CHAPTER 2
More Splines

2.1 Splines on the circle.

Splines on the circle can be obtained by imposing periodic boundary conditions on splines in W_m, but it is more instructive to describe splines on the circle from the beginning since the eigenfunctions and eigenvalues of the associated reproducing kernel have a particularly simple form.

Let W_m^0 (per) be the collection of all functions on $[0, 1]$ of the form

$$f(t) \sim \sqrt{2}\sum_{\nu=1}^{\infty} a_\nu \cos 2\pi\nu t + \sqrt{2}\sum_{\nu=1}^{\infty} b_\nu \sin 2\pi\nu t$$

with

$$\sum_{\nu=1}^{\infty}(a_\nu^2 + b_\nu^2)(2\pi\nu)^{2m} < \infty. \tag{2.1.1}$$

Since

$$\frac{d^m}{dt^m}\left\{ \begin{array}{c} \cos 2\pi\nu t, \\ \sin 2\pi\nu t \end{array} \right\} = (2\pi\nu)^m \times \left\{ \begin{array}{c} \pm\sin 2\pi\nu t \\ \pm\cos 2\pi\nu t \end{array} \right\}, \tag{2.1.2}$$

then if (2.1.1) holds, we have

$$\sum_{v=1}^{\infty}(a_\nu^2 + b_\nu^2)(2\pi\nu)^{2m} = \int_0^1 (f^{(m)}(u))^2\,du. \tag{2.1.3}$$

It is easy to see that the r.k. $R^1(s,t)$ for W_m^0 (per) is

$$\begin{aligned} R^1(s,t) &= \sum_{\nu=1}^{\infty} \frac{2}{(2\pi\nu)^{2m}} [\cos 2\pi\nu s \cos 2\pi\nu t \\ &\qquad + \sin 2\pi\nu s \sin 2\pi\nu t] \qquad (2.1.4) \\ &= \sum_{\nu=1}^{\infty} \frac{2}{(2\pi\nu)^{2m}} \cos 2\pi\nu(s-t). \qquad (2.1.5) \end{aligned}$$

The eigenvalues of the reproducing kernel are all of multiplicity 2 and are $\lambda_\nu = (2\pi\nu)^{-2m}$, and the eigenfunctions are $\sqrt{2}\ \sin 2\pi\nu t$ and $\sqrt{2}\cos 2\pi\nu t$.

Elements in W_m^0 (per) satisfy the boundary conditions

$$\begin{aligned} \int_0^1 f(u)\,du &= 0, \\ \int_0^1 f^{(k)}(u)\,du &= f^{(k-1)}(1) - f^{(k-1)}(0) = 0, \\ &\quad k = 1, \dots, m. \end{aligned} \tag{2.1.6}$$

To remove the condition $\int_0^1 f(u)\,du = 0$, we may adjoin the one-dimensional subspace $\mathcal{H}_0$ spanned by $\{1\}$, and let

$$W_m(\text{per}) = \{1\} \oplus W_m^0(\text{per}).$$

W_m (per), endowed with the norm

$$\|f\|^2 = \left[\int_0^1 f(u)\,du\right]^2 + \int_0^1 (f^{(m)}(u))^2\,du,$$

has the r.k.

$$R(s,t) = 1 + R^1(s,t) = 1 + 2\sum_{\nu=1}^{\infty} \frac{1}{(2\pi\nu)^{2m}} \cos 2\pi\nu(s-t). \tag{2.1.7}$$

W_m (per) is the subspace of W_m satisfying the periodic boundary conditions $f^{(\nu)}(1) = f^{(\nu)}(0)$, $\nu = 0, 1, 2, \dots, m-1$.

A closed form expression for $R^1(s,t)$ of (2.1.5) using Bernoulli polynomials was given by Craven and Wahba (1979). Recall that the Bernoulli polynomials $B_r(t)$, $r = 0, 1, \dots$, $t \in [0,1]$ satisfy the recursion relations

$$B_0(t) = 1$$

$$\frac{1}{r}\frac{d}{dt}B_r(t) = B_{r-1}(t), \quad \int_0^1 B_r(u)\,du = 0, \quad r = 1, 2, \dots .$$

Abramowitz and Stegun (1965) give the formula

$$B_{2m}(x) = (-1)^{m-1} 2(2m)! \sum_{\nu=1}^{\infty} \frac{\cos 2\pi\nu x}{(2\pi\nu)^{2m}}, \qquad x \in [0,1]$$

so that R^1 of (2.1.5) is given by

$$R^1(s,t) = \frac{(-1)^{m-1}}{(2m)!} B_{2m}([s-t])$$

where $[s-t]$ is the fractional part of $s - t$. $R^1(s,t)$ is a stationary covariance on the circle, whose associated stochastic process $X(t)$, $t \in [0,1]$ possess exactly $m-1$ quadratic-mean derivatives and satisfies $D^m X = dW$, and the periodic boundary conditions $X^{(\nu)}(0) = X^{(\nu)}(1)$, $\nu = 0, 1, 2, \dots, m-1$.

It is instructive to look at the "frequency response" of the smoothing spline in this case. Let n be even and consider

$$y_i = f\left(\frac{i}{n}\right) + \epsilon_i, \quad i = 1, 2, \ldots, n$$

with $f \in W_m$ (per) and ϵ as before. To simplify the argument, we will look at an approximation to the original minimization problem, namely, find f_λ of the form

$$f_\lambda(t) = a_0 + \sum_{\nu=1}^{n/2-1} a_\nu \sqrt{2}\cos 2\pi\nu t + \sum_{\nu=1}^{n/2-1} b_\nu \sqrt{2}\sin 2\pi\nu t + a_{n/2}\cos \pi n t \quad (2.1.8)$$

to minimize

$$\frac{1}{n}\sum_{i=1}^{n}\left(y_i - f\left(\frac{i}{n}\right)\right)^2 + \lambda \int_0^1 (f^{(m)}(u))^2\, du. \quad (2.1.9)$$

Using the orthogonality relations

$$\begin{aligned}
\frac{2}{n}\sum_{i=1}^{n} \cos 2\pi\nu\frac{i}{n}\cos 2\pi\mu\frac{i}{n} &= 1 \qquad \mu = \nu = 1, \ldots, n/2-1,\\
&= 0 \qquad \mu \neq \nu, \mu, \nu = 0, 1, \ldots, n/2,\\
\frac{2}{n}\sum_{i=1}^{n} \sin 2\pi\nu\frac{i}{n}\sin 2\pi\mu\frac{i}{n} &= 1 \qquad \mu = \nu = 1, \ldots, n/2-1,\\
&= 0 \qquad \mu \neq \nu, \mu, \nu = 1, \ldots, n/2-1,\\
\frac{1}{n}\sum_{i=1}^{n}\left(\cos 2\pi\nu\frac{i}{n}\right)^2 &= 1, \qquad \nu = 0, n/2,\\
\frac{1}{n}\sum_{i=1}^{n} \cos 2\pi\nu\frac{i}{n}\sin 2\pi\mu\frac{i}{n} &= 0
\end{aligned}$$

we have

$$\begin{aligned}
a_\nu &= \frac{\sqrt{2}}{n}\sum_{i=1}^{n}\cos 2\pi\nu\frac{i}{n} f\left(\frac{i}{n}\right), \qquad \nu = 1, 2, \ldots, n/2-1,\\
b_\nu &= \frac{\sqrt{2}}{n}\sum_{i=1}^{n}\sin 2\pi\nu\frac{i}{n} f\left(\frac{i}{n}\right), \qquad \nu = 1, 2, \ldots, n/2-1,\\
a_0 &= \frac{1}{n}\sum_{i=1}^{n} f\left(\frac{i}{n}\right),\\
a_{n/2} &= \frac{1}{n}\sum_{i=1}^{n}\cos \pi i f\left(\frac{i}{n}\right),
\end{aligned}$$

and letting

$$\hat{a}_\nu = \frac{\sqrt{2}}{n}\sum_{i=1}^{n} y_i \cos 2\pi\nu\frac{i}{n},$$

$$\hat{b}_\nu = \frac{\sqrt{2}}{n}\sum_{i=1}^{n} y_i \sin 2\pi\nu\frac{i}{n},$$

$$\hat{a}_0 = \frac{1}{n}\sum_{i=1}^{n} y_i,$$

$$\hat{a}_{n/2} = \frac{1}{n}\sum_{i=1}^{n} y_i \cos \pi i,$$

(2.1.9) becomes

$$\sum_{\nu=0}^{n/2}(a_\nu - \hat{a}_\nu)^2 + \sum_{\nu=1}^{n/2-1}(b_\nu - \hat{b}_\nu)^2 + \lambda\left[\sum_{\nu=1}^{n/2-1}(a_\nu^2 + b_\nu^2)(2\pi\nu)^{2m} + \frac{1}{2}a_{n/2}^2(\pi n)^{2m}\right].$$

The minimizing values are

$$\begin{aligned} a_\nu &= \hat{a}_\nu/(1+\lambda(2\pi\nu)^{2m}), \\ & \qquad\qquad\qquad \nu = 1, 2, \ldots, n/2-1, \\ b_\nu &= \hat{b}_\nu/(1+\lambda(2\pi\nu)^{2m}), \\ a_0 &= \hat{a}_0, \\ a_{n/2} &= \hat{a}_{n/2}/(1+\tfrac{1}{2}\lambda(2\pi\nu)^{2m}), \end{aligned}$$

and

$$\begin{aligned} f_\lambda(t) = \hat{a}_0 &+ \sum_{\nu=1}^{n/2-1}\frac{\hat{a}_\nu}{(1+\lambda(2\pi\nu)^{2m})}\cos 2\pi\nu t \\ &+ \sum_{\nu=1}^{n/2-1}\frac{\hat{b}_\nu}{(1+\lambda(2\pi\nu)^{2m})}\sin 2\pi\nu t \\ &+ \frac{\hat{a}_{n/2}}{(1+\frac{1}{2}\lambda(\pi n)^{2m})}\cos \pi n t. \end{aligned}$$

Thus, the smoothing spline obtained with the penalty functional $\int_0^1 (f^{(m)}(u))^2\,du$ may be viewed as a generalization of the so-called Butterworth filter, which smooths the data by downweighting the component at frequency ν by the weight $\omega(\nu) = (1+\lambda(2\pi\nu)^{2m})^{-1}$.

2.2 Splines on the sphere, the role of the iterated Laplacian.

We will see that the iterated Laplacian plays a role in splines on the circle, the sphere, the line, and the plane and other index sets on which the Laplacian operator commutes with the group operation. In d dimensions the Laplacian is

$$\Delta f = \left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \ldots + \frac{\partial^2}{\partial x_d^2}\right) f \tag{2.2.1}$$

and the (surface) Laplacian on the (unit) sphere is

$$\Delta f = \frac{1}{\cos^2 \phi} f_{\theta\theta} + \frac{1}{\cos \phi}(\cos \phi f_\phi)_\phi \tag{2.2.2}$$

where θ is the longitude ($0 \le \theta \le 2\pi$) and ϕ is the latitude ($-\pi/2 \le \phi \le \pi/2$). Here we use subscripts θ and ϕ to indicate derivatives with respect to θ and ϕ, not to be confused with a subscript λ that indicates dependence on the smoothing parameter λ. On the circle we have

$$\Delta f = \frac{\partial^2}{\partial x^2} f$$

and if $f \in W_m$ (per) then we can integrate by parts to obtain

$$\begin{aligned} \int_0^1 (f^{(m)}(u))^2 \, du &= (-1)^m \int_0^1 f(u) f^{(2m)}(u) \, du \\ &= (-1)^m \int_0^1 f(u) \Delta^m f(u) \, du. \end{aligned} \tag{2.2.3}$$

The eigenfunctions $\{\sqrt{2}\cos 2\pi\nu t,\ \sqrt{2}\sin 2\pi\nu t\}$ of the r.k. R^1 of W_m (per) are the eigenfunctions of the mth iterated Laplacian Δ^m on the circle, while the eigenvalues $\{\lambda_\nu = (2\pi\nu)^{-2m}\}$ are the inverses of the eigenvalues of Δ^m:

$$\Delta^m \Phi_\nu = (-1)^m (2\pi\nu)^{2m} \Phi_\nu,$$

that is,

$$\begin{aligned} D^{2m} \cos 2\pi\nu t &= (-1)^m (2\pi\nu)^{2m} \cos 2\pi\nu t, \\ D^{2m} \sin 2\pi\nu t &= (-1)^m (2\pi\nu)^{2m} \sin 2\pi\nu t. \end{aligned}$$

The generalization to the sphere is fairly immediate. The eigenfunctions of the (surface) Laplacian on the sphere are the spherical harmonics $Y_{\ell s}$, $s = -\ell, \ldots, \ell$, $\ell = 0, 1, \ldots$, where

$$\begin{aligned} Y_{\ell s}(\theta, \phi) &= \theta_{\ell s} \cos s\theta P_\ell^s(\sin \phi), \quad 0 < s \le \ell, \quad \ell = 0, 1, \ldots \\ &= \theta_{\ell s} \sin s\theta P_\ell^{|s|}(\sin \phi), \quad -\ell \le s < 0 \\ &= \theta_{\ell 0} P_\ell(\sin \phi), \quad s = 0 \end{aligned}$$

where

$$\begin{aligned} \theta_{\ell s} &= \sqrt{2} \sqrt{\frac{2\ell + 1}{4\pi} \frac{(\ell - |s|)!}{(\ell + |s|)!}}, \quad s \ne 0 \\ &= \sqrt{\frac{2\ell + 1}{4\pi}}, \quad s = 0. \end{aligned}$$

$P_\ell, \ell = 0, 1, \ldots$ are the Legendre polynomials and P_ℓ^s are the Legendre functions,

$$P_\ell^s(z) = (1 - z^2)^{s/2} \left(\frac{\partial^s}{\partial z^s} \right) P_\ell(z).$$

The spherical harmonics are the eigenfunctions of the mth iterated Laplacian,

$$\Delta^m Y_{\ell s} = (-1)^m [\ell(\ell+1)]^m Y_{\ell s}, \quad s = -\ell, \ldots, \ell, \ \ell = 0, 1, \ldots \tag{2.2.4}$$

and provide a complete orthonormal sequence for $\mathcal{L}_2(\mathcal{S})$, where $\mathcal{S}$ is the unit sphere (see Sansone (1959)).

Let $P = (\theta, \phi)$ and let

$$\begin{aligned} R(P, P') &= 1 + \sum_{\ell=1}^{\infty} \sum_{s=-\ell}^{\ell} \left[\frac{1}{\ell(\ell+1)}\right]^m Y_{\ell s}(P) Y_{\ell s}(P') \\ &= 1 + R^1(P, P'), \text{ say.} \end{aligned} \tag{2.2.5}$$

Letting

$$f_{\ell s} = \int_S f(P) Y_{\ell s}(P)\, dP,$$

the Hilbert space $\mathcal{H}$ with r.k. R of (2.2.5) is the collection of all functions on the sphere with $f_{00} < \infty$ and

$$\begin{aligned} \|P_1 f\|^2 &= \sum_{\ell=1}^{\infty} \sum_{s=-\ell}^{\ell} f_{\ell s}^2 [\ell(\ell+1)]^m \\ &\equiv \int_S (\Delta^{m/2} f)^2\, dP < \infty. \end{aligned} \tag{2.2.6}$$

Splines and generalized splines on the sphere have been studied by Wahba (1981d, 1982a), Freeden (1981), and Shure, Parker, and Backus (1982).

There is an addition formula for spherical harmonics analogous to the addition formula for sines and cosines

$$\cos 2\pi\nu s \cos 2\pi\nu t + \sin 2\pi\nu s \sin 2\pi\nu t = \cos 2\pi\nu(t - s),$$

it is

$$\sum_{s=-\ell}^{\ell} Y_{\ell s}(P) Y_{\ell s}(P') = \frac{2\ell+1}{4\pi} P_\ell(\cos \gamma(P, P')) \tag{2.2.7}$$

where γ is the angle between P and P' (see Sansone (1959)). Thus $R^1(P, P')$ collapses to

$$R^1(P, P') = \sum_{\ell=1}^{\infty} \frac{2\ell+1}{4\pi} \frac{1}{[\ell(\ell+1)]^m} P_\ell(\cos \gamma(P, P')), \tag{2.2.8}$$

with the stationarity (dependence only on $\gamma(P, P')$) being evident. Closed form expressions for R^1 for $m = 2$ and (in terms of the dilogarithms) for $m = 3$ were given by Wendelberger (1982), but it appears that closed form expressions for

$$\sum_{\ell=1}^{\infty} \frac{2\ell+1}{[\ell(\ell+1)]^m} P_\ell(z) \tag{2.2.9}$$

are not available for larger m. Reproducing kernels Q^1 that approximate R^1 for $m = 2, 3, \ldots$, and for which closed form expressions are available have been found in Wahba (1981d, 1982a). The eigenvalues $\lambda_{\ell s} = (\ell(\ell+1))^{-m}$ that appear in (2.2.5) and (2.2.8) are replaced by $\xi_{\ell s} = [(\ell + \frac{1}{2})(\ell+1)(\ell+2)\ldots(\ell+2m-1)]^{-1}$, to get

$$\begin{aligned} Q^1(P,P') &= \sum_{\ell=1}^{\infty} \sum_{s=-\ell}^{\ell} \xi_{\ell s} Y_{\ell s}(P) Y_{\ell s}(P') \\ &= \frac{1}{2\pi} \sum_{\ell=1}^{\infty} \frac{1}{(\ell+1)(\ell+2)\ldots(\ell+2m-1)} P_\ell(\cos\gamma(P,P')). \end{aligned}$$

Since

$$\begin{aligned} \xi_{\ell s} &\le \lambda_{\ell s} \le m^{2m} \xi_{\ell s}, \\ Q^1 &\preceq R^1 \preceq m^{2m} Q^1 \end{aligned}$$

where $A \preceq B$ means $B - A$ is nonnegative definite, and $\mathcal{H}_{Q^1}$ and $\mathcal{H}_{R^1}$ are topologically equivalent. Closed form expressions for Q^1 for $m = 2, 5/2, 3, \ldots$, were obtained via the symbol manipulation program MACSYMA.

Another way of computing an approximate spline on the sphere, given noisy data from the model

$$Y_i = f(P_i) + \epsilon_i,$$

is to let f be of the form

$$f = f_{00} + \sum_{\ell=1}^{N} \sum_{s=-\ell}^{\ell} f_{\ell s} Y_{\ell s}, \tag{2.2.10}$$

and choose the $f_{\ell s}$ to minimize

$$\frac{1}{n} \sum_{i=1}^{n} \left(y_i - \sum_{\ell=0}^{N} \sum_{s=-\ell}^{\ell} f_{\ell s} Y_{\ell s}(P_i) \right)^2 + \lambda \sum_{\ell=1}^{N} \sum_{s=-\ell}^{\ell} [(\ell)(\ell+1)]^m f_{\ell s}^2. \tag{2.2.11}$$

Arranging the index set $\{(\ell, s)\}$ in a convenient order, and letting f be the vector of $f_{\ell s}$ and X be the matrix with $i, \ell s$th entry $Y_{\ell s}(P_i)$, we have that (2.2.11) becomes

$$\frac{1}{n} \|y - Xf\|^2 + \lambda f' D^{-1} f \tag{2.2.12}$$

where D is the diagonal matrix with $\ell s, \ell s$th entry $[(\ell)(\ell+1)]^{-m}$. The minimizing vector f_λ is

$$f_\lambda = (X'X + \lambda D^{-1})^{-1} X' y.$$

Methods of computing f_λ for large n and N will be discussed later. Splines on the sphere have found application to the interpolation and smoothing of geophysical and meteorological data. Historical data can be used to choose the $\lambda_{\ell s}$ (see, for example, Stanford (1979)).

2.3 Vector splines on the sphere.

Vector splines on the sphere, for use in smoothing vector fields on the sphere (such as horizontal wind, or magnetic fields), can also be defined. Let the vector field be $\mathbf{V} = (U, V)$ where $U = U(P)$ is the eastward component and $V = V(P)$ is the northward component at P. By the Helmholtz theorem, there exist two functions Ψ and Φ defined on $\mathcal{S}$, called the stream function and the velocity potential, respectively, with the property that

$$U = \frac{1}{a}\left(-\frac{\partial\Psi}{\partial\phi} + \frac{1}{\cos\phi}\frac{\partial\Phi}{\partial\theta}\right), \tag{2.3.1}$$

$$V = \frac{1}{a}\left(\frac{1}{\cos\phi}\frac{\partial\Psi}{\partial\theta} + \frac{\partial\Phi}{\partial\phi}\right),$$

where a is the radius of the sphere. Furthermore, letting the vorticity ζ and the divergence D of $\mathbf{V}$ be defined (as usual) by

$$\zeta = \frac{1}{a\cos\phi}\left[-\frac{\partial}{\partial\phi}(U\cos\phi) + \frac{\partial V}{\partial\theta}\right], \tag{2.3.2}$$

$$D = \frac{1}{a\cos\phi}\left[\frac{\partial U}{\partial\theta} + \frac{\partial}{\partial\phi}(V\cos\phi)\right],$$

we have

$$\zeta = \Delta\Psi, \qquad D = \Delta\Phi, \tag{2.3.3}$$

where now the (surface) Laplacian on the sphere of radius a is

$$\Delta f = \frac{1}{a^2}\left[\frac{1}{\cos^2\phi}f_{\theta\theta} + \frac{1}{\cos\phi}(\cos\phi f_\phi)_\phi\right]. \tag{2.3.4}$$

Ψ and Φ are uniquely determined up to a constant, which we will take to be determined by

$$\int_S \Psi(P)\,dP = \int_S \Phi(P)\,dP = 0. \tag{2.3.5}$$

Given data (U_i, V_i) from the model

$$\begin{aligned} U_i &= U(P_i) + \epsilon_i^U, \quad i = 1, 2, \ldots, n, \\ V_i &= V(P_i) + \epsilon_i^V \end{aligned} \tag{2.3.6}$$

where the ϵ_i^U and ϵ_i^V are random errors, one can define a vector smoothing spline for this data as $\mathbf{V}_{\lambda\delta} = (U_{\lambda,\delta}, V_{\lambda,\delta})$ where

$$\begin{aligned} U_{\lambda,\delta} &= \frac{1}{a}\left(-\frac{\partial\Psi_{\lambda,\delta}}{\partial\phi} + \frac{1}{\cos\phi}\frac{\partial\Phi_{\lambda,\delta}}{\partial\theta}\right), \\ V_{\lambda,\delta} &= \frac{1}{a}\left(\frac{1}{\cos\phi}\frac{\partial\Psi_{\lambda,\delta}}{\partial\theta} + \frac{\partial\Phi_{\lambda\delta}}{\partial\phi}\right), \end{aligned} \tag{2.3.7}$$

and $\Psi_{\lambda,\delta}, \Phi_{\lambda,\delta}$ are the minimizers of

$$\frac{1}{n}\sum_{i=1}^{n}\left(U_i - \frac{1}{a}\left[-\frac{\partial\Psi}{\partial\phi}(P_i) + \frac{1}{\cos\phi}\frac{\partial\Phi}{\partial\theta}(P_i)\right]\right)^2 \tag{2.3.8}$$

$$+\frac{1}{n}\sum_{i=1}^{n}\left(V_i - \frac{1}{a}\left[\frac{1}{\cos\phi}\frac{\partial\Psi}{\partial\theta}(P_i) + \frac{\partial\Phi}{\partial\phi}(P_i)\right]\right)^2$$

$$+\lambda\left[\int_S(\Delta^{m/2}\Psi)^2dP + \frac{1}{\delta}\int_S(\Delta^{m/2}\Phi)^2dP\right].$$

$\Delta^{m/2}\Psi$ can, of course, be defined for noninteger $m/2$. If

$$\Psi \sim \sum_{\ell s}\Psi_{\ell s}Y_{\ell s}$$

then

$$\Delta^{m/2}\Psi \sim \sum_{\ell s}[\ell(\ell+1)]^{m/2}\Psi_{\ell s}Y_{\ell s}$$

whenever the sum converges in quadratic mean.

An approximation to the minimizer of (2.3.8) may be obtained by letting Ψ and Φ be of the form

$$\Psi = \sum_{\ell=1}^{N}\sum_{s=-\ell}^{\ell}\alpha_{\ell s}Y_{\ell s}, \tag{2.3.9}$$

$$\Phi = \sum_{\ell=1}^{N}\sum_{s=-\ell}^{\ell}\beta_{\ell s}Y_{\ell s}.$$

Now let X_1 be the matrix with $i,\ell s$th entry $(\partial/\partial\phi)Y_{\ell s}(P)|_{P=P_i}$ and X_2 be the matrix with $i,\ell s$th entry $(1/\cos\phi)(\partial/\partial\theta)Y_{\ell s}(P)|_{P=P_i}$, and let $U = (U_1,\ldots,U_n)$; $V = (V_1,\ldots,V_n)$, then (2.3.8) becomes

$$\frac{1}{n}\|U - \frac{1}{a}(-X_1\alpha + X_2\beta)\|^2$$

$$+\frac{1}{n}\|V - \frac{1}{a}(X_2\alpha + X_1\beta)\|^2$$

$$+\lambda\left[\alpha' D\alpha + \frac{1}{\delta}\beta' D\beta\right]$$

where D is as in (2.2.12).

Given the noisy data U and V, one may estimate the vorticity and divergence as

$$\Delta\Psi_{\lambda,\delta} = \sum_{\ell=1}^{N}\sum_{s=-\ell}^{\ell}\ell(\ell+1)\alpha_{\ell s}^{\lambda,\delta}Y_{\ell s},$$

$$\Delta\Phi_{\ell,s} = \sum_{\ell=1}^{N}\sum_{s=-\ell}^{\ell}\ell(\ell+1)\beta_{\ell s}^{\lambda,\delta}Y_{\ell s}$$

where $\alpha^{\lambda,\partial}$ and $\beta^{\lambda,\partial}$ are the minimizing values of α and β in (2.3.10). This method was proposed in Wahba (1982b); see also Swarztrauber (1981).

2.4 The thin-plate spline on E^d.

The theoretical foundations for the thin-plate spline were laid by Duchon (1975, 1976, 1977) and Meinguet (1979), and some further results and applications to meteorological problems were given in Wahba and Wendelberger (1980). Other applications can be found in Hutchinson and Bischof (1983) and Seaman and Hutchinson (1985). In two dimensions ($d = 2$, $m = 2$, $f = f(x_1, x_2)$), the thin-plate penalty functional is

$$J_2(f) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (f_{x_1x_1}^2 + 2f_{x_1x_2}^2 + f_{x_2x_2}^2)\, dx_1\, dx_2 \tag{2.4.1}$$

and, in general,

$$J_m(f) = \sum_{\nu=0}^{m} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \binom{m}{\nu} \left(\frac{\partial^m f}{\partial x_1^\nu \partial x_2^{m-\nu}}\right)^2 dx_1\, dx_2. \tag{2.4.2}$$

For $d = 3$, $m = 2$, the thin-plate penalty functional is

$$\begin{aligned} J_2(f) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (f_{x_1x_1}^2 + f_{x_2x_2}^2 + f_{x_3x_3}^2 \\ &\quad + 2[f_{x_1x_2}^2 + f_{x_1x_3}^2 + f_{x_2x_3}^2])\, dx_1\, dx_2\, dx_3 \end{aligned} \tag{2.4.3}$$

and the form for general d, m is

$$\begin{aligned} J_m^d(f) = &\sum_{\alpha_1+\ldots+\alpha_d=m} \frac{m!}{\alpha_1!\ldots\alpha_d!} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \\ &\cdot \left(\frac{\partial^m f}{\partial x_1^{\alpha_1}\ldots\partial x_d^{\alpha_d}}\right)^2 \prod_j dx_j. \end{aligned} \tag{2.4.4}$$

A formula analogous to (2.2.3) holds here.

Letting

$$\begin{aligned} < f, g > = &\sum_{\alpha_1+\ldots+\alpha_d=m} \frac{m!}{\alpha_1!\ldots\alpha_d!} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \\ &\cdot \left(\frac{\partial^m f}{\partial x_1^{\alpha_1}\ldots dx_d^{\alpha_d}}\right)\left(\frac{\partial^m g}{\partial x_1^{\alpha_1}\ldots\partial x_d^{\alpha_d}}\right) \prod_j dx_j, \end{aligned} \tag{2.4.5}$$

we note that a formal integration by parts results in

$$< f, g > = (-1)^m \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f \cdot \Delta^m g + \mathcal{B}, \tag{2.4.6}$$

where $\mathcal{B}$ represents boundary values at infinity.

We will suppose $f \in \chi$, a space of functions whose partial derivatives of total order m are in $\mathcal{L}_2(E^d)$ (see Meinguet (1979) for more details on χ). We want χ

endowed with the seminorm $J_m^d(f)$ to be an r.k.h.s., that is, we want to have the evaluation functionals be bounded in $\mathcal{X}$. For this it is necessary and sufficient that $2m - d > 0$.

Now, let the data model be

$$y_i = f(x_1(i), \ldots, x_d(i)) + \epsilon_i \ , \ i = 1, \ldots, n, \tag{2.4.7}$$

where $f \in \mathcal{X}$ and $\epsilon = (\epsilon_1, \ldots, \epsilon_n)' \sim \mathcal{N}(0, \sigma^2 I)$. A thin-plate smoothing spline is the solution to the following variational problem. Find $f \in \mathcal{X}$ to minimize

$$\frac{1}{n}\sum_{i=1}^{n}(y_i - f(x_1(i), \ldots, x_d(i)))^2 + \lambda J_m^d(f). \tag{2.4.8}$$

We will use the notation $t = (x_1, \ldots, x_d)$ and $t_i = (x_1(i), \ldots, x_d(i))$. The null space of the penalty functional J_m^d is the $M = \begin{pmatrix} d+m-1 \\ d \end{pmatrix}$-dimensional space spanned by the polynomials in d variables of total degree $\le m-1$. For example, for $d = 2$, $m = 2$, then $M = 3$ and the null space is spanned by ϕ_1, ϕ_2, and ϕ_3 given by

$$\phi_1(t) = 1, \ \phi_2(t) = x_1, \ \phi_3(t) = x_2.$$

In general, we will denote the M monomials of total degree less than m by $\phi_1, \ldots, \phi_M$.

Duchon (1977) showed that, if $t_1, \ldots, t_n$ are such that least squares regression on $\phi_1, \ldots, \phi_M$ is unique, then (2.4.8) has a unique minimizer f_λ, with representation

$$f_\lambda(t) = \sum_{\nu=1}^{M} d_\nu \phi_\nu(t) + \sum_{i=1}^{n} c_i E_m(t, t_i), \tag{2.4.9}$$

where E_m is a Green's function for the m-iterated Laplacian. Letting

$$\begin{aligned} E(\tau) &= \theta_{m,d}|\tau|^{2m-d}\ln|\tau| \quad \text{if } 2m-d \text{ an even integer,} \\ &= \theta_{m,d}|\tau|^{2m-d} \quad \text{otherwise,} \end{aligned} \tag{2.4.10}$$

where

$$\theta_{m,d} = \frac{(-1)^{d/2+1+m}}{2^{2m-1}\pi^{d/2}(m-1)!(m-d/2)!} \quad \text{if } 2m-d \text{ is an even integer,}$$

$$\theta_{m,d} = \frac{\Gamma(d/2-m)}{2^{2m}\pi^{d/2}(m-1)!} \quad \text{otherwise,}$$

and letting $|t - t_i| = (\sum_{j=1}^{d}(x_j - x_j(i))^2)^{1/2}$, $E_m(s,t)$ is given by

$$E_m(s,t) = E(|s-t|). \tag{2.4.11}$$

Formally,

$$\Delta^m E_m(\cdot, t_i) = \delta_{t_i}, \tag{2.4.12}$$

where δ_{t_i} is the Dirac delta function, so that

$$\Delta^m f_\lambda(t) = 0 \quad \text{for } t \neq t_i,\ i = 1, \ldots, n, \tag{2.4.13}$$

analogous to the univariate polynomial spline case where $(\partial^{2m}/\partial x^{2m}) f_\lambda(x) = 0$ for $x \neq x_1, \ldots, x_n$. The functions $E_m(t, t_i)$, $i = 1, \ldots, n$, play the same role as $\xi_i(t) = R^1(t_i, t)$ in Section 1.3, except that $E_m(\cdot,\cdot)$ is not positive definite. $E_m(\cdot,\cdot)$ is *conditionally positive definite*, a property that turns out to be enough. To explain the notion of conditional positive definiteness, we need the notion of a generalized divided difference. Given $t_1, \ldots, t_n \in E^d$, let T be the $n \times M$ matrix with $i\nu$th entry $\phi_\nu(t_i)$. In one dimension, T is always of full column rank if the t_i's are distinct. In two and higher dimensions it must be an explicit assumption that T is of full column rank, which we will always make. If, for example, $t_1, \ldots, t_n$ fall on a straight line on the plane this assumption will fail to hold. Now let $c \in E^n$ be any vector satisfying $T'c = 0$. Then $(c_1, \ldots, c_n)$, associated with $t_1, \ldots, t_n$, is called a generalized divided difference (g.d.d.) of order m, since it annihilates all polynomials of total degree less than m, that is, $\sum_{i=1}^n c_i \phi_\nu(t_i) = 0$, $\nu = 1, 2, \ldots, M$. Recall that the ordinary first-order divided differences are of the form $(f(t_{i+1}) - f(t_i))/(t_{i+1} - t_i)$ and annihilate constants, second-order divided differences are of the form

$$\left(\frac{f(t_{i+2}) - f(t_{i+1})}{t_{i+2} - t_{i+1}} - \frac{f(t_{i+1}) - f(t_i)}{t_{i+1} - t_i} \right) / (t_{i+2} - t_i)$$

and annihilate constants and linear functions, and so forth, thus a g.d.d. is a generalization of an ordinary divided difference.

Duchon (1977) and Matheron (1973) both have proved the following: Given $t_1, \ldots, t_n$ such that T is of rank M, let $K_{n \times n}$ be the $n \times n$ matrix with ijth entry $E_m(t_i, t_j)$. Then

$$c'Kc > 0 \tag{2.4.14}$$

for any g.d.d. c of order m, that is, for any c such that $T'c = 0$. E_m is then called m-conditionally (strictly) positive definite.

Now, let $E_t(\cdot) = E_m(t, \cdot)$ and write

$$\begin{aligned} < E_t, E_s > \quad = & \sum_{\alpha_1 + \ldots + \alpha_d = m} \frac{m!}{\alpha_1! \ldots \alpha_d!} \\ & \cdot \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \frac{\partial^m E_t}{\partial x_1^{\alpha_1} \ldots \partial x_d^{\alpha_d}} \cdot \frac{\partial^m E_s}{\partial x_1^{\alpha_1} \ldots \partial x_d^{\alpha_d}} \prod_j dx_j. \end{aligned} \tag{2.4.15}$$

A formal integration by parts yields

$$\begin{aligned} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} E_t(u) \Delta^m E_s(u)\, du + \mathcal{B} &= E_t(s) + \mathcal{B} \\ &= E_m(s, t) + \mathcal{B}. \end{aligned} \tag{2.4.16}$$

This calculation is not legitimate, since the boundary values at ∞ will be infinite. However, it is known from the work of Duchon and Meinguet that if we let

$$g(s) = \sum_{i=1}^{n} c_i E_m(s, t_i)$$

where $c = (c_1, \ldots, c_n)'$ is a g.d.d., then g has appropriate behavior at infinity and we can write

$$\begin{aligned} < g, g > &= < \sum_{i=1}^{n} c_i E_m(\cdot, t_i), \sum_{j=1}^{n} c_j E_m(\cdot, t_j) > \\ &= \sum_{i,j} c_i c_j \; E_m(t_i, t_j) \\ &= c'Kc > 0. \end{aligned} \tag{2.4.17}$$

By substituting (2.4.9) into (2.4.8) and using (2.4.17), we obtain that c, d are the minimizers of

$$\frac{1}{n}\|y - Td - Kc\|^2 + \lambda c'Kc \tag{2.4.18}$$

subject to $T'c = 0$. To find the minimizers c and d of this expression, we let the QR decomposition of T be

$$T = (Q_1 : Q_2) \begin{pmatrix} R \\ 0 \end{pmatrix} \tag{2.4.19}$$

where $(Q_1 : Q_2)$ is orthogonal and R is lower triangular. Q_1 is $n \times M$ and Q_2 is $n \times (n - M)$. Since $T'c = 0$, c must be in the column space of Q_2, $c = Q_2\gamma$ for some $m - M$ vector γ. By the orthogonality of $(Q_1 : Q_2)$ we have $\|x\|^2 = \|Q_1'x\|^2 + \|Q_2'x\|^2$ for any $x \in E^n$. Using this and substituting $Q_2\gamma$ for c in (2.4.18) gives

$$\frac{1}{n}\|Q_2'y - Q_2'KQ_2\gamma\|^2 + \frac{1}{n}\|Q_1'y - Rd - Q_1'KQ_2\gamma\|^2 + \lambda\gamma'Q_2'KQ_2\gamma. \tag{2.4.20}$$

It is seen that the minimizers d and γ satisfy

$$Rd = Q_1'(y - KQ_2\gamma) \tag{2.4.21}$$

and

$$Q_2'y = (Q_2'KQ_2 + n\lambda I)\gamma. \tag{2.4.22}$$

These relations can be seen to be equivalent to

$$Mc + Td = y, \tag{2.4.23}$$

$$T'c = 0 \tag{2.4.24}$$

with $M = K + n\lambda I$, by multiplying (2.4.23) by Q_2' and letting $c = Q_2\gamma$. The columns of Q_2 are all the g.d.d.'s.

It is possible to come to the same result for the minimizer of (2.4.8) via reproducing kernels. Let $s_1, \ldots, s_M$ be any M fixed points in E^d such that least squares regression on the M-dimensional space of polynomials of total degree less than m at the points $s_1, \ldots, s_M$ is unique, that is, the $M \times M$ matrix S, with $i\nu$th entry $\phi_\nu(s_i)$ is of full rank. (In this case we call the points $s_1, \ldots, s_M$ unisolvent.) Let $p_1, \ldots, p_M$ be the (unique) polynomials of total degree less than m satisfying $p_i(s_j) = 1,\ i = j, = 0, i \neq j$, and let

$$\begin{aligned} R^1(s,t) = E_m(s,t) \quad &- \sum_{\nu=1}^{M} p_\nu(t) E_m(s_\nu, s) \\ &- \sum_{\mu=1}^{M} p_\mu(s) E_m(t, s_\mu) \\ &+ \sum_{\mu,\nu=1}^{M} p_\mu(s) p_\nu(t) E_m(s_\mu, s_\nu). \end{aligned} \tag{2.4.25}$$

Letting

$$R_t^1(\cdot) = R^1(\cdot, t) \tag{2.4.26}$$

we have

$$R_t^1(s) = E_m(s,t) - \sum_{\nu=1}^{M} p_\nu(t) E_m(s_\nu, s) + \pi_t(s) \tag{2.4.27}$$

where for fixed t, $\pi_t(\cdot)$ is a polynomial of degree $m-1$ in s. Now for fixed t, consider the points $(t, s_1, \ldots, s_M)$ and coefficients $(1, -p_1(t), \ldots, -p_M(t))$. These coefficients together with the points $(t, s_1, \ldots, s_M)$ constitute a g.d.d., since

$$\phi(t) - \sum_{\nu=1}^{M} \phi(s_\nu) p_\nu(t) \equiv 0 \tag{2.4.28}$$

for any polynomial ϕ of total degree less than m. Equation (2.4.28) follows since the sum in (2.4.28) is a polynomial of degree less than m that interpolates to ϕ at a set of M unisolvent points, therefore it must be zero. In fact,

$$< R_t^1, R_s^1 >= R^1(s,t).$$

$R^1(s_\nu, s_\nu) = 0,\ \nu = 1, 2, \ldots, M$, but R^1 is positive semidefinite. R^1 is an r.k. for $\mathcal{H}_1$, the subspace of $\mathcal{X}$ of codimension M of functions satisfying $f(s_\nu) = 0,\ \nu = 1, 2, \ldots, M$, and $\mathcal{X}$ is the direct sum of $\mathcal{H}_0 = \text{span}\,\{p_1, \ldots, p_M\}$ and $\mathcal{H}_1$ with $J_m^d(f) = \|P_1 f\|^2$. It follows from Section 1.3 that f_λ has a representation

$$f_\lambda(t) = \sum_{\nu=1}^{M} d_\nu \phi_\nu(t) + \sum_{i=1}^{n} c_i R_{t_i}^1(t) \tag{2.4.29}$$

for some d, c, with $T'c = 0$. The end result from (2.4.29) can be shown to be the same as (2.4.9) since $\sum_{i=1}^n c_i E_m(t, t_i)$ and $\sum_{i=1}^n c_i R_{t_i}^1(t)$ differ by a polynomial of total degree less than m in t if c is a g.d.d.

2.5 Another look at the Bayes model behind the thin-plate spline.

Returning to (1.5.7), consider the "fixed effects" model

$$Y_i = F(t_i) + \epsilon_i, \; i = 1, 2, \ldots, n$$

where

$$F(t) = \sum_{\nu=1}^{M} \theta_\nu \phi_\nu(t) + b^{1/2} X(t), \; t \in \mathcal{T}. \tag{2.5.1}$$

One such model that will result in the thin-plate spline is: $\{\phi_1, \ldots, \phi_M\}$ span $\mathcal{H}_0$, the space of polynomials of total degree less than m, and

$$EX(s)X(t) = R^1(s,t),$$

with $R^1(s,t)$ given by (2.4.25). The points $s_1, \ldots, s_M$ used in defining R^1 were arbitrary, and it can be seen that it is not, in fact, necessary to know the entire covariance of $X(t)$, $t \in \mathcal{T}$. Any covariance for which the g.d.d.'s of X satisfy

$$E\sum_l c_l X(t_l) \sum_k c_k X(t_k) = \sum_{lk} c_l c_k E_m(t_l, t_k)$$

whenever $\sum_\ell c_\ell \phi_\nu(t_l) = 0, \nu = 1, 2, \ldots, M$, will result in the *same* thin-plate spline. Looking at this phenomenon from another point of view, one can replace $X(t), t \in \mathcal{T}$ in the model (2.5.1) by $\tilde{X}(t) = X(t) + \sum_{\nu=1}^{M} \tilde{\theta}_\nu \phi_\nu(t)$ where the $\tilde{\theta}_\nu$ are arbitrary, without changing the model. The estimation procedure assigns as much of the "explanation" of the data vector as possible to θ in (2.5.1) and not $\tilde{\theta}$. This kind of reasoning was behind the development of "kriging" due to a South African mining engineer, David Krige (see the references in Delfiner (1975)). The motivation for Krige's work was to estimate the total ore content of a volume of earth from observations from core samples. It was assumed that the ore density was a random process $Y(t), t \in E^d$, whose generalized divided differences were stationary, and that it had a so-called variogram $\tilde{E}(\tau)$ with the property that

$$E\sum_l c_l Y(s_l) \sum_k c_k Y(s_k) = \sum_l \sum_k c_l c_k \tilde{E}|s_l - s_k|$$

whenever the $\{c_l, s_l\}$ constituted a g.d.d., and a "drift," or mean-value function, of the form

$$\sum_{\nu=1}^{M} \theta_\nu \phi_\nu.$$

The kriging estimate $\hat{Y}(t)$, $t \in T$, was defined as the minimum variance, conditionally unbiased (with respect to θ) linear estimate of $Y(t)$ given $Y_i = y_i$, and if $\hat{Y}(t)$ is the estimate of $Y(t)$, then $\int_\Omega \hat{Y}(t)\,dt$ is the conditionally unbiased, minimum variance estimate of $\int_\Omega Y(t)\,dt$ (compare Section 1.5).

This connection between spline estimation and kriging was demonstrated in Kimeldorf and Wahba (1971, §7,) although the word kriging was never mentioned. We had not heard of it at the time. Duchon (1975, 1976) gave a general version of this result in French, and various connections between the two lines of research, which have been carried out fairly independently until the last few years, have been rediscovered a number of times.

Matheron (1973) characterized the class of k-conditionally positive-definite functions on E^d, in particular, letting

$$K(\tau) = \sum_{p=0}^{k} (-1)^{p+1} a_p \tau^{2p+1} \tag{2.5.2}$$

where the coefficients a_p satisfy

$$\sum_{p=0}^{k} \frac{a_p}{\pi^{2p+2+d/2}} \frac{\Gamma(\frac{1}{2}(2p+1+d))}{\Gamma[1+\frac{1}{2}(2p+1)]} \rho^{-d-2p+1} \geq 0 \tag{2.5.3}$$

for any $\rho \geq 0$. Matheron showed that

$$E(s,t) = K(|s-t|)$$

is k-conditionally positive definite. (Note that if $E(s,t)$ is k-conditionally positive definite, it is $k+1$ conditionally positive definite.) Much of the work on kriging involves variograms of the form

$$E(s,t) = |s-t|^3 - \beta|s-t|$$

or

$$E(s,t) = |s-t|^5 - \beta_1|s-t|^3 + \beta_2|s-t|,$$

where the β's are estimated from the data. See Delfiner (1975), Journel and Huijbregts (1978), and Cressie and Horton (1987).

We will now make some remarks concerning the variational problem associated with generalized covariances of the form $E(s,t) = K(|s-t|)$, where K is as in (2.5.2). It is easy to see what happens to the analogous case on the circle. Letting $s,t \in [0,1]$, let

$$\mathcal{E}_m(s,t) = 2\sum_{\nu=1}^{\infty} \frac{\cos 2\pi\nu(s-t)}{(2\pi\nu)^{2m}}$$

and let

$$R(s,t) = \sum_{l=m}^{m+k} \alpha_{l-m}\mathcal{E}_l(s,t) \tag{2.5.4}$$

with $\alpha_0 = 1$. Then the eigenvalues of $\mathcal{E}_m$ are $\lambda_\nu(\mathcal{E}_m) = (2\pi\nu)^{-2m}$ and the eigenvalues $\lambda_\nu(R)$ of R are

$$\lambda_\nu(R) = \sum_{l=m}^{m+k} \frac{\alpha_{l-m}}{(2\pi\nu)^{2l}} = \frac{1}{(2\pi\nu)^{2(m+k)}} \sum_{j=0}^{k} \alpha_{k-j}(2\pi\nu)^{2j}. \tag{2.5.5}$$

In order for R of (2.5.4) to be a covariance, it is necessary that $\lambda_\nu(R) \geq 0$, for this it is sufficient that $\sum_{j=0}^{k} \alpha_{k-j}\rho^{2j} \geq 0$ for all $\rho > 0$. In the discussion below, we will assume that the α_j's are such that $\lambda_\nu(R) > 0$. Letting $f_\nu^2 = c_\nu^2 + s_\nu^2$ where $c_\nu = \sqrt{2}\int f(t)\cos 2\pi\nu t$, $s_\nu = \sqrt{2}\int f(t)\sin 2\pi\nu t$, then we have that the squared norms associated with $\mathcal{E}_m$ and R are, respectively,

$$\sum_{\nu=1}^{\infty}(2\pi\nu)^{2m} f_\nu^2 \tag{2.5.6}$$

and

$$\sum_{\nu=1}^{\infty} \frac{(2\pi\nu)^{2(m+k)}}{\sum_{j=0}^{k}\alpha_{k-j}(2\pi\nu)^{2j}} f_\nu^2 = \sum_{\nu=1}^{\infty}(2\pi\nu)^{2m} H(\nu) f_\nu^2 \tag{2.5.7}$$

where

$$H(\nu) = \frac{(2\pi\nu)^{2k}}{\sum_{j=0}^{k}\alpha_{k-j}(2\pi\nu)^{2j}} = \left(1 + \sum_{j=1}^{k} \frac{\alpha_j}{(2\pi\nu)^{2j}}\right)^{-1}$$

As $\nu \to \infty$, $H(\nu) \to 1$. If $\sum_{j=1}^{k}(\alpha_j/(2\pi\nu)^{2j})$ is bounded strictly above -1, then the two norms satisfy

$$a\|f\|^2_{\mathcal{H}_{\mathcal{E}_m}} \leq \|f\|^2_{\mathcal{H}_R} \leq b\|f\|^2_{\mathcal{H}_{\mathcal{E}_m}}$$

for some $0 < a \leq b < \infty$. Then $f \in \mathcal{H}_R$ if and only if $f \in \mathcal{H}_{\mathcal{E}_m}$ and the two spaces are topologically equivalent.

We now return to E^d and the thin plate penalty functional. Since we are running out of symbols we will use $(\alpha_1, \ldots, \alpha_d)$ as a multi-index below, not to be confused with the α's in the definition of R above. Observing that the Fourier transform of $\partial^m f/\partial x_1^{\alpha_1}\ldots\partial x_d^{\alpha_d}$ is

$$\widehat{\frac{\partial^m f}{\partial x_1^{\alpha_1}\ldots\partial x_d^{\alpha_d}}} = \prod_{l=1}^{d}(2\pi i w_l)^{\alpha_l}\hat{f}(w_1,\ldots,w_d)$$

where "$\hat{f}$" denotes Fourier transform and that

$$(|w_1|^2 + \ldots + |w_d|^2)^m = \sum_{\alpha_1+\ldots+\alpha_d} \frac{m!}{\alpha_1!\ldots\alpha_d!}\prod_{l=1}^{d}|w_l|^{2\alpha_l}$$

we have

$$J_m^d(f) = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} \|2\pi w\|^{2m}|\hat{f}(w)|^2 \prod_l dw_l. \tag{2.5.8}$$

The argument below is loosely adapted from a recent thesis by Thomas-Agnan (1987), who considered general penalty functionals of the form

$$\int \alpha^2(w)|\hat{f}(w)|^2\,dw.$$

The penalty functional associated with the variogram $\sum_{l=m}^{m+k} \alpha_{l-m} E_l$ where E_l is defined as in (2.4.11) must be

$$\tilde{J}_m^d(f) = \int \dots \int \frac{\|2\pi w\|^{2(m+k)}}{\sum_{j=0}^k \alpha_{k-j} \|2\pi w\|^{2j}} |\hat{f}(w)|^2 \prod_l dw_l. \qquad (2.5.9)$$

This can be conjectured by analogy to (2.5.7). For rigorous details in the $d = 1$ case, see Thomas-Agnan (1987).

In going from the circle to E^d we must be a little more careful, however. Letting

$$H(w) = \frac{\|2\pi w\|^{2k}}{\|2\pi w\|^{2k} + \alpha_1 \|2\pi w\|^{2(k-1)} + \dots + \alpha_k},$$

we have

$$\tilde{J}_m^d(f) = \int \dots \int H(w) \|2\pi w\|^{2m} |\hat{f}(w)|^2 dw,$$

since $H(w) \to 1$ as $w \to \infty$, the tail behavior of the Fourier transforms of functions for which J_m^d and $\tilde{J}_m^d$ are finite will be the same. This tail behavior ensures that the total derivatives of order m are in $\mathcal{L}_2$. However, we have $\lim_{w \to 0} H(w) = 0$, in particular, $\lim_{w \to 0}(H(w)\|2\pi w\|^{2m})/\|2\pi w\|^{2(m+k)} \to 1/\alpha_k$. It can be argued heuristically that the polynomials of total degree less than $m+k$ are in the null space of $\tilde{J}_m^d(f)$, by writing

$$\begin{aligned} \tilde{J}_m^d(f) &= \int \dots \int \frac{\|2\pi w\|^{2(m+k)} |\hat{f}(w)|^2}{\|2\pi w\|^{2k} + \alpha_1 \|2\pi w\|^{2(k-1)} + \dots + \alpha_k} \prod_l dw_l \\ &= \int \dots \int \frac{\|\widehat{\partial^{2(m+k)} f}\|^2}{\|2\pi w\|^{2k} + \alpha_1 \|2\pi w\|^{2(k-1)} + \dots + \alpha_k} \prod_l dw_l \end{aligned} \qquad (2.5.10)$$

where $\widehat{\partial^{2(m+k)} f}$ is the Fourier transform of the $2(m+k)$th total derivative of f. If (2.5.10) is valid then $\tilde{J}_m^d(f) = 0$ for f a polynomial of total degree less than $m+k$.

Let $\tilde{s}_1, \dots, \tilde{s}_{\tilde{M}}$ be a unisolvent set of

$$\tilde{M} = \binom{d + (m+k) - 1}{d}$$

points in E^d, and let

$$E(s,t) = \sum_{l=m}^{m+k} \alpha_{l-m} E_l(s,t)$$

where the α's satisfy conditions ensuring that $E(s,t)$ is $m+k$ conditionally positive definite. Let $p_1, \dots, p_{\tilde{M}}$ be the $\tilde{M}$ polynomials satisfying $p_i(\tilde{s}_j) = 1$, $i =$

j, and 0, $i \neq j$, and let

$$\begin{aligned} R^1(s,t) = E(s,t) &- \sum_{\nu=1}^{\tilde{M}} \tilde{p}_\nu(t) E(\tilde{s}_\nu, s) \\ &- \sum_{\mu=1}^{\tilde{M}} \tilde{p}_\mu(s) E(t, \tilde{s}_\mu) \\ &+ \sum_{\mu,\nu=1}^{\tilde{M}} \tilde{p}_\mu(s)\tilde{p}_\nu(t) E(\tilde{s}_\mu, \tilde{s}_\nu). \end{aligned}$$

One can argue analogously to Section 2.4 that R^1 must be a positive-definite function that has the reproducing kernel property under the norm defined by $\tilde{J}_m^d(f)$.