Addendum to:

Integrated Conditional Moment Tests for Parametric Conditional Distributions

1. Introduction

The main purpose of this addendum to Bierens and Wang (2012) [BW hereafter] is to provide a complete proof of Mercer's theorem (Lemma 3 in BW), including the underlying Hilbert-Schmidt theorem regarding the existence of eigenvalues and eigenfunction of complex-valued continuous symmetric positive semidefinite kernels. For the proof of Lemma 3 BW referred to an unpublished paper, Hadinejad-Mahram et al. (2002), which has disappeared from the internet, and a published paper by Krein (1998). The latter author derived the complex Hilbert-Schmidt and Mercer theorems as by-products of linear operator theory, which I am not familiar with, and likely the same applies to most of my fellow econometricians. Therefore, in this addendum I will provide my own proofs of the complex Hilbert-Schmidt and Mercer theorems.

The same problem occurred with the (real) version of Mercer theorem in Bierens and Ploberger (1997), for which the proof was a reference to an exercise in a textbook. Therefore, I decided to derive this proof myself. See the addendum to Bierens and Ploberger (1997) in Chapter 5. The proofs in the current addendum are complex adaptations of the ones in the former addendum.

Also, it appears that Lemma 4 is incorrect. Therefore, in this addendum I will provide a revised version of Lemma 4.

On the basis of the results in BW and in this addendum I have been able to update my first consistent model specification testing paper, Bierens (1982), by deriving the asymptotic null distribution of the test, which is similar to BW, and upper bounds of the critical values on the basis of the revised Lemma 4. See the addendum to Bierens (1982) in Chapter 2 and section 7 below.

In this addendum I will use the same notations as in BW, except that in integrals with respect to a probability measure μ on set $\mathbf{B}$ I will use the notation $\mathrm{d}\mu(\beta)$ instead of $\mu(d\beta)$ because $\mu(\beta)$ can be interpreted as a distribution function.

The proofs in this addendum employ Hilbert space theory at the level of Bierens (2014a), linear algebra at the level of Bierens (2004, Appendix I), complex calculus at the level of Bierens (2004, Appendix III), and measure and probability theory at the level of Bierens (2004, Ch. 1-3).

2. Complex Covariance Functions

In Lemma 3 in BW the covariance function is of the form

$$\begin{aligned}\Gamma(\beta_1,\beta_2) &= E[Z(\beta_1)\overline{Z(\beta_2)}] \\ &= E\left[\left(\mathrm{Re}[Z(\beta_1)] + \mathbf{i}.\,\mathrm{Im}[Z(\beta_1)]\right)\left(\mathrm{Re}[Z(\beta_2)] - \mathbf{i}.\,\mathrm{Im}[Z(\beta_2)]\right)\right] \\ &= E\left[\mathrm{Re}[Z(\beta_1)].\,\mathrm{Re}[Z(\beta_2)]\right] + E\left[\mathrm{Im}[Z(\beta_1)].\,\mathrm{Im}[Z(\beta_2)]\right] \\ &\quad +\mathbf{i}.E\left[\mathrm{Im}[Z(\beta_1)].\,\mathrm{Re}[Z(\beta_2)]\right] - \mathbf{i}.E\left[\mathrm{Re}[Z(\beta_1)].\,\mathrm{Im}[Z(\beta_2)]\right], \quad (2.1)\end{aligned}$$

where $Z(\beta)$ is a complex-valued zero-mean continuous Gaussian process on a compact subset $\mathbf{B}$ of a Euclidean space. This covariance function is symmetric positive semidefinite in the following sense.

First, symmetry in the complex case means that

$$\Gamma(\beta_1,\beta_2) = \overline{\Gamma(\beta_2,\beta_1)}, \tag{2.2}$$

where the bar denotes the complex conjugate, which follows straightforwardly from (2.1). In particular, writing

$$\Gamma(\beta_1,\beta_2) = \mathrm{Re}[\Gamma(\beta_1,\beta_2)] + \mathbf{i}.\,\mathrm{Im}[\Gamma(\beta_1,\beta_2)],$$

the symmetry condition (2.2) implies that

$$\mathrm{Re}[\Gamma(\beta_1,\beta_2)] = \mathrm{Re}[\Gamma(\beta_2,\beta_1)],\ \mathrm{Im}[\Gamma(\beta_1,\beta_2)] = -\,\mathrm{Im}[\Gamma(\beta_2,\beta_1)] \tag{2.3}$$

and thus

$$\mathrm{Im}[\Gamma(\beta,\beta)] = 0. \tag{2.4}$$

Second, positive semidefiniteness with respect to a probability measure μ on $\mathbf{B}$ means that

$$\int\int \overline{\varphi(\beta_1)}\Gamma(\beta_1,\beta_2)\varphi(\beta_2)\mathrm{d}\mu(\beta_1)\mathrm{d}\mu(\beta_2) \geq 0, \tag{2.5}$$

for all $\varphi \in L^2_{\mathbb{C}}(\mu)$, where:

Definition 2.1. $L^2_{\mathbb{C}}(\mu)$ denotes the Hilbert space of all Borel measurable complex-valued functions φ on $\mathbf{B}$ satisfying $\int |\varphi(\beta)|^2 \mathrm{d}\mu(\beta) < \infty$, endowed with the inner-product $\langle \varphi_1, \varphi_2 \rangle = \int \overline{\varphi_1(\beta)} \varphi_2(\beta) \mathrm{d}\mu(\beta)$ and associated norm

$$||\varphi|| = \sqrt{\langle \varphi, \varphi \rangle} = \sqrt{\int \overline{\varphi(\beta)} \varphi(\beta) \mathrm{d}\mu(\beta)} = \sqrt{\int |\varphi(\beta)|^2 \mathrm{d}\mu(\beta)}$$

and metric $||\varphi_1 - \varphi_2||$.

In particular, in the case (2.1) it can be shown, after some tedious but straightforward complex calculations, that

$$\begin{aligned}
\int \int & \overline{\varphi(\beta_1)} \Gamma(\beta_1, \beta_2) \varphi(\beta_2) \mathrm{d}\mu(\beta_1) \mathrm{d}\mu(\beta_2) \\
&= E\left[\left(\int \mathrm{Re}[\varphi(\beta)] \, \mathrm{Re}[Z(\beta)] \mathrm{d}\mu(\beta)\right)^2\right] \\
&\quad + E\left[\left(\int \mathrm{Im}[\varphi(\beta)] \, \mathrm{Im}[Z(\beta)] \mathrm{d}\mu(\beta)\right)^2\right] \\
&+ E\left[\left(\int \mathrm{Im}[\varphi(\beta)] \, \mathrm{Re}[Z(\beta)] \mathrm{d}\mu(\beta)\right.\right. \\
&\quad \left.\left. + \int \mathrm{Re}[\varphi(\beta)] \, \mathrm{Im}[Z(\beta)] \mathrm{d}\mu(\beta)\right)^2\right] \geq 0.
\end{aligned}$$

Moreover, the covariance function $\Gamma(\beta_1, \beta_2)$ in (2.1) is continuous because $\mathrm{Re}[Z(\beta)]$ and $\mathrm{Im}[Z(\beta)]$ are a.s. continuous.

In the mathematical literature such a function $\Gamma(\beta_1, \beta_2)$ is called a *kernel*, and I will do so too.

The interpretation of $\Gamma(\beta_1, \beta_2)$ as a covariance function of a continuous complex-valued Gaussian process is irrelevant for the Hilbert-Schmidt and Mercer theorems. All we need to require is that:

Assumption 2.1. *The kernel* $\Gamma(\beta_1, \beta_2)$ *on* $\mathbf{B} \times \mathbf{B}$ *is complex-valued, continuous, and symmetric positive semidefinite with respect to a probability measure* μ *on* $\mathbf{B}$, *where* $\mathbf{B}$ *is a compact subset of a Euclidean space,* and

Assumption 2.2. $\int\int|\Gamma(\beta_1,\beta_2)|^2\mathrm{d}\mu(\beta_1)\mathrm{d}\mu(\beta_2)>0$.

The latter excludes the case that $\Gamma(\beta_1,\beta_2)$ is identical zero.

3. The Hilbert-Schmidt Theorem for Complex Kernels

3.1. The eigenvalue-eigenfunction problem

The general eigenvalue problem is to find a (real) scalar λ and a function $\psi \in L^2_{\mathbb{C}}(\mu)$ normalized to unit norm, $||\psi||=1$, such that[1]

$$\int \Gamma(\beta_1,\beta_2)\psi(\beta_2)\mathrm{d}\mu(\beta_2)=\lambda\psi(\beta_1) \text{ for all } \beta_1\in\mathbf{B}. \tag{3.1}$$

Taking complex conjugates in (3.1), the latter is equivalent to

$$\int \overline{\psi(\beta_2)}.\Gamma(\beta_2,\beta_1)\mathrm{d}\mu(\beta_2)=\lambda.\overline{\psi(\beta_1)}. \tag{3.2}$$

If such a pair (λ,ψ) exists then by (2.5), (3.1), (3.2) and the normalization $||\psi||=1$,

$$\begin{aligned} 0 &\le \int\int \overline{\psi(\beta_1)}\Gamma(\beta_1,\beta_2)\psi(\beta_2)\mathrm{d}\mu(\beta_1)\mathrm{d}\mu(\beta_2) \\ &= \lambda\int \overline{\psi(\beta_1)}\psi(\beta_1)\mathrm{d}\mu(\beta_1)=\lambda\int|\psi(\beta)|^2\mathrm{d}\mu(\beta)=\lambda. \end{aligned}$$

Thus, eigenvalues of positive semidefinite kernels are real valued and nonnegative. Moreover, if a solution of (3.1) exists with $\lambda>0$ then the corresponding eigenfunction ψ is continuous on $\mathbf{B}$ because $\Gamma(\beta_1,\beta_2)$ is continuous on $\mathbf{B}\times\mathbf{B}$.

Now denote

$$\begin{aligned} G(\lambda,\psi) &= \int\left|\int\Gamma(\beta_1,\beta_2)\psi(\beta_2)\mathrm{d}\mu(\beta_2)-\lambda\psi(\beta_1)\right|^2\mathrm{d}\mu(\beta_1) \\ &\quad \text{for } \lambda\in(0,\infty),\ \ \psi\in L^2_{\mathbb{C}}(\mu) \text{ with } ||\psi||=1. \end{aligned}$$

[1] In Lemma 3 in BW, (3.1) was incorrectly stated as $\lambda\psi(\beta_1)=\int\Gamma(\beta_1,\beta_2)\overline{\psi(\beta_2)}\mathrm{d}\mu(\beta_2)$.

Then eigenvalue problem (3.1) for $\lambda > 0$ is equivalent to the problem:

$$\textit{Find a } \lambda > 0 \textit{ and a } \psi \in L^2_{\mathbb{C}}(\mu) \textit{ with } ||\psi|| = 1 \textit{ such that } G(\lambda, \psi) = 0. \qquad (3.3)$$

See Remark 4.1 in the addendum to Bierens and Ploberger (1997) in Chapter 5.

Observe that

$$\begin{aligned}
G(\lambda, \psi) &= \int \left(\int \overline{\psi(\beta_2)}.\Gamma(\beta_2, \beta)\mathrm{d}\mu(\beta_2) - \lambda\overline{\psi(\beta)} \right) \\
&\quad \times \left(\int \Gamma(\beta, \beta_2)\psi(\beta_2)\mathrm{d}\mu(\beta_2) - \lambda\psi(\beta) \right) \mathrm{d}\mu(\beta) \\
&= \int \left(\int \overline{\psi(\beta_1)}.\Gamma(\beta_1, \beta)\mathrm{d}\mu(\beta_1) - \lambda\overline{\psi(\beta)} \right) \\
&\quad \times \left(\int \Gamma(\beta, \beta_2)\psi(\beta_2)\mathrm{d}\mu(\beta_2) - \lambda\psi(\beta) \right) \mathrm{d}\mu(\beta) \\
&= \int \int \overline{\psi(\beta_1)}\Gamma_2(\beta_1, \beta_2)\psi(\beta_2)\mathrm{d}\mu(\beta_1)\mathrm{d}\mu(\beta_2) \\
&\quad -2\lambda \int \overline{\psi(\beta_1)}\Gamma(\beta_1, \beta_2)\psi(\beta_2)\mathrm{d}\mu(\beta_2) + \lambda^2,
\end{aligned}$$

where

$$\Gamma_2(\beta_1, \beta_2) = \int \Gamma(\beta_1, \beta)\Gamma(\beta, \beta_2)\mathrm{d}\mu(\beta).$$

Minimizing $G(\lambda, \psi)$ to λ yields

$$\lambda = \int \int \overline{\psi(\beta_1)}\Gamma(\beta_1, \beta_2)\psi(\beta_2)\mathrm{d}\mu(\beta_1)\mathrm{d}\mu(\beta_2) \qquad (3.4)$$

and substituting this solution in $G(\lambda, \psi)$ yields

$$\begin{aligned}
\underline{G}(\psi) &= \min_{\lambda > 0} G(\lambda, \psi) = \int \int \overline{\psi(\beta_1)}\Gamma_2(\beta_1, \beta_2)\psi(\beta_2)\mathrm{d}\mu(\beta_1)\mathrm{d}\mu(\beta_2) \\
&\quad - \left(\int \int \overline{\psi(\beta_1)}\Gamma(\beta_1, \beta_2)\psi(\beta_2)\mathrm{d}\mu(\beta_1)\mathrm{d}\mu(\beta_2) \right)^2. \qquad (3.5)
\end{aligned}$$

Thus, the eigenfunction problem (3.3) is equivalent to the following problem:

Find a $\psi \in L^2_{\mathbb{C}}(\mu)$ *with* $||\psi|| = 1$ *such that* $\underline{G}(\psi) = 0$.
Then the corresponding eigenvalue λ *is given by* (3.4).

3.2. The maximum eigenvalue problem

The solution (3.4) suggests that, possibly, the eigenfunction ψ_1 corresponding to the largest eigenvalue λ_1, and λ_1 itself, can be determined by

$$\psi_1 = \arg \max_{\psi \in L^2_{\mathbb{C}}(\mu), ||\psi||=1} \int\int \overline{\psi(\beta_1)}\Gamma(\beta_1,\beta_2)\psi(\beta_2)\mathrm{d}\mu(\beta_1)\mathrm{d}\mu(\beta_2), \qquad (3.6)$$

$$\lambda_1 = \int\int \overline{\psi_1(\beta_1)}\Gamma(\beta_1,\beta_2)\psi_1(\beta_2)\mathrm{d}\mu(\beta_1)\mathrm{d}\mu(\beta_2). \qquad (3.7)$$

If so, we must have that $\underline{G}(\psi_1) = 0$, or equivalently,

$$\int\int \overline{\psi_1(\beta_1)}\Gamma_2(\beta_1,\beta_2)\psi_1(\beta_2)\mathrm{d}\mu(\beta_1)\mathrm{d}\mu(\beta_2) = \lambda_1^2.$$

Indeed, this conjecture is correct.

Theorem 3.1.
(*a*) *Under Assumption* 2.1 *the pair* λ_1, ψ_1 *determined by* (3.6) *and* (3.7) *is the maximum eigenvalue with corresponding eigenfunction of the kernel* $\Gamma(\beta_1,\beta_2)$ *in Assumption* 2.1.
(*b*) *Under the additional Assumption* 2.2, $\lambda_1 > 0$.
(*c*) *Otherwise,* $\lambda_1 = 0$ *implies that* $\Gamma(\beta_1,\beta_2) \equiv 0$ *on* $\mathbf{B} \times \mathbf{B}$.

This theorem will be proved by converting the maximum eigenvalue problem involved in real terms as in the addendum to Bierens and Ploberger (1997) in Chapter 5 for real-valued kernels, using the properties of the real Hilbert space $L^2(\mu)$.

Definition 3.1. $L^2(\mu)$ is the Hilbert space of Borel measurable real functions f on $\mathbf{B}$ satisfying $\int f(\beta)^2\mathrm{d}\mu(\beta) < \infty$, endowed with innerproduct

$$\langle f, g\rangle = \int f(\beta)g(\beta)\mathrm{d}\mu(\beta)$$

and associated norm $||f|| = \sqrt{\langle f, f\rangle}$ and metric $||f-g||$.

Lemma 3.1. *The Hilbert space* $L^2(\mu)$ *has an orthonormal base, say* $\{\varphi_j(\beta)\}_{j=1}^{\infty}$, *so that every* $f \in L^2(\mu)$ *has the series representation*

$$f(\beta) = \sum_{j=1}^{\infty} c_j\varphi_j(\beta), \qquad (3.8)$$

where $c_j = \langle f, \varphi_j \rangle$ *satisfying* $\sum_{j=1}^{\infty} c_j^2 = \int f(\beta)^2 \mathrm{d}\mu(\beta) < \infty$.

Note that the series representation (3.8) holds with probability 1, in the sense that

$$\mu\left(\left\{\beta \in \mathbf{B} : \lim_{n\to\infty} \sum_{j=1}^{n} c_j \varphi_j(\beta) = f(\beta)\right\}\right) = 1,$$

rather than exactly for all $\beta \in \mathbf{B}$. C.f. Bierens (2004) and the addendum to Bierens and Ploberger (1997) in Chapter 5.

Since the solution ψ_1 of (3.6) is an element of $L^2_{\mathbb{C}}(\mu)$, and therefore $\mathrm{Re}[\psi_1]$ and $\mathrm{Im}[\psi_1]$ are elements of $L^2(\mu)$, ψ_1 has the series representation

$$\begin{aligned} \psi_1(\beta) &= \sum_{j=1}^{\infty} (c_j + \mathbf{i}.d_j)\, \varphi_j(\beta), \text{ where} \\ c_j &= \langle \mathrm{Re}[\psi_1], \varphi_j \rangle, \ d_j = \langle \mathrm{Im}[\psi_1], \varphi_j \rangle \\ \sum_{j=1}^{\infty} \left(c_j^2 + d_j^2\right) &= \int \psi_1(\beta)\overline{\psi_1(\beta)} \mathrm{d}\mu(\beta) = 1. \end{aligned} \tag{3.9}$$

In particular, denoting for $n \in \mathbb{N}$,

$$\begin{aligned} \psi_{1,n}(\beta) &= \sum_{j=1}^{n} (c_{n,j} + \mathbf{i}.d_{n,j})\, \varphi_j(\beta), \text{ where} \\ c_{n,j} &= \frac{c_j}{\sqrt{\sum_{i=1}^{n}(c_i^2 + d_i^2)}}, \ d_{n,j} = \frac{d_j}{\sqrt{\sum_{i=1}^{n}(c_i^2 + d_i^2)}}, \end{aligned} \tag{3.10}$$

it follows that

$$\lim_{n\to\infty} \int |\psi_1(\beta) - \psi_{1,n}(\beta)|^2 \mathrm{d}\mu(\beta) = 0, \tag{3.11}$$

as is not hard to verify. See the addendum to Bierens and Ploberger (1997) in Chapter 5 for a similar result. This implies the following result.

Lemma 3.2. *Let* ψ_1 *in* (3.9) *be a solution of* (3.6) *and let* $\psi_{1,n}$ *be defined by* (3.10). *Then*

$$\begin{aligned} &\lim_{n\to\infty} \int\int \overline{\psi_{1,n}(\beta_1)} \Gamma(\beta_1, \beta_2) \psi_{1,n}(\beta_2) \mathrm{d}\mu(\beta_1) \mathrm{d}\mu(\beta_2) \\ &\qquad = \int\int \overline{\psi_1(\beta_1)} \Gamma(\beta_1, \beta_2) \psi_1(\beta_2) \mathrm{d}\mu(\beta_1) \mathrm{d}\mu(\beta_2). \end{aligned}$$

Proof. To prove Lemma 3.2, observe first that

$$\begin{aligned}
&\Bigg|\int\int \overline{\psi_1(\beta_1)}\Gamma(\beta_1,\beta_2)\psi_1(\beta_2)\mathrm{d}\mu(\beta_1)\mathrm{d}\mu(\beta_2) \\
&\qquad\qquad - \int\int \overline{\psi_{1,n}(\beta_1)}\Gamma(\beta_1,\beta_2)\psi_{1,n}(\beta_2)\mathrm{d}\mu(\beta_1)\mathrm{d}\mu(\beta_2)\Bigg| \\
&\leq \Bigg|\int\int \left(\overline{\psi_1(\beta_1)} - \overline{\psi_{1,n}(\beta_1)}\right)\Gamma(\beta_1,\beta_2)\psi_1(\beta_2)\mathrm{d}\mu(\beta_1)\mathrm{d}\mu(\beta_2) \\
&\qquad\qquad + \int\int \overline{\psi_{1,n}(\beta_1)}\Gamma(\beta_1,\beta_2)\psi_1(\beta_2)\mathrm{d}\mu(\beta_1)\mathrm{d}\mu(\beta_2) \\
&\qquad\qquad - \int\int \overline{\psi_{1,n}(\beta_1)}\Gamma(\beta_1,\beta_2)\psi_{1,n}(\beta_2)\mathrm{d}\mu(\beta_1)\mathrm{d}\mu(\beta_2)\Bigg| \\
&\leq \left|\int \left(\overline{\psi_1(\beta_1)} - \overline{\psi_{1,n}(\beta_1)}\right)\left(\int \Gamma(\beta_1,\beta_2)\psi_1(\beta_2)\mathrm{d}\mu(\beta_2)\right)\mathrm{d}\mu(\beta_1)\right| \\
&\qquad + \left|\int \left(\int \overline{\psi_{1,n}(\beta_1)}\Gamma(\beta_1,\beta_2)\mathrm{d}\mu(\beta_1)\right)\left(\psi_1(\beta_2) - \psi_{1,n}(\beta_2)\right)\mathrm{d}\mu(\beta_2)\right| \\
&\leq \sqrt{\int \left|\overline{\psi_1(\beta_1)} - \overline{\psi_{1,n}(\beta_1)}\right|^2 \mathrm{d}\mu(\beta_1)} \\
&\qquad\qquad\qquad \times \sqrt{\int \left|\int \Gamma(\beta_1,\beta_2)\psi_1(\beta_2)\mathrm{d}\mu(\beta_2)\right|^2 \mathrm{d}\mu(\beta_1)} \\
&+ \sqrt{\int \left|\int \overline{\psi_{1,n}(\beta_1)}\Gamma(\beta_1,\beta_2)\mathrm{d}\mu(\beta_1)\right|^2 \mathrm{d}\mu(\beta_2)} \\
&\qquad\qquad\qquad \times \sqrt{\int \left|\psi_1(\beta_2) - \psi_{1,n}(\beta_2)\right|^2 \mathrm{d}\mu(\beta_2)} \\
&\leq 2 \sup_{(\beta_1,\beta_2)\in\mathbf{B}\times\mathbf{B}} \sqrt{|\Gamma(\beta_1,\beta_2)|} \times \sqrt{\int \left|\psi_1(\beta) - \psi_{1,n}(\beta)\right|^2 \mathrm{d}\mu(\beta)},
\end{aligned}$$

where the third inequality follows from the Cauchy-Schwartz inequality for inner products, $|\langle x, y\rangle| \leq ||x||.||y||$, which also holds for the complex case $\langle x, y\rangle = \overline{\langle y, x\rangle}$.

To prove the last inequality, observe that by the same Cauchy-Schwartz inequality,

$$\begin{aligned}
\left|\int \Gamma(\beta_1,\beta_2)\psi_1(\beta_2)\mathrm{d}\mu(\beta_2)\right| &\leq \sqrt{\int |\Gamma(\beta_1,\beta_2)|^2\mathrm{d}\mu(\beta_2)}\sqrt{\int |\psi_1(\beta_2)|^2\mathrm{d}\mu(\beta_2)} \\
&= \sqrt{\int |\Gamma(\beta_1,\beta_2)|^2\mathrm{d}\mu(\beta_2)} \\
&\leq \sup_{(\beta_1,\beta_2)\in\mathbf{B}\times\mathbf{B}} |\Gamma(\beta_1,\beta_2)| < \infty
\end{aligned}$$

where the last inequality follows from the uniform continuity of $\Gamma(\beta_1,\beta_2)$ on $\mathbf{B}\times\mathbf{B}$. Hence,

$$\sqrt{\int \left|\int \Gamma(\beta_1,\beta_2)\psi_1(\beta_2)\mathrm{d}\mu(\beta_2)\right|^2 \mathrm{d}\mu(\beta_1)} \leq \sup_{(\beta_1,\beta_2)\in\mathbf{B}\times\mathbf{B}} \sqrt{|\Gamma(\beta_1,\beta_2)|} < \infty$$

and similarly

$$\sqrt{\int \left|\int \overline{\psi_{1,n}(\beta_1)}\Gamma(\beta_1,\beta_2)\mathrm{d}\mu(\beta_1)\right|^2 \mathrm{d}\mu(\beta_2)} \leq \sup_{(\beta_1,\beta_2)\in\mathbf{B}\times\mathbf{B}} \sqrt{|\Gamma(\beta_1,\beta_2)|} < \infty.$$

Lemma 3.2 follows now from (3.11). ■

The following lemma is also a well-known result, related to Lemma 3.1.

Lemma 3.3. *Given the orthonormal base $\{\varphi_j(\beta)\}_{j=1}^{\infty}$ of $L^2(\mu)$ in Lemma 3.1, every Borel measurable real function $g(\beta_1,\beta_2)$ on $\mathbf{B}\times\mathbf{B}$ satisfying*

$$\int\int g(\beta_1,\beta_2)^2\mathrm{d}\mu(\beta_1)\mathrm{d}\mu(\beta_2) < \infty$$

has the series representation

$$g(\beta_1,\beta_2) = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} c_{i,j}\varphi_i(\beta_1)\varphi_j(\beta_2), \tag{3.12}$$

where

$$\begin{aligned}
c_{i,j} &= \int\int \varphi_i(\beta_1)g(\beta_1,\beta_2)\varphi_j(\beta_2)\mathrm{d}\mu(\beta_1)\mathrm{d}\mu(\beta_2), \text{ with} \\
\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} c_{i,j}^2 &= \int\int g(\beta_1,\beta_2)^2\mathrm{d}\mu(\beta_1)\mathrm{d}\mu(\beta_2) < \infty.
\end{aligned}$$

Similar to Lemma 3.1 the series representation (3.12) holds with probability 1, in the sense that

$$\mu\times\mu\left(\left\{(\beta_1,\beta_2)\in\mathbf{B}\times\mathbf{B}:\lim_{\min(n_1,n_2)\to\infty}\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}c_{i,j}\varphi_i(\beta_1)\varphi_j(\beta_2)=g(\beta_1,\beta_2)\right\}\right)=1$$

rather than exactly for all $(\beta_1,\beta_2)\in\mathbf{B}\times\mathbf{B}$, where $\mu\times\mu$ is the product measure defined as follows.

Definition 3.2. Let $\widetilde{\beta}_1$ and $\widetilde{\beta}_2$ be independent random drawings from the distribution of μ. Then the product measure $\mu\times\mu$ is the probability measure on $\mathbf{B}\times\mathbf{B}$ induced by $(\widetilde{\beta}_1,\widetilde{\beta}_2)$.

Lemma 3.3 implies that

$$\begin{aligned}
\operatorname{Re}[\Gamma(\beta_1,\beta_2)] &= \sum_{i=1}^{\infty}\sum_{j=1}^{\infty}\alpha_{i,j}\varphi_i(\beta_1)\varphi_j(\beta_2),\ \text{where}\\
\alpha_{i,j} &= \int\int\varphi_i(\beta_1)\operatorname{Re}[\Gamma(\beta_1,\beta_2)]\varphi_j(\beta_2)\mathrm{d}\mu(\beta_1)\mathrm{d}\mu(\beta_2),\\
\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}\alpha_{i,j}^2 &= \int\int\left(\operatorname{Re}[\Gamma(\beta_1,\beta_2)]\right)^2\mathrm{d}\mu(\beta_1)\mathrm{d}\mu(\beta_2)<\infty. \qquad (3.13)\\
\operatorname{Im}[\Gamma(\beta_1,\beta_2)] &= \sum_{i=1}^{\infty}\sum_{j=1}^{\infty}\gamma_{i,j}\varphi_i(\beta_1)\varphi_j(\beta_2),\ \text{where}\\
\gamma_{i,j} &= \int\int\varphi_i(\beta_1)\operatorname{Im}[\Gamma(\beta_1,\beta_2)]\varphi_j(\beta_2)\mathrm{d}\mu(\beta_1)\mathrm{d}\mu(\beta_2),\\
\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}\gamma_{i,j}^2 &= \int\int\left(\operatorname{Im}[\Gamma(\beta_1,\beta_2)]\right)^2\mathrm{d}\mu(\beta_1)\mathrm{d}\mu(\beta_2)<\infty. \qquad (3.14)
\end{aligned}$$

Note that by (2.3) and (2.4),

$$\alpha_{i,j}=\alpha_{j,i},\ \gamma_{i,j}=-\gamma_{j,i},\ \gamma_{i,i}=0. \qquad (3.15)$$

Hence,

$$\Gamma(\beta_1,\beta_2) = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty}\left(\alpha_{i,j}+\mathbf{i}.\gamma_{i,j}\right)\varphi_i(\beta_1)\varphi_j(\beta_2). \qquad (3.16)$$

$$
\begin{aligned}
\Gamma_2(\beta_1,\beta_2) &= \int \Gamma(\beta_1,\beta)\Gamma(\beta,\beta_2)\mathrm{d}\mu(\beta) \\
&= \int \left(\sum_{i_1=1}^{\infty}\sum_{j_1=1}^{\infty}\left(\alpha_{i_1,j_1}+\mathbf{i}.\gamma_{i_1,j_1}\right)\varphi_{i_1}(\beta_1)\varphi_{j_1}(\beta)\right) \\
&\quad \times \left(\sum_{j_2=1}^{\infty}\sum_{i_2=1}^{\infty}\left(\alpha_{j_2,i_2}+\mathbf{i}.\gamma_{j_2,i_2}\right)\varphi_{j_2}(\beta)\varphi_{i_2}(\beta_2)\right)\mathrm{d}\mu(\beta) \\
&= \sum_{i_1=1}^{\infty}\sum_{j_1=1}^{\infty}\sum_{j_2=1}^{\infty}\sum_{i_2=1}^{\infty}\left(\alpha_{i_1,j_1}+\mathbf{i}.\gamma_{i_1,j_1}\right)\left(\alpha_{j_2,i_2}+\mathbf{i}.\gamma_{j_2,i_2}\right)\varphi_{i_1}(\beta_1)\varphi_{i_2}(\beta_2) \\
&\quad \times \int \varphi_{j_1}(\beta)\varphi_{j_2}(\beta)\mathrm{d}\mu(\beta) \\
&= \sum_{i_1=1}^{\infty}\sum_{i_2=1}^{\infty}\left(\sum_{j=1}^{\infty}\left(\alpha_{i_1,j}+\mathbf{i}.\gamma_{i_1,j}\right)\left(\alpha_{j,i_2}+\mathbf{i}.\gamma_{j,i_2}\right)\right)\varphi_{i_1}(\beta_1)\varphi_{i_2}(\beta_2) \\
&= \sum_{i_1=1}^{\infty}\sum_{i_2=1}^{\infty}\sum_{j=1}^{\infty}\left(\alpha_{i_1,j}\alpha_{j,i_2}-\gamma_{i_1,j}\gamma_{j,i_2}\right)\varphi_{i_1}(\beta_1)\varphi_{i_2}(\beta_2) \\
&\quad +\mathbf{i}.\sum_{i_1=1}^{\infty}\sum_{i_2=1}^{\infty}\sum_{j=1}^{\infty}\left(\alpha_{i_1,j}\gamma_{j,i_2}+\gamma_{i_1,j}\alpha_{j,i_2}\right)\varphi_{i_1}(\beta_1)\varphi_{i_2}(\beta_2). \qquad (3.17)
\end{aligned}
$$

Combining (3.9) and (3.16) it follows that

$$
\begin{aligned}
&\int\int \overline{\psi_1(\beta_1)}\Gamma(\beta_1,\beta_2)\psi_1(\beta_2)\mathrm{d}\mu(\beta_1)\mathrm{d}\mu(\beta_2) \\
&= \int\int \left(\sum_{m=1}^{\infty}\left(c_m-\mathbf{i}.d_m\right)\varphi_m(\beta_1)\right)\left(\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}\left(\alpha_{i,j}+\mathbf{i}.\gamma_{i,j}\right)\varphi_i(\beta_1)\varphi_j(\beta_2)\right) \\
&\times \left(\sum_{k=1}^{\infty}\left(c_k+\mathbf{i}.d_k\right)\varphi_k(\beta_2)\right)\mathrm{d}\mu(\beta_1)\mathrm{d}\mu(\beta_2) \\
&= \sum_{m=1}^{\infty}\sum_{k=1}^{\infty}\left(c_m-\mathbf{i}.d_m\right)\left(c_k+\mathbf{i}.d_k\right)\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}\left(\alpha_{i,j}+\mathbf{i}.\gamma_{i,j}\right) \\
&\times \int \varphi_i(\beta_1)\varphi_m(\beta_1)\mathrm{d}\mu(\beta_1)\int \varphi_j(\beta_2)\varphi_k(\beta_2)\mathrm{d}\mu(\beta_2) \\
&= \sum_{m=1}^{\infty}\sum_{k=1}^{\infty}\left(c_m-\mathbf{i}.d_m\right)\left(c_k+\mathbf{i}.d_k\right)\left(\alpha_{m,k}+\mathbf{i}.\gamma_{m,k}\right)
\end{aligned}
$$

$$
\begin{aligned}
&= \sum_{m=1}^{\infty}\sum_{k=1}^{\infty} \left((c_m c_k + d_m d_k) + \mathbf{i}.(c_m d_k - d_m c_k)\right)(\alpha_{m,k} + \mathbf{i}.\gamma_{m,k}) \\
&= \sum_{m=1}^{\infty}\sum_{k=1}^{\infty} c_m \alpha_{m,k} c_k + \sum_{m=1}^{\infty}\sum_{k=1}^{\infty} d_m \alpha_{m,k} d_k - \sum_{m=1}^{\infty}\sum_{k=1}^{\infty} c_m \gamma_{m,k} d_k \\
&\qquad\qquad + \sum_{m=1}^{\infty}\sum_{k=1}^{\infty} d_m \gamma_{m,k} c_k \\
&+\mathbf{i}.\sum_{m=1}^{\infty}\sum_{k=1}^{\infty} (c_m \alpha_{m,k} d_k - d_m \alpha_{m,k} c_k) + \mathbf{i}.\sum_{m=1}^{\infty}\sum_{k=1}^{\infty} (c_m \gamma_{m,k} c_k + d_m \gamma_{m,k} d_k) \\
&= \sum_{m=1}^{\infty}\sum_{k=1}^{\infty} c_m \alpha_{m,k} c_k + \sum_{m=1}^{\infty}\sum_{k=1}^{\infty} d_m \alpha_{m,k} d_k - 2\sum_{m=1}^{\infty}\sum_{k=1}^{\infty} c_m \gamma_{m,k} d_k \\
&= \sum_{m=1}^{\infty}\sum_{k=1}^{\infty} (c_m, d_m) \begin{pmatrix} \alpha_{m,k} & -\gamma_{m,k} \\ \gamma_{k,m} & \alpha_{m,k} \end{pmatrix} \begin{pmatrix} c_k \\ d_k \end{pmatrix},
\end{aligned}
$$

where the last two equalities follows from the fact that by (3.15),

$$
\begin{aligned}
\sum_{m=1}^{\infty}\sum_{k=1}^{\infty} c_m \alpha_{m,k} d_k &= \sum_{m=1}^{\infty}\sum_{k=1}^{\infty} d_m \alpha_{m,k} c_k, \\
\sum_{m=1}^{\infty}\sum_{k=1}^{\infty} d_m \gamma_{m,k} c_k &= -\sum_{m=1}^{\infty}\sum_{k=1}^{\infty} c_m \gamma_{m,k} c_k, \\
\sum_{m=1}^{\infty}\sum_{k=1}^{\infty} c_m \gamma_{m,k} c_k &= 0, \ \sum_{m=1}^{\infty}\sum_{k=1}^{\infty} d_m \gamma_{m,k} d_k = 0,
\end{aligned}
$$

and from the easy equality

$$
\begin{aligned}
&\sum_{m=1}^{\infty}\sum_{k=1}^{\infty} c_m \alpha_{m,k} c_k + \sum_{m=1}^{\infty}\sum_{k=1}^{\infty} d_m \alpha_{m,k} d_k - 2\sum_{m=1}^{\infty}\sum_{k=1}^{\infty} c_m \gamma_{m,k} d_k \\
&= \sum_{m=1}^{\infty}\sum_{k=1}^{\infty} (c_m, d_m) \begin{pmatrix} \alpha_{m,k} & -\gamma_{m,k} \\ \gamma_{k,m} & \alpha_{m,k} \end{pmatrix} \begin{pmatrix} c_k \\ d_k \end{pmatrix}.
\end{aligned}
$$

Thus, denoting

$$
A_{m,k} = \begin{pmatrix} \alpha_{m,k} & -\gamma_{m,k} \\ \gamma_{k,m} & \alpha_{m,k} \end{pmatrix}, \ b_m = (c_m, d_m)' \tag{3.18}
$$

we have

$$\int\int\overline{\psi_1(\beta_1)}\Gamma(\beta_1,\beta_2)\psi_1(\beta_2)\mathrm{d}\mu(\beta_1)\mathrm{d}\mu(\beta_2)=\sum_{m=1}^{\infty}\sum_{k=1}^{\infty}b_m'A_{m,k}b_k.$$

Similarly,

$$\int\int\overline{\psi_{1,n}(\beta_1)}\Gamma(\beta_1,\beta_2)\psi_{1,n}(\beta_2)\mathrm{d}\mu(\beta_1)\mathrm{d}\mu(\beta_2)=\sum_{m=1}^{n}\sum_{k=1}^{n}b_{n,m}'A_{m,k}b_{n,k}, \tag{3.19}$$

where $b_{n,m}=(c_{n,m},d_{n,m})'$. C.f. (3.10).

Now stack the $b_{n,m}$'s in an $2n\times 1$ vector x_n, and recall from (3.10) that $x_n'x_n=1$. Moreover, denote

$$A_n=\begin{pmatrix} A_{1,1} & A_{1,2} & \cdots & A_{1,n-1} & A_{1,n}\\ A_{2,1} & A_{2,2} & \cdots & A_{2,n-1} & A_{2,n}\\ \vdots & \vdots & \ddots & \vdots & \vdots\\ A_{n-1,1} & A_{n-1,2} & \cdots & A_{n-1,n-1} & A_{n-1,n}\\ A_{n,1} & A_{n,2} & \cdots & A_{n,n-1} & A_{n,n}\end{pmatrix}, \tag{3.20}$$

which is a symmetric $2n\times 2n$ matrix.[2] Then (3.19) reads

$$\int\int\overline{\psi_{1,n}(\beta_1)}\Gamma(\beta_1,\beta_2)\psi_{1,n}(\beta_2)\mathrm{d}\mu(\beta_1)\mathrm{d}\mu(\beta_2)=x_n'A_nx_n, \tag{3.21}$$

whereas obviously,

$$x_n'A_nx_n\le\int\int\overline{\psi_1(\beta_1)}\Gamma(\beta_1,\beta_2)\psi_1(\beta_2)\mathrm{d}\mu(\beta_1)\mathrm{d}\mu(\beta_2). \tag{3.22}$$

Similarly, replacing $\psi_{1,n}(\beta)$ by $\psi_n(\beta)=\sum_{j=1}^n(y_{1,j}+\mathbf{i}.y_{2,j})\varphi_j(\beta)$, where the $y_{i,j}$'s are arbitrary subject to the restriction $\sum_{j=1}^n y_{1,j}^2+\sum_{j=1}^n y_{2,j}^2=1$, and denoting $y_n=(y_{1,1},y_{2,1},y_{1,2},y_{2,2},...,y_{1,n},y_{2,n})$, we have

$$\sup_{y_n\in\mathbb{R}^{2n}.:y_n'y_n=1}y_n'A_ny_n\le\int\int\overline{\psi_1(\beta_1)}\Gamma(\beta_1,\beta_2)\psi_1(\beta_2)\mathrm{d}\mu(\beta_1)\mathrm{d}\mu(\beta_2). \tag{3.23}$$

[2] The symmetry follows from

$$A_{k,m}=\begin{pmatrix}\alpha_{k,m} & -\gamma_{k,m}\\ \gamma_{k,m} & \alpha_{k,m}\end{pmatrix}=\begin{pmatrix}\alpha_{m,k} & \gamma_{m,k}\\ -\gamma_{m,k} & \alpha_{m,k}\end{pmatrix}=A_{m,k}'.$$

Recall from linear algebra that the maximum eigenvalue $\overline{\lambda}_n$ of A_n is equal to

$$\overline{\lambda}_n = \sup_{y \in \mathbb{R}^{2n}: y'y=1} y'A_n y, \tag{3.24}$$

so that by (3.21), (3.22) and (3.23),

$$\begin{aligned} &\int\int \overline{\psi_{1,n}(\beta_1)}\Gamma(\beta_1,\beta_2)\psi_{1,n}(\beta_2)\mathrm{d}\mu(\beta_1)\mathrm{d}\mu(\beta_2) \\ &\qquad \le \overline{\lambda}_n \le \int\int \overline{\psi_1(\beta_1)}\Gamma(\beta_1,\beta_2)\psi_1(\beta_2)\mathrm{d}\mu(\beta_1)\mathrm{d}\mu(\beta_2). \end{aligned} \tag{3.25}$$

Consequently, it follows from Lemma 3.2 that

Lemma 3.4. $\int\int \overline{\psi_1(\beta_1)}\Gamma(\beta_1,\beta_2)\psi_1(\beta_2)\mathrm{d}\mu(\beta_1)\mathrm{d}\mu(\beta_2) = \lim_{n\to\infty}\overline{\lambda}_n$, *where* $\overline{\lambda}_n$ *is the maximum eigenvalue of the matrix* A_n *in* (3.20).

Next, let us focus on the case $\Gamma_2(\beta_1,\beta_2)$. Obviously, Lemma 3.2 carries over if we replace Γ by Γ_2.

Lemma 3.5. *Let* ψ_1 *in* (3.9) *be a solution of* (3.6) *and let* $\psi_{1,n}$ *be defined by* (3.10). *Then*

$$\begin{aligned} &\lim_{n\to\infty}\int\int \overline{\psi_{1,n}(\beta_1)}\Gamma_2(\beta_1,\beta_2)\psi_{1,n}(\beta_2)\mathrm{d}\mu(\beta_1)\mathrm{d}\mu(\beta_2) \\ &\qquad = \int\int \overline{\psi_1(\beta_1)}\Gamma_2(\beta_1,\beta_2)\psi_1(\beta_2)\mathrm{d}\mu(\beta_1)\mathrm{d}\mu(\beta_2). \end{aligned}$$

Moreover, after some tedious complex calculus exercises it can be shown from (3.17) and (3.9) that

$$\begin{aligned} &\int\int \overline{\psi_1(\beta_1)}\Gamma_2(\beta_1,\beta_2)\psi_1(\beta_2)\mathrm{d}\mu(\beta_1)\mathrm{d}\mu(\beta_2) \\ &= \sum_{k=1}^{\infty}\sum_{m=1}^{\infty}\sum_{j=1}^{\infty}(c_k,d_k) \\ &\quad \times \begin{pmatrix} \alpha_{k,j}\alpha_{j,m}-\gamma_{k,j}\gamma_{j,m} & -\alpha_{k,j}\gamma_{j,m}-\gamma_{k,j}\alpha_{j,m} \\ \alpha_{k,j}\gamma_{j,m}+\gamma_{k,j}\alpha_{j,m} & \alpha_{k,j}\alpha_{j,m}-\gamma_{k,j}\gamma_{j,m} \end{pmatrix}\begin{pmatrix} c_m \\ d_m \end{pmatrix} \\ &\qquad = \sum_{k=1}^{\infty}\sum_{m=1}^{\infty} b_k'\left(\sum_{j=1}^{\infty}A_{k,j}A_{j,m}\right)b_m, \end{aligned}$$

where $A_{k,j}$, $A_{j,m}$ and b_m are defined in (3.18). Furthermore, similar to (3.19) we have

$$\int\int \overline{\psi_{1,n}(\beta_1)}\Gamma_2(\beta_1,\beta_2)\psi_{1,n}(\beta_2)\mathrm{d}\mu(\beta_1)\mathrm{d}\mu(\beta_2) = \sum_{k=1}^{n}\sum_{m=1}^{n} b'_{n,k}\left(\sum_{j=1}^{\infty} A_{k,j}A_{j,m}\right) b_{n,m}, \qquad (3.26)$$

where $b_{n,m} = (c_{n,m}, d_{n,m})'$. C.f. (3.10).

Stacking the $b_{n,m}$'s in a $2n \times 1$ vector x_n as before, and denoting

$$C_n = \begin{pmatrix} C_{1,1}(n) & \cdots & C_{1,n}(n) \\ \vdots & \ddots & \vdots \\ C_{n,1}(n) & \cdots & C_{n,n}(n) \end{pmatrix}, \text{ where } C_{k,m}(n) = \sum_{j=n+1}^{\infty} A_{k,j}A_{j,m},$$

we can write the right-hand side of (3.26) as

$$\sum_{k=1}^{n}\sum_{m=1}^{n} b'_{n,k}\left(\sum_{j=1}^{n} A_{k,j}A_{j,m}\right) b_{n,m} + \sum_{k=1}^{n}\sum_{m=1}^{n} b'_{n,k}\left(\sum_{j=n+1}^{\infty} A_{k,j}A_{j,m}\right) b_{n,m}$$
$$= x'_n A_n^2 x_n + x'_n C_n x_n.$$

Because $x'_n x_n = 1$ the term $x'_n A_n^2 x_n$ is dominated by the maximum eigenvalue of A_n^2, which is the square of the maximum eigenvalue $\overline{\lambda}_n$ of A_n, and $x'_n C_n x_n$ is dominated by the trace of C_n, where

$$\begin{aligned} \mathrm{trace}(C_n) &= \sum_{m=1}^{n} \mathrm{trace}\left(C_{m,m}(n)\right) \\ &= 2\sum_{m=1}^{n}\sum_{j=n+1}^{\infty} \alpha_{m,j}\alpha_{j,m} - 2\sum_{m=1}^{n}\sum_{j=n+1}^{\infty} \gamma_{m,j}\gamma_{j,m} \\ &= 2\sum_{m=1}^{n}\sum_{j=n+1}^{\infty} \left(\alpha_{m,j}^2 + \gamma_{m,j}^2\right) \\ &\leq 2\left(\sum_{j=n+1}^{\infty}\sum_{m=1}^{\infty} \alpha_{m,j}^2 + \sum_{j=n+1}^{\infty}\sum_{m=1}^{\infty} \gamma_{m,j}^2\right). \end{aligned}$$

Due to (3.13) and (3.14) the latter converges to zero as $n \to \infty$.

Thus, it has been shown that

$$\int\int \overline{\psi_{1,n}(\beta_1)}\Gamma_2(\beta_1,\beta_2)\psi_{1,n}(\beta_2)\mathrm{d}\mu(\beta_1)\mathrm{d}\mu(\beta_2) \leq \overline{\lambda}_n^2 + o(1).$$

It follows now from Lemmas 3.4 and 3.5 that

$$\int\int \overline{\psi_1(\beta_1)}\Gamma_2(\beta_1,\beta_2)\psi_1(\beta_2)\mathrm{d}\mu(\beta_1)\mathrm{d}\mu(\beta_2) \\ \leq \left(\int\int \overline{\psi_1(\beta_1)}\Gamma(\beta_1,\beta_2)\psi_1(\beta_2)\mathrm{d}\mu(\beta_1)\mathrm{d}\mu(\beta_2)\right)^2. \quad (3.27)$$

However, note that the function $\underline{G}(\psi)$ in (3.5) is always nonnegative, which implies that

$$\int\int \overline{\psi_1(\beta_1)}\Gamma_2(\beta_1,\beta_2)\psi_1(\beta_2)\mathrm{d}\mu(\beta_1)\mathrm{d}\mu(\beta_2) \\ \geq \left(\int\int \overline{\psi_1(\beta_1)}\Gamma(\beta_1,\beta_2)\psi_1(\beta_2)\mathrm{d}\mu(\beta_1)\mathrm{d}\mu(\beta_2)\right)^2. \quad (3.28)$$

Part (a) of Theorem 3.1 now follows from (3.27) and (3.28).

Finally, observe from (3.24) and (3.25) that the sequence $\overline{\lambda}_n$ is monotonic non-decreasing and bounded, hence $\lim_{n\to\infty}\overline{\lambda}_n = \sup_{n\geq 1}\overline{\lambda}_n$. If the latter supremum is zero, then for all $n \geq 1$, $A_n = O_{2n,2n}$, hence the $\alpha_{i,j}$'s and $\gamma_{i,j}$'s are all zero and thus by (3.16), $\Gamma(\beta_1,\beta_2) \equiv 0$ on $\mathbf{B}\times\mathbf{B}$. This proves parts (b) and (c) of Theorem 3.1. ■

3.3. The other eigenvalues and eigenfunctions

Given that $\lambda_1 > 0$, let

$$\Gamma^{(2)}(\beta_1,\beta_2) = \Gamma(\beta_1,\beta_2) - \lambda_1\psi_1(\beta_1)\overline{\psi_1(\beta_2)},$$

which is symmetric and continuous on $\mathbf{B}\times\mathbf{B}$. To prove that $\Gamma^{(2)}(\beta_1,\beta_2)$ is positive semidefinite, let $\phi \in L^2_{\mathbb{C}}(\mu)$ be arbitrary. We can write $\phi = \langle\phi,\psi_1\rangle\,\psi_1 + r$, where $\langle r,\psi_1\rangle = 0$, hence

$$\int\int \phi(\beta_1)\Gamma^{(2)}(\beta_1,\beta_2)\overline{\phi(\beta_2)}\mathrm{d}\mu(\beta_1)\mathrm{d}\mu(\beta_2) \\ = \int\int r(\beta_1)\Gamma(\beta_1,\beta_2)\overline{r(\beta_2)}\mathrm{d}\mu(\beta_1)\mathrm{d}\mu(\beta_2) \geq 0,$$

as is not hard to verify. Then the maximum eigenvalue λ_2 of $\Gamma^{(2)}(\beta_1,\beta_2)$ with corresponding eigenfunction ψ_2 can be derived in the same way as before, which are an eigenvalue and corresponding eigenfunction of $\Gamma(\beta_1,\beta_2)$ as well, with

$$\int \overline{\psi_1(\beta)}\psi_2(\beta)\mathrm{d}\mu(\beta) = 0 \text{ if } \lambda_2 > 0.$$

To see this, note that $\int \overline{\psi_1(\beta_1)}\Gamma^{(2)}(\beta_1,\beta_2)\mathrm{d}\mu(\beta_1) = 0$, hence

$$\begin{aligned}\lambda_2 \int \overline{\psi_1(\beta_1)}\psi_2(\beta_1)\mathrm{d}\mu(\beta_1) &= \int\int \overline{\psi_1(\beta_1)}\Gamma^{(2)}(\beta_1,\beta_2)\psi_2(\beta_2)\mathrm{d}\mu(\beta_1)\mathrm{d}\mu(\beta_2) \\ &= 0.\end{aligned}$$

Now

$$\begin{aligned}\lambda_2\psi_2(\beta_1) &= \int \Gamma^{(2)}(\beta_1,\beta_2)\psi_2(\beta_2)\mathrm{d}\mu(\beta_2) \\ &= \int \Gamma(\beta_1,\beta_2)\psi_2(\beta_2)\mathrm{d}\mu(\beta_2) - \lambda_1\psi_1(\beta_1)\int \overline{\psi_1(\beta_2)}\psi_2(\beta_2)\mathrm{d}\mu(\beta_2) \\ &= \int \Gamma(\beta_1,\beta_2)\psi_2(\beta_2)\mathrm{d}\mu(\beta_2).\end{aligned}$$

Repeating this construction n times yield eigenvalues $\lambda_1 \geq \lambda_2 \geq .... \geq \lambda_n$ of $\Gamma(\beta_1,\beta_2)$ with corresponding orthonormal eigenfunctions ψ_m, $m = 1,2,...,n$. At this point the next eigenvalue $\lambda_{n+1} \leq \lambda_n$ and eigenfunction ψ_{n+1} are the maximum eigenvalue and corresponding eigenfunction of the kernel

$$\Gamma^{(n)}(\beta_1,\beta_2) = \Gamma(\beta_1,\beta_2) - \sum_{m=1}^{n} \lambda_m\psi_m(\beta_1)\overline{\psi_m(\beta_2)}. \tag{3.29}$$

Suppose that for some n, $\lambda_n > 0$ but $\lambda_{n+1} = 0$. Then the maximum eigenvalue of $\Gamma^{(n)}(\beta_1,\beta_2)$ is zero, hence by part (c) of Theorem 3.1, $\Gamma^{(n)}(\beta_1,\beta_2) \equiv 0$ on $\mathbf{B}\times\mathbf{B}$ and thus

$$\Gamma(\beta_1,\beta_2) \equiv \sum_{m=1}^{n} \lambda_m\psi_m(\beta_1)\overline{\psi_m(\beta_2)}.$$

Then any function ϕ in the orthogonal complement of $\mathrm{span}(\{\psi_m\}_{m=1}^{n})$, i.e., the space

$$\mathcal{U}_n = \left\{\phi \in L^2_{\mathbb{C}}(\mu) : \langle\phi,\psi_m\rangle = 0 \text{ for } m = 1,2,...,n\right\},$$

is an eigenfunction of $\Gamma(\beta_1, \beta_2)$ with zero eigenvalue. Since $\mathcal{U}_n$ is a Hilbert space itself contained in $L^2_{\mathbb{C}}(\mu)$, it is possible to select an orthonormal basis $\{\psi_m\}_{m=n+1}^{\infty}$ for $\mathcal{U}_n$, and the extended orthonormal sequence $\{\psi_m\}_{m=1}^{\infty}$ is then an orthonormal basis of $L^2_{\mathbb{C}}(\mu)$, i.e.,

$$L^2_{\mathbb{C}}(\mu) = \text{span}(\{\psi_m\}_{m=1}^{\infty}). \tag{3.30}$$

If there does not exists an n such that $\lambda_n = 0$ then we can repeat this construction indefinitely, yielding a non-increasing sequence $\{\lambda_n\}_{n=1}^{\infty}$ of positive eigenvalues of $\Gamma(\beta_1, \beta_2)$ with corresponding orthonormal sequence $\{\psi_m\}_{m=1}^{\infty}$ of eigenfunctions. However, that does not mean that $\Gamma(\beta_1, \beta_2)$ has only positive eigenvalues. If the Hilbert space $\mathcal{U}_\infty = \{\phi \in L^2_{\mathbb{C}}(\mu) : \langle \phi, \psi_m \rangle = 0 \text{ for all } m \in \mathbb{N}\}$ is non-trivial, in the sense that it is larger than the singleton $\{0\}$, then by Mercer's theorem below, all the function in $\mathcal{U}_\infty$ are eigenfunctions of $\Gamma(\beta_1, \beta_2)$ with zero eigenvalues.

3.4. The Hilbert-Schmidt theorem

Summarizing, the following main result has been proved.

Theorem 3.2. (Hilbert-Schmidt Theorem) *Under Assumption 2.1 the eigenvalue problem: "Find a scalar λ and a function $\psi \in L^2_{\mathbb{C}}(\mu)$ normalized to unit norm, $||\psi|| = 1$, such that $\int \Gamma(\beta_1, \beta_2)\psi(\beta_2)\mathrm{d}\mu(\beta_2) = \lambda\psi(\beta_1)$ for all $\beta_1 \in \mathbf{B}$" has countable many solutions $\{\lambda_m, \psi_m\}_{m=1}^{\infty}$, i.e.,*[3]

$$\int \Gamma(\beta_1, \beta_2)\psi_m(\beta_2)\mathrm{d}\mu(\beta_2) \equiv \lambda_m\psi_m(\beta_1), \tag{3.31}$$

where the eigenvalues λ_m are real valued and nonnegative and the eigenfunctions ψ_m are orthonormal. Moreover, the eigenfunctions corresponding to the positive eigenvalues are continuous on $\mathbf{B}$. If all the eigenvalues are zero then $\Gamma(\beta_1, \beta_2) \equiv 0$ on $\mathbf{B} \times \mathbf{B}$.

This theorem is called after David Hilbert and his Ph.D. student Erhard Schmidt who published a series of papers in the period 1904-1908 regarding the existence of eigenvalues and corresponding eigenfunctions for real-valued kernels on a rectangle $[a, b] \times [a, b]$. Their results are nowadays referred to as the Hilbert-Schmidt theorem. See for example Bernkopf (1966) and Siegmund-Schultze (1986) and the references therein.

[3] In Lemma 3 in BW, (3.31) was incorrectly stated as $\lambda_m\psi_m(\beta_1) = \int \Gamma(\beta_1, \beta_2)\overline{\psi_m(\beta_2)}\mathrm{d}\mu(\beta_2)$.

The complex case in Theorem 3.2 is not a new results, of course. See for example Krein (1998). However, the proof of Theorem 3.2 is different from what I have seen in the literature, where the proof is based on operator theory.

4. Mercer's Theorem for Complex Kernels

The original Mercer's theorem is due to Mercer (1909), who proved it for real-valued kernels on the rectangle $[a,b]\times[a,b]$, with μ the Lebesgue measure.

The current version of Mercer' theorem can be restated somewhat shorter than in Lemma 3 in BW because some parts are already covered by Theorem 3.2.

Theorem 4.1. *(*Mercer's Theorem*) Under Assumptions* 2.1 *and* 2.2 *the complex kernel* $\Gamma(\beta_1,\beta_2)$ *involved has the series representation*

$$\Gamma(\beta_1,\beta_2)=\sum_{m=1}^{\infty}\lambda_m\psi_m(\beta_1)\overline{\psi_m(\beta_2)}, \tag{4.1}$$

where the λ_m*'s are the eigenvalues of* $\Gamma(\beta_1,\beta_2)$ *and the* ψ_m*'s are the corresponding orthonormal eigenfunctions. Then in addition to the results in Theorem 3.2 the following hold.*
(a) The eigenvalues satisfy $\sum_{m=1}^{\infty}\lambda_m<\infty$.
(b) The convergence of the right-hand side of (4.1) *is uniform on* $\mathbf{B}\times\mathbf{B}$, *i.e.,*

$$\lim_{n\to\infty}\sup_{(\beta_1,\beta_2)\in\mathbf{B}\times\mathbf{B}}\left|\Gamma(\beta_1,\beta_2)-\sum_{m=1}^{n}\lambda_m\psi_m(\beta_1)\overline{\psi_m(\beta_2)}\right|=0.$$

(c) The orthonormal sequence $\{\psi_m\}_{m=1}^{\infty}$ *of eigenfunctions, including the eigenfunctions with zero eigenvalues, is complete in* $L^2_{\mathbb{C}}(\mu)$, *i.e.,* $L^2_{\mathbb{C}}(\mu)=\text{span}(\{\psi_m\}_{m=1}^{\infty})$.

Proof. Let $\{\psi_m\}_{m=1}^{\infty}$ be the sequence of *all* eigenfunctions, thus including those with zero eigenvalues. Denote

$$S_2=\text{span}(\{\psi_k(\beta_1)\psi_m(\beta_2)\}_{k,m=1}^{\infty}),$$

which is a subspace of the Hilbert space $L^2_{\mathbb{C}}(\mu\times\mu)$ of all square-integrable complex-valued Borel measurable functions on $\mathbf{B}\times\mathbf{B}$ endowed with the usual innerproduct and associated norm and metric, and note that $\Gamma \in L^2_{\mathbb{C}}(\mu \times \mu)$. By the projection theorem, the projection of Γ on S_2 takes the form

$$\begin{aligned}
\underline{\Gamma}(\beta_1, \beta_2) &= \sum_{m=1}^{\infty}\sum_{k=1}^{\infty} c_{k,m}\psi_m(\beta_1)\overline{\psi_k(\beta_2)}, \text{ where} \\
c_{k,m} &= \int\int \overline{\psi_m(\beta_1)}\Gamma(\beta_1, \beta_2)\psi_k(\beta_2)\mathrm{d}\mu(\beta_1)\mathrm{d}\mu(\beta_2) \\
&= \lambda_k \int \overline{\psi_m(\beta_1)}\psi_k(\beta_1)\mathrm{d}\mu(\beta_1) = \lambda_k \mathbf{1}(k = m),
\end{aligned}$$

where, as in BW, $\mathbf{1}(.)$ denotes the indicator function. Hence, the projection of Γ on S_2 is

$$\underline{\Gamma}(\beta_1, \beta_2) = \sum_{m=1}^{\infty} \lambda_m \psi_m(\beta_1)\overline{\psi_m(\beta_2)},$$

with projection residual

$$R(\beta_1, \beta_2) = \Gamma(\beta_1, \beta_2) - \underline{\Gamma}(\beta_1, \beta_2) \in S_2^{\perp},$$

where $S_2^{\perp}$ is the orthogonal complement of S_2.

If $R(\beta_1, \beta_2)$ is continuous and symmetric positive semidefinite then by Theorem 3.2, R has an eigenfunction $\varphi \in S_2^{\perp}$. But then φ is also an eigenfunction of Γ, and is therefore already contained in S_2. As in the proof of Mercer's theorem in the real-valued case in the addendum to Bierens and Ploberger (1997) in Chapter 5, we then must have that $R(\beta_1, \beta_2) = 0$ on $\mathbf{B} \times \mathbf{B}$.

It is obvious that $R(\beta_1, \beta_2)$ is symmetric. To prove that $R(\beta_1, \beta_2)$ is positive semidefinite, let $f \in L^2_{\mathbb{C}}(\mu)$ be arbitrary. Project f on $S_1 = \text{span}(\{\psi_m\}_{m=1}^{\infty})$, and let $f_1 \in S_1$ be the projection and $f_2 \in S_1^{\perp}$ be the projection residual. Then $\int R(\beta_1, \beta_2)\ f_1(\beta_2)\mathrm{d}\mu(\beta_2) = 0$ and $\int \underline{\Gamma}(\beta_1, \beta_2) f_2(\beta_2)\mathrm{d}\mu(\beta_2) = 0$, hence

$$\begin{aligned}
&\int\int \overline{f(\beta_1)}R(\beta_1, \beta_2)f(\beta_2)\mathrm{d}\mu(\beta_1)\mathrm{d}\mu(\beta_2) \\
&\quad = \int\int \overline{f_2(\beta_1)}R(\beta_1, \beta_2)f_2(\beta_2)\mathrm{d}\mu(\beta_1)\mathrm{d}\mu(\beta_2) \\
&\quad = \int\int \overline{f_2(\beta_1)}\Gamma(\beta_1, \beta_2)f_2(\beta_2)\mathrm{d}\mu(\beta_1)\mathrm{d}\mu(\beta_2) \geq 0. \quad (4.2)
\end{aligned}$$

as is not hard to verify. Thus, $R(\beta_1, \beta_2)$ is positive semidefinite.

To prove that $R(\beta_1,\beta_2)$ is continuous it suffices to prove that $\underline{\Gamma}(\beta_1,\beta_2)$ is continuous, as follows. Note that (4.2) implies that $R(\beta,\beta) \geq 0$ for all $\beta \in \mathbf{B}$, which in its turn implies that for all $\beta \in \mathbf{B}$,

$$\begin{aligned} \sum_{m=1}^{\infty} \lambda_m |\psi_m(\beta)|^2 &= \sum_{m=1}^{\infty} \lambda_m \psi_m(\beta)\overline{\psi_m(\beta)} = \underline{\Gamma}(\beta,\beta) \\ &\leq \Gamma(\beta,\beta) \leq \sup_{\beta\in\mathbf{B}} \Gamma(\beta,\beta) < \infty. \end{aligned} \tag{4.3}$$

Integrating β out yields $\sum_{m=1}^{\infty} \lambda_m < \infty$, which is just part (a) of Theorem 4.1.

To prove part (b) of Theorem 4.1, observe that

$$\begin{aligned} |\psi_m(\beta_1)\overline{\psi_m(\beta_2)}| &= |\psi_m(\beta_1)|.|\psi_m(\beta_2)| \\ &\leq \frac{1}{2}|\psi_m(\beta_1)|^2 + \frac{1}{2}|\psi_m(\beta_2)|^2. \end{aligned} \tag{4.4}$$

It follows from (4.3) and (4.4) that for all $(\beta_1,\beta_2) \in \mathbf{B}\times\mathbf{B}$,

$$\begin{aligned} \sum_{m=1}^{\infty} \lambda_m \left|\psi_m(\beta_1)\overline{\psi_m(\beta_2)}\right| &\leq \frac{1}{2}\sum_{m=1}^{\infty} \lambda_m |\psi_m(\beta_1)|^2 + \frac{1}{2}\sum_{m=1}^{\infty} \lambda_m |\psi_m(\beta_2)|^2 \\ &\leq \sup_{\beta\in\mathbf{B}} \Gamma(\beta,\beta) < \infty. \end{aligned} \tag{4.5}$$

By the same argument as in the proof of Mercer's theorem for real-valued kernels in the addendum to Bierens and Ploberger (1997) in Chapter 5 it follows that (4.5) implies

$$\lim_{n\to\infty} \sup_{(\beta_1,\beta_2)\in\mathbf{B}\times\mathbf{B}} \left| \underline{\Gamma}(\beta_1,\beta_2) - \sum_{m=1}^{n} \lambda_m \psi_m(\beta_1)\overline{\psi_m(\beta_2)} \right| = 0, \tag{4.6}$$

which in its turn implies, by the continuity of $\sum_{m=1}^{n} \lambda_m \psi_m(\beta_1)\overline{\psi_m(\beta_2)}$ for all $n \in \mathbb{N}$, that $\underline{\Gamma}(\beta_1,\beta_2)$ is continuous on $\mathbf{B}\times\mathbf{B}$, and so is $R(\beta_1,\beta_2)$. But then $R(\beta_1,\beta_2) \equiv 0$ on $\mathbf{B}\times\mathbf{B}$, hence

$$\underline{\Gamma}(\beta_1,\beta_2) \equiv \Gamma(\beta_1,\beta_2). \tag{4.7}$$

Part (b) of Theorem 4.1 follows now from (4.6) and (4.7).

As to part (c), suppose that $\{\psi_m\}_{m=1}^{\infty}$ is not complete in $L^2_{\mathbb{C}}(\mu)$. Then the orthogonal complement $S_1^{\perp}$ of $S_1 = \text{span}(\{\psi_m\}_{m=1}^{\infty})$ contains at least one nonzero function φ with unit norm. Since now (4.1) holds exactly on $\mathbf{B} \times \mathbf{B}$ this φ is an eigenvalue of Γ with zero eigenvalue, but then φ is already included in $\{\psi_m\}_{m=1}^{\infty}$. Therefore, $S_1^{\perp} = \{0\}$ and thus $\{\psi_m\}_{m=1}^{\infty}$ is complete in $L_{\mathbb{C}}^2(\mu)$.

This completes the proof of Theorem 4.1. ■

Remark 4.1. Similar to Remark 5.2 in the addendum to Bierens and Ploberger (1997) in Chapter 5, the condition in Assumption 2.1 that $\mathbf{B}$ is compact is only used in the proof of Theorem 3.2 to guarantee

$$\int\int |\Gamma(\beta_1, \beta_2)|^2 \mathrm{d}\mu(\beta_1)\mathrm{d}\mu(\beta_2) < \infty, \tag{4.8}$$

and is only used in the proof of Theorem 4.1 to guarantee that $\sup_{\beta \in \mathbf{B}} \Gamma(\beta, \beta) < \infty$. Therefore, Mercer's theorem carries over to probability measures μ on unbounded domains $\mathbf{B}$ as long as (4.8) holds and $\Gamma(\beta, \beta)$ is uniformly bounded.

5. Lemma 4 Revised

Lemma 4 in BW claims that, with $Z(\beta)$ a complex-valued continuous Gaussian process on a compact subset $\mathbf{B}$ of a Euclidean space and μ a probability measure on $\mathbf{B}$,

$$\int |Z(\beta)|^2 \mathrm{d}\mu(\beta) = \sum_{m=1}^{\infty} \lambda_m e_m' e_m, \tag{5.1}$$

where the λ_m's are the eigenvalues of the covariance kernel

$$\Gamma(\beta_1, \beta_2) = E\left[Z(\beta_1)\overline{Z(\beta_2)}\right]$$

and the e_m's are independently $N_2[0, I_2]$ distributed.

However, it follows from Mercer's theorem that

$$E\left[\int |Z(\beta)|^2 \mathrm{d}\mu(\beta)\right] = \int \Gamma(\beta, \beta)\mathrm{d}\mu(\beta) = \sum_{m=1}^{\infty} \lambda_m,$$

whereas (5.1) implies $E\left[\int |Z(\beta)|^2 \mathrm{d}\mu(\beta)\right] = 2\sum_{m=1}^{\infty} \lambda_m$. Apart from this impossibility result, the main flaw in the original proof of Lemma 4 is due to Equation (A.6) in BW, which reads

$$E[Z_2(\beta_1)Z_2(\beta_2)'] = \sum_{m=1}^{\infty} \lambda_m Q_m(\beta_1)Q_m(\beta_2),$$

where

$$Z_2(\beta) = \begin{pmatrix} \mathrm{Re}[Z(\beta)] \\ \mathrm{Im}[Z(\beta)] \end{pmatrix},$$

instead of the correct expression

$$E[Z_2(\beta_1)Z_2(\beta_2)' + Z_2^*(\beta_1)Z_2^*(\beta_2)'] = \sum_{m=1}^{\infty} \lambda_m Q_m(\beta_1)Q_m(\beta_2),$$

where

$$Z_2^*(\beta) = \begin{pmatrix} \mathrm{Im}[Z(\beta)] \\ -\mathrm{Re}[Z(\beta)] \end{pmatrix}.$$

Actually, the following corrected version of Lemma 4 is related to Bierens and Ploberger (1997, Theorem 3):

Theorem 5.1. (Revised Lemma 4 in BW) *Let $Z(\beta)$ be a complex-valued zero-mean continuous Gaussian process on a compact subs*et $\mathbf{B}$ *of a Euclidean space and let μ be a probability measure on* $\mathbf{B}$. *Then there exists a nonnegative sequence ω_m satisfying $\sum_{m=1}^{\infty} \omega_m < \infty$ such that*

$$\int |Z(\beta)|^2 \mathrm{d}\mu(\beta) = \sum_{m=1}^{\infty} \omega_m \varepsilon_m^2,$$

where the ε_m's are independent standard normally distributed.

Proof. Let $\{\lambda_m\}_{m=1}^{\infty}$ be the sequence of eigenvalues of the covariance kernel

$$\Gamma(\beta_1, \beta_2) = E\left[Z(\beta_1)\overline{Z(\beta_2)}\right]$$

with corresponding sequence $\{\psi_m(\beta)\}_{m=1}^{\infty}$ of orthonormal eigenfunctions (relative to μ). By the completeness of $\{\psi_m(\beta)\}_{m=1}^{\infty}$ we can write $Z(\beta) = \sum_{m=1}^{\infty} g_m \psi_m(\beta)$ a.e. μ,[4] where $g_m = \int Z(\beta)\overline{\psi_m(\beta)}\mathrm{d}\mu(\beta)$. Consequently

$$\int |Z(\beta)|^2 \mathrm{d}\mu(\beta) = \sum_{m=1}^{\infty} |g_m|^2. \tag{5.2}$$

[4] I.e., $\mu\left(\{\beta \in \mathbf{B} : Z(\beta) = \sum_{m=1}^{\infty} g_m \psi_m(\beta)\}\right) = 1.$

Since $Z(\beta)$ is zero-mean Gaussian, the g_m's are jointly zero-mean complex-valued normally distributed. Moreover, by Mercer's theorem we have

$$\begin{aligned}
E\left[\overline{g_k} g_m\right] &= \int\int \psi_k(\beta_2) E\left[\overline{Z(\beta_2)} Z(\beta_1)\right] \overline{\psi_m(\beta_1)} \mathrm{d}\mu(\beta_1)\mathrm{d}\mu(\beta_2) \\
&= \int\int \psi_k(\beta_2)\Gamma(\beta_1,\beta_2)\overline{\psi_m(\beta_1)}\mathrm{d}\mu(\beta_1)\mathrm{d}\mu(\beta_2) \\
&= \sum_{j=1}^{\infty}\lambda_j \int\int \psi_k(\beta_2)\psi_j(\beta_1)\overline{\psi_j(\beta_2)}.\overline{\psi_m(\beta_1)}\mathrm{d}\mu(\beta_1)\mathrm{d}\mu(\beta_2) \\
&= \sum_{j=1}^{\infty}\lambda_j \int \psi_k(\beta_2)\overline{\psi_j(\beta_2)}\mathrm{d}\mu(\beta_2) \int \psi_j(\beta_1)\overline{\psi_m(\beta_1)}\mathrm{d}\mu(\beta_1) \\
&= \sum_{j=1}^{\infty}\lambda_j \mathbf{1}(k=j)\mathbf{1}(m=j) \\
&= \lambda_m \mathbf{1}(k=m). \qquad (5.3)
\end{aligned}$$

By joint normality, (5.3) implies that the sequence $\{g_m\}_{m=1}^{\infty}$ is independent, and so is the sequence $G_m = (\mathrm{Re}[g_m], \mathrm{Im}[g_m])'$. This is a well-known result, but will be proved in Lemma 5.1 below because I could not find a formal proof in the literature.

Each G_m is bivariate zero mean normally distributed, i.e., $G_m \sim N_2[0,\Sigma_m]$. Using the well-known decomposition $\Sigma_m = Q_m\Omega_m Q_m'$, where $\Omega_m = \mathrm{diag}(\omega_{1,m},\omega_{2,m})$ is the diagonal matrix of eigenvalues of Σ_m and Q_m is the orthogonal matrix of the two corresponding eigenvectors, we can write

$$Q_m' G_m = \begin{pmatrix} \sqrt{\omega_{1,m}} e_{1,m} \\ \sqrt{\omega_{2,m}} e_{2,m} \end{pmatrix},$$

were the sequence $(e_{1,m}, e_{2,m})'$ is i.i.d. $N_2[0, I_2]$. Now

$$|g_m|^2 = g_m\overline{g_m} = G_m' G_m = G_m' Q_m Q_m' G_m = \omega_{1,m} e_{1,m}^2 + \omega_{2,m} e_{2,m}^2$$

where $\omega_{1,m} + \omega_{2,m} = \lambda_m$, and by Mercer's theorem,

$$\sum_{m=1}^{\infty}\omega_{1,m} + \sum_{m=1}^{\infty}\omega_{2,m} = \sum_{m=1}^{\infty} E[g_m\overline{g_m}] = \sum_{m=1}^{\infty}\lambda_m < \infty.$$

Thus (5.2) now reads

$$\int |Z(\beta)|^2 \mathrm{d}\mu(\beta) = \sum_{m=1}^{\infty} \omega_{1,m} e_{1,m}^2 + \sum_{m=1}^{\infty} \omega_{2,m} e_{2,m}^2.$$

Finally, denoting for $m \in \mathbb{N}$, $\omega_{2m-1} = \omega_{1,m}$, $\omega_{2m} = \omega_{2,m}$, $\varepsilon_{2m-1} = e_{1,m}$, $\varepsilon_{2m} = e_{2,m}$, for example, the result of Theorem 5.1 follows. ■

As said before, the claim that (5.3) implies that the sequence $\{g_m\}_{m=1}^{\infty}$ is independent, and so is the sequence $G_m = (\mathrm{Re}[g_m], \mathrm{Im}[g_m])'$, is a well-known result. However, as appears from the proof of the following lemma, this result is far from obvious.

Lemma 5.1. *Let* $\{g_m\}_{m=1}^{\infty}$ *be a sequence of zero-mean complex-valued jointly Gaussian random variables satisfying*

$$E\left[\overline{g_m} g_k\right] = 0 \text{ for } k \neq m. \tag{5.4}$$

Then the sequence $G_m = (\mathrm{Re}[g_m], \mathrm{Im}[g_m])'$, $m \in \mathbb{N}$, *is independent.*

Proof. By the joint normality of the sequence $\{G_m\}_{m=1}^{\infty}$ it suffices to verify that for $m \neq k$, $E\left[G_k G_m'\right] = O$, as follows. First, note that

$$\begin{aligned} \overline{g_m} g_k &= (\mathrm{Re}[g_m]\,\mathrm{Re}[g_k] + \mathrm{Im}[g_m]\,\mathrm{Im}[g_k]) \\ &\quad + \mathbf{i}.\,(\mathrm{Re}[g_m]\,\mathrm{Im}[g_k] - \mathrm{Im}[g_m]\,\mathrm{Re}[g_k]), \end{aligned}$$

hence (5.4) implies

$$\begin{array}{l} E\,(\mathrm{Re}[g_m]\,\mathrm{Re}[g_k]) + E\,(\mathrm{Im}[g_m]\,\mathrm{Im}[g_k]) = 0, \\ E\,(\mathrm{Re}[g_m]\,\mathrm{Im}[g_k]) - E\,(\mathrm{Im}[g_m]\,\mathrm{Re}[g_k]) = 0. \end{array} \tag{5.5}$$

It follows straightforwardly from (5.5) that

$$E\left[\begin{pmatrix} \mathrm{Re}[g_k] & -\mathrm{Im}[g_k] \\ \mathrm{Im}[g_k] & \mathrm{Re}[g_k] \end{pmatrix} \begin{pmatrix} \mathrm{Re}[g_m] & \mathrm{Im}[g_m] \\ -\mathrm{Im}[g_m] & \mathrm{Re}[g_m] \end{pmatrix}\right] = O,$$

which can be written as

$$E\left[G_k G_m'\right] + P_2 E\left[G_k G_m'\right] P_2' = E\left[(G_k, P_2 G_k) \begin{pmatrix} G_m' \\ G_m' P_2' \end{pmatrix}\right] = O, \tag{5.6}$$

where

$$P_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Next, observe from (5.5) that

$$\begin{aligned} E[G_k G'_m] &= \begin{pmatrix} E(\mathrm{Re}[g_k]\,\mathrm{Re}[g_m]) & E(\mathrm{Re}[g_k]\,\mathrm{Im}[g_m]) \\ E(\mathrm{Im}[g_k]\,\mathrm{Re}[g_m]) & E(\mathrm{Im}[g_k]\,\mathrm{Im}[g_m]) \end{pmatrix} \\ &= \begin{pmatrix} E(\mathrm{Re}[g_k]\,\mathrm{Re}[g_m]) & E(\mathrm{Re}[g_k]\,\mathrm{Im}[g_m]) \\ E(\mathrm{Re}[g_k]\,\mathrm{Im}[g_m]) & -E(\mathrm{Re}[g_k]\,\mathrm{Re}[g_m]) \end{pmatrix}, \end{aligned}$$

hence $E[G_k G'_m]$ is symmetric, with eigenvalues

$$\begin{aligned} \lambda_1 &= \sqrt{(E(\mathrm{Re}[g_m]\,\mathrm{Re}[g_k]))^2 + (E(\mathrm{Re}[g_m]\,\mathrm{Im}[g_k]))^2}, \\ \lambda_2 &= -\lambda_1, \end{aligned}$$

Therefore, $E[G_m G'_k]$ can be written as

$$E[G_m G'_k] = \lambda_1 Q_{k,m} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} Q'_{k,m}, \tag{5.7}$$

where $Q_{k,m}$ is an orthogonal 2×2 matrix. Substituting the expression (5.7) in (5.6) yields

$$\lambda_1 \left(\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + Q'_{k,m} P_2 Q_{k,m} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} Q'_{k,m} P'_2 Q_{k,m} \right) = O. \tag{5.8}$$

Since P_2 and $Q_{k,m}$ are orthogonal, the matrix $Q'_{k,m} P_2 Q_{k,m}$ is orthogonal, hence it follows from (5.8) that

$$\lambda_1 \left(Q'_{k,m} P'_2 Q_{k,m} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} Q'_{k,m} P'_2 Q_{k,m} \right) = O. \tag{5.9}$$

As is well-known, the 2×2 orthogonal matrix $Q'_{k,m} P_2 Q_{k,m}$ can be written as

$$Q'_{k,m} P_2 Q_{k,m} = \begin{pmatrix} \cos(\phi) & \sin(\phi) \\ -\sin(\phi) & \cos(\phi) \end{pmatrix} \tag{5.10}$$

for some $\phi \in [0, 2\pi]$. Since

$$\begin{aligned} \begin{pmatrix} \cos(\phi) & \sin(\phi) \\ -\sin(\phi) & \cos(\phi) \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} &= \begin{pmatrix} \cos(\phi) & -\sin(\phi) \\ -\sin(\phi) & -\cos(\phi) \end{pmatrix}, \\ \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \cos(\phi) & \sin(\phi) \\ -\sin(\phi) & \cos(\phi) \end{pmatrix} &= \begin{pmatrix} \cos(\phi) & \sin(\phi) \\ \sin(\phi) & -\cos(\phi) \end{pmatrix}, \end{aligned}$$

it follows that Equation (5.9) reads

$$2\lambda_1 \cos(\phi) \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = O,$$

hence $\lambda_1 \cos(\phi) = 0$, so that either

$$\lambda_1 = 0 \text{ or } \phi \in \{\pi/2, 3\pi/4\} \text{ or both.} \tag{5.11}$$

Suppose that $\phi = \pi/2$. Then by (5.10),

$$Q'_{k,m} P_2 Q_{k,m} = P_2. \tag{5.12}$$

Again, without loss of generality we may assume that for some $\theta \in [0, 2\pi]$,

$$Q_{k,m} = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}.$$

Then it is easy to verify that

$$Q'_{k,m} P_2 Q_{k,m} = \begin{pmatrix} 1 - 2\sin^2(\theta) & -2\cos(\theta)\sin(\theta) \\ -2\cos(\theta)\sin(\theta) & 1 - 2\cos^2(\theta) \end{pmatrix},$$

which is obviously unequal to P_2 for all $\theta \in [0, 2\pi]$. Thus, the equality (5.12) is not possible, hence $\phi \neq \pi/2$. Similarly, $\phi = 3\pi/4$ implies $Q'_{k,m} P_2 Q_{k,m} = P'_2$, which is also not possible, so that $\phi \neq 3\pi/4$ as well. Consequently, it follows from (5.11) that $\lambda_1 = 0$, which by (5.7) implies that $E\left[G_m G'_k\right] = O$. ■

6. Upper Bounds of the Critical Values of the SICM Test

Similar to Bierens and Ploberger (1997, Theorem 7) it follows from Theorem 5.1 that the following result holds.

Theorem 6.1. *Let the conditions of Theorem 5.1 hold, and let*

$$\overline{\chi}_1^2 = \sup_{n \geq 1} \frac{1}{n} \sum_{m=1}^{n} \varepsilon_m^2.$$

Then

$$\begin{aligned} \Pr\left[\left(\int \Gamma(\beta, \beta) \mathrm{d}\mu(\beta)\right)^{-1} \int |Z(\beta)|^2 \mathrm{d}\mu(\beta) > t\right] &= \Pr\left[\frac{\sum_{m=1}^{\infty} \omega_m \varepsilon_m^2}{\sum_{m=1}^{\infty} \omega_m} > t\right] \\ &\leq \Pr[\overline{\chi}_1^2 > t] \end{aligned}$$

for all $t > 0$. Therefore, for $\alpha \in (0,1)$ and $t(\alpha)$ such that $\Pr[\overline{\chi}_1^2 > t(\alpha)] = \alpha$,

$$\Pr\left[\int |Z(\beta)|^2 \mathrm{d}\mu(\beta) > t(\alpha). \int \Gamma(\beta,\beta)\mathrm{d}\mu(\beta)\right] \leq \alpha.$$

The values of $t(\alpha)$ for $\alpha = 0.01$, $\alpha = 0.05$ and $\alpha = 0.10$ have been calculated in Bierens and Ploberger (1997), i.e.,

$$t(0.01) = 6.81,\ t(0.05) = 4.26,\ t(0.10) = 3.23 \tag{6.1}$$

To apply these upper bounds of the asymptotic critical values to the SICM test we need a consistent estimate of $\int \Gamma(\beta,\beta)\mathrm{d}\mu(\beta)$. Recall from Section 3 in BW that the empirical process on which the SICM test is based takes the form

$$\widehat{Z}_n^{(s)}(\tau,\xi) = \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\left(\exp(\mathbf{i}.\tau'Y_j) - \exp(\mathbf{i}.\tau'\tilde{Y}_j)\right)\exp(\mathbf{i}.\xi'X_j)$$

where Y_j and X_j are (made) bounded random vectors, $\tilde{Y}_j$ is a random drawing from the estimated conditional distribution $F(y|X_j;\hat{\theta})$ of Y_j, and $(\tau,\xi) \in \Upsilon \times \Xi$. The corresponding estimated covariance function takes the form

$$\begin{aligned}
&\widehat{\Gamma}_n^{(s)}\left((\tau_1,\xi_1),(\tau_2,\xi_2)\right) \\
&= \frac{1}{n}\sum_{j=1}^{n}\left(\exp(\mathbf{i}.\tau_1'Y_j) - \exp(\mathbf{i}.\tau_1'\tilde{Y}_j)\right)\exp(\mathbf{i}.\xi_1'X_j) \\
&\qquad \times \left(\exp(-\mathbf{i}.\tau_2'Y_j) - \exp(-\mathbf{i}.\tau_2'\tilde{Y}_j)\right)\exp(-\mathbf{i}.\xi_2'X_j) \\
&= \frac{1}{n}\sum_{j=1}^{n}\Big(\exp(\mathbf{i}.(\tau_1-\tau_2)'Y_j) + \exp(\mathbf{i}.(\tau_1-\tau_2)'\tilde{Y}_j) \\
&\qquad - \exp(\mathbf{i}.\tau_1'Y_j)\exp(-\mathbf{i}.\tau_2'\tilde{Y}_j) - \exp(\mathbf{i}.\tau_1'\tilde{Y}_j)\exp(-\mathbf{i}.\tau_2'Y_j)\Big) \\
&\qquad\qquad \times \exp(\mathbf{i}.(\xi_1-\xi_2)'X_j),
\end{aligned}$$

hence

$$\widehat{\Gamma}_n^{(s)}\left((\tau,\xi),(\tau,\xi)\right)$$

$$= \frac{1}{n}\sum_{j=1}^{n}\Big(2 - \exp(\mathbf{i}.\tau'(Y_j - \tilde{Y}_j)) - \exp(-\mathbf{i}.\tau'(Y_j - \tilde{Y}_j))\Big)$$
$$= 2 - 2\frac{1}{n}\sum_{j=1}^{n}\cos(\tau'(Y_j - \tilde{Y}_j)).$$

Now let $\Upsilon = [-c, c]^m$ and $\Xi = [-c, c]^k$ and let μ be the uniform probability measure on $\Upsilon\times\Xi$. Then the expression for $\widehat{T}_n^{(s)}(c) = \int\left|\widehat{Z}_n^{(s)}(\tau,\xi)\right|^2 \mathrm{d}\mu(\tau,\xi)$ is given in Equation (31) in BW, whereas

$$\widehat{R}_n^{(s)}(c) = \int \widehat{\Gamma}_n^{(s)}\left((\tau,\xi),(\tau,\xi)\right)\mathrm{d}\mu(\tau,\xi) = 2 - 2\frac{1}{n}\sum_{j=1}^{n}\prod_{i=1}^{m}\frac{\sin\Big(c(Y_{i,j} - \tilde{Y_{i,j}})\Big)}{c(Y_{i,j} - \tilde{Y_{i,j}})},$$

where $Y_{i,j}$ and $\tilde{Y_{i,j}}$ are component i of Y_j and $\tilde{Y}_j$, respectively. The upper bounds (6.1) are now applicable to the standardized SICM test statistic $\widehat{T}_n^{(s)}(c)/\widehat{R}_n^{(s)}(c)$, i.e., under the null hypothesis,

$$\begin{aligned}
\limsup_{n\to\infty} \Pr\left[\widehat{T}_n^{(s)}(c)/\widehat{R}_n^{(s)}(c) > 6.81\right] &\leq 0.01\\
\limsup_{n\to\infty} \Pr\left[\widehat{T}_n^{(s)}(c)/\widehat{R}_n^{(s)}(c) > 4.36\right] &\leq 0.05\\
\limsup_{n\to\infty} \Pr\left[\widehat{T}_n^{(s)}(c)/\widehat{R}_n^{(s)}(c) > 3.23\right] &\leq 0.10
\end{aligned}$$

7. Standardization

In Assumption 3 we required that for each component V_i of $V = (Y', X')'$, $\sqrt{n}(\mu_{i,n} - \mu_i) = O_p(1)$ and $\sqrt{n}\,(\sigma_{i,n} - \sigma_i) = O_p(1)$, where $\mu_{i,n}$ is the sample mean of the V_i's, $\mu_i = E[V_i]$, $\sigma_{i,n}$ is the sample standard error of the V_i's, and $\sigma_i = \sqrt{E[(V_i - \mu_i)^2]}$. These conditions hold if $E[V_i^4] < \infty$. However, using the approach in the addendum to Bierens (1982) in Chapter 2 it can be shown that the conditions $\mu_{i,n} - \mu_i = o_p(1)$, $\sigma_{i,n} - \sigma_i = o_p(1)$ suffice, which only require that $E[V_i^2] < \infty$.

8. Concluding Remarks

In the mathematical literature the Hilbert-Schmidt and Mercer theorems are nowadays usually derived as by-products of linear operator theory,[5] which however is way over my head. Therefore, in this addendum to BW I have presented alternative proofs which only require elementary Hilbert space theory and basic knowledge of linear algebra and complex calculus. I do not claim any originality. It seems likely that these proofs are already done in this way somewhere, but if so I am not aware of any references. The only originality claim I can make is that I figured out these proofs all by myself.

Note that we could have used moment generating functions rather than characteristic functions because all the variables involved are bounded or made bounded. Then we could have used the Hilbert-Schmidt and Mercer theorems in Bierens and Ploberger (1997) and its addendum in Chapter 5.

Finally, note that recently, in Bierens and Wang (2016), the SICM test has been generalized to parametric conditional distributions of stationary time series models, by combining the weighted ICM testing idea in Bierens (1984) with the approach in BW.

References

Bernkopf, M. (1966): "The Development of Function Spaces with Particular Reference to their Origins in Integral Equation Theory", *Archive for History of Exact Sciences* 3, 1-96.

Bierens, H. J. (1982): "Consistent Model Specification Tests", *Journal of Econometrics* 20, 105-134, reprinted in Chapter 2.

Bierens, H. J. (1984): "Model Specification Testing of Time Series Regressions", *Journal of Econometrics* 26, 323-353, reprinted in Chapter 3.

Bierens, H. J. (1990): "A Consistent Conditional Moment Test of Functional Form", *Econometrica* 58, 1443-1458, reprinted in Chapter 4.

Bierens, H. J. (2004): *Introduction to the Mathematical and Statistical Foundations of Econometrics*, Cambridge University Press.

Bierens, H. J., and W. Ploberger (1997): "Asymptotic Theory of Integrated Conditional Moment Tests", *Econometrica* 65, 1129-1151, reprinted in Chapter 5.

[5] See for example Krein (1998), among others.

Bierens, H. J., and L. Wang (2012): "Integrated Conditional Moment Tests for Parametric Conditional Distributions", *Econometric Theory* 28, 328-362, reprinted in Chapter 6.

Bierens, H. J., and L. Wang (2016): "Weighted Simulated Integrated Conditional Moment Tests for Parametric Conditional Distributions of Stationary Time Series Processes", forthcoming in *Econometric Reviews.*

Hadinejad-Mahram, H., D. Dahlhaus and D. Blomker (2002): "Karhunen-Loeve Expansion of Vector Random Processes", Technical Report No. IKT-NT 1019, Communications Technological Laboratory, Swiss Federal Institute of Technology, Zurich.

Krein, M. G. (1998): "Compact Linear Operators on Functional Spaces with Two Norms", *Integral Equations and Operator Theory* 30, 140-162.

Mercer, J. (1909): "Functions of Positive and Negative Type and their Connection with the Theory of Integral Equations", *Philosophical Transactions of the Royal Society A* 209, 415-446.

Siegmund-Schultze, R. (1986): "Der Beweis des Hilbert-Schmidt-Theorems", *Archive for History of Exact Sciences* 36, 251-270.