axioms

MDPI

Article

A Goodness-of-Fit Test for Uniformity on Distance-Regular Graphs with Applications to Grouped Directional Data

Jean-Renaud Pycke

Laboratoire LaMME (UMR 8071), Departement of Mathematics, University of Paris-Saclay Evry, 91000 Evry-Courcouronnes, France; jeanrenaud.pycke@univ-evry.fr

Abstract

One of the main tasks in the field of directional statistics is to build goodness-of-fit tests for uniformity on circles, spheres or more abstract manifolds. We discuss a goodness-of-fit test for uniformity on a distance-regular graph. Our main tool is the pseudo-inverse of the combinatorial Laplacian, for which we give explicit expressions in terms of the intersection array of the graph. Such a test can be used as a test for uniformity of data from the circle or the sphere, grouped on the tiles of a regular tessellation associated with some finite group of isometries. We describe the cases of Platonic graphs and provide examples based on real circular and spherical data, vectorial or axial.

Keywords: directional statistics; V-statistics; grouped data; spectra of graphs; regular tessellations; Platonic graphs

MSC: 33C45

1. Introduction

The aim of this paper is to discuss some goodness-of-fit tests for uniformity of data from the sphere

$$S^2 = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1\},$$

or the unit circle

$$S^1 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}.$$

Our tests may be carried out after having grouped these data by means of a regular tessellation of S^2. These tests may take place during the exploratory phase of a statistical study, in order to detect departures from uniformity as clusters.

Grouped data occur in many ways in the fields of linear, directional or spatial statistics. First, grouping circular data into k intervals of equal lengths amounts to using the tessellation induced by a regular k-gon; see (Tables 1.1–1.4, [1]) for some examples. Time series consisting of counts indexed by days of the week, or months, are built in a similar way. Second, in the field of spatial statistics, points to be counted are in certain cases so numerous (insects, or trees in a tropical forest) that point patterns are recorded as local totals by the method of *quadrat counts*, see (p. 3, [2]). Third, a continuous spatial distribution can be seen as uniform, but only from a certain scale. For instance in the field of cosmology, the visible Universe is considered as homogeneous, but, according to (p. 1, [3]), only if we look out to distances larger than about 300 million light years. A popular model for such spatial data combines a *Poisson process* describing the distribution of parent objects, with a *cluster process* by which daughter objects are distributed about parent objects. Such models

Check for updates

Academic Editor: Tonghui Wang

Received: 7 April 2026
Revised: 12 May 2026
Accepted: 15 May 2026
Published: 19 May 2026

Copyright: © 2026 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

are used in geography, archaeology, and ecology, see (Chapter 8, [2]). Even if the parent objects are uniformly distributed, a test for uniformity will usually reject the hypothesis of a uniform distribution. Grouping data into "boxes" significantly larger than a typical cluster can help assess this type of uniformity in the exploratory phase of analysis.

On the sphere, the uniformity of a distribution is equivalent to its invariance under the action of the continuous group of isometries $O(3)$. Therefore, a natural way to discretize tests for continuous uniformity is to construct tests associated with finite subgroups of isometries. Some of these finite groups naturally arise from the use of particular coordinate systems, suited for directional data. Points of S^2 can be seen as unit vectors emanating from the origin, and represented by Cartesian coordinates (x, y, z), polar coordinates (θ, ϕ), geographical coordinates (L, ℓ), and geological coordinates $(\text{Dec.}, \text{Inc.})$, among others. Direction cosines (x, y, z), latitude L, longitudes ϕ and ℓ, colatitude θ, declination Dec. and inclination Inc. satisfy relations

$$\begin{cases} -1 \leq x, y, z \leq 1 \\ (\theta, \phi) \in [0, 180^\circ] \times [0, 360^\circ] \\ (L, \ell) \in [-90^\circ, 90^\circ] \times [0, 360^\circ] \\ (\text{Dec.}, \text{Inc.}) \in [0, 360^\circ] \times [-90^\circ, 90^\circ] \end{cases}, \begin{cases} \theta = 90^\circ - L = \text{Inc.} + 90^\circ \\ \phi = \ell = 360^\circ - \text{Dec.} \end{cases}, \tag{1}$$

with

$$\begin{cases} x = \sin\theta\cos\phi = \cos L \cos\ell = \cos(\text{Inc.})\cos(\text{Dec.}) \\ y = \sin\theta\sin\phi = \cos L \sin\ell = -\cos(\text{Inc.})\sin(\text{Dec.}) \\ z = \cos\theta = \sin L = -\sin(\text{Inc.}) \end{cases} \tag{2}$$

(see (§2.2, [4])). The two isometries $(x, y, z) \mapsto \pm(x, y, z)$ form a group of order 2, the orbit of the North Pole $(0, 0, 1)$ consisting of the antipodal pair $(0, 0, \pm 1)$. The Voronoi cells associated with these two points induce a regular tessellation of S^2 with two tiles, the Northern and Southern Hemispheres. If we extend this group by adding rotations as $(x, y, z) \mapsto (-y, x, z)$ or $(x, -z, y)$, the orbit of the North Pole becomes the vertex set of a regular octahedron, with elements $\pm(0, 0, 1)$, $\pm(0, 1, 0)$ and $\pm(1, 0, 0)$. The Voronoi cells associated with these six points induce the regular tessellation of S^2 by faces of a spherical cube. In polar coordinates, given a positive integer m, the orbit of the equatorial point $(\theta, \phi) = (\pi/2, 0)$ under the action of the cyclic group of isometries $\{(\theta, \phi) \mapsto (\theta, \phi + 2k\pi/m) : 0 \leq k \leq m - 1\}$ is a set of m points equally spaced around the equator.

To a regular tessellation of S^2, we can associate a graph $\Gamma = (V, E)$. The vertices $v \in V$ are the Voronoi cells (or as geometric points, their centroids). Two vertices are adjacent if the corresponding cells have an edge in common. We obtain graphs as the complete graph K_2 (two hemispheres), the circular graphs C_m for $m \geq 3$ (lunes with angles $2\pi/m$ at the North and South Poles), as well as the skeletons of the five Platonic solids: $T = K_4$ (tetrahedral graph), O (octahedral), C (hexahedral or cubical), I (icosahedral) and D (dodecahedral). In the terminology of (Chapter 5, [5]), all these graphs, with the exception of K_2, are called Platonic graphs.

From a sample of points $P_1, P_2, \ldots$ belonging to S^2, we obtain a sequence $X_1, X_2, \ldots$ of vertices of Γ, by setting $X_i = v \in V$ whenever $P_i \in v$. It is easily checked that if $P_1, P_2, \ldots$ is a sample drawn from the uniform spherical distribution, then $X_1, X_2, \ldots$ will be a sample from the discrete uniform distribution over $V(\Gamma)$. Therefore, if a goodness-of-fit test rejects the null hypothesis of uniformity for the latter, the same conclusion should hold for the former.

https://doi.org/10.3390/axioms15050379

All the graphs already mentioned are distance-transitive, but they belong to the much wider class of distance-regular graphs. Our main statistical result concerning these graphs is the following.

Theorem 1. *Let $\Gamma = (V, E)$ be a distance-regular graph, with diameter d, intersection array $\{b_0, \ldots, b_{d-1}; c_1, \ldots, c_d\}$, incidence matrices $A_0, A_1, \ldots, A_d$ and (combinatorial) Laplacian spectrum*

$$\mathbf{Sp}_C(\Gamma) = \begin{pmatrix} \lambda_0 & \lambda_1 & \cdots & \lambda_d \\ f_0 & f_1 & \cdots & f_d \end{pmatrix}. \tag{3}$$

(1) The Moore–Penrose inverse of the Laplacian $L := A_0 - b_0 A_1$ is the square matrix of order $|V|$ given by

$$G = \sum_{k=0}^{d} \tilde{g}(k) A_k, \text{ with } s_d \tilde{g} = s_{d-1}\, \tilde{g}_0 + \sum_{k=3}^{d+1} (-1)^k s_{d+1-k}\, \tilde{L}^{k-3} \tilde{e}, \tag{4}$$

where $s_0 = 1$, $s_1 = \lambda_1 + \cdots + \lambda_d, \ldots, s_d = \lambda_1 \cdots \lambda_d$ and

$$\tilde{g}_0 = \frac{1}{|V|} \begin{pmatrix} |V|-1 \\ -1 \\ \vdots \\ \vdots \\ -1 \end{pmatrix}, \quad \tilde{L} = \begin{pmatrix} b_0 & -b_0 & 0 & & & \\ -c_1 & b_1 + c_1 & -b_1 & & O & \\ 0 & -c_2 & \cdot & \cdot & & \\ & & \cdot & \cdot & \cdot & 0 \\ & O & & \cdot & \cdot & -b_{d-1} \\ & & & 0 & -c_d & c_d \end{pmatrix}, \quad \tilde{e} = \begin{pmatrix} b_0 \\ -c_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

(2) Assume that $X_1, X_2, \ldots$ is a sequence of independent random variables, following the discrete uniform distribution over the vertex-set V, i.e., $P(X_1 = v) = 1/|V|$ for every $v \in V$. The V-statistic defined by

$$V_n = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} G(X_i, X_j) \quad (n = 1, 2, \ldots) \tag{5}$$

satisfies the convergence in distribution

$$\lim_{n \to \infty} (|V| n V_n) = \sum_{j=1}^{d} \frac{\chi_j^2(f_j)}{\lambda_j} =: T_\Gamma, \tag{6}$$

where $(\chi_j^2(f_j))_{j \geq 1}$ is a sequence of independent chi-square random variables, with $f_1, f_2, \ldots$ degrees of freedom.

Before proceeding further, let us mention that the random variable on the right-hand side of (6) has a distribution function difficult to explicit, except in the exceptional cases when $d = 1$, or when all the f_j are even. However, for practitioners, a recent achievement is the R package "CompQuadForm", enabling us to compute it numerically in all cases. See a detailed example in Section 8.1 below. This fact is one of the motivations of our paper, which is organized as follows.

In Section 2, we recall basic facts and notations about distance-regular graphs, following the fundamental reference (Chapters 2, 4, [6]), in connection with our statistical goals. In Sections 3–6, we derive formulas for particular families of graphs. In Section 7, we state some properties of our statistic, and complete the proof of Theorem 1. In Section 8, we exemplify our results with real data. We use circular data from [1,7], and spherical data of Tables B1, B5 (vectorial) and B15 (axial) from [4].

To end this introduction, let us note that we consider distance-regular graphs as geometric objects enabling us to group data from a manifold in which it is embedded. Let us mention for potential future applications that distance-regular graphs can also be used as models in physics (see, e.g., [8–10]).

2. Distance Graphs and Their Centered Green Function

Throughout this paper, $\Gamma = (V, E)$ will denote a distance-regular graph, with vertex set $V = V(\Gamma)$, edge set $E = E(\Gamma)$, and diameter d. Recall that such a graph is connected, without loops or multiple edges, and determines an intersection array of the form

$$\iota(\Gamma) = \begin{Bmatrix} 0 & c_1 & \cdots & c_{d-1} & c_d \\ a_0 & a_1 & \cdots & a_{d-1} & a_d \\ b_0 & b_1 & \cdots & b_{d-1} & 0 \end{Bmatrix}. \tag{7}$$

For every column we have $a_i + b_i + c_i = b_0$, so that this array is usually written in the shorter form $\iota(\Gamma) = \{b_0, \ldots, b_{d-1}; c_1, \ldots, c_d\}$. The geometric meaning of these integers is the following. Given $(u, v) \in V \times V$, at distance $\delta(u, v) = i$ from each other, the integers

$$a_i = a_i(u, v) = \mathrm{Card}\{w : w \sim v, \delta(u, w) = i\}, \tag{8}$$

$$b_i = b_i(u, v) = \mathrm{Card}\{w : w \sim v, \delta(u, w) = i + 1\}, \tag{9}$$

$$c_i = c_i(u, v) = \mathrm{Card}\{w : w \sim v, \delta(u, w) = i - 1\}, \tag{10}$$

do not depend upon the particular pair (u, v), but only upon $i = \delta(u, v)$. In particular, each vertex has exactly b_0 neighbours. The Bose–Mesner algebra, say $\mathcal{A}_\Gamma$, is generated by the incidence matrices $A_0, \ldots, A_d$, symmetric of order $|V|$, with entries defined by $(A_i)_{u,v} = 1$ if $\delta(u, v) = i$, zero otherwise, for $u, v \in V$. In particular, A_0 is the identity matrix I, and A_1 is the ordinary adjacency matrix. Matrix A_1 has exactly $(d + 1)$ distinct eigenvalues $b_0 = \theta_0 > \theta_1 > \ldots > \theta_d$. The multiplicity of θ_j is denoted by f_j. We always have $f_0 = 1$, due to the connectedness of Γ. There exists a sequence of projection matrices $(E_i)_{0 \le i \le d}$ generating $\mathcal{A}_\Gamma$, called the minimal idempotents, satisfying relations

$$I = \sum_{j=0}^{d} E_j, \quad A_1 = \sum_{j=0}^{d} \theta_j E_j, \tag{11}$$

$$E_j^2 = E_j, \quad E_i E_j = E_j E_i = 0, \ \mathrm{Tr}\, E_j = f_j \quad (0 \le i \ne j \le d). \tag{12}$$

The image vector space of E_i is an eigenspace of A_j, for $0 \le i, j \le d$. The combinatorial Laplacian associated with Γ is the matrix

$$L = b_0 I - A_1 = \theta_0 \sum_{j=0}^{d} E_j - \sum_{j=0}^{d} \theta_j E_j = \sum_{j=1}^{d} \lambda_j E_j,$$

with eigenvalues $0 = \lambda_0 < \lambda_1 < \ldots < \lambda_d$ given by $\lambda_j = \theta_0 - \theta_j$.

The Laplacian can be seen as the operator mapping $f : V \to \mathbb{R}$ into Lf, defined by

$$Lf(v) = b_0 f(v) - \sum_{w \sim v} f(w) \quad (v \in V). \tag{13}$$

https://doi.org/10.3390/axioms15050379

Following [11,12], the ordinary spectrum and the Laplacian spectrum of Γ will be denoted by

$$\mathbf{Sp}_P(\Gamma) = \begin{pmatrix} \theta_0 & \theta_1 & \cdots & \theta_d \\ f_0 & f_1 & \cdots & f_d \end{pmatrix} = \theta_0^{f_0}\theta_1^{f_1}\cdots\theta_d^{f_d}, \tag{14}$$

$$\mathbf{Sp}_C(\Gamma) = \begin{pmatrix} \lambda_0 & \lambda_1 & \cdots & \lambda_d \\ f_0 & f_1 & \cdots & f_d \end{pmatrix} = \lambda_1^{f_1}\cdots\lambda_d^{f_d}, \tag{15}$$

respectively.

Remark 1. *In (Appendix, [12]), tables give (14), from which (15) is easily inferred. But it should be mentioned that in this reference, the numbers $\lambda_1, \lambda_2, \ldots$ denote eigenvalues belonging not to the Laplacian spectrum (15), but to the ordinary spectrum (14), each eigenvalue repeated with its multiplicity. For instance, the spectrum $(\lambda_1, \lambda_2, \lambda_3, \lambda_4) = (2, 0, 0, -2)$ given for the cycle graph C_4 (1.7 p. 272–273), becomes $\theta_0^{f_0}\theta_1^{f_1}\theta_2^{f_2} = 2^1 0^2 (-2)^1$ and $\lambda_1^{f_1}\lambda_2^{f_2} = 2^2 4^1$ in our notations.*

The minimal polynomial of A_1 is, by definition,

$$\Pi(X) := \prod_{j=0}^{d}(X - \theta_j),$$

but it can be computed directly from the intersection-array, via

Proposition 1. *The minimal polynomial satisfies $\Pi = w_{d+1} \in \mathbb{Z}[X]$, where $(w_i)_{0\le i\le d+1}$ is the sequence of polynomials defined recursively by*

$$\begin{aligned} w_0(x) &:= 1, \quad w_1(x) := x, \\ w_{i+1}(x) &= (x - a_i)w_i(x) - c_i b_{i-1} w_{i-1}(x) \quad (i = 0, \cdots, d). \end{aligned}$$

Proof. First it is clear, by induction, that w_i is a monic polynomial with degree i and integral coefficients. In particular, w_{d+1} has degree $d + 1$. Second, we know (from 4.4.1 Proposition, p. 129, [6]), that the roots of w_{d+1} are exactly $\theta_0, \cdots, \theta_d$, and the announced result is established. □

For the definition and basic properties of the Moore–Penrose inverse, see (§6.4, Definition 6.6 p. 287, [13]).

Proposition 2. *The Moore–Penrose inverse of L is given by*

$$G := \Pi'(b_0)^{-1}\left(\sum_{k=2}^{d+1}\frac{(-1)^k\Pi^{(k)}(b_0)}{k!}L^{k-2} - \frac{\Pi^{(2)}(b_0)}{2}E_0\right) = \sum_{j=1}^{d}\frac{1}{\lambda_j}E_j. \tag{16}$$

Proof. Let us first prove the second equality by using relation $G = GI = \sum_{j=0}^{d} GE_j$. On the first hand, from relations $(L^0 - E_0)E_0 = E_0 - E_0^2 = 0$ and $L^{k-2}E_0 = 0$ for $k \ge 3$, we obtain $GE_0 = 0$. On the second hand, consider a non-zero eigenvalue λ_j, for $1 \le j \le d$. As $\Pi(b_0) = 0$, Taylor formula at b_0 gives

$$\Pi(X) = \sum_{k=0}^{d+1}\frac{\Pi^{(k)}(b_0)}{k!}(X - b_0)^k = \sum_{k=1}^{d+1}\frac{\Pi^{(k)}(b_0)}{k!}(X - b_0)^k,$$

with $\Pi'(b_0) \neq 0$ because Π has only simple roots. Now, recall that $\Pi(\theta_j) = 0$ and $\theta_j - b_0 = -\lambda_j$. Therefore, this formula, written for $X = \theta_j$, gives

$$0 = \sum_{k=1}^{d+1} \frac{\Pi^{(k)}(b_0)}{k!} (-\lambda_j)^k \Rightarrow \lambda_j = \Pi'(b_0)^{-1} \sum_{k=2}^{d+1} \frac{(-1)^k \Pi^{(k)}(b_0)}{k!} \lambda_j^k.$$

This equality, combined with $L^{k-2}E_j = \lambda_j^{k-2}E_j$ and $E_0E_j = 0$, leads to

$$GE_j = \Pi'(b_0)^{-1} \left(\sum_{k=2}^{d+1} \frac{(-1)^k \Pi^{(k)}(b_0)}{k!} \lambda_j^{k-2} - 0 \right) E_j = \frac{1}{\lambda_j} E_j.$$

We have proved the second equality.

This equality, combined with relations (12), easily implies that $LGL = L$ and $GLG = G$. As G and L are real, symmetric, and commute, the two conditions $(LG)' = LG$ and $(GL)' = GL$ are also satisfied, and the result is established. □

From the preceding proposition we see that $G \in \mathcal{A}_\Gamma$; in other words, it is a linear combination of $A_0, A_1, \ldots, A_d$. It means that the value of $G(v, w)$ depends only upon $\delta(v, w)$, taking at most $(d + 1)$ values, far less than $\mathrm{Card}(V \times V) = |V|^2$. For example, all diagonal coefficients of G are equal. By evaluating the trace of both members in the equality proved in the preceding proposition, we obtain

$$G(v, v) = \frac{1}{|V|} \sum_{j=1}^{d} \frac{f_j}{\lambda_j} \quad (v \in V).$$

From the point of view of applications requiring explicit formulas, this point is crucial. Let us introduce a matrix of lower size, but containing all information required to compute entries of L^{k-2} for $k \geq 4$, appearing in the preceding proposition.

We call *radial Laplacian* associated with Γ the operator $\tilde{L}$ mapping a function $\phi : \mathbb{Z} \to \mathbb{R}$ with support in $[\![0, d]\!]$ to $\tilde{L}\phi$, with support in $[\![0, d]\!]$, and defined by

$$\tilde{L}\phi(i) = -b_i\phi(i + 1) + (b_i + c_i)\phi(i) - c_i\phi(i - 1) \quad (0 \leq i \leq d).$$

It can be viewed as the $(d + 1) \times (d + 1)$ tridiagonal matrix introduced in Theorem 1. The connection between the Laplacian and its radial part is given by

Proposition 3. *Assume $v_0 \in V$. Consider functions $f : V \to \mathbb{R}$, and $\tilde{f} : \mathbb{Z} \to \mathbb{R}$ with support in $[\![0, d]\!]$. We have*

$$f(v) = \tilde{f}(\delta(v, v_0)) \quad (v \in V) \Rightarrow Lf(v) = \tilde{L}\tilde{f}(\delta(v, v_0)) \quad (v \in V).$$

Proof. Assume $v \in V$ with $\delta(v_0, v) = i$. From definitions (8)–(10) of a_i, b_i and c_i, taken with $u = v_0$, relation (13) becomes

$$Lf(v) = b_0\tilde{f}(i) - a_i\tilde{f}(i) - b_i\tilde{f}(i + 1) - c_i\tilde{f}(i - 1),$$

and the result follows from equality $b_0 - a_i = b_i + c_i$. □

The following result will enable us to simplify the relation from Proposition 2.

Lemma 1. *Let $\tilde{e}$ be as in Theorem 1. We have, for $k = 1, 2, \ldots$*

$$[L^k](v, v_0) = [\tilde{L}^{k-1}\tilde{e}](\delta(v, v_0)) \quad (v_0, v \in V).$$

Proof. Let us proceed by induction. Initialization ($k = 1$): recall that $L = b_0 I - A_1$ where A_1 is the adjacency matrix, so that if $\delta(v, v_0) = 0$ we have $L(v, v_0) = b_0 - 0 = \tilde{e}(0)$; if $\delta(v, v_0) = 1$ we have $L(v, v_0) = 0 - 1 = \tilde{e}(1)$; if $\delta(v, v_0) = i \geq 2$, we have $L(v, v_0) = 0 - 0 = \tilde{e}(i)$. Heredity: with notations of the preceding proposition, let us suppose that for some $k \geq 1$ we have

$$f(v) := [L^k](v, v_0) = [\tilde{L}^{k-1}\tilde{e}(\delta(v, v_0)) =: \tilde{f}(\delta(v, v_0)).$$

We obtain

$$[L^{k+1}](v, v_0) = Lf(v) = \tilde{L}f(\delta(v, v_0)) = [\tilde{L}^k\tilde{e}](\delta(v, v_0)),$$

and the result is proved. □

Computation of the entries of the $|V| \times |V|$ matrix G reduces to computation of at most $(d + 1)$ different values, given by

Lemma 2. *We have* $G(v, v_0) = \tilde{g}(\delta(v, v_0))$ *with*

$$w'_{d+1}(b_0)\tilde{g}(i) := \frac{w^{(2)}_{d+1}(b_0)}{2}\tilde{g}_0(i) + \sum_{k=3}^{d+1} \frac{(-1)^k w^{(k)}_{d+1}(b_0)}{k!} [\tilde{L}^{k-3}\tilde{e}](i) \quad (i = 0, 1, \ldots, d), \tag{17}$$

where $\tilde{g}_0$ *and* $\tilde{e}$ *are as in Theorem 1.*

Proof. In virtue of Proposition 1 and of the first equality from Proposition 2, we have

$$w'_{d+1}(b_0)G = \Pi'(b_0)G = \frac{\Pi^{(2)}(b_0)}{2}(I - E_0) + \sum_{k=3}^{d+1} \frac{(-1)^k \Pi^{(k)}(b_0)}{k!} L^{k-2} \tag{18}$$

$$= \frac{w^{(2)}_{d+1}(b_0)}{2}(I - E_0) + \sum_{k=3}^{d+1} \frac{(-1)^k w^{(k)}_{d+1}(b_0)}{k!} L^{k-2}. \tag{19}$$

On the first hand, we know (from p. 2, [6]), that for every distance-regular graph, $|V|E_0$ is a matrix of ones. In particular, $E_0(v, v_0) = 1/|V|$, so that

$$I(v, v_0) - E_0(v, v_0) = \begin{cases} 1 - 1/|V| & \text{if } v = v_0, \\ -1/|V| & \text{if } v \neq v_0. \end{cases}$$

In all cases we have

$$I(v, v_0) - E_0(v, v_0) = g_0(\delta(v, v_0)). \tag{20}$$

On the other hand, we know from Lemma 1 that

$$\sum_{k=3}^{d+1} \frac{(-1)^k w^{(k)}_{d+1}(b_0)}{k!} [L^{k-2}](v, v_0) = \sum_{k=3}^{d+1} \frac{(-1)^k w^{(k)}_{d+1}(b_0)}{k!} [\tilde{L}^{k-3}\tilde{e}](\delta(v, v_0)). \tag{21}$$

From (19)–(21), we deduce

$$w'_{d+1}(b_0)G(v_0, v) = \frac{w^{(2)}_{d+1}(b_0)}{2}\tilde{g}_0(\delta(v_0, v)) + \sum_{k=3}^{d+1} \frac{(-1)^k w^{(k)}_{d+1}(b_0)}{k!} [\tilde{L}^{k-3}\tilde{e}](\delta(v_0, v)),$$

and our result is established. □

https://doi.org/10.3390/axioms15050379

Before proceeding further, let us notice that we can now write

$$G = \sum_{i=0}^{d} \tilde{g}(i) A_i. \tag{22}$$

In the case when the Laplacian eigenvalues are available and have simple expressions, a useful complement to the preceding result is the following.

Lemma 3. *If the Laplacian spectrum is given by* (15)*, relation* (17) *is equivalent to*

$$s_d \tilde{g}(i) = s_{d-1}\, \tilde{g}_0(i) + \sum_{k=3}^{d+1} (-1)^k s_{d+1-k}\, [\tilde{L}^{k-3}\tilde{e}](i) \quad (0 \le i \le d), \tag{23}$$

where the symmetric sums

$$s_0 = 1, \quad s_1 = \lambda_1 + \cdots + \lambda_d, \quad \cdots \quad , s_d = \lambda_1 \cdots \lambda_d$$

are defined via one of the two equivalent relations

$$\prod_{j=1}^{d} (X - \lambda_j) = (-1)^d s_d + \cdots - s_1 X^{d-1} + s_0 X^d, \tag{24}$$

$$w_{d+1}(b_0 + X) = s_d X + s_{d-1} X^2 + \cdots + s_1 X^d + s_0 X^{d+1}. \tag{25}$$

Proof. Let us first prove the equivalence between (24) and (25). The equivalence we have to prove can be rewritten as

$$w_{d+1}(b_0 + X) = (-1)^d X \prod_{j=1}^{d} (-X - \lambda_j) = X \prod_{j=1}^{d} (X + \lambda_j).$$

Recalling that $\lambda_j = b_0 - \theta_j = \theta_0 - \theta_j$, it is equivalent to

$$w_{d+1}(X) = (X - \theta_0) \prod_{j=1}^{d} (X - \theta_0 + \lambda_j) = (X - \theta_0) \prod_{j=1}^{d} (X - \theta_j),$$

which is known to be true from Proposition 1. In view of $w_{d+1}(b_0) = 0$, Taylor formula reads $w_{d+1}(b_0 + X) = \sum_{k=1}^{d} w_{d+1}^{(k)}(b_0) X^k / k!$, and we obtain, by identification, $w_{d+1}^{(k)}(b_0)/k! = s_{d-k+1}$ for $k = 1, \ldots, d$. Relation (23) then follows from (17). □

In view of these results, we can affirm that assertion (1) in Theorem 1 is established.

3. Complete Graphs and Chi-Square Tests

We know (from p. 17 [11] and 7.5.2 Theorem [6]) that the complete graph over $|V| = m$ vertices, denoted by K_m, is distance-regular with

$$\iota(K_m) = \begin{Bmatrix} 0 & c_1 \\ a_0 & a_1 \\ b_0 & 0 \end{Bmatrix} = \begin{Bmatrix} 0 & 1 \\ 0 & m-2 \\ m-1 & 0 \end{Bmatrix}, \quad \mathbf{Sp}_C(K_m) = \begin{pmatrix} \lambda_0 & \lambda_1 \\ f_0 & f_1 \end{pmatrix} = \begin{pmatrix} 0 & m \\ 1 & m-1 \end{pmatrix} \tag{26}$$

The symmetric sum is $s_1 = m$ so that relation (23) becomes

$$m\tilde{g}_{K_m} = \begin{pmatrix} \frac{m-1}{m} \\ -\frac{1}{m} \end{pmatrix}$$

Therefore, our statistic satisfies, with obvious notations,

$$mnV_n = \frac{1}{mn}\left((m-1)\sum_{i=1}^{m} n_i^2 - \sum_{1\le i\ne j\le m} n_i n_j\right) = \frac{1}{m}\sum_{i=1}^{m}\frac{(n_i - \frac{n}{m})^2}{\frac{n}{m}} \to \frac{\chi_1^2(m-1)}{m} \tag{27}$$

and our test reduces to a chi-square test. It is well known that in a chi-square test, it is customary to use the asymptotic distribution under the null hypothesis as a good approximation for that of the statistic. As a working hypothesis, we will use the same approximation in all further cases.

It is worth noting at this stage that it would be useless to introduce a sophisticated mathematical apparatus borrowed from graph theory, only to reintroduce already well-known goodness-of-fit procedures, such as the chi-square test. But firstly, note that in view of (6), a chi-square limiting distribution, as in (27), will occur under the null hypothesis of uniformity if and only if the diameter of Γ is $d = 1$. In other words, it will happen only in the case of a complete graph. Secondly, recall that a core result from the theory of graphs is that their spectra give insights into some of their geometric features as cycles, girth, connectedness and so on. Our aim is to use tools from algebraic graph theory in order to take into account such features in the analysis of data. The subsequent examples will illustrate this aspect.

4. Cyclic Graphs

For Table 1, see (p. 17, [11]), (p. 1, [6]) and graphs 1.7 and 1.27 p. 272–275, 4.1 p. 306–307 from [12]. Constant $\phi = (1+\sqrt{5})/2$ is the golden ratio, and we will use repeatedly identities derived from the fundamental relation $\phi^2 = 1 + \phi$.

In view of Table 1, Formula (23) gives

$$8\tilde{g}_{C_4} = 6\begin{pmatrix}\frac{3}{4}\\ -\frac{1}{4}\\ -\frac{1}{4}\end{pmatrix} - \begin{pmatrix}2\\ -1\\ 0\end{pmatrix} \Rightarrow \tilde{g}_{C_4} = \frac{1}{16}\begin{pmatrix}5\\ -1\\ -3\end{pmatrix}, \tag{28}$$

$$5\tilde{g}_{C_5} = 5\begin{pmatrix}\frac{4}{5}\\ -\frac{1}{5}\\ -\frac{1}{5}\end{pmatrix} - \begin{pmatrix}2\\ -1\\ 0\end{pmatrix} \Rightarrow \tilde{g}_{C_5} = \frac{1}{5}\begin{pmatrix}2\\ 0\\ -1\end{pmatrix}, \tag{29}$$

and for C_6 we find

$$\tilde{L} = \begin{pmatrix}2 & -2 & 0 & 0\\ -1 & 2 & -1 & 0\\ 0 & -1 & 2 & -1\\ 0 & 0 & -2 & -2\end{pmatrix}, \quad \tilde{e} = \begin{pmatrix}2\\ -1\\ 0\\ 0\end{pmatrix}, \quad \tilde{L}\,\tilde{e} = \begin{pmatrix}6\\ -4\\ 1\\ 0\end{pmatrix}$$

so that

$$12\tilde{g}_{C_6} = 19\begin{pmatrix}\frac{5}{6}\\ -\frac{1}{6}\\ -\frac{1}{6}\\ -\frac{1}{6}\end{pmatrix} - 8\begin{pmatrix}2\\ -1\\ 0\\ 0\end{pmatrix} + \begin{pmatrix}6\\ -4\\ 1\\ 0\end{pmatrix} \Rightarrow \tilde{g}_{C_6} = \frac{1}{72}\begin{pmatrix}35\\ 5\\ -13\\ -19\end{pmatrix}$$

Table 1. Cycle graphs.

	C_4	C_5	C_6
$\iota(\Gamma)$	$\{2,1;1,2\}$	$\{2,1;1,1\}$	$\{2,1,1;1,1,2\}$
$\theta_0^{f_0}\cdots\theta_d^{f_d}$	$2^1 0^2 (-2)^1$	$2^1(1/\phi)^2(-\phi)^2$	$2^1 1^2 (-1)^2 (-2)^1$
$\lambda_1^{f_1}\cdots\lambda_d^{f_d}$	$2^2 4^1$	$(3-\phi)^2(2+\phi)^2$	$1^2 3^2 4^1$
$s_1,\ldots,s_d$	$6,8$	$5,5$	$8,19,12$

Example 1. *Let us illustrate computations related to our statistic with the case* $\Gamma = C_4$. *Assume that the vertex-set is given by* $V(C_4) = \{v_1, v_2, v_3, v_4\}$, *with* $v_1 \sim v_2 \sim v_3 \sim v_4 \sim v_1$. *In this setting, the incidence matrices generating the Bose–Mesner algebra are*

$$A_0 = I = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad A_1 = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

Relation (22) *and values* (28) *yield*

$$16G = 5A_0 - A_1 - 3A_2 = \begin{pmatrix} 5 & -1 & -3 & -1 \\ -1 & 5 & -1 & -3 \\ -3 & -1 & 5 & -1 \\ -1 & -3 & -1 & 5 \end{pmatrix}.$$

Let n_i *denote the number of observations equal to* v_i *in a sample of size* $n = \sum_{i=1}^4 n_i$. *This sample, denoted by* S, *can be seen as an element of* V^n. *We might compute* V_n *via the matrix formula*

$$V_n = n^{-2} d_n' G d_n = \frac{1}{16n^2} (n_1, n_2, n_3, n_4) \begin{pmatrix} 5 & -1 & -3 & -1 \\ -1 & 5 & -1 & -3 \\ -3 & -1 & 5 & -1 \\ -1 & -3 & -1 & 5 \end{pmatrix} \begin{pmatrix} n_1 \\ n_2 \\ n_3 \\ n_4 \end{pmatrix},$$

where $d_n = (n_1, n_2, n_3, n_4)'$. *But matrix calculus is time- and space-consuming in the general case. Let us, therefore, describe an approach suited for larger values of* n. *Let us partition* $V \times V$ *into the three subsets* $R_i = \{(v,w) \in V \times V : \delta(v,w) = i\}$, $i = 0, 1, 2$. *They are called the classes of the association scheme and are given by*

$$\begin{aligned} R_0 &= \{(v_1, v_1); (v_2, v_2); (v_3; v_3); (v_4, v_4)\}, \\ R_1 &= \{(v_1, v_2); (v_2, v_1); (v_2; v_3); (v_3, v_2); (v_3; v_4); (v_4, v_3)(v_4; v_1); (v_1, v_4)\}, \\ R_2 &= \{(v_1, v_3); (v_3, v_1); (v_2; v_4); (v_4, v_2)\} \end{aligned}$$

(see (§2.1 p. 44, [6])). We have $(v,w) \in R_{\delta(v,w)}$. *Therefore,* $G(v,w) = \tilde{g}(i)$ *if and only if* $(v,w) \in R_i$. *The number of times the ordered pair* (v_i, v_j) *occurs in* $S \times S$ *is* $n_i n_j$. *We can reduce the number of summands in* (5) *by using relation* $G(v_i, v_j) + G(v_j, v_i) = 2\tilde{g}(\delta(v_i, v_j))$ *whenever* $i \neq j$. *Thus, our statistic is given by*

$$\begin{aligned} n^2 V_n = \tilde{g}(0)(n_1^2 + n_2^2 + n_3^2 + n_4^2) + 2\tilde{g}(1)(n_1 n_2 + n_2 n_3 + n_3 n_4 + n_4 n_1) \\ + 2\tilde{g}(2)(n_1 n_3 + n_2 n_4). \end{aligned} \quad (30)$$

In view of (28), *this expression becomes*

$$16V_n = \frac{5(n_1^2+n_2^2+n_3^2+n_4^2)-2(n_1n_2+n_2n_3+n_3n_4+n_4n_1)-6(n_1n_3+n_2n_4)}{n^2} \quad (31)$$

$$= \frac{[(n_1-n_2)^2+(n_2-n_3)^2+(n_3-n_4)^2+(n_4-n_1)^2]+3[(n_1-n_3)^2+(n_2-n_4)^2]}{n^2}. \quad (32)$$

It is evidently essentially different from the chi-square statistic given by

$$\chi^2 = \frac{(4n_1-n)^2+(4n_2-n)^2+(4n_3-n)^2+(4n_4-n)^2}{4n}.$$

Let us give a geometric interpretation of their dissimilarity. The isomorphism group of C_4 *contains the* 8 *elements of the dihedral group* D_4, *generated by* $(v_1,v_2,v_3,v_4) \mapsto (v_2,v_3,v_4,v_1)$ *and* $(v_1,v_2,v_3,v_4) \mapsto (v_3,v_2,v_1,v_4)$. *Transposition* $(v_1,v_2,v_3,v_4) \mapsto (v_2,v_1,v_3,v_4)$ *does not belong to this group. Our statistic* V_n *enjoys the same dihedral symmetries, as, for example,*

$$V_n(n_1,n_2,n_3,n_4) = V_n(n_2,n_3,n_4,n_1) = V_n(n_3,n_2,n_1,n_4) \neq V_n(n_2,n_1,n_3,n_4).$$

On the contrary, the chi-square statistic is invariant under the 24 *permutations of the four vertices. This omnibus test does not take into account the geometry of* C_4, *whereas our new test based upon* V_n *does.*

5. The Petersen and Dodecahedral Graphs

For the Petersen graph P and the dodecahedral graph D, we obtain Table 2 by using (p. 1, [6]), (3B p. 20, [11]), graphs 3.26 p. 294–295 and 4.20 p. 308–309 in [12].

Table 2. Petersen and dodecahedral graphs.

	Petersen	**Dodecahedron**
$\iota(\Gamma)$	$\{3,2;1,1\}$	$\{3,2,1,1,1;1,1,1,2,3\}$
$\theta_0^{f_0}\cdots\theta_d^{f_d}$	$3^1 1^5 (-2)^4$	$3^1(\sqrt{5})^3 1^5 0^4 (-2)^4 (-\sqrt{5})^3$
$\lambda_1^{f_1}\cdots\lambda_d^{f_d}$	$2^5 5^4$	$(3-\sqrt{5})^3 2^5 3^4 5^4 (3+\sqrt{5})^3$
$s_1,\ldots,s_d$	7,10	16, 95, 256, 304, 120

The dodecahedral graph D is drawn in Figure 1, and Petersen's graph P is drawn in Figure 2. D is a planar graph that can be represented without line-crossings. The numerous line-crossings occurring when P is drawn in the plane hide the fact that D and P are topologically closely related. We can partition the vertex set of D into 10 antipodal pairs $(\{v_i, v_i'\})_{1\leq i\leq 10}$. Each class $\{v_i, v_i'\}$ can be seen as a vertex (simply denoted by v_i) of Petersen's graph, with $\{v_i, v_i'\} \sim \{v_j, v_j'\}$ if $v_i \sim v_j$ or $v_i \sim v_j'$. In other words, D is an antipodal 2-cover of P (see (4.2.8 Proposition p. 144, [6])). In Section 8.3, we will see that D and P are closely related to the icosahedral graph I, with the three edges emanating from v_i or v_i' being identified with three adjacent faces of I. The fact that the dodecahedron and the icosahedron are two dual Platonic solids lies, of course, at the heart of all these connections between D, I and P.

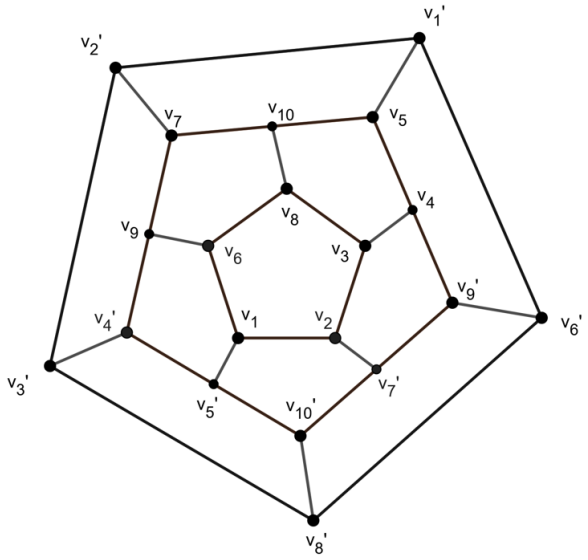


Figure 1. The dodecahedral graph D. Vertices $\mathrm{v_i}$ and $\mathrm{v_i'}$ form antipodal pairs for $1 \leq i \leq 10$.

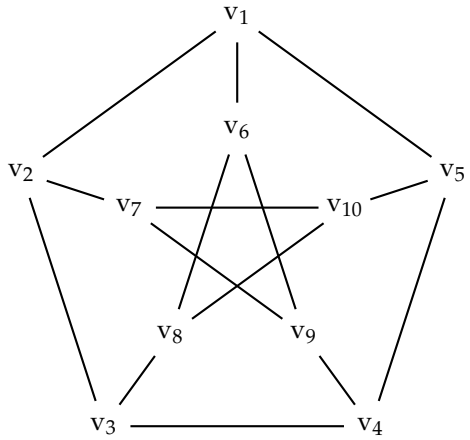


Figure 2. Petersen's graph P as a quotient graph of D. Vertex $\mathrm{v_i}$ from P is identified with the antipodal pair $\{\mathrm{v_i}, \mathrm{v_i'}\}$ from D.

Let us now compute all functions associated with these two graphs. For Petersen's graph, we easily get

$$\tilde{g}_{P,0} = \begin{pmatrix} \frac{9}{10} \\ -\frac{1}{10} \\ -\frac{1}{10} \end{pmatrix}, \quad \tilde{e}_P = \begin{pmatrix} 3 \\ -1 \\ 0 \end{pmatrix},$$

so that

$$10\tilde{g}_P = 7\begin{pmatrix} \frac{9}{10} \\ -\frac{1}{10} \\ -\frac{1}{10} \end{pmatrix} - \begin{pmatrix} 3 \\ -1 \\ 0 \end{pmatrix} = \frac{1}{10}\begin{pmatrix} 33 \\ 3 \\ -7 \end{pmatrix}$$

And for the dodecahedron, we find

$$\tilde{g}_{D,0} = \begin{pmatrix} \frac{19}{20} \\ -\frac{1}{20} \\ -\frac{1}{10} \\ -\frac{1}{20} \\ -\frac{1}{10} \\ -\frac{1}{20} \end{pmatrix}, \quad \tilde{e}_D = \begin{pmatrix} 3 \\ -1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \tilde{L} = \begin{pmatrix} 3 & -3 & 0 & 0 & 0 & 0 \\ -1 & 3 & -2 & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 \\ 0 & 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & 0 & -2 & 3 & -1 \\ 0 & 0 & 0 & 0 & -3 & 3 \end{pmatrix}$$

and

$$\tilde{L}\,\tilde{e}_D = \begin{pmatrix} 12 \\ -6 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \tilde{L}^2\,\tilde{e}_D = \begin{pmatrix} 54 \\ -32 \\ 8 \\ -1 \\ 0 \\ 0 \end{pmatrix} \quad \tilde{L}^3\,\tilde{e}_D = \begin{pmatrix} 258 \\ -166 \\ 49 \\ -10 \\ 2 \\ 0 \end{pmatrix}$$

From these values, we infer

$$120\tilde{g}_D = 304 \begin{pmatrix} \frac{19}{20} \\ -\frac{1}{20} \\ -\frac{1}{10} \\ -\frac{1}{20} \\ -\frac{1}{10} \\ -\frac{1}{20} \end{pmatrix} - 256 \begin{pmatrix} 3 \\ -1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + 95 \begin{pmatrix} 12 \\ -6 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} - 16 \begin{pmatrix} 54 \\ -32 \\ 8 \\ -1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 258 \\ -166 \\ 49 \\ -10 \\ 2 \\ 0 \end{pmatrix} = \frac{1}{5} \begin{pmatrix} 274 \\ 84 \\ 4 \\ -46 \\ -66 \\ -76 \end{pmatrix}.$$

Proposition 4. *The pseudo-inverse of the Laplacian is given, for Petersen's and the dodecahedral graphs, by*

$$G_\Gamma(v,w) = \tilde{g}_\Gamma(\delta(v,w)), \quad \textit{with} \tag{33}$$

$$\tilde{g}_P = \frac{1}{100}\begin{pmatrix} 33 \\ 3 \\ -7 \end{pmatrix}, \quad \tilde{g}_D = \frac{1}{300}\begin{pmatrix} 137 \\ 42 \\ 2 \\ -23 \\ -33 \\ -38 \end{pmatrix}, \tag{34}$$

respectively.

6. The Octahedral, Cubical and Icosahedral Graphs

For Table 3, see (p. 1, [6]), and graphs 3.7, 4.37, 4.49 p. 292, 310–311 in [12].

Table 3. Graphs of three regular polyhedra.

	Octahedron	**Cube**	**Icosahedron**
$\iota(\Gamma)$	$\{4,1;1,4\}$	$\{3,2,1;1,2,3\}$	$\{5,2,1;1,2,5\}$
$\theta_0^{f_0}\cdots\theta_d^{f_d}$	$4^1 0^3 (-2)^2$	$3^1 1^3 (-1)^3 (-3)^1$	$5^1\sqrt{5}^3(-1)^5(-\sqrt{5})^3$
$\lambda_1^{f_1}\cdots\lambda_d^{f_d}$	$4^3 6^2$	$2^3 4^3 6^1$	$(5-\sqrt{5})^5 6^5$ $(5+\sqrt{5})^5$
$s_1,\ldots,s_d$	$10, 24$	$12, 44, 48$	$16, 80, 120$

First, we have

$$\tilde{g}_{O,0} = \frac{1}{6}\begin{pmatrix} 5 \\ -1 \end{pmatrix}, \quad \tilde{g}_{C,0} = \frac{1}{8}\begin{pmatrix} 7 \\ -1 \\ -1 \end{pmatrix}, \quad \tilde{g}_{I,0} = \frac{1}{12}\begin{pmatrix} 11 \\ -1 \\ -1 \end{pmatrix}$$

Then,

$$\tilde{e}_0 = \begin{pmatrix} 4 \\ -1 \\ 0 \end{pmatrix} \tag{35}$$

$$\tilde{L}_C = \begin{pmatrix} 3 & -3 & 0 & 0 \\ -1 & 3 & -2 & 0 \\ 0 & -2 & 3 & -1 \\ 0 & 0 & -3 & 3 \end{pmatrix}, \quad \tilde{e}_C = \begin{pmatrix} 3 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \quad \tilde{L}_C\,\tilde{e}_C = \begin{pmatrix} 12 \\ -6 \\ 2 \\ 0 \end{pmatrix} \tag{36}$$

$$\tilde{L}_I = \begin{pmatrix} 5 & -5 & 0 & 0 \\ -1 & 3 & -2 & 0 \\ 0 & -2 & 3 & -1 \\ 0 & 0 & -5 & 5 \end{pmatrix}, \quad \tilde{e}_I = \begin{pmatrix} 5 \\ -1 \\ 0 \\ 0 \end{pmatrix} \quad \tilde{L}_I\,\tilde{e}_I = \begin{pmatrix} 30 \\ -8 \\ 2 \\ 0 \end{pmatrix}. \tag{37}$$

From the values of the symmetric functions $s_1, s_2 \ldots$ we get, from (23),

$$24\tilde{g}_O = 10\begin{pmatrix} \frac{5}{6} \\ -\frac{1}{6} \\ -\frac{1}{6} \end{pmatrix} - \begin{pmatrix} 4 \\ -1 \\ 0 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 13 \\ -2 \\ -5 \end{pmatrix}$$

$$48\tilde{g}_C = 44\begin{pmatrix} \frac{7}{8} \\ -\frac{1}{8} \\ -\frac{1}{8} \\ -\frac{1}{8} \end{pmatrix} - 12\begin{pmatrix} 3 \\ -1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 12 \\ -6 \\ 2 \\ 0 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 29 \\ 1 \\ -7 \\ -11 \end{pmatrix}$$

$$120\tilde{g}_I = 80\begin{pmatrix} \frac{11}{12} \\ -\frac{1}{12} \\ -\frac{1}{12} \\ -\frac{1}{12} \end{pmatrix} - 16\begin{pmatrix} 5 \\ -1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 30 \\ -8 \\ 2 \\ 0 \end{pmatrix} = \frac{2}{3}\begin{pmatrix} 35 \\ 2 \\ -7 \\ -10 \end{pmatrix}$$

so that we have proved

Proposition 5. *The pseudo-inverse of the Laplacian is given, for the octahedral, cubical, and icosahedral graphs, by*

$$G_\Gamma(v,w) = \tilde{g}_\Gamma(\delta(v,w)), \quad \textit{with} \tag{38}$$

$$\tilde{g}_O = \frac{1}{72}\begin{pmatrix} 13 \\ -2 \\ -5 \end{pmatrix} \quad \tilde{g}_C = \frac{1}{96}\begin{pmatrix} 29 \\ 1 \\ -7 \\ -11 \end{pmatrix}, \quad \tilde{g}_I = \frac{1}{180}\begin{pmatrix} 35 \\ 2 \\ -7 \\ -10 \end{pmatrix}, \tag{39}$$

respectively.

7. Statistical Meaning of Some Elements in $\mathcal{A}_\Gamma$

Thus far, we have used properties of the Bose–Mesner algebra only as handy computational tools. Let us now explain in which way algebraic identities can be translated into the language of statistics.

7.1. The Values of G as Hitting Times

In all examples above it can be noticed that

$$\tilde{g}(0) > \tilde{g}(1), \cdots, \tilde{g}(d). \tag{40}$$

In other words, $G(v,w)$ reaches its maximal value if and only if $v = w$. This inequality is essential to determine the maximum values of V_n, because it entails

$$V_n = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} G(X_i, X_j) \leq \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} \tilde{g}(0) = \tilde{g}(0),$$

with equality if and only if $X_1 = \ldots = X_n$. In order to get a result about these maxima, let us prove that (40) holds for all distance-regular graphs.

A method of proof based on random walks will provide us with interesting features of our function G. Consider a random walk defined by Γ, i.e., a discrete time Markov chain $\{X_0, X_1, \ldots\}$ with state space $V(\Gamma)$, and transition probability matrix $P_1 = A_1/b_0$. The stationary distribution is the uniform distribution π_u which satisfies relations

$$\pi_u := \frac{1}{|V|}\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}, \quad P_1\pi_u = \pi_u.$$

We will use the standard notation $P(X_k = w | X_0 = v) = P_v(X_k = w)$.

Let us show that G is closely connected to the positive stopping times

$$T(w) = \inf\{n \geq 1 : X_n = w\}.$$

A corollary of the following result will be (40).

Proposition 6. *The entries of G satisfy the identities*

$$|V|b_0[G(v,v) - G(v,w)] = E_v T(w) = E_w T(v) \geq 1 \quad (v \neq w \in V). \tag{41}$$

Proof. From $GL = LG = I - E_0$, $GE_0 = E_0 L = 0$ and $E_0^2 = E_0$, we infer that

$$(b_0 G + E_0)(b_0^{-1} L + E_0) = I - E_0 + 0 + 0 + E_0 = I.$$

Consequently, as $L/b_0 = I - P_1$, we have

$$F := b_0 G + E_0 = (I - (P_1 - E_0))^{-1},$$

which in the terminology of [14] is the fundamental matrix denoted by Z; see (6.1) p. 226. In the same reference, relation (6.14) reads

$$E_w[T(v)] = \frac{F(v,v) - F(w,v)}{\pi_u(v)} = |V|b_0[G(v,v) - G(w,v)] \quad (v \neq w),$$

and the result follows, by using the symmetry of G. □

7.2. Principal Components

We summarize the dataset corresponding to our sample $X_1, \ldots, X_n \in V$ by introducing the column-vector

$$D(n) := (n_v)_{v \in V}, \quad \text{with} \quad n_v = \sum_{i=1}^{n} \mathbb{1}_{\{X_i = v\}} \quad (v \in V).$$

Decomposition $I = \sum_{j=0}^{d} E_j$, applied to our data, gives

$$D(n) = \sum_{j=0}^{d} D_j(n) \quad \text{with} \quad D_j(n) = E_j D(n).$$

In virtue of Proposition 2, our statistic can be written in the form

$$n^2 V_n = D(n)' G D(n) = \sum_{j=1}^{n} \frac{D_j(n)' D_j(n)}{\lambda_j}.$$

For $j = 1, \ldots, n$, $D_j(n)$ can be interpreted as the j-th principal component, measuring one of the d components of the departure of $D(n)$ from the uniform component. The latter is the same for all n-samples, given by

$$D_0(n) = \begin{pmatrix} n/|V| \\ \vdots \\ n/|V| \end{pmatrix} = n\pi_u \quad (n = 1, 2, \ldots).$$

We have $D_j(n)' D_j(n) \geq 0$ for $j = 1, \ldots, d$, implying that $V_n \geq 0$. The set of equalities $D_j(n)' D_j(n) = 0$ for $j = 1, \ldots, d$ will take place if and only if $D_j(n) = 0$ for $j = 1, \ldots, d$, which can happen only if $D(n) = D_0(n) = E_0 D(n) = n\pi_u$, meaning that the sample distribution is uniform. Note that this can occur only in cases when the sample size n is a multiple of the order of the graph $|V|$.

7.3. Extreme Values of V_n

The results of the two preceding subsections allow us to state

Proposition 7. *For the sample $X_1, \ldots, X_n$, the inequalities*

$$0 \leq V_n \leq \tilde{g}(0) = \frac{1}{|V|} \sum_{j=1}^{d} \frac{f_j}{\lambda_j}$$

hold. Furthermore, $V_n = 0$ if only if $D(n) = n\pi_u$, and $V_n = \tilde{g}(0)$ if only if $X_1 = \ldots = X_n$.

7.4. Asymptotic Distribution

Given v and w two distinct vertices, we have

$$E(\mathbb{1}^2_{\{X_i = v\}}) = E(\mathbb{1}_{\{X_i = v\}}) = \frac{1}{|V|}, \quad E(\mathbb{1}_{\{X_i = v\}} \mathbb{1}_{\{X_i = w\}}) = E(0) = 0.$$

https://doi.org/10.3390/axioms15050379

In other words, the column-vector $Y_i = \left(\mathbb{1}_{\{X_i=v\}}\right)_{v\in V}$ is a random vector with mean vector π_u and variance–covariance matrix $I/|V|$. Furthermore, we have $D(n) = \sum_{i=1}^{n} Y_i$. The multivariate central limit theorem tells us that we have the convergence in distribution

$$\sqrt{\frac{|V|}{n}}\sum_{i=1}^{n}(Y_i - \pi_u) = \sqrt{\frac{|V|}{n}}(D(n) - n\pi_u) \to Z,$$

where Z has a multivariate normal distribution with mean 0 and variance–covariance matrix the identity (see, e.g., (Theorem 3.12 p. 147, [15])). In view of the equality $G\pi_u = 0$, this means that our statistic satisfies

$$n|V|V_n = \frac{|V|}{n}D(n)'GD(n) \tag{42}$$

$$= \left[\sqrt{\frac{|V|}{n}}(D(n) - n\pi_u)\right]' G \left[\sqrt{\frac{|V|}{n}}(D(n) - n\pi_u)\right] \tag{43}$$

$$\to Z'GZ = \sum_{j=1}^{d}\frac{Z'E_jZ}{\lambda_j}. \tag{44}$$

Next, decomposition $I = \sum_{j=0}^{d} E_j$ is the assumption in Cochran's theorem, the conclusion being that if Z is a centered normal random vector with covariance matrix I, then the random variables $Z'E_jZ$, $j = 0, 1, \ldots, d$, are independent and satisfy, in law, $Z'E_jZ = \chi_j^2(f_j)$. Thus, assertion (2) in Theorem 1 is proved.

8. Examples with Real Data

The aim of our examples is twofold. First, we want to show how our previous theoretical results can be applied to real data. Second, our aim is to check whether the conclusions based on our tests agree with those of some standard tests.

For the sake of simplicity, we will follow the rule indicated by (p. 32, [16]): if our sample size is n, we shall consider $|V| \approx \sqrt{n}$ as the optimal order of the graph we should choose to group data.

8.1. Circular Data

The first example (Example 1.1 p. 2) of real circular data in the seminal book [7] provides us with a sample consisting of the 9 angles at which a roulette wheel stopped: $43°, 45°, 52°, 61°, 75°, 88°, 88°, 279°$, and $357°$. As $\sqrt{9} = 3$, we choose to use the cycle graph C_3, with vertices equivalent to classes $[0°, 120°)$, $[120°, 240°)$ and $[240°, 360°)$. Respective frequencies are 7, 0 and 2. Recalling that our test associated with C_3 is a chi-square test, we obtain the approximate p-value

$$P(\chi^2(2) \geq \frac{(7-3)^2 + (0-3)^2 + (2-3)^2}{3}) = 1.3\%,$$

and we reject uniformity at level 5%. This p-value is in agreement with the p-value between 1% and 5% obtained with the range test applied to the same data; see Example 6.8 p. 107 in [1] (noting that $188.1°$ should be read instead of $88.1°$ for the 5% value, see Appendix 2.8 p. 367).

In Example 6.7 (p. 106, [1]), the authors report a pigeon-homing experiment providing a sample of circular measurements consisting of the 15 angles $115°, 120°, 120°, 130°, 135°, 140°, 150°, 150°, 150°, 165°, 185°, 210°, 235°, 270°, 345°$. As $\sqrt{15} \approx 4$, we choose to use the cycle graph C_4 (see Section 4 above), with vertices represented by intervals $v_1 = [0°, 90°)$, $v_2 = [90°, 180°)$, $v_3 = [180°, 270°)$ and $v_4 = [270°, 360°)$. We have the

adjacency relations $v_1 \sim v_2 \sim v_3 \sim v_4 \sim v_1$; antipodal pairs are (v_1, v_3) and (v_2, v_4); frequencies are $n_1 = 0$, $n_2 = 10$, $n_3 = 3$ and $n_4 = 2$. Let us use (28) or Example 1 above. We obtain for our statistic the equalities

$$V_{15} = \frac{5\sum_{i=1}^{4} n_i^2 - (2n_1n_2 + 2n_2n_3 + 2n_3n_4 + 2n_4n_1) - 3(2n_1n_3 + 2n_2n_4)}{16 \times 15^2} \quad (45)$$

$$= \frac{5 \times 113 - 72 - 3 \times 40}{16 \times 15^2} = \frac{373}{16 \times 15^2}, \quad (46)$$

$$4 \times 15 \times V_{15} = \frac{373}{4 \times 15} \approx 6.22 \quad (47)$$

The Laplacian spectrum of C_4 is $\lambda_1^{f_1}\lambda_2^{f_2} = 2^2 4^1$. Therefore, the limiting distribution of our statistic is that of the random variable

$$T_{C_4} = \frac{\chi_1^2(2)}{2} + \frac{\chi_1^2(1)}{4}.$$

Let us use the R package mentioned in the Introduction:

```
> install.packages("CompQuadForm");
>library("CompQuadForm");
> davies(6.22, c(1/2, 1/4), c(2,1));
 $Qq
[1] 0.002814588
```

This output means that

$$P(\frac{\chi_1^2(2)}{2} + \frac{\chi_1^2(1)}{4} \geq 6.22) \approx 0.28\%,$$

and we reject uniformity at the level 5%. This p-value is in agreement with the p-value close to 1% obtained with the Hodges–Ajne test; see Example 6.7 (p. 106, [1]).

8.2. Vectorial Spherical Data

We refer to the datasets B1 and B5 from [4]. The sizes of samples are 50 and 52, respectively. As $\lceil\sqrt{50}\rceil = \lceil\sqrt{52}\rceil = 8$, the use of a regular tessellation of S^2 by the eight faces of a regular octahedron is appropriate. The eight centers of the faces form a cube. Another reason motivating this choice is that in geographical or geological coordinates, the eight faces of a regular spherical octahedron have simple parametric equations. Therefore, they are suited for grouping data from B1 (measurements in geographical coordinates) in Table 4, and data from B5 (measurements in geological coordinates) in Table 5.

Table 4. Data B1, of the form $(L, \ell) \in (-90^\circ, 90^\circ) \times (0, 360^\circ)$, grouped on the vertices of a cubical graph $\Gamma = C$.

Vertex	Frequency
$L > 0,\ \ell < 90^\circ$	$n_1 = 0$
$L > 0,\ 90^\circ < \ell < 180^\circ$	$n_2 = 0$
$L > 0,\ 180^\circ < \ell < 270^\circ$	$n_3 = 0$
$L > 0,\ 270^\circ < \ell < 360^\circ$	$n_4 = 0$
$L < 0,\ \ell < 90^\circ$	$n_5 = 17$
$L < 0,\ 90^\circ < \ell < 180^\circ$	$n_6 = 19$
$L < 0,\ 180^\circ < \ell < 270^\circ$	$n_7 = 9$
$L < 0,\ 270^\circ < \ell < 360^\circ$	$n_8 = 5$

Table 5. Data B5, of the form $(\text{Dec.}, \text{Inc.}) \in (0, 360^\circ) \times (-90^\circ, 90^\circ)$, grouped on the vertices of a cubical graph $\Gamma = C$.

Vertex	Frequency
$\text{Dec.} < 90, \text{Inc.} > 0$	$n_1 = 7$
$90 < \text{Dec.} < 180, \text{Inc.} > 0$	$n_2 = 10$
$180 < \text{Dec.} < 240, \text{Inc.} > 0$	$n_3 = 3$
$240 < \text{Dec.} < 360, \text{Inc.} > 0$	$n_4 = 8$
$\text{Dec.}. < 90, \text{Inc.} < 0$	$n_5 = 9$
$90 < \text{Dec.}. < 180, \text{Inc.} < 0$	$n_6 = 5$
$180 < \text{Dec.}. < 240, \text{Inc.} < 0$	$n_7 = 2$
$240 < \text{Dec.}. < 360, \text{Inc.} < 0$	$n_8 = 8$

The vertices of C, say $(v_i)_{1 \le i \le 8}$, are numbered according to the rows of the two tables. Adjacency relations are evident. Mutual distances between vertices are given by the matrix

$$\delta_C = \begin{pmatrix} 0 & 1 & 2 & 1 & 1 & 2 & 3 & 2 \\ 1 & 0 & 1 & 2 & 2 & 1 & 2 & 3 \\ 2 & 1 & 0 & 1 & 3 & 2 & 1 & 2 \\ 1 & 2 & 1 & 0 & 2 & 3 & 2 & 1 \\ 1 & 2 & 3 & 2 & 0 & 1 & 2 & 1 \\ 2 & 1 & 2 & 3 & 1 & 0 & 1 & 2 \\ 3 & 2 & 1 & 2 & 2 & 1 & 0 & 1 \\ 2 & 3 & 2 & 1 & 1 & 2 & 1 & 0 \end{pmatrix}.$$

In view of Proposition 5, the Green matrix is

$$G_C = \frac{1}{96} \begin{pmatrix} 29 & 1 & -7 & 1 & 1 & -7 & -11 & -7 \\ 1 & 29 & 1 & -7 & -7 & 1 & -7 & -11 \\ -7 & 1 & 29 & 1 & -11 & -7 & 1 & -7 \\ 1 & -7 & 1 & 29 & -7 & -11 & -7 & 1 \\ 1 & -7 & -11 & -7 & 29 & 1 & -7 & 1 \\ -7 & 1 & -7 & -11 & 1 & 29 & 1 & -7 \\ -11 & -7 & 1 & -7 & -7 & 1 & 29 & 1 \\ -7 & -11 & -7 & 1 & 1 & -7 & 1 & 29 \end{pmatrix}.$$

From a cursory inspection of Table 4, we are inclined to conclude that data from $B1$ are drawn from a population located exclusively in the region $L < 0$. No test is required to conclude that the null hypothesis of uniformity has to be rejected. However, for the sake of completeness, let us verify that our statistic leads to this conclusion.

Computations give

$$8 \times 50V_{50} \approx 33$$

The Laplacian spectrum of C is $\lambda_1^{f_1}\lambda_2^{f_2}\lambda_3^{f_3} = 2^3 4^3 6^1$, and from the R package we obtain

$$P(\frac{\chi_1^2(3)}{2} + \frac{\chi_1^2(3)}{4} + \frac{\chi_1^2(1)}{6} \geq 33) = 0\%.$$

The p-value equals, as expected, 0.

As for B5, Table 5 suggests, for consideration, a population uniformly distributed. We find

$$8 \times 50V_{50} \approx 3 \cdot 35,$$

and from the R package, we get

$$P(\frac{\chi_1^2(3)}{2} + \frac{\chi_1^2(3)}{4} + \frac{\chi_1^2(1)}{6} \geq 3 \cdot 35) = 21\%$$

as an approximation of the p-value. We accept uniformity, in agreement with (Example 5.7 p. 106, [4]).

8.3. Axial Spherical Data

An axial observation is a unit vector x, emanating from the origin. Thus, $-x$ and x are indistinguishable, as they define the same axis. In other words, it is $\pm x$ which is observed. Therefore, for grouping axial data, we have to use regular tessellations enjoying the antipodal symmetry.

Dataset B15 (from p. 295, [4]) forms a sample of 94 axial data. As $\sqrt{94} \approx 10$, the ten axes of a regular dodecahedron are appropriate to group these data. Given a regular tessellation of the sphere by the 20 faces of a regular icosahedron, the centres of antipodal faces joined by a straight line will form the 10 axes of the dual polyhedron, namely, a dodecahedron.

Therefore, let us first draw the icosahedral graph, as in Figure 3. This graph can be embedded in a sphere in the following way. Recall that ϕ denotes the golden ratio. The twelve points $(0, \pm 1, \pm\phi)$, $(\pm 1, \pm\phi, 0)$ and $(\pm\phi, 0, \pm 1)$ belong to the sphere centred at the origin with radius $(\phi^2 + 1)^{1/2}$. They are the vertices of a regular icosahedron with edges of length 2 (compare with (3.75 p. 52, [17])).

If $|\varepsilon_1| = |\varepsilon_2| = 1$, in view of equality $1 + \phi^2 + (\phi - 1)^2 = 4$, the five neighbours of $(0, \varepsilon_1, \varepsilon_2\phi)$ are easily seen to be $(0, -\varepsilon_1, \varepsilon_2\phi)$, $(\pm 1, \varepsilon_1\phi, 0)$ and $(\pm\phi, 0, \varepsilon_2)$. For other vertices, apply cyclic permutations of coordinates. We place $w_1 = (0, 1, \phi)$, $w_2 = (1, \phi, 0)$, $w_3 = (-1, \phi, 0)$, $w_4 = (-\phi, 0, 1)$, $w_5 = (0, -1, \phi)$,$w_6 = (\phi, 0, 1)$, $w_7 = (\phi, 0, -1)$, $w_8 = (0, 1, -\phi)$, $w_9 = (-\phi, 0, -1)$, $w_{10} = (-1, -\phi, 0)$, $w_{11} = (1, -\phi, 0)$, $w_{12} = (0, -1, -\phi)$, and will label vectors and matrices accordingly.

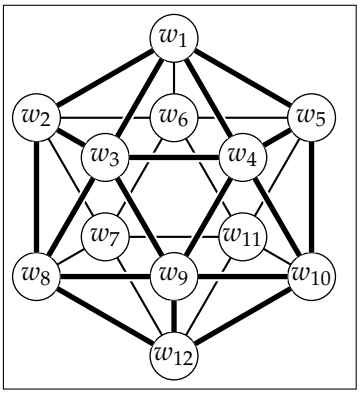


Figure 3. The icosahedral graph.

By duality, the centre of the faces of an icosahedron are the vertices of a dodecahedron. For instance, the centre of the face $w_1w_2w_3$ has Cartesian coordinates $3^{-1}(0, 2\phi + 1, \phi) = 3^{-1}\phi^2(0, \phi, \phi^{-1})$, and we will denote the corresponding vertex of a dodecahedron by $w_1w_2w_3 = (0, \phi, \phi^{-1})$. In this setting, the 20 vertices of our dodecahedron (of edge $2\phi^{-1}$), are $(0, \pm\phi, \pm\phi^{-1})$, $(\pm\phi, \pm\phi^{-1}, 0)$, $(\pm\phi^{-1}, 0, \pm\phi)$ and $(\pm1, \pm1, \pm1)$ (compare with (3.76 p. 53, [17])). Two adjacent faces of the icosahedron give rise to two adjacent vertices of the dodecahedron. The graph we obtain is isomorphic to the graph in Figure 1.

The antipodal quotient graph is constructed as follows. Each face $w_iw_jw_k$ of the icosahedron is a vertex of our new graph. If $(w_i, w_{i'})$, $(w_j, w_{j'})$ and $(w_k, w_{k'})$ are three antipodal pairs of vertices of I, then $w_iw_jw_k$ and $w_{i'}w_{j'}w_{k'}$ represent the same vertex of the new graph. If $w_k \sim w_\ell$, then $w_iw_jw_k$ and $w_iw_jw_\ell$ are neighbours. For instance, we have $w_5w_{10}w_{11} = w_2w_3w_8 \sim w_1w_2w_3$, so that our new graph can be represented as in Figure 4.

We recognize Petersen's graph; see Figure 2 above.

The 10 vertices of our 2-folded dodecahedron can be chosen to be (each set of coordinates being taken jointly with its opposite) $v_1 = w_1w_2w_3 = \pm(0, \phi, \phi^{-1})$, $v_2 = w_1w_3w_4 = \pm(-1, 1, 1)$, $v_3 = w_1w_4w_5 = \pm(-\phi^{-1}, 0, \phi)$, $v_4 = w_4w_5w_{10} = \pm(-1, -1, 1)$, $v_5 = w_5w_{10}w_{11} = \pm(0, -\phi, \phi^{-1})$, $v_6 = w_1w_2w_6 = \pm(1, 1, 1)$, $v_7 = w_6w_7w_{11} = \pm(\phi, -\phi^{-1}, 0)$, $v_8 = w_1w_5w_6 = \pm(\phi^{-1}, 0, \phi)$, $v_9 = w_2w_6w_7 = \pm(\phi, \phi^{-1}, 0)$, and $v_{10} = w_5w_6w_{11} = \pm(1, -1, 1)$.

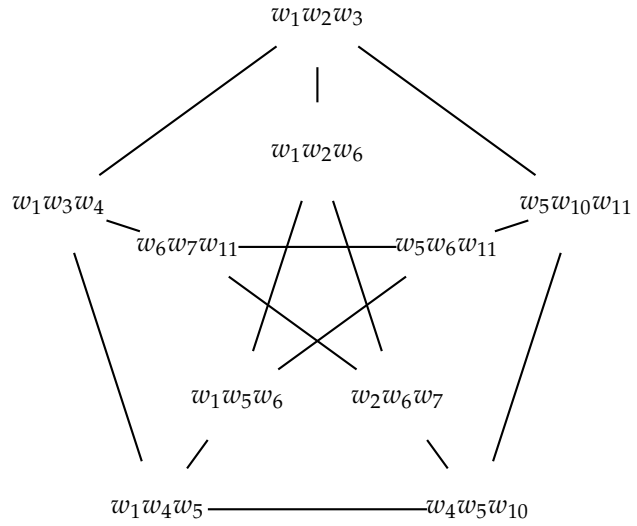


Figure 4. Petersen's graph with pairs of antipodal faces of the icosahedron as vertices.

Given, in Cartesian coordinates, an axial observation $\pm(x, y, z)$, we assign it to vertex $v_i = \pm(x_i, y_i, z_i)$, provided $|xx_i + yy_i + zz_i| \geq |xx_j + yy_j + zz_j|$ for $j = 1, \ldots, 10$. Computations yield Table 6.

With these labels, the distances between vertices are given by

$$\delta = \begin{array}{c} \\ v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \\ v_7 \\ v_8 \\ v_9 \\ v_{10} \end{array} \begin{array}{c} \begin{array}{cccccccccc} v_1 & v_2 & v_3 & v_4 & v_5 & v_6 & v_7 & v_8 & v_9 & v_{10} \end{array} \\ \begin{pmatrix} 0 & 1 & 2 & 2 & 1 & 1 & 2 & 2 & 2 & 2 \\ 1 & 0 & 1 & 2 & 2 & 2 & 1 & 2 & 2 & 2 \\ 2 & 1 & 0 & 1 & 2 & 2 & 2 & 1 & 2 & 2 \\ 2 & 2 & 1 & 0 & 1 & 2 & 2 & 2 & 1 & 2 \\ 1 & 2 & 2 & 1 & 0 & 2 & 2 & 2 & 2 & 1 \\ 1 & 2 & 2 & 2 & 2 & 0 & 2 & 1 & 1 & 2 \\ 2 & 1 & 2 & 2 & 2 & 2 & 0 & 2 & 1 & 1 \\ 2 & 2 & 1 & 2 & 2 & 1 & 2 & 0 & 2 & 1 \\ 2 & 2 & 2 & 1 & 2 & 1 & 1 & 2 & 0 & 2 \\ 2 & 2 & 2 & 2 & 1 & 2 & 1 & 1 & 2 & 0 \end{pmatrix} \end{array}$$

and the Green matrix, in view of results from Section 5, is

$$G_P = \frac{1}{100} \begin{pmatrix} 33 & 3 & -7 & -7 & 3 & 3 & -7 & -7 & -7 & -7 \\ 3 & 33 & 3 & -7 & -7 & -7 & 3 & -7 & -7 & -7 \\ -7 & 3 & 33 & 3 & -7 & -7 & -7 & 3 & -7 & -7 \\ -7 & -7 & 3 & 33 & 3 & -7 & -7 & -7 & 3 & -7 \\ 3 & -7 & -7 & 3 & 33 & -7 & -7 & -7 & -7 & 3 \\ 3 & -7 & -7 & -7 & -7 & 33 & -7 & 3 & 3 & -7 \\ -7 & 3 & -7 & -7 & -7 & -7 & 33 & -7 & 3 & 3 \\ -7 & -7 & 3 & -7 & -7 & 3 & -7 & 33 & -7 & 3 \\ -7 & -7 & -7 & 3 & -7 & 3 & 3 & -7 & 33 & -7 \\ -7 & -7 & -7 & -7 & 3 & -7 & 3 & 3 & -7 & 33 \end{pmatrix}.$$

Computations give

$$10 \times 94 V_{94} = 8.9,$$

and from the R package we get, for B5,

$$P\left(\frac{\chi_1^2(5)}{2} + \frac{\chi_1^2(4)}{5} \geq 8.9\right) = 0.7\%$$

as an approximation of the p-value. We reject uniformity, in agreement with Example 6.34 (p. 194, [4]).

Table 6. Data B15 grouped on the ten vertices of $\Gamma = P$.

Vertex	Frequency
$v_1 = w_1 w_2 w_3 = \pm(0, \phi, \phi^{-1})$	7
$v_2 = w_1 w_3 w_4 = \pm(-1, 1, 1)$	9
$v_3 = w_1 w_4 w_5 = \pm(-\phi^{-1}, 0, \phi)$	19
$v_4 = w_4 w_5 w_{10} = \pm(-1, -1, 1)$	12
$v_5 = w_5 w_{10} w_{11} = \pm(0, -\phi, \phi^{-1})$	7

Table 6. *Cont.*

Vertex	Frequency
$v_6 = w_1w_2w_6 = \pm(1,1,1)$	7
$v_7 = w_6w_7w_{11} = \pm(\phi,-\phi^{-1},0)$	6
$v_8 = w_1w_5w_6 = \pm(\phi^{-1},0,\phi)$	14
$v_9 = w_2w_6w_7 = \pm(\phi,\phi^{-1},0)$	5
$v_{10} = w_5w_6w_{11} = \pm(1,-1,1)$	8
Total	94

9. Discussion

9.1. Other Families of Graphs

Whereas our general result is based on the use of distance-regular graphs, all our examples were based exclusively upon distance-transitive graphs. Two questions arise. Firstly, are there examples of regular tessellations associated with distance-regular graphs that are not distance-transitive? More generally, can we include in our framework graphs that are not distance-regular?

To the first question, an answer is provided by Shrikande graph; see (§3.11 (vi) p. 104–105, [6]). It is the unique graph with order $|V| \leq 16$, which is not distance-transitive, but is distance-regular. The vertices of this graph can be used as the centres of 16 hexagons, providing a regular tessellation of a flat torus. See, e.g., Figures 5.3 p. 89 and 5.7 p. 98 in [18]. A host of similar examples is available.

To the second question, the answer is positive from the theoretical point of view. We have built a V-statistic with kernel G, defined to be the Moore–Penrose inverse of the combinatorial Laplacian. This inverse always exists. The assumption of distance-regularity only enabled us to obtain an explicit expression for G, in terms of integers from the intersection array, which is most important in view of computations in applications. But in one of our proofs we mentioned a connection with the theory of random walks on graphs. This tool provides another interpretation of G, in terms of hitting-times. Expressions of hitting times are available for many kinds of graphs.

9.2. Robustness

Recall from [19] that robust statistics is concerned with the fact that most theoretical results from probability theory are, at most, approximations to reality.

In regard to robustness, a fundamental remark about goodness-of-fit tests can be found in [20]: if a practitioner employs a large enough sample size, almost any goodness-of-fit test will result in rejection of the null hypothesis. In other words, the law of large numbers almost surely holds in theory, and almost surely fails in practice. Grouping data from large samples may overcome this problem.

On the other hand, the way we defined classes in order to group circular or spherical data may be seen as arbitrary and determining some lack of robustness in our tests.

However, once a system of coordinates has been chosen, the subdivisions of the range of coordinates we set forth are quite natural. Thus, the arbitrariness in the choice of classes, is, in most cases, only a consequence of the choice of a coordinate system. This issue is not new; see, e.g., (§3.8(ii), [4]) and references therein. In addition to this, a second, somewhat arbitrary, choice we had to make is the number of classes. We followed the rule of thumb based upon the square root of the sample size, but it is not consensual.

Therefore, an evaluation of our new statistical procedure, via simulations as described in (§3.6.1, [4]), remains necessary.

10. Conclusions

Whereas the issue of symmetry, in connection with regular polygons or polyhedra, has fascinated mathematicians of all times, some elementary goodness-of-fit tests for uniformity based on these familiar objects are, surprisingly, not as popular as one could expect. The aim of the present paper was to fill this gap. Secondly, in addition to the circle and the sphere, the case of manifolds as the flat torus may be discussed in a similar way. The Shrikhande graph is a famous example of a distance-regular graph that is not distance-transitive, and the tessellation of a torus based on this graph is a direction for further research. Thirdly, we have only outlined the statistical properties of our new family of statistics, and tools from Bahadur's theory of efficiency seem best suited to the algebraic framework or the Bose–Mesner algebra. Our aim is to discuss these matters in a forthcoming paper.

Funding: This research received no external funding.

Data Availability Statement: All datasets used in the present article are available in references [1,4,7].

Conflicts of Interest: The author declares no conflicts of interest.

References

1. Mardia, K.V.; Jupp, P.E. *Directional Statistics*; John Wiley & Sons: Hoboken, NJ, USA, 2009.
2. Ripley, B.D. *Spatial Statistics*; John Wiley & Sons: Hoboken, NJ, USA, 2005.
3. Weinberg, S. *Cosmology*; OUP Oxford: Oxford, UK, 2008.
4. Fisher, N.I.; Lewis, T.; Embleton, B.J. *Statistical Analysis of Spherical Data*; Cambridge University Press: Cambridge, UK, 1993.
5. Trudeau, R.J. *Introduction to Graph Theory*; Courier Corporation: North Chelmsford, MA, USA, 2013.
6. Brouwer, A.; Cohen, A.; Neumaier, A. *Distance-Regular Graphs*; Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge; Springer: Berlin/Heidelberg, Germany, 1989.
7. Mardia, K. *Statistics of Directional Data*; Probability and Mathematical Statistics: A Series of Monographs and Textbooks; Academic Press: Cambridge, MA, USA, 1972; Volume 13, p. Chs-6.
8. Curtin, B.; Nomura, K. Some formulas for spin models on distance-regular graphs. *J. Comb. Theory Ser. B* **1999**, *75*, 206–236. [CrossRef]
9. Jafarizadeh, M.A.; Sufiani, R. Perfect state transfer over distance-regular spin networks. *Phys. Rev. A—At. Mol. Opt. Phys.* **2008**, *77*, 022315. [CrossRef]
10. Kim, S.; Monterde, H.; Ahmadi, B.; Chan, A.; Kirkland, S.; Plosker, S. A generalization of quantum pair state transfer: S. Kim et al. *Quantum Inf. Process.* **2024**, *23*, 369. [CrossRef]
11. Biggs, N. *Algebraic Graph Theory*; Cambridge University Press: Cambridge, UK, 1974.
12. Cvetković, D.; Doob, M.; Sachs, H. *Spectra of Graphs: Theory and Applications*; Academic Press: Cambridge, London, 1980.
13. Basilevsky, A. *Applied Matrix Algebra in the Statistical Sciences*; Courier Corporation: North Chelmsford, MA, USA, 2013.
14. Brémaud, P. *Markov Chains: Gibbs Fields, Monte Carlo Simulation, and Queues*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 1999; Volume 31.
15. Knight, K. *Mathematical Statisitcs*; Chapman and Hall/CRC: Boca Raton, FL, USA, 2000.
16. Sheskin, D.J. *Handbook of Parametric and Nonparametric Statistical Procedures*; Chapman and Hall/CRC: Boca Raton, FL, USA, 2003.
17. Coxeter, H.S.M. *Regular Polytopes*; Macmillan: London, UK, 1973.
18. Royle, G.F.; Godsil, C. *Algebraic Graph Theory*; Springer: New York, NY, USA, 2001; Volume 207.
19. Hampel, F.; Ronchetti, E.; Rousseeuw, P.; Stahel, W. *Robust Statistics, Wiley*; John Wiley & Sons: Hoboken, NJ, USA, 1986.
20. Conover, W.J. *Practical Nonparametric Statistics*; John Wiley & Sons: Hoboken, NJ, USA, 1999.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.