Math. Proc. Camb. Phil. Soc. (2004), **137**, 725 © 2004 *Cambridge Philosophical Society*
DOI: 10.1017/S0305004104007765 *Printed in the United Kingdom*

The quantization error of self-similar distributions

By KLAUS PÖTZELBERGER

Department of Statistics, Vienna University of Economics and Business Administration, Augasse 2-6, *A*-1090 *Vienna, Austria.*
e-mail: Klaus.Poetzelberger@wu-wien.ac.at

(*Received* 12 *March* 2003)

Abstract

We prove results on the asymptotic behavior of the quantization error for self-similar distributions P when n, the number of prototypes, goes to infinity. In particular, we show that the suitably scaled quantization error converges, if P is nonarithmetic and that it is asymptotically periodic in the arithmetic case. The results hold also for distributions that are absolutely continuous with respect to a self-similar distribution.

1. *Introduction*

The aim of this paper is to prove a result on the asymptotic behavior of the quantization error for self-similar distributions.

Let an iterated function system (IFS), i.e. set of similarities $S_1, \dots, S_m$ on $\mathbb{R}^d$, be given. Denote by $\sigma_1, \dots, \sigma_m$ their scaling numbers, defined by

$$\|S_i(x) - S_i(y)\| = \sigma_i \|x - y\|.$$

We assume that the strong separation condition (SSC) holds, that is, $S_i(A) \cap S_j(A) = \varnothing$ for $i \neq j$, where A is the attractor, the unique compact set A with $A = S_1(A) \cup \dots \cup S_m(A)$.

Let $\pi = (\pi_1, \dots, \pi_m)$ with $\pi_i > 0$ and $\pi_1 + \dots + \pi_m = 1$. Each probability vector π defines a self-similar probability distribution P, a Borel probability on $\mathbb{R}^d$ satisfying

$$P = \sum_{i=1}^{m} \pi_i P \circ S_i^{-1}. \tag{1·1}$$

Let $n \in \mathbb{N}$ and $\mathcal{O} \subseteq \mathbb{R}^d$ with $|\mathcal{O}| = n$. The quantization error of $\mathcal{O}$ of order r is defined as

$$\Delta_{n,r}(P, \mathcal{O}) = \int \min_{p \in \mathcal{O}} \|x - p\|^r dP(x) \tag{1·2}$$

and

$$\Delta_{n,r}(P) := \inf_{|\mathcal{O}|=n} \Delta_{n,r}(P, \mathcal{O}) \tag{1·3}$$

is the quantization error of order r of the probability distribution P. A set $\mathcal{O}$ of size n is called optimal, if $\Delta_{n,r}(P, \mathcal{O}) = \Delta_{n,r}(P)$, its elements are called prototypes or

https://doi.org/10.1017/S0305004104007765 Published online by Cambridge University Press

principal points. The quantization dimension s_r of a distribution P is defined as

$$s_r = \lim_{n\to\infty} \frac{-r \log n}{\log \Delta_{n,r}(P)}. \tag{1·4}$$

Graf and Luschgy [**4**] prove that if the IFS satisfies the open set condition, the quantization dimension s_r of self-similar distributions P is the unique solution of

$$\sum_{i=1}^{m} \left(\pi_i \sigma_i^r\right)^{\frac{s_r}{r+s_r}} = 1. \tag{1·5}$$

If the IFS satisfies the strong separation condition, then a more accurate analysis of the asymptotic behavior of the quantization error is possible. It is easy to show that

$$\liminf_{n\to\infty} n^{r/s_r} \Delta_{n,r}(P) > 0 \tag{1·6}$$

and

$$\limsup_{n\to\infty} n^{r/s_r} \Delta_{n,r}(P) < \infty, \tag{1·7}$$

for s_r the solution of (1·5). The inequalities (1·6) and (1·7) thus identify the quantization dimension. The aim of this paper is to provide a result on what could be called the microstructure of the asymptotics of the scaled quantization errors.

Let $\pi^* = (\pi_1^*, \ldots, \pi_m^*)$ with

$$\pi_i^* = \left(\pi_i \sigma_i^r\right)^{\frac{s_r}{r+s_r}}. \tag{1·8}$$

π^* is a probability vector. A vector $x = (x_1, \ldots, x_m)$ with $x_i > 0$ is called arithmetic, if there is a $\theta > 0$ such that $x_i/\theta \in \mathbb{N}$. θ is not unique. It is unique if the greatest common divisor of $x_1/\theta, \ldots, x_m/\theta$ is 1. In this case θ is called the period of x. The following theorem is the main result of this paper.

THEOREM 1. *If* $(-\log \pi_1^*, \ldots, -\log \pi_m^*)$ *is arithmetic with period* $-\log \tau$, *then there exists a periodic function* $f : \mathbb{R}_+ \to \mathbb{R}_+$ *with period* $-\log \tau$, *such that*

$$\lim_{n\to\infty} \left(n^{r/s_r} \Delta_{n,r}(P) - f(\log n)\right) = 0. \tag{1·9}$$

If $(-\log \pi_1^*, \ldots, -\log \pi_m^*)$ *is nonarithmetic, then there exists a constant* $c(P) \in]0, \infty[$, *such that*

$$\lim_{n\to\infty} n^{r/s_r} \Delta_{n,r}(P) = c(P). \tag{1·10}$$

Existence and properties of the quantization constant $c(P)$ are a central theme in the theory of quantization of distributions, which are absolutely continuous with respect to d-dimensional Lebesgue measure. In contrast to that, there are only few results on the existence of quantization constants for singular distributions. Graf and Luschgy [**2**] show that the scaled quantization errors do not converge, if P is the Hausdorff measure on the Cantor set. The book [**3**] contains some results. For instance, for the covering radius, a limiting case for $r \to \infty$, the existence of a quantization constant $c(P)$ is proved for the nonarithmetic case. This result is given also in Falconer [**1**] as a result on the covering number of the attractor of an IFS.

The proof of Theorem 1 is given in Section 2. In Section 3 we generalize a result of Zador [**7**] on distributions G, which are absolutely continuous with respect to P. If π is nonarithmetic, then the limits of the scaled quantization errors factors into a

https://doi.org/10.1017/S0305004104007765 Published online by Cambridge University Press

product of the quantization constant $c(P)$ and a functional of the Radon–Nikodym derivative.

We recommend Graf and Luschgy [**3**] for information on problems concerning the existence of quantization constants and on the mathematical foundations of quantization in general.

2. *Limits of scaled quantization errors*

Throughout the paper we assume that the SSC holds. A always denotes the attractor.

The following simple observation is the basis of the identification of the quantization dimension and of the analysis of the quantization error of self-similar distributions. Let $\mathcal{O}$ be a set of prototypes of size n and let $\{V(p) \mid p \in \mathcal{O}\}$ denote the Voronoi partition corresponding to $\mathcal{O}$. It is easy to see (a proof is given in [**3**]) that if n is large enough, then for all $p \in \mathcal{O}$ there is exactly one $i \leqslant m$ such that $P(V(p) \cap S_i(A)) > 0$. Therefore, there exists a n_0 such that for $n \geqslant n_0$, $\mathcal{O} = \mathcal{O}_1 \cup \cdots \cup \mathcal{O}_m$, where $\mathcal{O}_i$ is the set of prototypes p for which $P(V(p) \cap S_i(A)) > 0$. Let $n_i = |\mathcal{O}_i|$. Then we get from (1·1)

$$\begin{aligned} n^{r/s_r}\Delta_{n,r}(P) &= \sum_{i=1}^{m} n^{r/s_r} \int_{\bigcup_{p\in\mathcal{O}_i} V(p)} \min_{p\in\mathcal{O}_i} \|x-p\|^r dP(x) \\ &= \sum_{i=1}^{m} n^{r/s_r}\pi_i\sigma_i^r\Delta_{n_i,r}(P). \end{aligned}$$

We abbreviate $n^{r/s_r}\Delta_{n,r}(P)$ by $\Psi(n)$ and have for $n \geqslant n_0$

$$\begin{aligned} \Psi(n) &= \sum_{i=1}^{m} \frac{n^{r/s_r}}{n_i^{r/s_r}}\pi_i\sigma_i^r\Psi(n_i) \\ &= \sum_{i=1}^{m} \left(\frac{\pi_i^*}{n/n_i}\right)^{r/s_r} \pi_i^*\Psi(n_i) \end{aligned}$$

and therefore

$$\Psi(n) = \min\left\{\sum_{i=1}^{m} \left(\frac{\pi_i^*}{n/n_i}\right)^{r/s_r} \pi_i^*\Psi(n_i) \mid n_i \geqslant 1, n_1+\cdots+n_m \leqslant n\right\}. \tag{2·1}$$

We will show that a function Ψ that satisfies equation (2·1) is asymptotically periodic if $-\log \pi^*$ is arithmetic and constant in the nonarithmetic case.

In order to analyze the asymptotics of a Ψ we approximate Ψ by a function that is defined on $\mathbb{R}_+$ and satisfies an equation analogous to (2·1).

Definition 1. Let $h\colon \mathbb{R}_+ \to \mathbb{R}_+$ and $x^* > 0$. We say that h satisfies the renewal equation with penalized optimization (REPO) on $[x^*, \infty[$, if for all $x \geqslant x^*$

$$h(x) = \inf\left\{\sum_{i=1}^{m} (\pi_i^*/q_i)^{r/s_r}\pi_i^* h(xq_i) \mid q_i > 0, q_1+\cdots+q_m \leqslant 1\right\}. \tag{2·2}$$

Theorem 2. *Let $h\colon \mathbb{R}_+ \to \mathbb{R}_+$ and $x^* > 0$ such that h is bounded on $]0, x^*]$ and satisfies the REPO-equation on $[x^*, \infty[$.*

https://doi.org/10.1017/S0305004104007765 Published online by Cambridge University Press

If $(-\log \pi_1^*, \dots, -\log \pi_m^*)$ *is arithmetic with period* $-\log \tau$, *then there exists a function* $f : \mathbb{R}_+ \to \mathbb{R}_+$, *which is periodic with period* $-\log \tau$, *such that*

$$\lim_{x\to\infty} (h(x) - f(\log x)) = 0. \tag{2·3}$$

If $(-\log \pi_1^*, \dots, -\log \pi_m^*)$ *is nonarithmetic, then the limit* $\lim_{x\to\infty} h(x)$ *exists.*

First, let us provide some useful properties and estimates concerning the function h. Without loss of generality, let $\pi_1^* \geqslant \pi_2^* \geqslant \cdots \geqslant \pi_m^*$.

LEMMA 1. *Let* $h_-, h_+ \in \mathbb{R}$ *such that* $0 < h_- \leqslant h(x) \leqslant h_+$ *for* $x \in]0, x^*]$. *Then* $0 < h_- \leqslant h(x) \leqslant h_+$ *for* $x \in \mathbb{R}_+$.

Proof. Let $x \in \mathbb{R}_+$ and $N \in \mathbb{N}$ with $x{\pi_1^*}^N \leqslant x^*$. For weights $q_i = \pi_i^*$ we get

$$\begin{aligned} h(x) &\leqslant \sum_{i=1}^m \pi_i^* h(\pi_i^* x) \\ &\leqslant \sup\{h(y) \mid y \leqslant x\pi_1^*\} \\ &\quad \cdots \\ &\leqslant \sup\left\{h(y) \mid y \leqslant x{\pi_1^*}^N\right\} \\ &\leqslant h_+. \end{aligned}$$

Since $\sum_{i=1}^m (\pi_i^*/q_i)^{r/s_r} \pi_i^* \geqslant 1$ we have for all weights q_i

$$\begin{aligned} \sum_{i=1}^m (\pi_i^*/q_i)^{r/s_r} \pi_i^* h(xq_i) &\geqslant \sum_{i=1}^m (\pi_i^*/q_i)^{r/s_r} \pi_i^* \inf\{h(y) \mid y \leqslant x\pi_1^*\} \\ &\geqslant \inf\{h(y) \mid y \leqslant x\pi_1^*\} \\ &\quad \cdots \\ &\geqslant \inf\left\{h(y) \mid y \leqslant x{\pi_1^*}^N\right\} \\ &\geqslant h_-. \end{aligned}$$

Let $\epsilon > 0$. Weights $q_1, \dots, q_m$ are called ϵ-optimal for x if

$$\sum_{i=1}^m (\pi_i^*/q_i)^{r/s_r} \pi_i^* h(q_i x) < h(x) + \epsilon.$$

They are called optimal for x, if they are ϵ-optimal for for all $\epsilon > 0$.

LEMMA 2. *Let* $\epsilon > 0$. *There exists a* $c > 0$ *such that* $q_i \leqslant c\pi_i^*$ *and* $1 - q_i \leqslant c(1 - \pi_i^*)$, *for all* x, *all weights* $q_1, \dots, q_m$ *that are* ϵ*-optimal for* x *and* $i \leqslant m$. *Moreover, there are constants* $0 < q_- \leqslant q_+ < \infty$, *such that for all* x *and all weight* $q_1, \dots, q_m$ *that are* ϵ*-optimal for* x *and all* $i \leqslant m$,

$$0 < q_- \leqslant q_i \leqslant q_+ < \infty. \tag{2·4}$$

By q_- and q_+ we denote real numbers that satisfy (2·4) for all x and a fixed $\epsilon > 0$.

Proof. Since

$$\sum_{i=1}^m (\pi_i^*/q_i)^{r/s_r} \pi_i^* h(q_i x) < h(x) + \epsilon \leqslant h_+ + \epsilon$$

https://doi.org/10.1017/S0305004104007765 Published online by Cambridge University Press

and

$$\sum_{i=1}^{m}(\pi_i^*/q_i)^{r/s_r}\pi_i^* h(q_i x) \geqslant \sum_{i=1}^{m}(\pi_i^*/q_i)^{r/s_r}\pi_i^* h_-$$

we have

$$\sum_{i=1}^{m}(\pi_i^*/q_i)^{r/s_r}\pi_i^* < \frac{h_+ + \epsilon}{h_-} < \infty.$$

Thus the weights q_i are uniformly bounded from 0 and from 1.

LEMMA 3. *If $x^* \leqslant x \leqslant y$, then*

$$h(y) \leqslant h(x)(y/x)^{r/s_r}. \tag{2·5}$$

Therefore $\limsup_{y\to x^+} h(y) \leqslant h(x)$ *and* $\liminf_{y\to x-} h(y) \geqslant h(x)$.

Proof. Let $x^* \leqslant x \leqslant y$, $\epsilon > 0$ and let $q_1, \dots, q_m$ ϵ-optimal for x. Since $\sum_i q_i x/y \leqslant 1$ we have

$$\begin{aligned} h(y) &\leqslant \sum_{i=1}^{m}\left(\frac{\pi_i^*}{q_i x/y}\right)^{r/s_r}\pi_i^* h(q_i(x/y)y) \\ &\leqslant (y/x)^{r/s_r}(h(x)+\epsilon). \end{aligned}$$

LEMMA 4. *Let $y > x^*$, $\epsilon > 0$ and let $q_1, \dots, q_m$ be ϵ-optimal for y. If $j \leqslant m$ is such that $y(1-q_j) > x^*$, then for $y(1-q_j) < x \leqslant y$*

$$(y/x)^{r/s_r}h(x) \leqslant h(y) + \epsilon + \left(\frac{\pi_j^*}{x/y - 1 + q_j}\right)^{r/s_r}\pi_j^* h(q_j y + x - y) - \left(\frac{\pi_j^*}{q_j}\right)^{r/s_r}\pi_j^* h(q_j y). \tag{2·6}$$

Thus for

$$\delta(y) := \limsup_{x\to y} |h(x) - h(y)| \tag{2·7}$$

we have the estimate

$$\delta(y) \leqslant \pi_1^* \sup\{\delta(x) \mid x \leqslant yq_+\}. \tag{2·8}$$

Proof. Define for x weights $\tilde{q}_i$ by $\tilde{q}_i = q_i y/x$ if $i \neq j$ and $\tilde{q}_j = 1 - (1-q_j)y/x$. Then

$$\begin{aligned} h(x) &\leqslant \sum_{i=1}^{m}(\pi_i^*/\tilde{q}_i)^{r/s_r}\pi_i^* h(\tilde{q}_i x) \\ &\leqslant (x/y)^{r/s_r}\left(h(y) + \epsilon + \left(\frac{\pi_j^*}{x/y - 1 + q_j}\right)^{r/s_r}\pi_j^* h(q_j y + x - y) - \left(\frac{\pi_j^*}{q_j}\right)^{r/s_r}\pi_j^* h(q_j y)\right). \end{aligned}$$

LEMMA 5. *If h is continuous on $]0, x^*]$, then h is continuous on $\mathbb{R}_+$. If h is differentiable on $]0, x^*]$, then h is differentiable on $\mathbb{R}_+$. If $|h'|$ is also bounded on $]0, x^*]$, then there exists a constant $c_1 > 0$ with*

$$c_1 \leqslant \max\{h_+ r/s_r, \sup\{|xh'(x)| \mid 0 < x \leqslant x^*\}\} \tag{2·9}$$

https://doi.org/10.1017/S0305004104007765 Published online by Cambridge University Press

such that for all $x \in \mathbb{R}_+$,

$$|xh'(x)| \leqslant c_1. \tag{2·10}$$

Proof. The inequalities (2·5) and (2·6) imply that h is continuous, if it is on $]0, x^*]$ and that it is differentiable if it is on $]0, x^*]$. A constant c_1 exists, for which (2·10) holds, if $|h'|$ is bounded on $]0, x^*]$: From (2·5) and (2·6) we see that for any $t > 0$,

$$\sup\{|xh'(x)| \mid x \in]0, t]\} \leqslant \max\{h_+ r/s_r, \sup\{|xh'(x)| \mid x \in]0, tq_+]\}\},$$

so that (2·9) holds on $]0, t]$ if it holds on $]0, tq_+]$ and $c_1 \geqslant h_+ r/s_r$.

Proof of Theorem 2, *arithmetic case.* Let $\pi_i^* = \tau^{n_i}$ where $n_i \in \mathbb{N}$ and the g.c.d. of $n_1, \ldots, n_m$ is 1. W.l.g. let $n_1 \leqslant n_2 \leqslant \cdots \leqslant n_m$.

We define for $v \in \mathbb{N}$ Markov kernels M_v on $\{0, \ldots, n_m - 1\} \times \{vn_m, \ldots, (v+1)n_m - 1\}$. Let $\tilde{M}_v$ denote the Markov kernel on $\{1, \ldots, (v+1)n_m\} \times \{1, \ldots, (v+1)n_m\}$ defined by $\tilde{M}_v(i, j) = \pi^*_{j-i}$ if $i \leqslant vn_m$ and $j - i \in \{n_1, \ldots, n_m\}$, $\tilde{M}_v(i, j) = 1$ if $i > vn_m$ and $i = j$ and $\tilde{M}_v(i, j) = 0$ in the remaining cases. According to this kernel a particle moves from i to j with probability π^*_{j-i} until it reaches $\{vn_m + 1, \ldots, (v+1)n_m\}$. The states $j \in \{vn_m + 1, \ldots, (v+1)n_m\}$ are absorbing. There is a minimal k such that for all $i \leqslant n_m$ and all $j \leqslant vn_m$, $\tilde{M}_v^k(i, j) = 0$, i.e. where all particles starting in $\{1, \ldots, n_m\}$ have reached $\{vn_m + 1, \ldots, (v+1)n_m\}$ with probability 1. We define $M_v(i, j) = \tilde{M}_v^k(i + 1, j + 1)$. Note that $M_v(i, j)$ is a product of elements of $\{\pi_1^*, \ldots, \pi_m^*\}$ if $M_v(i, j) > 0$.

From the construction of $\tilde{M}_v$ and M_v, we know that for all $i = 0, \ldots, n_m - 1$,

$$h(x) \leqslant \sum_{j=vn_m}^{(v+1)n_m - 1} M_v(i, j) h(xM_v(i, j)). \tag{2·11}$$

Let $\hat{x} \geqslant x^*$ arbitrary. We define a sequence of mappings $f_k : [1, 1/\tau[\to \mathbb{R}$ by

$$f_k(x) = \max\{h(x\hat{x}\tau^{-n_m k + i}) \mid i = 0, \ldots, n_m - 1\}.$$

Let $f_k(x) = h(x\hat{x}\tau^{-n_m k + i_k})$ with $i_k = i_k(x) \in \{0, \ldots, n_m - 1\}$. $(f_k(x))$ is nonincreasing (in k), since

$$\begin{aligned}
f_k(x) &= h(x\hat{x}\tau^{-n_m k + i_k}) && (2·12)\\
&\leqslant \sum_{j=vn_m}^{(v+1)n_m - 1} M_v(i_k, j) h(x\hat{x}\tau^{-n_m k + i_k} M_v(i_k, j)) && (2·13)\\
&= \sum_{j=vn_m}^{(v+1)n_m - 1} M_v(i_k, j) h(x\hat{x}\tau^{-n_m k + i_k + j - i_k}) && (2·14)\\
&= \sum_{j=vn_m}^{(v+1)n_m - 1} M_v(i_k, j) h(x\hat{x}\tau^{-n_m (k-v) + j - n_m v}) && (2·15)\\
&\leqslant \sum_{j=vn_m}^{(v+1)n_m - 1} M_v(i_k, j) f_{k-v}(x) && (2·16)\\
&= f_{k-v}(x). && (2·17)
\end{aligned}$$

https://doi.org/10.1017/S0305004104007765 Published online by Cambridge University Press

Let $\tilde{f}(x) := \lim_{k\to\infty} f_k(x)$. By definition,

$$\tilde{f}(x) = \lim_{k\to\infty} h(x\hat{x}\tau^{-n_m k+i_k}). \tag{2·18}$$

We show that (2·18) holds if we replace (i_k) by any sequence $(\tilde{i}_k)$ with $\tilde{i}_k \in \{0,\dots,n_m-1\}$. From (2·12) – (2·17) we infer that

$$\sum_{j=vn_m}^{(v+1)n_m-1} M_v(i_k,j)(f_{k-v}(x) - h(xM_v(i_k,j))) \leqslant f_{k-v}(x) - f_k(x) \leqslant f_{k-v}(x) - \tilde{f}(x).$$

Therefore, if $M_v(i_k,j) > 0$,

$$\tilde{f}(x) - h(xM_v(i_k,j)) \leqslant \frac{1}{M_v(i_k,j)}(f_{k-v}(x) - \tilde{f}(x)). \tag{2·19}$$

Thus for any sequence (i_k) for which (2·18) holds, then (2·18) holds also for the sequence $(i_k - j_k)$ if $M_v(i_k,j_k) > 0$.

The important step is to prove that there exists a $v \in \mathbb{N}$ such that for all $i \in \{0,\dots,n_m-1\}$ and all $j \in \{vn_m,\dots,(v+1)n_m-1\}$, $M_v(i,j) > 0$. Here we need the assumption that the greatest common divisor of $\{n_1,\dots,n_m\}$ is 1. Consider the semigroup $\{0,1,\dots,n_m-1\}$ with addition modulo n_m. This semigroup is then generated by $\{0,n_1,n_2,\dots,n_{m-1}\}$. Therefore, for all $i,j \in \{0,1,\dots,n_m-1\}$ there is a $v \in \mathbb{N}$ with $M_v(i,vn_m+j) > 0$.

Note that since $\pi_m = \tau^{n_m} > 0$, $M_v(i,vn_m+j) > 0$ implies $M_{v+1}(i,(v+1)n_m+j) > 0$. Therefore, for all i there is a v_i such that for all $j \in \{0,1,\dots,n_m-1\}$, $M_{v_i}(i,v_in_m+j) > 0$. Let v denote the smallest common multiple of $v_0,\dots,v_{n_m-1}$. Then for all $i \in \{0,\dots,n_m-1\}$ and all $j \in \{vn_m,\dots,(v+1)n_m-1\}$, $M_v(i,j) > 0$.

Thus (2·18) holds for all sequences (i_k) with $i_k \in \{0,\dots,n_m-1\}$.

$\tilde{f}$ is the pointwise limit of (f_k). The convergence is uniform on compact subsets of $[1,1/\tau[$, since (f_k) is nonincreasing. As $\hat{x}$ is arbitrary, we can show that (f_k) converges to $\tilde{f}$ uniformly on $[1,1/\tau[$.

To complete the proof of the nonarithmetic case, let $[y]$ denote the greatest integer less or equal than y. W.l.g. let $\hat{x} = 1$. We define

$$f(x) = \tilde{f}\big(\tau^{[-x/\log\tau]}e^x\big). \tag{2·20}$$

Then for $n = n(x) := [-\log x/\log\tau]$,

$$\begin{aligned}\lim_{x\to\infty} |h(x) - f(\log x)| &= \lim_{x\to\infty} |h(x) - \tilde{f}(x\tau^n)| \\ &= \lim_{x\to\infty} |h(x\tau^n\tau^{-n}) - \tilde{f}(x\tau^n)| \\ &\leqslant \lim_{n\to\infty} \sup_{y\in[1,1/\tau[} |h(y\tau^{-n}) - \tilde{f}(y)| \\ &= 0.\end{aligned}$$

Nonarithmetic case. The idea of the proof of the nonarithmetic case is similar to the arithmetic case. Let $\pi_1^* \geqslant \cdots \geqslant \pi_m^*$ and let $\hat{x} > 0$ be arbitrary. We define $I_k = [\hat{x}(1/\pi_m^*)^{k-1}, \hat{x}(1/\pi_m^*)^k[$ and $f_k = \sup\{h(x) \mid x \in I_k\}$. Furthermore, we define Markov kernels $(\mu_v(x,j))_{j=1}^{\ell_x^v}$ analogous to the arithmetic case: A particle starts in $x \in I_k$ and moves from y to $y\pi_i^*$ with probability π_i^* until it reaches I_{k-v}. I_{k-v} is

https://doi.org/10.1017/S0305004104007765 Published online by Cambridge University Press

absorbing. The states in I_{k-v} reached by the particle are of the form $x\mu_v(x,j)$ with probability $\mu_v(x,j)$. ℓ_x^v is the number of states in I_{k-v} reached from x with positive probability. There exists a number $a>0$ such that for all v,k and $x\in I_k$, $\ell_x^v \leqslant a^v$.

Since $\mu_v(x,j)$ is a product of elements of $\{\pi_1^*,\dots,\pi_m^*\}$,

$$h(x) \leqslant \sum_{j=1}^{\ell_x^v} \mu_v(x,j) h(x\mu_v(x,j)), \tag{2·21}$$

and thus $f_k \leqslant f_{k-v}$. Let $f_\infty := \lim_{k\to\infty} f_k$. Again, if (x_k) is a sequence with $x_k\in I_k$ and $\lim_{k\to\infty} h(x_k) = f_\infty$, then

$$\lim_{k\to\infty} h(\mu_v(x_k,j)x_k) = f_\infty \tag{2·22}$$

for all choices of $j\leqslant \ell_{x_k}^v$.

Since $(-\log\pi_1^*,\dots,-\log\pi_m^*)$ is nonarithmetic, the sets $\{x\mu_v(x_k,j)\,|\,j\leqslant \ell_{x_k}^v\}$ become dense in the following sense. For all $\epsilon>0$ there exists a $v\in\mathbb{N}$ such that for all $k>v$ and all $y\in I_{k-v}$ with $y+\epsilon y\in I_{k-v}$ there exists a $j\leqslant \ell_{x_k}^v$ such that

$$y < x_k\mu_v(x_k,j) < y+\epsilon y.$$

Then, by Lemma 3 we have for such y

$$h(y) \leqslant (1+\epsilon)^{r/s_r} h(x_k\mu_v(x_k,j)).$$

Now suppose there exist an $\epsilon'>0$ and a sequence (y_i) with $y_i\to\infty$ such that $h(y_i) < f_\infty - \epsilon'$. Then for all $\hat{x}$ there are $k_i\in\mathbb{N}$ and $\epsilon_i>0$ with $\epsilon_i\to 0$, such that

$$y_i < \hat{x}(1/\pi_m^*)^{k_i} \leqslant y_i(1+\epsilon_i). \tag{2·23}$$

However, (2·23) is not possible: Choose $\hat{x}_1$ and $\hat{x}_2$ with $\hat{x}_2/\hat{x}_1 = \sqrt{\pi_m^*}$. Let $k_{j,i}\in\mathbb{N}$ and $\epsilon_{j,i}$ correspond to $\hat{x}_j$ and assume that $1+\epsilon_{j,i} < (1/\pi_m^*)^{1/4}$. Then

$$\begin{aligned} 1 &= y_i/y_i \\ &\leqslant \frac{\hat{x}_1(1/\pi_m^*)^{k_{1,i}}}{\hat{x}_2(1/\pi_m^*)^{k_{2,i}}/(1+\epsilon_{2,i})} \\ &\leqslant {\pi_m^*}^{k_{2,i}-k_{1,i}-3/4} \end{aligned}$$

and

$$\begin{aligned} 1 &= y_i/y_i \\ &\geqslant \frac{\hat{x}_1(1/\pi_m^*)^{k_{1,i}}/(1+\epsilon_{1,i})}{\hat{x}_2(1/\pi_m^*)^{k_{2,i}}} \\ &\geqslant {\pi_m^*}^{k_{2,i}-k_{1,i}-1/4}. \end{aligned}$$

Thus $1/4\leqslant k_{2,i}-k_{1,i}\leqslant 3/4$, contradicting $k_{2,i}-k_{1,i}\in\mathbb{N}$.

Remark 1. Note that Theorem 2 has been stated in the form that is most appropriate for the analysis of the quantization error. We can prove a more general version of the REPO-equation (2·2) exactly in the same way. Let the penalty functions $q\mapsto(\pi_i^*/q)^{r/s_r}$ be replaced by general penalty functions $\varphi_i,\dots,\varphi_m$: Let $\mathcal{Q} = \{(q_1,\dots,q_m)\,|\,q_i>0, q_1+\dots+q_m\leqslant 1\}$ and let $\varphi_i :]0,1[\to\mathbb{R}_+$ satisfy $\varphi_1(q_1)\pi_1^*+\dots+\varphi_m(q_m)\pi_m^*\geqslant 1$ for $(q_1,\dots,q_m)\in\mathcal{Q}$ and $\varphi_1(q_1)\pi_1^*+\dots+\varphi_m(q_m)\pi_m^* = 1$

https://doi.org/10.1017/S0305004104007765 Published online by Cambridge University Press

if and only if $(q_1, \dots, q_m) = (\pi_1^*, \dots, \pi_m^*)$. Then Theorem 2 holds if regularity conditions hold. An example of such a set of conditions is: $\varphi_i(q)$ is unbounded for $q \to 0$ and for $q \to 1$ and $q \mapsto \varphi_i(q)$ is continuous at π_i^*, uniformly in $q \in [a, b]$ for all $0 < a < b < 1$.

Let h be continuous and satisfy the relation (2·2) for $x > x^*$. We define a function $\beta : \mathbb{R}_+ \to \mathbb{R}_+$ recursively as $\beta(x) = 1$ if $x \leqslant x^*$ and for $x > x^*$ by

$$\beta(x) = \sup \left\{ \sum_{i=1}^{m} (\pi_i^*/q_i)^{r/s_r} \pi_i^* \beta(q_i x) \mid q_1, \dots, q_m \right\}, \tag{2·24}$$

where the supremum is taken over all weights $q_1, \dots, q_m$ that are optimal for x. Note that there is a constant $\beta_\infty \leqslant h_+/h_-$ such that for all x, $1 \leqslant \beta(x) \leqslant \beta_\infty$.

For the proof of Theorem 1 we need the following lemma on the approximation of Ψ by a function h.

Lemma 6. *Let $n^* \leqslant \hat{n} \in \mathbb{N}$ and $c_1 > 0$, such that Ψ satisfies the relation* (2·1) *for $n \geqslant n^*$, h is differentiable with $x|h'(x)| \leqslant c_1$ for $x > 0$ and satisfies* (2·2) *for $x \geqslant \hat{n}$. Furthermore, assume that $\hat{n} \geqslant 2/q_-$.*

If $h(n) = \Psi(n)$ for all $n \leqslant \hat{n}$, then there is a constant c_6 such that for all $n \in \mathbb{N}$,

$$\Psi(n) - h(n) \leqslant \frac{c_6}{\hat{n}}. \tag{2·25}$$

The constant c_6 depends on $\pi_1^*, \dots, \pi_m^*$, h_+, h_-, c_1 and $\Psi_+ := \sup\{\Psi(n) \mid n \in \mathbb{N}\}$ only.

Proof. Note that $h(n) \leqslant \Psi(n)$ for all $n \in \mathbb{N}$. Let for $x \in \mathbb{R}_+$

$$\alpha(x) = \max\{\Psi(n) - h(n) \mid n \leqslant x, n \in \mathbb{N}\}.$$

For $k \in \mathbb{N}$, let $I_k =]\hat{n} {q_+}^{-(k-1)}, \hat{n} {q_+}^{-v}]$. We prove that there exists a constant c_5, such that for all k and $x \in I_k$,

$$\alpha(x) \leqslant \frac{\beta(x)}{\hat{n}} \frac{1 - {q_+}^{k+1}}{1 - q_+} c_5. \tag{2·26}$$

(2·26) implies (2·25) with

$$c_6 = \frac{\beta_\infty}{\hat{n}} \frac{c_5}{1 - q_+}.$$

First note that there exists a positive constant c_4, such that for all k and $x \in I_k$,

$$\left(\frac{q_i x}{q_i x - 1} \right)^{r/s_r} - 1 \leqslant \frac{c_4 {q_+}^k}{\hat{n}}. \tag{2·27}$$

We define

$$c_5 := \max \left\{ \Psi_+ c_4 + \frac{2c_1}{q_+ q_-}, \sup\{\alpha(x)\hat{n}/(1 + q_+)\beta(x) \mid x \in I_1\} \right\}. \tag{2·28}$$

We prove (2·26) by induction on k. Since by construction,

$$c_5 \geqslant \alpha(x)\hat{n}/(1 + q_+)\beta(x)$$

for $x \in I_1$, (2·26) holds for $k = 1$.

https://doi.org/10.1017/S0305004104007765 Published online by Cambridge University Press

Suppose, (2·26) holds for all $k' < k$ and all $x \in I_{k'}$. Let $x \in I_k$ and let $q_1, \dots, q_m$ denote optimal weights for x. Let $n = [x]$, $n_i = [q_i n]$ and $\tilde{q}_i = n_i/n$. Then

$$\begin{aligned}\Psi([x]) - h(x) &\leqslant \sum_{i=1}^{m} (\pi_i^*/\tilde{q}_i)^{r/s_r} \pi_i^* \Psi(n_i) - h(x) \\ &= A + B + C\end{aligned}$$

with

$$\begin{aligned}A &= \sum_{i=1}^{m} \left((\pi_i^*/\tilde{q}_i)^{r/s_r} - (\pi_i^*/q_i)^{r/s_r} \right) \pi_i^* \Psi(n_i), \\ B &= \sum_{i=1}^{m} (\pi_i^*/q_i)^{r/s_r} \pi_i^* \left(h(n_i) - h(q_i x) \right), \\ C &= \sum_{i=1}^{m} (\pi_i^*/q_i)^{r/s_r} \pi_i^* \left(\Psi(n_i) - h(n_i) \right).\end{aligned}$$

We have

$$\begin{aligned}A &= \sum_{i=1}^{m} (\pi_i^*/q_i)^{r/s_r} \pi_i^* \Psi(n_i) \left(\left(\frac{q_i}{\tilde{q}_i} \right)^{r/s_r} - 1 \right) \\ &\leqslant \sum_{i=1}^{m} (\pi_i^*/q_i)^{r/s_r} \pi_i^* \Psi_+ \left(\left(\frac{q_i x}{q_i x - 1} \right)^{r/s_r} - 1 \right) \\ &\leqslant \sum_{i=1}^{m} (\pi_i^*/q_i)^{r/s_r} \pi_i^* \Psi_+ \frac{c_4 {q_+}^k}{\hat{n}} \\ &\leqslant \beta(x) \Psi_+ \frac{c_4 {q_+}^k}{\hat{n}}.\end{aligned}$$

An upper bound for B is

$$\begin{aligned}B &\leqslant \sum_{i=1}^{m} (\pi_i^*/q_i)^{r/s_r} \pi_i^* (q_i x - [q_i x]) \sup \{ |h'(y)| \mid [q_i x] \leqslant y \leqslant q_i x \} \\ &\leqslant \sum_{i=1}^{m} (\pi_i^*/q_i)^{r/s_r} \pi_i^* \frac{c_1}{[q_i x]} \\ &\leqslant \sum_{i=1}^{m} (\pi_i^*/q_i)^{r/s_r} \pi_i^* \frac{c_1}{q_i x - 1} \\ &\leqslant \sum_{i=1}^{m} (\pi_i^*/q_i)^{r/s_r} \pi_i^* \frac{c_1}{q_- \hat{n}/{q_+}^{(k-1)} - 1} \\ &= \sum_{i=1}^{m} (\pi_i^*/q_i)^{r/s_r} \pi_i^* \frac{c_1 {q_+}^k}{\hat{n}} \frac{1}{q_- q_+ - {q_+}^k/\hat{n}} \\ &\leqslant \beta(x) \frac{c_1 {q_+}^k}{\hat{n}} \frac{2}{q_- q_+},\end{aligned}$$

since $\hat{n} \geqslant 2/q_-$.

https://doi.org/10.1017/S0305004104007765 Published online by Cambridge University Press

Finally $q_i x \in \bigcup_{k'<k} I_{k'}$ implies

$$\begin{aligned} C &\leqslant \sum_{i=1}^{m} (\pi_i^*/q_i)^{r/s_r} \pi_i^* \alpha(q_i x) \\ &\leqslant \sum_{i=1}^{m} (\pi_i^*/q_i)^{r/s_r} \pi_i^* \frac{\beta(q_i x)}{\hat{n}} \frac{1-{q_+}^k}{1-q_+} c_5 \\ &= \beta(x) \frac{c_5}{\hat{n}} \frac{1-{q_+}^k}{1-q_+}. \end{aligned}$$

Combining these three estimates gives

$$\begin{aligned} \Psi([x]) - h(x) &\leqslant \frac{\beta(x)}{\hat{n}} \left({q_+}^k \left(\Psi_+ c_4 + \frac{2c_1}{q_- q_+} \right) + \frac{1-{q_+}^k}{1-q_+} c_5 \right) \\ &\leqslant \frac{\beta(x)}{\hat{n}} \left({q_+}^k c_5 + \frac{1-{q_+}^k}{1-q_+} c_5 \right) \\ &= \frac{\beta(x)}{\hat{n}} \frac{1-{q_+}^{k+1}}{1-q_+} \quad c_5. \end{aligned}$$

Proof of Theorem 1. Let Ψ satisfy the relation (2·1) for $n > n^*$. For all $\hat{n} > n^*$ we choose functions $h_{\hat{n}}$ that satisfy $h_{\hat{n}}(n) = \Psi(n)$ for all $n \leqslant \hat{n}$ and (2·2) for $x > \hat{n}$. These functions are chosen such that

$$h_{\hat{n}+1}(x) = h_{\hat{n}}(x)$$

for $x \leqslant \hat{n}$ and

$$h_{\hat{n}+1}(x) \geqslant h_{\hat{n}}(x)$$

for $\hat{n} \leqslant x \leqslant \hat{n}+1$. They may be chosen to be differentiable with $x|h'(x)|$ uniformly bounded by a constant c_1. Note that $h_{\hat{n}+1}(x) \geqslant h_{\hat{n}}(x)$ for all $x > 0$ and $\Psi(n) \geqslant h_{\hat{n}+1}(n) \geqslant h_{\hat{n}}(n)$ for all $n \in \mathbb{N}$ follows.

By Lemma 6 there exists a constant c_6 for which (2·25) holds uniformly, i.e.

$$\Psi(n) - h_{\hat{n}}(n) \leqslant c_6/\hat{n}.$$

Let $(-\log \pi_1^*, \ldots, -\log \pi_m^*)$ be arithmetic with period $-\log \tau$. Let $f_{\hat{n}}$ denote the $-\log \tau$-periodic limit of $h_{\hat{n}}$ in the sense of (2·3). For fixed x, $f_{\hat{n}}(x)$ is nondecreasing. Let $f(x) := \lim_{\hat{n}\to\infty} f_{\hat{n}}(x)$. f is periodic with period $-\log \tau$. Since $f_{\hat{n}} \leqslant f$, we have for all $\hat{n}$,

$$\begin{aligned} \limsup_{n\to\infty} |\Psi(n) - f(\log n)| &= \limsup_{n\to\infty} (\Psi(n) - f(\log n)) \\ &\leqslant \limsup_{n\to\infty} (\Psi(n) - h_{\hat{n}}(n)) \\ &\quad + \limsup_{n\to\infty} (h_{\hat{n}}(n) - f_{\hat{n}}(\log n)) \\ &\quad + \limsup_{n\to\infty} (f_{\hat{n}}(\log n) - f(\log n)) \\ &\leqslant c_6/\hat{n}. \end{aligned}$$

In the nonarithmetic case the functions $f_{\hat{n}}$ are constant and thus f is constant.

https://doi.org/10.1017/S0305004104007765 Published online by Cambridge University Press

3. *Absolutely continuous distributions*

Let G be absolutely continuous with respect to the self-similar distribution P. Furthermore, for $r \geqslant 1$, let P^* denote the self-similar distribution defined by the probability vector $\pi^* = (\pi_1^*, \dots, \pi_m^*)$.

To generalize Theorem 1 for distributions, which are absolutely continuous with respect to Lebesgue measure, we have to discuss the problem of choosing a suitable Radon–Nikodym derivative dG/dP. According to Theorem 1 we expect that the scaled quantization error $n^{r/s_r}\Delta_{n,r}(G)$ is asymptotically periodic, if $-\log \pi^* = (-\log \pi_1^*, \dots, -\log \pi_m^*)$ is arithmetic and converges to a constant if $-\log \pi^*$ is nonarithmetic. Let for $k \in \mathbb{N}$ and $i = (i_1, \dots, i_k) \in \{1, \dots, m\}^k$,

$$A_{k,i} = S_{i_k} \circ \cdots \circ S_{i_1}(A). \tag{3·1}$$

Suppose there is a $k \in \mathbb{N}$ such that $g = dG/dP$ is constant on the sets $A_{k,i}$. Then a simple argument of optimal allocation of n prototypes to the m^k sets $A_{k,i}$ identifies the limit (3·8), for instance, if $-\log \pi^*$ is nonarithmetic. Let D denote the similarity dimension of the IFS. Except when $\pi_i = \sigma_i^D$, $i = 1, \dots, m$, P and P^* are singular. We even have $\dim_{\mathcal{H}}(P) < s_r < \dim_{\mathcal{H}}(P^*)$. Thus (3·8) can hold only for certain versions of dG/dP in the general case. Therefore, we have to choose a specific Radon–Nikodym derivative g and prove continuity of limits of scaled quantization errors for sequences of piecewise constant derivatives g_k approximating g.

Let

$$g_k(x) = \sum_{i \in \{1,\dots,m\}^k} I_{A_{k,i}}(x) \frac{G(A_{k,i})}{P(A_{k,i}) + P^*(A_{k,i})}. \tag{3·2}$$

$(g_k)_{k=1}^{\infty}$ is a positive martingale with respect to $\tilde{P} = 1/2P + 1/2P^*$ and thus convergent. I.e. there exists a function g such that $g_k \to g$ $\tilde{P}$-a.e. g is a version of dG/dP is $P \neq P^*$. If $P = P^*$, then $2g$ is a derivative. By Scheffé's lemma (g_k) converges to g also in the $L_1(\tilde{P})$-norm and therefore both in $L_1(P)$ and $L_1(P^*)$-norm. Stronger assumptions are necessary to prove the continuity of the quantization errors.

Let for a Borel set B with $G(B) > 0$, $G_{\restriction B}$ denote the conditional distribution of G given B, i.e. $G_{\restriction B}(E) = G(E \cap B)/G(B)$.

LEMMA 7. *Let $\pi_m^* \leqslant \cdots \leqslant \pi_1^*$. Assume there exists a positive Borel measure Q, such that for all $k \in \mathbb{N}$ and $i \in \{1, \dots, m\}^k$,*

$$\frac{G(A_{k,i})P^*(A_{k,i})}{P(A_{k,i})} \leqslant Q(A_{k,i}). \tag{3·3}$$

Then there exists a constant $b > 0$ such that for all Borel sets B with $G(B) > 0$

$$\limsup_{n\to\infty} n^{r/s_r}\Delta_{n,r}(G_{\restriction B}) \leqslant b\frac{Q(B)}{G(B)}. \tag{3·4}$$

An upper bound for b is $diam(A)^r {\pi_m^*}^{-r/s_r}$, where $diam(A)$ denotes the diameter of the attractor A.

Remark 2. We consider two cases. First, let $P = P^*$. Then we can choose $Q = G$ to prove that $n^{r/s_r}\Delta_{n,r}(G_{\restriction B})$ is uniformly bounded in B with $G(B) > 0$.

https://doi.org/10.1017/S0305004104007765 Published online by Cambridge University Press

If $P \neq P^*$ but dG/dP is bounded by a constant g_∞, we have to choose $Q = g_\infty P^*$ to get

$$\limsup_{n\to\infty} n^{r/s_r}\Delta_{n,r}(G_{\restriction B}) \leqslant b g_\infty \frac{P^*(B)}{G(B)}. \tag{3·5}$$

Proof of Lemma 7. Let $\epsilon > 0$. There is a $k \in \mathbb{N}$ and a set of multiindices $J_k \subseteq \{1,\dots,m\}^k$, such that $B \subseteq \tilde{B} := \bigcup_{i\in J_k} A_{k,i}$ and $Q(\tilde{B}) \leqslant Q(B)(1+\epsilon)$. Let n be large enough such that for all $i \in J_k$, $nP^*(A_{k,i}) \geqslant 1$. We split sets $A_{k,i}$ into $S_j(A_{k,i})$, $j = 1,\dots,m$ if $nP^*(S_m(A_{k,i})) \geqslant 1$ and proceed by splitting the obtained sets of the form $A_{k+1,i'}$ until we get components $A_{k',i'}$ with $nP^*(S_m(A_{k',i'})) < 1$. There are a $k_1 \geqslant k$ and sets of multiindices $J_{k'} \subseteq \{1,\dots,m\}^{k'}$, such that

$$\tilde{B} = \bigcup_{k'=k}^{k_1} \bigcup_{i\in J_{k'}} A_{k',i}.$$

If i is a multiindex in $J_{k'}$, we set $n_i = [nP^*(A_{k',i})]$ and choose a set $\mathcal{O}_i$ of n_i elements of $A_{k',i}$. Note that $\mathcal{O}$, the union of all $\mathcal{O}_i$, has size at most n. Furthermore, since the process of splitting components has stopped in $A_{k',i}$ and $P^*(S_m(A_{k',i})) = \pi_m^* P^*(A_{k',i})$, we have

$$nP^*(A_{k',i}) < 1/\pi_m^*. \tag{3·6}$$

For a multiindex $i = (i_1,\dots,i_{k'})$ we define $\sigma_i = \sigma_{i_1}\cdots\sigma_{i_{k'}}$. Then we have $\operatorname{diam}(A_{k',i}) = \sigma_i \operatorname{diam}(A)$ and

$$\begin{aligned}
n^{r/s_r}\Delta_{n,r}(G_{\restriction B})G(B) &\leqslant n^{r/s_r}\Delta_{|\mathcal{O}|,r}(G_{\restriction B},\mathcal{O})G(B)\\
&\leqslant n^{r/s_r}\bigcup_{k'=k}^{k_1}\bigcup_{i\in J_{k'}} \sigma_i^r \operatorname{diam}(A)^r G(A_{k',i})\\
&= \operatorname{diam}(A)^r n^{r/s_r}\bigcup_{k'=k}^{k_1}\bigcup_{i\in J_{k'}} \sigma_i^r P(A_{k',i})\frac{G(A_{k',i})}{P(A_{k',i})}\\
&= \operatorname{diam}(A)^r \bigcup_{k'=k}^{k_1}\bigcup_{i\in J_{k'}} n^{r/s_r} P^*(A_{k',i})^{r/s_r}\frac{P^*(A_{k',i})G(A_{k',i})}{P(A_{k',i})}\\
&\leqslant \operatorname{diam}(A)^r \bigcup_{k'=k}^{k_1}\bigcup_{i\in J_{k'}} {\pi_m^*}^{-r/s_r} Q(A_{k',i})\\
&= \operatorname{diam}(A)^r {\pi_m^*}^{-r/s_r} Q(\tilde{B})\\
&\leqslant \operatorname{diam}(A)^r {\pi_m^*}^{-r/s_r} Q(B)(1+\epsilon).
\end{aligned}$$

The following result is a generalization both of Theorem 1 and of the results on distributions, which are absolutely continuous with respect to Lebesgue measure or with respect to Hausdorff measure on regular sets (cf. [**5**, **7**]).

Theorem 3. *Let $G \ll P$. Assume that G has a bounded bounded Radon–Nikodym derivative dG/dP if $P \neq P^*$.*

https://doi.org/10.1017/S0305004104007765 Published online by Cambridge University Press

If $(-\log \pi_1^*, \dots, -\log \pi_m^*)$ *is arithmetic with period* $-\log \tau$, *then there exists a periodic function* $f : \mathbb{R}_+ \to \mathbb{R}_+$ *with period* $-\log \tau$, *such that*

$$\lim_{n\to\infty} \left(n^{r/s_r} \Delta_{n,r}(G) - f(\log n)\right) = 0. \tag{3·7}$$

If $(-\log \pi_1^*, \dots, -\log \pi_m^*)$ *is nonarithmetic, then*

$$\lim_{n\to\infty} n^{r/s_r} \Delta_{n,r}(G) = c(P) \left(\int g^{s_r/(r+s_r)} dP^* \right)^{1+r/s_r}. \tag{3·8}$$

$c(P)$ *denotes the quantization constant of* P *and* g *is a version of* dG/dP. *If* $P \neq P^*$, *then* g *is* $\tilde{P}$ *a.e. the limit of the sequence* (g_k) *defined by* (3·2).

Proof. First we establish (3·7) and (3·8) for piecewise constant Radon–Nikodym derivatives g. Suppose, there exists a $k \in \mathbb{N}$ and a version g of the Radon–Nikodym derivative of dG/dP that is constant on the sets $A_{k,i}$, $i \in \{1, \dots, m\}^k$. Thus for $g_{k,i} = G(A_{k,i})/P(A_{k,i})$, $g(x) = \sum_i g_{k,i} I_{A_{k,i}}(x)$. Define for $i = (i_1, \dots, i_k)$, $\sigma_i = \sigma_{i_1} \cdots \sigma_{i_k}$. Since SSC holds, there are, for n large enough, $n_i \in \mathbb{N}$, $\sum_i n_i = n$, such that

$$\begin{aligned} n^{r/s_r} \Delta_{n,r}(G) &= \sum_{i\in\{1,\dots,m\}^k} (n/n_i)^{r/s_r} \sigma_i^r \Psi(n_i) g_{k,i} P(A_{k,i}) \\ &= \sum_{i\in\{1,\dots,m\}^k} (n/n_i)^{r/s_r} g_{k,i} P^*(A_{k,i})^{1+r/s_r} \Psi(n_i). \end{aligned}$$

It is obvious that if Ψ is asymptotically periodic (in $\log n$), (3·7) holds with f periodic and continuous.

If $\Psi(n)$ converges to $c(P)$, then, since asymptotically optimal weights n_i/n satisfy $n_i/n = P^*(A_{k,i}) g_{k,i}^{s_r/(r+s_r)} / \sum_j P^*(A_{k,j}) g_{k,j}^{s_r/(r+s_r)} + o(1)$, we get

$$\begin{aligned} \lim_{n\to\infty} n^{r/s_r} \Delta_{n,r}(G) &= \sum_{i\in\{1,\dots,m\}^k} \left(\frac{\sum_j P^*(A_{k,j}) g_{k,j}^{s_r/(r+s_r)}}{P^*(A_{k,i}) g_{k,i}^{s_r/(r+s_r)}} \right)^{r/s_r} g_{k,i} P^*(A_{k,i})^{1+r/s_r} c(P) \\ &= c(P) \left(\sum_{i\in\{1,\dots,m\}^k} g_{k,j}^{s_r/(r+s_r)} P^*(A_{k,i}) \right)^{1+r/s_r} \\ &= c(P) \left(\int g^{s_r/(r+s_r)} dP^* \right)^{1+r/s_r}. \end{aligned}$$

Now let in the case $P \neq P^*$ g be $\tilde{P}$-a.e. the limit of (g_k), where g_k is defined by (3·2) and $Q = g_\infty P^*$, where g_∞ is a upper bound of g. If $P = P^*$, then g is P-a.e. the limit of $g_k := \sum_i I_{A_{k,i}}(x) G(A_{k,i})/P(A_{k,i})$ and $Q = G$.

Let $G_k \ll P$ with $dG_k/dP = g_k$. We shall prove that

$$\limsup_{k\to\infty} \limsup_{n\to\infty} n^{r/s_r} |\Delta_{n,r}(G) - \Delta_{n,r}(G_k)| = 0. \tag{3·9}$$

(3·9) implies (3·7) and (3·8) respectively: In the arithmetic case $n^{r/s_r} \Delta_{n,r}(G_k)$ is asymptotically periodic with period $-\log \tau$ and therefore $n^{r/s_r} \Delta_{n,r}(G)$ is. In the nonarithmetic case, since $\lim_{k\to\infty} \|g_k - g\|_{L_1(P^*)} = 0$, the sequence (g_k) is uniformly

https://doi.org/10.1017/S0305004104007765 Published online by Cambridge University Press

integrable and hence

$$\begin{aligned}\lim_{n\to\infty} n^{r/s_r}\Delta_{n,r}(G) &= \lim_{k\to\infty}\lim_{n\to\infty} n^{r/s_r}\Delta_{n,r}(G_k)\\ &= \lim_{k\to\infty}\left(\int g_k^{s_r/(r+s_r)}dP^*\right)^{1+r/s_r}\\ &= \left(\int g^{s_r/(r+s_r)}dP^*\right)^{1+r/s_r}.\end{aligned}$$

To prove

$$\limsup_{k\to\infty}\limsup_{n\to\infty} n^{r/s_r}(\Delta_{n,r}(G)-\Delta_{n,r}(G_k))\leqslant 0, \tag{3·10}$$

let $\epsilon_1,\epsilon_2>0$ and define $B=\{x\mid g(x)\leqslant(1+\epsilon_2)g_k(x)\}$. B^c is its complement. Let $\mathcal{O}_1$ be a set of prototypes for $G_{\restriction B^c}$ of size $n_1=[\epsilon_1 n]$ and $\mathcal{O}_2$ a set of prototypes for G_k of size $n_2=n-n_1$. Then

$$\begin{aligned}\Delta_{n,r}(G) &\leqslant \Delta_{n,r}(G,\mathcal{O}_1\cup\mathcal{O}_2)\\ &\leqslant \Delta_{n_1,r}(G_{\restriction B^c},\mathcal{O}_1)G(B^c)+\int_B \min_{p\in\mathcal{O}_2}\|x-p\|^r dG(x)\\ &\leqslant \Delta_{n_1,r}(G_{\restriction B^c},\mathcal{O}_1)G(B^c)+(1+\epsilon_2)\Delta_{n_2,r}(G_k)\\ &\leqslant n_1^{-r/s_r}bQ(B^c)+(1+\epsilon_2)\Delta_{n_2,r}(G_k),\end{aligned}$$

where the last inequality follows from Lemma 7. Thus

$$\begin{aligned}\limsup_{n\to\infty} n^{r/s_r}(\Delta_{n,r}(G)-\Delta_{n,r}(G_k)) &\leqslant \epsilon_1^{-r/s_r}bQ(B^c)+\left((1+\epsilon_2)(1-\epsilon_1)^{-r/s_r}-1\right)\\ &\quad\times\limsup_{n\to\infty} n^{r/s_r}\Delta_{n,r}(G_k)\\ &\quad+\limsup_{n\to\infty}\left(n_2^{r/s_r}\Delta_{n_2,r}(G_k)-n^{r/s_r}\Delta_{n,r}(G_k)\right).\end{aligned}$$

Note that $\limsup_{k\to\infty}\limsup_{n\to\infty} n^{r/s_r}\Delta_{n,r}(G_k)<\infty$. An upper bound is $\Psi_+(\int g^{s_r/(r+s_r)}\,dP^*)^{1+r/s_r}$. Moreover,

$$\limsup_{n\to\infty}\left(n_2^{r/s_r}\Delta_{n_2,r}(G_k)-n^{r/s_r}\Delta_{n,r}(G_k)\right)=0$$

in the nonarithmetic case or, if $\epsilon_2=\tau^\ell$ with $\ell\in\mathbb{N}$, in the arithmetic case.

Let $\epsilon>0$ be given. Let $(1-\epsilon_1)^{-r/s_r}=1+\epsilon$ and $\epsilon_2=\tau^\ell$ such that $\epsilon_2\leqslant\epsilon$. Let $k_0\in\mathbb{N}$ such that for all $k\geqslant k_0$, $Q(B^c)<\epsilon$. Then

$$\limsup_{n\to\infty} n^{r/s_r}(\Delta_{n,r}(G)-\Delta_{n,r}(G_k))\leqslant b(1-\epsilon)\epsilon+(2\epsilon+\epsilon^2)\limsup_{n\to\infty} n^{r/s_r}\Delta_{n,r}(G_k).$$

This completes the proof of (3·10). The proof of the inequality

$$\limsup_{k\to\infty}\limsup_{n\to\infty} n^{r/s_r}(\Delta_{n,r}(G_k)-\Delta_{n,r}(G))\leqslant 0$$

is similar.

https://doi.org/10.1017/S0305004104007765 Published online by Cambridge University Press

REFERENCES

[1] K. J. FALCONER. *Techniques in Fractal Geometry* (John Wiley, 1997).
[2] S. GRAF and H. LUSCHGY. The quantization of the Cantor distribution. *Math. Nachr.* **183** (1997), 113–133.
[3] S. GRAF and H. LUSCHGY. *Foundations of quantization for probability distributions.* Lecture Notes in Mathematics 1730 (Springer, 2000).
[4] S. GRAF and H. LUSCHGY. The quantization dimension of self-similar probabilities. *Math. Nachr.* **241** (2002), 103–109.
[5] K. PÖTZELBERGER. The general quantization problem for distributions with regular support. *Math. Methods Statist.* **9** (2000), 176–198.
[6] K. PÖTZELBERGER. The quantization dimension of distributions. *Math. Proc. Camb. Phil. Soc.* **131** (2001), 507–519.
[7] P. L. ZADOR. *Development and Evaluation of procedures for quantizing multivariate distributions.* Ph.D. Thesis (Stanford University, 1964).

https://doi.org/10.1017/S0305004104007765 Published online by Cambridge University Press