A Decomposition for Invariant Tests of Uniformity on the Sphere

Jean-Renaud Pycke

ABSTRACT. We introduce a $U-$statistic on which can be based a test for uniformity on the sphere. It is a simple function of the geometric mean of distances between points of the sample and consistent against all alternatives. We show that this type of $U-$ statistics, whose kernel is invariant by isometries, can be separated into a set of statistics whose limiting random variables are independent. This decomposition is obtained via the so-called canonical decomposition of a group representation. The distribution of the limiting random variables of the components under the null hypothesis is given. We propose an interpretation of Watson type identities between quadratic functionals of Gaussian processes in the light of this decomposition.

1. Introduction

There are various problems in the field of directional statistics where the observations are directions in three dimensions. The surface of a unit sphere may then be used as the sample space for directions in space, each measurement being thought of as a point on a sphere of unit radius. One of the most important hypotheses about a distribution on a sphere is that of uniformity. We introduce in Theorem 2.1 a new $U-$statistic appropriate for testing uniformity on the sphere. A general survey and references concerning tests of uniformity for spherical data are given in [**8**], chapter 9-10.

The algebraic, geometrical, topological structures of the sphere give rise to particular problems that necessitate the use of special tools. For example the uniform distribution on the sphere does not have an extrinsic mean and therefore the theory of distributions with extrinsic mean (see [**1**] and [**2**]) though generic, cannot be applied. In the delicate area of spherical data that do not necessarily have a mean, the invariance under the action of a group can therefore play an important role. The uniform distribution is characterized by its invariance by $O(3)$, the group of isometries of the sphere. Several of the important theoretical distributions occurring in directional statistics are also characterized by invariance under the action of a group. Distributions with *rotational symmetry* are invariant by the group $SO(2)$ of rotations around a given direction. See [**8**] p.179 for examples and references about

1991 *Mathematics Subject Classification.* Primary 62G10, 62H11; Secondary 47G10, 20C15.
Key words and phrases. Goodness of fit test, $U-$ statistics, group representations.

models with rotational symmetry, particularly the celebrated von Mises-Fisher distribution. When the observations are not directions but axes the sample space is the set of couples of antipodal points of the sphere. Axial distributions correspond to spherical distributions invariant by *antipodal symmetry*. Different distributions (as those of Watson, Binghan) and tests of uniformity or rotational symmetry for axial distributions are discussed in [**8**]§ 9.4 and 10.7. We provide an example of utilization of group theory in section 3. Theorem 3.1 gives a method for deriving the decomposition of $U-$statistics whose kernel Φ is $G-$invariant with respect to a compact subgroup $G \subseteq O(3)$, i.e.

$$\forall g \in G \;:\; \Phi(g \cdot \xi_1, g \cdot \xi_2) = \Phi(\xi_1, \xi_2). \tag{1.1}$$

The interest of breaking-down a statistic into a set of uncorrelated components, each measuring some distinctive aspects of the data, has been exemplified in the basic papers [**3**], [**4**]. We show in section 4 how the statistic $U_{\Gamma,n}$ introduced in Theorem 2.1 can be decomposed in order to build goodness of fit tests whose hypotheses, given by (4.2) are related to invariance under the action of a group. The consistency of these components under certain alternatives is stated in Proposition 4.1 and Proposition 4.2. Example 4.1 deals with the case of rotational symmetry, Example 4.2 with antipodal symmetry. Example 4.3 illustrates the use of the character table of a finite group.

Our decomposition is obtained by combining two different tools, from spectral and group theory respectively. We first use classical spectral methods in order to obtain the well-known decompositions $(3.3)-(3.5)$. In the case where (1.1) holds we obtain a refinement of these decomposition by means of the *canonical decomposition* (following Serre's terminology in [**15**] §2.7) of the linear representations of G given by (3.7).

Interestingly, Watson's identity and bivariate generalizations introduced in [**9**] can be interpreted in the light of this approach, see Remark 4.4. Consequently, it seems to provide an efficient tool for deriving quadratic functionals of Gaussian processes arising as the limits in distribution of invariant $U-$statistics.

As it is underlined in the recent paper [**6**], the problem of finding systematical methods for building goodness of fit tests on the sphere and other manifolds remains widely opened. Giné established in [**5**] a general framework for testing uniformity on a wide family of sample spaces including the sphere. The eigenfunctions and eigenspaces of the Laplacian play a central role in this framework. Interestingly, the new test of uniformity introduced Theorem 2.1 is also closely related to the Laplacian, more precisely to the zero-mean Green's function of this operator given by (2.1). Natural extensions of this method to other manifolds and distribution will be discussed in a forthcoming paper.

Throughout this paper $C_\ell(k),\ (k, \ell \geq 1)$ will denote a sequence of independent random variables such that hold the equalities in law $C_\ell(k) = \chi^2(\ell) - \mathbb{E}\chi^2(\ell) = \chi^2(\ell) - \ell$, where $\chi^2(\ell)$ is a chi squared random variable having ℓ degrees of freedom.

2. A test of uniformity based on the geometric mean of spacings

Let $S^2 = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1\}$ be the unit sphere of the Euclidean space E_3. A point $\xi \in S^2$ is specified by spherical coordinates

$$(\text{colatitude, longitude}) = (\theta, \phi) \in [0, \pi] \times [0, 2\pi]$$

which are related to the Cartesian coordinates given by $x = \sin\theta \sin\phi$, $y = \sin\theta\cos\phi$, and $z = \cos\theta$. We consider a population specified by a probability density function $f(\xi) = f(\theta, \phi)$ with respect to the surface element

$$d\xi = \sin\theta d\theta d\phi.$$

Suppose that we wish to test the null hypothesis

> H_0 : $\xi_i(\theta_i, \phi_i)$, $1 \le i \le n$, is a sample of n independent observations from the uniform distribution $f(\xi) = f_0(\xi) := 1/(4\pi)$;

against the alternative hypothesis $H_1 : f \neq f_0$. Consider the kernel

$$\Gamma(\xi_1, \xi_2) := -\frac{1}{4\pi} \log \frac{e}{2}(1 - \overrightarrow{\xi_1} \cdot \overrightarrow{\xi_2}) \quad (\xi_1, \xi_2 \in S^2, \xi_1 \neq \xi_2) \tag{2.1}$$

where for each $\xi \in S^2$, $\overrightarrow{\xi}$ denotes the unit vector emanating from the origin of the Cartesian system. The idea underlying the use of Γ for testing uniformity on S^2 arises naturally from the interpretation of the celebrated Watson's statistic U_n, introduced in [**17**] in order to test uniformity on the circle S^1. In brief outline, U_n is a degenerate $V-$statistic with kernel the zero-mean Green's function of the Laplacian on S^1. Our kernel Γ can be shown to be the zero-mean Green's function of the Laplacian on S^2. In the following Theorem

$$\delta(\xi_1, \xi_2) := \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2} \text{ and } \gamma_n := \prod_{1 \le i < j \le n} \delta(\xi_i, \xi_j)^{\frac{2}{n(n-1)}}$$

are defined to be the distance between two points in E_3, and given n points $\xi_1, ..., \xi_n$, the geometric mean of distances between these points.

THEOREM 2.1. *Under H_0, the $U-$statistic*

$$U_{\Gamma,n}(\xi_1, \dots, \xi_n) := \frac{2}{n-1} \sum_{1 \le i < j \le n} \Gamma(\xi_i, \xi_j) = -\frac{n}{4\pi} \log\left(\frac{e}{4}\gamma_n^2\right) \tag{2.2}$$

converges in law, as $n \to \infty$, toward the random variable

$$U_\Gamma := \sum_{\ell=1}^{\infty} \frac{C_\ell(2\ell + 1)}{\ell(\ell + 1)}. \tag{2.3}$$

PROOF. First noticing that $1 - \overrightarrow{\xi_1} \cdot \overrightarrow{\xi_2} = \delta^2(\xi_1, \xi_2)/2$, we obtain

$$\Gamma(\xi_1, \xi_2) = -(4\pi)^{-1} \log[4\delta^2(\xi_1, \xi_2)/e]$$

and the second equality in (2.2) is readily checked. For the following basic facts about spherical harmonics see, e.g., [**13**], chapter III. Let P_ℓ and P_ℓ^m denote the Legendre polynomials and associated Legendre functions defined, for $\ell = 0, 1, ...,$ by

$$P_\ell(x) := \frac{1}{\ell! 2^\ell} \frac{d^\ell}{dx^\ell}(x^2 - 1)^\ell, \; P_\ell^m(x) := (1 - x^2)^{\frac{m}{2}} \frac{d^m P_\ell(x)}{dx^m} \quad (1 \le m \le \ell),$$

(see [**13**], formulas (7) p.174 and (1) p.246). An orthonormal basis of the Hilbert space $L^2(S^2)$ equipped with the inner product (3.1) is provided by the set of spherical harmonics $\cup_{\ell=0}^{\infty}\{f_\ell^m : -\ell \le m \le \ell\}$ with (see [**13**], $(II') - (IV)$ p.262-264)

$$f_\ell^0(\xi) := [(2\ell + 1)/(4\pi)]^{1/2} P_\ell(\cos\theta) \quad (\ell \ge 0) \tag{2.4}$$

$$f_\ell^m(\xi) := \alpha_\ell^m \cos(m\phi) P_\ell^m(\cos\theta) \qquad (1 \le m \le \ell) \tag{2.5}$$

$$f_\ell^m(\xi) := \alpha_\ell^m \sin(m\phi) P_\ell^{|m|}(\cos\theta) \qquad (-\ell \le m \le -1) \tag{2.6}$$

with $\alpha_\ell^m = \alpha_\ell^{-m} = [(2\ell+1)(\ell-m)!/(2\pi(\ell+m)!)]^{1/2}$ for $m > 0$. From formula (105) p.311 in [**11**] we derive the pointwise convergent expansion

$$\log 2 - 1 - \log\left(1 - \overrightarrow{\xi_1}\cdot\overrightarrow{\xi_2}\right) = \sum_{\ell=1}^{\infty} \frac{(2\ell+1)P_\ell(\overrightarrow{\xi_1}\cdot\overrightarrow{\xi_2})}{\ell(\ell+1)} \quad (\xi_1,\xi_2\in S^2, \xi_1\neq\xi_2). \tag{2.7}$$

On dividing both sides of (2.7) by $1/(4\pi)$ and in view of (2.1) and (2.4) it becomes

$$\Gamma(\xi_1,\xi_2) = \sum_{\ell=1}^{\infty} \frac{[(2\ell+1)/(4\pi)]^{1/2}\ f_\ell^0(\overrightarrow{\xi_1}\cdot\overrightarrow{\xi_2})}{\ell(\ell+1)}, \quad (\xi_1,\xi_2\in S^2, \xi_1\neq\xi_2). \tag{2.8}$$

In spherical coordinates $\overrightarrow{\xi_1}\cdot\overrightarrow{\xi_2} = \cos\theta_1\cos\theta_2 + \sin\theta_1\sin\theta_2\cos(\phi_1-\phi_2)$ and the addition formula for Legendre polynomials (see (VII') p.268 in [**13**]) may be written

$$[(2\ell+1)/(4\pi)]^{1/2} f_\ell^0(\overrightarrow{\xi_1}\cdot\overrightarrow{\xi_2}) = \sum_{-\ell\le m\le\ell} f_\ell^m(\xi_1)f_\ell^m(\xi_2). \tag{2.9}$$

Furthermore, from [**12**], inequality (35) p.87 involving Legendre functions, we infer that there exists $a > 0$ such that $|f_\ell^m(\xi_1)f_\ell^m(\xi_2)| \le a\ell^{1/2}$ if $(\theta_1,\theta_2,m,\ell)\neq(0,0,0,0)$. Consequently the general term of the series on the right-hand side of (2.10) is of order $a/\sqrt{\ell}$, hence converges to 0. When combined with (2.8) and (2.9) this fact ensures the pointwise convergence in the expansion

$$\Gamma(\xi_1,\xi_2) = \sum_{\ell=1}^{\infty}\sum_{-\ell\le m\le\ell} \frac{f_\ell^m(\xi_1)f_\ell^m(\xi_2)}{\ell(\ell+1)} \quad (\xi_1,\xi_2\in S^2, \xi_1\neq\xi_2). \tag{2.10}$$

The convergence is also valid in $L^2(S^2\times S^2)$ and the convergence in law of $U_{\Gamma,n}$ toward the random variable (2.3) is a consequence of Theorem 4.3.1 p.138 in [**7**]. $\square$

3. Decomposition of $G-$invariant $U-$statistics

For basic definitions and facts about groups and their representations the reader is referred to [**15**], Part I. Let $\mathcal{G}$ denote the set of compact subgroups of $O(3)$. These groups (as the cyclic and dihedral groups or the symmetry groups of Platonic solids) are of particular interest in mathematical physics. Some of them are discussed in [**15**] §5.1 − 5.6. Let $G\in\mathcal{G}$. An isometry $g\in G$ maps a point $\xi\in S^2$ onto $g\xi\in S^2$. An *action* of G on functions f and Φ defined on S^2 and $S^2\times S^2$ respectively is given by the shift operators

$$g\cdot f(\xi) := f(g^{-1}\xi), \quad g\cdot\Phi(\xi,\eta) := \Phi(g^{-1}\xi, g^{-1}\eta),\ g\in G.$$

A function f (resp. a set of functions $\mathcal{F}$) is said to be $G-$*invariant* if for each $g\in G$, one has $g\cdot f = f$ (resp. $g\cdot f\in\mathcal{F}$ for each $f\in\mathcal{F}$).

Consider the Hilbert space $L^2(S^2)$ of square integrable functions $f: S^2\to\mathbb{R}$ equipped with the usual inner product and the corresponding norm

$$(f_1|f_2) := \int_{S^2} f_1(\xi)f_2(\xi)d\xi, \quad \|f\| := (f|f)^{1/2}. \tag{3.1}$$

Consider a $U-$ statistic defined as

$$U_n(\xi_1,\dots,\xi_n) = \frac{2}{(n-1)}\sum_{1\le i<j\le n}\Phi(\xi_i,\xi_j) \tag{3.2}$$

where Φ is a real valued kernel satisfying the four following conditions.

- C1 : $\Phi(\xi_1,\xi_2) = \Phi(\xi_2,\xi_1)$ and $\sup_{\xi_1\in S^2}\int_{S^2}\Phi(\xi_1,\xi_2)^2 d\xi_2 < \infty$;

- C2 : $\int_{S^2} \Phi(\xi_1, \xi_2) d\xi_2 = 0$ (*degeneracy condition*);
- C3 : Φ is $G-$invariant;
- C4 : for each $\xi_1 \in S^2$, the mean convergent expansion (3.5) is pointwise convergent for each $\xi_2 \in S^2$ except maybe for $\xi_2 \in C_{\xi_1}$ where $C_{\xi_1} \subseteq S^2$ is a finite set.

Note that it is readily checked from (2.1) (for C1 and C3) and the expansion (2.10) (for C2 and C4) that our kernel Γ satisfies these four conditions. Let us briefly motivate these assumptions. C1 and C2 are standard hypotheses allowing to define a degenerate $U-$statistic with kernel Φ to whom we can apply the method of orthogonal decomposition (see e.g. [**14**] §5.5.2 or [**7**] §4.3). C4 is convenient in order to obtain a pointwise decomposition in (3.16) and not only an equality in L^2. Classical results from spectral theory in Hilbert spaces state that under C1 the integral operator A defined in $L^2(S^2)$ by

$$A : f(.) \mapsto Af(.) := \int_{S^2} \Phi(., \eta) f(\eta) d\eta$$

is a Hilbert-Schmidt operator. Consequently we have the orthogonal decomposition

$$L^2(S^2) = \ker A \oplus (\oplus_{\ell \geq 1} E_\ell), \quad \text{with } \dim E_\ell < \infty, \tag{3.3}$$

where $(E_\ell)_{\ell \geq 1}$ are the eigenspaces of A corresponding to the sequence $(\lambda_\ell)_{\ell \geq 1}$ of pairwise distinct non-null eigenvalues which satisfy $\sum_\ell (\dim E_\ell) \lambda_\ell^2 < \infty$. For each $\ell \geq 1$ let $\{\phi_\ell^m : 1 \leq m \leq \dim E_\ell\}$ denote an orthonormal basis of E_ℓ. For each $\xi \in S^2$ the kernel $\Phi_\ell(\xi, .) := \sum_m \phi_\ell^m(\xi) \phi_\ell^m(.) \in L^2(S^2)$ does not depend on the choice of the basis, since $\Phi_\ell(\xi, .)$ is the projection of $\Phi(\xi, .)$ into E_ℓ. Thus the *eigenfunction expansion* of Φ usually given in the form

$$\Phi(\xi_1, \xi_2) \stackrel{L^2}{=} \sum_{\ell=1}^{\infty} \lambda_\ell [\sum_{m=1}^{\dim E_\ell} \phi_\ell^m(\xi_1) \phi_\ell^m(\xi_2)] \tag{3.4}$$

(see (2) p.196 in [**14**] or (4.2.6) p.138 in [**7**]), can be rewritten in the canonical form (independent of the choice of orthonormal bases in the spaces $E_\ell, \ell \geq 1$)

$$\Phi(\xi_1, \xi_2) \stackrel{L^2}{=} \sum_{\ell=1}^{\infty} \lambda_\ell \Phi_\ell(\xi_1, \xi_2). \tag{3.5}$$

It will be noted that (3.4) and (3.5) correspond to (2.10) and (2.8) for $\Phi = \Gamma$, with $\lambda_\ell = [\ell(\ell+1)]^{-1}$.

Proposition 3.1. *Assume Φ satisfies C1,C2 and C3. Then : (a) each eigenspace of the associated integral operator is $G-$invariant, (b) each kernel Φ_ℓ satisfies C1,C2 and C3.*

Proof. (a) follows from the equalities, valid for f satisfying $Af = \lambda f$ and $g \in G$,

$$\begin{aligned} A(g \cdot f)(\xi) &= \int_{S^2} \Phi(\xi, \eta) f(g^{-1}\eta) d\eta = \int_{S^2} \Phi(g^{-1}\xi, g^{-1}\eta) f(g^{-1}\eta) d\eta \\ &= \int_{S^2} \Phi(g^{-1}\xi, g^{-1}\eta) f(g^{-1}\eta) d(g^{-1}\eta) = \int_{S^2} \Phi(g^{-1}\xi, \eta) f(\eta) d\eta \\ &= (Af)(g^{-1}\xi) = \lambda f(g^{-1}\xi) = \lambda (g \cdot f)(\xi) \end{aligned}$$

where for the second and third equalities we used the $G-$invariance of Φ and $d\eta$ respectively. For (b), C1 follows readily from the fact that for each $\ell \geq 1$, $\Phi_\ell(\xi, .)$ is the orthogonal projection of $\Phi(\xi,.)$ into E_ℓ. Next remark that C2 is fulfilled for Φ means that the constant function $\mathbf{1}:\xi\mapsto 1$ belongs to $\ker A$. From (3.3) we obtain $\Phi_\ell(\xi,.)\in E_\ell\subseteq(\ker A)^\perp$ which implies $(\Phi_\ell(\xi,.)|\mathbf{1})=0$ and C2 is satisfied. By invariance of the inner product under $G-$ action it is clear that for each orthonormal basis $(\phi_\ell^m)_m$ the set $(g\cdot\phi_\ell^m)_m$ is another orthonormal basis of E_ℓ. The remark made before (3.4) then allows us to write $\Phi_\ell(g\cdot\xi_1,g\cdot\xi_2)=\sum_m g\cdot\phi_\ell^m(\xi_1)g\cdot\phi_\ell^m(\xi_2)=\sum_m\phi_\ell^m(\xi_1)\phi_\ell^m(\xi_2)=\Phi_\ell(\xi_1,\xi_2)$ and C3 is satisfied. □

For each $G\in\mathcal{G}$, let $\widehat{G}$ denote the set of characters corresponding to a complete set of mutually nonisomorphic irreducible unitary representations of G. For each $\chi\in\widehat{G}$, d_χ denotes the degree of χ. The set $\widehat{G}$ is finite whenever G is finite, or countably infinite otherwise. If G is not finite, let dg be the invariant measure (or Haar measure) of the group G. We are only concerned with real valued statistics and random variables. We shall therefore consider a set of characters, say Υ, such that

$$G\in\mathcal{G},\ \Upsilon\subseteq\widehat{G}\text{ and each }\chi\in\Upsilon\text{ has values in }\mathbb{R}. \tag{3.6}$$

Assertion (a) in Proposition 3.1 enables us to define a linear representation T_ℓ of G in E_ℓ by putting, for each $g\in G$,

$$T_\ell(g):\begin{array}{ccc}E_\ell&\to&E_\ell\\ f&\mapsto&g\cdot f.\end{array} \tag{3.7}$$

For any function $f:S^2\to\mathbb{R}$ and $\chi\in\widehat{G}$ we set

$$P^\chi f(\xi):=\begin{cases}\frac{d_\chi}{|G|}\sum_{g\in G}\chi(g)f(g^{-1}\xi) & \text{if } G \text{ is finite,}\\ d_\chi\int_G\chi(g)f(g^{-1}\xi)dg & \text{if } G \text{ is not finite.}\end{cases} \tag{3.8}$$

THEOREM 3.1. *Assume* (3.6) *is fulfilled. Let* Φ *be a kernel satisfying the four conditions C1-C4. Let* χ_ℓ *denote the character of the representation* (3.7)*. For each* $\chi\in\widehat{G}$*, the mapping* $f\in E_\ell\mapsto P^\chi f$ *is an orthogonal projection of* E_ℓ *into* $E_\ell^\chi\subseteq E_\ell$*. If* $\chi_1,\chi_2\in\Upsilon$ *and* $\chi_1\neq\chi_2$ *then the spaces* $E_\ell^{\chi_1}$ *and* $E_\ell^{\chi_2}$ *are orthogonal. The space* E_ℓ^χ *is* G*-invariant and of dimension*

$$\dim E_\ell^\chi=d_\chi<\chi_\ell|\chi>,\ \textit{where}\ <\chi_\ell|\chi>:=\begin{cases}|G|^{-1}\sum_{g\in G}\chi_\ell(g)\chi(g) & \textit{if } G \textit{ is finite,}\\ \int_G\chi(g)\chi_\ell(g)dg & \textit{if } G \textit{ is not finite.}\end{cases}$$

If furthermore (3.6) *is satisfied with* $\Upsilon=\widehat{G}$*, then*

$$E_\ell=\bigoplus_{\chi\in\widehat{G}}E_\ell^\chi\ ,\ \textit{hence}\ (\ker A)^\perp=\bigoplus_{\ell\geq1}\bigoplus_{\chi\in\widehat{G}}E_\ell^\chi. \tag{3.9}$$

PROOF. This Theorem follows from results proved in [**15**]. See Theorem 8 p.21 for the case where G is finite. For the case where G is not finite, we use the extensions of the preceding Theorems stated in assertions (a) and (e) in § 4.3. □

We are now equipped to deal with the decomposition our $G-$invariant $U-$statistic. From a kernel Φ, we obtain a new kernel by setting

$$\Phi^\chi(\xi,.):=P^\chi\Phi(\xi,.). \tag{3.10}$$

PROPOSITION 3.2. *Assume* (3.6) *is satisfied. Let* Φ *be a kernel fulfilling the four conditions C1-C4. Then for each* $\chi \in \Upsilon$ *the latter are also satisfied by* Φ^χ*, the expansion* (3.5) *referred to in C4 being replaced for* Φ^χ *by*

$$\Phi^\chi(\xi_1, \xi_2) \overset{L^2}{=} \sum_{\ell=1}^{\infty} \lambda_\ell \Phi_\ell^\chi(\xi_1, \xi_2). \tag{3.11}$$

For each $\xi_1 \in S^2$*, the convergence in* (3.11) *is pointwise for each* $\xi_2 \in S^2$*, except maybe for* ξ_2 *belonging to a finite or countable set.*

PROOF. When G satisfies (3.6), we know from [**15**], assertion (ii) of Proposition 1 p.10 that $\chi(g^{-1}) = \chi(g)$ for each $g \in G$. To avoid notational cumbersomeness, we restrict the proof concerning the symmetry in C1 to the case where G is finite. We have

$$P^\chi\Phi(\xi,\eta) = \frac{d_\chi}{|G|}\sum_{g\in G}\chi(g)\Phi(\xi, g^{-1}\eta) \qquad \text{by definition of } P^\chi(\Phi)$$

$$= \frac{d_\chi}{|G|}\sum_{g\in G}\chi(g)\Phi(\xi, g\eta) \quad \text{by changing } g \text{ into } g^{-1} \text{ and using } \chi(g^{-1}) = \chi(g)$$

$$= \frac{d_\chi}{|G|}\sum_{g\in G}\chi(g)\Phi(g^{-1}\xi, \eta) = \frac{d_\chi}{|G|}\sum_{g\in G}\chi(g)\Phi(\eta, g^{-1}\xi) = P^\chi\Phi(\eta,\xi)$$

where for the last equalities we used the $G-$invariance and the symmetry of Φ. Thus Φ^χ is symmetric. The second assertion in C1 is a direct consequences of the spectral decomposition (3.3) and the fact that the restriction of P^χ to each E_ℓ is an orthogonal projection. C2 follows readily from assertion (b) in Proposition 3.1. For C3 we first notice that (iii) in Proposition 1 p.10 in [**15**] implies $\chi(h^{-1}gh) = \chi(g)$ for $g, h \in G$. And this enables us to obtain, for any $h \in G$, the relations

$$P^\chi\Phi(h\xi, h\eta) = \frac{d_\chi}{|G|}\sum_{g\in G}\chi(g)\Phi(h\xi, g^{-1}h\eta) = \frac{d_\chi}{|G|}\sum_{g\in G}\chi(hgh^{-1})\Phi(h\xi, (hgh^{-1})^{-1}h\eta)$$

$$= \frac{d_\chi}{|G|}\sum_{g\in G}\chi(g)\Phi(h\xi, hg^{-1}\eta) = \frac{d_\chi}{|G|}\sum_{g\in G}\chi(g)\Phi(\xi, g^{-1}\eta) = P^\chi\Phi(\xi,\eta)$$

which proves C3. We omit details for C4. □

The preceding lemma enables us to define new $G-$invariant degenerate $U-$statistics with kernels Φ^χ defined by (3.10) and

$$\Phi^\Upsilon(\xi_1,\xi_2) := \sum_{\chi\in\Upsilon}\Phi^\chi(\xi_1,\xi_2), \quad \overline{\Phi}^\Upsilon(\xi_1,\xi_2) := \Phi(\xi_1,\xi_2) - \Phi^\Upsilon(\xi_1,\xi_2) \tag{3.12}$$

and the corresponding $U-$statistics

$$U_n^\chi(\xi_1,...,\xi_n) := \frac{2}{(n-1)}\sum_{1\le i<j\le n}\Phi^\chi(\xi_i,\xi_j), \quad (\chi\in\Upsilon) \tag{3.13}$$

$$U_n^\Upsilon(\xi_1,...,\xi_n) := \sum_{\chi\in\Upsilon}U_n^\chi(\xi_1,...,\xi_n), \tag{3.14}$$

$$\overline{U}_n^\Upsilon(\xi_1,...,\xi_n) := U_n(\xi_1,...,\xi_n) - \sum_{\chi\in\Upsilon}U_n^\chi(\xi_1,...,\xi_n). \tag{3.15}$$

THEOREM 3.2. *Assume* (3.6) *is satisfied and let* Φ *be a kernel fulfilling the four conditions C1-C4. Then one has for each* $n \geq 2$ *and almost every* n*-tuple* ξ_i, $(1 \leq i \leq n)$,

$$U_n(\xi_1, \ldots, \xi_n) = \sum_{\chi \in \Upsilon} U_n^{\chi}(\xi_1, \ldots, \xi_n) + \overline{U}_n^{\Upsilon}(\xi_1, ..., \xi_n). \tag{3.16}$$

Under H_0, *the statistics* $U_n^{\chi}, (\chi \in \Upsilon)$ *and* $\overline{U}_n^{\Upsilon}$ *are asymptotically pairwise independent and one has the convergence in law*

$$\lim U_n^{\chi} = \sum_{\ell \geq 1} \lambda_\ell C_\ell(d\chi < \chi_\ell | \chi >), \quad \lim U_n^{\Upsilon} = \sum_{\ell \geq 1} \lambda_\ell C_\ell(\sum_{\chi \in \Upsilon} d\chi < \chi_\ell | \chi >),$$

$$\text{and } \lim \overline{U}_n^{\Upsilon} = \sum_{\ell \geq 1} \lambda_\ell C_\ell(\dim E_\ell - \sum_{\chi \in \Upsilon} d\chi < \chi_\ell | \chi >)$$

where χ_ℓ *is the character of representation* (3.7).

PROOF. This Theorem is a consequence of Theorem 3.1 combined with basic results from the theory of orthogonal expansions applied to $U-$ statistics, see, e.g., [**7**], Theorem 4.3.1 p.138.

□

4. Application to goodness of fit tests with $G-$invariant hypotheses

We first discuss the consistency of tests based on U_n or the statistics (3.13) − (3.15). Let $\mathcal{F}_{S^2} \subseteq L^2(S^2)$ denote the set of probability density functions on the sphere.

PROPOSITION 4.1. *Suppose that* $\mathcal{F}_0, \mathcal{F}_1 \subseteq \mathcal{F}_{S^2}$ *and* $f_0 \in \mathcal{F}_0$. *The test based on rejecting* $H_0 : f \in \mathcal{F}_0$, *against* $H_1 : f \in \mathcal{F}_1$, *for large absolute values of the* $U-$*statistic defined by* (3.2) *is consistent when*

$$\mathcal{F}_0 \subseteq \ker A, \quad \mathcal{F}_1 \cap \ker A = \emptyset, \tag{4.1}$$

A being the integral operator associated with Φ. In particular the test of uniformity based on $U_{\Gamma,n}$ is consistent against all alternatives.

PROOF. If $f \in \mathcal{F}_1$ holds, the $U-$statistic U_n with kernel Φ is non degenerate, and we know from [**7**], Theorem 4.2.1, or [**14**], Theorem A p.192 that there exist $\mu \in \mathbb{R}$ and $\sigma > 0$ such that the convergence in law $n^{1/2}(U_n/n - \mu) \to \mathcal{N}(0, \sigma^2)$ holds. When compared with (2.3), this convergence implies the desired result. In the particular case where $\Gamma = \Phi$ and $\mathcal{F}_0 = \{f_0\}$, we use the fact that the kernel of the integral operator associated with Γ is the set of constant functions whose orthogonal is generated by the set of nonconstant spherical harmonics. □

We now fix $\Phi = \Gamma$. Recall that the trivial representation of a group G denoted by χ_0 is the representation of degree one defined by $\chi_0(g) = 1$ for each $g \in G$. In this case we shall use the notations Γ^G, $U_{\Gamma,n}^G$, $\overline{U}_{\Gamma,n}^G$ and E_ℓ^G instead of Γ^{χ_0}, $U_n^{\chi_0}$, $\overline{U}_n^{\chi_0}$ and $E_\ell^{\chi_0}$. Let $\mathcal{F}_G$ denote the set of $G-$ invariant distributions on the sphere. The cases where $\mathcal{F}_0 = \{f_0\}, \mathcal{F}_1 = \mathcal{F}_G \setminus \{f_0\}$ and $\mathcal{F}_0 = \mathcal{F}_G, \mathcal{F}_1 = \mathcal{F}_{S^2} \setminus \mathcal{F}_G$ in Proposition 4.1 correspond to the two goodness of fit tests

$$\text{T} : \begin{cases} H_0 : f \text{ is uniform,} \\ H_1 : f \text{ is } G- \text{ invariant but not uniform.} \end{cases} \qquad \text{T'}: \begin{cases} H_0' : f \text{ is } G-\text{invariant,} \\ H_1' : f \text{ is not } G- \text{ invariant.} \end{cases} \tag{4.2}$$

Proposition 4.2. *Assume* (3.6). *A test based on rejecting* H_0 *(resp.* H_0'*) for large values of* $|U^G_{\Gamma,n}|$ *(resp.*$|\overline{U}^G_{\Gamma,n}|$*) is consistent against* H_1 *(resp.* H_1'*). One has under* H_0 *the convergence in law*

$$U^G_{\Gamma,n} \to \sum_{\ell\geq 1} \frac{C_\ell(\dim E^G_\ell)}{\ell(\ell+1)} \text{ with } E^G_\ell = \{f \in E_\ell : f \text{ is } G\text{-invariant}\} \tag{4.3}$$

$$\text{and under } H_0' : \overline{U}^G_{\Gamma,n} \to \sum_{\ell\geq 1} \frac{C_\ell(2\ell+1-\dim E^G_\ell)}{\ell(\ell+1)} \tag{4.4}$$

Proof. Except for the consistency, the results of this Proposition are a restatement of Theorem 3.2 in two particular cases. Concerning the consistency, the fact that (4.1) is fulfilled in both cases is easily seen after noticing that $f \mapsto P^{\chi_0} f =: f^G$ is an orthogonal projection into $\mathcal{F}_G$. This implies the equivalences

$$(f|\Gamma^G(\xi,.)) = 0 \iff (f^G|\Gamma(\xi,.)) = 0,$$
$$(f|\Gamma(\xi,.) - \Gamma^G(\xi,.)) = 0 \iff (f - f^G|\Gamma^G(\xi,.)) = 0.$$

□

Example 4.1. Assume $G = SO(2)$ is the group of rotations g through an angle $\phi \in [0, 2\pi]$ around the polar axis, with Haar measure $dg = d\phi/(2\pi)$. We obtain

$$\Gamma^{SO(2)}(\xi_1,\xi_2) = \int_0^{2\pi} \Gamma(\xi_1,\xi_2)\frac{d\phi_2}{2\pi} = -\frac{1}{4\pi}\log\frac{e(1-\cos\theta_1)(1+\cos\theta_2)}{4}.$$

Under H_0: $U^{SO(2)}_{\Gamma,n} \to \sum_{\ell=1}^{\infty} \frac{C_\ell(1)}{\ell(\ell+1)}$ and under H_0': $\overline{U}^{SO(2)}_{\Gamma,n} \to \sum_{\ell=1}^{\infty} \frac{C_\ell(2\ell)}{\ell(\ell+1)}$

Example 4.2. Antipodal symmetry is invariance under the action of the group $\{I, \sigma\}$ where I is the identity and σ the reflection through the origin. The corresponding kernel is

$$\Gamma^{I,\sigma}(\xi_1,\xi_2) = \frac{\Gamma(\xi_1,\xi_2) + \Gamma(\xi_1,\sigma\cdot\xi_2)}{2} = -\frac{1}{4\pi}\log\frac{e}{2}(1 - |\overrightarrow{\xi_1}\cdot\overrightarrow{\xi_2}|^2)$$

Under H_0: $U^{I,\sigma}_n \to \sum_{\ell \; even} \frac{C_\ell(2\ell+1)}{\ell(\ell+1)}$ and under H_0': $\overline{U}^{I,\sigma}_{\Gamma,n} \to \sum_{\ell \; odd} \frac{C_\ell(2\ell)}{\ell(\ell+1)}$.

Example 4.3. The aim of this example is to show how the character table of a finite group can be used in order to write (3.8) explicitly. We follow the notations introduced in [**15**] §5.8. If G is the symmetry group of a regular tetrahedron it has 24 elements partitioned into 5 equivalence classes denoted 1, (ab), $(ab)(cd)$, (abc) and $(abcd)$. Furthermore we have $\widehat{G} = \{\chi_0, \varepsilon, \theta, \psi\varepsilon, \psi\}$ and these five characters are real valued. Hence Γ can be decomposed into five components. For example one of them corresponds to $\chi = \theta$, character of degree $d_\theta = \theta(1) = 2$. Therefore (3.8) is written, in view of the character table of G,

$$\Gamma^\theta(\xi_1,\xi_2) = \frac{2}{24}[2\Gamma(\xi_1,\xi_2) + \sum_{g\in(ab)(cd)} 2\Gamma(\xi_1, g\xi_2) - \sum_{g\in(abc} \Gamma(\xi_1, g\xi_2)]$$

Remark 4.4. We are now in a position, as claimed in the introduction, to show that Watson's identity and generalizations given in [**9**], Theorem 3 are related to the canonical decomposition of a group representation. These identities correspond to a decomposition of the form (3.16), applied to the covariance function or trajectories of the Gaussian processes appearing in these identities in the following way. Consider a group G acting on a set S, on which is defined a Gaussian process

$f(x,\omega)=f(x),\ x\in S$. Assume moreover that the latter has a covariance function Φ satisfying the invariance property,

$$\Phi(x,y)=\Phi(g\cdot x,g\cdot y)\quad \text{whence } f(x)\overset{\text{(in law)}}{=}f(g\cdot x)\quad (x\in S, g\in G).$$

We restrict ourselves to the case of Watson's identity given in [**16**], relation (7), with a new proof and references for different proofs. In [**10**] we gave an elementary proof of this identity, based on the decomposition of a function $f: S=[0,1]\to\mathbb{R}$ in the form $f(x)=f_1(x)+f_2(x):=[f(x)+f(1-x)]/2+[f(x)-f(1-x)]/2$. The group G of isometries of $[0,1]$ is $\{\iota,s\}$ with $\iota(x)=x,\ s(x)=1-x$. One has $\widehat{G}=\{\chi_1,\chi_2\}$ with $\chi_1(\iota)=\chi_1(s)=1$ and $\chi_2(\iota)=-\chi_2(s)=1$ hence $d_{\chi_1}=d_{\chi_2}=1$. In this setting the decomposition $f=f_1+f_2$ becomes $f=P^{\chi_1}f+P^{\chi_2}f$ where the projections are defined by (3.8).

Acknowledgment. The author would like to express his gratitude to an anonymous referee for his helpful comments and suggestions.

References

[1] R. N. Bhattacharya and V.Patrangenaru. *Large sample theory of intrinsic and extrinsic sample means on manifolds. I.* Ann.Statist. 31 (2003) no.1, 1–29.
[2] —- *Large sample theory of intrinsic and extrinsic sample means on manifolds. II.* Ann.Statist. vol. 33 (2005), no.3, 1225–1259.
[3] J. Durbin, M. Knott and C.C. Taylor. *Components of Cramér-von Mises statistics. I.* J. Roy. Statist. Soc. Ser. B 34 (1972), 290–307.
[4] —- *Components of Cramér-von Mises statistics. II.* J. Roy. Statist. Soc. Ser. B 37 (1975), 216–237.
[5] M. E. Giné. *Invariant tests for uniformity on compact Riemannian manifolds based on Sobolev norms.* Ann. Statist. vol. 3 (1975), no. 6, 1243–1266.
[6] P. E. Jupp *Sobolev tests of goodness of fit of distributions on compact riemannian manifolds.* Ann. Statist. vol. 33, (2005), no. 6, 2957–2966.
[7] V. S. Koroljuk and Yu. V. Borovskich. *Theory of U-statistics.* Mathematics and its Applications, 273. Kluwer Academic Publishers Group, Dordrecht, 1994.
[8] V. M. Mardia and P.E. Jupp. *Directional Statistics.* Wiley Series in Probability and Statistics. John Wiley, 2000.
[9] G. Peccati and M. Yor. *Identities in law between quadratic functionals of bivariate Gaussian processes, through Fubini Theorems and symmetric projections.* Preprint.
[10] J.-R. Pycke *Sur une identité en loi entre deux fonctionnelles quadratiques du pont brownien.* C.R. Acad. Sci. Paris, Ser. I 340 (2005).
[11] L. Robin *Fonctions sphériques de Legendre et fonctions sphéroïdales.* Tome II. Collection Technique et Scientifique du C. N. E. T. Gauthier-Villars, Paris 1959.
[12] —- *Fonctions sphériques de Legendre et fonctions sphéroïdales.* Tome III. Collection Technique et Scientifique du C. N. E. T. Gauthier-Villars, Paris 1959.
[13] G. Sansone. *Orthogonal functions.* Interscience Publishers, 1959.
[14] R.J. Serfling. *Approximation heorems of mathematical statistics.* John Wiley, 1980.
[15] J.-P. Serre. *Linear Representations of Finite Groups.* Graduate Texts in Mathematics, Vol. 42. Springer-Verlag, 1977.
[16] Z. Shi and M. Yor. *On an identity in law for the variance of the Brownian bridge.* Bull. London Math. Soc. 29 (1997), no. 1, 103–108.
[17] G.S. Watson *Goodness-of-fit tests on a circle.* Biometrika, 48 (1961) p. 109-114.

Université d'Évry Val d'Essone
Current address: Département de Mathématiques, Université d'Évry, Boulevard F. Mitterrand - 91025 Evry Cedex, France
E-mail address: jrpycke@univ-evry.fr, pycke@ccr.jussieu.fr