

Simplification and Strengthening of Weyl's Definition of Asymptotic Equal Distribution of Two Families of Finite Sets

William F. Trench

Mailing address: 95 Pine Lane, Woodland Park, CO 80863 USA
Trinity University, San Antonio, Texas, USA
wtrench@trinity.edu

ABSTRACT

Suppose that $-\infty < a < b < \infty$, $a \leq u_{1n} \leq u_{2n} \leq \dots \leq u_{nn} \leq b$, and $a \leq v_{1n} \leq v_{2n} \leq \dots \leq v_{nn} \leq b$, $n \geq 1$. We simplify and strengthen Weyl's definition of asymptotic equal distribution of $\mathbf{U} = \{\{u_{in}\}_{i=1}^n\}_{n \geq 1}$ and $\mathbf{V} = \{\{v_{in}\}_{i=1}^n\}_{n \geq 1}$ by showing that the following statements are equivalent:

- (i) $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n (F(u_{in}) - F(v_{in})) = 0$ for all $F \in C[a, b]$.
- (ii) $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n |u_{in} - v_{in}| = 0$.
- (iii) $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n |F(u_{in}) - F(v_{in})| = 0$ for all $F \in C[a, b]$.

RESUMEN

Suponemos que $-\infty < a < b < \infty$, $a \leq u_{1n} \leq u_{2n} \leq \dots \leq u_{nn} \leq b$, y $a \leq v_{1n} \leq v_{2n} \leq \dots \leq v_{nn} \leq b$, $n \geq 1$. Simplificamos y fortalecemos la definición de Weyl de la distribución de igualdad asintótica de $\mathbf{U} = \{\{u_{in}\}_{i=1}^n\}_{n \geq 1}$ y $\mathbf{V} = \{\{v_{in}\}_{i=1}^n\}_{n \geq 1}$ mostrando que las siguientes afirmaciones son equivalentes:

- (i) $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n (F(u_{in}) - F(v_{in})) = 0$ para todo $F \in C[a, b]$.
- (ii) $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n |u_{in} - v_{in}| = 0$.
- (iii) $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n |F(u_{in}) - F(v_{in})| = 0$ para todo $F \in C[a, b]$.

Key words and phrases: asymptotic equal distribution
Math. Subj. Class.: 15A18, 15A57

1 Introduction

The following definition is due to H. Weyl [1, p. 62].

Definition 1.1 Suppose that $-\infty < a < b < \infty$,

$$\{u_{in}\}_{i=1}^n \subset [a, b], \quad \text{and} \quad \{v_{in}\}_{i=1}^n \subset [a, b], \quad n \geq 1.$$

Then $\mathbf{U} = \{\{u_{in}\}_{i=1}^n\}_{n \geq 1}$ and $\mathbf{V} = \{\{v_{in}\}_{i=1}^n\}_{n \geq 1}$ are *asymptotically equally distributed* if

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n (F(u_{in}) - F(v_{in})) = 0, \quad F \in C[a, b].$$

We present a simple necessary and sufficient condition for asymptotic equal distribution and point out that a stronger conclusion is implicit in Definition 1.1.

Without loss of generality, we may assume that

$$a \leq u_{1n} \leq u_{2n} \leq \cdots \leq u_{nn} \leq b, \quad a \leq v_{1n} \leq v_{2n} \leq \cdots \leq v_{nn} \leq b, \quad n \geq 1. \quad (1)$$

Theorem 1.2 *If (1) holds then the following assertions are equivalent:*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n (F(u_{in}) - F(v_{in})) = 0, \quad F \in C[a, b]; \quad (2)$$

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n |u_{in} - v_{in}| = 0; \quad (3)$$

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n |F(u_{in}) - F(v_{in})| = 0, \quad F \in C[a, b]. \quad (4)$$

Obviously, (4) implies (2). The proof that (3) implies (4) (Section 2) is straightforward. Our main effort is devoted to showing that (2) implies (3).

Theorem 1.2 is a special case of more general results in [4] concerning asymptotic relationships between the eigenvalues or singular values of two infinite sequences of matrices $\{A_n\}_{n=1}^{\infty}$ and $\{B_n\}_{n=1}^{\infty}$ related in some way that it is not necessary to specify here. However, [4] is quite technical and of interest mainly to the linear algebra community. We think it is worthwhile to present Theorem 1.2 in this expository article addressed to a larger audience.

Given Theorem 1.2, we suggest replacing Definition 1.1 by the following simpler definition while bearing in mind that (3) implies (4).

Definition 1.3 $\mathbf{U} = \{\{u_{in}\}_{i=1}^n\}_{n \geq 1}$ and $\mathbf{V} = \{\{v_{in}\}_{i=1}^n\}_{n \geq 1}$ are *asymptotically equally distributed* if (1) holds and

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n |u_{in} - v_{in}| = 0.$$

2 Proof that (3) implies (4)

Suppose that $F \in C[a, b]$ and $\epsilon > 0$. By the Weierstrass approximation theorem, there is a polynomial P such that

$$|F(x) - P(x)| < \epsilon/2, \quad a \leq x \leq b.$$

By the triangle inequality,

$$\begin{aligned} |F(u_{in}) - F(v_{in})| &\leq |F(u_{in}) - P(u_{in})| + |P(u_{in}) - P(v_{in})| \\ &\quad + |P(v_{in}) - F(v_{in})| \\ &< |P(u_{in}) - P(v_{in})| + \epsilon. \end{aligned} \tag{5}$$

Let $M = \max_{a \leq x \leq b} |P'(x)|$. By the mean value theorem,

$$|P(u_{in}) - P(v_{in})| \leq M|u_{in} - v_{in}|.$$

This and (5) imply that

$$\frac{1}{n} \sum_{i=1}^n |F(u_{in}) - F(v_{in})| < \epsilon + \frac{M}{n} \sum_{i=1}^n |u_{in} - v_{in}|.$$

From this and (3),

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n |F(u_{in}) - F(v_{in})| \leq \epsilon.$$

Since ϵ is arbitrary, this implies (4).

3 Four Required Lemmas

We need the following lemmas to show that (2) implies (3).

Lemma 3.1 (Helly's First Theorem) *Let $\{\phi_m\}_{m=1}^\infty$ be an infinite sequence of functions on $[a, b]$ and suppose that there is a finite number K such that*

$$|\phi_m(x)| \leq K, \quad a \leq x \leq b, \quad \text{and} \quad V_a^b(\phi_m) \leq K, \quad m \geq 1.$$

Then there is a subsequence of $\{\phi_m\}_{m=1}^\infty$ that converges at every point of $[a, b]$ to a function of bounded variation on $[a, b]$.

Lemma 3.2 (Helly's Second Theorem) *Let $\{\phi_m\}_{m=1}^\infty$ be an infinite sequence of functions on $[a, b]$ such that $V_a^b(\phi_m) \leq K < \infty$, $m \geq 1$, and*

$$\lim_{m \rightarrow \infty} \phi_m(x) = \phi(x), \quad a \leq x \leq b.$$

Then $V_a^b(\phi) \leq K$ and

$$\lim_{m \rightarrow \infty} \int_a^b F(x) d\phi_m(x) = \int_a^b F(x) d\phi(x), \quad F \in C[a, b].$$

Lemma 3.3 Suppose that $\phi(a) = \phi(b) = 0$, ϕ is of bounded variation on $[a, b]$, and

$$\int_a^b F(x) d\phi(x) = 0, \quad F \in C[a, b].$$

Then $\phi(x) = 0$ at all points of continuity of ϕ . Thus, $\phi(x) \neq 0$ for at most countably many values of x .

For proofs of Lemmas 3.1–3.3, see [2, p. 222], [2, p. 233], and [3, p. 111].

The following lemma is also known [5, p. 108], but we include its short proof for convenience.

Lemma 3.4 Suppose that $x_1 \leq x_2 \leq \dots \leq x_n$ and $y_1 \leq y_2 \leq \dots \leq y_n$. Let $\{\ell_1, \ell_2, \dots, \ell_n\}$ be a permutation of $\{1, 2, \dots, n\}$ and define

$$Q(\ell_1, \ell_2, \dots, \ell_n) = \sum_{i=1}^n (x_i - y_{\ell_i})^2.$$

Then

$$Q(\ell_1, \ell_2, \dots, \ell_n) \geq Q(1, 2, \dots, n). \quad (6)$$

Proof. The proof is by induction. Let P_n be the stated proposition. P_1 is trivial. Suppose that $n > 1$ and P_{n-1} is true. If $\ell_n = n$, P_{n-1} implies P_n . If $\ell_n = s < n$, choose r so that $\ell_r = n$, and define

$$\ell'_i = \begin{cases} \ell_i & \text{if } i \neq r \text{ and } i \neq n, \\ s & \text{if } i = r, \\ n & \text{if } i = n. \end{cases}$$

Then

$$\begin{aligned} Q(\ell_1, \ell_2, \dots, \ell_n) - Q(\ell'_1, \ell'_2, \dots, \ell'_n) &= (x_n - y_s)^2 + (x_r - y_n)^2 \\ &\quad - (x_n - y_n)^2 - (x_r - y_s)^2 \\ &= 2(x_n - x_r)(y_n - y_s) \geq 0. \end{aligned} \quad (7)$$

Since $\ell'_n = n$, P_{n-1} implies that

$$Q(\ell'_1, \ell'_2, \dots, \ell'_n) \geq Q(1, 2, \dots, n).$$

Therefore (7) implies (6), which completes the induction. ■

4 Proof that (2) implies (3)

We will show that if (2) holds then

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n (u_{in} - v_{in})^2 = 0. \quad (8)$$

From Schwarz's inequality,

$$\frac{1}{n} \sum_{i=1}^n |u_{in} - v_{in}| \leq \left(\frac{1}{n} \sum_{i=1}^n (u_{in} - v_{in})^2 \right)^{1/2},$$

so (8) implies (3).

The proof of (8) is by contradiction. If (8) is false, there is an $\epsilon_0 > 0$ and an increasing sequence $\{\ell_k\}_{k=1}^\infty$ of positive integers such that

$$\frac{1}{\ell_k} \sum_{i=1}^{\ell_k} (u_{i\ell_k} - v_{i\ell_k})^2 \geq \epsilon_0, \quad k \geq 1. \quad (9)$$

However, we will show that if (2) holds, then any increasing sequence $\{\ell_k\}_{k=1}^\infty$ of positive integers has a subsequence $\{n_k\}_{k=1}^\infty$ such that

$$\lim_{k \rightarrow \infty} \frac{1}{n_k} \sum_{i=1}^{n_k} (u_{in_k} - v_{in_k})^2 = 0, \quad (10)$$

contradicting (9).

If S is a set, let $\text{card } S$ be the cardinality of S . For $a \leq x \leq b$, let

$$\nu_n(x; \mathbf{U}) = \text{card} \{i \mid u_{in} < x\} \quad \text{and} \quad \nu_n(x; \mathbf{V}) = \text{card} \{i \mid v_{in} < x\}. \quad (11)$$

Define

$$\rho_n(x; \mathbf{U}) = \begin{cases} \nu_n(x; \mathbf{U})/n, & a \leq x < b, \\ 1, & x = b, \end{cases} \quad (12)$$

and

$$\rho_n(x; \mathbf{V}) = \begin{cases} \nu_n(x; \mathbf{V})/n, & a \leq x < b, \\ 1, & x = b. \end{cases} \quad (13)$$

If $F \in C[a, b]$, then

$$\frac{1}{n} \sum_{i=1}^n F(u_{in}) = \int_a^b F(x) d\rho_n(x; \mathbf{U}) \quad (14)$$

and

$$\frac{1}{n} \sum_{i=1}^n F(v_{in}) = \int_a^b F(x) d\rho_n(x; \mathbf{V}) \quad (15)$$

[2, p. 231]. The sequences $\{\rho_n(\cdot; \mathbf{U})\}_{n=1}^\infty$ and $\{\rho_n(\cdot; \mathbf{V})\}_{n=1}^\infty$ both satisfy the hypotheses of Lemma 3.1. Therefore, there is a subsequence $\{m_k\}_{k=1}^\infty$ of $\{\ell_k\}_{k=1}^\infty$ such that

$$\gamma(x; \mathbf{U}) := \lim_{k \rightarrow \infty} \rho_{m_k}(x; \mathbf{U}) \quad (16)$$

exists for $a \leq x \leq b$, and there is a subsequence $\{n_k\}_{k=1}^\infty$ of $\{m_k\}_{k=1}^\infty$ such that

$$\gamma(x; \mathbf{V}) := \lim_{k \rightarrow \infty} \rho_{n_k}(x; \mathbf{V}) \quad (17)$$

exists for $a \leq x \leq b$. Clearly, (16) implies that

$$\gamma(x; \mathbf{U}) = \lim_{k \rightarrow \infty} \rho_{n_k}(x; \mathbf{U}), \quad a \leq x \leq b. \quad (18)$$

From (11)–(13), $\gamma(\cdot; \mathbf{U})$ and $\gamma(\cdot; \mathbf{V})$ are nondecreasing,

$$\gamma(a; \mathbf{U}) = \gamma(a; \mathbf{V}) = 0, \quad \text{and} \quad \gamma(b; \mathbf{U}) = \gamma(b; \mathbf{V}) = 1. \quad (19)$$

Therefore, (17), (18), and Lemma 3.2 imply that

$$\lim_{k \rightarrow \infty} \int_a^b F(x) d\rho_{n_k}(x; \mathbf{U}) = \int_a^b F(x) d\gamma(x; \mathbf{U}), \quad F \in C[a, b], \quad (20)$$

and

$$\lim_{k \rightarrow \infty} \int_a^b F(x) d\rho_{n_k}(x; \mathbf{V}) = \int_a^b F(x) d\gamma(x; \mathbf{V}), \quad F \in C[a, b]. \quad (21)$$

Now (2), (14), (15) (20), and (21) imply that

$$\int_a^b F(x) d\gamma(x; \mathbf{U}) = \int_a^b F(x) d\gamma(x; \mathbf{V}), \quad F \in C[a, b].$$

This, (19), and Lemma 3.3 with $\phi = \gamma(\cdot; \mathbf{U}) - \gamma(\cdot; \mathbf{V})$ imply that

$$\gamma(x; \mathbf{U}) = \gamma(x; \mathbf{V})$$

except for at most countably many values of x in $[a, b]$.

If $\epsilon > 0$, choose a_0, a_1, \dots, a_m so that

$$\begin{aligned} a &= a_0 < a_1 < \dots < a_m = b, \\ a_j - a_{j-1} &< \sqrt{\epsilon}, \quad 1 \leq j \leq m, \end{aligned} \quad (22)$$

and

$$\gamma(a_j; \mathbf{U}) = \gamma(a_j; \mathbf{V}), \quad 1 \leq j \leq m. \quad (23)$$

Let

$$I_j = [a_{j-1}, a_j], \quad 1 \leq j \leq m-1, \quad I_m = [a_{m-1}, a_m].$$

Define

$$U_{jk} = \begin{cases} \nu_{n_k}(a_1; \mathbf{U}), & j = 1, \\ \nu_{n_k}(a_j; \mathbf{U}) - \nu_{n_k}(a_{j-1}; \mathbf{U}), & 2 \leq j \leq m-1, \\ n_k - \nu_{n_k}(a_{m-1}; \mathbf{U}), & j = m, \end{cases}$$

and

$$V_{jk} = \begin{cases} \nu_{n_k}(a_1; \mathbf{V}), & j = 1, \\ \nu_{n_k}(a_j; \mathbf{V}) - \nu_{n_k}(a_{j-1}; \mathbf{V}), & 2 \leq j \leq m-1, \\ n_k - \nu_{n_k}(a_{m-1}; \mathbf{V}), & j = m. \end{cases}$$

Then

$$U_{jk} = \text{card} \{i \mid u_{in_k} \in I_j\}, \quad V_{jk} = \text{card} \{i \mid v_{in_k} \in I_j\},$$

and

$$\lim_{k \rightarrow \infty} \frac{U_{jk} - V_{jk}}{n_k} = 0, \quad 1 \leq j \leq m, \quad (24)$$

from (12), (13), (17), (18), and (23). Since

$$\min(U_{jk}, V_{jk}) = \frac{U_{jk} + V_{jk} - |U_{jk} - V_{jk}|}{2},$$

and

$$\sum_{j=1}^m U_{jk} = \sum_{j=1}^m V_{jk} = n_k,$$

it follows that

$$\sum_{j=1}^m \min(U_{jk}, V_{jk}) = n_k - r_k, \quad (25)$$

where

$$r_k = \frac{1}{2} \sum_{j=1}^m |U_{jk} - V_{jk}|.$$

From (24),

$$\lim_{k \rightarrow \infty} \frac{r_k}{n_k} = 0. \quad (26)$$

From (22) and (25), there is a permutation τ_{n_k} of $\{1, \dots, n_k\}$ such that

$$(u_{in_k} - v_{\tau_k(i), n_k})^2 < \epsilon$$

for $n_k - r_k$ values of i ; hence

$$\sum_{i=1}^{n_k} (u_{in_k} - v_{\tau_k(i), n_k})^2 < n_k \epsilon + r_k (b-a)^2.$$

Now Lemma 3.4 implies that

$$\sum_{i=1}^{n_k} (u_{in_k} - v_{in_k})^2 < n_k \epsilon + r_k (b-a)^2.$$

Hence, from (26),

$$\limsup_{k \rightarrow \infty} \frac{1}{n_k} \sum_{i=1}^{n_k} (u_{in_k} - v_{in_k})^2 \leq \epsilon.$$

Since ϵ is arbitrary, this implies (10), which completes the proof.

5 Acknowledgment

I thank Professor Paolo Tilli for a suggestion that enabled me to complete the proof in Section 4.

Received: March 2003. Revised: June 2003.

References

- [1] U. Grenander, G. Szegő, *Toeplitz Forms and Their Applications*, Univ. of California Press, Berkeley and Los Angeles, 1958.
- [2] I. P. Natanson, *Theory of Functions of a Real Variable*, Frederick Ungar Publishing Co., New York, 1955.
- [3] F. Riesz and B. Sz.-Nagy, *Functional Analysis*, Frederick Ungar Publishing Co., New York, 1955.
- [4] W. F. Trench, Absolute equal distribution of families of finite sets, *Linear Algebra Appl.* 367 (2003), 131-146.
- [5] J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford, 1965.