A modern exposition of Kakutani's criterion for equivalence of product measures

Sasha Bell and Owen Rodgers

November 2023

1 Introduction

In [Kak48], S. Kakutani gives a characterization of when two infinite product measures are equivalent or orthogonal. We present a modern exposition of Kakutani's characterization, including many of the details that were left out of the original publication.

Acknowledgements

This exposition was written as part of a summer research project at McGill University in 2023, under the supervision of Professor Anush Tserunyan. S. Bell was partially supported by an NSERC Undergraduate Student Research Award, with supplemental funding from FRQNT. O. Rodgers was partially supported by a McGill Science Undergraduate Research Award. Both authors were partially supported by A. Tserunyan's NSERC Discovery Grant RGPIN-2020-07120.

2 Preliminaries

Every measure we refer to is a probability measure defined on an arbitrary set X. We assume without loss of generality that the measure is complete. For any two such measures μ and ν, we say that μ is **absolutely continuous** with respect to ν, denoted $\mu \ll \nu$, if $\nu(A) = 0$ implies $\mu(A) = 0$ for all measurable $A \subseteq X$.

If $\mu \ll \nu$, there exists a unique measurable function $\frac{d\nu}{d\mu} : X \to \mathbb{R}^+$, called the **Radon-Nikodym derivative**, such that

$$\nu(A) = \int_A \frac{d\nu}{d\mu} d\mu$$

for all measurable $A \subseteq X$.

μ and ν are called **orthogonal**, and we write $\mu \perp \nu$, if there exists a measurable set A such that $\mu(A) = 0$ and $\nu(X \setminus A) = 0$. Equivalently, $\mu \perp \nu$ if for all $\varepsilon > 0$ there exists a measurable set A such that $\mu(A) < \varepsilon$ and $\nu(X \setminus A) < \varepsilon$.

μ and ν are **equivalent**, denoted $\mu \sim \nu$, if $\mu \ll \nu$ and $\nu \ll \mu$.

We would like to determine whether two probability measures μ and ν are equivalent or orthogonal, by comparing their Radon-Nikodym derivatives using inner products. We consider

$$\sqrt{\frac{d\nu}{d\mu}}$$

so it is an element of $L^2(X, \mu)$. We will also use the L^2 norm to obtain a convenient expression for the Radon-Nikodym derivatives of product measures in terms of the derivatives of the marginal measures.

Let $\lambda := \mu + \nu$ so that $\mu \ll \lambda$ and $\nu \ll \lambda$.

Note that $\left\|\sqrt{\frac{d\mu}{d\lambda}}\right\|_2 = \left\|\sqrt{\frac{d\nu}{d\lambda}}\right\|_2 = 1$. For instance,

$$\left\langle \sqrt{\frac{d\mu}{d\lambda}}, \sqrt{\frac{d\mu}{d\lambda}} \right\rangle = \int_X \frac{d\mu}{d\lambda} d\lambda = \mu(X) = 1,$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product in $L^2(X, \lambda)$.

For ease of notation, we define

$$\rho(\mu, \nu) := \left\langle \sqrt{\frac{d\mu}{d\lambda}}, \sqrt{\frac{d\nu}{d\lambda}} \right\rangle = \int_X \sqrt{\frac{d\mu}{d\lambda}} \sqrt{\frac{d\nu}{d\lambda}} d\lambda.$$

Notice that $\rho(\mu, \nu) = \rho(\nu, \mu)$. Also, we have that $\rho(\mu, \nu) = 0$ if and only if $\mu \perp \nu$.

One can view the space of probability measures on X which are absolutely continuous with respect to λ as an embedding into $L^2(X, \lambda)$, with the corresponding metric

$$d(\mu, \nu) := \left\| \sqrt{\frac{d\mu}{d\lambda}} - \sqrt{\frac{d\nu}{d\lambda}} \right\|_2$$

$$= \sqrt{\left\| \sqrt{\frac{d\mu}{d\lambda}} \right\|_2^2 - 2 \left\langle \sqrt{\frac{d\mu}{d\lambda}}, \sqrt{\frac{d\nu}{d\lambda}} \right\rangle + \left\| \sqrt{\frac{d\nu}{d\lambda}} \right\|_2^2} = (2(1 - \rho(\mu, \nu)))^{1/2}.$$

It then follows by the Cauchy-Schwartz inequality, for $\mu \sim \nu$, that $0 < \rho(\mu, \nu) \leq 1$, and $\rho(\mu, \nu) = 1$ if and only if $\mu = \nu$. With this, we may view the space of probability measures on X as a subspace of the unit sphere in $L^2(X, \mu)$.

In the case where $\mu \sim \nu$, it follows from the chain rule that

$$\rho(\mu, \nu) = \int_X \sqrt{\frac{d\nu}{d\mu}}\, d\mu. \tag{2.1}$$

Remark 2.2. Let λ' be another measure such that $\mu \ll \lambda'$ and $\nu \ll \lambda'$. Then clearly $\lambda \ll \lambda'$, so it follows from the chain rule that

$$\int_X \sqrt{\frac{d\mu}{d\lambda}}\sqrt{\frac{d\nu}{d\lambda}} d\lambda = \int_X \sqrt{\frac{d\mu}{d\lambda'}}\sqrt{\frac{d\nu}{d\lambda'}} d\lambda'.$$

Hence, $\rho(\mu, \nu)$ does not depend on the choice of measure λ.

Now let $\{X_n, \mathcal{B}_n\}_{n\in\mathbb{N}}$ be a countable collection of measurable spaces, and let μ_n and ν_n be equivalent probability measures on X_n for each $n \in \mathbb{N}$. We define product measures

$$\mu := \prod_{n\in\mathbb{N}} \mu_n \text{ and } \nu := \prod_{n\in\mathbb{N}} \nu_n.$$

Theorem 2.3. *For $\{(X_n, \mathcal{B}_n)\}_{n\in\mathbb{N}}, \mu_n, \nu_n, \mu$ and ν as above, the following are equivalent:*

(1) μ and ν are not orthogonal.

(2) $\mu \sim \nu$.

(3) $\prod_{n\in\mathbb{N}} \rho(\mu_n, \nu_n) > 0$.

(4) $\sum_{n\in\mathbb{N}} - \log \rho(\mu_n, \nu_n) < \infty$.

(5) $\sum_{n\in\mathbb{N}} d^2(\mu_n, \nu_n) < \infty$.

Proof. We show that

$$(3) \implies (1) \implies (2) \implies (3) \iff (4) \iff (5).$$

(2) $\implies$ (1) and (3) $\iff$ (4) are trivial.

(4) $\iff$ (5) : Recall that

$$\sum_{n\in\mathbb{N}} d^2(\mu_n, \nu_n) = 2\sum_{n\in\mathbb{N}} (1 - \rho(\mu_n, \nu_n)).$$

We show that for a sequence $(a_n)_{n\in\mathbb{N}}$ with $a_n \in (0,1]$ for all $n \in \mathbb{N}$, the series $\sum_{n\in\mathbb{N}}(1-a_n)$ converges if and only if the series $\sum_{n\in\mathbb{N}} -\log a_n$ converges.

Observe that if either $\sum_{n\in\mathbb{N}}(a_n - 1)$ or $\sum_{n\in\mathbb{N}} -\log(a_n)$ converges then

$$\lim_{n\to\infty}(a_n - 1) = 0.$$

For any $x > 0$, we have that

$$\lim_{x\to 0} \frac{\log(1+x)}{x} = 1,$$

so

$$\lim_{n\to\infty} \frac{\log a_n}{a_n - 1} = 1.$$

Whenever one of the sequences converges. Thus we have by the limit comparison test that $\sum_{n\in\mathbb{N}}(1-a_n)$ converges if and only if $-\sum_{n\in\mathbb{N}} \log a_n$ converges.

The result follows since we can take $a_n := \rho(\mu_n, \nu_n)$ for all $n \in \mathbb{N}$.

To prove (3) $\implies$ (2) and (1) $\implies$ (3), completing the proof, we use the following lemmas.

Lemma 2.4.

$$R_n := \frac{d\nu_n}{d\mu_n} : X_n \to \mathbb{R}^+$$

is a real-valued measurable function for each n.

For $x := (x_1, ..., x_k) \in \prod_{n\leq k} X_n$, define

$$R_n(x) := R_n(x_n).$$

Then $(R_n)_{n\leq k}$ is a system of real-valued measurable functions defined over $\prod_{n\leq k} X_n$, which are independent when viewed as random variables over the space of measurable subsets of $\prod_{n\leq k} X_n$.

This extends to the infinite case.

Proof. For all $n, R_n : X_n \to \mathbb{R}^+$ is $\mathcal{B}_n$-measurable by the Radon-Nikodym theorem. Hence, as a function over $\prod_{n\leq k} X_n$ or $\prod_{n\in\mathbb{N}} X_n, R_n$ is also measurable.

$(R_n)_{n\in\mathbb{N}}$ is a system of independent functions since for all $n \in \mathbb{N}, x \in X, R_n(x)$ depends only on the n-th coordinate of x. □

Lemma 2.5. *Let $\mu_{\leq k} := \mu_1 ... \mu_k$ and let $\nu_{\leq k} := \nu_1 ... \nu_k$.*

If $\mu_n \sim \nu_n$ for all $n \leq k$, then $\mu_{\leq k} \sim \nu_{\leq k}$. Moreover we have that

$$\frac{d\nu_{\leq k}}{d\mu_{\leq k}} = \prod_{n \leq k} R_n \tag{2.6}$$

and

$$\rho(\mu_{\leq k}, \nu_{\leq k}) = \prod_{n \leq k} \rho(\mu_n, \nu_n).$$

Proof. Assume $\mu_n \sim \nu_n$ for all $n \leq k$.

Let $E := \prod_{n \leq k} E_n \subseteq \prod_{n \leq k} X_n$ be an elementary set. Then

$$\nu_{\leq k}(E) = \nu_1(E_1)...\nu_k(E_k) = \prod_{n \leq k} \int_{E_n} R_n \mu_n.$$

By the independence of the R_n's, and an application of Tonelli's Theorem,

$$\nu_{\leq k}(E) = \prod_{n \leq k} \int_{E_n} R_n \mu_{\leq k} = \int_{E_1} ... \int_{E_k} \prod_{n \leq k} R_n \mu_{\leq k}$$

$$= \int_{E_1 \times ... E_k} \prod_{n \leq k} R_n \mu_{\leq k}.$$

So for all elementary sets E,

$$\nu_{\leq k}(E) = \int_E \prod_{n \leq k} R_n \mu_{\leq k}.$$

Notice that the function λ defined by

$$\lambda(B) := \int_B \prod_{n \leq k} R_n \mu_{\leq k},$$

for all measurable $B \subseteq \prod_{n \leq k} X_n$, is a measure on $\prod_{n \leq k} X_n$.

Since λ and ν agree on the elementary subsets of $\prod_{n \leq k} X_n$, they agree on all measurable sets by the Caratheodory extension theorem.

Hence, $\nu_{\leq k} \ll \mu_{\leq k}$. One shows that $\mu_{\leq k} \ll \nu_{\leq k}$ by a completely symmetrical argument, so we have that $\mu_{\leq k} \sim \nu_{\leq k}$.

(2.6) follows from the uniqueness of the Radon-Nikodym derivative.

Moreover, the independence of the R_n's gives that

$$\rho(\mu_{\leq k}, \nu_{\leq k}) = \int_{X_{\leq k}} \sqrt{R_1 ... R_k} d\mu_{\leq k} = \prod_{n \leq k} \int_{X_n} \sqrt{R_n} d\mu_n = \prod_{n \leq k} \rho(\mu_n, \nu_n).$$

□

Lemma 2.7. *For $k \in \mathbb{N}$, define*

$$\psi_k \coloneqq \prod_{n \leq k} \sqrt{R_n}.$$

Then $(\psi_k)_{k \geq 1}$ is a sequence of elements of $L^2(X, \mu)$ such that $\|\psi_k\|_2 = 1$ for all k.

Moreover, for any $l > k \geq 1$ with $k < l$, we have that

$$\|\psi_k - \psi_l\|_2^2 = 2(1 - \prod_{n=k+1}^{l} \int_{X_n} \sqrt{R_n} d\mu_n)$$

$$= 2(1 - \prod_{n=k+1}^{l} \rho(\mu_n, \nu_n)).$$

Proof.

$$\|\psi_k\|_2^2 = \int_X \prod_{n \leq k} R_n d\mu$$

$$= \int_{X_1 \times ... \times X_k} \prod_{n \leq k} R_n d\mu = \int_{X_k} ... \int_{X_1} \prod_{n \leq k} R_n d\mu$$

by Tonelli's Theorem.

By the independence of the R_n's, and since each R_n only depends on the n-th coordinate of $x \in X$, we have that

$$\|\psi_k\|_2^2 = \prod_{n \leq k} \int_{X_n} R_n d\mu = \prod_{n \leq k} \int_{X_n} R_n d\mu_n$$

$$= \prod_{n \leq k} \nu_n(X_n) = 1.$$

So $\psi_k \in L^2(X, \mu)$ and $\|\psi_k\|_2 = 1$ for all $k \in \mathbb{N}$.

Now let $k, l \geq 1$ be such that $k < l$. Then

$$\langle \psi_k, \psi_l \rangle = \int_X \psi_k \psi_l d\mu = \int_{X_1 \times \ldots X_l} (\prod_{n \leq k} R_n)(\prod_{n=k+1}^{l} \sqrt{R_n}) d\mu.$$

So by Tonelli's Theorem and the independence of the R_n's,

$$\langle \psi_k, \psi_l \rangle = \left(\prod_{n \leq k} \int_{X_n} R_n d\mu \right) \left(\prod_{n=k+1}^{l} \int_{X_n} \sqrt{R_n} d\mu \right)$$

$$= \prod_{n=k+1}^{l} \int_{X_n} \sqrt{R_n(x_n)} d\mu_n = \prod_{n=k+1}^{l} \rho(\mu_n, \nu_n).$$

Combining these results, we have that

$$\|\psi_k - \psi_l\|_2 = \|\psi_k\|_2^2 + \|\psi_l\|_2^2 - 2\langle \psi_k, \psi_l \rangle = 2(1 - \prod_{n=k+1}^{l} \rho(\mu_n, \nu_n)).$$ □

Lemma 2.8. *If* $\prod_{n \in \mathbb{N}} \rho(\mu_n, \nu_n) > 0$, *then* $(\psi_k)_{k \in \mathbb{N}}$ *is a Cauchy sequence in* $L^2(X, \mu)$, *and therefore has a limit* ψ.

Proof. Suppose $\prod_{n \in \mathbb{N}} \rho(\mu_n, \nu_n) > 0$. Recall from Lemma 2.7 that

$$\|\psi_k - \psi_l\|_2 = 2(1 - \prod_{n=k+1}^{l} \rho(\mu_n, \nu_n)).$$

We show that $1 - \prod_{n=k+1}^{l} \rho(\mu_n, \nu_n)$ approaches 0 as l and k approach infinity.

Let $(p_n)_{n \in \mathbb{N}}$ be the sequence of partial products of $\prod_{n \in \mathbb{N}} \rho(\mu_n, \nu_n)$. Then since the product is convergent, $\lim_{n \to \infty} p_n$ exists and is strictly positive. $(p_n)_{n \in \mathbb{N}}$ being convergent implies that $(p_n)_{n \in \mathbb{N}}$ is Cauchy. Moreover, since the product is convergent, there exists $M > 0$ such that $p_n > M$ for each $n \in \mathbb{N}$.

So we have that for all $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that for all $l > k \geq N$,

$$|p_k - p_l| = |\prod_{n \leq k} \rho(\mu_n, \nu_n)||1 - \prod_{n=k+1}^{l} \rho(\mu_n, \nu_n)| < \varepsilon M.$$

So

$$|1 - \prod_{n=k+1}^{l} \rho(\mu_n, \nu_n)| < \varepsilon M (|\prod_{n \leq k} \rho(\mu_n, \nu_n)|)^{-1} \leq \varepsilon M M^{-1} = \varepsilon.$$

So for all $\varepsilon > 0$ there exists N such that for all $k, l > N$, $\|\psi_k - \psi_l\|_2 < \varepsilon$, so the sequence $(\psi_n)_{n\in\mathbb{N}}$ is Cauchy.

□

For (3) $\Longrightarrow$ (2),

Suppose $\prod_{n\in\mathbb{N}} \rho(\mu_n, \nu_n) > 0$. We will show that

$$\nu(B) = \int_B \psi^2 d\mu,$$

for any measurable $B \subseteq X$, and hence that $\nu \ll \mu$ and $\frac{d\nu}{d\mu} = \psi^2$.

We first show this for an elementary set $E \subseteq X$. For some $k \in \mathbb{N}$, we have that $\nu(E) = \nu_{\leq k}(E)$. By the independence of the R_n's,

$$\nu(E) = \nu_{\leq k}(E_{\leq k}) = \int_{E_{\leq k}} \prod_{n\leq k} R_n d\mu_{\leq k} = \int_E \prod_{n=1}^{k} R_n d\mu = \int_E \prod_{n=1}^{l} R_n d\mu = \int_E \psi_l^2 d\mu$$

for any $l > k$.

Using that $\lim_{l\to\infty} \|\psi_l - \psi\|_2 = 0$, we get that

$$\nu(E) = \int_E \psi^2 d\mu. \tag{2.9}$$

Recall that $\mathcal{E}$ is the ring of elementary subsets of X. Notice that the function $\lambda : \mathcal{B} \to \mathbb{R}^+$ defined by

$$\lambda(B) := \int_B \psi^2 d\mu,$$

for measurable $B \subseteq X$, is a measure on $\mathcal{B}$. Also, λ agrees with ν on $\mathcal{E}$, by (2.9). By the Caratheodory Extension Theorem, λ and ν will also agree on $\mathcal{B}$.

Hence,

$$\nu(B) = \int_B \psi^2 d\mu,$$

for any measurable $B \subseteq X$.

So $\nu \ll \mu$ and $\frac{d\nu}{d\mu} = \psi^2$. One shows that $\mu \ll \nu$ by a completely symmetrical argument, so μ and ν are equivalent.

(1) $\Longrightarrow$ (3) is proven by contrapositive: assuming $\prod_{n\in\mathbb{N}} \rho(\mu_n, \nu_n) = 0$, we will show that $\mu \perp \nu$.

Fix $\varepsilon > 0$ and $k \in \mathbb{N}^+$ such that

$$\prod_{n \le k} \rho(\mu_n, \nu_n) < \varepsilon.$$

Let $B_{\le k}$ be the set of all $x = (x_1, ..., x_k) \in \prod_{n \le k} X_n$ such that

$$\prod_{n \le k} R_n(x) > 1.$$

Note that $B_{\le k}$ is measurable since the R_n's are measurable functions.

Then

$$\mu_{\le k}(B_{\le k}) = \int_{B_{\le k}} d\mu_{\le k} \le \int_{B_{\le k}} \prod_{n \le k} \sqrt{R_n} d\mu_{\le k} \le \rho(\mu_{\le k}, \nu_{\le k}) < \varepsilon.$$

Also,

$$\nu_{\le k}(X_{\le k} \setminus B_{\le k}) = \int_{X_{\le k} \setminus B_{\le k}} \prod_{n \le k} R_n d\mu_{\le k}.$$

Since the R_n's are less than or equal to 1 on $X_{\le k} \setminus B_{\le k}$,

$$\nu_{\le k}(X_{\le k} \setminus B_{\le k}) \le \int_{X_{\le k} \setminus B_{\le k}} \prod_{n \le k} \sqrt{R_n} d\mu_{\le k}$$
$$\le \rho(\mu_{\le k}, \nu_{\le k}) < \varepsilon.$$

So if we let $B \coloneqq B_{\le k} \times \prod_{n=k+1}^{\infty} X_n$, then

$$\mu(B) = \mu_{\le k}(B_{\le k}) < \varepsilon$$

and

$$\nu(X \setminus B) = \nu_{\le k}(X \setminus B_{\le k}) < \varepsilon.$$

So for any $\varepsilon > 0$, there exists a measurable $B \subseteq X$ such that $\mu(B) < \varepsilon$ and $\nu(X \setminus B) < \varepsilon$. Hence, μ and ν are orthogonal. $\square$

Corollary 2.10. *For μ and ν defined as above, we have a dichotomy: μ and ν are either equivalent or perpendicular.*

Corollary 2.11. *For μ and ν defined as above,*

$$\rho(\mu, \nu) = \prod_{n \in \mathbb{N}} \rho(\mu_n, \nu_n).$$

Proof. By Theorem 2.3, μ and ν are either equivalent or orthogonal, and if they are orthogonal then

$$\prod_{n\in\mathbb{N}} \rho(\mu_n, \nu_n) = \rho(\mu, \nu) = 0.$$

If μ and ν are equivalent, recall that

$$\psi_k = \prod_{n\leq k} \sqrt{R_n}$$

for all k. It follows from the proof of Theorem 2.3 (specifically (3)⇒(2)) that

$$\rho(\mu, \nu) = \int_X \psi d\mu = \lim_{n\to\infty} \int_X \psi_n d\mu$$
$$= \lim_{n\to\infty} \rho(\mu_{\leq n}, \nu_{\leq n}) = \lim_{n\to\infty} \prod_{k=1}^{n} \rho(\mu_n, \nu_n).$$

□

Lemma 2.8 stated that $(\psi_k)_{k\in\mathbb{N}}$ converges in $L^2(X, \mu)$. The next corollary gives a stronger convergence which will allow us to express the Radon-Nikodym derivative of the product measure in terms of the derivatives of the marginal measures:

Corollary 2.12. *In the context of Theorem 2.3, $(\psi_k)_{k\in\mathbb{N}}$ converges to ψ μ-almost everywhere on X, and so*

$$\frac{d\mu}{d\nu} = \prod_{n\in\mathbb{N}} \frac{d\mu_n}{d\nu_n}$$

μ-almost everywhere on X.

Proof. $(\psi_k)_{k\in\mathbb{N}}$ converges to ψ in $L^2(X, \mu)$, so for any $\varepsilon > 0$, we have that

$$\lim_{k\to\infty} \mu(\{x : |\psi_k(x) - \psi(x)| > \varepsilon\}) = 0. \tag{2.13}$$

Also, $\psi > 0$ μ-a.e. Hence, the sequence

$$(\log \psi_k)_{k\in\mathbb{N}} = (\frac{1}{2} \sum_{n\leq k} R_n)_{k\in\mathbb{N}}$$

satisfies

$$\lim_{k\to\infty} \mu(\{x : |\log \psi_k(x) - \log \psi(x)| > \varepsilon\}) = 0. \tag{2.14}$$

Since $(\log R_n)_{n\in\mathbb{N}}$ is a system of independent functions, (2.14) implies that $(\log \psi_k)_{k\in\mathbb{N}}$ converges to $\log \psi$ μ-almost everywhere. This is a general fact from [Kam40] which states that any series of independent functions on an infinite product space that are converging in measure also converge μ almost-everywhere.

Consequently, $(\psi_k)_{k\in\mathbb{N}}$ converges to ψ μ-almost everywhere on X. □

3 Application to the Cantor Space

Setting $X := 2^{\mathbb{N}}$, we show how the sum convergence can be simplified. We consider families of measures $(\mu_n)_{n\in\mathbb{N}}$ and $(\nu_n)_{n\in\mathbb{N}}$, and set $\mu_n(0) =: \alpha_n$ and $\nu_n(0) =: \beta_n$ for all n. Then, a quick calculation yields

$$\rho(\mu_n, \nu_n) = \sqrt{\alpha_n\beta_n} - \sqrt{(1-\alpha_n)(1-\beta_n)}.$$

We now show a convenient formulation of $d^2(\mu_n, \nu_n)$.

$$\begin{aligned} d^2(\mu_n,\nu_n) &= 2(1-\rho(\mu_n,\nu_n)) \\ &= 2(1-(\sqrt{\alpha_n\beta_n} - \sqrt{(1-\alpha_n)(1-\beta_n)})) \\ &= \alpha_n - 2\sqrt{\alpha_n\beta_n} + \beta_n + (1-\alpha_n) - 2\sqrt{1-\alpha_n}\sqrt{1-\beta_n} + (1-\beta_n) \\ &= (\sqrt{\alpha_n} - \sqrt{\beta_n})^2 + (\sqrt{1-\alpha_n} - \sqrt{1-\beta_n})^2. \end{aligned}$$

Thus, the equivalence or orthogonality of the product measure is determined by the convergence or divergence respectively of the sum

$$\sum_{n\in\mathbb{N}} (\sqrt{\alpha_n} - \sqrt{\beta_n})^2 + (\sqrt{1-\alpha_n} - \sqrt{1-\beta_n})^2. \tag{3.1}$$

Proposition 3.2. *If there exists a $\gamma \in (0,1)$ such that*

$$\gamma \le \alpha_n, \beta_n \le 1-\gamma, \tag{3.3}$$

it is sufficient to check the convergence or divergence of the sum

$$\sum_{n\in\mathbb{N}} (\alpha_n - \beta_n)^2 \tag{3.4}$$

to determine whether μ and ν are equivalent or orthogonal.

Proof. Suppose (3.3) holds. We show that $\sum_{n\in\mathbb{N}} d^2(\mu_n, \nu_n)$ converges if and only if (3.4) converges via the limit comparison test on the individual summands of (3.1).

$$\lim_{n\to\infty} \frac{(\alpha_n-\beta_n)^2}{(\sqrt{\alpha_n}-\sqrt{\beta_n})^2} = \lim_{n\to\infty} \frac{(\alpha_n-\beta_n)^2}{(\sqrt{\alpha_n}-\sqrt{\beta_n})^2} \cdot \frac{(\sqrt{\alpha_n}+\sqrt{\beta_n})^2}{(\sqrt{\alpha_n}+\sqrt{\beta_n})^2}$$
$$= \lim_{n\to\infty} \frac{(\alpha_n-\beta_n)^2 \cdot (\sqrt{\alpha_n}+\sqrt{\beta_n})^2}{(\alpha_n-\beta_n)^2} = \lim_{n\to\infty} (\sqrt{\alpha_n}+\sqrt{\beta_n})^2 \le (1+1)^2 = 4.$$

Thus since the limit exists and is positive, the two series both converge or both diverge.

We now apply the limit comparison test to the other summand.

$$\lim_{n\to\infty} \frac{(\alpha_n - \beta_n)^2}{(\sqrt{1-\alpha_n} - \sqrt{1-\beta_n})^2} = \lim_{n\to\infty} \frac{(\alpha_n - \beta_n)^2}{(\sqrt{1-\alpha_n} - \sqrt{1-\beta_n})^2} \cdot \frac{(\sqrt{1-\alpha_n} + \sqrt{1-\beta_n})^2}{(\sqrt{1-\alpha_n} + \sqrt{1-\beta_n})^2}$$

$$= \lim_{n\to\infty} \frac{(\alpha_n - \beta_n)^2(\sqrt{1-\alpha_n} - \sqrt{1-\beta_n})^2}{(\alpha_n - \beta_n)^2} = \lim_{n\to\infty} (\sqrt{1-\alpha_n} - \sqrt{1-\beta_n})^2.$$

Now by (3.3), this limit is a finite positive number, so both series converge or diverge. Thus, since both summands convergence is governed by (3.4), we have that the whole of (3.1) is as well. □

Proposition 3.5. *If $\mu \sim \nu$, then*

$$\frac{d\nu}{d\mu}(x) = \lim_{n\to\infty} \prod_{k\leq n} \frac{\nu_n(x)}{\mu_n(x)}.$$

Proof. Fix $n \in \mathbb{N}$. Then

$$\int_{\{x\}} \frac{\nu_n}{\mu_n} d\mu_n = \frac{\nu_n(x)}{\mu_n(x)} \mu_n(x) = \nu_n(x)$$

for $x \in \{0, 1\}$, so by the uniqueness of the Radon-Nikodym derivative,

$$\frac{d\nu_n}{d\mu_n} = \frac{\nu_n}{\mu_n}$$

for all $n \in \mathbb{N}$. The result then follows from Corollary 2.12. □

References

[Kak48] Shizuo Kakutani. On equivalence of infinite product measures. *Annals of Mathematics*, 49(1):214–224, 1948.

[Kam40] E. R. Van Kampen. Infinite product measures and infinite convolutions. *American Journal of Mathematics*, 62(1):417–448, 1940.