) \leq \frac{1}{\theta - x_1} \tag{2.17}$$

$$I([x_2,1];\theta) \leq \frac{1}{x_2-\theta}$$

$$I([x_1,x_2];\theta) \leq \left[1-(\theta-x_1)\int_{x_1}^{\theta}\frac{dy}{g(y)}\right]^{-1}\int_{x_1}^{\theta}\frac{dy}{g(y)} + \left[1-(x_2-\theta)\int_{\theta}^{x_2}\frac{dy}{g(y)}\right]^{-1}\int_{\theta}^{x_2}\frac{dy}{g(y)}.$$

(The last inequality is understood to apply only when both denominators are positive.)

PROOF. Substitute (2.9) into (2.15) to obtain

$$I([a,b];\theta) = \int_a^b \frac{\partial}{\partial x}[g(x)\mu_\theta^g(x)]\frac{dx}{\theta - x}. \tag{2.18}$$

On the integration area we have $(\theta-a)^{-1} \leq (\theta-x)^{-1} \leq (\theta-b)^{-1}$. By Lemma 3 we have $\frac{\partial}{\partial x}[g(x)\mu_\theta^g(x)] \geq 0$ for all $x \in [0,\theta]$. Hence

$$\frac{1}{\theta-a}[g(b)\mu_\theta^g(b) - g(a)\mu_\theta^g(a)] \leq I([a,b];\theta) \leq \frac{1}{\theta-b}[g(b)\mu_\theta^g(b) - g(a)\mu_\theta^g(a)] \quad \text{for all } 0 \leq a < b \leq \theta. \tag{2.19}$$

Now substitute $a = 0$ and $b = x_1, \theta$, and use (2.10) together with $g\mu_\theta^g \leq 1$, to get the first inequalities in (2.16) and (2.17). The second inequalities are derived similarly.

To prove the third inequality in (2.17), split

$$I([x_1,x_2];\theta) = I([x_1,\theta];\theta) + I([\theta,x_2];\theta). \tag{2.20}$$

For the first integral write

$$\begin{aligned} I([x_1,\theta];\theta) &= \int_{x_1}^{\theta} g(x)\mu_\theta^g(x)\frac{dx}{g(x)} \\ &= \left[g(x)\mu_\theta^g(x)\int_\theta^x \frac{dy}{g(y)}\right]_{x=x_1}^{\theta} - \int_{x_1}^{\theta}\frac{\partial}{\partial x}[g(x)\mu_\theta^g(x)]\left\{\int_\theta^x \frac{dy}{g(y)}\right\}dx \\ &= g(x_1)\mu_\theta^g(x_1)\int_{x_1}^{\theta}\frac{dy}{g(y)} + \int_{x_1}^{\theta}(\theta-x)\mu_\theta^g(x)\left\{\int_x^\theta \frac{dy}{g(y)}\right\}dx \\ &= \big[g(x_1)\mu_\theta^g(x_1) + (\theta-x_1)I([x_1,\theta];\theta)\big]\int_{x_1}^{\theta}\frac{dy}{g(y)} \end{aligned} \tag{2.21}$$

where the third equality uses (2.9). Since $g\mu_\theta^g \leq 1$ this gives the first half of the upper bound in (2.17). The second half is similar. ∎

LEMMA 6. *For all $g \in \mathcal{H}'$ (recall (0.34))*

$$\liminf_{\theta \downarrow 0} \theta \int_0^1 \mu_\theta^g(x)\,dx \geq 1. \tag{2.22}$$

If $\int_0^{1/2} \frac{x}{g(x)}\,dx < \infty$ then

$$\lim_{\theta \downarrow 0} \theta \int_0^1 \mu_\theta^g(x)\,dx = 1. \tag{2.23}$$

PROOF. By (2.16)

$$\theta \int_0^1 \mu_\theta^g(x)\,dx \geq \theta \left[\frac{1}{\theta} + \frac{1 - r(\theta)}{1 - \theta} \right] \tag{2.24}$$

where we use that $l(\theta) = 0$ by Lemma 4. Let $\theta \downarrow 0$ to get (2.22).

Pick $0 < \alpha < 1 < \beta < \infty$, $x_1 = \alpha\theta$ and $x_2 = \beta\theta$ with θ sufficiently small. Then summing the upper bounds in (2.17) we obtain

$$\theta \int_0^1 \mu_\theta^g(x)\,dx \leq \frac{1}{1-\alpha} + \frac{1}{\beta - 1} + \frac{A_\alpha(\theta)}{1 - (1-\alpha)A_\alpha(\theta)} + \frac{B_\beta(\theta)}{1 - (\beta - 1)B_\beta(\theta)} \tag{2.25}$$

with the abbreviations

$$A_\alpha(\theta) = \theta \int_{\alpha\theta}^{\theta} \frac{dx}{g(x)} \tag{2.26}$$

$$B_\beta(\theta) = \theta \int_\theta^{\beta\theta} \frac{dx}{g(x)}.$$

Since $\int_0^{1/2} \frac{x}{g(x)}\,dx < \infty$ implies that $A_\alpha(\theta), B_\beta(\theta) \longrightarrow 0$ as $\theta \downarrow 0$ we have

$$\limsup_{\theta \downarrow 0} \theta \int_0^1 \mu_\theta^g(x)\,dx \leq \frac{1}{1-\alpha} + \frac{1}{\beta - 1}. \tag{2.27}$$

Finally, let $\alpha \downarrow 0$ and $\beta \uparrow \infty$ and combine with (2.22) to arrive at (2.23). ∎

We can now prove Proposition 5. From (2.8) we have

$$\theta^{-1}(Fg)(\theta) = \frac{\int_0^1 g(x)\mu_\theta^g(x)\,dx}{\theta \int_0^1 \mu_\theta^g(x)\,dx}. \tag{2.28}$$

Because $g\mu_\theta^g \leq 1$, the nominator converges to

$$\int_0^1 \exp\left[- \int_0^x \frac{y}{g(y)}\,dy \right] dx \tag{2.29}$$

by Lebesgue's the monotone convergence theorem. Therefore (2.2) is immediate from Lemma 6.

2.3. *Proof of Proposition 6.* Abbreviate the quotient in (2.8) by $N(\theta)/D(\theta)$. If $g(x) = dx^2$ then

$$g(x)\mu_\theta^g(x) = \exp\Big[-\int_\theta^x \frac{y-\theta}{dy^2}\,dy\Big] = e^{\frac{1}{d}}\Big(\frac{\theta}{x}\Big)^{\frac{1}{d}}\exp\Big[-\frac{\theta}{dx}\Big]. \tag{2.30}$$

For $D(\theta)$ this gives

$$\lim_{\theta\downarrow 0}\theta D(\theta) = e^{\frac{1}{d}}\,d^{\frac{1}{d}}\lim_{\theta\downarrow 0}\int_{\theta/d}^\infty z^{\frac{1}{d}}e^{-z}\,dz = e^{\frac{1}{d}}d^{\frac{1}{d}}\Gamma\Big(\frac{1}{d}+1\Big). \tag{2.31}$$

For $N(\theta)$, on the other hand, the behavior depends on d, namely

$$\begin{aligned} &0<d<1 && \lim_{\theta\downarrow 0}\theta^{-1}N(\theta) = e^{\frac{1}{d}}\,d^{\frac{1}{d}-1}\lim_{\theta\downarrow 0}\int_{\theta/d}^\infty z^{\frac{1}{d}-2}e^{-z}\,dz = e^{\frac{1}{d}}\,d^{\frac{1}{d}-1}\Gamma\Big(\frac{1}{d}-1\Big)\\ &d=1 && \lim_{\theta\downarrow 0}\Big[\theta\log\Big(\frac{1}{\theta}\Big)\Big]^{-1}N(\theta) = e\lim_{\theta\downarrow 0}\Big[\log\Big(\frac{1}{\theta}\Big)\Big]^{-1}\int_\theta^\infty z^{-1}e^{-z}\,dz = e\\ &d>1 && \lim_{\theta\downarrow 0}\theta^{-\frac{1}{d}}N(\theta) = e^{\frac{1}{d}}\int_0^\infty x^{-\frac{1}{d}}\,dx = e^{\frac{1}{d}}\frac{d}{d-1}. \end{aligned} \tag{2.32}$$

By combining (2.31) and (2.32), and writing $\frac{1}{d}\Gamma(\frac{1}{d}-1) = \frac{d}{1-d}\Gamma(\frac{1}{d}+1)$, we get (2.3).

2.4. *Proof of Lemma 2.* The monotonicity of F expressed by Lemma 2 is a consequence of the following property.

LEMMA 7. *For any $0\le a\le\theta\le b\le 1$, if $g_1\le g_2$ on $[a,b]$ then*

$$g_1\mu_\theta^{g_1} \le g_2\mu_\theta^{g_2} \text{ on } [a,b] \tag{2.33)(i}$$

$$\int_a^b \mu_\theta^{g_1}(x)\,dx \ge \int_a^b \mu_\theta^{g_2}(x)\,dx. \tag{2.33)(ii}$$

PROOF. Part (i) is evident from (2.7) because $-g_1^{-1}\le -g_2^{-1}$ on $[a,b]$. Part (ii) follows from Part (i) and the representations

$$\begin{aligned} \int_\theta^b \mu_\theta^g(x)\,dx &= \frac{1-g(b)\mu_\theta^g(b)}{b-\theta} + \int_\theta^b \frac{1-g(x)\mu_\theta^g(x)}{(x-\theta)^2}\,dx \quad (b>\theta)\\ \int_a^\theta \mu_\theta^g(x)\,dx &= \frac{1-g(a)\mu_\theta^g(a)}{\theta-a} + \int_a^\theta \frac{1-g(x)\mu_\theta^g(x)}{(\theta-x)^2}\,dx \quad (a<\theta) \end{aligned} \tag{2.34}$$

which are obtained by partial integration using (2.9). (First exclude an ϵ-neighborhood of θ to avoid the pole of $x\longrightarrow(x-\theta)^{-1}$ and then let $\epsilon\downarrow 0$.) ∎

The inequalities in (2.33)(i), (ii) go in the *opposite* direction and therefore Lemma 2 now follows from (2.8) by setting $a=0, b=1$.

2.5. *Proof of Proposition 7.* Fix $g\in\mathcal{H}'$ and $\eta\in(0,1)$. If $\theta<\eta$ then it follows from the second inequality in (2.17) and the second expression in (2.34) that

$$\begin{aligned} \int_\eta^1 \mu_\theta^g(x)\,dx &\le \frac{1}{\eta-\theta}\\ \int_0^\eta \mu_\theta^g(x)\,dx \ge \int_0^\theta \mu_\theta^g(x)\,dx &= \frac{1}{\theta} \end{aligned} \tag{2.35}$$

where we use Lemma 4. This implies

$$\lim_{\theta\downarrow 0} \frac{\int_\eta^1 \mu_\theta^g(x)\,dx}{\int_0^\eta \mu_\theta^g(x)\,dx} = 0. \tag{2.36}$$

Next, $g(x)\mu_\theta^g(x)$ can be written as

$$g(x)\mu_\theta^g(x) = \exp\left[-\int_\theta^\eta \frac{y-\theta}{g(y)}\,dy\right]\exp\left[-\int_\eta^x \frac{y-\theta}{g(y)}\,dy\right], \tag{2.37}$$

where the first factor does not depend on x. It follows from the monotone convergence theorem that

$$\lim_{\theta\downarrow 0} \frac{\int_\eta^1 g(x)\mu_\theta^g(x)\,dx}{\int_0^\eta g(x)\mu_\theta^g(x)\,dx} = C_g(\eta) \tag{2.38}$$

with

$$C_g(\eta) = \frac{\int_\eta^1 \exp\left[-\int_\eta^x \frac{y}{g(y)}\,dy\right]dx}{\int_0^\eta \exp\left[\int_x^\eta \frac{y}{g(y)}\,dy\right]dx} \tag{2.39}$$

(cancel out the first factor of (2.37) before passing to the limit $\theta \downarrow 0$). Combining (2.36) and (2.38) with (2.8) we get (2.5).

The bounds on $C_g(\eta)$ in (2.6i) are obvious from (2.39). To get (2.6ii) note that $g(x) \leq x^2$ for $x \in [0,\eta)$ gives $C_g(\eta) \leq 1/\int_0^\eta (\frac{\eta}{x})\,dx = 0$.

2.6. *Proof of Theorem 5.* The proof will be by brute force. For notational convenience, let us write $x \longrightarrow \nu_\theta^g(x)$ to denote the density function of the probability measure ν_θ^g defined in (0.1). Recall from (2.7) that

$$\nu_\theta^g(x) = \frac{\mu_\theta^g(x)}{\int_0^1 \mu_\theta^g(x)}. \tag{2.40}$$

LEMMA 8. *For all $g \in \mathcal{H}$*
(i) $\theta \longrightarrow (Fg)(\theta)$ is C^∞ on $(0,1)$.
(ii) The k-th derivative has the representation

$$\left(\frac{d}{d\theta}\right)^k (Fg)(\theta) = \int_{[0,1]^{k+1}} \left\{\prod_{l=1}^{k+1} dx_l\,\nu_\theta^g(x_l)\right\} f\big((x_l)_{l=1}^{k+1}\big) \tag{2.41}$$

with

$$f\big((x_l)_{l=1}^{k+1}\big) = g(x_1)\prod_{n=1}^{k}\sum_{m=1}^{n}\int_{x_{n+1}}^{x_m}\frac{dz}{g(z)} \quad (k \geq 1). \tag{2.42}$$

(iii) For $\theta \in (0,1)$ there exists $R_\theta^g < \infty$ such that

$$\left|\left(\frac{d}{d\theta}\right)^k (Fg)(\theta)\right| \leq \|g\|_\infty (k!)^k (R_\theta^g)^{\frac{1}{2}k(k+1)} \quad (k \geq 1). \tag{2.43}$$

PROOF. (i) will follow once we have proved (ii) and (iii).

We begin with (ii). The proof is by induction and makes use of the following identity:

$$\frac{\partial}{\partial\theta}\nu_\theta^g(x) = \nu_\theta^g(x)\int_0^1 dy\,\nu_\theta^g(y)\int_y^x \frac{dz}{g(z)}. \tag{2.44}$$

This relation is a straightforward combination of (2.40) and the identity $\frac{\partial}{\partial\theta}\mu_\theta^g(x) = \mu_\theta^g(x)\int_\theta^x \frac{dz}{g(z)}$, which follows from (2.7).

To check (2.41) for $k = 1$, we compute using (2.44)

$$\begin{aligned}\frac{d}{d\theta}(Fg)(\theta) &= \frac{d}{d\theta}\int_0^1 g(x)\nu_\theta^g(x)\,dx \\ &= \int_0^1 dx\int_0^1 dy\,\nu_\theta^g(x)\nu_\theta^g(y)\left\{g(x)\int_y^x \frac{dz}{g(z)}\right\}.\end{aligned} \tag{2.45}$$

The induction step, on the other hand, is easily verified by differentiating (2.41), using (2.44) and the recursion relation

$$f\big((x_l)_{l=1}^{k+2}\big) = f\big((x_l)_{l=1}^{k+1}\big)\left\{\sum_{m=1}^{k+1}\int_{x_{k+2}}^{x_m}\frac{dz}{g(z)}\right\}. \tag{2.46}$$

Next we prove (iii). The difficulty to handle here is that $f\big((x_l)_{l=1}^{k+1}\big)$ diverges as one or more of its arguments tend to 0 or 1. Define

$$J_\theta^g(k) = \int_0^1\left|\int_{1/2}^x \frac{dz}{g(z)}\right|^k \nu_\theta^g(x)\,dx \quad (k \geq 0). \tag{2.47}$$

The key to (iii) is the following estimate

$$J_\theta^g(k) \leq k!\,(\hat{R}_\theta^g)^{k+1} \quad \text{for some } \hat{R}_\theta^g < \infty \text{ and all } k \geq 0, \tag{2.48}$$

which will be proved below. Continuing from (2.48), we have via (2.42) that

$$\big|f\big((x_l)_{l=1}^{k+1}\big)\big| \leq \|g\|_\infty \prod_{n=1}^{k}\sum_{m=1}^{n}\left\{\left|\int_{1/2}^{x_m}\frac{dz}{g(z)}\right| + \left|\int_{1/2}^{x_{n+1}}\frac{dz}{g(z)}\right|\right\}. \tag{2.49}$$

Substitution of (2.49) into (2.41) yields

$$\left|\left(\frac{d}{d\theta}\right)^k (Fg)(\theta)\right| \leq \|g\|_\infty \sum_{(p_l)_{l=1}^k,\,\sum_{l=1}^k p_l = k} \sigma\big((p_l)_{l=1}^k\big)\prod_{l=1}^k [J_\theta^g(l)]^{p_l}, \tag{2.50}$$

where $\sigma\big((p_l)_{l=1}^k\big) \geq 0$ are certain integer coefficients that add up to $\prod_{n=1}^k \sum_{m=1}^n 2 = 2^{k(k+1)/2}$. Since $[J_\theta^g(l)]^{1/l}$ is nondecreasing in l, the latter observation immediately gives

$$\left|\left(\frac{d}{d\theta}\right)^k (Fg)(\theta)\right| \leq \|g\|_\infty 2^{\frac{1}{2}k(k+1)}[J_\theta^g(k)]^k. \tag{2.51}$$

Together with (2.48) this completes (iii).

It remains to check (2.48), which goes as follows. First we show that

$$\int_0^{\theta/2}\left\{\int_x^{\theta/2}\frac{dz}{g(z)}\right\}^k \mu_\theta^g(x)\,dx \le k!\left(\frac{\theta}{2}\right)^{-(k+1)}. \tag{2.52}$$

Indeed, define

$$h_\theta(x)=\int_x^{\theta/2}\frac{dz}{g(z)}\quad\left(0\le x\le\frac{\theta}{2}\right). \tag{2.53}$$

One easily sees from (2.7) that

$$\mu_\theta^g(x)\le\left(-\frac{\partial}{\partial x}h_\theta(x)\right)\exp\left[-\frac{\theta}{2}h_\theta(x)\right]\quad\left(0\le x\le\frac{\theta}{2}\right). \tag{2.54}$$

Substitution of (2.54) into (2.52) gives

$$\begin{aligned}\text{l.h.s. (2.52)} &\le \int_0^{\theta/2}\{h_\theta(x)\}^k\left(-\frac{\partial}{\partial x}h_\theta(x)\right)\exp\left[-\frac{\theta}{2}h_\theta(x)\right]\\ &= \int_0^{\infty}u^k\exp\left[-\frac{\theta}{2}u\right]du = k!\left(\frac{\theta}{2}\right)^{-(k+1)}\end{aligned} \tag{2.55}$$

because $h_\theta(\theta/2)=0$ and $h_\theta(0)=\infty$ (recall (0.2iii)). This proves (2.52). A similar argument gives

$$\int_{\frac{1+\theta}{2}}^{1}\left\{\int_{\frac{1+\theta}{2}}^{x}\frac{dz}{g(z)}\right\}^k \mu_\theta^g(x)\,dx\le k!\left(\frac{1-\theta}{2}\right)^{-(k+1)}. \tag{2.56}$$

Combining now (2.52) with (2.56) and using that g is bounded away from 0 on $[\frac{\theta}{2},\frac{1+\theta}{2}]$, we get (2.48) for the integral in (2.47) but with ν_θ^g replaced by μ_θ^g. Finally, note that the denominator in (2.40) is finite on $(0,1)$. ■

2.7. *Proof of Lemma 1.* Because $K_{g,c}f=K_{\frac{1}{c}g,1}f$ it suffices to consider $c=1$. Abbreviate $K_g=K_{g,1}$. Recall (1.2) which reads

$$(K_gf)(\theta)=\int_0^1 f(x)\nu_\theta^g(x)\,dx. \tag{2.57}$$

Using (2.44) we obtain (compare with (2.41))

$$\begin{aligned}\frac{d}{d\theta}(K_gf)(\theta)&=\int_0^1 dx\int_0^1 dy\,\nu_\theta^g(x)\nu_\theta^g(y)\left\{f(x)\int_y^x\frac{dz}{g(z)}\right\}\\ &=\frac{1}{2}\int_0^1 dx\int_0^1 dy\,\nu_\theta^g(x)\nu_\theta^g(y)\big(f(x)-f(y)\big)\int_y^x\frac{dz}{g(z)}.\end{aligned} \tag{2.58}$$

Suppose that f has Lipschitz constant L, *i.e.*,

$$\left|\frac{f(x)-f(y)}{x-y}\right|\le L\quad\text{for all } x,y\in(0,1). \tag{2.59}$$

Then, since $(x-y)\int_y^x\frac{dz}{g(z)}\ge 0$, it follows from (2.58) that

$$\left|\frac{d}{d\theta}(K_gf)(\theta)\right|=\frac{1}{2}L\int_0^1 dx\int_0^1 dy\,\nu_\theta^g(x)\nu_\theta^g(y)(x-y)\int_y^x\frac{dz}{g(z)}. \tag{2.60}$$

We complete the proof by showing that the triple integral in (2.60) equals 2.

Split $x-y=(x-\theta)-(y-\theta)$ and use symmetry to write

(2.61)
$$\int_0^1 dx \int_0^1 dy\, \mu_\theta^g(x)\mu_\theta^g(y)(x-y)\int_y^x \frac{dz}{g(z)} = 2\int_0^1 dx\int_0^1 dy\,\{(x-\theta)\mu_\theta^g(x)\}\mu_\theta^g(y)\int_y^x \frac{dz}{g(z)}.$$

Substitute (2.9) to rewrite (2.61) as follows:

$$\begin{aligned}(2.62)\ 2\int_0^1 dx\int_0^1 dy&\Big\{-\frac{\partial}{\partial x}[g(x)\mu_\theta^g(x)]\Big\}\mu_\theta^g(y)\int_y^x\frac{dz}{g(z)}\\ &= 2\int_0^1 dy\,\mu_\theta^g(y)\left\{\left[-g(x)\mu_\theta^g(x)\int_y^x\frac{dz}{g(z)}\right]_{x=0}^1+\int_0^1 dx\, g(x)\mu_\theta^g(x)\frac{1}{g(x)}\right\}\\ &= 2\int_0^1 dx\,\mu_\theta^g(x)\int_0^1 dy\mu_\theta^g(y).\end{aligned}$$

In the last equality we use a slightly stronger version of Lemma 4 to get that the boundary terms at $x=0$ and $x=1$ vanish. This easily follows from the estimates in (2.52) and (2.56). Recall (2.40) to see that (2.62) proves the claim.

3. **Proof of Theorems 3 and 4.** The main step in the proof of Theorems 3 and 4 is the following lemma. For $g\in\mathcal{H}$ define

(3.1)
$$s(g)=\limsup_{x\downarrow 0} x^{-2}g(x)$$
$$i(g)=\liminf_{x\downarrow 0} x^{-2}g(x).$$

Define the map $T\colon [0,\infty)\longrightarrow[0,\infty]$ by (recall (2.3))

(3.2)
$$T(d)=\begin{cases}\frac{d}{1-d} & \text{if } 0\le d<1\\ \infty & \text{if } d\ge 1.\end{cases}$$

LEMMA 9. *For all* $g\in\mathcal{H}$
(i) $s(Fg)\le T\big(s(g)\big)$
(ii) $i(Fg)\ge T\big(i(g)\big)$.
In particular, if $s(g)=i(g)=d$ *then* $s(Fg)=i(Fg)=T(d)$.

PROOF. Let $f_d(d>0)$ be the function $f_d(x)=dx^2$. By writing the composition

(3.3)
$$\theta^{-2}(Fg)(\theta)=\theta^{-2}(Ff_d)(\theta)\frac{(G^\eta f_d)(\theta)}{(Ff_d)(\theta)}\frac{(G^\eta g)(\theta)}{(G^\eta f_d)(\theta)}\frac{(Fg)(\theta)}{(G^\eta g)(\theta)}$$

one can combine Propositions 6, 7 and Lemma 7 to get the following statement:

(3.4i)
$$\text{if } g\le f_d \text{ on } [0,\eta) \text{ then } s(Fg)\le T(d)\frac{1+C_g(\eta)}{1+C_{f_d}(\eta)}$$

(3.4ii)
$$\text{if } g\ge f_d \text{ on } [0,\eta) \text{ then } i(Fg)\ge T(d)\frac{1+C_g(\eta)}{1+C_{f_d}(\eta)}.$$

Here the inequalities come from the middle factor $(G^{\eta}g)(\theta)/(G^{\eta}f_d)(\theta)$ in (3.3), which by (2.4) and (2.38) has a lim sup ≤ 1 resp. a lim inf ≥ 1 as $\theta \downarrow 0$.

There are now four cases:

1. $0 \leq s(g) < 1$: Then there exist $\epsilon > 0$ and $\eta \in (0, 1)$ such that

$$
\begin{aligned}
& s(g) + \epsilon < 1 \\
& g \leq f_{s(g)+\epsilon} \text{ on } [0, \eta).
\end{aligned}
\tag{3.5}
$$

Therefore, by (3.4i) and (2.6ii), $s(Fg) \leq T\big(s(g) + \epsilon\big)$. Let $\epsilon \downarrow 0$.

2. $s(g) \geq 1$: Since now $T\big(s(g)\big) = \infty$, the claim is void.

3. $0 \leq i(g) < 1$: Since $T(0) = 0$ it suffices to consider $0 < i(g) < 1$. Then there exist $\epsilon > 0$ and $\eta \in (0, 1)$ such that

$$
\begin{aligned}
& i(g) - \epsilon > 0 \\
& g \geq f_{i(g)-\epsilon} \text{ on } [0, \eta).
\end{aligned}
\tag{3.6}
$$

Therefore, by (3.4ii) and (2.6ii), $i(Fg) \geq T(i(g) - \epsilon)[1 + C_g(\eta)] \geq T\big(i(g) - \epsilon\big)$. Let $\epsilon \downarrow 0$.

4. $i(g) \geq 1$: The same argument as in 3 gives

$$
i(Fg) \geq T\big(i(g) - \epsilon\big)\frac{1 + C_g(\eta)}{1 + C_{f_{i(g)-\epsilon}}(\eta)}. \tag{3.7}
$$

By (2.6i) the quotient is bounded below by η for all $\epsilon > 0$, and the limit is ∞ as $\epsilon \downarrow 0$. ■

We can now prove Theorems 3 and 4. First, to apply Lemma 9 to the transformation F_c (remember that at the beginning of Section 2 we had put $c = 1$), we use (0.19) which shows that Lemma 9 also holds with F, T replaced by F_c, T_c, where T_c is the map defined by

$$
T_c(d) = cT\Big(\frac{1}{c}\, d\Big) = \begin{cases} \frac{d}{1 - \frac{1}{c}d} & \text{if } 0 \leq d < c \\ \infty & \text{if } d \geq c. \end{cases} \tag{3.8}
$$

Next, putting (recall (0.17))

$$
T^{(n)} = T_{c_{n-1}} \circ \cdots \circ T_{c_0} \quad (n \geq 0), \tag{3.9}
$$

we get

$$T^{(n)}(d) = \frac{d}{1 - a_n d} \tag{3.10}$$

with a_n defined in (0.22). Now we argue:

(i) Suppose that $s(g) = d \leq a_\infty^{-1}$. Then by iteration of Lemma 9(i) we have $s(F^{(n)}g) \leq T^{(n)}(d) < \infty$ for all n. A similar result holds for the right boundary. Hence $g \in \mathcal{H}_2(a_\infty^{-1})$ implies $F^{(n)}g \in \cup_{d<\infty}\mathcal{H}_2(d)$. Since the latter class is disjoint from $\bar{\mathcal{H}} \supset \mathcal{H}^*$ we get Theorems 3(b),4(b).

(ii) Suppose that $i(g) = d > a_\infty^{-1}$ (and assume that the lim inf at the right boundary is at least as large). By iteration of Lemma 9(ii) we have $i(F^{(n)}g) \geq T^{(n)}(d)$ for all n, which becomes ∞ as soon as $a_n d \geq 1$. Now, the same monotonicity argument as in the proof of Lemma 9(ii) shows that

$$\begin{aligned} &\text{if } i(g) > c \text{ then } F_c g \in \bar{\mathcal{H}} \\ &\text{if } i(g) = c \text{ then } i(F_c g) = \infty \end{aligned} \tag{3.11}$$

(use Lemma 8 and the analogue of (2.3) for F_c). Hence $\bar{n}(g) < \infty$.

Finally, suppose that $s(g) = i(g) = d > a_\infty^{-1}$. Then $s(F^{(n)}g) = i(F^{(n)}g) = T^{(n)}(d)$ for all n. This proves (0.32) and completes the proof of Theorems 3(a), 4(a).

ACKNOWLEDGMENT. The research in this paper is supported by NATO-grant no. CRG920680. Part of the work was carried out at the Institut für Mathematische Stochastik of the Universität Göttingen and at the Institut de Mathématiques et Informatique of the Université de Lyon I.

REFERENCES

[B] L. Breiman, *Probability*, Addison-Wesley, Reading, 1968.

[CG] J. T. Cox and A. Greven, *On the long term behavior of finite particle systems: a critical dimension example*. In: Random Walks, Brownian Motion and Interacting Particle Systems, A Festschrift in Honor of Frank Spitzer, (eds. R. Durrett and H. Kesten), Progr. Probab. **28**, Birkhäuser, Boston, 1991, 203–213.

[CGS] J. T. Cox, A. Greven and T. Shiga, *Finite and infinite systems of interacting diffusions*, Probab. Theory Related Fields, to appear.

[DG1] D. A. Dawson and A. Greven, *Multiple time scale analysis of hierarchically interacting systems*. In: Stochastic Processes, A Festschrift in Honour of Gopinath Kallianpur, (eds. S. Cambanis, J. K. Ghosh, R. L. Karandikar and P. K. Sen), Springer, New York, 1993, 41–50.

[DG2] ———, *Multiple time scale analysis of interacting diffusions*, Probab. Theory Related Fields **95**(1993), 467–508.

[DG3] ———, *Hierarchical models of interacting diffusions: multiple time scale phenomena, phase transition and pattern of cluster-formation*, Probab. Theory Related Fields **96**(1993), 435–473.

[DGV] D. A. Dawson, A. Greven and J. Vaillancourt, *Equilibria and quasi-equilibria for infinite collections of interacting Fleming-Viot processes*, Trans. Amer. Math. Soc., to appear.

[D] R. Durrett, *Lecture Notes on Particle Systems and Percolation*, Wadsworth Brooks-Cole, Pacific Grove, 1988.

[FG] K. Fleischmann and A. Greven, *Diffusive clustering in an infinite system of hierarchically interacting diffusions*, IAAS Berlin, 1992, Probab. Theory Related Fields **98**(1994), 517–566.

[G] H.-O. Georgii, *Gibbs Measures and Phase Transitions*, Walter de Gruyter, Berlin, 1988.

[L] T. Liggett, *Interacting Particle Systems*, Springer, New York, 1985.

[OK] T. Ohta and M. Kimura, *A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a finite population*, Genet. Res. Camb. **22**(1973), 201–204.

[S] T. Shiga, *An interacting system in population genetics* J. Math. Kyoto Univ. **20**(1980), I: 213–243, II: 723–733.

[SF] S. Sawyer and J. Felsenstein, *Isolation by distance in a hierarchically clustered population*, J. Appl. Probab. **20**(1983), 1–10.

Institut de Mathématiques et Informatique
Université de Lyon I
43 Bd du 11 novembre 1918
F-69622 Villeurbanne Cedex
France

Faculteit der Technische Wiskunde en Informatica
Technische Universiteit Delft
Mekelweg 4
NL-2600 GA Delft
The Netherlands

Mathematisches Institut
Universität Erlangen-Nürnberg
Bismarckstrasse $1\frac{1}{2}$
D-91054 Erlangen
Germany

Mathematisch Instituut
Universiteit Nijmegen
Toernooiveld 1
NL-6525 ED Nijmegen
The Netherlands