A PROOF OF THE ERGODIC THEOREM USING NONSTANDARD ANALYSIS

TRISTRAM DE PIRO

ABSTRACT. The following paper follows on from [2] and gives a rigorous proof of the Ergodic Theorem, using nonstandard analysis.

1. The Ergodic Theorem

This paper gives a nonstandard proof of the Ergodic Theorem. Although, a standard proof of this result is known, we consider the approach taken to be useful. The use of nonstandard analysis is important in Physics, where we can represent diffusions, either classically as in the heat equation, or in electrodynamics, using infinitesimals. The use of nonstandard analysis, probability and the method of the Ergodic Theorem, finds an application in the computation of the entropy of transformations, involved in the work of Boltzmann. An important result in this area is the Maxwell-Boltzmann distribution for the probabilities of occupying a given quantum state. A refinement in the context of electrochemistry could be based on a recent paper [5] where the energies of atomic systems are calculated using the wave equation.

There are many versions of the Ergodic Theorem. The one we will prove in this paper, using nonstandard analysis, is given in Theorem 1.1. The nonstandard method of proof, in particular the proof of Theorem 1.3, is mainly due to Kamae, see [2], but there are a number of details omitted from his original paper, in particular the proofs of Lemma 1.11 and Theorem 1.15, using the combinatorial footnote 11. We show, in Theorem 1.3, that the Ergodic Theorem holds for a nonstandard probability space $(K, \mathfrak{B}, P, \phi)$, with a result from the Appendix. Lemmas 1.4, 1.6, 1.7 and Theorem 1.9 reduce the main result Theorem 1.1 to the problem of finding a typical element for the space $([0,1]^{\mathcal{N}}, \mathfrak{C}, \rho, \sigma)$, a result originally due to de Ville. This result, under a certain hypothesis of weak convergence of probability measures, is proved in Lemma 1.11. The hypothesis is proved in Lemmas 1.12, 1.14

and Theorem 1.15, by finding a probability measure ρ_β supported at a periodic element $\beta \in [0,1]^{\mathcal{N}}$.

Theorem 1.1. *Ergodic Theorem*

Let $(\Omega, \mathfrak{C}, \mu)$ be a probability space, and let T be a measure preserving transformation, then, if $g \in L^1(\Omega, \mathfrak{C}, \mu)$;

$\diamond g(\omega) = lim_{n\rightarrow\infty}\frac{1}{n}\sum_{i=0}^{n-1} g(T^i\omega)$

exists for almost all $\omega \in \Omega$, with respect to μ, and, $\diamond g \in L^1(\Omega, \mathfrak{C}, \mu)$, with;

$\int_\Omega \diamond g d\mu = \int_\Omega g d\mu$

Remarks 1.2. *There are a number of good standard proofs of this result. A particular good reference is* [3]*. However, the reader should be aware that it is assumed there that $\mathfrak{C}$ is complete and T is*invertible, *in the sense that T is one-one and onto, and both T and T^{-1} are measurable. A m.p.t is then required to satisfy $\mu(C) = \mu(T^{-1}C)$ for all $C \in \mathfrak{C}$. We will not require these assumption in the proofs of this section, in the sense that we only require a m.p.t to be a measurable T with $\mu(C) = \mu(T^{-1}C)$ for all $C \in \mathfrak{C}$. In* [3]*, a seemingly stonger result is shown, (under the above assumptions), namely that if $C \in \mathfrak{C}$, with $T^{-1}(C) = C$, then;*

$\int_C \diamond g d\mu = \int_C g d\mu$ $(*)$

from which it easily follows that if $\mathfrak{C}'$ is the sub σ-algebra of all T-invariant sets, where a set C is T invariant in [3]*, if $T^{-1}C = C$ a.e $d\mu$, then $\diamond g = E(g|\mathfrak{C}')$, $(**)$. In the particular case when T is ergodic, that is every T invariant set has measure* 0 *or* 1*, we obtain the well known result that $\diamond g = E(g)$ a.e $d\mu$, $(***)$. However, this result $(*)$ follows easily from our Theorem 1.1. as we can, wlog, assume that $\mu(C) > 0$, and then restrict and rescale the measure. Of course, we even obtain a slight strengthening of $(*)$, by our weaker assumption on a m.p.t, and obtain similar strengthenings of $(**)$ and $(***)$. (It is not necessary to restrict attention to real valued functions, in the statement of the theorem, the complex version follows immediately from the real case).*

As usual, we work in an $\aleph_1$-saturated model. Let $k \in {}^*\mathcal{N}_{>0}$ be infinite, and let $K = \{x \in {}^*\mathcal{N} : 0 \leq x < k\}$. We let $\mathfrak{K}$ be the algebra of all internal subsets of K. Observe that as K is hyperfinite, $\mathfrak{K}$ is a hyperfinite ${}^*\sigma$-algebra. We let ν denote the counting measure, defined by setting $\nu(A) = \frac{Card(A)}{k}$, for $A \in \mathfrak{K}$. We adopt some of the notation of Section 3 in [4], and let $P = {}^\circ\nu$. By Theorem 3.4, and remarks before Lemma 3.15 of [4], P extends uniquely to the completion $\mathfrak{B}$ of the σ-algebra, $\sigma(\mathfrak{K})$, generated by $\mathfrak{K}$. It is clear that $(K, \mathfrak{B}, P)$ is a probability space, it is also the Loeb space associated to $(K, \mathfrak{K}, \nu)$. We let $\phi : K \rightarrow K$ denote the map defined by;

$\phi(x) = x + 1$, if $0 \leq x < k - 1$

$\phi(x) = 0$, if $x = k - 1$

Clearly, ϕ is invertible, internal, preserves the counting measure ν, and $\phi^{-1}(\sigma(\mathfrak{K})) = \sigma(\mathfrak{K})$. Then $P \circ \phi^{-1}$ defines a measure on $(K, \sigma(\mathfrak{K}), P)$, extending ν. By Theorem 3.4(ii) of [4], it agrees with P. By definition of the completion, $P \circ \phi^{-1}$ agrees with P on $(K, \mathfrak{B}, P)$, so ϕ, and similarly ϕ^{-1} are m.p.t's. We will first prove the following;

Theorem 1.3. *The ergodic theorem, as stated in Theorem 1.1, holds for $(K, \mathfrak{B}, P, \phi)$.*

Proof. Let $g \in L^1(K, \mathfrak{B}, P)$, without loss of generality, we can assume that $g \geq 0$. For $x \in K$, we let;

$\overline{g}(x) = limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} g(\phi^i x)$

$\underline{g}(x) = liminf_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} g(\phi^i x)$

In order to prove the theorem, it is sufficient to show that $\overline{g}$ is integrable and;

$\int_K \overline{g} dP \leq \int_K g dP \leq \int_K \underline{g} dP$ (†)

Then, as $\underline{g} \leq \overline{g}$, we must have equality in (†), so $\underline{g} = \overline{g}$ a.e dP, that is $\diamond g$ exists a.e dP, and;

$\int_K \diamond g dP = \int_K g dP$

as required.

Now let $M \in \mathcal{N}_{>0}$, then, as $\overline{g}$ is $\mathfrak{B}$-measurable, see [7], $min(\overline{g}, M)$ is integrable with respect to P. Let $\epsilon > 0$ be standard, then we can apply Theorem 2.1 in the Appendix to this paper, and Definition 3.9 and Remarks 3.10 of [4], to obtain internal functions $F, G : K \rightarrow^* \mathcal{R}$, with $g \leq F$ and $G \leq min(\overline{g}, M)$, such that;

$|\int_A g dP - \frac{1}{k}{}^*\sum_{x\in A} F(x)| < \epsilon$

$|\int_A min(\overline{g}, M) dP - \frac{1}{k}{}^*\sum_{x\in A} G(x)| < \epsilon$, for all internal $A \subset K$, (††).

Now observe that $\overline{g}$ is ϕ-invariant,([1]). Fixing $x \in K$, by the definition of $\overline{g}$, we can find $n \in \mathcal{N}_{>0}$ such that;

$min(\overline{g}(x), M) \leq \frac{1}{n}\sum_{i=0}^{n-1} g(\phi^i x) + \epsilon$ $(*)$

Then, if $0 \leq m \leq n-1$, we have;

$G(\phi^m x) \leq min(\overline{g}(\phi^m x), M)$, by definition of G

$= min(\overline{g}(x), M)$, by ϕ invariance of $\overline{g}$

$\leq \frac{1}{n}\sum_{i=0}^{n-1} g(\phi^i x) + \epsilon$, by $(*)$

$\leq \frac{1}{n}\sum_{i=0}^{n-1} F(\phi^i x) + \epsilon$, by definition of F

Therefore,

$\sum_{i=0}^{n-1} G(\phi^i x) \leq n(\frac{1}{n}\sum_{i=0}^{n-1} F(\phi^i x) + \epsilon) = \sum_{i=0}^{n-1} F(\phi^i x) + n\epsilon$ $(**)$

Now let $S_G : [1, k) \times K \rightarrow {}^*\mathcal{R}$ be defined by;

$S_G(n, x) = {}^*\sum_{i=0}^{n-1} G(\phi^i x)$

and, similarly, define S_F. By Definition 2.19 of [4], and using the facts that K is $*$-finite, and G, F are internal, S_G and S_F are internal. Then, the relation $(**)$ becomes the internal relation on $[1, k) \times K$, given

[1]There is a probably a proof of this result in the literature, but we supply one here. Fix $x \in K$. Let $A_m = \frac{1}{m}\sum_{i=0}^{m-1} g(\phi^i x)$ and let $B_m = \frac{1}{m}\sum_{i=0}^{m-1} g(\phi^{i+1} x)$. Then a simple calculation shows that $\frac{mB_m + g(x)}{m+1} = A_{m+1}$. Hence, $|B_m - A_{m+1}| = |\frac{A_{m+1} - g(x)}{m}|$, $(*)$. Suppose that $\overline{g}(x) = t < \infty$, $(**)$, (the case when $\overline{g}(x) = \infty$ is similar), and $\overline{g}(\phi x) < t$, $(***)$, (the case $\overline{g}(\phi x) > t$ is again similar). Then, by $(***)$, there exists $\delta > 0$, such that, for $m \geq m_0$, $B_m < t - \delta$. By $(*)$ and $(**)$, we can find $m_1 \geq m_0$, such that $|B_m - A_{m+1}| < \frac{\delta}{2}$, for $m \geq m_1$. Again, by $(*)$, we can find $m_2 \geq m_1 \geq m_0$, such that $A_{m_2+1} > t - \frac{\delta}{2}$. This clearly gives a contradiction.

by $R(n,x)$ iff $S_G(n,x) \leq S_F(n,x)+n\epsilon$. Using the fact above, that the fibres of R over K are non-empty, by transfer of the corresponding standard result, we can find an internal function $T : K \rightarrow [1,k)$, which assigns to $x \in K$, the least $n \in [1,k)$, for which $(**)$ holds. Moreover, as we have observed in $(*)$, $T(x)$ is standard, for all $x \in K$. By Lemma 3.11, $r = max_{x\in K}T(x)$ exists and is standard. Now, define T_j hyper inductively by;

$T_0 = 0$ and $T_j = T_{j-1} + T(T_{j-1})$

and let J be the first j such that $k - r \leq T_j < k$.([2])

Observe that T_j defines an internal partition of the interval $[0, T_{J-1}] \subset [0,k)$, into $J-1$ blocks of step size $T_j - T_{j-1} = T((T_{j-1})$. Hence, we can write;

$\frac{1}{k}{}^*\sum_{x=0}^{T_J-1} G(x) = \frac{1}{k}{}^*\sum_{j=0}^{J-1}{}^*\sum_{i=0}^{T(T_j)-1} G(\phi^i T_j)$

$\leq \frac{1}{k}{}^*\sum_{j=0}^{J-1}{}^*\sum_{i=0}^{T(T_j)-1} F(\phi^i T_j) + T(T_j)\epsilon$,by definition of T and $(**)$.

Now we can rearrange this last sum as;

$\frac{1}{k}{}^*\sum_{x=0}^{T_J-1} F(x) + \frac{\epsilon}{k}{}^*\sum_{j=0}^{J-1} T(T_j)$

$= \frac{1}{k}{}^*\sum_{x=0}^{T_J-1} F(x) + \frac{T_J\epsilon}{k}$

$< \frac{1}{k}{}^*\sum_{x=0}^{T_J-1} F(x) + \epsilon$

using the facts that ${}^*\sum_{j=0}^{J-1} T(T_j) = {}^*\sum_{j=0}^{J-1}(T_{j+1} - T_j) = T_J$, and $T_J < k$. Therefore, we have that;

$\frac{1}{k}{}^*\sum_{x=0}^{T_J-1} G(x) < \frac{1}{k}{}^*\sum_{x=0}^{T_J-1} F(x) + \epsilon$ $(***)$

[2]This perhaps requires some explanation. Define $I = \{m \in {}^*\mathcal{N}_{>0} : \exists! S(dom(S) = [0,m] \wedge S(0) = 0 \wedge (\forall 1 \leq j \leq m) S(j) = S(j-1) + T(S(j-1)_{modk}))\}$, $(*)$, then it is easy to see that I is internal, $I(1)$ holds, and $I(m)$ implies $I(m+1)$. Applying Lemma 2.12 of [4], $I = {}^*\mathcal{N}_{>0}$. Hence there exists an internal function f, defined on ${}^*\mathcal{N}_{>0}$, such that $f(m)$ is the unique S satisfying $(*)$. We can then define $T_j = f(j)(j)$, and clearly $T_j - T_{j-1} \leq r$. Let $V = \{j \in {}^*\mathcal{N}_{>0} : T_j < k\}$. Then, as $T \geq 1$, V is the interval $[1,t]$ for some infinite $t < k$. Then $k - r \leq T_t < k$, otherwise $T_{t+1} < k$. Then $U = \{j \in {}^*\mathcal{N}_{>0} : k - r \leq T_j < k\}$ is internal and non empty. Therefore, by transfer, it contains a first element J.

Now, observing that $\nu([T_J,k))\leq\frac{r}{k}\simeq 0$, as r is standard, we have $P([T_J,k))=0$. Hence, using $(\dagger\dagger)$, $(***)$;

$$\int_X min(\overline{g},M)dP=\int_{[0,T_J)}min(\overline{g},M)dP<\frac{1}{k}{}^*\sum_{x=0}^{T_J-1}G(x)+\epsilon$$

$$<\frac{1}{k}{}^*\sum_{x=0}^{T_J-1}F(x)+2\epsilon<\int_{[0,TJ)}gdP+3\epsilon=\int_X gdP+3\epsilon$$

Now, letting $M\rightarrow\infty$ and $\epsilon\rightarrow 0$, we can apply the MCT, to obtain;

$$\int_X\overline{g}dP\leq\int_X gdP$$

As g is integrable with respect to P, so is $\overline{g}$, and a similar argument to the above demonstrates that $\int_X gdP\leq\int_X\underline{g}dP$. Therefore, $(\dagger)$ is shown and the theorem is proved.

□

We now generalise Theorem 1.3, to obtain Theorem 1.1. We let $\mathcal{P}$ consist of spaces of the form $(\mathcal{R}^{\mathcal{N}},\mathfrak{D},\lambda,\sigma)$, where $\mathfrak{D}$ is the Borel field on $\mathcal{R}^{\mathcal{N}}$, σ is the left shift on $\mathcal{R}^{\mathcal{N}}$, and λ is a shift invariant probability measure. Note that σ is not invertible, but we require that $\lambda=\sigma_*\lambda$, so σ is a m.p.t, with respect to λ. Similarly, we let $\mathcal{Q}$ consist of spaces of the form $([0,1]^{\mathcal{N}},\mathfrak{E},\rho,\sigma)$, where $\mathfrak{E}$ is the Borel field on $[0,1]^{\mathcal{N}}$, σ is again the left shift, and ρ is a shift invariant probability measure.

We first require the following simple lemma;

Lemma 1.4. *Theorem 1.1 is true iff the Ergodic Theorem holds for all spaces in $\mathcal{P}$.*

Proof. One direction is obvious. For the other direction, let $(\Omega,\mathfrak{C},\mu,T)$ and $g\in L^1(\Omega,\mathfrak{C},\mu)$ be given. Define a map $\tau:\Omega\rightarrow\mathcal{R}^{\mathcal{N}}$ by $\tau(\omega)(n)=g(T^n\omega)$. Clearly, as g is measurable with respect to $\mathfrak{C}$ and T is a m.p.t, using the definition of the Borel field on $\mathcal{R}^m$, for finite m, we have that for a cylinder set $U\in\mathfrak{D}$, $\tau^{-1}(U)\in\mathfrak{C}$. By the definition of the Borel field on $\mathcal{R}^{\mathcal{N}}$, $\tau^{-1}(\mathfrak{D})\subset\mathfrak{C}$, ([3]). Let λ be the probability measure $\tau_*\mu$. Then λ is σ invariant, as clearly, using the fact that T is a m.p.t, $\lambda=\sigma_*\lambda$ on the cylinder sets in $\mathfrak{D}$. Using the definition of the Borel field and Caratheodory's Theorem, we obtain that $\lambda=\sigma_*\lambda$. Let $\pi:\mathcal{R}^{\mathcal{N}}\rightarrow\mathcal{R}$ be the projection onto the $0'th$ coordinate. Then $g=\pi\circ\tau$, and, so $\pi\in L^1(\mathcal{R}^{\mathcal{N}},\mathfrak{D},\lambda)$ by the change of variables formula,

[3] As $\{V\in\mathfrak{D}:\tau^{-1}(V)\in\mathfrak{C}\}$ is a σ-algebra containing the cylinder sets.

([4]). Moreover, $g(T^i\omega) = \pi(\sigma^i\tau(\omega))$, so applying the Ergodic Theorem for $(\mathcal{R}^{\mathcal{N}}, \mathfrak{D}, \lambda, \sigma)$, with the change of variables formula, we have that $\diamond g$ exists and $\diamond g = \diamond\pi \circ \tau$ a.e $d\mu$, and $\int_\Omega \diamond g d\mu = \int_\Omega (\diamond\pi \circ \tau) d\mu = \int_{\mathcal{R}^{\mathcal{N}}} \diamond\pi d\lambda = \int_{\mathcal{R}^{\mathcal{N}}} \pi d\lambda = \int_\Omega g d\mu$ as required.

□

We make the following definition;

Definition 1.5. *We say that* $(\mathcal{R}^{\mathcal{N}}, \mathfrak{D}, \lambda, \sigma) \in \mathcal{P}$ *is a factor of* $(K, \mathfrak{B}, P, \phi)$ *if there exists;*

$\Gamma : (K, \mathfrak{B}, P) \rightarrow (\mathcal{R}^{\mathcal{N}}, \mathfrak{D}, \lambda)$

which is measurable and measure preserving, such that;

$\Gamma(\phi x) = \sigma(\Gamma x)$ *a.e* $(x \in K)$ dP.

We make the same definition if $([0,1]^{\mathcal{N}}, \mathfrak{E}, \rho, \sigma) \in \mathcal{Q}$.

Lemma 1.6. *Suppose that* $(\mathcal{R}^{\mathcal{N}}, \mathfrak{D}, \lambda, \sigma) \in \mathcal{P}$ *is a factor of* $(K, \mathfrak{B}, P, \phi)$*, then, if the Ergodic Theorem holds for* $(K, \mathfrak{B}, P, \phi)$*, it holds for* $(\mathcal{R}^{\mathcal{N}}, \mathfrak{D}, \lambda, \sigma)$.

Proof. The proof is similar to Lemma 1.4. If $h \in L^1(\mathcal{R}^{\mathcal{N}}, \mathfrak{D}, \lambda)$, then, by change of variables, $\Gamma^*h \in L^1(K, \mathfrak{B}, P)$. Applying the Ergodic Theorem for $(K, \mathfrak{B}, P, \phi)$ and the definition of a factor, we have that $\diamond\Gamma^*h$ exists and $\diamond\Gamma^*h = \Gamma^* \diamond h$, a.e d$P$, $(*)$. So $\diamond h$ exists a.e dλ, and, again, by change of variables, $(*)$, and the Ergodic theorem for $(K, \mathfrak{B}, P, \phi)$;

$$\int_{\mathcal{R}^{\mathcal{N}}} \diamond h d\lambda = \int_K \Gamma^*(\diamond h) dP = \int_K \diamond(\Gamma^* h) dP = \int_K (\Gamma^* h) dP = \int_{\mathcal{R}^{\mathcal{N}}} h d\lambda$$

□

We now claim the following;

Lemma 1.7. *Every space in* $\mathcal{P}$ *is isomorphic, in the sense of dynamical systems, ([5]), to a space in* $\mathcal{Q}$.

[4] This states that if $\tau : (X_1, \mathfrak{C}_1, \mu_1) \rightarrow (X_2, \mathfrak{C}_2, \mu_2)$ is measurable and measure preserving, so $\mu_2 = \tau_*\mu_1$, then a function $\theta \in L^1(X_2, \mathfrak{C}_2, \mu_2)$ iff $\tau^*\theta \in L^1(X_1, \mathfrak{C}_1, \mu_1)$ and $\int_C \theta d\tau_*\mu_1 = \int_{\tau^{-1}(C)} \tau^*\theta d\mu_1$.

[5] By which I mean there exists measurable and measure preserving maps $r : (\mathcal{R}^{\mathcal{N}}, \mathfrak{D}, \lambda) \rightarrow ([0,1]^{\mathcal{N}}, \mathfrak{E}, \rho)$ and $s : ([0,1]^{\mathcal{N}}, \mathfrak{E}, \rho) \rightarrow (\mathcal{R}^{\mathcal{N}}, \mathfrak{D}, \lambda)$ such that $s \circ r = Id$ and $r \circ \sigma = \sigma \circ r$ a.e dλ, $r \circ s = Id$ and $s \circ \sigma = \sigma \circ s$ a.e dρ

Proof. There exists an isomorphism, in the sense of measure spaces, $\Phi : (\mathcal{R}^{\mathcal{N}}, \mathfrak{D}, \lambda) \rightarrow ([0,1], \mathfrak{E}', \rho')$, where $\mathfrak{E}'$ is the Borel field and ρ' is a probability measure, see [3], Theorem 1.4.4. Now define $r : \mathcal{R}^{\mathcal{N}} \rightarrow [0,1]^{\mathcal{N}}$ by $r(\omega)(n) = \Phi(\sigma^n \omega)$. Again, using the argument above and the fact that Φ and σ are measurable, $r^{-1}(\mathfrak{E}) \subset \mathfrak{D}$, where is the Borel field on $[0,1]^{\mathcal{N}}$. Let ρ be the probability measure $r_* \lambda$, so $r : (\mathcal{R}^{\mathcal{N}}, \mathfrak{D}, \lambda) \rightarrow ([0,1]^{\mathcal{N}}, \mathfrak{E}, \rho)$ is also measure preserving. We have that $r(\sigma\omega)(n) = \Phi(\sigma^{n+1}\omega) = (r\omega)(n+1) = \sigma(r\omega)(n)$, so $r \circ \sigma = \sigma \circ r$, for all $\omega \in \mathcal{R}^{\mathcal{N}}$. This also shows that ρ is σ invariant, as λ is σ invariant. Hence, $([0,1]^{\mathcal{N}}, \mathfrak{E}, \rho, \sigma)$ belongs to $\mathcal{Q}$. Define $s : ([0,1]^{\mathcal{N}}, \mathfrak{E}, \rho) \rightarrow (\mathcal{R}^{\mathcal{N}}, \mathfrak{D}, \lambda)$, by, $s(\omega') = \Phi^{-1}(\pi(\omega'))$, where again π is the 0'th coordinate projection, clearly s is measurable. Then $(s \circ r)(\omega) = \Phi^{-1} \circ \pi \circ r(\omega)$, and $\pi \circ r(\omega) = r(\omega)(0) = \Phi(\omega)$, so $(s \circ r) = Id$ a.e, and, similarly $r \circ \sigma = \sigma \circ r$ a.e dλ. This clearly shows that s is measure preserving, and that $(r \circ s) = Id$, $s \circ \sigma = \sigma \circ s$,$(*)$, hold, restricted to $r(U)$, where $\lambda(U) = 1$. As, by definition, $\rho(\lambda(U)) = 1$, and the conditions in $(*)$ are measurable, we obtain the result. (Note that the map s need not be invertible in the ordinary sense.) □

We now make the following;

Definition 1.8. *Let* $([0,1]^{\mathcal{N}}, \mathfrak{E}, \rho, \sigma)$ *belong to* $\mathcal{Q}$*, then we say that* α *is typical for* ρ *if;*

$lim_{n\rightarrow\infty} \sum_{i=0}^{n-1} g(\sigma^i \alpha) = \int_{[0,1]^{\mathcal{N}}} g d\rho$

for any $g \in C([0,1]^{\mathcal{N}})$.

We now show;

Theorem 1.9. *Let* $([0,1]^{\mathcal{N}}, \mathfrak{E}, \rho, \sigma)$ *belong to* $\mathcal{Q}$*, possessing a typical element* α*. Then* $([0,1]^{\mathcal{N}}, \mathfrak{E}, \rho, \sigma)$ *is a factor of* $(K, \mathfrak{B}, P, \phi)$ *in the sense of Definition 1.5.*

Proof. Define $\Gamma : K \rightarrow [0,1]^{\mathcal{N}}$ by $\Gamma(x) = {}^{\circ}(\sigma^x\alpha)$, ([6]). Now suppose that $g \in C([0,1]^{\mathcal{N}})$, so, as $[0,1]^{\mathcal{N}}$ is compact, g is bounded,$(*)$, then;

${}^{\circ}g(\sigma^x\alpha) = g(\Gamma(x))$ for all $x \in K$, $(**)$ ([7]).

This implies that Γ is measurable, as if B is an open set for the product topology on $[0,1]^{\mathcal{N}}$, then, taking g to be a continuous function with support B, $\Gamma^* g$ is measurable with respect to P, by Theorem 3.8 (Lemma 3.15) of [4]. This clearly implies that $\Gamma^{-1}(B)$ is measurable. By previous arguments, we obtain the result. Moreover;

$\int_{[0,1]^{\mathcal{N}}} g d\rho$

$= lim_{n\rightarrow\infty}\frac{1}{n}\sum_{i=0}^{n-1} g(\sigma^i\alpha)$, (by definition of a typical element α)

$=^{\circ} (\frac{1}{k}{}^{*}\sum_{x=0}^{k-1} g(\sigma^x\alpha))$, ([8]).

$= {}^{\circ} \int_K g(\sigma^x\alpha)d\nu$ (using Definition 3.9 of [4] and Remarks 3.10 of [4])

$= \int_K g(\Gamma(x))dP$, (using $(*)$, $(**)$ and Theorem 3.12 of [4] (Lemma 3.15 of [4]))

$(***)$

The result of $(***)$ implies that Γ is measure preserving. The probability measure $\Gamma_* P$ defines a bounded linear functional on $C([0,1]^{\mathcal{N}})$, which agrees with ρ. Using the fact that $[0,1]^{\mathcal{N}}$ is a compact Hausdorff

---

[6]Here, $(\sigma^x\alpha) = {}^*H(x)$ for the internal function ${}^*H : {}^*\mathcal{N} \rightarrow {}^*([0,1]^{\mathcal{N}}) = ({}^*[0,1])^{{}^*\mathcal{N}}$, obtained by transferring the standard function $H : \mathcal{N} \rightarrow [0,1]^{\mathcal{N}}$, defined by $H(n) = \sigma^n(\alpha)$. Observe that $[0,1]^{\mathcal{N}}$ is compact and Haussdorff in the product topology, so, by Theorem 2.34 of [4], there exists a unique standard part mapping ${}^{\circ} : {}^*([0,1]^{\mathcal{N}}) \rightarrow [0,1]^{\mathcal{N}}$. In fact, see [6], this mapping is defined by setting ${}^{\circ}s = ({}^{\circ}s(n))_{n\in\mathcal{N}}$ where $s : {}^*\mathcal{N} \rightarrow^* [0,1]$ is internal.

[7]I have also denoted by g, the transfer of g to ${}^*C({}^*([0,1]^{\mathcal{N}}))$. Observe that $\sigma^x(\alpha) \simeq \Gamma(x)$ by definition of Γ, it is then straightforward to adapt Theorem 2.25 of [4], using the fact that g is continuous, to show that $g(\sigma^x\alpha) \simeq g(\Gamma(x))$.

[8]Observe that $s(n) = \frac{1}{n}\sum_{i=0}^{n-1} g(\sigma^i\alpha)$ is a standard sequence, with limit $s = \int_{[0,1]^{\mathcal{N}}} g d\rho$. By Theorem 2.22 of [4], using the fact that k is infinite, $s \simeq s(k)$. Using Definition 2.19 of [4], it is clear that $s(k)$ is the hyperfinite sum $\frac{1}{k}{}^*\sum_{x=0}^{k-1} g(\sigma^x\alpha)$

space, and $\rho, \Gamma_* P$ are regular, see [7] Theorem 2.18, ([9]), we can apply the uniqueness part of the Riesz Representation Theorem, see [7] Theorem 6.19, to conclude that $\Gamma_* P = \rho$, we will discuss this further below. Now, as σ is continuous with respect to $\mathfrak{E}$, ([10]),;

$$\sigma(\Gamma x) = \sigma(^\circ(\sigma^x \alpha) =^\circ (\sigma(\sigma^x \alpha)) =^\circ (\sigma^{x+1}\alpha) = \Gamma(x+1) = \Gamma(\phi(x))$$

except for $x = k-1$, so a.e dP. Hence, the result follows.

□

We now address the problem of finding a typical element for a space $([0,1]^{\mathcal{N}}, \mathfrak{E}, \rho, \sigma) \in \mathcal{Q}$. By Theorem 1.3, Lemma 1.4, Lemma 1.6, Lemma 1.7 and Theorem 1.9, we then obtain the Ergodic Theorem 1.1. The proof of this result does *not* require the Ergodic Theorem, and is originally due to de Ville, see [2].

Definition 1.10. *We say that a sequence of measures $(\rho_n)_{n\in\mathcal{N}}$ converges weakly to ρ if, for all $g \in C([0,1]^{\mathcal{N}})$;*

$$lim_{n\rightarrow\infty}\left(\int_{[0,1]^{\mathcal{N}}} g d\rho_n\right) = \int_{[0,1]^{\mathcal{N}}} g d\rho.$$

We require the following lemma;

Lemma 1.11. *Let $(\alpha_n)_{n\in\mathcal{N}}$ be a sequence of periodic, with respect to σ, elements in $[0,1]^{\mathcal{N}}$, such that the sequence of probability measures $(\rho_{\alpha_n})_{n\in\mathcal{N}}$ converges weakly to ρ, where;*

$$\rho_{\alpha_n} = \frac{1}{c_n}(\delta_{\alpha_n} + \delta_{\sigma\alpha_n} + \ldots + \delta_{\sigma^{c_n-1}\alpha_n})$$

δ_{α_n} denotes the probability measure supported on α_n and c_n denotes the period of α_n. Then there exists a sequence $(r_n)_{n\in\mathcal{N}}$ of positive integers, such that if $(T_n)_{n\in\mathcal{N}}$ is defined by $T_0 = 0$ and $T_{n+1} - T_n = c_n r_n$, the element $\alpha \in [0,1]^{\mathcal{N}}$, defined by $\alpha(m) = \alpha_n(m - T_n)$, for $T_n \leq m < T_{n+1}$, is typical for ρ.

---

[9] It is easy to see that $[0,1]^{\mathcal{N}}$ is σ-compact. This follows from the fact that finite intersections of cylinder sets form a basis for the topology on $[0,1]^{\mathcal{N}}$. Any open set in U in $[0,1]^m$ is a countable union of closed sets, as every $x \in U$ lies inside a closed box B with rational corners, such that $B \subset U$. Hence, any cylinder set is a countable union of such closed sets $\pi_m^{-1}(B)$.

[10] Again I have denoted by σ the transfer of the standard shift σ to $^*([0,1]^{\mathcal{N}})$. The fact that $\sigma(\sigma^x \alpha) = \sigma^{x+1}(\alpha)$ follows immediately by transferring the standard fact that $\sigma(\sigma^n(\alpha)) = \sigma^{n+1}(\alpha)$ for $n \in \mathcal{N}$.

Proof. The proof is intuitively clear, but hard to write down rigorously. As ρ_{α_n} converges weakly to ρ, we have that;

$lim_{n\to\infty}(\int_X f d\rho_{\alpha_n}) = \int_X f d\rho$

By definition of ρ_{α_n};

$\int_X f d\rho_{\alpha_n} = \frac{1}{c_n}(f(\alpha_n) + \ldots + f(\sigma^{c_n-1}\alpha_n)$

So it is sufficient to prove that;

$lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1} f(\sigma^i\alpha) = lim_{n\to\infty}\frac{1}{c_n}(f(\alpha_n) + \ldots + f(\sigma^{c_n-1}\alpha_n))$ $(*)$

We first claim that, if $f \in C([0,1]^{\mathcal{N}})$, there exists an increasing sequence $\{m_n\}_{n\in\mathcal{N}}$ of positive integers, such that if $b, c \in [0,1]^{\mathcal{N}}$, and agree up to the m_n'th coordinate, then $|f(b) - f(c)| < \frac{1}{n}$, $(**)$. In order to see this, for $x \in [0,1]^{\mathcal{N}}$, let $U_x = \{y : |f(x) - f(y)| < \frac{1}{2n}\}$. As f is continuous, U_x is open in the Borel field, hence there exists $V_x \subset U_x$, containing x, of the form $\pi^{-1}(W_x)$, where $W_x \subset \mathcal{R}^{n_x}$ is open, and π is the projection onto the first n_x coordinates. Then, if $y, z \in U_x$, $|f(y) - f(z)| \leq |f(y) - f(x)| + |f(z) - f(x)| < \frac{1}{n}$. The sets $\{V_x : x \in X\}$ form an open cover of $[0,1]^{\mathcal{N}}$, which is compact in the product topology. Hence, there exists a finite subcover $V_{x_1} \cup \ldots \cup V_{x_r}$. We can choose m_n such that each V_{x_j} is of the form $\pi^{-1}(W_{x_j})$, for $W_{x_j} \subset \mathcal{R}^{m_n}$. Then, if b and c agree up to the m_n'th coordinate, we have that $b \in V_{x_j}$ iff $c \in V_{x_j}$, so $|f(b) - f(c)| < \frac{1}{n}$, showing $(**)$. Now let $\{g_n\}_{n\in\mathcal{N}}$ be any increasing sequence of positive integers, such that if $Q_n = sup\{|f(b) - f(c)| : \pi_{g_n}(b) = \pi_{g_n}(c)\}$, then $\{Q_n\}_{n\in\mathcal{N}}$ is decreasing and $lim_{n\to\infty}Q_n = 0$. Clearly such a sequence exists by $(**)$. Without loss of generality, we can choose $\{g_n\}_{n\in\mathcal{N}}$, such that the periods $c_n|g_n$, $(\sharp)$. Now choose $\{T_i\}_{i\in\mathcal{N}}$ as follows;

(i). $T_{i+1} \geq 2^i T_i$

(ii). $g_i|T_{i+1} - T_i$ (so $c_i|T_{i+1} - T_i$)

(iii). $C_i = \frac{T_{i+1}-T_i}{g_i} \gneq C_{i-1} = \frac{T_i - T_{i-1}}{g_{i-1}}$ $(i \geq 1)$.

(iv). $T_i \geq 2^i c_i$ $(i \geq 1)$.

We now claim there exists a decreasing sequence $\{b_n\}_{n\in\mathcal{N}_{>0}}$ of positive reals, such that;

$|\frac{1}{T_n}\sum_{i=0}^{T_n-1} f(\sigma^i\alpha)-t_n|\leq b_n$ $(***)$

where $lim_{n\rightarrow\infty}b_n=0$, and $t_n=\frac{1}{c_n}(f(\alpha_n)+\ldots+f(\sigma^{c_n-1}\alpha_n)$, for $n\geq 1$. For ease of notation, we let;

$A_n=\frac{1}{n}\sum_{i=0}^{n-1} f(\sigma^i\alpha)$

$A_{m,n}=\frac{1}{n-m}\sum_{i=m}^{n-1} f(\sigma^i\alpha)$

Recall the law of weighted averages, $A_n=\frac{mA_m+(n-m)A_{m,n}}{n}$. We first estimate $|A_{T_n}-A_{T_{n-1},T_n}|$. We have;

$A_{T_n}=\frac{T_{n-1}A_{T_{n-1}}+(T_n-T_{n-1})A_{T_{n-1},T_n}}{T_n}$

$|A_{T_n}-A_{T_{n-1},T_n}|$

$=|\frac{T_{n-1}}{T_n}A_{T_{n-1}}+\frac{T_n-T_{n-1}}{T_n}A_{T_{n-1},T_n}-A_{T_{n-1},T_n}|$

$\leq\frac{|A_{T_{n-1}}|}{2^{n-1}}+\frac{|A_{T_{n-1},T_n}|}{2^{n-1}}$ by (i)

$\leq\frac{M}{2^{n-2}}$, where $|f|\leq M$, (A)

We now estimate the average A_{T_{n-1},T_n}. The idea is to divide the interval between T_{n-1} and T_n into C_{n-1} blocks of length g_{n-1}, where the period $c_{n-1}|g_{n-1}$, using $(\sharp)$ and (ii). We estimate $|A_{T_{n-1},T_n}-A_{T_{n-1},T_n-g_n}|$;

$A_{T_{n-1},T_n}=\frac{C_{n-1}-1}{C_{n-1}}A_{T_{n-1},T_n-g_{n-1}}+\frac{1}{C_{n-1}}A_{T_n-g_{n-1},T_n}$

$|A_{T_{n-1},T_n}-A_{T_{n-1},T_n-g_{n-1}}|$

$=|\frac{A_{T_n-g_{n-1},T_n}}{C_{n-1}}-\frac{A_{T_{n-1},T_n-g_{n-1}}}{C_{n-1}}|\leq\frac{2M}{C_{n-1}}$ (B)

We now let;

$B_{T_{n-1},m}=\frac{1}{m-T_{n-1}}\sum_{i=0}^{m-T_{n-1}-1} f(\sigma^i\alpha_{n-1})$, for $m\leq n$.

We estimate $|A_{T_{n-1},T_n-g_{n-1}}-B_{T_{n-1},T_n-g_n}|$. We have that $\sigma^{T_{n-1}+i}\alpha$ and $\sigma^i\alpha_{n-1}$ agree up to the g_{n-1}'th coordinate, for $0\leq i<T_n-T_{n-1}-g_{n-1}$.

Therefore, for such i, $|f(\sigma^i\alpha_{n-1})-f(\sigma^{T_{n-1}+i}\alpha)|\leq Q_{n-1}$, and so;

$|A_{T_{n-1},T_n-g_{n-1}}-B_{T_{n-1},T_{n-1}-g_{n-1}}|\leq Q_{n-1}$ (C)

Now, by the same argument as in (B);

$|B_{T_{n-1},T_n}-B_{T_{n-1},T_n-g_n}|\leq\frac{2M}{C_{n-1}}$ (D)

Finally, by periodicity;

$B_{T_{n-1},T_n}=\frac{1}{c_{n-1}}(f(\alpha_{n-1})+\ldots+f(\sigma^{c_{n-1}-1}\alpha_{n-1}))=t_n$ (E)

Now, combining the estimates $(A),(B),(C),(D),(E)$, we have;

$|A_{T_n}-t_n|\leq\frac{M}{2^{n-2}}+\frac{2M}{2^{n-2}}+Q_{n-1}+\frac{2M}{C_{n-1}}=b_n$

Clearly $\{b_n\}_{n\in\mathcal{N}}$ is decreasing. Moreover, $lim_{n\to\infty}b_n=0$, as $lim_{n\to\infty}C_n=\infty$, (iii), and by the choice of $\{Q_n\}_{n\in\mathcal{N}}$. This shows $(***)$. We now have to estimate the averages up to place between the critical points T_n and T_{n+1}.

Case 1. The place v is a periodic point of the form;

T_n+mg_n, where $0\leq m\leq C_n-1$

We have $A_v=\lambda A_{T_n}+(1-\lambda)A_{T_n,v}$ $(0\leq\lambda\leq 1)$, where $|A_{T_n,v}-t_{n+1}|\leq Q_n$, by $(C),(E)$, and $|A_{T_n}-t_n|\leq b_n$, by $(***)$. Now, let $t=lim_{n\to\infty}t_n$. Given $\epsilon>0$, choose $N(\epsilon)$, such that $|t_n-t|<\epsilon$, for all $n\geq N(\epsilon)$. Then;

$|A_v-t|\leq max\{|A_{T_n}-t|,|A_{T_n,v}-t|\}$

$\leq max\{b_n+\frac{\epsilon}{2},Q_n+\frac{\epsilon}{2}\}$

Choose $N_1(\epsilon)\geq N(\epsilon)$, such that $max\{b_n,Q_n\}<\frac{\epsilon}{2}$, for all $n\geq N_1(\epsilon)$, then $|A_v-t|<\epsilon$, for all $n\geq N_1(\epsilon)$.

Case 2. The place v is a possibly non-periodic point of the form;

T_n+w, where $0\leq w\leq T_{n+1}-T_{n-1}-g_n$.

Choose periodic points v_1 and v_2, with $T_n \leq v_1 \leq v \leq v_2 \leq T_{n+1} - g_n$, and $v_2 - v_1 = c_n$, so $0 \leq v - v_1 = e \leq c_n$. Then $A_v = \frac{v_1}{v_1+e} A_{v_1} + \frac{e}{v_1+e} A_{v_1,v}$. As $v_1 \geq T_n$, we have;

$\frac{e}{v_1+e} \leq \frac{e}{T_n+e} \leq \frac{c_n}{T_n} \leq \frac{1}{2^n}$ by (iv).

Therefore;

$|A_v - A_{v_1}| = |(1-\delta)A_{v_1} + \delta A_{v_1,v} - A_{v_1}|$, $(\delta \leq \frac{1}{2^n})$

$\leq \delta(|A_{v_1}| + |A_{v_1,v}|) \leq \frac{M}{2^{n-1}}$

For $n \geq N_1(\frac{\epsilon}{2})$, $|A_{v_1} - t| < \frac{\epsilon}{2}$, by Case 1, so $|A_v - t| < \epsilon$, for $n \geq N_2(\epsilon)$, where $N_2(\epsilon) = max\{N_1(\frac{\epsilon}{2}), log(\frac{2M}{\epsilon}) + 2\}$.

Case 3. The place v is of the form;

$T_n + w$, where $T_{n+1} - T_n - g_n \leq w \leq T_{n+1} - T_n$.

We have;

$A_v = \lambda A_{T_n} + (1-\lambda) A_{T_n,v}$, $(0 \leq \lambda \leq 1)$, $(\dagger)$,

$A_{T_n,T_{n+1}} = \mu A_{T_n,v} + (1-\mu) A_{v,T_{n+1}}$, $\frac{C_n - 1}{C_n} \leq \mu \leq 1$

Therefore;

$|A_{T_n,T_{n+1}} - A_{T_n,v}| \leq \frac{2M}{C_n}$

$|A_{T_n,T_{n+1}} - t_{n+1}| \leq b_{n+1}$, by $(B), (C), (D), (E)$

$|A_{T_n,v} - t_{n+1}| \leq \frac{2M}{C_n} + b_{n+1}$

$|A_{T_n} - t_n| \leq b_n$ by $(***)$

$|A_v - t| \leq max\{|A_{T_n} - t|, |A_{T_n,v} - t|\}$ by $(\dagger)$

$\leq max\{b_n + |t_n - t|, \frac{2M}{C_n} + b_{n+1} + |t_{n+1} - t|\}$, $(\dagger\dagger)$

We have, for $n \geq N(\frac{\epsilon}{2})$, $max\{|t_n - t|, |t_{n+1} - t|\} < \frac{\epsilon}{2}$. Choose $N_3(\epsilon)$, such that $max\{b_n, \frac{2M}{C_n} + b_{n+1}\} < \frac{\epsilon}{2}$, for all $n \geq N_3(\epsilon)$. Then, for

$n \geq N_3(\epsilon)$, $|A_v - t| < \epsilon$.

To complete the proof, let $N_4(\epsilon) = max\{N_1(\epsilon), N_2(\epsilon), N_3(\epsilon)\}$. Then, for $n \geq N_4(\epsilon)$, $|A_m - t| < \epsilon$, for all $m \geq T_n$, by Cases 1,2 and 3. Therefore;

$lim_{m\rightarrow\infty}\frac{1}{m}\sum_{i=0}^{m-1} f(\sigma^i\alpha) = \int_X f d\rho$

so α is typical, as required.

□

We now formulate the following criteria.

Lemma 1.12. *Suppose that for every $g \in C([0,1]^{\mathcal{N}})$, and $\epsilon > 0$, there exists a periodic element $\beta \in [0,1]^{\mathcal{N}}$, with;*

$|\int_{[0,1]^{\mathcal{N}}} g d\rho_\beta - \int_{[0,1]^{\mathcal{N}}} g d\rho| < \epsilon$

then there exists a sequence of periodic elements $(\alpha_n)_{n\in\mathcal{N}}$, with $(\rho_{\alpha_n})_{n\in\mathcal{N}}$ converging weakly to ρ.

Proof. We abbreviate $[0,1]^{\mathcal{N}}$ to X. Let $\mathcal{M}$ denote the vector space of real valued regular measures on $(X, \mathfrak{C})$. As we observed every probability measure belongs to $\mathcal{M}$. $\mathcal{M}$ is a Banach space, with norm defined by total variation, see [7]. Using the Riesz Representation Theorem, $\mathcal{M}$ can be identified with the dual space $C(X)^*$. It is easy to see that then $\mathcal{M} \cong C(X)^*$, as Banach spaces, however, we will not require this fact. The weak $*$-topology, see [1], on $\mathcal{M}$, is the coursest topology for which all the elements $\hat{g} \in C(X)^{**}$, where $g \in C(X)$, are continuous. Formally, we define a set $U \subset \mathcal{M}$ to be open if for all $\rho \in U$, there exist $\{g_1, \ldots, g_n\} \subset C(X)$, and positive reals $\{\epsilon_1, \ldots, \epsilon_n\}$ such that;

$\{\rho' \in \mathcal{M} : |\rho'(g_i) - \rho(g_i)| < \epsilon_i\} \subset U$

Fixing ρ, let Ω_ρ denote the open sets containing ρ. We show that Ω_ρ has a countable base, $(*)$. Using the compactness argument, given in Lemma 1.11, and the Stone-Weierstrass Theorem, see [1], it is easy to show that the space V of pullbacks of polynomial functions on $[0,1]^n$, for some n, is dense in $C(X)$. Clearly V has a countable basis, which shows that $C(X)$ is separable, that is, contains a countable dense subset Y. Now suppose that $g \in C(X)$, $\epsilon > 0$. Let $U_{g,\epsilon} = \{\rho' : |\rho'(g) - \rho(g)| < \epsilon\}$, and $D \in \mathcal{Q}$. Choose $\delta \in \mathcal{Q}$ with

$\delta < \frac{\epsilon}{2(D+2|\rho(X)|)}$, and $\gamma \in \mathcal{Q}$ with $\gamma < \frac{\epsilon}{2}$. Choose $h \in Y$ with $||g-h||_{C(X)} < \delta$. Then $U_{h,\gamma} \cap U_{1,D} \subset U_{g,\epsilon}$, $(**)$, as if $|\rho'(h)-\rho(h)| < \gamma$, then;

$|\rho'(g)-\rho(g)| = |\rho'(g-h)+\rho'(h)-\rho(g-h)-\rho(h)| \leq \delta(|\rho'(X)| + |\rho(X)|) + \gamma$

and, if $|\rho'(1)-\rho(1)| < D$, then $|\rho'(X)| + |\rho(X)| < D + 2|\rho(X)|$, so $|\rho'(g)-\rho(g)| < \epsilon$. This clearly shows $(**)$. As sets of the form $U_{h,q} \in \Omega_\rho$, for $h \in Y$, and $q \in \mathcal{Q}$, are countable, we clearly have $(*)$. Let $I : \mathcal{N} \rightarrow \Omega_\rho$ be an enumeration of the sets $U_{h,q}$, and let $J : \mathcal{N} \rightarrow \Omega_\rho$ define the intersection of the first n elements in I. If the assumption in the lemma is satisfied, we can define a sequence of probability measures $(\rho_{\alpha_n})_{n\in\mathcal{N}}$, by taking ρ_{α_n} to lie inside the open set $J(n)$. Then clearly such a sequence converges to ρ in the weak $*$-topology, hence, for any $g \in C(X)$, as g is continuous for this topology $lim_{n\rightarrow\infty}\rho_{\alpha_n}(g) = \rho(g)$. Therefore, the sequence $(\rho_{\alpha_n})_{n\in\mathcal{N}}$ converges weakly to ρ. □

We refine this criteria further;

Definition 1.13. *Given a positive integer m, we define the partition E_m of $[0,1]$ to consist of the sets;*

$E_{j,m} = [\frac{j}{m}, \frac{j+1}{m})$ *for j an integer between 0 and $m-2$*

$E_{m-1,m} = [\frac{m-1}{m}, 1]$

Given positive integers m,n, we define the partition $B_{m,n}$ of $[0,1]^n$ to consist of the sets;

$B_{\bar{j},m,n} = E_{j_0,m} \times E_{j_1,m} \times \ldots \times E_{j_{n-1},m}$

where $\bar{j} = (j_0, j_1, \ldots, j_{n-1})$ and $\{j_0, \ldots, j_{n-1}\}$ are integers between 0 and $m-1$.

We define the partition $C_{m,n}$ of $[0,1]^{\mathcal{N}}$ to consist of the sets;

$C_{\bar{j},m,n} = \pi_n^{-1}(B_{\bar{j},m,n})$

where π_n is the projection onto the first n coordinates.

Lemma 1.14. *Let $\epsilon > 0$, $g \in C(X)$ be given as in Lemma 1.12, and let ρ' be a regular Borel measure, then there exist positive integers m, n, and $\delta > 0$, such that, if;*

$$|\rho'(C_{\bar{j},m,n}) - \rho(C_{\bar{j},m,n})| < \delta$$

for all sets $C_{\bar{j},m,n}$ belonging to $C_{m,n}$, then;

$$|\int_{[0,1]^{\mathcal{N}}} g d\rho' - \int_{[0,1]^{\mathcal{N}}} g d\rho| < \epsilon$$

Proof. For a positive integer n, let W_n consist of the inverse images in X (from the projection π_n) of open boxes in $[0,1]^n$, with rational corners. Let $W = \bigcup_{n \in \mathcal{N}} W_n$. It is clear that W forms a countable basis for the topology on $[0,1]^{\mathcal{N}}$. Adapting the compactness argument, given above in Lemma 1.11, for any $\gamma > 0$ and $g \in C(X)$, we can find a positive integer n, and finitely many sets $\{W_{1,n}, \ldots, W_{r,n}\}$ in W_n, covering X, such that $|g(x) - g(y)| < \gamma$ for all x, y in $W_{j,n}$, $1 \leq j \leq r$. Now choose m such that each set of the partition $C_{m,n}$ lies inside one of the $W_{j,n}$. Then $|g(x) - g(y)| < \gamma$ on each $C_{\bar{j},m,n}$, belonging to $C_{m,n}$. Now, for given $\delta > 0$, suppose we choose ρ' such that $|\rho'(C_{\bar{j},m,n}) - \rho(C_{\bar{j},m,n})| < \delta$, $(*)$. Then;

$$|\int_X g d\rho' - \int_X g d\rho| = |\sum_{\bar{j}} \int_{C_{\bar{j},m,n}} g d\rho' - \sum_{\bar{j}} \int_{C_{\bar{j},m,n}} g d\rho|$$

$$\leq \sum_{\bar{j}} |\int_{C_{\bar{j},m,n}} g d\rho' - \int_{C_{\bar{j},m,n}} g d\rho|, \ (**)$$

Without loss of generality, assuming ρ' is positive, by definition of the integral, see [7], we have that;

$$c_{\bar{j}} \rho'(C_{\bar{j},m,n}) \leq \int_{C_{\bar{j},m,n}} g d\rho' \leq d_{\bar{j}} \rho'(C_{\bar{j},m,n})$$

$$c_{\bar{j}} \rho(C_{\bar{j},m,n}) \leq \int_{C_{\bar{j},m,n}} g d\rho \leq d_{\bar{j}} \rho(C_{\bar{j},m,n})$$

where $c_{\bar{j}} = inf_{C_{\bar{j},m,n}} g$ and $d_{\bar{j}} = sup_{C_{\bar{j},m,n}} g$. Then;

$$c_{\bar{j}} \rho'(C_{\bar{j},m,n}) - d_{\bar{j}} \rho(C_{\bar{j},m,n}) \leq \int_{C_{\bar{j},m,n}} g d\rho' - \int_{C_{\bar{j},m,n}} g d\rho$$

$$\leq d_{\bar{j}} \rho'(C_{\bar{j},m,n}) - c_{\bar{j}} \rho(C_{\bar{j},m,n})$$

Therefore, again, without loss of generality;

$|\int_{C_{\bar{j},m,n}} gd\rho' - \int_{C_{\bar{j},m,n}} gd\rho|$

$\leq (d_{\bar{j}} - c_{\bar{j}})\rho'(C_{\bar{j},m,n}) + |c_{\bar{j}}||\rho'(C_{\bar{j},m,n}) - \rho(C_{\bar{j},m,n})| \leq \gamma\rho'(C_{\bar{j},m,n}) + |c_{\bar{j}}|\delta$

$(***)$

By $(*)$, $\rho'(X) = \sum_{\bar{j}} \rho'(C_{\bar{j},m,n}) \leq \sum_{\bar{j}} \rho(C_{\bar{j},m,n}) + \delta m^n = 1 + \delta m^n$, so using $(**)$, $(***)$, and the fact that $|g| \leq M$;

$|\int_X gd\rho' - \int_X gd\rho| \leq \gamma(1 + \delta m^n) + \delta M m^n$

So if we choose $0 < \gamma < \frac{\epsilon}{2}$ and $0 < \delta < \frac{\epsilon}{2(\gamma+M)m^n}$, we obtain;

$|\int_X gd\rho' - \int_X gd\rho| < \epsilon$

as required.

□

We finally claim;

Theorem 1.15. *If $C_{m,n}$ is a partition, as in Definition 1.13 and $\delta > 0$, then there exists a periodic element β, such that;*

$|\rho_\beta(C_{\bar{j},m,n}) - \rho(C_{\bar{j},m,n})| < \delta$

for all sets $C_{\bar{j},m,n}$ belonging to $C_{m,n}$.

Proof. Let $\Sigma = \{\frac{1}{2m}, \frac{3}{2m}, \ldots, \frac{2m-1}{2m}\}$. Define $\kappa : \Sigma^n \rightarrow \mathcal{R}$ by;

$\kappa((\frac{2j_0+1}{2m}, \ldots, \frac{2j_{n-1}+1}{2m})) = \rho(C_{\bar{j},m,n})$

As $C_{m,n}$ is a partition of X and ρ is a probability measure, κ is a probability measure on Σ^n. Moreover, using the partition property and the fact that ρ is σ-invariant;

$\sum_{\xi_0 \in \Sigma} \kappa((\xi_0, \ldots, \xi_{n-1})) = \rho(\pi_n^{-1}([0,1] \times E_{j_1,m} \times \ldots \times E_{j_{n-1},m}))$

$= \rho(\pi_n^{-1}(E_{j_1,m} \times \ldots \times E_{j_{n-1},m}))$

$= \rho(\pi_n^{-1}(E_{j_1,m} \times \ldots \times E_{j_{n-1},m} \times [0,1]))$

$= \sum_{\xi_0 \in \Sigma} \kappa((\xi_1, \ldots, \xi_{n-1}, \xi_0))$ $(*)$

Now let $N > 0$ be a sufficiently large positive integer, then we claim that we can find a probability measure κ' on Σ^n such that;

(i). $|\kappa'(\bar{\xi}) - \kappa(\bar{\xi})| < \delta$

(ii). The condition $(*)$ still holds.

(iii). $N\kappa'(\bar{\xi})$ is a non-negative integer, for all $\bar{\xi} \in \Sigma^n$

This follows from a simple linear algebra argument. We can identify the set of real measures on Σ^n with the real vector space V of dimension m^n. The condition $(*)$ then defines a subspace $W \subset V$. The condition of being a probability measure requires that;

$\sum_{\xi_0, \ldots, \xi_{n-1} \in \Sigma^n} \kappa((\xi_1, \ldots, \xi_{n-1}, \xi_0)) = 1$, $(**)$

which defines an affine space $S_{aff} \subset V$. $S_{aff} \cap W$ contains a rational point q, corresponding to the probability measure with coordinates m^{-n}. It is straightforward to see that $(S_{aff} \cap W) = [(S_{aff} - q) \cap W] + q$. Moreover, $(S_{aff} - q) \cap W$ is a vector space defined by rational coefficients, so it has a rational basis. This shows that rational points are dense in $S_{aff} \cap W$. We can, without loss of generality, assume that all the coordinates of κ are strictly greater than zero. If not, consider instead the space $S_{aff} \cap W \cap W'$, where $W' = Ker(\pi)$ is the kernel of the projection onto the non-zero coordinates of κ. The same argument shows that rational points are dense in $S_{aff} \cap W \cap W'$. We can now obtain a probability measure κ', satisfying conditions $(i) - (iii)$, by finding a rational vector sufficiently close to κ in $S_{aff} \cap W$, and choosing N large enough.

Now take a longest sequence $\{\xi^0, \ldots, \xi^{r-1}\}$ of elements in Σ^n, such that;

(1). $(\xi_1^i, \ldots, \xi_{n-1}^i) = (\xi_0^{i+1}, \ldots, \xi_{n-2}^{i+1})$.

(2). $Card(\{i : 0 \leq i < r, \xi^i = \xi\}) \leq N\kappa'(\xi)$ for any $\xi \in \Sigma^n$

where $\xi^i = (\xi_0^i, \ldots, \xi_{n-1}^i)$, for $0 \leq i \leq r$, and $\xi^r = \xi^0$.

Then, by graph theoretical considerations, ([11]), one can show that equality holds in the above inequality in (2), for any $\xi \in \Sigma^n$, $(***)$.

[11] The graph theory argument proceeds as follows. We construct a tree. For every $\xi' \in \Sigma^{n-1}$, where $\xi' = (\xi_1, \ldots, \xi_{n-1})$, associate a vertex $v_{\xi'}$ (the trunk). Similarly, for every $\xi \in \Sigma^n$, where $\xi = (\xi_0, \ldots, \xi_{n-1})$, associate two vertices l_ξ (left) and r_ξ (right). Attach the vertex l_ξ to $v_{\xi'}$ iff $\pi(\xi) = \xi'$, where π is the projection onto the last $n-1$ coordinates, and, attach l_ξ to $v_{\xi'}$. iff $\pi'(\xi) = \xi'$, where π' is the projection onto the first $n-1$ coordinates. In this way, we obtain a tree, having $m^{n-1}(2m+1)$ vertices, $m^{n-1}(2m)$ branches, and m_{n-1} components. Each element $\xi \in \Sigma^n$ corresponds to two vertices, one on the left and one on the right of the tree. Now attach weights $m_\xi = n_\xi$ to the left vertices and right vertices respectively, by assigning the vertices l_ξ and r_ξ, the weights $m_\xi = N\kappa'(\xi)$ and $n_\xi = N\kappa'(\xi)$ respectively. Observe that, by the condition $(*)$ in the main text, for any given ξ';

$$m_{\xi'} = \sum_{\xi \in \Sigma^n : \pi(\xi) = \xi'} m_\xi = n_{\xi'} = \sum_{\xi \in \Sigma^n : \pi'(\xi) = \xi'} n_\xi \ (\dagger)$$

Now, given a sequence $\{\xi^0, \xi^1, \ldots, \xi^k\}$ of elements in Σ^n, where $\xi^i = (\xi^i_0, \ldots, \xi^i_{n-1})$, for $0 \leq i \leq k$, we attach sets L_ξ to each vertex l_ξ, by requiring that, $\xi^i \in L_\xi$ iff $\xi^i = \xi$, and, similarly, we attach sets R_ξ to each vertex r_ξ. We call a sequence allowed if (i). For each $\xi \in \Sigma^n$, $Card(L_\xi) = Card(R_\xi) \leq m_\xi = n_\xi$ and (ii). For each $1 \leq i \leq k$, if ξ^i appears in the set R_ξ, then ξ^{i-1} appears in a set $L_{\xi''}$, where $l_{\xi''}$ and r_ξ are attached to the same vertex $v_{\xi'}$, so that $\pi(\xi'') = \pi'(\xi) = \xi'$. Clearly, all allowed sequences are bounded in length by $N\kappa'(X)$, so there exists a longest allowed sequence $s = (\xi^i)_{0 \leq i \leq t}$. Let ξ^t be the final element in the sequence, and suppose that $\xi^t \in L_{\xi''}$, then, we claim that ξ^0 belongs to a set R_ξ, where $\pi(\xi'') = \pi'(\xi) = \xi'$, $(\dagger\dagger)$. If not, all such sets R_ξ, with $\pi'(\xi) = \pi(\xi'')$, consists of elements ξ^i with $i \geq 1$. If, for one of these sets R_ξ, $Card(R_\xi) \lneq n_\xi$, then we can extend the sequence by setting $\xi^{t+1} = \xi$, clearly such a sequence is allowed, contradicting maximality. So we can assume that $Card(R_\xi) = n_\xi$. By condition (ii), for every element ξ^i, $i \geq 1$, appearing in R_ξ, there exists an element ξ^{i-1} appearing in an $L_{\xi''}$, with $\pi(\xi'') = \pi(\xi^t)$. This provides a total of $w+1$ elements appearing in such $L_{\xi''}$, where $w = \sum_{\xi \in \Sigma^n : \pi'(\xi) = \xi'} n_\xi$. By $(\dagger)$, this is greater than $\sum_{\xi \in \Sigma^n : \pi(\xi) = \xi'} m_\xi$. Clearly, this contradicts condition (i) of an allowed path. Hence, $(\dagger\dagger)$ is shown. Observe also that if $\xi' \in \Sigma^{n-1}$, and $s_{r,\xi'}$ denotes the total number of elements from the sequence s, appearing in sets to the right of ξ', $s_{l,\xi'}$, to the left, then $s_{l,\xi'} = s_{r,\xi'}$, In particular, by $(\dagger)$, $m_{\xi'} - s_{l,\xi'} = n_{\xi'} - s_{r,\xi'} \geq 0$, so the number of "vacant slots" (if there are any), is the same on both sides of a given ξ', $(\dagger\dagger\dagger)$. In order to see this, we can, without loss of generality, assume that $\pi'(\xi^0) \neq \xi'$, then just note that an element ξ^{i+1} belongs to a set on the right of ξ' iff ξ^i belongs to a set on the left of ξ', by condition (ii) of an allowed path. We now claim that for all $\xi \in \Sigma^n$, $Card(R_\xi) = n_\xi$, $(\dagger\dagger\dagger\dagger)$, (so there are no vacant slots). We have already shown this in the particular case when $\pi'(\xi) = \pi'(\xi^0)$. We define an element ξ to be cyclic if $\pi(\xi) = \pi'(\xi)$, so cyclic elements are just constant sequences. We define an element ξ to be free if $Card(R_\xi) \lneq n_\xi$. No free cyclic element ξ_{cyc} can encounter the sequence s, for suppose that there exists a ξ^i, for some $0 \leq i \leq t$, with $\pi(\xi^i) = \pi'(\xi_{cyc})$, then we can extend the sequence s to $s' = \{\xi^0, \ldots, \xi^i, \xi_{cyc}, \xi^{i+1}, \ldots, \xi^t\}$, and still obtain an allowed path, contradicting maximality. So we have that, if ξ is free cyclic, with $\pi_\xi = \xi'$, then $s_{l,\xi'} = s_{r,\xi'} = 0$, $(\dagger\dagger\dagger\dagger\dagger)$. Now suppose there exists a free element ξ_{free}. Choose the largest k, with $0 \leq k \leq t$, such that ξ^k appears

Now let β be the periodic element in $[0,1]^{\mathcal{N}}$, with period $n+r-1$, defined by;

$(\beta(0),\beta(1),\ldots,\beta(n+r-2))=(\xi_0^0,\xi_1^0,\ldots,\xi_{n-1}^0,\xi_{n-1}^1,\xi_{n-1}^2,\ldots,\xi_{n-1}^{r-1})$

By (i), it is sufficient to prove that, for each $\bar{j}\in m^n$;

$|\rho_\beta(C_{\bar{j},m,n})-\kappa'(\xi_{\bar{j}})|<\epsilon$, $(****)$,

where $\epsilon=min_{\bar{i}}(\delta-|\kappa'(\xi_{\bar{i}})-\kappa(\xi_{\bar{i}})|)$, and $\xi_{\bar{j}}$ is the unique element of Σ^n lying inside $C_{\bar{j},m,n}$. By definition of ρ_β, $\rho_\beta(C_{\bar{j},m,n})=\frac{c_{\bar{j}}}{n+r-1}$, where;

$c_{\bar{j}}=Card(\{k:0\leq k<n-r-1,\pi_n(\sigma^k(\beta))=\xi_{\bar{j}}\})$.

By definition of β, and $(***)$, $\frac{c_{\bar{j}}}{n+r-1}=\frac{N\kappa'(\xi_{\bar{j}})+y}{n+r-1}$, where $0\leq y\leq n$. As κ' is a probability measure, again by $(***)$, we have that $r-1=N$. Hence;

$\frac{c_{\bar{j}}}{n+r-1}=\frac{N\kappa'(\xi_{\bar{j}})+y}{N+n}=\kappa'(\xi_{\bar{j}})+\frac{y-n\kappa'(\xi_{\bar{j}})}{N+n}$.

Therefore,

$|\rho_\beta(C_{\bar{j},m,n})-\kappa'(\xi_{\bar{j}})|\leq\frac{n}{N+n}<\epsilon$.

---

in $L_{\xi''}$ with $\pi(\xi'')=\pi'(\xi_{free})$, $(\sharp)$. As we have observed, $k\lneq t$. We construct a forward path from ξ_{free} as follows. Define $\eta^0=\xi_{free}$, add the element η^0 to $R_{\xi_{free}}$ and $L_{\xi_{free}}$, and call the new sets $R_{0,\xi}$ and $L_{0,\xi}$, for $\xi\in\Sigma^n$. Having defined η^j, there are four cases. If $\pi(\eta^j)=\pi'(\eta^0)$, terminate the sequence. Otherwise, if $\pi(\eta^j)=\pi(\xi_{cyc})$ for some cyclic element with $Card(R_{j,\xi_{cyc}})\lneq n_{\xi_{cyc}}$, then define $\eta^{j+1}=\xi_{cyc}$, add the element η^{j+1} to $R_{j,\xi_{cyc}}$ and $L_{i,\xi_{cyc}}$, calling the new sets $R_{j+1,\xi}$ and $L_{j+1,\xi}$, for $\xi\in\Sigma^n$. If there is no such cyclic element, and there exists a free element ξ' with $\pi(\eta^j)=\pi'(\xi')$ and $Card(R_{j,\xi'})\lneq n_{\xi'}$, then define $\eta^{j+1}=\xi'$ (so there is some choice here), and, as before, redefine the sets $R_{j,\xi}$ and $L_{j,\xi}$ to $R_{j+1,\xi}$ and $L_{j+1,\xi}$, for $\xi\in\Sigma^n$. If there is no free element of this form, then terminate the sequence. It is straightforward to see, using $(\dagger\dagger\dagger)$, $(\dagger\dagger\dagger\dagger\dagger)$, and the fact that η^0 is not cyclic, that the sequence $\{\eta^0,\ldots,\eta^j\}$ terminates after a finite number of steps l, with $l>0$, and $\pi(\eta^l)=\pi'(\eta^0)$. Moreover, for all $k<i<t$, and $0\leq j\leq l$, we have that $\pi(\xi^i)\neq\pi'(\eta^j)$, by $(\sharp)$. Hence, we can construct an allowed sequence $s''=\{\xi^0,\ldots,\xi^k,\eta^0,\ldots,\eta^l\xi^{k+1},\ldots,\xi^t\}$, contradicting maximality of s. This shows $(\dagger\dagger\dagger\dagger)$. It is clear that the sequence $s'''=\{\xi^0,\ldots,\xi^{r-1}\}$, as defined in the main text, is a longest allowed sequence, as defined in this footnote, using $(\dagger\dagger)$. Hence, by $(\dagger\dagger\dagger\dagger)$, we have equality in (2) as required.

if we choose N sufficiently large. Hence, $(****)$ and the theorem are shown.

$\square$

We summarise what we have done;

Theorem 1.16. *The Ergodic Theorem 1.1 holds and admits a non-standard proof.*

Proof. Combine Theorems 1.3,1.9,1.15, and Lemmas 1.4,1.6,1.7,1.11,1.12,1.14.

$\square$

Remarks 1.17. *There are some outstanding questions in Ergodic Theory, which one might hope to solve using nonstandard methods, similar to the above. One of these is Ornstein's Isomorphism Theorem, I hope to investigate this direction further.*

2. Appendix

Theorem 2.1. *Suppose $g : X \rightarrow \mathcal{R}$ is integrable with respect to μ_L, $\mu_L(X) < \infty$, and $\epsilon > 0$ is standard, then there exist $F, G : X \rightarrow^* \mathcal{R}$, which are $\mathfrak{A}$-measurable, such that;*

(i). $G \leq g \leq F$.

(ii). $|\int_A g d\mu_L - \int_A G d\nu| < \epsilon$, $|\int_A g d\mu_L - \int_A F d\nu| < \epsilon$

for all $A \in \mathfrak{A}$.

Proof. Consider, first, the case when $g \geq 0$.

Upper Bound. As g is integrable, by Theorem 3.31 of [4], it has an S-integrable lifting F', such that ${}^\circ F' = g$ a.e μ_L, and;

${}^\circ \int_X F' d\nu = \int_X g d\mu_L$

Without loss of generality, we cam assume that $F' \geq 0$. Now let $\epsilon > 0$ be given and choose $\delta > 0$ such that $\mu_L(X)\delta < \frac{\epsilon}{2}$. Then $F' + \delta$ is S-integrable and $F' + \delta \geq f$ a.e μ_L, $(*)$, $F' + \delta > 0$. Moreover;

${}^\circ \int_X (F' + \delta) d\nu = \int_X g d\mu_L + \delta \mu_L(X) < C + \frac{\epsilon}{2}$, $(**)$

where $C = \int_X g d\mu_L$. Let $N \in \mathfrak{M}_L$, with $\mu_L(N) = 0$, such that $(*)$ holds on N^c. Let $N_n = N \cap g^{-1}((n-1, n])$, for $n \in \mathcal{N}_{>0}$, $N_0 = N \cap g^{-1}(0)$. Then $N = \bigcup_{n \geq 0} N_n$, and $\mu_L(N_n) = 0$. By Lemma 3.15 (3.4(i)) of [4], we can choose $U_n \supset N_n$, with $U_n \in \mathfrak{A}$, such that $\mu_L(U_n) < \frac{\epsilon}{4(n+1)^3}$. Inductively, define $F_0 = F' + \delta$, and, having defined F_n, let $F_{n+1} = F_n$ on U_{n+1}^c, and $F_{n+1} = F_n + n + 1$ on U_{n+1}. Then $\{F_n\}$ is an increasing sequence of $\mathfrak{A}$-measurable functions. Moreover;

$\int_X F_{n+1} d\nu$

$= \int_{U_{n+1}^c} F_n d\nu + \int_{U_{n+1}} (F_n + (n+1)) d\nu$

$\simeq \int_X F_n d\nu + (n+1)\mu_L(U_{n+1})$

$< \int_X F_n d\nu + \frac{\epsilon}{4(n+1)^2}$

$\int_X F_n d\nu < C + \frac{\epsilon}{2} + \sum_{m=1}^{n} \frac{\epsilon}{4m^2} < C + \epsilon$ (using $(**)$)

We clearly have that for all $x \in N_n$, $g(x) \leq F_n$. Now, by countable comprehension, we can find an internal sequence $\{F_n\}_{n \in {}^*\mathcal{N}}$ extending the sequence $\{F_n\}_{n \in \mathcal{N}}$. By overflow, there exists an infinite ω, such that $F_n \leq F_\omega$, for all $n \in \mathcal{N}$, $F_\omega > 0$, and;

$\int_X F_\omega d\nu < C + \epsilon$, $(\dagger)$

Clearly $g(x) \leq F_\omega(x)$, for all $x \in X$. Now, if $A \in \mathfrak{A}$, with;

$\int_A F_\omega d\nu - \int_A g d\mu_L > \epsilon$

then, using Theorem 3.16 of [4];

$\int_X F_\omega d\nu$

$= \int_A F_\omega d\nu + \int_{A^c} F_\omega d\nu$

$> \epsilon + \int_A g d\mu_L + \int_{A^c} g d\mu_L = C + \epsilon$

contradicting $(\dagger)$. Setting $F = F_\omega$ gives an upper bound.

Lower Bound. Again choose $\delta > 0$, with $\mu_L(X)\delta < \frac{\epsilon}{2}$. Let F' be as before, then $F' - \delta$ is S-integrable, $F' - \delta \leq g$ a.e μ_L, and:

$\int_X (F' - \delta) d\nu > C - \frac{\epsilon}{2}$

Again choose N, with $\mu_L(N) = 0$, such that $F' - \delta \leq g$ on N^c. Using Lemma 3.15(3.4(i)) of [4] again, we can choose a decreasing sequence of sets $\{U_n\}_{n \in \mathcal{N}_{>0}}$, belonging to $\mathfrak{A}$, with $U_n \supset N$, and $\mu_L(U_n) < \frac{1}{n}$. By S-integrability;

${}^\circ \int_{U_n} (F' - \delta) d\nu = \int_{U_n} {}^\circ(F' - \delta) d\mu_L$

and;

$lim_{n \rightarrow \infty} (\int_{U_n} {}^\circ(F' - \delta) d\mu_L) = 0$

by the DCT, as ${}^\circ(F' - \delta)\chi_{U_n}$ converges to 0 a.e μ_L. Hence, for sufficiently large n, we can assume that;

$\int_{U_n} (F' - \delta) d\nu < \frac{\epsilon}{2}$

Now let $G = (F' - \delta)$ on U_n^c, and $G = 0$ on U_n. Clearly $G(x) \leq g(x)$, for all $x \in X$. Moreover;

$\int_X G d\nu$

$= \int_{U_n^c} (F' - \delta) d\nu$

$= \int_X (F' - \delta) d\nu - \int_{U_n} (F' - \delta) d\nu > C - \epsilon$

The same argument as above shows that, for all $A \in \mathfrak{A}$;

$\int_A g d\mu_L - \int_A G d\nu \leq \epsilon$

Hence, G is a lower bound.

Now, if g is integrable μ_L, we can write $g = g^+ - g^-$, with $\{g^+, g^-\}$ integrable μ_L. Choosing $G \geq g^+$ and $H \leq g^-$, $G - H \geq (g^+ - g^-) = g$, choosing $G' \leq g^+$ and $H' \geq g^-$, $G' - H' \leq (g^+ - g^-) = g$, and, clearly, we can obtain the integral condition, using $\frac{\epsilon}{2}$.

□

References

[1] Linear Analysis, Bela Bollobas, Cambridge Mathematical Textbooks, (1990).

[2] A Simple Proof of the Ergodic Theorem Using Nonstandard Analysis, Teturo Kamae, Isreal Journal of Mathematics, (1982).

[3] Ergodic Theory, Karl Petersen, Cambridge Studies in Advanced Mathematics, (1983).

[4] Applications of Nonstandard Analysis to Probability Theory, Tristram de Piro, Msc Thesis, University of Exeter, available at http://www.magneticstrix.net, (2013).

[5] Arguments for the Wave Equation in Quantum Theory I, Tristram de Piro, available at http://www.magneticstrix.net, (2021).

[6] Non-Standard Analysis, Abraham Robinson, Studies in Logic and the Foundation of Mathematics, North-Holland, (1974).

[7] Real and Complex Analysis, Walter Rudin, McGraw Hill Third Edition, (1987).

Mathematics Department, The University of Exeter, Exeter
E-mail address: t.de-piro@exeter.ac.uk