4 Place Selections Revisited

4.1 Introduction Now that we have definitions of randomness based on two entirely different ideas, to wit, place selections (Chapter 2) and statistical tests (Chapter 3), we must investigate the relations between these definitions. The main *philosophical* differences are summarized in the Introduction to Chapter 3 and we shall not repeat them here. In this chapter, we shall be interested primarily in the *extensional* relation between von Mises' proposal and that of Martin-Löf. Prima facie, an obstacle to a mathematical investigation of this relation is that, as it stands, von Mises' definition is not formal and does not lead to a well-defined set of random sequences, whereas Martin-Löf's definition does determine such a set. We therefore cannot in any literal sense determine the extensional relationship, but we may ask, for example, how one could introduce admissible place selections in Martin-Löf's framework (note that Martin-Löf's definition as such accords no privileged position to place selections). We shall do so in two steps: sections 4.2-5 contain a quantitative study of the behaviour of random sequences under place selections and 5.6 adds admissibility.

It is perhaps best to view these investigations along the following lines: we take some mathematical model for Kollektivs, in this case random sequences (according to any of the definitions of Chapter 3) and we investigate their adequacy for the expression of von Mises' ideas. In a similar vein, Kamae [40] chooses as a formalisation of Kollektivs the Bernoulli sequences (definition 2.5.1.3) and investigates how these sequences behave under a special class of admissible place selections, the entropy zero sequences (see section 5.6). We do not claim finality for any of these formalisations; we are interested in constructing mathematical models for some of von Mises' ideas, even if these models are only partial or in some respects defective.

While the results of 4.5 show that random sequences share many of the desiderata of Kollektivs, section 4.6 elaborates on Ville's theorem (2.6.2.2) and shows that there are some properties of random sequences which need not be satsfied by Kollektivs, when these are defined using some countable set of place selections. The law of the iterated logarithm is one such property, but not the only one. The novelty of the argument of 4.6 is mainly that it is based as directly as possible on the *philosophical* differences between strict frequentism and the propensity interpretation uncovered in Chapter 2.

We now give an outline of the contents of this chapter. In sections 4.2-5 we state precisely and prove the "principle of homogeneity" first mentioned in 2.5: *if* x *is a Kollektiv with respect to* (1–p,p), *so is almost every subsequence of* x. The main result is Theorem 4.5.2, the version of the principle adapted to Martin-Löf's definition of randomness. The really hard part is 4.4, where we prove various effective versions of Fubini's theorem. In section 4.6 we give a new proof of Ville's theorem, which says that for any countable set of place selections $\mathcal{H}$, one can construct a Kollektiv x with respect to $\mathcal{H}$ which approaches its limiting relative frequency $\frac{1}{2}$ from above, thus contradicting the law of the iterated logarithm. The philosophical significance of Ville's theorem was discussed at length in 2.6.2.2. The idea of the new proof is to construct a non-atomic measure μ on 2^ω such that $\mu(C(\frac{1}{2})\cap R_w(\frac{1}{2})^c) = 1$, where $C(\frac{1}{2})$ denotes Church-randomness (with parameter $\frac{1}{2}$). We then have at one stroke continuously many Church-random sequences which are not (weakly) random, but the main advantage of the proof is that it also provides an explanation of this phenomenon.

4.2 Place selections from a modern perspective The starting point of our investigations is proposition 2.3.2.2 (von Mises [67,58]):

An admissibly chosen subsequence of a Kollektiv is again a Kollektiv, with the same distribution.

Using recursive place selections one obtains countably many subsequences of a Kollektiv which are themselves Kollektivs, but we noted in 2.5.2 that a "true" Kollektiv was likely to satisfy a stronger property, dubbed the "principle of homogeneity":

If x *is a Kollektiv with respect to* (1–p,p)*, then so is almost every subsequence of* x.

To put the conjecture in a form susceptible to mathematical analysis, we recall some notation from Chapter 2.

4.2.1 Definition Let x, y $\in 2^\omega$ and suppose that y contains infinitely many ones. Then $^x/y \in 2^\omega$ is determined by

$$(^x/y)_k = x_m \text{ if } m \text{ is the index of the } k^{th}\ 1 \text{ in y.}$$

One may now state the principle of homogeneity as follows:

If x *is a Kollektiv with respect to distribution* (1–p,p), *then* $\mu_p\{x|\ x/y$ is a Kollektiv w.r.t. $\mu_p\} = 1$.

This statement is still only semi-formal, since we have not said what we mean by "Kollektiv". We now examine two possible formalizations.

It seems that the first attempt to prove a principle of homogeneity was Steinhaus' [94,305]. He showed (curiously, without mentioning either von Mises or Kollektivs):

4.2.2 Theorem x $\in$ LLN(p) iff for all q $\in$ (0,1): $\mu_q\{y|\ ^x/y \in \text{LLN(p)}\} = 1$.

While this interesting in itself and will be useful to us later, it is defective as a formulation of the principle of homogeneity. It would be satisfactory only if LLN(p) could be replaced by, say, C($\mathcal{H}$,p), for arbitrary countable sets of place selections $\mathcal{H}$; but the proof does not yield this. Hence typical Kollektiv-like behaviour is not incorporated in the theorem. Indeed, we know of no probabilistic proof which accomplishes this (except for the slightly differently oriented work of Kamae).

In Martin-Löf's set-up, we identify Kollektivs with random sequences and we may prove the principle of homogeneity in the following form (Theorem 4.5.2):

> *Let* $p \in (0,1)$ *be computable and suppose that* ν *is a non-atomic computable measure on* 2^{ω}*. Then for* $x \in R(\mu_p)$, $\nu\{y/\ {}^{x}/y \notin R(\mu_p)\} = 0$.

The "almost all" clause in the principle of homogeneity thus refers, not to some specific measure, but to all computable non-atomic measures, indicating (at least for the constructivist) the extreme smallness of the set of subsequences which are not themselves Kollektivs.
But note that the theorem itself does not speak of admissibility (unless we *define*: y is admissible with respect to x if ${}^{x}/y \in R(\mu_p)$); it has a purely quantitative character. A direct formulation of admissibility must wait until 5.6, when we have at our disposal the notion of Kolmogorov-complexity. There, the techniques used in proving the above theorem will be helpful. One final remark on the principle of homogeneity: it will be observed that the principle states a necessary condition for randomness, whereas Steinhaus' theorem (4.2.2) states a necessary and sufficient condition. We comment on the difference in 4.5.

For completeness' sake, we prove the principle of homogeneity not only for (Martin-Löf) randomness, but for all notions of randomness introduced in Chapter 3. In the case of weak randomness this leads to considerable complexities, but this part of 4.4 can be skipped: section 4.3, lemma 4.4.1 and Theorem 4.4.4 suffice to understand the proof of the main theorem (4.5.2).

4.3 Preliminaries Eventually, in section 4.5, we shall prove

> Let $p \in (0,1)$ be computable and suppose that ν is a non-atomic computable measure on 2^{ω}. Then for $x \in R(\mu_p)$, $\nu\{y/\ {}^{x}/y \notin R(\mu_p)\} = 0$.

Here $R(\mu_p)$ refers to Martin-Löf's definition of randomness (3.2.1.4), but the result holds as well if we replace $R(\mu_p)$ by $R_w(\mu_p)$ (definition 3.2.1.5). For the notions of Gaifman and Snir introduced in section 3.2.4 there is an analogous result if we replace "computable" by "strongly computable". In this section we present some preparatory lemmas and motivate the construction to follow.

The method used in the proof of the main theorem is based on the following observations. The first lemma was already mentioned in section 2.5.

4.3.1 Lemma (Doob [20]) Let $p \in (0,1)$. If $\Phi: 2^{\omega} \to 2^{\omega}$ is a place selection, A a Borel subset of 2^{ω}, then $\mu_p\{x | \Phi x \in A\} \leq \mu_p A$. If $\mu_p(\text{dom}\Phi) = 1$, then we have in fact equality for all A.

Proof See Schnorr [88,23]. □

4.3.2 Lemma For all $p \in (0,1)$, for all non-atomic measures ν on 2^ω, for all Borel subsets A in 2^ω: $\mu_p \times \nu\{<x,y> \mid {}^{x}/y \in A\} = \mu_p A$.

Proof If y contains infinitely many ones, $/y$: $2^\omega \to 2^\omega$ is a total place selection. Since ν is non-atomic, the set of y's having only finitely many ones has measure zero. We may therefore write, using the previous lemma and Fubini's theorem: $\mu_p \times \nu\{<x,y> \mid {}^{x}/y \in A\} =$

$$= \int 1_{\{<x,y> \mid {}^{x}/y \in A\}} d\mu_p \times \nu = \int \mu_p\{x \mid {}^{x}/y \in A\} d\nu(y) = \mu_p A.$$ □

4.3.3 Lemma If $O \subseteq 2^\omega$ is Σ_1, then the set $\{<x,y> \in 2^\omega \times 2^\omega \mid {}^{x}/y \in O\}$ is Σ_1, with Gödelnumber primitive recursive in the Gödelnumber for O.

Proof It suffices to prove the lemma for $O = [w]$. Now observe that the operation / is completely determined by the operation /': $\bigcup_n (2^n \times 2^n) \to 2^{<\omega}$, as follows:

$({}^{v}/'u)_k = v_m$ if m is the index of the k^{th} 1 in u;

and /' is primitive recursive. □

Lemma 4.3.1 suffices to show that for computable $p \in (0,1)$, $R_w(\mu_p)$ is closed under the action of recursive place selections with domain of full measure. Let Φ be a recursive place selection and suppose $\mu_p(\mathrm{dom}\Phi) = 1$. If $N = \bigcap_n O_n$ is a total recursive sequential test with respect to μ_p, then $\Phi^{-1}N = \bigcap_n \Phi^{-1}O_n$ is Π_2 and by lemma 4.3.1, $\mu_p\Phi^{-1}O_n = \mu_p O_n$, so that $\Phi^{-1}N$ is a total recursive sequential test with respect to μ_p. Obviously, for Martin-Löf's $R(\mu_p)$ we have also invariance under recursive place selections whose domain has measure less than one.

Now let μ,ν be computable measures on 2^ω. In Chapter 3 we defined (total) recursive sequential tests as subsets of 2^ω, but definitions 3.2.1.2-3 are easily generalized to the space $2^\omega \times 2^\omega$ and the measure $\mu \times \nu$. We may then state the most useful consequence of the preceding lemmas as follows:

4.3.4 Lemma Let $p \in (0,1)$ be computable and suppose ν is a computable measure on 2^ω. If N is a (total) recursive sequential test in 2^ω with respect to μ_p, then $\{<x,y> \mid {}^{x}/y \in N\}$ is a (total) recursive sequential test with respect to $\mu_p \times \nu$. Similarly, for $n \geq 2$, if N is Π_n μ_p-nullset in 2^ω, then $\{<x,y>| {}^x/y \in N\}$ is a Π_n $\mu_p \times \nu$-nullset in $2^\omega \times 2^\omega$.

This lemma suggests the following strategy for proving the main theorem. Since $R(\mu_p)^c$ is a recursive sequential test with respect to μ_p, the last lemma implies that for any computable measure ν, $\{<x,y>| {}^x/y \in R(\mu_p)\}$ is a recursive sequential test with respect to $\mu_p \times \nu$. By Fubini's theorem, $\mu_p\{x| \nu\{y| {}^x/y \in R(\mu_p)\} > 0\} = 0$. We are done if we can show that this set of x's is in fact contained in a recursive sequential test with respect to μ_p. That this is so, will be proven in the next section.

4.4 Effective Fubini theorems Let μ,ν be computable measures on 2^ω. This section addresses the following question: if $N \subseteq 2^\omega \times 2^\omega$ is a (total) recursive sequential test with respect to $\mu \times \nu$, is it possible to construct a (total) recursive sequential test M with respect to μ such that $\{x| \nu N_x > 0\} \subseteq M$? The answer is yes, but the construction is somewhat complicated, especially in the case of total recursive sequential tests. We also treat briefly the analogous question for Π_n $\mu \times \nu$-nullsets.

In the following pages we shall often use the phrase "[a real] $b_{n,..}$ is computable, uniformly in (the parameter(s)) n,...". This phrase should be interpreted as: "There exists a total recursive function g such that g(n,...) is a Gödelnumber for an algorithm which computes $b_{n,...}$".

The first lemma is in essence due to Sacks (see Sacks [87] or Kechris [42]). For the definition of strongly computable measures, the reader is referred to 3.2.1.1.

4.4.1 Lemma (i) Let ν be a computable measure on 2^ω and suppose that A is a Σ_0 subset of $2^\omega \times 2^\omega$. Then the function $x \to \nu A_x$ is of the form

$$\nu A_x = \sum_{k=1}^{n} c_k \cdot 1_{C_k}(x),$$

where C_k is a Σ_0 subset of 2^ω and c_k is a computable real. In addition, if ν is strongly computable, then the sets $\{a \in \mathbb{Q}| c_k < a\}$ and $\{a \in \mathbb{Q}| c_k > a\}$ are recursive. (ii) Let ν be a computable measure on 2^ω and suppose that A is a Σ_1 subset of $2^\omega \times 2^\omega$. Then the set $\{<a,x> \in \mathbb{Q} \times 2^\omega \mid \nu A_x > a\}$ is Σ_1. (iii) Let ν be a strongly computable measure on 2^ω. If $A \subseteq 2^\omega \times 2^\omega$ is Σ_n, then the set $\{<a,x> \in \mathbb{Q} \times 2^\omega| \nu A_x > a\}$ is Σ_n. If A is Π_n, then $\{<a,x> \in \mathbb{Q} \times 2^\omega| \nu A_x > a\}$ is Σ_{n+1}.

Proof (i) Using if necessary a suitable tiling of A, we may write A as a *disjoint* union

$$A = \bigcup_{i=1}^{m} ([w^i] \times [v^i])$$

such that all w^i have the same length n (hence the $[w^i]$ are either disjoint or identical). Then we have, for all x

$$\nu A_x = \sum_{x(n) = w^i} \nu[v^i].$$

If we define for $k \le 2^n$, $C_k := [u]$ for the k^{th} word u in 2^n and

$$c_k := \sum_{C_k = w^i} \nu[v^i] \ \ (\text{where } \sum_{\emptyset} \nu[v^i] = 0),$$

then c_k has the required properties and

$$\nu A_x = \sum_{k=1}^{2^n} c_k \cdot 1_{C_k}(x).$$

(ii) Let $A = \{<x,y> \mid \exists n\, R(n,x,y)\}$, where R is a recursive relation. Write $A^m := \{<x,y> \mid \exists n \le m\, R(n,x,y)\}$, then A^m is Σ_0. We have

$$\{<a,x> \mid \nu A_x > a\} = \{<a,x> \mid \exists m\, (\nu A_x^m > a)\},$$

and the result follows by (i).

(iii) If A is Π_1, then $A = \{<x,y> \mid \forall n\, R(n,x,y)\}$ for some recursive relation R. Put $A^m := \{<x,y> \mid \forall n \le m\, R(n,x,y)\}$, then A^m is Σ_0 and we may write

$$\{<a,x> \mid \nu A_x > a\} = \{<a,x> \mid \exists \delta \in \mathbb{Q}^+ \forall m\, (\nu A_x^m > a+\delta)\},$$

and for strongly computable measures ν this set is Σ_2, by (i). The result now follows by induction on n. ∎

4.4.2 Theorem Let μ,ν be strongly computable measures on 2^ω. Suppose that $N \subseteq 2^\omega \times 2^\omega$ is a Π_n $\mu\times\nu$-nullset. Then $\{x \mid \nu N_x > 0\}$ is a Σ_{n+1} μ-nullset.

Proof $\{x \mid \nu N_x > 0\} = \{x \mid \exists a \in \mathbb{Q}^+(\nu N_x > a)\}$ is Σ_{n+1} by lemma 4.4.1 and a μ-nullset by Fubini's theorem. ∎

Theorem 4.4.2 is slightly unsatisfactory, in that one would like to have "Π_n" instead of "Σ_{n+1}" in the conclusion of the theorem. We do not know whether the above estimate is exact. We can show, however, that in general "Σ_{n+1}" cannot be replaced by "Σ_n". Namely, we construct a Π_2 $\lambda\times\lambda$-nullset in $2^\omega\times 2^\omega$ such that $\{x \mid \lambda N_x > 0\}$ is not contained in a Σ_2 λ-nullset. Let M be a total recursive sequential test (with respect to λ) which contains $LLN(\frac{1}{2})^c$ (see section 3.3). Consider $N := \{<x,y> \mid {}^x/y \in M\}$. By lemma 4.3.3, N is Π_2 and by lemma 4.3.2, $\lambda\times\lambda N = 0$. Suppose $\{x \mid \lambda N_x > 0\}$ were contained in a Σ_2 set B with $\lambda B = 0$. If $x \in LLN(\frac{1}{2})^c$, then by Theorem 4.2.2, $\{y \mid {}^x/y \in LLN(\frac{1}{2})^c\} = 1$; hence $\lambda N_x = 1$ and thus $x \in B$. Therefore $B^c \subseteq$

LLN(½). But this is impossible since LLN(½) is first category while B^c is residual: the first statement is obvious and the second statement follows since B^c is a G_δ set which is dense by $\lambda B^c = 1$.

In what follows, we shall often refer to *computable* real-valued functions on 2^ω, the recursion-theoretic analogue of the *continuous* real-valued functions of constructive analysis (see e.g. Bishop–Bridges [6,38]). We therefore introduce

4.4.3 Definition $f: 2^\omega \to \mathbb{R}$ is *computable* if it is recursively uniformly continuous, i.e. if for some total recursive $h: \omega \to \mathbb{Q}$:

for all n, for all x,y: if $|x - y| < h(n)$, then $|f(x) - f(y)| < 2^{-n}$.

The first part of lemma 4.4.1 implies that if ν is a computable measure and $A \subseteq 2^\omega \times 2^\omega$ is Σ_0, then the function $x \to \nu A_x$ is computable.

The effective Fubini theorem for recursive sequential tests can fortunately be obtained easily by formalizing the proof of Theorem 14.1 in Oxtoby [80].

4.4.4 Theorem Let μ,ν be computable measures on 2^ω and suppose that $N \subseteq 2^\omega \times 2^\omega$ is a recursive sequential test with respect to $\mu\times\nu$. Then $\{x|\ \nu N_x > 0\}$ is contained in a recursive sequential test with respect to μ.

Proof Let $N = \bigcap_n O_n \subseteq 2^\omega \times 2^\omega$ be a recursive sequential test with respect to $\mu\times\nu$. Uniformly in n, we construct Σ_1 sets $B_n \subseteq 2^\omega$ such that $\mu B_n \leq 2^{-n}$ and $\{x|\ \nu N_x > 0\} \subseteq B_n$. Choose n. Clearly $\mu\times\nu \bigcup_{k>n} O_k \leq 2^{-n}$. $\bigcup_{k>n} O_k$ is of the form $\bigcup_i [w^i]\times[v^i]$ and the sequence $([w^i]\times[v^i])_i$ covers N infinitely often, that is, each $\langle x,y\rangle \in N$ is contained in infinitely many cylinders $[w^i]\times[v^i]$ of the sequence.

Define a sequence of functions f_k, $k \geq 0$, by

$$f_0(x) = 0 \text{ for all } x$$

$$f_k(x) = \sum_{\{i \leq k |\ x \in [w^i]\}} \nu[v^i], \text{ for } k \geq 1.$$

f_k is a computable stepfunction, $f_k: 2^\omega \to [0,1]$, $f_k \leq f_{k+1}$ and

$$f_{k+1}(x) - f_k(x) = \begin{cases} \nu[v^{k+1}] & \text{if } x \in [w^{k+1}] \\ 0 & \text{otherwise.} \end{cases}$$

Clearly

$$\int f_k d\mu \;=\; \sum_{i=1}^{k}\int (f_i - f_{i-1})d\mu \;=\; \sum_{i=1}^{k}\mu[w^i]\cdot\nu[v^i] \;\le\; 2^{-n}.$$

Define $B_n := \{x|\ \exists k\ f_k(x) > 1\}$ (remember that the f_k depend implicitly on n!). Obviously, B_n is Σ_1, uniformly in n. Moreover, $\{x|\ \nu N_x > 0\} \subseteq B_n$: choose x such that $\nu N_x > 0$, then a fortiori for some y, $<x,y> \in N$. Hence for infinitely many i: $<x,y> \in [w^i]\times[v^i]$. Let (i') be the sequence of indices for which $x \in [w^{i'}]$. For any $y \in N_x$, for infinitely many i': $y \in [v^{i'}]$. Hence the sequence $([v^{i'}])_{i'}$ covers N_x infinitely often, so $\sum_{i'}\nu[v^{i'}]$ must diverge (otherwise, we could cover N_x with open sets of arbitrarily small ν–measure). It follows that, still for this particular x, $\lim_{k\to\infty} f_k(x) = \infty$ and thus, for some k, $f_k(x) > 1$, i.e. $x \in B_n$.

Clearly then, $\bigcap_n B_n$ is the required recursive sequential test if we can show that $\mu B_n \le 2^{-n}$. Now if we put $A_m := \{x|\ \exists k\le m\ f_k(x) > 1\}$, B_n is the limit of the A_m. Since $f_k \le f_{k+1}$,

$$\mu A_m \;=\; \int 1_{A_m} d\mu \;<\; \int f_m d\mu \;\le\; 2^{-n} \text{ for all m,}$$

and so $\mu B_n \le 2^{-n}$. ∎

4.4.5 Corollary Let μ,ν be computable measures on 2^ω. Suppose that U is the universal recursive sequential test with respect to $\mu\times\nu$ and that U' is the universal recursive sequential test with respect to μ. Then $U' = \{x|\ \nu U_x > 0\}$.

Proof By the preceding theorem, $\{x|\ \nu U_x > 0\}\subseteq U'$. On the other hand, $U'\times 2^\omega \subseteq U$. ∎

Consequently, if N is a recursive sequential test, $\{x|\ \nu N_x > 0\}$ need not be contained in a *total* recursive sequential test, since such a test cannot be universal, as we have seen in Chapter 3. This fact necessitates a separate effective Fubini theorem for total recursive sequential tests. The reader not especiallly interested in total recursive sequential tests is free to stop here and may proceed directly to section 4.5.

Our next object is to prove

4.4.6 Theorem Let μ,ν be computable measures on 2^ω. Let $N \subseteq 2^\omega\times 2^\omega$ be a total recursive sequential test with respect to $\mu\times\nu$. Then $\{x|\ \nu N_x > 0\}$ is contained in a total recursive sequential test with respect to μ.

This theorem can presumably be proved by formalizing proofs of Fubini's theorem from constructive analysis. However, since we allowed ourselves the use of classical logic and mathematics, a more direct approach is possible. The key of the proof consists in the following observation:

> If $O \subseteq 2^\omega \times 2^\omega$ is a Σ_1 set such that $\mu\times\nu O$ is computable and if the image measure π is defined by $\pi[0,s] := \mu\{x|\ \nu O_x \le s\}$, for $0 \le s \le 1$, then the set of points of continuity of π has a Π_2 definition.

Since the set of points of continuity is dense, it follows from an effective version of the Baire Category Theorem, that π has a recursively enumerable dense set of computable points of continuity. From then on, the going is easy.

Our proof strategy is fairly opportunistic: whenever possible, we borrow the requisite algorithms from constructive analysis (e.g. the functions $g(u,v,\cdot)$ defined below, are taken from Bishop and Cheng [7]); but the proofs that these algorithms are in fact total are entirely classical (e.g. lemma 4.4.12).

We now proceed to the proof of Theorem 4.4.6. Write $N = \bigcap_n O^n$, $O^{n+1} \subseteq O^n$, $O^n \in \Sigma_1$, $\mu\times\nu O^n$ computable (uniformly in n) and $\le 2^{-n}$. Define on [0,1] the image measure π_n as follows:

$$\pi_n[0,s] := \mu\{x|\ \nu O^n_x \le s\},\ \ 0 \le s \le 1.$$

π_n need not be a computable measure, but nevertheless, as we shall see, some integrals with respect to π_n are computable. We use this fact to compute $\pi_n[0,s]$ for a recursively enumerable dense set of computable reals s.

4.4.7 Definition For $u,v \in [0,1]\cap\mathbb{Q}$, $u<v$, we determine a function $g(u,v,\cdot)$ as follows:

$$g(u,v,t) = \begin{cases} 1 & t<u \\ (v-t)/(v-u) & u\le t\le v \\ 0 & v<t. \end{cases}$$

Let $u_0<v_0<u_1<v_1$ be rationals. The functions $f(u_0,v_0,u_1,v_1,\cdot)$ are defined by

$$f(u_0,v_0,u_1,v_1,t) := \min\{1 - g(u_0,v_0,t),\ g(u_1,v_1,t)\}.$$

Before we can motivate the introduction of these auxiliary functions, we need a lemma.

4.4.8 Lemma The integrals

$$\int_{[0,1]} g(u,v,t)d\pi_n(t)\ ,\ \int_{[0,1]} f(u_0,v_0,u_1,v_1,t)d\pi_n(t)$$

are computable uniformly in the parameters n,u,v and n,u_0,v_0,u_1,v_1 respectively.

Proof For this lemma we are indebted to constructive analysis, and in particular to the constructive theory of integration developed in [6], [7] and [9]. Observe that

(i) $\int_{[0,1]} g(u,v,t)d\pi_n(t) = \int_{2^\omega} g(u,v,\nu O^n_x)d\mu(x);$

(ii) $g(u,v,\nu O^n_x) = \min\left\{1, \frac{(v - \min(\nu O^n_x, v))}{v-u}\right\};$

(iii) there exists a recursive family of Σ_0 sets $C^{n,k}$ such that each O^n can be written as a *disjoint* union $O^n = \bigcup_k C^{n,k}$. We then have, for all x:

$$\nu O^n_x = \sum_{k=1}^{\infty} \nu C^{n,k}_x \text{ and } \sum_{k=1}^{\infty} \int_{2^\omega} \nu C^{n,k}_x d\mu(x) = \int_{2^\omega} \nu O^n_x d\mu(x) = \mu\times\nu O^n \text{ is}$$

computable, uniformly in n.

(iv) the function $x \to \nu C^{n,k}_x$ is computable (by lemma 4.4.1) and

$\int_{2^\omega} \nu C^{n,k}_x d\mu(x)$ is computable, both uniformly in n and k.

Call a function h *integrable* (with respect to μ) if there exists a sequence (h_m) of computable functions such that $h = \sum_m h_m$ μ-a.e. and $\sum_m \int h_m d\mu$ is computable (cf. [6,226]). Then the function $x \to \nu O_x$ is integrable (by (iii) and (iv)) and Theorem 2.18 of Bishop-Bridges [6,230] may be translated to our recursion–theoretic setting to show that the operation $\min(\cdot,\cdot)$ preserves integrability. Hence f and g are integrable (by (i) and (ii)). ∎

Now consider a computable real s and rationals u_0,v_0,u_1,v_1 such that $u_0<v_0<s<u_1<v_1$. Obviously, $\int g(u_0,v_0,t)d\pi_n(t) \le \pi_n[0,s] \le \int g(u_1,v_1,t)d\pi_n(t)$, and by the preceding lemma the terms on the left hand side and on the right hand side are computable. What remains to be done, is to find a computable estimate of the difference

$$\int g(u_1,v_1,t)d\pi_n(t) - \int g(u_0,v_0,t)d\pi_n(t).$$

For certain s, this can be achieved using the functions $f(u_0,v_0,u_1,v_1,t)$.

4.4.9 Definition $s \in [0,1]$ is an *atom* of π_n if $\pi_n\{s\} > 0$. $s \in [0,1]$ is a *point of continuity* of π_n (abbreviated: s is p.c. of π_n) if $\pi_n\{s\} = 0$.

The key of the proof of Theorem 4.4.6 is that the set of p.c.'s of the π_n has a Π_2 definition.

4.4.10 Lemma $s \in [0,1]$ is p.c. of all π_n iff

$$(*)\ \forall n\ \forall \varepsilon{>}0\ \exists \delta{>}0\ \exists u_0, v_0, u_1, v_1\ \Big(v_0{<}s{-}\delta{<}s{+}\delta{<}u_1\ \ \&\ \int_{2^\omega} f(u_0,v_0,u_1,v_1,t)d\pi_n(t) < \varepsilon\Big),$$

where the quantifiers "$\forall\varepsilon$" and "$\exists\delta$" range over the rationals. Moreover, (*) is a $\prod_2$ statement.

Proof The first statement is obvious as soon as we realize that the condition "$v_0{<}s{-}\delta{<}s{+}\delta{<}u_1$" in (*) means that $f(u_0,v_0,u_1,v_1,t)$ equals 1 on $(s{-}\delta,s{+}\delta)$. The second statement follows from lemma 4.4.8. ∎

The $\prod_2$ definition of the property "s is p.c. of all π_n" enables us to apply the following effective version of the Baire Category Theorem:

4.4.11 Lemma Let G be a dense $\prod_2$ subset of [0,1]. Then G contains a recursively enumerable dense subset of computable reals.

Proof Formalize a proof of the Baire Category Theorem (e.g. Oxtoby [80,2]). ∎

Combining these lemmas, we get

4.4.12 Lemma There exists a recursively enumerable dense set D of computable points of continuity of all π_n.

Proof By lemma 4.4.10 the set of p.c. of all π_n has a $\prod_2$ definition. This set is dense in [0,1], since the set of s which are an atom for some π_n is countable (this argument is non-constructive). Now apply the preceding lemma. ∎

We are now almost done.

4.4.13 Lemma Let $s \in [0,1]$ be a computable point of continuity of all π_n. Then $\pi_n[0,s]$ is computable, uniformly in n.

Proof Choose $\varepsilon > 0$. We must effectively determine $u < v < u' < v'$ such that

(1) $\int g(u,v,t)d\pi_n(t) \le \pi_n[0,s] \le \int g(u',v',t)d\pi_n(t)$

(2) $\int g(u',v',t)d\pi_n(t) - \int g(u,v,t)d\pi_n(t) < \varepsilon$.

Choose recursively enumerable sequences of rationals (b_k), (c_k) such that for all k, $b_k < s < c_k$ and $c_k - b_k < 2^{-k}$. By lemma 4.4.10 there exist (for this particular ε) $\delta > 0$ and rationals $u_0 < v_0 < u_1 < v_1$ such that $v_0 < s-\delta < s+\delta < u_1$ and $\int f(u_0,v_0,u_1,v_1,t)d\pi_n(t) < \varepsilon$. Choose k large enough so that $s - b_k < \delta/4$ and $c_k - s < \delta/4$.
Define $u := b_k - \delta/4$, $v := b_k$, $u' := c_k$ and $v' = c_k + \delta/4$. Then $v_0 < u < v < s < u' < v' < u_1$, hence (1) holds and $\int g(u',v',t)d\pi_n(t) - \int g(u,v,t)d\pi_n(t) \leq \int f(u_0,v_0,u_1,v_1,t)d\pi_n(t) < \varepsilon$. ∎

Now let D be the set constructed in lemma 4.4.12. Theorem 4.4.6 follows if we can show that

$$\bigcup_{s\in D}\bigcap_n\{x \mid \nu O^n_x > s\}$$

is contained in a total recursive sequential test with respect to μ.
By lemma 4.4.1, for $s \in D$,

$$\{x \mid \nu O^n_x > s\} \in \Sigma_1.$$

Moreover, since

(i) $\mu\{x \mid \nu O^n_x > s\}$ is computable, uniformly in n (by lemma 4.4.13)

(ii) $\mu\bigcap_n\{x \mid \nu O^n_x > s\} = 0$ by Fubini's theorem,

we can determine a recursively enumerable infinite sequence (n_k) of natural numbers such that for all k

$$\mu\{x \mid \nu O^{n_k}_x > s\} < 2^{-k}.$$

Because $O^{n+1} \subseteq O^n$

$$\bigcap_n\{x \mid \nu O^n_x > s\} = \bigcap_k\{x \mid \nu O^{n_k}_x > s\};$$

and

$$\bigcap_k\{x \mid \nu O^{n_k}_x > s\}$$

is a total recursive sequential test with respect to μ. By lemma 3.2.3.8, the union of these tests over D is contained in a total recursive sequential test with respect to μ. But this union equals $\{x \mid \nu N_x > 0\}$. This concludes the proof of Theorem 4.4.6. ∎

4.5 Proof of the principle of homogeneity Classically, a subset E of 2^ω has *absolute measure zero* if for every finite non-atomic measure μ on 2^ω we can find a Borelset A such that $E \subseteq A$ and $\mu A = 0$. Hausdorff constructed an example of such a set of cardinality $\aleph_1$ (and

this is the best possible result).
This concept can be transferred to the constructive realm as follows: $E \subseteq 2^\omega$ is *recursively small* if for every *computable* finite non-atomic measure μ on 2^ω, we can find a Borelset A such that $E \subseteq A$ and $\mu A = 0$.
Theorems 4.5.2-3 will show that if $x \in R(\mu_p)$ ($R_w(\mu_p)$), then the set $\{y| {}^x/y \notin R(\mu_p)\}$ ($\{y| {}^x/y \notin R_w(\mu_p)\}$) is recursively small. (In another sense, these sets are quite large, since they are residual.) For completeness' sake, we begin with the corresponding result for n-randomness.
Strongly computable measures were defined in 3.2.1.1. We say that $p \in (0,1)$ is *strongly computable* if the sets $\{a\in \mathbb{Q}| a > p\}$ and $\{a\in \mathbb{Q}| a < p\}$ are both Δ_1. If $p \in (0,1)$ is strongly computable, then μ_p is a strongly computable measure.

4.5.1 Theorem Let ν be a non-atomic strongly computable measure on 2^ω and let $p \in (0,1)$ be strongly computable. For $n \geq 2$, if x is n–random with respect to μ_p, then $\nu\{y| {}^x/y$ is not n–random with respect to $\mu_p\} = 0$.

Proof It suffices to show that for each $\prod_n$ $\mu_p\times\nu$-nullset N, $\{x| \nu\{y| {}^x/y \in N\} > 0\}$ is contained in a $\sum_{n+1}$ μ_p-nullset. By lemma 4.3.4, $\{\langle x,y\rangle| {}^x/y \in N\}$ is a $\prod_n$ $\mu_p\times\nu$-nullset. Since μ_p is strongly computable, we may now apply Theorem 4.4.2. □

4.5.2 Theorem Let $p \in (0,1)$ be computable. If $x \in R(\mu_p)$, then $\{y| {}^x/y \notin R(\mu_p)\}$ is recursively small.

Proof Let ν be a non-atomic computable measure. Since $R(\mu_p)^c$ is a recursive sequential test with respect to μ_p, lemma 4.3.4 implies that $\{\langle x,y\rangle| {}^x/y \notin R(\mu_p)\}$ is a recursive sequential test with respect to $\mu_p\times\nu$. Now apply Theorem 4.4.4. □

4.5.3 Theorem Let $p \in (0,1)$ be computable. If $x \in R_w(\mu_p)$, then $\{y| {}^x/y \notin R_w(\mu_p)\}$ is recursively small.

Proof Let N be a total recursive sequential test with respect to μ_p. Let ν be a non-atomic computable measure. By lemma 4.3.4, $\{\langle x,y\rangle| {}^x/y \in N\}$ is a total recursive sequential test with respect to $\mu_p\times\nu$. Now apply Theorem 4.4.6. □

4.5.4 Remarks (i) The principle of homogeneity thus holds true for a wide class of definitions of randomness based on probabilistic laws, although we needed three different proofs to show this. The common core of these proofs is that the operation / is measure-preserving and also preserves arithmetical structure; the differences result from the fact that the Fubini-property needs a separate verification in each case.

(ii) Looking back on what we have accomplished, we see that, at least in a quantitative sense, von Mises' intuitions can be salvaged: if we provisionally identify Kollektivs with random sequences (in Martin–Löf's sense), then the set of subsequences of a Kollektiv which are not themselves Kollektivs is exceedingly small. Alternatively, we might say that Martin-Löf's definition and its variants capture at least some of von Mises' intentions. Observe that, from von Mises' point of view, the preceding theorems should not be interpreted as a result on the extremely small *probability* of the set $\{y|\ {}^{x}/y \notin R(\mu_p)\}$.
(iii) If we compare Theorem 4.2.2 with the preceding theorems, we see that the latter state a necessary condition for randomness, whereas the first is a necessary and sufficient condition for satisfying the law of large numbers. We doubt whether the preceding theorems admit a converse. Perhaps there is a converse if we strengthen the consequences using compositions of recursive selections and random selections, in the following sense.
If Φ is a recursive place selection (with generating function ϕ as in definition 2.5.1.1) such that $\mu_p \mathrm{dom}\Phi = 1$, define $/^{\Phi}$ by

$$({}^{x}/^{\Phi}y)_k = x_m \text{ if m is the index of the } k^{th} \text{ 1 in y and } \phi(x(m-1)) = 1.$$

Since $/^{\Phi}$ satisfies lemmas 4.3.1-3, the preceding theorems hold with / replaced by $/^{\Phi}$.

Various other theorems on the operation / can be derived along these lines, the most interesting of which is perhaps the following. Let ${}^{x}/^{\circ}y$ be defined as ${}^{x}/y$, except that we now look at the zeros of y. Hence, when viewed as sets of natural numbers, ${}^{x}/y \cup {}^{x}/^{\circ}y = \mathbb{N}$.

4.5.4 Theorem Let $p \in (0,1)$ be computable. If $x \in R(\mu_p)$, then the set $\{y|\ <{}^{x}/y, {}^{x}/^{\circ}y> \notin R(\mu_p \times \mu_p)\}$ is recursively small.

Proof Let ν be a computable non-atomic measure. We show first that $\mu_p \times \nu\{<x,y>|\ <{}^{x}/y, {}^{x}/^{\circ}y> \in A \times B\} = \mu_p A \cdot \mu_p B$. As in lemma 4.3.2, it suffices to show that for fixed y, $\mu_p\{x|\ {}^{x}/y \in A, {}^{x}/^{\circ}y \in B\} = \mu_p A \cdot \mu_p B$. We need only verify this equality for A = [w], B= [v]. But $\{x|\ {}^{x}/y \in [w], {}^{x}/^{\circ}y \in [v]\} = [u]$, where $|u| = |w| + |v|$ and u consists of w and v intertwined. Hence $\mu_p[u] = \mu_p[w] \cdot \mu_p[v]$. From here on, the argument is entirely similar to the arguments above. □

To interpret this theorem, recall that we defined two Kollektivs z^0, z^1 to be *independent* if the pair $<z^0,z^1>$ is a Kollektiv with respect to the product distribution (cf. 2.4.1). Having formalized Kollektivs as random sequences, it seems reasonable to formalize a pair of independent Kollektivs as an element of $R(\mu_p \times \mu_p)$ (such pairs are invariant under recursive place selections, they satisfy the law of the iterated logarithm etc.). We saw in 2.4.1 that a *lawlike* partition of a Kollektiv into two (or more) Kollektivs yields provably independent Kollektivs and we remarked that this feature reflects the assumed independence of successive tosses. We now see that also in this context a principle of homogeneity obtains: "almost every" partition, whether lawlike or not, produces independent Kollektivs.

4.6 New proof of a theorem of Ville In 2.6.2.2, we stated Ville's theorem as follows:

Given a countable set $\mathcal{H}$ of place selections $\Phi: 2^\omega \to 2^\omega$ we can construct $x \in 2^\omega$ such that

(i) $x \in \mathrm{dom}\Phi$ implies $\Phi x \in \mathrm{LLN}(\frac{1}{2})$, for all $\Phi \in \mathcal{H}$

(ii) for all n $\frac{1}{n}\sum_{k=1}^{n} x_k \geq \frac{1}{2}$.

C(p), the set of Church-random sequences with parameter p, was defined in 2.5.1.7. Since property (ii) contradicts the law of the iterated logarithm and all (weakly) random sequences satisfy the law of the iterated logarithm (as was shown in section 3.3), we have as a consequence $C(\frac{1}{2}) \cap R(\lambda)^c \neq \emptyset$ (although of course $R(\lambda) \subseteq C(\frac{1}{2})$). Thus $C(\frac{1}{2})$ and $R(\lambda)$, which have very different philosophical justifications, differ also extensionally.

We need not repeat here the discussion on the philosophical significance of Ville's theorem given in 2.6.2.2; in the present section we are concerned only with its proof. Ville's argument [99,55-69] has a combinatorial character and consists roughly speaking in replacing $\mathcal{H}$ by a different set $\mathcal{H}'$ of place selections Ψ such that if $\Psi, \Psi' \in \mathcal{H}'$, then Ψ and Ψ' are "disjoint". This notion of disjointness is best illustrated by means of an example. Let (p_n) be an enumeration of the prime numbers and let Ψ_n be the place selection that chooses all indices which are a power of p_n. Then no two Ψ_n choose the same index and in this case it is very easy to construct an x which satisfies specifications (i) and (ii). By adroitly manipulating place selections, Ville is able to reduce the general case to something very like the above example.

Without denying the ingenuity of Ville's construction, it seems worthwhile to try to derive the theorem from first principles, that is, as an expression of the philosophical differences between strict frequentism and the propensity interpretation uncovered in Chapter 2. In other words, we want to show that the different interpretations of probability underlying the definitions of Church-random sequences and (Martin-Löf) random sequences, namely probability as relative frequency and coordinate-wise probability respectively, *themselves* imply that $C(\frac{1}{2}) \cap R(\lambda)^c \neq \emptyset$.

In the introduction to Chapter 3 we observed that, from the point of view of strict frequentism, the distribution (1–p,p) on {0,1} should not be associated with a unique measure on 2^ω, to wit, μ_p, but rather with a whole class of measures, namely all those which in a certain sense determine the same limiting relative frequencies 1–p and p. Existence theorems should not be affected when we replace one measure from this class by another. We may therefore state

Conjecture 1 Let $\pi = \prod_n(1-p_n,p_n)$ be a product measure such that $\lim_{n\to\infty} p_n = p$ and let $\mathcal{H}$ be a countable set of place selections. Then not only $\mu_p C(\mathcal{H},p) = 1$, as was shown in Theorem 2.5.2.3, but also $\pi C(\mathcal{H},p) = 1$.

On the other hand, the definition of (weakly) random sequences at first sight seems to involve a unique measure, namely μ_p. This impression is confirmed by the discussion of Martingales in 3.4, where it was seen that their definition seemed to require (constant) probabilities at individual coordinates. One way to state this seeming dependence upon the underlying measure is as follows:

Conjecture 2 If $\pi = \prod_n(1-p_n,p_n)$ with $\lim_{n\to\infty} p_n = p$ but $p_n \neq p$ for all n, then $\pi R(\mu_p) = 0$. By the 0–1 law, $\pi R(\mu_p)$ is either one or zero and the first case seems to be excluded by the above argument.

Both conjectures, taken together, would give us the required proof of Ville's theorem from first principles, for if π satisfies the hypothesis of Conjecture 2, we would have $\pi(C(p) \cap R(\mu_p)^c) = 1$. But, although Conjecture 1 can indeed be proven (see corollary 4.6.3), Conjecture 2 is false. First impressions notwithstanding, $R(\mu_p)$ is not *that* sensitive to the choice of the underlying measure: there exist $\pi = \prod_n(1-p_n,p_n)$ such that $\lim_{k\to\infty} p_n = p$ and $p_n \neq p$ for all n, for which $\pi R(\mu_p) = 1$.

On the other hand, the idea that the extensional difference between C(p) and $R(\mu_p)$ is due to a difference in sensitivity to the choice of the measure is correct, but it should be formulated more carefully. Although for a computable product measure $\pi = \prod_n(1-p_n,p_n)$, $\pi R(\mu_p) = 1$ does not imply that $p_n = p$ for all n, it *does* imply that $\sum_n(p-p_n)^2 < \infty$, in other words, that p_n converges to p rather *fast*.We then get a proof of Ville's theorem if we take a computable product measure π for which the marginals p_n converge *slowly* to p, for in that case $\pi(C(p) \cap R(\mu_p)^c) = 1$ (Theorem 4.6.1).

The result we derive in this way differs from Ville's original formulation in two minor respects:

– not only is $C(p) \cap R(\mu_p)^c$ non-empty, it has the cardinality of the continuum;

– on the other hand, the proof does not yield that for *every* π such that $\pi(C(p) \cap R(\mu_p)^c) = 1$, already $\pi(C(p) \cap LIL(\mu_p)^c) = 1$, where $LIL(\mu_p)$ is the set of sequences which satisfy the law of the iterated logarithm for μ_p. Indeed, the proof *cannot* yield such a result, since it is false for some π with $\pi(C(p) \cap R(\mu_p)^c) = 1$. But for some very slowly converging π, we do have that $\pi(C(p) \cap LIL(\mu_p)^c) = 1$, thus strengthening Ville's theorem in its original formulation.

The reader may wonder why we persistently formulate these results for C(p) instead of for C($\mathcal{H}$,p), for arbitrary countable sets $\mathcal{H}$ of place selections. The answer is that R(μ_p), due to its recursion theoretic structure can only be reasonably compared with C(p). For very slowly converging π, however, we have, for arbitrary $\mathcal{H}$, $\pi(C(\mathcal{H},p)\cap LIL(\mu_p)^c) = 1$.

This section is organized as follows. We first prove Ville's theorem along the lines sketched above (Theorem 4.6.1) and we comment on the significance of the proof (Corollary 4.6.6 and following discussion). The reader may then proceed to Chapter 5; the rest of the section generalizes Corollary 4.6.6 to measures which are not product measures and is not essential to the main argument.

We prove Ville's theorem in the following form:

4.6.1 Theorem Let $p \in (0,1)$ be a computable real. There exists a non-atomic computable measure π such that $\pi(C(p)\cap R_w(\mu_p)^c) = 1$. A fortiori, $\pi(C(p)\cap R(\mu_p)^c) = 1$ and $C(p)\cap R_w(\mu_p)^c$ has the cardinality of the continuum.

The measure will be a computable product measure $\pi = \prod_n(1-p_n,p_n)$ such that $\lim_{k\to\infty} p_n = p$ and $\pi\perp\mu_p$. In fact, the proof will show that for *any* such measure π, $\pi(C(p)\cap R_w(\mu_p)^c) = 1$.

4.6.2 Lemma Let $\pi = \prod_n(1-p_n,p_n)$ be a computable product measure. Then $\pi C(p) = 1$ iff $\lim_{n\to\infty} p_n = p$.

Proof $\Rightarrow$ Suppose not. Then for some rational $\varepsilon > 0$, at least one of the sets $\{n|\ p_n > p+\varepsilon\}$, $\{n|\ p_n < p-\varepsilon\}$ is infinite, say the first set. By the computability of π this set is recursively enumerable, hence contains an infinite recursive subset. Using this subset, we can define a recursive place selection Φ such that $\pi\Phi^{-1}(LLN(p)) = 0$, a contradiction.
$\Leftarrow$ For this direction, no assumption of computability or recursiveness is needed. So let Φ be a place selection and π a measure of the form $\pi = \prod_n(1-p_n,p_n)$ such that $\lim_{n\to\infty} p_n = p$ and assume that $p_n \neq 0$ for all n (which is no essential restriction). We show that $\pi(\mathrm{dom}\Phi) = \pi(\mathrm{dom}\Phi\cap\Phi^{-1}(LLN(p)))$. Given Φ and its generating function ϕ (as in definition 2.5.1.1), we define a partial function θ: $2^\omega\times\omega \to \omega$ as follows:

(1) $\mathrm{dom}\theta = \mathrm{dom}\Phi$

(2) if $x \in \mathrm{dom}\Phi$, then $\theta(x,n) = 1 + \min\left\{k \mid n = \sum_{j=1}^{k}\phi(x(j))\right\}$.

Assume first that $\pi(\mathrm{dom}\Phi) = 1$ Define random variables $Z_n: 2^\omega \to \mathbb{R}$ by

$$Z_n(x) = \frac{x_{\theta(x,n)}}{p_{\theta(x,n)}} \text{ for } x \in \mathrm{dom}\Phi \text{ and } Z_n(x) = 1 \text{ otherwise.}$$

Let B_n denote the algebra generated by the cylinders of length n. Let $\mathbb{E}_\pi$ denote the expectation with respect to π and $\mathbb{E}_\pi(\ldots|B_n)$ the conditional expectation with respect to π and B_n. We then have

(3) $\mathbb{E}_\pi(Z_n) = 1$ for all n: $\mathbb{E}_\pi(Z_n) = \sum_{k=n}^{\infty} \int_{\{x|\theta(x,n) = k\}} Z_n d\pi =$

$$= \sum_{k=n}^{\infty} p_k^{-1}\cdot\pi\{x|\ \theta(x,n) = k \ \&\ x_k = 1\} = \sum_{k=n}^{\infty} p_k^{-1}\cdot p_k\cdot\pi\{x|\ \theta(x,n) = k\} =$$

$$= \sum_{k=n}^{\infty}\pi\{x|\ \theta(x,n) = k\} = 1.$$

The third equality is a consequence of the fact that $\{x|\ \theta(x,n) = k\} \in B_{k-1}$ and $\{x|\ x_k = 1\} \in B_k$, so that these events are independent with respect to π. The last equality follows from the assumption that $\pi(\mathrm{dom}\Phi) = 1$.

(4) $\mathbb{E}_\pi(Z_n|B_{n-1})(x) = 1$ for all x: by definition, $\mathbb{E}_\pi(Z_n|B_{n-1})$ is B_{n-1}–measurable and satisfies for $B \in B_{n-1}$ $\int_B Z_n d\pi = \int_B \mathbb{E}_\pi(Z_n|B_{n-1})d\pi$.

Now $\int_B Z_n d\pi = \sum_{k=n}^{\infty} \int_{\{x|\theta(x,n) = k\}\cap B} x_k\cdot p_k^{-1} d\pi = \sum_{k=n}^{\infty} p_k^{-1}\cdot\pi\{x|\ \theta(x,n) = k \ \&\ x_k = 1\}\cap B =$

$= \sum_{k=n}^{\infty} p_k^{-1}\cdot p_k\cdot\pi\{x|\ \theta(x,n) = k\}\cap B = \pi B$. Since $\mathbb{E}_\pi(Z_n|B_{n-1})$ is constant on cylinders of length n–1, this implies that $\mathbb{E}_\pi(Z_n|B_{n-1})$ equals 1 everywhere.

(5) Since $\lim_{n\to\infty} p_n = p \in (0,1)$, there is $\delta \in (0,1)$ and $n_0 \in \mathbb{N}$ such that for $n \geq n_0$: $\delta < p_n < 1-\delta$. Then, again for $n \geq n_0$: $0 \leq Z_n \leq p_n^{-1} < \delta^{-1}$, hence the Z_n are uniformly bounded.

By Theorem 3 in Feller [26,243]:

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n} Z_k(x) = \lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}\frac{x_{\theta(x,k)}}{p_{\theta(x,k)}} = \lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}\frac{\Phi(x)_k}{p_{\theta(x,k)}} = 1 \quad \pi\text{–a.e.}$$

But, generally, if (a_n) and (b_n) are sequences of positive reals,

$$\text{if } \lim_{n\to\infty} b_n = p \text{ and } \lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}\frac{a_k}{b_k} = 1, \text{ then } \lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n} a_k = p.$$

Hence, still under the assumption $\pi(\mathrm{dom}\Phi) = 1$:

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n} \Phi(x)_k = p.$$

We now drop the assumption. Note that $\pi(\mathrm{dom}\Phi\cap\Phi^{-1}(\mathrm{LLN}(p))) = \pi(\mathrm{dom}\Phi)$ is equivalent to $\pi(\Phi^{-1}\mathrm{LLN}(p)|\mathrm{dom}\Phi) = 1$ (under the assumption $\pi(\mathrm{dom}\Phi) > 0$, but otherwise there is nothing to prove), so for the general case it suffices to replace in the above proof π by $\pi(\ldots|\mathrm{dom}\Phi)$. ▯

4.6.3 Corollary Let $\mathcal{H}$ be a countable set of place selections and $\pi = \prod_n(1-p_n,p_n)$ a product measure such that $\lim_{n\to\infty} p_n = p$. Then $\pi C(\mathcal{H},p) = 1$.

We next investigate the sensitivity of $R(\mu_p)$ to the underlying measure.

4.6.4 Lemma Let μ,ν be computable measures on 2^ω. $\mu\perp\nu$ is equivalent to either of the following statements: (i) there exists a total recursive sequential test N with respect to μ such that $\nu N = 1$; (ii) for each rational $\varepsilon > 0$, there exists a Π_1 set A such that $\nu A > 1-\varepsilon$ and $\mu A = 0$.

Proof Trivially, (i) and (ii) imply $\mu\perp\nu$. For $\mu\perp\nu$ implies (i) we use the following equivalence

$$\mu\perp\nu \text{ iff } \forall\varepsilon>0\ \exists C\in\Sigma_0\ (\nu C > 1-\varepsilon \ \&\ \mu C < \varepsilon)$$

and we take advantage of the Π_2 statement on the right hand side. Let f: $\mathbb{Q}^+ \to\Sigma_0$ be a total recursive function which for each ε in $\mathbb{Q}^+$ gives $f(\varepsilon)$ in Σ_0 such that $\nu f(\varepsilon) > 1-\varepsilon$ and $\mu f(\varepsilon)<\varepsilon$. Such a function exists by the computability of μ and ν. Let $N = \bigcap_n\bigcup_i f(2^{-i-n-1})$. Obviously N is Π_2. Since for each n and i, $\mu f(2^{-i-n-1}) < 2^{-i-n-1}$, $\mu\bigcup_i f(2^{-i-n-1})$ is computable (see the proof of the first effective Borel–Cantelli lemma (3.3.1)). Hence N is a total recursive sequential test with respect to μ. On the other hand, for each n and all i, $\nu\bigcup_i f(2^{-i-n-1}) \geq \nu f(2^{-i-n-1}) \geq 1-2^{-i-n-1}$, so $\nu\bigcup_i f(2^{-i-n-1}) = 1$. For (i) implies (ii), reverse the roles of μ and ν in (i), obtaining N =

$\bigcap_n O_n$ such that $\mu N = 1$, $\nu N = 0$ and each O_n in Σ_1; then some $(O_n)^c$ will do. ∎

The following beautiful criterion for singularity of product measures is due to Kakutani [39][1].

4.6.5 Lemma Let $\mu = \prod_n(1-p_n,p_n)$, $\pi = \prod_n(1-q_n,q_n)$ be product measures on 2^ω such that for some $\delta > 0$ and all n, $\delta < p_n,q_n < 1-\delta$. If $\sum_n(p_n - q_n)^2$ diverges, then μ and π are mutually singular; on the other hand, if $\sum_n(p_n - q_n)^2$ converges, then μ and π are equivalent.

It follows from the zero-one law that product measures on 2^ω are either singular or equivalent, but Kakutani's theorem provides us with a criterion to distinguish these cases and this is what we shall use to finish the proof of Theorem 4.6.1.

Let $p_n := p\cdot(1 + (n+1)^{-\frac{1}{2}})$, $\pi = \prod_n(1-p_n,p_n)$, then π is computable and since $\sum_n(p - p_n)^2 = \sum_n p^2\cdot n^{-1} = \infty$, $\pi \perp \mu_p$. By corollary 4.6.3, $\pi C(p) = 1$. By lemma 4.6.4, $\pi R(\mu_p) = 0$. This completes the proof of Theorem 4.6.1. ∎

We may extract the following information from the proof of Theorem 4.6.1:

4.6.6 Corollary Let $\pi = \prod_n(1-p_n,p_n)$ be a computable product measure, $p \in (0,1)$ a computable real.

(i) $\pi C(p) = 1$ iff $\lim_{n\to\infty} p_n = p$

(ii) $\pi R(\mu_p) = 1$ iff $\sum_n(p - p_n)^2$ converges.

(iii) $\pi(C(p)\cap R(\mu_p)^c) = 1$ iff $\lim_{n\to\infty} p_n = p$ but $\sum_n(p - p_n)^2$ diverges.

4.6.7 Remark We saw in 2.6.2.2 that there exist countably many recursive place selections Φ such that Kollektivs of Ville's type can never belong to the domain of Φ. But if $x \notin \mathrm{dom}\Phi$, then the statement "$x \in \mathrm{dom}\Phi$ implies $\Phi x \in LLN(p)$" is uninformative. (A failure is significant only when preceded by a serious effort.) Similarly, although we have formally proved that $\pi(C(p)\cap R(\mu_p)^c) = 1$ if π satisfies the right hand side of (iii), the theorem and its corollary are interesting only for a *subclass* of the recursive place selections, namely for those Φ for which $\pi(\mathrm{dom}\Phi) = 1$ if π is a product measure whose marginals converge to p.

The reader will have noticed undoubtedly that Ville's theorem in its original formulation uses the law of the iterated logarithm essentially, whereas it is absent from our proof. This leads to the following question: is the difference between C(p) and $R(\mu_p)$ due *entirely* to the law of the iterated logarithm, in the sense that each sequence in $C(p)\cap R(\mu_p)^c$ fails to satisfy it?

Interestingly, it is a corollary of Theorem 4.6.1 that this is not so: if π is, e.g., the product measure $\prod_n(1-p_n,p_n)$ with $p_n = p\cdot(1 + (n+1)^{-\frac{1}{2}})$, then π assigns measure one to the set of sequences which satisfy the law of the iterated logarithm (for μ_p). To see this, we need a general form of the

Law of the iterated logarithm (Kolmogorov [45])

Let $\mu = \prod_n(1-q_n,q_n)$ be a product measure and define the variance s_n by $s_n := \sum_{k=1}^{n} q_n\cdot(1-q_n)$.

Then

$$\text{(1)}\quad \text{for } \beta > 1, \text{ for } \mu\text{–a.a. } x:\ \exists m\ \forall n\geq m\ \left|\sum_{k=1}^{n} x_k - \sum_{k=1}^{n} n\cdot q_k\right| < \beta\sqrt{2s_n \log\log s_n}$$

$$\text{(2)}\quad \text{for } \beta < 1, \text{ for } \mu\text{–a.a. } x:\ \forall m\ \exists n\geq m\ \sum_{k=1}^{n} x_k - \sum_{k=1}^{n} n\cdot q_k > \beta\sqrt{2s_n \log\log s_n}$$

$$\text{for } \beta < 1, \text{ for } \mu\text{–a.a. } x:\ \forall m\ \exists n\geq m\ \sum_{k=1}^{n} n\cdot q_k - \sum_{k=1}^{n} x_k > \beta\sqrt{2s_n \log\log s_n}.$$

If all q_n are equal to p, we get back the form of the law stated in 2.6.2.2. Let $LIL(\mu_p)$ denote the set of sequences which satisfy the law for the measure μ_p. Let π be the product measure constructed above. We show that $\pi LIL(\mu_p) = 1$.

If for instance for some $\alpha < 1$,

$$\pi\{x \mid \exists m\ \forall n\geq m \sum_{k=1}^{n} x_k > p\cdot n - \alpha\sqrt{2n\cdot p\cdot(1-p)\log\log n}\} = 1,$$

so that $\pi LIL(\mu_p) = 0$, then, by the general form of the law of the iterated logarithm, for $\beta > 1$ and n sufficiently large:

$$p\cdot n + p\cdot\sum_{k=1}^{n}\frac{1}{\sqrt{k+1}} - \beta\sqrt{2s_n \log\log s_n} > p\cdot n - \alpha\sqrt{2n\cdot p\cdot(1-p)\log\log n},$$

hence

$$p\cdot\sum_{k=1}^{n}\frac{1}{\sqrt{k+1}} > \beta\sqrt{2s_n \log\log s_n} - \alpha\sqrt{2n\cdot p\cdot(1-p)\log\log n};$$

but this is easily seen to be false, since the left hand side is $O(\sqrt{n})$, whereas the right hand side is $O(\sqrt{(n\log\log n)})$. An analogous argument for the upper bound then shows that $\pi LIL(\mu_p) = 1$.

On the other hand, it is possible to construct uncountably many Church-random sequences (with parameter p) which do not satisfy the law of the iterated logarithm (for μ_p) if we use product measures μ_p whose marginals converge to p slower than those of π. Choose a such that $-\frac{1}{2} < a < 0$ and put $q_n := p\cdot(1 + (n+1)^a)$, $\mu := \prod_n(1-q_n, q_n)$.
We now do have, for $\alpha < 1$,

$$\mu\{x \mid \exists m\ \forall n\geq m \sum_{k=1}^{n} x_k \ > \ p\cdot n \ - \ \alpha\sqrt{2n\cdot p\cdot(1-p)\mathrm{loglog}n}\ \} = 1;$$

by the general form of the law of the iterated logarithm, it suffices to show that for some $\beta > 1$ and all n sufficiently large:

$$p\cdot n \ + \ p\cdot\sum_{k=1}^{n}(k+1)^a \ - \ \beta\sqrt{2s_n\mathrm{loglog}s_n} \ > \ p\cdot n \ - \ \alpha\sqrt{2n\cdot p\cdot(1-p)\mathrm{loglog}n}\,;$$

in other words, that

$$p\cdot\sum_{k=1}^{n}(k+1)^a \ > \ \beta\sqrt{2s_n\mathrm{loglog}s_n} \ - \ \alpha\sqrt{2n\cdot p\cdot(1-p)\mathrm{loglog}n}\,.$$

But now the left hand side is $O(n^{a+1})$, with $a+1 > \frac{1}{2}$ and the right hand side is still $O(\sqrt{(n\mathrm{loglog}n)})$. Hence not only $\mu(C(p)\cap R(\mu_p)^c) = 1$ (since $\mu\perp\mu_p$), but also $\mu(C(p)\cap LIL(\mu_p)^c) = 1$.

We may thus conclude that a part of, but *only* a part of, the difference between C(p) and R(μp) is caused by the law of the iterated logarithm. The proof of Theorem 4.6.1 shows that Church–random sequences may also fail to satisfy properties which are essentially different from the law of the iterated logarithm.

The rest of this section is rather technical: we investigate what remains of Corollary 4.6.6 if we drop the assumption that π be a product measure. We now obtain a theorem which connects the different concepts of randomness with different types of convergence of measures.

4.6.8 Definition Let μ and ν be measures on 2^ω and let $T: 2^\omega \to 2^\omega$ be the left shift. We say that the sequence of measures $(\mu T^{-n})_{n\in\mathbb{N}}$ *converges strongly* to ν if for all Borel sets A, $\lim_{n\to\infty} \mu T^{-n}A = \nu A$. We say that $(\mu T^{-n})_{n\in\mathbb{N}}$ *converges weakly* to ν if for all Borel sets A such that $\nu\partial A = 0$ (where ∂A is the boundary of A), $\lim_{n\to\infty} \mu T^{-n}A = \nu A$.

The next lemma considerably simplifies the last condition:

4.6.9 Lemma (See Billingsley [4].) $(\mu T^{-n})_{n\in\mathbb{N}}$ converges weakly to ν if for all cylinders [w]: $\lim_{n\to\infty} \mu T^{-n}[w] = \nu[w]$.

Part (i) of Corollary 4.6.6 can now be restated thus: $\pi C(p) = 1$ iff $(\pi T^{-n})_{n\in\mathbb{N}}$ converges weakly to μ_p. We shall see presently that one half of this result can be salvaged even without the assumption that π be a product measure.

4.6.10 Theorem Let μ be a measure such that for all place selections Φ recursive in μ, if $\mu(\mathrm{dom}\Phi) = 1$, then $\mu(\Phi^{-1}LLN(p)) = 1$. Then $(\mu T^{-n})_{n\in\mathbb{N}}$ converges weakly to μ_p. In particular, if μ is computable and $\mu C(p) = 1$, then $(\mu T^{-n})_{n\in\mathbb{N}}$ converges weakly to μ_p.

Proof Suppose not; then there exists a smallest binary string s such that $\lim_{n\to\infty} \mu T^{-n}[s] \neq \mu_p[s]$. Without loss of generality we may suppose that for some rational $\varepsilon > 0$, for some sequence (N_i) recursive in μ and for all i:

$$\mu T^{-N_i}[s] > \mu_p[s] + \varepsilon.$$

Define *for this particular sequence* (N_i) and for *all* binary words v a place selection Ψ_v by

$$\psi_v(x(m)) = \begin{cases} 1 & \text{if } \exists i\,(N_i + |v| = m)\ \&\ \exists u\in 2^{<\omega}(uv = x(m)) \\ 0 & \text{otherwise}. \end{cases}$$

Recall that "v⊂w" means that v is a strict initial segment of w and that ‹› denotes the empty string.

Claim 1

$$\forall w\in 2^{<\omega}\ \Big(\forall v\subset w\ \lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} \Psi_v(x)_k = p \longrightarrow \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} 1_{[w]}(T^{N_i}x) = \mu_p[w]\Big).$$

Proof of claim 1 We use induction on w. If w = 1, the hypothesis of the claim implies that

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} \Psi_{‹›}(x)_k = p,$$

which is by definition of $\Psi_{‹›}$ equivalent to

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} 1_{[1]}(T^{N_i}x) = p.$$

Suppose the claim holds for w. Note that

$$\frac{\frac{1}{n}\sum_{i=1}^{n} 1_{[w1]}(T^{N_i}x)}{\frac{1}{n}\sum_{i=1}^{n} 1_{[w]}(T^{N_i}x)} = \frac{\sum_{k=1}^{|\Psi_w(x(N_n))|} \Psi_w(x)_k}{|\Psi_w(x(N_n))|}.$$

The hypothesis of the claim implies that the right hand side converges to p; the hypothesis of induction implies that the denominator of the left hand side converges to $\mu_p[w]$. It follows that the numerator of the left hand side must converge to $\mu_p[w1]$. This concludes the proof of claim 1.

Claim 2 Under the hypothesis of the theorem, for $\alpha = 0,1$:

$$\forall v \in 2^{<\omega}\, \mu(\mathrm{dom}\Psi_v) = 1 \;\&\; \mu\{x|\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} 1_{[v\alpha]}(T^{N_i}x) = \mu_p[v\alpha]\} = 1.$$

Proof of claim 2 We use induction on v. Trivially, $\mu(\mathrm{dom}\Psi_{\langle\rangle}) = 1$ and hence for $\alpha = 0,1$:

$$\mu\{x|\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} 1_{[\alpha]}(T^{N_i}x) = \mu_p[\alpha]\} = 1,$$

by claim 1 and the hypothesis of the theorem. Suppose the claim holds for $u \subset v$, then again by claim 1 and the hypothesis of the theorem:

$$\mu\{x|\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} 1_{[v]}(T^{N_i}x) = \mu_p[v]\} = 1.$$

It follows that μ–a.e. v occurs infinitely often at coordinates starting with an index N_i+1; hence $\mu(\mathrm{dom}\Psi_v) = 1$. Then, as a consequence of claim 1 and the hypothesis of the theorem:

$$\mu\{x|\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} 1_{[v\alpha]}(T^{N_i}x) = \mu_p[v\alpha]\} = 1.$$

This concludes the proof of claim 2.

Claim 2 implies that for the particular string s determined at the outset,

$$\mu\{x|\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} 1_{[s]}(T^{N_i}x) = \mu_p[s]\} = 1.$$

By the dominated convergence theorem,

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \mu T^{-N_i}[s] = \lim_{n\to\infty} \int_{2^\omega} \frac{1}{n}\sum_{i=1}^{n} 1_{[s]}(T^{N_i}x)d\mu(x) =$$
$$= \int_{2^\omega} \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} 1_{[s]}(T^{N_i}x)d\mu(x) = \mu_p[s],$$

a contradiction. □

A converse to the theorem is not to be expected. Indeed, the conclusion of the theorem is probably too weak; it is plausible that the hypothesis implies some kind of asymptotic independence condition.

We next generalize the second part of Corollary 4.6.6 to arbitrary computable measures.

4.6.11 Lemma Let μ, ν be computable measures such that μ is not absolutely continuous with respect to ν. Then for some total recursive sequential test N with respect to ν, $\mu N > 0$.

Proof We showed in Example 3.4.6 that one can define a recursive sequential test N with respect to ν such that $\mu N > 0$, using the likelihood ratio $\mu[w]/\nu[w]$. For reasons explained at length in 3.4, it is difficult, if not impossible, to prove that N is a *total* recursive sequential test with respect to ν. We therefore borrow an idea of Gaifman and Snir [34,518]. Choose $\varepsilon > 0$. Since μ is not absolutely continuous with respect to ν, there exists a sequence (C_i) of Σ_0 sets such that $\mu\bigcap_i C_i > \varepsilon$ and $\nu\bigcap_i C_i = 0$. Let (D_k) be a recursive enumeration of the Σ_0 sets. Define

$$f(n) := \min\{k > n \mid \mu D_k > \varepsilon \ \&\ \nu D_k < 2^{-k}\}.$$

Let $N = \bigcap_n \bigcup_{m\geq n} D_{f(m)}$, then $\mu N > \varepsilon$. That N is a total recursive sequential test is shown by an argument similar to the proof of the effective first Borel–Cantelli lemma, 3.3.1. □

Gaifman [34,519] asks whether μ and ν can already be separated by a Π_1 set. An affirmative answer would follow from lemma 4.6.4 in the unlikely event that the Lebesgue decomposition of μ with respect to ν, namely $\mu = \mu_0 + \mu_1$, where $\mu_0 << \nu$ and $\mu_1 \perp \nu$, can be achieved with computable μ_0, μ_1. It is more probable, however, that one can produce a counterexample to computable Lebesgue decomposition in this way.

4.6.11 Lemma μ is absolutely continuous with respect to μ_p iff $(\mu T^{-n})_{n\in\mathbb{N}}$ converges strongly to μ_p.

Proof $\Rightarrow$ T is *strongly mixing* with respect to μ_p, i.e. for f,g in $L^1(\mu_p)$, $\lim_{n\to\infty} \int (f\circ T^n)\cdot g d\mu_p$ $= (\int f d\mu_p)\cdot(\int g d\mu_p)$. Let g in $L^1(\mu_p)$ be a Radon–Nikodym derivative of μ with respect to μ_p, then for all Borel sets A: $\mu A = \int g\cdot 1_A d\mu_p$. Hence $\lim_{n\to\infty} \mu T^{-n}A = \lim_{n\to\infty} \int g\cdot(1_A\circ T^n)d\mu_p =$ $\mu_p A\cdot(\int g d\mu_p) = \mu_p A$.

$\Leftarrow$ (proof due to M.S. Keane) Suppose μ is not absolutely continuous with respect to μ_p. Let A be a Borel set with $\mu_p A = 0$ and such that $\mu A > 0$ is maximal. We construct a Borel set B such that $\mu_p B = 0$ and for all n, $\mu T^{-n}B = \mu A$. Let $B_1 := TA$. Claim: B_1 is also Borel. For we can split T into two homeomorphisms T_0: $[0] \to 2^\omega$, T_1: $[1] \to 2^\omega$ defined by $T_i(ix) = x$, for i = 0,1. Since the T_i are homeomorphisms, the sets $T_i(B\cap[i])$ are Borel; but $TB = T_0(B\cap[0])\cup T_1(B\cap[1])$. Clearly $\mu_p B = 0$. Since $T^{-1}B_1 \supseteq A$ and A was chosen to have maximal μ-measure, $\mu T^{-1}B_1 = \mu A$. For each n, repeat the above argument with T^n replacing T, yielding B_n. Put $B := \bigcup_n B_n$, then $\mu_p B = 0$ and $\mu T^{-n}B = \mu A$ for all n. □

4.6.12 Theorem Let μ be a computable measure. Then $\mu R(\mu_p) = 1$ iff $(\mu T^{-n})_{n\in\mathbb{N}}$ converges strongly to μ_p.

Proof By lemma 4.6.11, $\mu R(\mu_p) = 1$ implies that μ is absolutely continuous with respect to μ_p. The converse is trivial. Now apply the previous lemma. □

4.7 Digression: the difference between randomness and 2-randomness We are interested in the size of the difference between $R(\lambda)$ and $R_2(\lambda)$, the randomness notion that was defined in 3.2.4.1. We have seen in 3.2.4 that $R(\lambda)\cap R_2(\lambda)^c$ is non-empty. On the other hand, by lemma 4.6., there is no computable measure μ such that $\mu(R(\lambda)\cap R_2(\lambda)^c) = 1$: if $\mu R_2(\lambda) = 1$, then $\mu\perp\lambda$, which implies $\mu R(\lambda) = 0$. (Note that, for all we know, there might be a computable μ such that $\mu(R(\lambda)\cap R_2(\lambda)^c) > 0$.)

We now show, as an application of the techniques developed in 4.1-6, that $R(\lambda)\cap R_2(\lambda)^c$ is indeed large: there exists a non-atomic Δ_2 definable measure μ_x such that $\mu_x(R(\lambda)\cap R_2(\lambda)^c) = 1$.

To prove this, we need a random measure, that is, a family of measures $(\mu_x)_{x\in 2^\omega}$ defined as follows:

$$\mu_x = \prod_n (1-p_n^x, p_n^x), \text{ where } p_n^x = \begin{cases} 3/4 & \text{if } x_n = 1 \\ 1/4 & \text{if } x_n = 0. \end{cases}$$

It is easily shown that for each Borel set B, the mapping $x \to \mu_x B$ is measurable. Hence we may define a measure μ on 2^ω by

$$\mu(A\times B) = \int_A \mu_x B d\lambda(x).$$

μ is obviously computable, hence $R(\mu)$ is well defined. Using a construction exactly parallel to the Fubini theorem for recursive sequential tests (Theorem 4.4.4), one can demonstrate that for all $x \in R(\lambda)$, $\mu_x R(\mu)_x = 1$. For this, it suffices to show that for each recursive sequential test N with respect to μ, $\{x| \mu_x N_x > 0\}$ is contained in a recursive sequential test with respect to λ. This can be done if we change slightly the definition of the functions f_k occurring in the proof of Theorem 4.4.4. We now put

$$f_0(x) = 0 \text{ for all } x$$

$$f_k(x) = \sum_{\{i\leq k| x\in [w^i]\}} \mu_x[v^i], \text{ for } k\geq 1;$$

the rest of the proof then goes through almost literally.

We now show that for $x \in R(\lambda)$, $R(\mu)_x \subseteq R(\lambda)$. For this, it suffices to show that the mapping π_2: $2^\omega\times 2^\omega \to 2^\omega$ defined by $\pi_2<x,y> = y$ is such that for any recursive sequential test N with respect to λ, $\pi_2^{-1}N$ is a recursive sequential test with respect to μ, for in that case, $<x,y> \in R(\mu)$ implies $y \in R(\lambda)$. (Observe that $x \in R(\lambda)$ implies that $R(\mu)_x \neq \emptyset$.) Now $\pi_2^{-1}N$ is obviously Π_2 and is a recursive sequential test, since for all Borel sets A: $\mu\pi_2^{-1}A = \lambda A$.

We thus have that for each $x \in R(\lambda)$, $\mu_x R(\lambda) = 1$. In particular, this is true of the Δ_2 sequence constructed in 3.2.2.3. Fix such a Δ_2 definable μ_x; this μ_x is then recursive in $\emptyset'$. It is not difficult to see that $\mu_x \perp \lambda$; either by Kakutani's theorem (4.6.5) or by observing that $R(\mu) \subseteq R(\lambda\times\lambda)^c$ and applying the Fubini theorem 4.4.4 to conclude that for $x \in R(\lambda)$, $\lambda R(\mu)_x = 0$.

Since our μ_x is singular to λ, we may perform the construction of lemma 4.6.4 (ii) recursively in $\emptyset'$, to obtain a Δ_2 definable sequence (C_n) of Σ_0 sets C_n, such that $\lambda \bigcap_n C_n = 0$ and $\mu_x \bigcap_n C_n > 0$. Now $\bigcap_n C_n$ is Π_2, hence $\mu_x R_2(\lambda)^c > 0$ and since $R_2(\lambda)^c$ is a tailset and μ_x a product measure, we get in fact $\mu_x(R(\lambda)\cap R_2(\lambda)^c) = 1$.

Notes to Chapter 4

1. A simple proof of Kakutani's theorem has recently been published by S.D. Chatterji. See S.D. Chatterji, Martingale theory: An analytical formulation with some applications in

analysis, in: Letta (ed.), Probability and analysis, *Lecture Notes in Mathematics* **1206**, Springer-Verlag (1986).