SPLINES AND LOCAL APPROXIMATION OF THE EARTH'S GRAVITY FIELD

HERMANUS GERHARDUS VAN GYSEN

Submitted to the University of Cape Town
in fulfilment of the requirements for the
Degree of Doctor of Philosophy.

Cape Town

July 1988

SPLINES AND LOCAL APPROXIMATION OF THE EARTH'S GRAVITY FIELD

Herman van Gysen

Department of Surveying and Mapping

University of Natal

King George V Avenue

Durban, 4001, South Africa

ABSTRACT

The Hilbert space spline theory of Delvos and Schempp, and the reproducing kernel theory of L. Schwartz, provide the conceptual foundation and the construction procedure for rotation-invariant splines on Euclidean spaces, splines on the circle, and splines on the sphere and harmonic outside the sphere.

Spherical splines and surface splines such as multi-conic functions, Hardy's multiquadric functions, pseudo-cubic splines, and thin-plate splines, are shown to be largely as effective as least squares collocation in representing geoid heights or gravity anomalies. A pseudo-cubic spline geoid for southern Africa is given, interpolating Doppler-derived geoid heights and astro-geodetic deflections of the vertical. Quadrature rules are derived for the thin-plate spline approximation (over a circular disk, and to a planar approximation) of Stokes's formula, the formulae of Vening Meinesz, and the L_1 vertical gradient operator in the analytical continuation series solution of Molodensky's problem.

July 1988

ACKNOWLEDGEMENTS

I am indebted to Professor Charles Merry for years of encouragement and support.

Ms A. Allardice prepared many of the figures. The computations were carried out on the Unisys 1100 systems at the University of Cape Town and at the University of Natal.

In 1984 and 1985 I enjoyed the support of a CSIR/FRD postgraduate bursary.

For the data used in Chapter 6 I am grateful to Professor C.L. Merry; the Chief Director of Surveys and Mapping, Mowbray, and Mr D.P.M. Rousseau; the Surveyor-General, Harare; and the Director of Lands, Gaborone.

A final acknowledgement, and a dedication, is to Niek van Gysen († 11 August 1980), who taught me to take an equal interest in the natural world and the world of ideas.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

GUIDE TO NOTATION

2 INTRINSIC SPLINE THEORY

F	observation mapping, epimorphism $X \to Z$;
J_X, J_Y	canonical duality isomorphisms between X and X^*, Y and Y^*;
P	spline projector in X, $\|\cdot\|_X$, (2.20);
P_U	spline projector in Y, (2.13);
$P^*, P_U{}^*$	transpose mappings to P, P_U;
S	smoothing mapping, epimorphism $X \to W$, (2.8);
U	co-observation mapping, epimorphism $X \to Y$;
V	isomorphism between Y, Y^*, (2.27);
$W, \|\cdot\|_W$	product space $Y \times Z$, with norm derived from (2.7);
$X, \|\cdot\|$	spline ground space with given norm;
$X, \|\cdot\|_X$	spline ground space with 'initial' norm, (2.18);
X^*, Y^*	dual spaces of $X, \|\cdot\|_X$; $Y, \|\cdot\|_Y$;
$Y, \|\cdot\|_Y$	energy space, range space of U;
$Z, \|\cdot\|_Z$	observation space, range space of F.

3 CONSTRUCTIVE SPLINE THEORY

δ_x	Dirac measure at the point $x \in \Omega$;
Ω	a domain, an open bounded set;
θ	Riesz mapping $\mathbb{H}' \to H$, (3.4) or (3.7);
A	operator associated with the Sard system (X, Y, Z, U, F);
B	bilinear form on $E' \times E'$ induced by H;

E,E'	locally convex topological vector space over $\mathbb{R}$, quasi-complete; its dual;
G	Green function of A, (3.20);
H	kernel associated with $\mathbb{H}$;
$\mathbb{H}$	a Hilbert sub-space of E, or a semi-Hilbert sub-space of E;
$\mathrm{Hilb}(E)$	set of all Hilbert sub-spaces of E;
J,J^*	inclusion map of $\mathbb{H}$ into E; its transpose;
K	reproducing kernel $\Omega\times\Omega \to \mathbb{R}$, (3.12);
$L^+(E',E)$	space of positive continuous linear mappings from E' to E;
M,M°	linear span of the functionals making up F; its orthogonal in E;
N,N°	nullspace of the semi-Hilbert sub-space $\mathbb{H}$; its orthogonal in E';
Q	positive quadratic form on E' induced by H;
$\mathbb{R}^\Omega$	space of all real-valued functions on Ω given the topology of pointwise convergence;
U_0,U_0^*	restriction of U to X_0; its adjoint;
X_0	KerF, Section 3.6.

4 NATURAL ROTATION-INVARIANT SPLINES ON $\mathbb{R}^n$

$\alpha,\lvert\alpha\rvert$	multi-index, $\alpha \in (\mathbb{Z}^+)^n$; order of α;
β	multi-index, $\beta \in n^m$;
Ω	a domain in $\mathbb{R}^n$;
$\tilde{\phi}$	$\tilde{\phi}(y) = \phi(-y)$, (4.9);
Φ_n	space of rapidly decreasing functions on $\mathbb{R}^n$, Schwartz space;

Φ'_n space of tempered distributions;

τ_x shift operator, (4.9);

θ kernel associated with $D^{-m}\tilde{H}^s(\mathbb{R}^n)$, $m - \frac{n}{2} < s < \frac{n}{2}$, (4.57);

θ_p kernel associated with $\mathbb{P}_{m-1}$, (4.89);

θ_0 kernel, $\theta_0 = \theta - \theta_p$, (4.90);

$C(\Omega)$ space of complex-valued continuous functions on Ω;

$C_0(\mathbb{R}^n)$ continuous functions on $\mathbb{R}^n$ that vanish at infinity;

$C^\infty(\Omega)$ complex-valued functions on Ω with continuous derivatives of every order;

$C_0^\infty(\Omega)$ C^∞-functions on Ω with compact support;

dm_n normalised Lebesgue measure on $\mathbb{R}^n$, (4.19);

$d\mu_s$ measure on $\mathbb{R}^n$ given by (4.36);

D_α differential operator $(i)^{-|\alpha|}D^\alpha$, (4.33);

$D^{-m}H^s(\mathbb{R}^n)$ distributions u with $D^\alpha u \in H^s(\mathbb{R}^n)$, $|\alpha| = m$;

$D^{-m}\tilde{H}^s(\mathbb{R}^n)$ distributions u with $D^\alpha u \in \tilde{H}^s(\mathbb{R}^n)$, $|\alpha| = m$, (4.50);

$\mathbb{D}(\Omega)$ $C_0^\infty(\Omega)$ given the canonical Schwartz topology, the space of test functions on Ω;

$\mathbb{D}'(\Omega)$ dual space of $\mathbb{D}(\Omega)$, space of distributions on Ω;

e_t character function, (4.22);

$H^s(\mathbb{R}^n)$ Sobolev space (4.37), $s \in \mathbb{R}$;

$H_K^s(\mathbb{R}^n)$ distributions in $H^s(\mathbb{R}^n)$ with support in K, a closed set in $\mathbb{R}^n$;

$H^s(\Omega)$ restrictions to Ω of distributions in $H^s(\mathbb{R}^n)$, Ω an open set in $\mathbb{R}^n$;

$H_{\bar{\Omega}}^{-s}(\Omega)$ dual of $H^s(\Omega)$;

$H^s_{loc}(\mathbb{R}^n)$	distributions on $\mathbb{R}^n$ whose restrictions to a bounded open set Ω are in $H^s(\Omega)$;
$H^{-s}_{comp}(\mathbb{R}^n)$	dual of $H^s_{loc}(\mathbb{R}^n)$;
$\tilde{H}^s(\mathbb{R}^n)$	Sobolev space (4.42), $s < \frac{n}{2}$;
$K_{2m+2s-n}$	kernel function associated with $D^{-m}\tilde{H}^s(\mathbb{R}^n)$, (4.62);
$\overline{K}_{2m+2s-n}$	regularised kernel function, Section 4.12;
K_0	reproducing kernel function, (4.91);
$L^p(\mathbb{R}^n)$	Lebesgue space, $1 \leq p < \infty$, with norm (4.20);
$\tilde{p}_j$	element of orthonormal dual basis to $\mathbb{P}_{m-1}$, (4.87);
P	polynomial in n variables, Section 4.4;
Pf	pseudo-function or 'partie finie';
$\mathbb{P}_{m-1}$	space of polynomials on $\mathbb{R}^n$ of degree m-1 or less;
$\mathbb{Z}^+$	non-negative integers.

5 PERIODIC, SPHERICAL AND HARMONIC SPLINES

∇_*	surface gradient operator, (5.23);
Δ^*	Laplace-Beltrami operator, (5.24);
Ω	unit sphere;
Π	unit circle;
θ	kernel associated with $\mathbb{H}^m(\Pi)$, (5.10), or with $H^m(\Omega)$, (5.34);
$\mathbb{D}(\Pi)$	space of functions on Π whose 2π-periodic extensions to $\mathbb{R}$ are in $C^\infty(\mathbb{R})$;
$\mathbb{D}'(\Pi)$	space of distributions on Π;
$\mathbb{D}(\Omega)$	Schwartz space of C^∞-functions on Ω;
$\mathbb{D}'(\Omega)$	space of distributions on Ω;

$H^m(\Omega)$	space of distributions on Ω whose m'th surface gradients are square-integrable, (5.25);
$\mathbb{H}^m(\Pi)$	space of distributions on Π whose m'th derivatives are square-integrable, (5.9);
k_n	n'th degree variance of gravity field representation;
$K_{2m,0}$	reproducing kernel function for $\mathbb{H}^m(\Pi)$, (5.16);
K_{2m}	kernel function for $\mathbb{H}^m(\Pi)$, (5.11);
K_m	kernel function for $H^m(\Omega)$, (5.35);
$K_{m,0}$	reproducing kernel function for $H^m(\Omega)$, (5.41);
$K_{m,B}$	regularised kernel function for $H^m(\Omega)$, (5.45), (5.46);
K_0	Szegö kernel, (5.47);
$K_{K,B}$	Krarup kernel, (5.52);
P_n	Legendre polynomial of degree n;
$\mathbb{P}_m$	space of spherical harmonics of degree m or less;
R, R_B	radius of spherical earth; radius of Bjerhammer sphere;
s_n	spherical harmonic of degree n;
$s_{n,j}$	element of orthonormal basis for S_n;
S_n	space of spherical harmonics of degree n.

6 SPLINE APPROXIMATION OF THE LOCAL GRAVITY FIELD

Δ_g	spherical gravity anomaly operator, (6.12);

ζ_n n'th term in the Molodensky series for height anomaly, (6.23);

$\zeta_{n,I}$ contribution of circular innermost zone to ζ_n, (6.28);

$\zeta_{n,I,s}$ spline approximation of $\zeta_{n,I}$, (6.31);

ξ_n,η_n n'th terms in the Molodensky series for the deflection of the vertical, (6.24);

$\xi_{n,I},\eta_{n,I}$ contribution of circular innermost zone to ξ_n,η_n, (6.29);

$\xi_{n,I,s},\eta_{n,I,s}$ spline approximation of $\xi_{n,I},\eta_{n,I}$, (6.32);

$K_{col,B}$ kernel function for Tscherning and Rapp's 1974 geopotential degree variance model, (6.6);

L_1 vertical gradient operator (6.25);

$L_{1,I}$ contribution of circular innermost zone to L_1;

$L_{1,I,s}$ spline approximation of $L_{1,I}$, (6.33).

CHAPTER 1

INTRODUCTION

1.1 The Global Gravity Field

The external gravity field of the earth has long been an object of scientific and practical geodetic interest. The gravity field is inextricably part of all geodetic surveying measurements. This was true of the methods of the great age of geodetic triangulation; it is even more so of modern surveying techniques. Satellite altimetry, and satellite methods of geodetic positioning, for example, require information about the global gravity field - so that satellite orbits can be predicted - and at the other extreme, as has been made clear for us by *Engelis et al (1984)*, *Kearsley (1984)*, and *Schwarz et al (1987)*, very local gravity field information is required if orthometric height differences are to be obtained from differential GPS positioning.

The gravity field is also of interest in itself, for the information it contains about the shape, and internal structure of the earth. At a global level the flattening of the earth places constraints on its internal strength. At a regional level it is possible to infer something of the structure of the crust, and possibly something of the inhomogeneities of the upper mantle. Locally the gravity field responds to nearby, shallower features. Information about the dynamic behaviour of the earth - plate kinematics, post-glacial isostatic uplift, or the tidal response of the earth, for example - can also be sought in the gravity field.

Until recently, measurements from which the gravity field could be inferred were restricted to astro-geodetic methods, gravimetry and gravity gradiometry (using torsion balances). Each has in common that it is a point measurement responding very much to the local behaviour of the gravity field. Moreover, these measurements require for their utilisation additional geodetic data on position or elevation. It is difficult to establish a dense, world-wide system of measurements, all within a single, well-defined reference system.

From the point of view of global gravity field determination it is important, therefore, that satellite methods - both those already in use, and especially those now being planned: ranging to satellites, satellite radar altimetry (over ocean areas), satellite-to-satellite tracking, satellite gravity gradiometry - give precise, uniform data from which the global characteristics of the gravity field can be obtained. The most important limiting factor in satellite methods is their inability to resolve the finer details of the gravity field. The limiting resolution of NASA's Geopotential Research Mission, for example, will be 100 km *(Murphy, 1983)*, although there is promise of both finer resolution and greater precision with a satellite gravity gradiometer mission *(Rapp, 1986)*.

The success of satellite methods for global solutions, and their inabilities at short wavelengths, leaves local gravity field determination as a natural, distinct field of enquiry. Here too, there is a possibility for rapid, precise and consistent measurement - through airborne gravity gradiometry *(Jekeli, 1983)*. This possibilty apart,

it seems to be most appropriate for local gravity field determination to consider methods that can accommodate a wide range of point gravity measurements: geopotential differences, deflections of the vertical, gravity anomalies, gravity gradients - as well as other sources of gravity information: topographic elevations, rock densities and known geophysical constraints; but which could also accept as data other than point observations, e.g. mean gravity anomalies, or data profiles in the form of a continuous function. These data would be used purely for the purpose of obtaining local, detailed representations of the gravity field, i.e. of the geopotential or any of its functionals (in practice, the disturbing potential, its gradient vector, or gravity anomalies). Datum considerations, with respect to the zero and first degree spherical harmonics of the gravity field, and with respect to supplementary geodetic information, would play a reduced role.

The achievements of satellite methods for global gravity field determination thus leaves local gravity field determination as a pursuit worthy in itself (and conversely, it is perhaps no longer appropriate to use local terrestrial measurements to try to obtain global gravity information). In the next section the representation methods appropriate to this pursuit will be introduced.

1.2 The Local Gravity Field

By 'gravity field' will in the main be meant the disturbing potential - obtained by removing from the external gravitational potential of the earth a reference field made up of low order solid spherical harmonics, and the rotational potential - or its gradient, the disturbance vector,

or the gravity anomaly.

By 'representation' of the gravity field, some authors mean any numerical procedure capable of computing an approximation to the disturbing potential, or an approximation to some functional of the disturbing potential. We shall be a bit more restrictive, and call a 'representation' a finite linear combination of independent functions, approximating the gravity field, chosen either to interpolate, or smooth, or approximate in some sense, the given data, and satisfying some optimality criterion. The choice of suitable interpolating or smoothing functions is to be a central concern in the following chapters. We will not restrict ourselves exclusively to harmonic functions, and where we do work in spaces of harmonic functions, these will (on the whole) be functions harmonic in the exterior space of a sphere contained wholly within the earth - the Bjerhammer sphere. When considering functions harmonic outside a sphere, we appeal to the Runge-Krarup theorem: any harmonic function regular outside the earth's surface may be uniformly approximated by harmonic functions regular outside any sphere inside the earth *(Moritz, 1980, p 69)*.

The meaning of 'local' and 'global' representations of the gravity field perhaps also needs to be made clearer. *Schwarz (1984)* divides the gravity field spectrum into four frequency ranges:

1	$2 \leq n \leq 36$	low
2	$36 < n \leq 360$	medium
3	$360 < n \leq 3600$	high
4	$3600 < n$	very high,

where n is spherical harmonic degree.

Global solutions seek to determine the long wavelength, low and medium frequency components of the gravity field. Satellite solutions are most important in the low range, and will become increasingly important for the medium frequencies as well.

Local representations seek to resolve the remaining frequencies, with the 'very high' part of the spectrum becoming increasingly important in future. The word 'local' can have two senses. Firstly, it could indicate a solution for an area of limited extent (and 'regional' indicating a larger area). Alternatively, the word could be used to mean that little consideration is given to the datum problem, i.e. to constraining correctly the zero and first degree harmonics - to determining correctly the earth's centre, its rotation axis, and its mass - though Tscherning stresses on several occasions (e.g. *Tscherning, 1985)* the need not to neglect these datum constraints entirely (not least to ensure that data from different sources are compatible, e.g. terrestrial gravity anomalies, against anomalies derived from satellite altimetry). We shall use 'local' in both senses.

When considering possible methods for obtaining gravity field representations, *Tscherning's (1981)* review provides a useful starting point. The methods that Tscherning has grouped under the 'model approach' - e.g. integral formulae representing solutions of spherical boundary value problems - are not of direct concern, except in so far as the numerical integration implies some representation of the gravity field functional being integrated. Of interest are the methods grouped as being part of the 'operational approach', where the underlying philosophy is to attempt to make optimal use of all available observations, with minimal *a priori*

Lelgemann (1979, 1980), Moritz (1980, pp 169 et seq.), and *Freeden (1981a)* describe this analytic collocation.

Of special significance is the choice of the disturbing potential covariance function as reproducing kernel function, for then one can appeal to the theory of stochastic processes to show the collocation solution is an optimal estimate, one that minimises the error variance between the disturbing potential and its representation *(Kimeldorf and Wahba, 1970; Lauritzen, 1973).* The theoretical problem of the estimability of the covariance function will not concern us; it has not detracted from the practical significance and efficacy of least squares collocation. To the contrary, least squares collocation has become an almost universal geodetic tool. Entry points to the vast literature on the topic are *Moritz (1980),* and *Tscherning (1984).*

The class of approximation methods minimising norms in reproducing kernel Hilbert spaces can be extended considerably by including methods that minimise semi-norms in reproducing kernel semi-Hilbert spaces. A set of 'natural' constraints can be adjoined to fill out the nullspace of the semi-norm. In this way we obtain representations by natural spline functions - this is the real subject of the present work. Collocation is a special instance of a spline representation, as is made clear by *Freeden (1981a),* and by *Sünkel (1984a).* Freeden calls the collocation solution a 'representation by collocating spline functions'.

Interpolating spline functions in semi-Hilbert spaces have several attractive features, e.g. minimum semi-norm among all interpolating non-splines, and best approximation among all non-interpolating

splines. The next chapter will bring these properties clearly to light. As with collocation, there is a ready extension to representations that smooth the data, to least squares approximation to the data, and to the approximation of linear functionals. Not the least attraction is the possiblity of characterising explicitly the semi-norm being minimised as the norm in the range space of a linear operator measuring curvature or gradient in some general sense. It thus possible to supplement the dictum 'the smoother the function, the better' *(Tscherning, 1985)* by an exact statement of the sense in which the spline representation is smoothest.

1.3 Guide to Contents

One-dimensional splines have a well-developed theory, and not surprisingly, have found wide application; see, for example, *Böhmer (1974)*, *de Boor (1978)*, or *Schumaker (1981)*. One natural way of extending these to more dimensions is by taking tensor products. The bicubic spline is an example of a popular multivariate spline arising in this way (e.g. *Hayes and Halliday, 1974*); *Dierckx (1984)* describes spherical tensor-product splines. A disadvantage of these functions is that they are not rotation-invariant - they depend on the choice of co-ordinate system.

In looking for multivariate splines that minimise rotation-invariant norms, we face the question of what general properties a class of functions is to have to be regarded as spline functions. In the next chapter we consider the abstract properties intrinsic to the notion of 'spline function'. For this purpose we use *Atteia's (1965a, 1965b)* 'generalised splines', also Sard's theory of optimal approximation *(Sard, 1967)*, and the work of *Delvos and Schempp (1970, 1975, 1976)* on abstract spline

systems. Although it makes sense to consider splines within a wider setting, we only consider spline functions within the context of Hilbert spaces.

The theory of Chapter 2 is intrinsic; it tells us what constitutes a spline function, and the circumstances in which spline functions can be found; it gives no prescription for finding a solution, although there are clear indications that use should be made of orthogonal projections. Chapter 3 can be said to deal with the systematic construction of projectors; alternatively, it deals with the explicit construction of the Riesz map between a Hilbert space and its dual. For this purpose we use reproducing kernel functions in Hilbert spaces, and semi-Hilbert spaces, through the reproducing kernel theory of *Schwartz (1964)*, and its application to spline functions by *Schempp and Tippenhauer (1974)*.

The title of Chapter 3 is 'Constructive spline theory'. In a sense this is the subject of the remaining chapters as well. Chapter 4 deals with natural splines in $\mathbb{R}^n$, among which are rotation-invariant generalisations of the familiar polynomial splines of a single variable. The Fourier transform of spaces of tempered distributions will be a basic tool for constructing the kernel functions. Here we rely heavily on *Duchon (1977)*, although *Meinguet (1979a, 1979b, 1979c, 1979d, 1983)* also provides important insights. We see that the multiquadric functions introduced by *Hardy (1971)*, when subjected to a slight modification, are examples of multivariate polynomial splines. In fact, they provide a multivariate generalisation of linear interpolation - so their success for many surface interpolation problems is hardly surprising. We also look at thin-plate splines, and pseudo-cubic splines, and give a spline

interpretation to modelling with point masses.

The fifth chapter starts with a brief treatment of natural spline interpolation and approximation on the unit sphere. We take our cue from *Sünkel (1984)*, but instead of using Green's functions as construction tool, continue to use Fourier transforms to exhibit reproducing kernels associated with plausible smoothing operators. The chapter continues with spline functions on the unit sphere, and splines harmonic outside the sphere; here one must take cognisance of *Freeden's (1981a, 1982)* contribution to the topic.

Chapter 6 is given over to a number of numerical applications. In the first place we compare geoid interpolation using the spline functions from Chapter 4, and one spherical spline from Chapter 5, and least squares collocation, to show that with good data the different techniques give very similar results. In a second example the efficacy of auxiliary gravity anomaly data in improving the interpolation of geoid heights is assessed using multiquadric functions, spherical splines, and least squares collocation.

The third application points to the value of two-dimensional pseudo-cubic splines in representing the astro-geodetic geoid, taking as example a geoid for southern Africa interpolating both Doppler-derived geoid heights, and astro-geodetic deflections of the vertical.

Finally, we derive quadrature formulae, based on a thin-plate spline representation, for the numerical evaluation of geodetic integrals, when working to a planar approximation, and concerning ourselves only with

the innermost zone about the computation point. Example are given for Stokes's formula, for Vening-Meinesz's formula, for the analytic continuation integral that appears in the recursive formulation of Molodensky's solution of the geodetic boundary value problem, and for the formula for gravimetric terrain corrections.

The final chapter contains a brief summary, and attempts to draw a few conclusions. What I hope will have been achieved is some indication of the considerable unity underlying diverse methods of gravity field approximation, and that the language of spline approximation is adequate to describe interpolation and smoothing by collocation (in its various forms), by point masses and multipoles, by harmonic splines, by thin-plate splines, and by multiquadric functions. To this end, the presentation will be largely theoretical, dealing more with underlying structures than with the difficulties in solving real problems. Where we do touch on actual applications, the approach is somewhat direct, and the geodetic modelling possibly somewhat simplistic.

Many major areas of inquiry have been almost entirely neglected, such as the questions of error estimates or error bounds, of the degree of approximation, of convergence, of numerical stability, etc. On the question of convergence we draw some comfort that there are both some very general results, as given by *Delvos et al (1976)*, for example, on the convergence of abstract splines, and several very particular results, as for example, *Freeden (1984)* on the convergence of spherical splines. There is certainly much more to be said on the subject of spline representations of the earth's gravity field than I have attempted to tackle in this study.

CHAPTER 2

INTRINSIC SPLINE THEORY

2.1 Introduction

This chapter outlines an abstract theory of interpolation and smoothing by generalised spline functions. 'Spline interpolation is a process of orthogonal projection' *(de Boor and Lynch, 1966)*, thus stress will be laid on the properties of splines obtained by projection.

2.2 Interpolating and Smoothing Splines

We formulate the problems of spline interpolation and smoothing as the solution of extremal problems.

Let X, Y, Z be real separable Hilbert spaces, and let

$$U \in L(X,Y)$$

$$F \in L(X,Z)$$

be epimorphisms, i.e. U and F are continuous linear mappings that are open and onto.

$$\begin{array}{ccc} X & \xrightarrow{U} & Y = \mathrm{Im}U \cong X/\mathrm{Ker}U \\ & \searrow^{F} & \\ & & Z = \mathrm{Im}F \cong X/\mathrm{Ker}F. \end{array}$$

Further, let

$$\mathrm{Ker}U + \mathrm{Ker}F \text{ be closed in } X, \tag{2.1}$$

$$\mathrm{Ker}U \cap \mathrm{Ker}F = \{0\}. \tag{2.2}$$

KerU and KerF are the nullspaces of U and F, respectively; likewise ImU

and ImF are the images of X under U and F. (2.2) implies that U and F are collectively injective, i.e. collectively one-to-one.

Definition: $s \in X$ is an *interpolating spline* for $z \in Z$ when

$$Fs = z \tag{2.3}$$

(interpolating property), and

$$\| Us \|_Y \leq \|Ux\|_Y, \quad x \in X, \; Fx = z \tag{2.4}$$

(minimum property).

Remark: $\|\cdot\|_Y$ is the norm on Y induced by the inner product $(\cdot,\cdot)_Y$; it could equally well be regarded as a semi-norm on X.

Definition: Let $\sigma > 0$. $\tau \in X$ is a *smoothing spline* for $z \in Z$ if

$$\| U\tau \|_Y^2 + \sigma \| F\tau - z \|_Z^2 \leq \|Ux\|_Y^2 + \sigma \| Fx - z\|_Z^2, \quad x \in X \tag{2.5}$$

(Anselone and Laurent, 1968). This problem could also be reformulated as follows *(Böhmer, 1974, p 90)*: the space $W = Y \times Z$ with inner product

$$(w_1,w_2)_W = (y_1,y_2)_Y + \sigma(z_1,z_2)_Z, \; w_1 = (y_1,z_1), \; w_2 = (y_2,z_2) \in W, \tag{2.6}$$

and corresponding norm

$$\|w\|_W^2 = \|y\|_Y^2 + \sigma \|z\|_Z^2, \quad w = (y,z) \in W \tag{2.7}$$

is a Hibert space; moreover, $S: X \to W$ defined by

$$Sx = (Ux,Fx) \tag{2.8}$$

is a continuous linear mapping. Let $p = (0,z) \in W$; then

$$\| Ux\|_Y^2 + \sigma \| Fx - z\|_Z^2 = \| Sx - p\|_W^2, \qquad (2.9)$$

and $\tau \in X$ is a smoothing spline for $z \in Z$ if

$$\| S\tau - p \|_W \leq \| Sx - p \|_W, \; x \in X. \qquad (2.10)$$

The situation (for spline interpolation) can be pictured as follows (cf. Figure 2.1): let $x_0 \in X$; the only knowledge we have of x_0 is its image $z = Fx_0$ in Z. (F is the 'observation mapping', and z the 'observation'). Any element in the linear variety $x_0 + \mathrm{Ker}F$ interpolates z, i.e. $Fx = z$ for $x \in x_0 + \mathrm{Ker}F$. The optimal interpolant - the interpolant we call the interpolating spline - has minimum norm in Y. We shall see in a moment that, in consequence of (2.1), such an interpolating element s does exist, that (2.2) ensures that s is unique, and that Us is the orthogonal projection on $(U\mathrm{ker}F)^+$ of the image of $x_0 + \mathrm{Ker}F$ in Y.

Moreover, the space X can be given a Hilbert space structure with induced norm equivalent to that of the original space, in which the spline element s is the orthogonal projection of the linear variety $x_0 + \mathrm{Ker}F$ on $(\mathrm{Ker}F)^+$; cf. Figure 2.2.

Figure 2.3 shows in somewhat simplified form how spline smoothing can be pictured.

The space X is called the *spline ground space*. U is a mapping chosen to indicate the variability of $x \in X$. For example, U could measure gradient, or curvature, or the forces to which x is subject. The norm in Y would measure total curvature, or total energy. Spline elements minimise

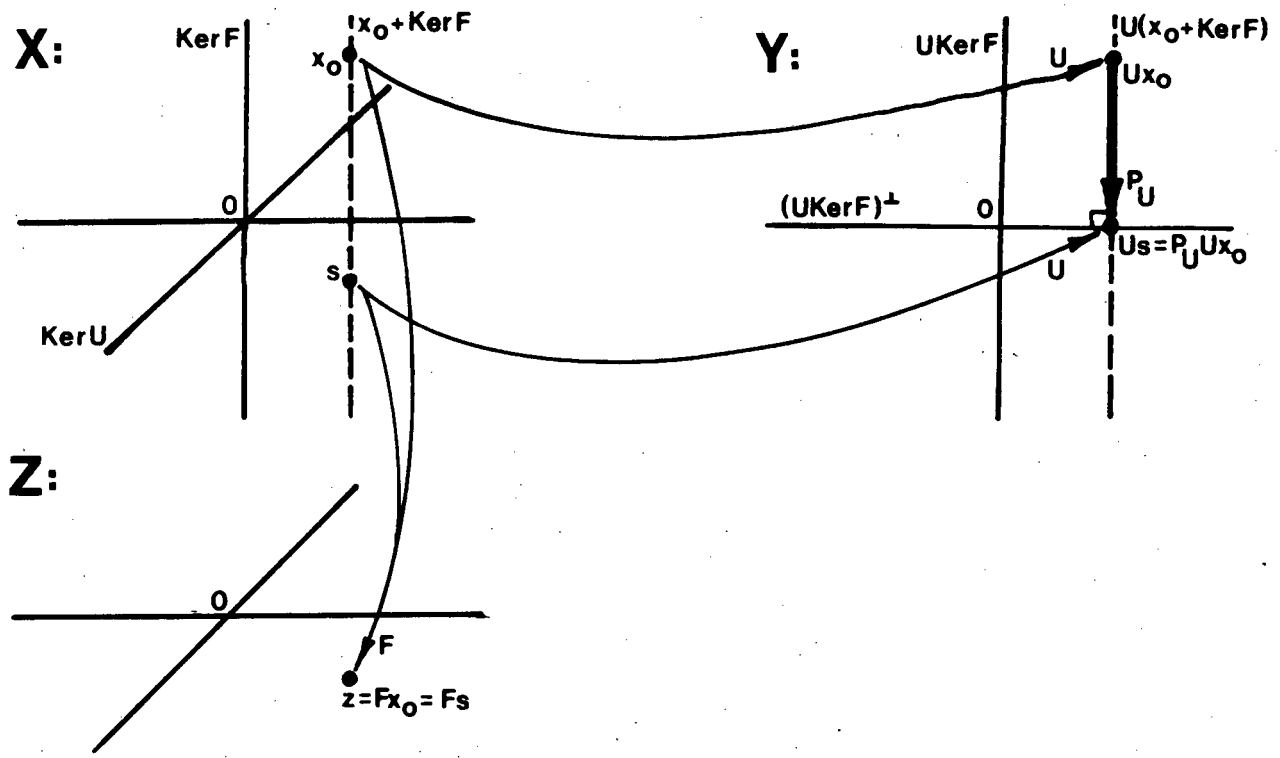


Figure 2.1: Spline interpolation as projection in the energy space Y.

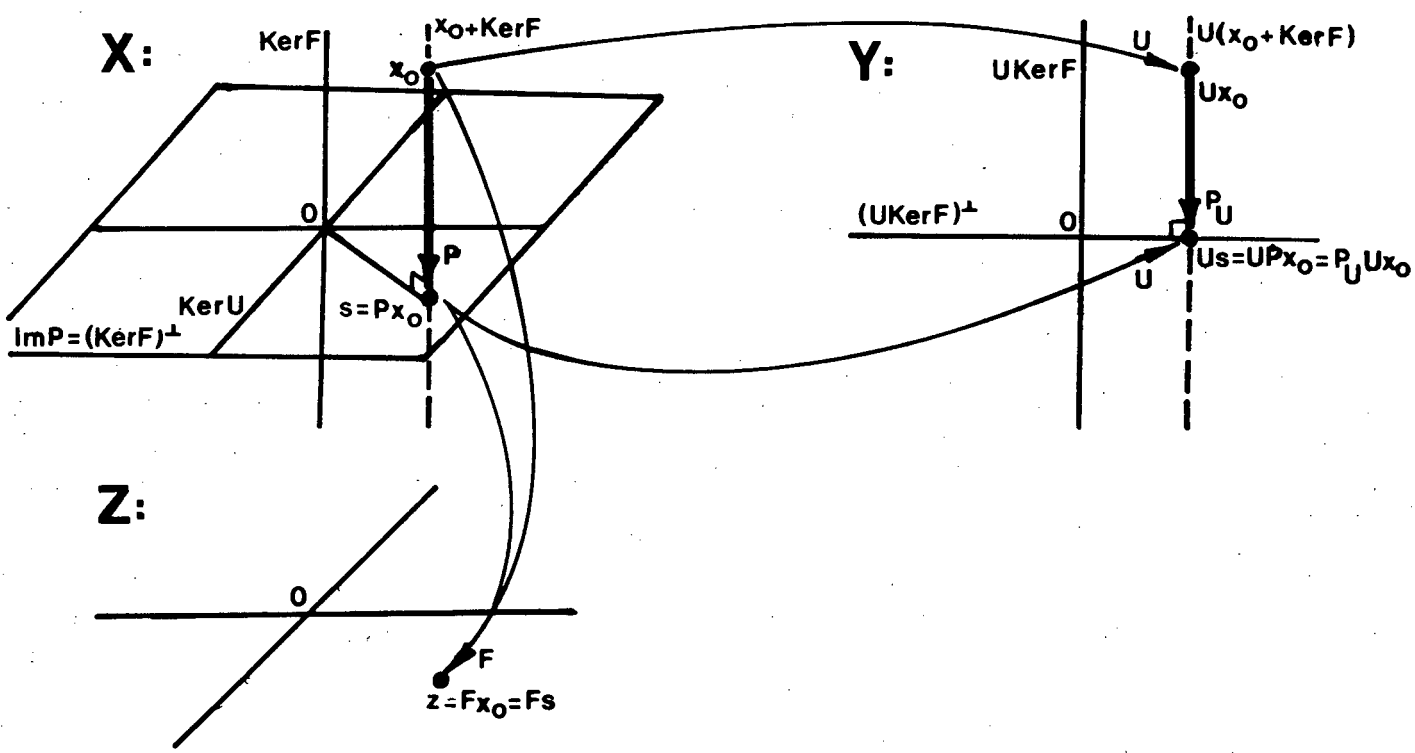


Figure 2.2: Spline interpolation as projection in the ground space X.

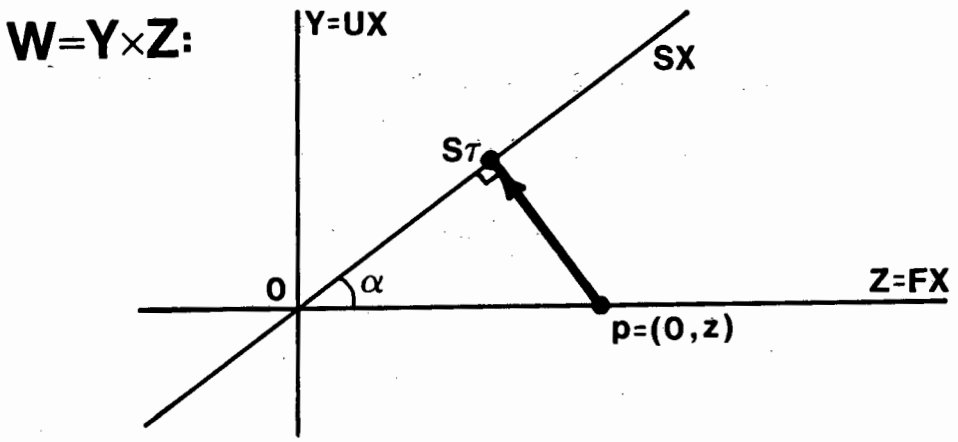


Figure 2.3: Spline smoothing as projection in Hilbert space W. The angle α depends on the choice of σ.

curvature subject to interpolating (or smoothing) 'boundary' conditions. The 'boundary constraints' are given by the observation mapping F. Often this will be a mapping onto $\mathbb{R}^n$ consisting of the cartesian product of n bounded linear functionals on X:

$$F = \ell_1 \times \ell_2 \times \dots \times \ell_n : X \to \mathbb{R}^n$$

$$F: x \mapsto (\langle x,\ell_1\rangle, \langle x,\ell_2\rangle, \dots, \langle x,\ell_n\rangle), \quad \ell_1,\ell_2,\dots,\ell_n \in X^*$$

($\langle\cdot,\cdot\rangle$ duality pairing between X and X^*). However, the theory is not restricted to finite observations (to observation operators with finite range); the operator F could, for example, take sections of a multivariate function.

2.3 Existence and Uniqueness

The assumption (2.1) that KerU + KerF is closed (in X) is equivalent to assuming UKerF closed in Y, and that the image of the linear variety of interpolants x_0 + KerF is also closed. By the projection theorem the

$$s_0 = s + x_0,\ x_0 \in \mathrm{Ker}U \cap \mathrm{Ker}F$$

also interpolates z, and moreover also satisfies (2.12), and is thus also a spline interpolating z.

Similarly, let τ be a spline smoothing z, and let

$$\tau_0 = \tau + x_0,\ x_0 \in \mathrm{Ker}U \cap \mathrm{Ker}F,$$

then $S\tau_0 = S\tau$, and τ_0 is also a smoothing spline.

$\mathrm{Ker}U \cap \mathrm{Ker}F = \{0\}$ is thus a necessary condition for the uniqueness of interpolating and smoothing splines.

We shall see in a moment that the condition $\mathrm{Ker}U \cap \mathrm{Ker}F = \{0\}$ means that X can be turned into a Hilbert space (with norm equivalent to the original norm on X), in which splines can be obtained by projection in X, rather than by projection in Y.

2.4 Minimum Properties

Let $x \in X$, s_x the spline that interpolates Fx, and τ_x the spline smoothFx. Let $p_x = (0, Fx)$ in W. It is a consequence of (2.12) and (2.14) that

$$\| Ux \|_Y^2 = \| Us_x \|_Y^2 + \| U(x - s_x) \|_Y^2, \tag{2.15}$$

$$\| Sx - p_x \|_W^2 = \| S\tau_x - p_x \|_W^2 + \| S(x - \tau_x) \|_W^2. \tag{2.16}$$

This is the *second minimum property*, also known as the *first integral*

relation. The *first minimum property* is also the consequence of projection in Y: let $x \in X$, and s_x its interpolating spline; then

$$\| U(x - s_x) \|_Y \leq \| U(x - s) \|_Y, \; Us \in (UKerF)^{\perp}. \tag{2.17}$$

Among all splines, i.e. among all $\{s \in X: Us \in (UKerF)^{\perp}\}$, the interpolating spline best approximates x.

Theorem: The space X with inner product

$$(x_1,x_2)_X = (Ux_1,Ux_2)_Y + (Fx_1,Fx_2)_Z, \; x_1,x_2 \in X \tag{2.18}$$

is a Hilbert space. The norm induced by $(\cdot,\cdot)_X$ is equivalent to the original norm $\|\cdot\|$ on X, i.e. there are constants $B_1,B_2 > 0$, such that

$$B_1\| x \| \leq \| x \|_X \leq B_2\| x \|, \; x \in X.$$

Proof: $(\cdot,\cdot)_X$ is clearly bilinear;

$$\| x \|_X^2 = (x,x)_X = \| Ux \|_Y^2 + \| Fx \|_Z^2 \geq 0,$$

and equals zero if (and only if) $x \in KerU \cap KerF$. It remains to show that X is complete with respect to the norm $\|\cdot\|_X$. Let $\{x_n\}_{n=1,2,\ldots}$ be a sequence in X, Cauchy with respect to $\|\cdot\|_X$; then $\{Ux_n\}_{n=1,2,\ldots}$ and $\{Fx_n\}_{n=1,2,\ldots}$ are Cauchy sequences in Y and Z, respectively. $\lim_{n\to\infty} Ux_n = y$, say, and $\lim_{n\to\infty} Fx_n = z$. Since both U and F are surjective, there exists $x_1,x_2 \in X$, such that

$$Ux_1 = y;\ Fx_2 = z.$$

In fact

$$U(x_1 + \mathrm{Ker}U) = y;\ F(x_2 + \mathrm{Ker}F) = z.$$

$x_1 + \mathrm{Ker}U$ is a translate of $\mathrm{Ker}U$; $x_2 + \mathrm{Ker}F$ is a translate of $\mathrm{Ker}F$; $(x_1 + \mathrm{Ker}U) \cap (x_2 + \mathrm{Ker}F)$ is a translate of $\mathrm{Ker}U \cap \mathrm{Ker}F = \{0\}$, and thus contains a single element $x \in X$. $\{x_n\}_{n=1,2,\ldots}$ converges (in the norm $\|\cdot\|_X$) to $x \in X$.

The two norms on X are equivalent: by the continuity of U and F,

$$\| x \|_X \leq \| Ux \|_Y + \| Fx \|_Z \leq B_2 \| x \|.$$

$S: X \to W$, $x \mapsto (Ux, Fx)$ is a continuous linear map. $\mathrm{Ker}S = \mathrm{Ker}U \cap \mathrm{Ker}F = \{0\}$. SX is a closed linear subspace of W. S is a homeomorphism between X and SX, and S^{-1}, when restricted to SX, is therefore also continuous. Put $\sigma = 1$; there is a constant $C > 0$ such that

$$\begin{aligned} \| x \|^2 &\leq C^2 \| Sx \|_W^2,\ x \in X \\ &= C^2 (\| Ux \|_Y^2 + \| Fx \|_Z^2) \\ &= C^2 \| x \|_X^2. \end{aligned}$$

Thus, with $B_1 = 1/C$,

$$B_1 \| x \| \leq \| x \|_X,\ x \in X.$$

Corollary: U and F are continuous with respect to $\|\cdot\|_X$.

Corollary: There is a constant C such that

$$\| x \|^2 \leq C^2(\| Ux \|_Y^2 + \| Fx \|_Z^2), \; x \in X. \tag{2.19}$$

Sard (1967), and *Delvos and Schempp (1975)* assume (2.19) at the outset, rather than $\mathrm{Ker}U \cap \mathrm{Ker}F = \{0\}$, as we have done. Either assumption ensures that U and F are collectively injective (see also *Aubin, 1979, p 99)*.

Let P be the orthogonal projector in X with nullspace KerF and range $(\mathrm{Ker}F)^{\perp}$ - orthogonal with respect to the inner product (2.18). Let P_U be the orthogonal projector we have used already, with nullspace UKerF and range $(U\mathrm{Ker}F)^{\perp}$. Let $x_0 \in X$. As we have seen, the interpolating spline s_{x_0} is the unique element in $x_0 + \mathrm{Ker}F$ satisfying

$$(Us_{x_0}, U\tilde{x})_Y = 0, \; \tilde{x} \in \mathrm{Ker}F.$$

In fact

$$Us_{x_0} = P_U Ux_0.$$

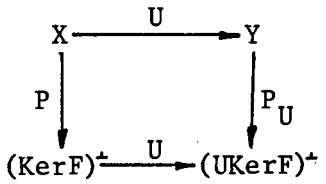
But Px_0 also interpolates Fx_0 (since $x_0 - Px_0 \in \mathrm{Ker}F$), moreover

$$0 = (Px_0, \tilde{x})_X = (UPx_0, U\tilde{x})_Y, \; \tilde{x} \in \mathrm{Ker}F.$$

Thus $Px_0 = s_{x_0}$, and $Us_{x_0} = UPx_0 = P_U Ux_0$. So

$$UPx = P_U Ux, \; x \in X. \tag{2.20}$$

The following commutative diagram thus holds:



One consequence is that the approximation property (2.17) can also be written as

$$\| U(x_0 - Px_0) \|_Y = \min_{x \in X} \| U(x_0 - Px) \|_Y. \tag{2.21}$$

In the abstract spline theory developed by *Delvos and Schempp (1970, 1972)*, the projector P plays a central role. In their terminology

$$(X, P, U, Y)$$

is a *spline system*. *Delvos and Schempp (1975, 1976)* extend their treatment to encompass Sard's theory of optimal interpolation *(Sard, 1967)*; they call

$$(X, Y, Z, U, F)$$

a *Sard system*. The orthogonal projector P is called the *spline projector*, and for $x_0 \in X$, Px_0 is called the *spline approximation* of x_0.

Remarks: 1. When $x_0 \in (KerF)^{\perp}$, the spline approximation is *exact:* $s_{x_0} = Px_0 = x_0$, and the representation error

$$\| U(x_0 - s_{x_0}) \|_Y = 0.$$

In particular, for $x_0 \in KerU$, the spline approximation is exact, since

$\mathrm{Ker}U \subset (\mathrm{Ker}F)^{\perp}$.

2. Smoothing splines also belong to $\mathrm{Im}P = (\mathrm{Ker}F)^{\perp}$: these splines minimise $\| Sx - p \|_W$, $x \in X$, but

$$\begin{aligned} \| SPx - p \|_W^2 &= \| UPx \|_Y^2 + \sigma \| FPx - p \|_Z^2 \\ &\leq \| Ux \|_Y^2 + \sigma \| Fx - p \|_Z^2 \\ &= \| Sx - p \|_W^2 \end{aligned}$$

(bearing in mind that $x - Px \in \mathrm{Ker}F$).

3. Connection with generalised operator inverses. Let X have inner product (2.18). Let $z \in Z$. We would like an optimal solution of

$$Fx = z \tag{2.22}$$

Since F is surjective, $F^{-1}\{z\} = \{x \in X: Fx = z\}$ is non-empty. Let $x \in F^{-1}\{z\}$, then

$$F^{-1}\{z\} = x + \mathrm{Ker}F.$$

We have seen that $x + \mathrm{Ker}F$ contains a unique element s_x of minimum norm:

$$s_x = Px,$$

where P is the orthogonal projector onto $(\mathrm{Ker}F)^{\perp}$. The map

$$F^{+}: Z \to X;\ z \mapsto PF^{-1}\{z\} \tag{2.23}$$

is the *right orthogonal inverse* of F, the type of inverse also sought in least squares estimation by the method of condition equations. In the least squares method of parameters, to the contrary, one looks for *left orthogonal inverses*. There is more than a formal correspondence between the familiar normal equations solution of the least squares problem, and (2.26) in the next section.

4. For the applications we have in mind

$$\dim(\mathrm{Ker}U) = q < \infty. \tag{2.24}$$

Should the observation mapping be a finite-rank operator - have finite-dimensional range -

$$\dim(Z) = \dim(\mathrm{Im}F) = n < \infty, \tag{2.25}$$

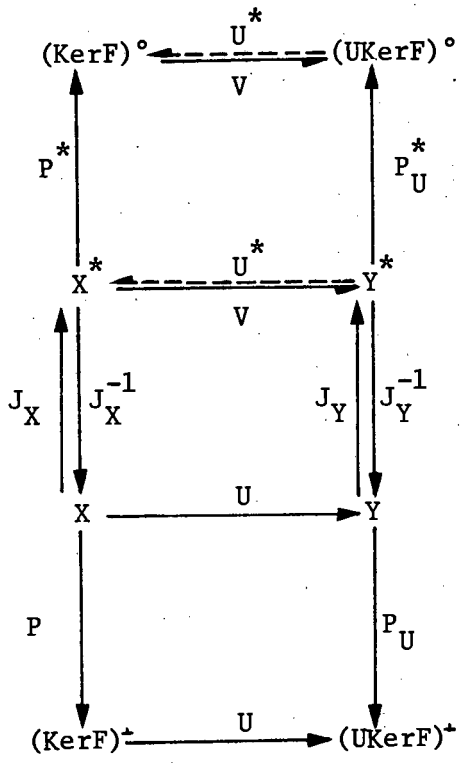
and since $(\mathrm{Ker}F)^{+} \cong \mathrm{Im}F$, $\dim(\mathrm{Ker}F)^{+} = n$. For a finite-rank observation mapping thus, the space $(\mathrm{Ker}F)^{+}$ from which spline approximations are drawn is finite-dimensional. This suggests that interpolating (and smoothing) splines be obtained as finite linear combinations of suitable base elements in $(\mathrm{Ker}F)^{+}$.

2.5 Approximation of Linear Functionals

Let X have the inner product (2.18). The spline approximation to $x_0 \in X$ is Px_0. The purpose of this section is to obtain a corresponding result in the dual space X^*: let $\ell_0 \in X^*$, then orthogonal projection onto $(\mathrm{Ker}F)^{\circ}$ provides the best approximation to ℓ_0, exact on $(\mathrm{Ker}F)^{+}$.

((KerF)° is the orthogonal of KerF in X^* - it consists of continuous linear functionals on X annihilating KerF).

Let X^*, Y^* be the duals of X, Y; let J_X, J_Y be the canonical duality isomorphisms between X, Y and X^*, Y^*, respectively; and let U^*, P^*, P_U^* be the transposed mappings to U, P, P_U. We show that the spline projection structure of the last section can be transposed to the dual spaces, i.e. that the mappings shown by solid lines in the diagram below are commutative.



U^* is injective (being the transpose of an epimorphism); its image $ImU^* = (KerU)°$ is closed in X^* (closed graph theorem). U^* thus has a *left orthogonal inverse:*

$$(U^*)^- = (UJ_X^{-1}U^*)^{-1}UJ_X^{-1} \tag{2.26}$$

(Aubin, 1979, p 89). Define

$$V = \| U \|^2 (U^*)^-;$$

then

$$V = J_Y U J_X^{-1}, \tag{2.27}$$

since $(UJ_X^{-1}U^*)^{-1}$, like J_Y, is an isomorphism between Y and Y^*, differing only in norm by $1/\|U\|^2$. Thus the images of U,V in Y,Y^* are isomorphic under J_Y.

P and P_U also have closed images ($(KerF)^\perp$ and $(UKerF)^\perp$), thus

$$Im P^* = (KerP)^\circ = (KerF)^\circ,$$
$$Im P_U^* = (KerP_U)^\circ = (UKerF)^\circ.$$

In fact P^*, P_U^* are orthogonal projections onto $(KerF)^\circ$ and $(UKerF)^\circ$, respectively: from

$$P^* = J_X P J_X^{-1};\ P_U^* = J_Y P_U J_Y^{-1} \tag{2.28}$$

(P,P_U are self-adjoint), and the fact that J_X, J_Y preserve inner products, it is easily seen that P^*, P_U^* are idempotent, and symmetric with respect to the inner products on X^*, Y^* transported by J_X, J_Y. In consequence

$$\begin{aligned} VP^* &= J_Y U J_X^{-1} J_X P J_X^{-1} \\ &= J_Y P_U U J_X^{-1} \\ &= J_Y J_Y^{-1} P_U^* J_Y J_Y^{-1} V J_X J_X^{-1} \\ &= P_U^* V. \end{aligned}$$

We thus have:

Theorem: Let $\ell_0 \in X^*$. There is a unique $\ell \in X^*$ satisfying the exactness condition

$$<x,\ell> = <x,\ell_0>, \; x \in (\mathrm{Ker}F)^{\perp}, \tag{2.29}$$

which minimises the functional $\| V(\cdot) \|_Y^*$.

Proof: The set $\ell_0 + (\mathrm{Ker}F)^{\perp 0}$ satisfies the exactness condition. By the correspondences that have been set up it can be seen that $P^* \ell_0$ satisfies the conditions of the theorem.

Remark: Frequently a simpler problem is posed. Instead of the initial assumptions (2.1), (2.2), it is assumed at the outset that

$$\mathrm{Ker}U \text{ is closed; } \mathrm{Ker}U = (\mathrm{Ker}F)^{\perp}.$$

The approximation found by applying the theorem will then be exact for $x \in \mathrm{Ker}U$, i.e. *exact in the sense of Sard.*

CHAPTER 3

CONSTRUCTIVE SPLINE THEORY

3.1 Introduction

The last chapter outlined the requirements for the existence and uniqueness of spline approximations in Hilbert space, and the attractive minimum properties obtained by orthogonal projection, but there was no hint as to how spline elements may be got in practice. In the present chapter we take a first step in that direction, with the aim of using kernel mappings as a systematic constructive tool. In Section 3.2 we introduce reproducing kernels in the context of L. Schwartz'z theory of kernel mappings *(Schwartz, 1964, 1973; Schempp and Tippenhauer, 1974)*. In the following section the spline interpolation problem is solved using kernels associated with semi-Hilbert sub-spaces, with a section looking briefly at the approximation of linear functionals. Section 3.5 is devoted to the reproducing kernel functions so widely used in geodetic approximation, and the final section gives brief consideration to Green's functions as an alternative constructive tool.

In several places reference will be made to topological notions, without the concepts being defined or elaborated; *Jarchow (1981)* is a general reference to the topological content of this chapter, and *Aubin (1979)* an elegant guide to much of the functional analytic background.

3.2 Hilbert Sub-Spaces and their Associated Kernels

Let E be a locally convex topological vector space over $\mathbb{R}$ (which we will assume at the outset to be at least quasi-complete); $\mathbb{H}$ is a

Hilbert sub-space of E if

(i) $\mathbb{H}$ is a linear sub-space of the vector space E;

(ii) $\mathbb{H}$ is equipped with a Hilbert space structure; and

(iii) the canonical injection of $\mathbb{H}$ into E (the inclusion map) is continuous.

The set of all Hilbert sub-spaces of E is denoted Hilb(E).

We can define on Hilb(E) a structure of multiplication by non-negative real scalars: let $\mathbb{H} \in$ Hilb(E); if $\lambda = 0$, put $\lambda\mathbb{H} = \{0\}$, otherwise if $\lambda > 0$, define the inner product on $\lambda\mathbb{H}$ as

$$(h,k)_{\lambda\mathbb{H}} = \frac{1}{\lambda}(h,k)_{\mathbb{H}}, \quad h,k \in \mathbb{H}, \tag{3.1}$$

then

$$\| h \|_{\lambda\mathbb{H}} = \frac{1}{\lambda^{1/2}} \| h \|_{\mathbb{H}}, \quad h \in \mathbb{H}. \tag{3.2}$$

Hilb(E) can also be made closed under addition: if $\mathbb{H}_1, \mathbb{H}_2 \in$ Hilb(E), their linear sum $\mathbb{H}_1 + \mathbb{H}_2$ is certainly a linear sub-space of E, and it is easily verified that

$$\| h \|_{\mathbb{H}_1+\mathbb{H}_2} = \inf_{\substack{h=h_1+h_2 \\ h_1\in\mathbb{H}_1, h_2\in\mathbb{H}_2}} (\|h_1\|^2_{\mathbb{H}_1} + \|h_2\|^2_{\mathbb{H}_2})^{1/2} \tag{3.3}$$

is a norm on $\mathbb{H}_1 + \mathbb{H}_2$. *Schwartz (1973, 331)* shows this norm arises from a Hilbert structure on $\mathbb{H}_1 + \mathbb{H}_2$.

Lastly, it is possible to define an order relation on Hilb(E): let $\mathbb{H}_1, \mathbb{H}_2 \in$ Hilb(E); we say $\mathbb{H}_1 < \mathbb{H}_2$ if

(i) $\mathbb{H}_1 \subset \mathbb{H}_2$; and

(ii) the inclusion map $\mathbb{H}_1 \to \mathbb{H}_2$ has norm at most 1, i.e.

$$\| h \|_{\mathbb{H}_2} \leq \| h \|_{\mathbb{H}_1}, \quad h \in \mathbb{H}_1.$$

If $\lambda\mathbb{H} \leq \mathbb{H}$, then

$$\| h \|_{\mathbb{H}} \leq \| h \|_{\lambda\mathbb{H}} = \frac{1}{\lambda^{1/2}} \| h \|_{\mathbb{H}},$$

i.e. $\lambda \leq 1$. Also, $\mathbb{H}_1 \leq \mathbb{H}_1 + \mathbb{H}_2$, for if $h_1 \in \mathbb{H}_1$,

$$\| h_1 \|_{\mathbb{H}_1 + \mathbb{H}_2} = \inf_{\substack{k_1 + k_2 = h_1 \\ k_1 \in \mathbb{H}_1, k_2 \in \mathbb{H}_2}} (\| k_1 \|^2_{\mathbb{H}_1} + \| k_2 \|^2_{\mathbb{H}_2})^{1/2} \leq \| h_1 \|_{\mathbb{H}_1},$$

since $h_1 + 0$ is one of the combinations $k_1 + k_2$ competing for the infinum.

The scalar multiplicative and additive structure given to Hilb(E) turns it into a *convex cone*, and more particularly, into a *salient* (sharp) and *regular* convex cone, for Hilb(E) satisfies the following axioms of a convex cone: let $\mathbb{H}_1, \mathbb{H}_2, \mathbb{H}_3 \in$ Hilb(E), and $\lambda_1, \lambda_2 > 0$, then

(i) addition is associative: $\mathbb{H}_1 + (\mathbb{H}_2 + \mathbb{H}_3) = (\mathbb{H}_1 + \mathbb{H}_2) + \mathbb{H}_3$;

(ii) addition is commutative: $\mathbb{H}_1 + \mathbb{H}_2 = \mathbb{H}_2 + \mathbb{H}_1$;

(iii) Hilb(E) contains a zero element: $0 + \mathbb{H}_1 = \mathbb{H}_1$;

(iv) $0\mathbb{H}_1 = \lambda_1 0 = 0$;

(v) $1\mathbb{H}_1 = \mathbb{H}_1$;

(vi) $(\lambda_1\lambda_2)\mathbb{H}_1 = \lambda_1(\lambda_2\mathbb{H}_1)$;

(vii) $(\lambda_1 + \lambda_2)\mathbb{H}_1 = \lambda_1\mathbb{H}_1 + \lambda_2\mathbb{H}_1$;

(viii) $\lambda_1(\mathbb{H}_1+\mathbb{H}_2) = \lambda_1\mathbb{H}_1+\lambda_1\mathbb{H}_2$.

Hilb(E) is salient since it can be shown that $\mathbb{H}_1+\mathbb{H}_2 = 0$ implies $\mathbb{H}_1 = \mathbb{H}_2 = 0 = \{0\}$, i.e. the only sub-space of Hilb(E) is the trivial one containing only the zero element, and Hilb(E) is regular since if $\mathbb{H}_1+\mathbb{H}_2 = \mathbb{H}_1+\mathbb{H}_3$, then $\mathbb{H}_2 = \mathbb{H}_3$ *(Schwartz, 1973, p 338)*.

Let $\mathbb{H} \in$ Hilb(E), and $\mathbb{H}'$ its dual; the Riesz mapping θ: $\mathbb{H}' \to \mathbb{H}$ defined by

$$(h,\theta k') = \langle h,k' \rangle, \quad h \in \mathbb{H}, \quad k' \in \mathbb{H}', \tag{3.4}$$

is an isomorphism between $\mathbb{H}'$ and $\mathbb{H}$ (θ is the inverse of the duality isomorphism introduced on p 25). We will now extend this to a map from E' (the dual of E) to E: using some of the notation from section 2.5 for a new purpose, let J: $\mathbb{H} \to$ E be the inclusion map of $\mathbb{H}$ into E, and J^*: E' $\to \mathbb{H}'$ its transpose; J^* is also a continuous linear map. Define H: E' $\to$ E by

$$H = J\theta J^*: E' \to \mathbb{H}' \to \mathbb{H} \to E. \tag{3.5}$$

H is called the *kernel associated with* $\mathbb{H}$. It is a positive map in the sense that, for $e' \in E'$,

$$\begin{aligned} \langle He',e' \rangle &= \langle J\theta J^* e',e' \rangle_{(E,E')} \\ &= \langle \theta J^* e', J^* e' \rangle_{(\mathbb{H},\mathbb{H}')} \\ &= (\theta J^* e', \theta J^* e')_{\mathbb{H}} \quad \geq 0. \end{aligned}$$

Moreover, H is the only map from E' to $\mathbb{H}$ such that

$$(h,He')_{\mathbb{H}} = \langle h,e' \rangle,\ h \in \mathbb{H},\ e' \in E'. \tag{3.6}$$

H is an element of $\mathbb{L}^+(E',E)$, the space of positive continuous linear mappings from E' to E - H is called a *positive kernel relative to E.* The correspondence (3.5) between $\mathbb{H}$ and H establishes a bijection between Hilb(E) and $\mathbb{L}^+(E',E)$ *(Schwartz, 1973, p 337)*, one moreover, that preserves the convex cone and order structure on both these sets:

(i) If $\mathbb{H} \in$ Hilb(E) and H is its associated kernel, then for $\lambda \geq 0$, the kernel associated with $\lambda\mathbb{H}$ is λH:

$$(h,\lambda He')_{\lambda\mathbb{H}} = \lambda(h,He')_{\lambda\mathbb{H}} = (h,He')_{\mathbb{H}} = \langle h,e' \rangle,\ h \in \mathbb{H},\ e' \in E'.$$

(ii) If $\mathbb{H}_1, \mathbb{H}_2 \in$ Hilb(E), and H_1, H_2 are their associated kernels, then $H_1 + H_2$ is the kernel associated with $\mathbb{H}_1 + \mathbb{H}_2$ *(Schwartz, 1964, p 160).*

(iii) If $\mathbb{H}_1, \mathbb{H}_2 \in$ Hilb(E), and H_1, H_2 are their associated kernels, then $\mathbb{H}_1 \leq \mathbb{H}_2$ is equivalent to $H_1 < H_2$ *(ibid.).*

Thus the map $\mathbb{H} \mapsto H$ implied by (3.5) is an isomorphism for the convex cone structure on Hilb(E) and $\mathbb{L}^+(E',E)$, and also an isomorphism for the order structure on these sets.

The kernel H: $E' \to E$ induces a bilinear form on $E' \times E'$ through

$$B(e',f') = \langle He',f' \rangle,$$

and in turn, a positive quadratic form on E' given by

$$Qe' = B(e',e') = \langle He',e' \rangle = \| He' \|^2_{\mathbb{H}}.$$

Conversely, it is also true that if Q is a positive quadratic form on E' such that the bilinear form on E' x E'

$$B(e',f') = \{\frac{1}{2} Q(e'+f') - Qe' - Qf'\}$$

is separately continuous, then there is an $\mathbb{H} \in$ Hilb(E) such that Q is induced by the kernel associated with $\mathbb{H}$ *(Schwartz, 1973, p 339)*.

The notions of Hilbert sub-spaces and their associated kernels can readily be extended to semi-Hilbert sub-spaces. As before, let E be a locally convex topological vector space, and E' its dual; let N be a finite-dimensional sub-space of E, and N° its orthogonal in E' (see (3.8) below). A linear sub-space $\mathbb{H}$ of E is a *semi-Hilbert sub-space of E with nullspace N* if $\mathbb{H}$ is equipped with a semi-norm $\|\cdot\|_{\mathbb{H}}$ (with nullspace N) deriving from a semi-inner product $(\cdot,\cdot)_{\mathbb{H}}$ on $\mathbb{H}$ such that $\mathbb{H}/N$ with the natural norm $\|x+N\|_{\mathbb{H}/N} = \|x\|_{\mathbb{H}}$ is a Hilbert sub-space of E/N.

A unique positive kernel relative to E/N can be associated with $\mathbb{H}/N$, i.e. the linear mapping H: $(E/N)' \to E/N$ satisfying

$$(h,He')_{\mathbb{H}} = \langle h+N, e' \rangle, \quad h \in \mathbb{H}, e' \in (E/N)'.$$

Since $(E/N)' \cong N^\circ$, we may instead consider the map

$$(h,\theta e')_{\mathbb{H}} = \langle h, e' \rangle, \quad h \in \mathbb{H}, e' \in N^\circ, \tag{3.7}$$

and call θ the *kernel associated with the semi-Hilbert space* $\mathbb{H}$.

A final remark that is required in the preparation for the next section, is that $\mathbb{H} \in$ Hilb(E) needs be dense in E if J^*, transpose of the inclusion map of $\mathbb{H}$ in E, is to be injective *(Köthe, 1979, p 84)*, and E' to be contained in $\mathbb{H}'$. (E' is then dense in $\mathbb{H}'$, and J^* restricts elements of E' to $\mathbb{H}'$).

3.3 Solution of the Spline Interpolation Problem

In this section kernels associated with semi-Hilbert sub-spaces are put to use to solve the spline interpolation problem.

Let (X, Y, Z, U, F) be a Sard system as defined in Section 2.4. The mapping U: $X \to Y$ produces a semi-inner product on X:

$$(x_1, x_2)_X = (Ux_1, Ux_2)_Y, \quad x_1, x_2 \in X.$$

The semi-norm induced by $(\cdot,\cdot)_X$ has nullspace $N = \mathrm{Ker}U$, which we assume to have finite dimension. The natural epimorphism associated with $\|\cdot\|_X$ maps X onto X/N which is isomorphic to $Y = \mathrm{Im}U$. It is thus not necessary to specify the 'co-observation' mapping U, as was done in Chapter 2, but specify only a semi-inner product on X. The properties of the interpolating (or smoothing) spline then derive from this semi-inner product.

Let E be a locally convex topological vector space, and X a semi-Hilbert sub-space with nullspace N. Let N° be the orthogonal of N in E, i.e.

$$N^\circ = \{e' \in E' : \langle x,e' \rangle = 0, \ \forall x \in N\}. \tag{3.8}$$

Although the theory of Chapter 2 is sufficiently general to accommodate 'observations' other than (real) scalar or vector quantities, we will henceforth consider only scalar observations, and then only a finite number of observations. The 'observation' mapping thus has the form

$$F = \ell_1 x \ \ell_2 x \ldots \ x \ell_m, \ \ell_i \in E', \ i = 1,\ldots,m.$$

Let M be the linear span of $\{\ell_1,\ell_2,\ldots,\ell_m\}$, then $\dim M \leq m$, and M is a closed linear sub-space of E'. Put

$$M^\circ = \{x \in E : \langle x,e' \rangle = 0, \ \forall e' \in M\}. \tag{3.9}$$

The range of F is $\mathbb{R}^m$; an inner product can be defined through

$$(w,v)_{\mathbb{R}^m} = \sum_{i=1}^{m} w_i v_i, \quad w,v \in \mathbb{R}^m.$$

Thus, for $x,y \in X$,

$$(Fx,Fy)_{\mathbb{R}^m} = \sum_{i=1}^{m} \langle x,\ell_i \rangle \langle y,\ell_i \rangle.$$

The condition on U and F for the existence of interpolating (and smoothing) splines, viz. KerU + KerF closed in X, now expresses itself as $N + M^\circ \cap X$ closed in X. We assume further, that X is dense in E; then $N^\circ \cong (E/N)' \subset (X/N)'$, and $N + M^\circ \cap X$ closed in X is equivalent to $N^\circ \cap M$ closed in $(X/N)'$, for which the finite-dimensionality of M is sufficient.

The requirement for a unique solution was $\mathrm{Ker}U \cap \mathrm{Ker}F = \{0\}$; this now becomes $N \cap M^\circ = \{0\}$, i.e. if $x \in N$ and $\langle x, e' \rangle = 0$ for all $e' \in M$, then $x = 0$.

It was shown in Chapter 2 how under these conditions a unique interpolating spline exists. This result is also obtained directly in the theorem below, and in the following theorem this spline is characterised using the kernel map θ defined in (3.7).

Theorem: *(Duchon, 1977)*. Let E, X, N, M be as above; let $x \in X$, then there is a unique $s \in X$ satisfying the interpolatory conditions

$$\langle s, e' \rangle = \langle x, e' \rangle, \quad e' \in M,$$

and having minimum semi-norm.

Proof: $x + M^\circ \cap X$ is the linear variety of interpolants to x on M. $x + M^\circ \cap X + N$ is a closed linear variety in the Hilbert space X/N. (X/N with norm $\| x+N \|_{X/N} = \| x \|_X$). It thus contains a unique element of minimum norm $s+N$. $s+N$ in turn, contains a unique element in $x + M^\circ \cap X$ (since $N \cap M^\circ = \{0\}$), which is s.

Figure 3.1 helps visualise the situation.

Theorem: *(Duchon, 1977)*. Let E, X, N, M be as above, and let θ be the kernel associated with X as a semi-Hilbert sub-space of E. Let $x \in X$; then the interpolating spline s coincides with the only $\xi \in \theta(M \cap N^\circ) + N$ satisfying

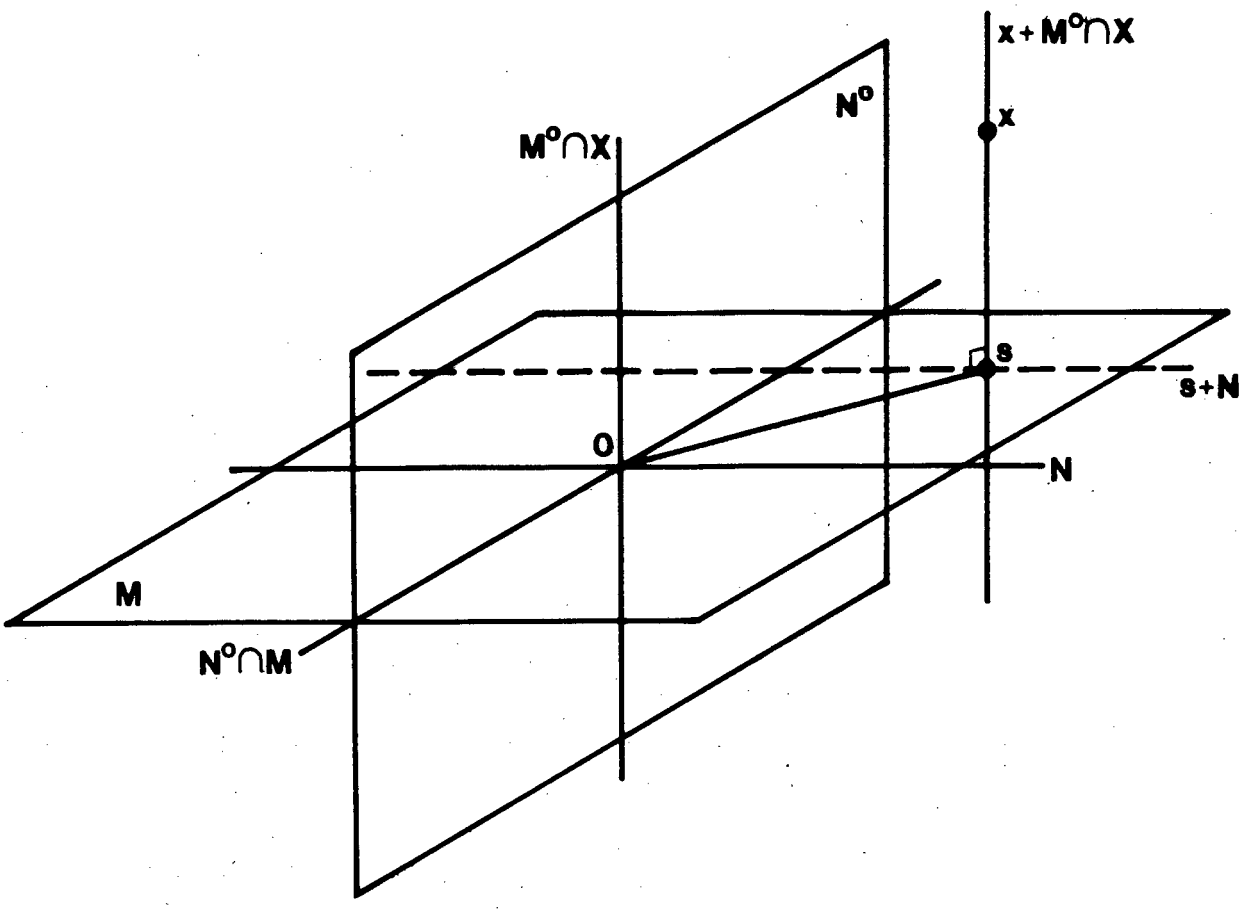


Figure 3.1: Spline interpolation in the semi-Hilbert ground space X. (For the purposes of illustration, X' has been identified with X).

$$\langle \xi, e' \rangle = \langle x, e' \rangle, \quad e' \in M.$$

Proof: The interpolants of x are $x+M^\circ \cap X$. $s+N$ as given by the previous theorem is the projection of 0 (in X/N) onto $x+M^\circ \cap X+N$, and is therefore the only element of $x+M^\circ \cap X+N$ belonging to the orthogonal of $M^\circ \cap X+N$ (orthogonal in X/N). The orthogonal of $M^\circ \cap X+N$ is the image of $(M^\circ \cap X+N)^\circ$ (orthogonal in $(X/N)'$) by the canonical isometry between $(X/N)'$ and X/N. But $M \cap N^\circ$ is closed in $(X/N)'$, therefore $(M^\circ \cap X+N)^\circ = M \cap N^\circ$. Its image in X/N is $\theta(M \cap N^\circ)+N$.

The interpolating spline therefore, is the unique interpolant of the

form $\theta(M \cap N)^\circ + N$; it is a linear combination of basis elements of the nullspace N, and basis elements of M under the kernel mapping θ, subject to the condition that these annihilate N. In a carry-over of the language of single variable splines, this last condition is referred to as the 'natural spline end condition'.

3.4 Approximation of Linear Functionals

In this section we take the problem of Section 2.5 a step further, to find a best approximation to a given linear functional, satisfying the condition of exactness for all spline elements.

Let (X, Y, Z, U, F) be a Sard system, X with inner-product (2.18) a dense Hilbert sub-space of the topological vector space E. As in the last section we will instead consider X as a semi-Hilbert sub-space of E, with nullspace $N = \mathrm{Ker}U$. $F = \ell_1 \times \ell_2 \times \ldots \times \ell_m$ as before, and M is the linear span of these ℓ_i. We suppose the conditions hold for the unique solution of the spline interpolation problem.

Theorem: Let $\ell_0 \in X'$; there is a unique $\ell \in M$ satisfying the exactness condition (2.29)

$$\langle x, \ell \rangle = \langle x, \ell_0 \rangle, \quad x \in (\mathrm{Ker}F)^{\perp},$$

with minimum semi-norm $\|\cdot\|_{X'}$, (with nullspace $(X/N)^\circ$).

Proof: The set $\ell_0 + (\mathrm{Ker}F)^{\perp\circ}$ satisfies the exactness condition (orthogonal here with respect to the innerproduct (2.18)). Spline elements

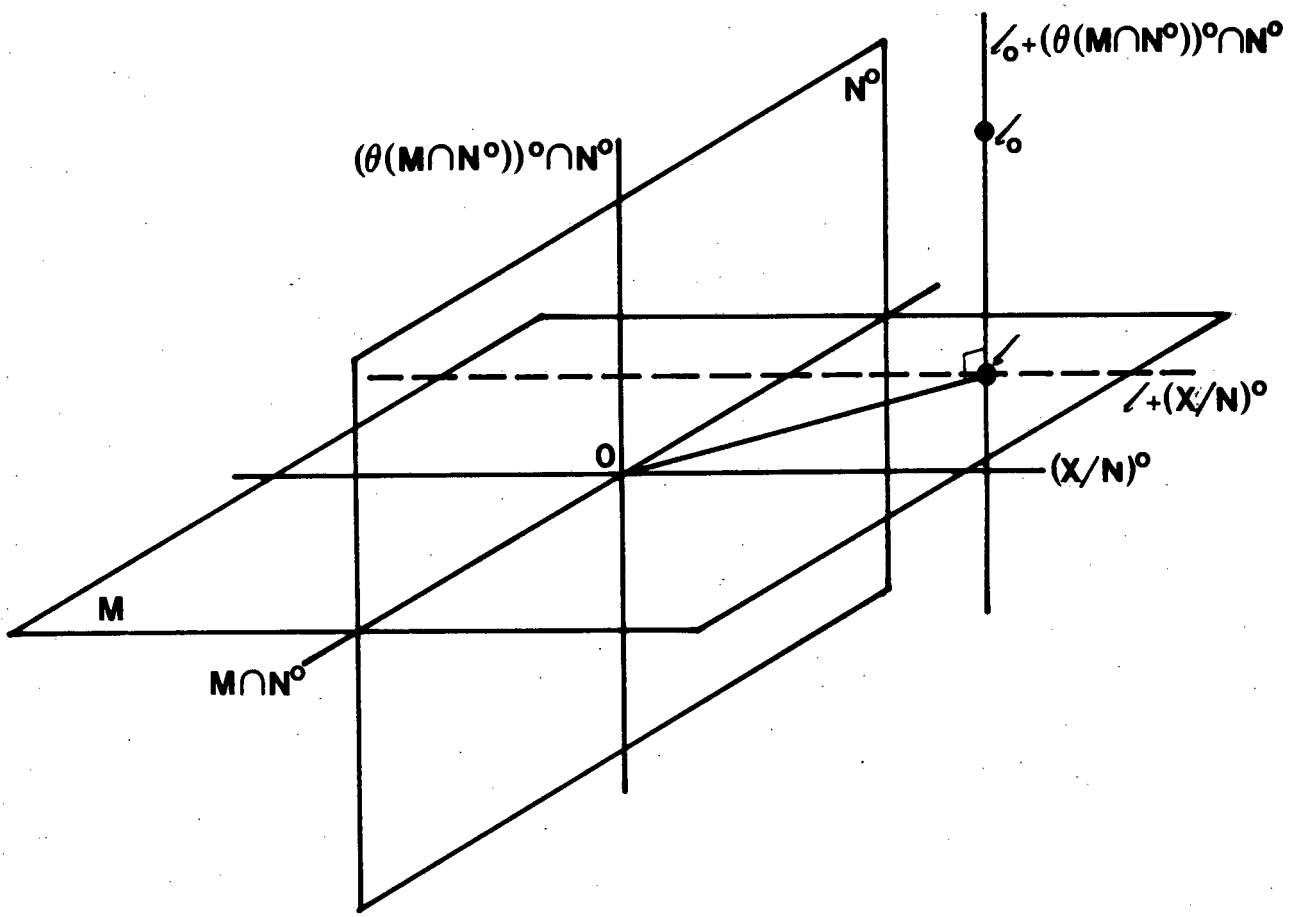


Figure 3.2: Best approximation of linear functionals in the dual space X'.

are elements of $(\mathrm{Ker}F)^+$; spline elements are equally elements of $\theta(M\cap N^\circ)+N$, thus the exactness is also satisfied by $\ell_0+(\theta(M\cap N^\circ)+N)^\circ$.

$$\ell_0+((\theta(M\cap N^\circ)+N)^\circ+(X/N)^\circ$$

is a closed linear variety in the Hilbert space $(X/N)' - (X/N)' = (\theta(N^\circ))^\circ$ - and thus contains a unique element of minimum quotient norm, $\ell+(X/N)^\circ$ say, which in turn contains a single element ℓ satisfying the exactness condition, since

$$(X/N)^\circ\cap N^\circ\cap(\theta(M\cap N^\circ))^\circ = \{0\}.$$

Figure 3.2 shows the geometry of best approximation in the dual space X'.

The best approximation to a given bounded linear functional ℓ_0 can thus be found as a linear combination of basis elements of M, which are exact on N, and exact on elements of $\theta(M)$ satisfying the natural spline end conditions. These requirements give rise to a system of linear equations identical in form to equations (6.38) of *Freeden (1981a), p 234* - as will become clear in Section 4.9.

3.5 Reproducing Kernel Functions

The last two sections give potent reason to continue the study of kernel mappings. This section points to correspondences between these mappings and reproducing kernel functions, as treated by *Meschkowski (1962)* and *Aronszajn (1950)*, and which have had a fruitful impact on geodetic estimation techniques.

It has been indicated that to every $\mathbb{H} \in$ Hilb(E) there is an associated kernel $H \in \mathbb{L}^+(E',E)$, and a positive bilinear form B(e',f') on $E' \times E'$. In the case that E is a space of functions on some domain Ω, we can go further, and associate with each $\mathbb{H} \in$ Hilb(E) a postive function on $\Omega \times \Omega$ called a *reproducing kernel function*. The treatment here follows *Schwartz (1964), §9*, a brief summary of which appears in §2 of *Schempp and Tippenhauer (1974)*.

Let Ω be a domain, a non-empty open set. Let $E = \mathbb{R}^{\Omega}$ be the space of all real-valued functions on Ω, and let E be given the topology of

pointwise convergence. The dual E' consists of measures with finite support on Ω, i.e. measures of the form

$$\mu = \sum_{x \in \Omega} c_x \delta_x, \tag{3.10}$$

where δ_x is the Dirac measure at the point x, and at most a finite number of the real coefficients c_x are non-zero. The duality between $f \in E = \mathbb{R}^{\Omega}$ and $\mu \in E'$ is expressed through

$$\langle f, \mu \rangle = \sum_{x \in \Omega} c_x f(x). \tag{3.11}$$

Let $H \in \mathbb{L}^{+}(E',E)$ be a positive kernel relative to E; it defines a *reproducing kernel* $K: \Omega \times \Omega \to \mathbb{R}$ via

$$K(x,y) = \langle H\delta_y, \delta_x \rangle, \quad x,y \in \Omega. \tag{3.12}$$

Conversely, if K is a postive function on $\mathbb{R}^{\Omega \times \Omega}$, then for $\mu, \upsilon \in E'$ with

$$\mu = \sum_{x \in \Omega} c_x \delta_x, \qquad \upsilon = \sum_{y \in \Omega} d_y \delta_y,$$

$H: E' \to E'$ can be defined by

$$\langle H\upsilon, \mu \rangle = \sum_{x,y \in \Omega} c_x d_y K(x,y). \tag{3.13}$$

$H\upsilon$ takes x to $\sum_{y \in \Omega} d_y K(x,y)$; in particular,

$$H\delta_y = K(\cdot, y), \tag{3.14}$$

thus

$$\langle H\delta_y, \delta_x \rangle = K(x,y), \quad x,y \in \Omega.$$

This one-to-one correspondence between $\mathbb{L}^+(E',E)$, and positive functions on $\Omega \times \Omega$ (suitably topologised) extends to an isomorphism between these sets, and thus an isomorphism also between $\mathrm{Hilb}(\mathbb{R}^{\Omega})$ and reproducing kernel functions *(Schwartz, 1973, p 339)*.

Let $\mathbb{H} \in \mathrm{Hilb}(\mathbb{R}^{\Omega})$, and let K be the unique reproducing kernel function associated with $\mathbb{H}$; let $h \in \mathbb{H}$, and $x \in \Omega$, then by (3.14) $H\delta_x = K(\cdot,x)$, and (3.6):

$$(h, K(\cdot,x))_{\mathbb{H}} = \langle H, \delta_x \rangle = h(x). \tag{3.15}$$

In particular,

$$(K(\cdot,y), K(\cdot,x))_{\mathbb{H}} = K(x,y),$$

$$\| K(\cdot,x) \|^2_{\mathbb{H}} = K(x,x),$$

$$|h(x)| \leq \| h \| K(x,x)^{1/2},$$

and

$$K(x,y) = B(\delta_y, \delta_x).$$

In (3.15) can be seen the reproducing property that gives reproducing kernel functions their name. Because they relate evaluation to the inner product of the Hilbert space, and thereby to the geometry of Hilbert space projection, reproducing kernels have played an important

$<g,\delta_{x_i}> = 0$, $i = 1,\ldots,n$, then $g = 0$, i.e. the condition for a unique spline solution holds. Let K be the reproducing kernel associated with X as Hilbert sub-space of E. Let $f \in X$; the interpolating spline s is the unique interpolant orthogonal to KerF - it is given by

$$s = \sum_{i=1}^{n} a_i K(\cdot,x_i), \tag{3.16}$$

where the coefficients a_i are obtained from the solution of the n linear equations

$$\sum_{i=1}^{n} a_i K(x_j,x_i) = f(x_j), \quad j = 1,\ldots,n. \tag{3.17}$$

From (3.11) and (3.6) it can be seen that, if δ_{x_i}, $i = 1,\ldots,n$, are linearly independent, then $K(\cdot,x_i)$ are independent too, and the Gram matrix with elements

$$\begin{aligned} <<K(\cdot,\cdot),\delta_{x_j}>,\delta_{x_i}> &= (K(\cdot,x_j),K(\cdot,x_i))_X \\ &= K(x_j,x_i) \end{aligned}$$

is non-singular; thus (3.17) can be solved for a_i, $i = 1,\ldots,n$, and (3.16) does interpolate f at x_j:

$$s(x_j) = \sum_{i=1}^{n} a_i <K(\cdot,x_i),\delta_{x_j}>$$

$$\sum_{i=1}^{n} a_i K(x_j,x_i) = f(x_j),$$

and, most important, if $g \in$ KerF, then $(s,g)_X = 0$, for

$$\left(\sum_{i=1}^{n} a_i K(\cdot, x_i), g \right)_X = \sum_{i=1}^{n} a_i g(x_i) = 0,$$

and so (3.16) does indeed represent an interpolating spline for the Sard system $(X, Y, \mathbb{R}^n, U, F)$.

A reproducing kernel for X (with inner product (2.18)) can be constructed from the sum of the reproducing kernels for KerU and $(\text{KerU})^+$ in the way shown by *Schempp and Tippenhauer (1974)*, and used by *Freeden (1981a)* (for harmonic splines). Here the topic of reproducing kernel functions will not be taken further; in the following two chapters we shall build instead on the results of Section 3.3.

3.6 Green's Functions

An explicit construction for the kernel mapping θ associated with a semi-Hilbert sub-space will be given in Chapter 4, and also in Chapter 5, for a number of spaces of interest for gravity field representation, and the theorem of Section 3.3 used to derive practical computational formulae for spline interpolation. These kernel mappings turn out to be simpler than their corresponding reproducing kernel functions.

The attraction of Green's functions is that they too, are simpler than reproducing kernel functions. It is for this reason, and also because Green's functions arise naturally in the solution of boundary value problems (in which light spline interpolation, or smoothing, can also be viewed), that some authors have adopted Green's functions in constructing spline representations, for example, *Freeden (1981b, 1982)*, and *Sünkel, 1984)*, in constructing spherical splines and harmonic splines.

This section gives a brief outline of the connection between Green's functions, reproducing kernel functions, and the spline theory developed thus far. The material of this section is based largely on *Delvos and Schempp (1976)*.

Let (X, Y, Z, U, F) be a Sard system, where the spaces X, Y, Z and mappings U, F are as defined in Section 2.2, and let P be the spline projector corresponding to this spline system. P is the orthogonal projector on X (with inner product (2.18)), with

$$\mathrm{KerP} = \mathrm{KerF}; \quad \mathrm{ImP} = (\mathrm{KerF})^{\perp}.$$

At this stage two further assumptions are made:

(i) X is a sub-space of Y, and the inclusion map $X \to Y$ is continuous;

(ii) $X_0 = \mathrm{KerF}$ is dense in Y.

Under these assumptions *Delvos and Schempp (1974)* call (X, Y, Z, U, F) an *extended Sard system* (or, in the language of differential equations, (X,Y,Z,U,F) possesses the *trace property (Aubin, 1979, p 390)*).

Let U_0 be the restriction of U to X_0; since X_0 is dense in Y, it makes sense to consider

$$A = U_0^* U_0, \tag{3.18}$$

where U_0^* is the adjoint of U_0. Since U_0 is closed in Y (as can easily be shown, *Delvos and Schempp, 1976)*, A is a self-adjoint positive

operator on Y *(Riesz and Sz-Nagy, 1955, p 312)*; moreover, A is postive-definite: $\mathrm{Dom}A \subset X_0$; let $x \in \mathrm{Dom}A$,

$$(Ax,x)_Y = (Ux,Ux)_Y = (x,x)_X \geq C(x,x)_Y,$$

for some $C > 0$. By the Friederichs extension theorem *(Riesz and Sz-Nagy, 1955, p 330)* $A = U_0^* U_0$ is the unique self-adjoint, positive-definite operator satisfying $\mathrm{Dom}A \subset X_0$ and

$$(Ax,y)_Y = (Ux,Uy)_Y = (x,y)_X, \quad x \in \mathrm{Dom}A, \; y \in X_0 = \mathrm{Ker}F. \tag{3.19}$$

Delvos and Schempp (1976) thus call A the *operator associated with the Sard system* (X, Y, Z, U, F).

Now suppose $E = \mathbb{R}^{\Omega}$, and let $K: \Omega \times \Omega \to \mathbb{R}$ be the reproducing kernel associated with X as a Hilbert sub-space of E (X Hilbert space with inner product $(\cdot,\cdot)_X$ given by (2.18)). The reproducing property is given by (3.15): for $x \in X$, $t \in \Omega$

$$x(t) = (x,K(\cdot,t))_X.$$

Define

$$G(\cdot,t) = K(\cdot,t) - PK(\cdot,t), \quad t \in \Omega. \tag{3.20}$$

Thus $G(\cdot,t) \in X_0$, since $X_0 = \mathrm{Ker}F$, and $\mathrm{Im}P = (\mathrm{Ker}F)^{\perp}$. For $x \in \mathrm{Dom}A$, $t \in \Omega$

$$x(t) = (x,K(\cdot,t))_X \qquad \text{(continued)}$$

$$= (x,G(\cdot,t))_X \qquad (x \in \mathrm{Ker}F)$$

$$= (Ux,UG(\cdot,t))_Y$$

$$= (Ax,G(\cdot,t))_Y.$$

Thus

$$x(t) = (Ax,G(\cdot,t))_Y, \quad x \in \mathrm{Dom}A, \ t \in \Omega.$$

Because of this result $G(\cdot,\cdot)$ is called the *Green function of A associated with the Sard system* (X,Y,Z,U,F). $G(\cdot,\cdot)$ is the reproducing kernel function of the space X_0. Thus the results (3.16) and (3.17) of Section 3.5 can be directly transposed to the present context, to obtain spline approximations to functions $x \in X_0$.

CHAPTER 4

NATURAL ROTATION-INVARIANT SPLINES ON $\mathbb{R}^n$

4.1 Introduction

Polynomial splines are well known for their efficacy in approximating functions of a single variable; the properties that make these splines so useful should also be present in multivariate splines. On the other hand, surface spline interpolation techniques based on linear combinations of radial functions - and, in geodetic and surveying applications, especially Hardy's multiquadric functions - have long been known to be efficient and easy to use. In the sections that follow we introduce spline functions that are multivariate generalisations of familiar single-variable splines. It turns out that these are linear combinations of simple radial functions, and include multi-conic functions (the simplest of Hardy's multiquadrics). They minimise rotation-invariant semi-norms on $\mathbb{R}^n$, and thus unlike multivariate splines constructed through tensor products of one-dimensional one's, are not restricted to rectangular regions and meshes, nor use 'somewhat unnatural' function spaces and norms *(Meinguet, 1979c)*.

The different splines to be encountered will be constructed by choosing specific semi-Hilbert sub-spaces, and then finding the specific form of the kernel mapping $\theta: N^\circ \to \mathbb{H}$, where N is the nullspace of the semi-Hilbert space $\mathbb{H}$. Within the family of splines constructed in this way will be multi-conic splines (seen to be a natural multivariate generalisation of linear interpolation), pseudo-cubic splines (a multivariate generalisation of cubic splines), thin-plate splines, and subject to some limitations, point mass splines as well.

Splines solve a variational problem, and the semi-norms that arise most naturally are those induced by differential operators. It is thus not surprising that the theory in its most general form needs to be developed in spaces of distributions (or generalised functions), nor that Fourier transforms play an important role. The first few sections that follow summarise the theory of distributions and their Fourier transforms; the reason for including this material is that I was hitherto unfamiliar with distributions. Sections 4.5 and 4.6 introduce Sobolev spaces and Beppo-Levi spaces, respectively - these are spaces naturally associated with differential operators - and in Section 4.7, where I draw heavily on *Duchon (1977)*, kernel mappings are constructed for the semi-Hilbert Beppo-Levi spaces. These mappings are of fundamental practical importance for the construction of interpolating splines (Section 4.8), concrete examples of which follow in Section 4.9. In Section 4.10 reproducing kernel functions are obtained from the kernel mappings, and an allusion made to their significance for stable numerical computations. Whereas most of this chapter deals with interpolating splines, the next section shows how the results are readily transposed to the case of smoothing splines as well.

Smoothing splines 'smooth' the data; an alternative form of smoothing is to 'smooth' (or regularise) the interpolating spline functions themselves, making them less variable. Regularisation is essential when the desired spline functions are of themselves insufficiently continuous, as happens with point mass splines, for example, at the mass points. Section 4.12 deals with the regularisation of splines on $\mathbb{R}^n$, though it falls short of a comprehensive and general theory of regularisation. A final short section on the approximation of linear functionals

completes the chapter.

4.2 Distributions

The purpose of this section is to give a brief overview of the theory of distributions; the material presented here is based largely on *Rudin (1973)*.

A *multi-index* is a vector in $(\mathbb{Z}^+)^n$ ($\mathbb{Z}^+$: non-negative integers):

$$\alpha \in (\mathbb{Z}^+)^n; \alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n).$$

The *order* of α is $|\alpha| = \alpha_1 + \alpha_2 + \ldots + \alpha_n$. The differential operator D^α (on some domain in $\mathbb{R}^n$) is

$$D^\alpha = (\frac{\partial}{\partial x_1})^{\alpha_1} (\frac{\partial}{\partial x_2})^{\alpha_2} \ldots (\frac{\partial}{\partial x_n})^{\alpha_n}.$$

The *monomial* x^α, $x \in \mathbb{R}^n$, is

$$x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \ldots x_n^{\alpha_n}.$$

Let Ω be domain (a non-empty, open, connected set) in $\mathbb{R}^n$. A complex-valued function (of a real vector variable) f on Ω is in $C^\infty(\Omega)$, if it has continuous derivatives of every order, i.e. if $D^\alpha f \in C^\infty(\Omega)$ for every multi-index α.

$C_0^\infty(\Omega)$ is the sub-space of $C^\infty(\Omega)$ consisting of C^∞-functions on Ω with compact support. There is a standard way of making $C_0^\infty(\Omega)$ a complete,

locally convex topological space - see *Rudin (1973), p 136.* $C_0^\infty(\Omega)$ when given this canonical *Schwartz topology* is denoted $\mathbb{D}(\Omega)$ and called the *space of test functions.* The topology on $\mathbb{D}(\Omega)$ is characterised by the following form of uniform convergence: if a sequence of functions $\phi_{i\ i=1,2,\ldots}$ in $\mathbb{D}(\Omega)$ converges to ϕ (in the topology of $\mathbb{D}(\Omega)$), then there is a compact set K in Ω such that the support $\text{supp}(\phi_i) \subset K$, $i = 1,2\ldots$, and $D^\alpha\phi_i$ converges to $D^\alpha\phi$, uniformly on K, for every multi-index α.

Every differential operator D^α is a continuous mapping of $\mathbb{D}(\Omega)$ to itself.

The space of continuous linear functionals on $\mathbb{D}(\Omega)$ is denoted $\mathbb{D}'(\Omega)$; it is called the *space of distributions on* Ω, and given the weak-* topology- the topology of pointwise convergence: a sequence $\{\ell_i\}_{i=1,2,\ldots}$ converges to ℓ in $\mathbb{D}'(\Omega)$ if (and only if)

$$\lim_{i\to\infty} \ell_i\phi = \ell\phi \tag{4.1}$$

for all $\phi\in\mathbb{D}(\Omega)$. $\ell \in \mathbb{D}'(\Omega)$ if (and only if) for each compact $K\subset\Omega$ there is $m\in\mathbb{N}$ and a constant $c<\infty$ such that

$$|\ell\phi| \leq c\|\phi\|_m \tag{4.2}$$

for all $\phi \in \mathbb{D}(\Omega)$ whose support lies in K, i.e. for all $\phi \in \mathbb{D}(\Omega)_K$, and where

$$\|\phi\|_m = \max_{|\alpha|\leq m} \sup_{x\in\Omega} |D^\alpha\phi(x)|. \tag{4.3}$$

Four examples are given of distributions:

(i) If $x \in \Omega$, the linear functional δ_x on $\mathbb{D}(\Omega)$ defined by

$$\delta_x \phi = \phi(x) \tag{4.4}$$

is, in view of what has just been said in characterising $\mathbb{D}'(\Omega)$, indeed an element of $\mathbb{D}'(\Omega)$ - m can be taken as 0. The *Dirac measures* on $\mathbb{R}^n$ therefore define distributions.

(ii) Locally integrable functions are distributions: suppose $f \in \mathbb{L}'_{loc}(\Omega)$, i.e. f is Lebesgue measurable, and

$$\int_K |f(x)| dm(x) < \infty$$

for every compact set K in Ω (and dm(x) is Lebesgue measure). Put

$$\ell_f \phi = \int_\Omega f(x)\phi(x) dm(x), \quad \phi \in \mathbb{D}(\Omega), \tag{4.5}$$

then, for K compact in Ω,

$$|\ell_f \phi| \leq (\int_\Omega |f(x)| dm(x)) \| \phi \|_0, \quad \phi \in \mathbb{D}_K;$$

thus $\ell_f \in \mathbb{D}'(\Omega)$. f is identified with ℓ_f.

(iii) If ℓ is a distribution, then all its derivatives are also distributions: let $\ell \in \mathbb{D}'(\Omega)$, and let α be a multi-index; define

$$(D^\alpha \ell)\phi = (-1)^{|\alpha|} \ell(D^\alpha \phi), \quad \phi \in \mathbb{D}(\Omega). \tag{4.6}$$

This definition for the distributional derivative requires some motivation - see *Rudin (1973), p 136,* or *Adams (1975), p 20.* It clearly defines a linear functional, and we check that it defines a bounded linear functional. Let $\ell \in \mathbb{D}'(\Omega)$; for K compact in Ω, there is an $m \in \mathbb{N}$ such that

$$|\ell\phi| \leq c\| \phi \|_m$$

for all $\phi \in \mathbb{D}_K$; then

$$|(D^\alpha \ell)\phi| = |\ell(D^\alpha \phi)| \leq c\| D^\alpha \phi \|_m \leq c\| \phi \|_{m+|\alpha|},$$

i.e. $D^\alpha \ell \in \mathbb{D}'(\Omega)$.

(iv) Distributions multiplied by smooth functions are again distributions: let $\ell \in \mathbb{D}'(\Omega)$ and $f \in C^\infty(\Omega)$. Define $f\ell$ by

$$(f\ell)\phi = \ell(f\phi), \quad \phi \in \mathbb{D}(\Omega). \tag{4.7}$$

The proof that $f\ell$ is also a distribution by this definition is based on Leibniz's formula for repeated differentiation of a product - see *Rudin (1973), p 144.*

The fact that distributions may be partially differentiated as often as one may wish is an important part of the theory, as is the converse result: every distribution is locally the distributional derivative of a continuous function, i.e. let $\ell \in \mathbb{D}'(\Omega)$, K compact in Ω; then there is $f \in C(\Omega)$, and a multi-index α such that

$$\ell\phi = (-1)^{|\alpha|}\int_{\Omega} f(x)D^{\alpha}\phi(x)dm(x), \quad \phi \in \mathbb{D}_K, \tag{4.8}$$

(Rudin, 1973, p 152).

Distributions with compact support: The continuity of distributions on Ω with compact support has an even simpler characterisation. Let $\ell \in \mathbb{D}'(\Omega)$ with

$$\text{supp}(\Omega) = \{\phi \in \mathbb{D}(\Omega), \ \ell\phi \neq 0\}^{-}$$

compact in $\mathbb{D}(\Omega)$; then there is an $m \in \mathbb{N}$, and a constant $c < \infty$ such that

$$|\ell\phi| \leq c\| \phi \|_m$$

for all $\phi \in \mathbb{D}(\Omega)$. More importantly, ℓ has a unique extension to a continuous linear functional on $C^{\infty}(\Omega)$.

Convolutions: Finally, as preparation for the next section on Fourier transforms, it is necessary to say a few words about convolutions - the convolution of a distribution and a test function, and the convolution of two distributions.

First, translation and reflection operators are defined on $\mathbb{D}(\mathbb{R}^n)$. If $\phi \in \mathbb{D}(\mathbb{R}^n)$ and $x \in \mathbb{R}^n$, define *(Rudin, 1973, p 155):*

$$\begin{aligned} \tau_x\phi(y) &= \phi(y - x) \\ \tilde{\phi}(y) &= \phi(-y), \quad y \in \mathbb{R}^n. \end{aligned} \tag{4.9}$$

If $\phi \in \mathbb{D}(\mathbb{R}^n)$ and $u \in \mathbb{D}'(\mathbb{R}^n)$, the convolution of u and ϕ is defined to be

$$u*\phi(x) = u(\tau_x\tilde{\phi}), \quad x \in \mathbb{R}^n. \tag{4.10}$$

(When u is locally integrable this definition corresponds to the familiar definition of convolution:

$$u*\phi(x) = \int_{\mathbb{R}^n} u(y)\phi(x-y)dm(y) \ ).$$

If $u \in \mathbb{D}'(\mathbb{R}^n)$, its translate by x, $x \in \mathbb{R}^n$, is defined to be

$$(\tau_x u)\phi = u(\tau_{-x}\phi), \quad \phi \in \mathbb{D}(\mathbb{R}^n). \tag{4.11}$$

Within the calculus of distributions we have the following results ($u \in \mathbb{D}'(\mathbb{R}^n)$; $\phi, \psi \in \mathbb{D}(\mathbb{R}^n)$):

$$\begin{aligned} &\tau_x(u*\phi) = (\tau_x u)*\phi = u*(\tau_x\phi), \quad x \in \mathbb{R}^n; \\ &D^\alpha(u*\phi) = (D^\alpha u)*\phi = u*(D^\alpha\phi), \quad \alpha \in (\mathbb{Z}^+)^n; \\ &u*(\phi*\psi) = (u*\phi)*\psi. \end{aligned} \tag{4.12}$$

In particular

$$\delta_x*\phi(y) = \delta_x(\tau_y\tilde{\phi})(x) = \tilde{\phi}(x-y) = \phi(y-x),$$

or

$$\delta_x*\phi = \tau_x\phi.$$

Convolution is a smoothing or spreading operation: $u*\phi \in C(\mathbb{R}^n)$;

moreover, for a given $u \in \mathbb{D}'(\mathbb{R}^n)$, convolution defines a continuous linear mapping $L_u: \mathbb{D}(\mathbb{R}^n) \to C^\infty(\mathbb{R}^n)$ through

$$L_u\phi = u*\phi, \quad \phi \in \mathbb{D}(\mathbb{R}^n), \tag{4.13}$$

and further, L_u commutes with translation:

$$\tau_x L_u = L_u \tau_x, \quad x \in \mathbb{R}^n. \tag{4.14}$$

If u has compact support, it has a unique extension to a continuous linear functional on $C^\infty(\mathbb{R}^n)$, so if u is such a distribution, and $\phi \in C^\infty(\mathbb{R}^n)$, one can define the convolution

$$(u*\phi)(x) = u(\tau_x\tilde{\phi}), \quad x \in \mathbb{R}^n, \tag{4.15}$$

in the same way as before. Further, if $u \in \mathbb{D}'(\mathbb{R}^n)$ with compact support, $\phi \in C^\infty(\mathbb{R}^n)$ and $\psi \in \mathbb{D}(\mathbb{R}^n)$, then the three equations (4.12) hold, and in particular $u*\phi \in \mathbb{D}(\mathbb{R}^n)$. It thus makes sense to define convolution of two distributions, provided one of the two has compact support: let $u,v \in \mathbb{D}'(\mathbb{R}^n)$, and let at least one of u and v have compact support; define

$$(u*v)*\phi = u*(v*\phi), \quad \phi \in C^\infty(\mathbb{R}^n), \tag{4.16}$$

for if v has compact support, $v*\phi \in \mathbb{D}(\mathbb{R}^n)$ and $u*(v*\phi)$ is defined, and if u has compact support, $v*\phi \in C^\infty(\mathbb{R}^n)$ and $u*(v*\phi)$ is again defined. Extending the calculus of distributions to include convolutions of

distributions, we have, for $u,v \in \mathbb{D}'(\mathbb{R}^n)$, and at least one of u,v with compact support:

$$\begin{aligned} &u*v = v*u; \\ &D^\alpha(u*v) = (D^\alpha u)*v = u*(D^\alpha v), \qquad (\mathbb{Z}^+)^n; \\ &D^\alpha u = (D^\alpha \delta)*u; \end{aligned} \tag{4.17}$$

where δ is the Dirac measure at 0, $\delta\phi = \phi(0)$. In particular this means

$$u = \delta*u,$$

and, more generally that

$$\tau_x u = \delta_x * u, \quad x \in \mathbb{R}^n, \tag{4.18}$$

since

$$\begin{aligned} (\delta_x * u)*\phi &= \delta_x*(u*\phi) \\ &= \tau_x(u*\phi) \\ &= (\tau_x u)*\phi, \quad \phi \in \mathbb{D}(\mathbb{R}^n). \end{aligned}$$

4.3 The Fourier Transform and Rapidly Decreasing Functions

This section introduces the Fourier transform, and the Schwartz space of rapidly decreasing functions on which the Fourier transform is a homeomorphism. The material presented here can be found in many texts; I have followed *Rudin (1973), Chapter 7*, rather closely.

The notation is simplified by using *normalised Lebesgue measure* on $\mathbb{R}^n$,

$$dm_n(x) = (2\pi)^{-n/2} dm(x), \tag{4.19}$$

as a means of norming the Lebesgue spaces $\mathbb{L}^p(\mathbb{R}^n)$ $(1 \leq p < \infty)$:

$$\| f \|_p = \{\int_{\mathbb{R}^n} |f(x)|^p dm_n(x)\}^{1/p}, \quad f \in \mathbb{L}^p(\mathbb{R}^n). \tag{4.20}$$

The convolution of two functions on $\mathbb{R}^n$ now becomes

$$f*g(x) = \int_{\mathbb{R}^n} f(x-y)g(y)dm_n(y), \tag{4.21}$$

provided the integral exists.

The introduction of the *character function (Dym and McKean, 1972, p 205)* also helps to simplify the notation. For $t \in \mathbb{R}^n$, the *character* e_t is the function defined by

$$e_t(x) = e^{it\cdot x}, \quad x \in \mathbb{R}^n, \tag{4.22}$$

where $t\cdot x = \sum_{i=1}^{n} t_i x_i$. The character function satisfies

$$e_t(x+y) = e_t(x)e_t(y),$$

and e_t is thus a homomorphism of $\mathbb{R}^n$ (given its additive group structure) onto the multiplicative group $\{z \in \mathbb{C}: |z| = 1\}$.

The *Fourier transform* of $f \in \mathbb{L}^1(\mathbb{R}^n)$ is the function $\hat{f}$ defined by

$$\hat{f}(t) = \int_{\mathbb{R}^n} f(y) e_{-t}(y) dm_n(y), \quad t \in \mathbb{R}^n. \tag{4.23}$$

Thus

$$\hat{f}(t) = f * e_t(0) \tag{4.24}$$

(since $e_t(-y) = e_{-t}(y)$). Translation can be defined as before, (4.9): if f is a function on $\mathbb{R}^n$, and $x \in \mathbb{R}^n$,

$$\tau_x f(y) = f(x - y), \quad y \in \mathbb{R}^n.$$

Calculus of Fourier transforms of integrable functions: let $f, g \in \mathbb{L}^1(\mathbb{R}^n)$, and $x, t \in \mathbb{R}^n$; we have the following identities:

$$\begin{aligned} (\tau_x f)\hat{} &= e_{-x}\hat{f}; \\ (e_x f)\hat{} &= \tau_x \hat{f}; \\ (f * g)\hat{} &= \hat{f}\hat{g}. \end{aligned} \tag{4.25}$$

Moreover, if $\lambda > 0$, and $h(x) = f(x/\lambda)$, then $\hat{h}(t) = \lambda^n \hat{f}(\lambda t)$.

The Fourier transform is a map $\mathbb{F}$: $\mathbb{L}^1(\mathbb{R}^n) \to C_0(\mathbb{R}^n)$ to the space of complex-valued continuous functions on $\mathbb{R}^n$ that vanish at infinity. Furthermore

$$\| \hat{f} \|_\infty \leq \| f \|_1,$$

where

$$\| g \|_\infty = \sup_{t \in \mathbb{R}^n} |g(t)|, \quad g \in C_0(\mathbb{R}^n).$$

Rapidly decreasing functions: These are $f \in C^{\infty}(\mathbb{R}^n)$ such that

$$\max_{|\alpha| \leq m} \sup_{x \in \mathbb{R}^n} (1 + |x|^2)^m |(D^{\alpha} f)(x)| < \infty \qquad (4.26)$$

for all $m \in \mathbb{Z}^+$. Rapidly decreasing functions form a linear vector space Φ_n (called the *Schwartz space)*, for which the countable collection of norms (4.26) defines a locally convex topology. With this topology Φ_n can be shown to be a Frechet space, i.e. there is a complete translation-invariant metric inducing the locally convex topology on Φ_n *(Rudin, 1973, p 168)*.

$\mathbb{D}(\mathbb{R}^n)$ is dense in Φ_n; the canonical injection of $\mathbb{D}(\mathbb{R}^n)$ into Φ_n is continuous *(Rudin, 1973, p 173)*, and Φ_n, in turn, is dense in $\mathbb{L}^1(\mathbb{R}^n)$, as also in $\mathbb{L}^2(\mathbb{R}^n)$. A rapidly decreasing function, when multiplied by a polynomial, remains rapidly decreasing. On Φ_n the Fourier transform is invertible, for if $g \in \Phi_n$, then

$$g(x) = \int_{\mathbb{R}^n} \hat{g}(y) e_x(y) \, dm_n(y). \qquad (4.27)$$

(hence $\hat{\hat{g}}(x) = g(-x)$). The Fourier transform is a continuous linear one-to-one mapping of Φ_n onto Φ_n, with an inverse that is also continuous - the Fourier transform is thus a homeomorphism on Φ_n *(Rudin, 1973, p 170; Weidman, 1980, p 291)*.

Convolution on Φ_n: If $f, g \in \Phi_n$, then $f*g \in \Phi_n$, and

$$(fg)\hat{} = \hat{f}*\hat{g}; \qquad \text{(continued)}$$

$$\hat{f}\hat{g} = (f*g)\hat{}. \tag{4.28}$$

A classical result is that the Fourier transform is an isometry of $\mathbb{L}^2(\mathbb{R}^n)$ onto itself.

4.4 Tempered Distributions

The previous sections provide the background necessary for the introduction of tempered distributions. Let $i: \mathbb{D}(\mathbb{R}^n) \to \Phi_n$ be the canonical injection of $\mathbb{D}(\mathbb{R}^n)$ into Φ_n; it is continuous. Let ℓ be a continuous linear functional on Φ_n, and put

$$u_\ell = \ell i, \tag{4.29}$$

then u_ℓ is also continuous, with $u_\ell \in \mathbb{D}'(\mathbb{R}^n)$. Since $\mathbb{D}(\mathbb{R}^n)$ is dense in Φ_n, u_ℓ has a unique continuous extension to a linear functional on Φ_n. (4.29) thus expresses a vector space isomorphism between Φ_n' and a sub-space of $\mathbb{D}'(\mathbb{R}^n)$, the space of *tempered distributions* made up of those distributions in $\mathbb{D}'(\mathbb{R}^n)$ with continuous extensions to Φ_n. If u_ℓ is identified with ℓ, we can say Φ_n' is the space of tempered distributions.

Examples of tempered distributions:

(i) Every distribution with compact support is tempered.

(ii) Let $1 \leq p < \infty$ and $N > 0$, and suppose g is a measurable function on $\mathbb{R}^n$ such that

$$\int_{\mathbb{R}^n} |(1+|x|^2)^{-N} g(x)|^p dm_n(x) < \infty, \tag{4.30}$$

then g is a tempered distribution *(Rudin, 1973, p 175)*.

(iii) In consequence of (ii), every $g \in \mathbb{L}^p(\mathbb{R}^n)$, $1 \leq p < \infty$, is a tempered distribution.

(iv) Every polynomial is a tempered distribution.

(v) If $\alpha \in (\mathbb{Z}^+)^n$, $g \in \Phi_n$ and $u \in \Phi_n'$, then the distributions $D^\alpha u$ and gu are also tempered, where

$$(D^\alpha u)f = (-1)^{|\alpha|} u(D^\alpha f);$$
$$(gu)f = u(gf), \quad f \in \Phi_n. \tag{4.31}$$

(vi) A distribution each of whose derivatives is tempered, is itself tempered: from (4.26) it can be seen that a C^∞-function on $\mathbb{R}^n$ each of whose derivatives is rapidly decreasing, is itself rapidly decreasing, and the same is thus true also for distributions with tempered derivatives.

The Fourier transform of a tempered distribution: Let $u \in \Phi_n'$, and define

$$\hat{u}\phi = u\hat{\phi}, \quad \phi \in \Phi_n. \tag{4.32}$$

With this definition, the Fourier transform is a homeomorphism of Φ_n' on Φ_n'. The following results on Fourier transforms of differential operators on tempered distributions are needed for future reference. (I continue to use *Rudin (1973), Chapter 7)*. If α is a multi-index, define

$$D_\alpha = (i)^{-|\alpha|} D^\alpha = (\frac{1}{i}\frac{\partial}{\partial x_i})^{\alpha_1} \ldots \quad (\frac{1}{i}\frac{\partial}{\partial x_i})^{\alpha_n}. \tag{4.33}$$

Let P be a polynomial in n variables, with complex coefficients, thus

$$P(\xi) = \Sigma c_\alpha \xi^\alpha = \Sigma c_\alpha \xi_1{}^{\alpha_1} \ldots \quad \xi_n{}^{\alpha_n},$$

and define the differential operators P(D) and P(-D) to be

$$P(D) = \Sigma c_\alpha D_\alpha; \quad P(-D) = \Sigma (-1)^{|\alpha|} c_\alpha D_\alpha.$$

If e_t is a character, (4.22), then

$$D_\alpha e_t = t^\alpha e_t, \text{ and } P(D)e_t = P(t)e_t.$$

A result required later is that, if $u \in \Phi_n'$, and P is a polynomial, then

$$(P(D)u)\hat{} = P\hat{u} \tag{4.34}$$

(Rudin, 1973, p 176), and in particular

$$(D_\alpha u)\hat{} = t^\alpha \hat{u},$$

or

$$(D^\alpha u)\hat{} = (i)^{-|\alpha|} t^\alpha \hat{u}. \tag{4.35}$$

4.5 Sobolev Spaces

We now come a step closer to describing the spaces in which solutions to the spline interpolation and approximation problems are to be sought. The notation in this section will be further simplified by introducing, for $s \in \mathbb{R}$, a positive measure on $\mathbb{R}^n$ through

$$d\mu_s(y) = (1 + |y|^2)^s dm_n(y), \quad y \in \mathbb{R}^n. \tag{4.36}$$

If $f \in \mathbb{L}^2_{\mu_s}(\mathbb{R}^n)$, i.e. if

$$\int_{\mathbb{R}^n} |f(x)|^2 d\mu_s(x) < \infty,$$

then f is a tempered distribution, and hence the Fourier transform of a tempered distribution, $f = \hat{u}$; the vector space of all tempered distributions u arising in this way is the (fractional order) *Sobolev space* $H^s(\mathbb{R}^n)$:

$$H^s(\mathbb{R}^n) = \{u \in \Phi'_n : \int_{\mathbb{R}^n} (1 + |y|^2)^s |\hat{u}(y)|^2 dm_n(y) < \infty\} \qquad (4.37)$$

(Lions and Magenes, 1972, p 30). $H^s(\mathbb{R}^n)$ is thus the space of inverse Fourier transforms of measurable functions in $\mathbb{L}^2_{\mu_s}(\mathbb{R}^n)$. By Plancherel's theorem (the Fourier transform is an $\mathbb{L}^2(\mathbb{R}^n)$-isometry), $H^0(\mathbb{R}^n) = \mathbb{L}^2(\mathbb{R}^n)$. For $r < s$, $H^s(\mathbb{R}^n) \subset H^r(\mathbb{R}^n)$, and the inclusion is continuous.

$H^s(\mathbb{R}^n)$ is a Hilbert space with inner product

$$(u,v)_{H^s(\mathbb{R}^n)} = \int_{\mathbb{R}^n} \hat{u}(y)\overline{\hat{v}(y)} d\mu_s(y), \quad u,v \in H^s(\mathbb{R}^n), \qquad (4.38)$$

and norm

$$\| u \|^2_{H^s(\mathbb{R}^n)} = \int_{\mathbb{R}^n} |\hat{u}(y)|^2 d\mu_s(y), \quad u \in H^s(\mathbb{R}^n). \qquad (4.39)$$

$H^s(\mathbb{R}^n)$ is thus isometrically isomorphic to $\mathbb{L}^2_{\mu_s}(\mathbb{R}^n)$, and indeed a Hilbert space. For $s > 0$, the dual of $H^s(\mathbb{R}^n)$ is identified with $H^{-s}(\mathbb{R}^n)$, thus

$$(H^s(\mathbb{R}^n))' = H^{-s}(\mathbb{R}^n) \tag{4.40}$$

(Lions and Magenes, 1972, p 31). For $s \in \mathbb{Z}^+$, i.e. a non-negative integer, the definition of a Sobolev space (4.37) is equivalent to

$$H^s(\mathbb{R}^n) = \{u \in \mathbb{L}^2(\mathbb{R}^n) : D^\alpha u \in \mathbb{L}^2(\mathbb{R}^n),\ |\alpha| \leq s\}. \tag{4.41}$$

We denote by $D^{-m}(H^s(\mathbb{R}^n))$ $(m \in \mathbb{Z}^+)$ the distributions $u \in \Phi_n'$ such that $D^\alpha u \in H^s(\mathbb{R}^n)$ for $|\alpha| = m$. For K a closed set in $\mathbb{R}^n$, $H_K^s(\mathbb{R}^n)$ is the set of distributions in $H^s(\mathbb{R}^n)$ whose support is contained in K. (The definition of support implied here is that given by *Rudin (1973), p 149)*. $H_K^s(\mathbb{R}^n)$ is a closed linear sub-space of $H^s(\mathbb{R}^n)$. If Ω is an open set in $\mathbb{R}^n$, we denote by $H^s(\Omega)$ the set of restrictions to Ω of distributions in $H^s(\mathbb{R}^n)$. $H^s(\Omega)$ is isomorphic to $H^s(\mathbb{R}^n)/H^s_{\mathbb{R}^n/\Omega}(\mathbb{R}^n)$ *(Lions and Magenes, 1972, p 40)*, and hence also a Hilbert space. The dual of $H^s(\Omega)$ can be identified with $H_{\overline{\Omega}}^{-s}(\mathbb{R}^n)$.

If some condition is placed on Ω, e.g. if Ω is $\mathbb{R}^n$, or a half-space in $\mathbb{R}^n$, or Ω is bounded with a sufficiently regular boundary (circumstances made more explicit by *Adams (1975), p 66, p 214)*, the functions u in $H^s(\Omega)$, for s non-negative and non-integral, are those satisfying

$$D^\alpha u \in \mathbb{L}^2(\Omega), \quad |\alpha| \leq [\![s]\!]$$

and

$$\sum_{|\alpha| = [\![s]\!]} \iint_{\Omega \times \Omega} \frac{|D^\alpha u(x) - D^\alpha u(y)|^2}{|x - y|^{n+2(s-[\![s]\!])}}\, dm(x)\, dm(y) < \infty,$$

where $[\![s]\!]$ is the integral part of s. For integral $s \geq 0$, we have

$$D^{\alpha}u \in \mathbb{L}^2(\Omega), \quad |\alpha| \leq s.$$

$H^s_{loc}(\mathbb{R}^n)$ is the set of distributions on $\mathbb{R}^n$ (whether tempered or not) whose restrictions to a bounded open set Ω is in $H^s(\Omega)$. Locally, $D^{-m}(H^s) = H^{m+s}$ *(Rudin, 1973, p 201)*. $C^k(\mathbb{R}^n) \subset H^s_{loc}(\mathbb{R}^n)$ for $k \geq s$. The converse result, for $s > k + \frac{n}{2}$, $H^s_{loc}(\mathbb{R}^n) \subset C^k(\mathbb{R}^n)$, with the inclusion being continuous, is part of *Sobolev's embedding theorem (Lions and Magenes, 1972, p 45; Adams, 1975, p 217)*. As *Duchon (1977), p 88,* indicates, $H^s_{loc}(\mathbb{R}^n)$ can be given a Frechet space structure, and its dual be identified with $H^{-s}_{comp}(\mathbb{R}^n)$, restrictions to compact sets.

The spaces $\tilde{H}^s(\mathbb{R}^n)$, $s < \frac{n}{2}$ *(Duchon, 1977, p 89):* Put

$$\tilde{H}^s(\mathbb{R}^n) = \{u \in \Phi'_n : \int_{\mathbb{R}^n} |y|^{2s} |\hat{u}(y)|^2 dm_n(y) < \infty\} \tag{4.42}$$

(in contrast to $H^s(\mathbb{R}^n)$ which was defined to be, (4.37)

$$H^s(\mathbb{R}^n) = \{u \in \Phi'_n : \int_{\mathbb{R}^n} (1 + |y|^2)^s |\hat{u}(y)|^2 dm_n(y) < \infty\}\ ).$$

$\tilde{H}^s(\mathbb{R}^n)$ can be normed by

$$\| u \|^2_{\tilde{H}^s(\mathbb{R}^n)} = \int_{\mathbb{R}^n} |y|^{2s} |\hat{u}(y)|^2 dm_n(y). \tag{4.43}$$

We shall see that $\tilde{H}^s(\mathbb{R}^n)$ is a Hilbert space (always with $s < \frac{n}{2}$), but note for the moment that it is clear from (4.42), that for

$$s > 0: \quad \tilde{H}^s(\mathbb{R}^n) \subset H^s(\mathbb{R}^n); \qquad \text{(continued)}$$

$$s = 0: \quad H^0(\mathbb{R}^n) = \tilde{H}^0(\mathbb{R}^n) = \mathbb{L}^2(\mathbb{R}^n);$$
$$s < 0: \quad \tilde{H}^s(\mathbb{R}^n) \subset H^s(\mathbb{R}^n).$$

In general, $H^s(\mathbb{R}^n) \neq \tilde{H}^s(\mathbb{R}^n)$, since it can be seen from the definitions (4.37), (4.42) that for $s > 0$, $H^s(\mathbb{R}^n) \subset \mathbb{L}^2(\mathbb{R}^n)$, but $\tilde{H}^s(\mathbb{R}^n)$ is not contained in $\mathbb{L}^2(\mathbb{R}^n)$. But it can be shown that on bounded sets the two spaces coincide:

Theorem: *(Duchon, 1977, p 89).* $\tilde{H}^s(\mathbb{R}^n)$ is a Hilbert space if (and only if) $s < \frac{n}{2}$; moreover, $\tilde{H}^s(\mathbb{R}^n)$ is a Hilbert sub-space of Φ_n' (and hence also of $\mathbb{D}'(\mathbb{R}^n)$).

Proof: Let $f \in \mathbb{L}^2(\mathbb{R}^n)$, and let K be a bounded set in $\mathbb{R}^n$.

$$\int_K |y|^{-s} |f(y)| dm_n(y) \leq \left(\int_K |y|^{-2s}\right)\left(\int_K |f(y)|^2 dm_n(y)\right)^{1/2}.$$

$|y|^\lambda$ is locally integrable iff $\lambda > -n$; therefore $|y|^{-s} f(y)$ is locally integrable, hence a distribution, and in fact a tempered distribution (by (4.30)) iff $s < \frac{n}{2}$. Now consider the mapping from $\mathbb{L}^2(\mathbb{R}^n)$ onto $\tilde{H}^s(\mathbb{R}^n)$ defined by

$$f \mapsto (|y|^{-s} f)^\wedge. \tag{4.44}$$

$$\| (|y|^{-s} f)^\wedge \|^2_{\tilde{H}^s(\mathbb{R}^n)} = \int_{\mathbb{R}^n} |y|^{2s} |(|y|^{-s} f(y))^\wedge|^2 dm_n(y)$$

$$= \int_{\mathbb{R}^n} |f(-y)|^2 dm_n(y) = \| f \|^2_{\mathbb{L}^2(\mathbb{R}^n)}.$$

Therefore the map $f \mapsto (|y|^{-s} f)^\wedge$ is an isometry, and $\tilde{H}^s(\mathbb{R}^n)$ a Hilbert

space.

It will now be shown that $\tilde{H}^s(\mathbb{R}^n)$ is a Hilbert sub-space of Φ_n', i.e. that the inclusion map in Φ_n' is continuous; let $\{f_j\}_{j=1,2,\ldots}$ be a sequence in $\mathbb{L}^2(\mathbb{R}^n)$ with $\hat{f}_j \to 0$ (in $\mathbb{L}^2(\mathbb{R}^n)$). Under the mapping (4.44) $\hat{f}_j \mapsto |y|^{-s} f_j(-y)$ and $|y|^{-s} f_j \to 0$ in $\tilde{H}^s(\mathbb{R}^n)$. We show it does the same in Φ_n' by showing

$$\ell_j \phi = \int_{\mathbb{R}^n} |y|^{-s} f_j(-y)\phi(y)\, dm_n(y) \to 0$$

for all $\phi \in \Phi_n$. Now

$$\begin{aligned} |\ell_j \phi| &\leq \left(\int_{\mathbb{R}^n} |f_j(-y)|^2 dm_n(y)\right)^{1/2} \left(\int_{\mathbb{R}^n} |y|^{-2s} |\phi(y)|^2 dm_n(y)\right)^{1/2} \\ &\leq \| f_j \|_{\mathbb{L}^2(\mathbb{R}^n)} \left(\int_{\mathbb{R}^n} \frac{|y|^{-2s}}{(1+|y|^2)^{2N}} dm_n(y)\right)^{1/2} \cdot \\ &\qquad \cdot \left(\sup_{x \in \mathbb{R}^n} \{(1+|y|^2)^N |\phi(y)|\}\right) \\ &= \| \hat{f}_j \|_{\mathbb{L}^2(\mathbb{R}^n)} \cdot B \cdot C \to 0, \end{aligned}$$

where N is chosen so large that

$$\int_{\mathbb{R}^n} \frac{|y|^{-2s}}{(1+|y|^2)^{2N}} dm_n(y) = B^2 < \infty,$$

and where (cf. (4.26))

$$C = \sup_{x \in \mathbb{R}^n} (1+|y|^2)^N |\phi(y)| < \infty.$$

As preparation for the theorem to come on the equality of $H^s(\Omega)$ and

and $\tilde{H}^s(\Omega)$, there is the following lemma giving a characterisation for distributions in $\tilde{H}^s(\mathbb{R}^n)$.

Lemma: Let $u \in \tilde{H}^s(\mathbb{R}^n)$, with $s > 0$, then u can be represented as

$$u = u_1 + u_2 \tag{4.45}$$

with $u_1 \in H^s(\mathbb{R}^n)$, and $u_2 \in C^\infty(\mathbb{R}^n)$.

Proof: Let $B_1 = \{x \in \mathbb{R}^n: |x| \leq 1\}$ be the unit ball in $\mathbb{R}^n$. If $y \in \mathbb{R}^n$, $y \notin B_1$, we have

$$(1 + |y|^2)^s \leq 2|y|^{2s}. \tag{4.46}$$

Let $u \in H^s(\mathbb{R}^n)$, then $|y|^s \hat{u} \in \mathbb{L}^2(\mathbb{R}^n)$, and $\hat{u}$ is a Lebesgue measurable function. Put

$$\hat{u} = v_1 + v_2,$$

where $v_2 = \hat{u}$ almost everywhere on B, and is zero elsewhere. We can say of v_1 that

$$\begin{aligned}\int_{\mathbb{R}^n}(1 + |y|^2)^s |v_1(y)|^2 dm_n(y) &= \int_{\mathbb{R}^n/B_1}(1 + |y|^2)^s |v_1(y)|^2 dm_n(y) \\ &\leq 2\int_{\mathbb{R}^n/B_1}|y|^{2s}|v_1(y)|^2 dm_n(y) \\ &\leq 2\int_{\mathbb{R}^n}|y|^{2s}|\hat{u}(y)|^2 dm_n(y) < \infty,\end{aligned}$$

which means that the inverse Fourier transform of v_1 is in $H^s(\mathbb{R}^n)$, i.e. if $\hat{u}_1 = v_1$, then $u_1 \in H^s(\mathbb{R}^n)$. On the other hand v_2 has compact support. Its Fourier transform (or inverse Fourier transform) is a function in $C^\infty(\mathbb{R}^n)$ - the Paley-Wiener theorem *(Rudin, 1973, p 180)*. Thus $\tilde{H}^s(\mathbb{R}^n) \subset H^s(\mathbb{R}^n) + C^\infty(\mathbb{R}^n)$.

Corollary: Let $s < 0$, and let $u \in H^s(\mathbb{R}^n)$, then u can be represented as

$$u = u_1 + u_2, \tag{4.47}$$

where $u_1 \in \tilde{H}^s(\mathbb{R}^n)$, and $u_2 \in C^\infty(\mathbb{R}^n)$.

Proof: The inequality (4.46) is reversed; for the rest the proof follows the same lines as the lemma.

Theorem: *(Duchon, 1973, p 89)*. Let $-\frac{n}{2} < s < \frac{n}{2}$, and let Ω be an open bounded set in $\mathbb{R}^n$; then

$$\tilde{H}^s(\Omega) = H^s(\Omega). \tag{4.48}$$

Proof: For the case $s > 0$: it has already been remarked that $C^\infty(\mathbb{R}^n) \subset H^s_{loc}(\mathbb{R}^n)$, i.e. $C^\infty(\Omega) \subset H^s(\Omega)$. On the other hand we have just seen in the lemma that $\tilde{H}^s(\Omega) \subset H^s(\Omega) + C^\infty(\Omega)$, therefore $\tilde{H}^s(\Omega) = H^s(\Omega)$. The case $s < 0$ will follow in a similar fashion, using the corollary to the lemma, if $C^\infty(\mathbb{R}^n) \subset H^s_{loc}(\mathbb{R}^n)$. For $s < -\frac{n}{2}$, $|y|^{2s}$ is locally integrable. The functions $\mathbb{D}(\mathbb{R}^n)$ are continuous and have compact support. Thus $\mathbb{D}(\mathbb{R}^n) \subset \tilde{H}^s(\mathbb{R}^n)$, which is the same as saying

$$C_0^\infty(\mathbb{R}^n) \subset \tilde{H}^s(\mathbb{R}^n)$$

and from which we infer that

$$C^\infty(\mathbb{R}^n) \subset \tilde{H}^s_{loc}(\mathbb{R}^n).$$

Finally, since $H^s(\mathbb{R}^n) \subset \tilde{H}^s(\mathbb{R}^n)$, for $s > 0$, with continuous inclusion (and $\tilde{H}^s(\mathbb{R}^n) \subset H^s(\mathbb{R}^n)$ for $s < 0$, also with continuous inclusion), the two Hilbert space structures on $H^s(\Omega) = \tilde{H}^s(\Omega)$ are comparable, and therefore equivalent *(Pryce, 1973, p 145)*.

4.6 Beppo-Levi Spaces

The spaces to be looked at in this section are spaces of distributions all of whose m'th derivatives are in $\tilde{H}^s(\mathbb{R}^n)$, $s < \frac{n}{2}$; these are the semi-Hilbert sub-spaces in which spline solutions will soon be obtained.

It will be convenient to extend the index notation beyond the multi-index notation introduced in Section 4.2. For non-negative integers m, let

$$\beta = (i_1, i_2, \ldots, i_m)$$

be an m-vector taking values in $\{1,2,\ldots,n\}$. The number of such m-vectors is n^m. For $x \in \mathbb{R}^n$, $x = (x_1, x_2, \ldots, x_n)$, let

$$x^\beta = x_{i_1} x_{i_2} \ldots x_{i_m}.$$

x^β is thus a monomial of degree m (as was x^α, $|\alpha| = m$). x^m is defined

to be the n^m-tuple of all monomials x^β, with

$$|x^m|^2 = \sum_{\beta \in n^m} |x^\beta|^2 = |x|^{2m}, \tag{4.49}$$

whereas

$$\sum_{|\alpha|=m} |x^\alpha| \leq |x|^{2m}.$$

Multiplication is carried out co-ordinatewise, treating x^m as an n^m-vector. Differential operators are similarly defined:

$$D^\beta = \frac{\partial}{\partial x_{i_1}} \frac{\partial}{\partial x_{i_2}} \cdots \frac{\partial}{\partial x_{i_m}}.$$

Thus $D^m u$ is the n^m-tuple of derivatives $D^\beta u$, and

$$|D^m u|^2 = \sum_{\beta \in n^m} |D^\beta u|^2.$$

In the notation just introduced the gradient and Laplacian are, for example

$$\nabla = D^1 = \frac{\partial}{\partial x_1} \underline{e}_1 + \frac{\partial}{\partial x_2} \underline{e}_2 + \cdots + \frac{\partial}{\partial x_n} \underline{e}_n;$$

$$\nabla^2 = D^1 \cdot D^1 = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \cdots + \frac{\partial^2}{\partial x_n^2}.$$

Let m be a non-negative integer; put

$$D^{-m}\tilde{H}^s(\mathbb{R}^n) = \{u \in \mathbb{D}'(\mathbb{R}^n) : D^\alpha u \in \tilde{H}^s(\mathbb{R}^n), \ |\alpha| = m\}. \tag{4.50}$$

Since a distribution whose derivatives are tempered is itself tempered,

$D^{-m}\tilde{H}^s(\mathbb{R}^n)$ is a linear sub-space of Φ'_n; it is given the semi-norm

$$\| u \|^2_{D^{-m}H^s(\mathbb{R}^n)} = \int_{\mathbb{R}^n} |y|^{2s} |(D^m u(y))\hat{}\,|^2 dm_n(y). \tag{4.51}$$

The nullspace of this semi-norm is $\mathbb{P}_{m-1}$, the space of polynomials on $\mathbb{R}^n$ of degree $m-1$, and less. The semi-norm (4.51) is translation invariant, and because of the way D^m is defined, also rotation invariant; further, should scale be changed from x to λx, $\lambda > 0$, and should $u(\lambda x) = v(x)$, then $\hat{u}(t) = \lambda^n \hat{v}(\lambda t)$, and (4.51) changes by a constant factor. These are attractive attributes.

The spaces $D^{-m}\tilde{H}^s(\mathbb{R}^n)$ and $D^{-m}\tilde{H}^s(\mathbb{R}^n)/\mathbb{P}_{m-1}$ (given its natural quotient norm) are called *Beppo-Levi spaces*.

$$D^{-m}\tilde{H}^s(\mathbb{R}^n)/\mathbb{P}_{m-1} \subset \Phi'_n/\mathbb{P}_{m-1},$$

and reasoning similar to that used on page 69 shows that the inclusion is continuous, and hence that $D^{-m}\tilde{H}^s(\mathbb{R}^n)/\mathbb{P}_{m-1}$ is a Hilbert sub-space of $\Phi'_n/\mathbb{P}_{m-1}$, or $D^{-m}\tilde{H}^s(\mathbb{R}^n)$ is a semi-Hilbert sub-space of Φ'_n.

The next significant result that requires to be proven, is that on bounded sets, $(D^{-m}\tilde{H}^s)(\Omega) = H^{m+s}(\Omega)$:

Theorem: *(Duchon, 1977, p 91)* Let Ω be a bounded domain in $\mathbb{R}^n$, and let $-m - \frac{n}{2} < s < \frac{n}{2}$, then

$$(D^{-m}\tilde{H}^s)(\Omega) = H^{m+s}(\Omega). \tag{4.52}$$

Proof: Let $u \in H^{m+s}(\Omega)$; u extends to an element Pu (prolongation of u) in $H^{m+s}(\mathbb{R}^n)$ with compact support. Recall (4.35), for α a multi-index of order m,

$$(D^{\alpha}Pu)\hat{} = (i)^{-|\alpha|}y^{\alpha}(Pu)\hat{}.$$

Pu compact means $(Pu)\hat{} \in C^{\infty}(\mathbb{R}^n)$; thus

$$\int_{\mathbb{R}^n}|y|^{2s}|(D^{\alpha}Pu(y))\hat{}|^2 dm_n(y) = \int_{\mathbb{R}^n}|y|^{2s}|y^{\alpha}(Pu(y))\hat{}|^2 dm_n(y)$$

$$\leq \int_{\mathbb{R}^n}|y|^{2s+2m}|(u(y))\hat{}|^2 dm_n(y), \quad (4.53)$$

since $|y^{\alpha}|^2 \leq |y|^{2m}$. Now, $|y|^{2s+2m}$ is locally integrable for $2s+2m > -n$ (or, $s > -m - \frac{n}{2}$), and

$$|y|^{2s+2m} < 2(1+|y|^2)^{m+s}$$

outside a large enough ball in $\mathbb{R}^n$. Thus from (4.53)

$$\int_{\mathbb{R}^n}|y|^{2s}|(D^{\alpha}Pu(y))\hat{}|^2 dm_n(y) < \infty$$

for all α, $|\alpha| = m$, and $Pu \in D^{-m}\tilde{H}^s(\mathbb{R}^n)$. Thus every element of $H^{m+s}(\Omega)$ extends to an element of $D^{-m}\tilde{H}^s(\mathbb{R}^n)$ from which the inference follows that $H^{m+s}(\Omega) \subset (D^{-m}\tilde{H}^s)(\Omega)$, the restrictions to Ω of elements of $D^{-m}\tilde{H}^s(\mathbb{R}^n)$. Conversely

$$(D^{-m}\tilde{H}^s)(\Omega) \subset D^{-m}(\tilde{H}^s(\Omega)) \subset D^{-m}(H^s(\Omega)),$$

since $\tilde{H}^s(\Omega) \subset H^s(\Omega)$ for $s < \frac{n}{2}$. On the other hand, it has been noted that $D^{-m}(H^s(\Omega)) = H^{m+s}(\Omega)$, and so the result follows.

Corollary: $D^{-m}\tilde{H}^s(\mathbb{R}^n)$ is a dense semi-Hilbert sub-space of $H^{m+s}_{loc}(\mathbb{R}^n)$.

Proof: It follows from the theorem that $D^{-m}\tilde{H}^s(\mathbb{R}^n) \subset H^{m+s}_{loc}(\mathbb{R}^n)$, and hence that

$$D^{-m}\tilde{H}^s(\mathbb{R}^n)/\mathbb{P}_{m-1} \subset H^{m+s}_{loc}(\mathbb{R}^n)/\mathbb{P}_{m-1},$$

and by application of the closed graph theorem (in the same way as on page 72) the inclusion can be seen to be continuous, and dense, since any continuous linear functional on $H^{m+s}_{loc}(\mathbb{R}^n)$ that annihilates $D^{-m}H^s(\mathbb{R}^n)$ also annihilates $H^{m+s}_{loc}(\mathbb{R}^n)$.

Finally, since $(D^{-m}\tilde{H}^s)(\Omega)/\mathbb{P}_{m-1}$ and $H^{m+s}(\Omega)/\mathbb{P}_{m-1}$ are equal, and since both are Hilbert sub-spaces of $\mathbb{D}'(\Omega)/\mathbb{P}_{m-1}$, there is a continuous bijection between these two sub-spaces, making them topologically isomorphic. Hence we can say that $H^{-m-s}_{\overline{\Omega}}(\mathbb{R}^n) \cap \mathbb{P}^\circ_{m-1}$ (which is the dual space of $H^{m+s}(\Omega)/\mathbb{P}_{m-1}$) is closed in $(D^{-m}\tilde{H}^s(\mathbb{R}^n)/\mathbb{P}_{m-1})'$ *(Duchon, 1977, p 91)*.

It is in the Beppo-Levi spaces introduced in this section that solutions will be sought to the spline interpolation and smoothing problems. A step that still needs to be taken is to construct the kernel mappings between these spaces and their duals.

4.7 Kernel Mappings on Beppo-Levi Spaces

In this section we will make use of the two theorems of Section 3.3 to solve the spline problem through an explicit construction of the kernel mapping. Throughout this section $-m-\frac{n}{2}<s<\frac{n}{2}$, so that we are able to draw upon the results of the last section. For the space X (in the theorems of Section 3.3) $D^{-m}\tilde{H}^{s}(\mathbb{R}^n)$ will be used, with semi-norm (4.51), and nullspace $\mathbb{P}_{m-1}$. As has just been seen, $D^{-m}\tilde{H}^{s}(\mathbb{R}^n)$ is a dense semi-Hilbert sub-space of $H^{m+s}_{loc}(\mathbb{R}^n)$, which is a Frechet space, and thus a complete locally convex topological vector space; its dual is $H^{-m-s}_{comp}(\mathbb{R}^n)$.

Theorem: *(Duchon, 1977, p 93).* Let Ω be a bounded domain in $\mathbb{R}^n$. Let M be a finite-dimensional (and therefore closed) linear sub-space of $H^{-m-s}_{\overline{\Omega}}(\mathbb{R}^n)$ such that, if $p\in\mathbb{P}_{m-1}$ and

$$<p,\mu>_{(H^{m+s}_{loc},H^{-m-s}_{comp})} = 0$$

for all $\mu\in M$, then $p=0$. Let $f\in H^{m+s}(\Omega)$. There exists a unique $f^M\in D^{-m}\tilde{H}^{s}(\mathbb{R}^n)$ satisfying

$$<f^M,\mu> = <f,\mu>$$

for all $\mu\in M$ with minimum semi-norm $\| f^M \|_{D^{-m}\tilde{H}^{s}(\mathbb{R}^n)}$. Moreover, let θ be the kernel associated with $D^{-m}\tilde{H}^{s}(\mathbb{R}^n)$; then the only g of the form $\theta\mu+p$ with $\mu\in M\cap\mathbb{P}^{\circ}_{m-1}$ and $p\in\mathbb{P}_{m-1}$ satisfying $<g,\mu>=<f,\mu>$ for $\mu\in M$, is precisely f^M.

Proof: This is a restatement of the theorems of Section 3.3, with

$X = D^{-m}\tilde{H}^s(\mathbb{R}^n)$, $E = H^{m+s}_{loc}(\mathbb{R}^n)$, and $N = \mathbb{P}_{m-1}$. M finite-dimensional means that $M \cap \mathbb{P}^\circ_{m-1}$ is closed in $(D^{-m}\tilde{H}^s(\mathbb{R}^n)/\mathbb{P}_{m-1})'$. M can be made finite-dimensional since this is the situation most likely in practice; however, the theorem still holds if M is not finite-dimensional, but closed in $H^{-m-s}_{\overline{\Omega}}(\mathbb{R}^n)$ *(Duchon, 1977, p 93)*.

The task that remains is to determine θ, the kernel associated with the semi-Hilbert sub-space $D^{-m}\tilde{H}^s(\mathbb{R}^n)$, alternatively the kernel L associated with $D^{-m}\tilde{H}^s(\mathbb{R}^n)/\mathbb{P}_{m-1}$ as a Hilbert sub-space of $H^{m+s}_{loc}(\mathbb{R}^n)/\mathbb{P}_{m-1}$, the positive continuous linear mapping

$$L: H^{-m-s}_{comp}(\mathbb{R}^n) \cap \mathbb{P}^\circ_{m-1} \to H^{m+s}_{loc}(\mathbb{R}^n)$$

satisfying (cf. (3.6))

$$(w, L\mu)_{D^{-m}\tilde{H}^s(\mathbb{R}^n)/\mathbb{P}_{m-1}} = \langle w, \mu \rangle$$

for $w \in D^{-m}\tilde{H}^s(\mathbb{R}^n)/\mathbb{P}_{m-1}$, $\mu \in H^{-m-s}_{comp}(\mathbb{R}^n) \cap \mathbb{P}^\circ_{m-1}$, and

$$\begin{aligned}(w, L\mu)_{D^{-m}\tilde{H}^s(\mathbb{R}^n)/\mathbb{P}_{m-1}} &= \int_{\mathbb{R}^n} |y|^{2s} (D^m L\mu)\hat{}(y) \cdot \overline{(D^m w)\hat{}(y)}\, dm_n(y) \\ &= \int_{\mathbb{R}^n} |y|^{2s} (y^m (L\mu)\hat{})(y) \cdot \overline{(y^m \hat{w})(y)}\, dm_n(y).\end{aligned} \tag{4.55}$$

In preparation for the theorem to come three lemmas are needed:

Lemma 1: $$|y|^{2m} \hat{u} = |y|^{-2s} \hat{\mu}, \tag{4.56}$$

where $u \in L\mu$ ($L\mu$ is an equivalence class modulo $\mathbb{P}_{m-1}$).

Proof: $s > -m - \frac{n}{2}$ implies $\Phi_n \subset D^{-m}\tilde{H}^s(\mathbb{R}^n)$, for $\phi \in \Phi_n$ means $\hat{\phi}$ is integrable, any polynomial times $\hat{\phi}$ is integrable (cf. Section 4.3), and $s > -m - \frac{n}{2}$ means $|y|^{2s+2m}$ is locally integrable, and thus that

$$\int_{\mathbb{R}^n} |y|^{2s+2m} \hat{\phi}(y) dm_n(y) < \infty.$$

Thus (4.55) may be applied with $w = \hat{\phi}$, $\phi \in \mathbb{D}(\mathbb{R}^n)$ $(\subset \Phi_n)$, and using (4.32),

$$\int_{\mathbb{R}^n} |y|^{2s} (y^m \hat{u})(y) \cdot \overline{(y^m \hat{\phi})(y)} dm_n(y) = \langle \hat{\phi}, \mu \rangle = \langle \phi, \hat{\mu} \rangle$$

for all $\phi \in \mathbb{D}(\mathbb{R}^n)$, and so

$$|y|^{2s} y^m \cdot (y^m \hat{u})(y) = \hat{\mu}(y) \text{ almost everywhere,}$$

i.e.

$$y^m \cdot (y^m \hat{u})(y) = |y|^{-2s} \hat{\mu}(y) \text{ almost everywhere,}$$

or

$$|y|^{2m} \hat{u} = |y|^{-2s} \hat{\mu}$$

as distributions.

Lemma 2: Let $\mu \in \mathbb{D}'(\mathbb{R}^n)$ have compact support, and suppose μ annihilates $\mathbb{P}_{m-1}$, then the distribution $y^\alpha \hat{\mu} \mathrm{Pf} |y|^{-2m-2s} \in \mathbb{L}^1_{loc}(\mathbb{R}^n)$, $|\alpha| = m$, i.e. is locally integrable.

Pf stands for *pseudo-function*, or 'partie finie', and refers to the manner in which $|y|^\lambda$, $\lambda \in \mathbb{C}$, has been regularised, to make it a regular (locally integrable) distribution *(Schwartz, 1950, p 38; Constantinescu,*

1980, p 69).

Proof: It is enough to show that the function $y^{\alpha}\hat{\mu}(y)|y|^{-2m-2s}$ is locally integrable. μ has compact support; it thus has a unique extension to a continuous linear functional on $C^{\infty}(\mathbb{R}^n)$. Further, $\hat{\mu} \in C^{\infty}(\mathbb{R}^n)$. Suppose $|\alpha| < m$, then

$$D^{\alpha}\hat{\mu}(x) = (i)^{-|\alpha|}(y^{\alpha}\mu)\hat{}(x) = (i)^{-|\alpha|}\int_{\mathbb{R}^n} y^{\alpha}\mu(y)e_{-x}\,dm_n(y).$$

At $x = 0$,

$$\mu y^{\alpha} = \int_{\mathbb{R}^n} y^{\alpha}\mu(y)\,dm_n(y)$$

expresses μ as a linear functional on $y^{\alpha} \in C^{\infty}(\mathbb{R}^n)$, and is zero by hypothesis. Thus $D^{\alpha}\hat{\mu}$ has vanishing derivatives of order $|\alpha| < m$ at 0, which in turn means that

$$|\hat{\mu}(y)| \leq c|y|^{m}$$

in a neighbourhood of the origin (c some constant), and that ($|\alpha| = m$ now)

$$|y^{\alpha}\hat{\mu}(y)||y|^{-2m-2s} \leq c|y|^{-2s}$$

in the same neighbourhood, and C^{∞} elsewhere. Since $s < \frac{n}{2}$ $(-2s > -n)$

$$y^{\alpha}\hat{\mu}(y)|y|^{-2m-2s} \in \mathbb{L}^1_{loc}(\mathbb{R}^n).$$

Lemma 3: Let $f \in \Phi_n'$, and suppose

$$|y|^{2m}\hat{f} = 0$$

and

$$y^{\alpha}\hat{f} \in \mathbb{L}^1_{loc}(\mathbb{R}^n), \quad |\alpha| = m,$$

then $f \in \mathbb{P}_{m-1}$.

Proof: $|y|^{2m}$ vanishes only for $y = 0$, so the only support $\hat{f}$ can have is $\{0\}$, and the only support $y^{\alpha}\hat{f}$ can have is also $\{0\}$, but $y^{\alpha}\hat{f}$ is locally integrable, so $y^{\alpha}\hat{f} = 0$, $D^{\alpha}f = 0$ for all $|\alpha| = m$, i.e. $f \in \mathbb{P}_{m-1}$.

We are now in a position to tackle the main result of this section:

Theorem: *(Duchon, 1977, p 94)* As before, $-m - \frac{n}{2} < s < \frac{n}{2}$.

$$\theta: H^{-m-s}_{comp}(\mathbb{R}^n) \cap \mathbb{P}^{\circ}_{m-1} \to D^{-m}\tilde{H}^{s}(\mathbb{R}^n);$$
$$\theta: \mu \mapsto \mu^*(Pf|y|^{-2m-2s})\hat{} \tag{4.57}$$

is the kernel of $D^{-m}\tilde{H}^{s}(\mathbb{R}^n)$ as a semi-Hilbert sub-space of $H^{m+s}_{loc}(\mathbb{R}^n)$.

Proof: Let $v = \mu^*(Pf|y|^{-2m-2s})\hat{}$, then $\hat{v} = \hat{\mu}Pf|y|^{-2m-2s}$, so

$$|y|^{2m}\hat{v} = |y|^{2m}\hat{\mu}Pf|y|^{-2m-2s} = \hat{\mu}|y|^{-2s},$$

since $|y|^{2m}Pf|y|^{-2m-2s} = |y|^{-2s}$. By Lemma 2, $y^{\alpha}\hat{v} \in \mathbb{L}^1_{loc}(\mathbb{R}^n)$ for all $|\alpha| = m$. On the other hand, if $u \in L\mu$, Lemma 1 shows (4.56), $|y|^{2m}\hat{u} = \hat{\mu}|y|^{-2s}$; thus $|y|^{2m}(\hat{u} - \hat{v}) = 0$ and $y^{\alpha}(\hat{u} - \hat{v}) \in \mathbb{L}^1_{loc}(\mathbb{R}^n)$ for $|\alpha| = m$. Thus

Lemma 3 shows $u - v \in \mathbb{P}_{m-1}$, i.e. $v \in L\mu$, or

$$L\mu = \mu^*(Pf|y|^{-2m-2s})\hat{} + \mathbb{P}_{m-1},$$

alternatively, that

$$\theta\mu = \mu^*(Pf|y|^{-2m-2s})\hat{}$$

is indeed the kernel associated with $D^{-m}\tilde{H}^s(\mathbb{R}^n)$ as semi-Hilbert sub-space of $H^{m+s}_{loc}(\mathbb{R}^n)$.

In consequence we need to know the Fourier transforms of pseudo-functions $Pf|y|^{\lambda}$. Three cases need to be distinguished *(Schwartz, 1951, p 114; Jones, 1982, pp 251 et seq.; Constantinescu, 1980, p 90):*

(i) For $\lambda = 2k$, $k \in \mathbb{Z}^+$:

$$(|x|^{2k})\hat{} = (-1)^k \Delta^k \delta. \tag{4.58}$$

(ii) For $\lambda = -n - 2k$, $k \in \mathbb{Z}^+$:

$$(Pf|x|^{-n-2k})\hat{} = c_1|\xi|^{2k}\log_e|\xi| + c_2|\xi|^{2k}, \tag{4.59}$$

where the constants c_1 and c_2 are given, for example, by *Jones (1982), p 530.*

(iii) For all other λ:

$$(Pf|x|^{\lambda})\hat{} = c_3 Pf|\xi|^{-n-\lambda}, \tag{4.60}$$

with

$$c_3 = 2^{\frac{n}{2}+\lambda}\,\frac{\Gamma(\frac{\lambda+n}{2})}{\Gamma(-\frac{\lambda}{2})}.$$

If $m+s>0$, the first case is excluded (and in a moment we shall go further and make $m+s>\frac{n}{2}$ in order to ensure $H^{m+s}_{loc}(\mathbb{R}^n)$ is a space of continuous functions). Thus, $0<m+s<m+\frac{n}{2}$,

$$(Pf|x|^{-2m-2s})\hat{} = \begin{cases} c_1|\xi|^{2m+2s-n}\log_e|\xi| + c_2|\xi|^{2m+2s-n} \\ \qquad \text{if } 2m+2s-n \text{ is even and non-negative;} \\ c_3|\xi|^{2m+2s-n} \qquad \text{otherwise.} \end{cases} \tag{4.61}$$

Further, if $\mu \in \mathbb{P}^{\circ}_{m-1}$, and $2m+2s-n$ is even and non-negative, then

$$\mu^*|\xi|^{2m+2s-n} \in \mathbb{P}_{m-1},$$

for $2m+2s-n<2m$, μ reduces $|\xi|^{2m+2s-n}$ by degree m, and what remains is in the nullspace of the semi-inner product (4.55).

So, by putting

$$K_{\lambda}(x) = \begin{cases} |x|^{\lambda}\log_e|x|, & \lambda \text{ even, non-negative,} \\ |x|^{\lambda} & \text{otherwise,} \end{cases} \tag{4.62}$$

the map

$$\mu \mapsto \mu^* K_{2m+2s-n}$$

is proportional to the kernel associated with $D^{-m}\tilde{H}^{s}(\mathbb{R}^n)$.

When $m+s>\frac{n}{2}$, $H^{m+s}_{loc}(\mathbb{R}^n)$ is a space of continuous functions, and more generally, if

$$m+\quad > \frac{n}{2}+k,$$

then

$$H^{m+s}_{loc}(\mathbb{R}^n) \subset C^k(\mathbb{R}^n)$$

by the Sobolev embedding theorem *(Lions and Magenes, 1972, p 45)*. By appropriate choice of m it is possible to find spline functions that are at least k times differentiable.

How to compute interpolating natural spline functions from given point data is made clear in the next section.

4.8 Interpolating Rotation-Invariant Splines on $\mathbb{R}^n$

A requirement of the fundamental spline theorem of the previous section is that, if p is a polynomial of degree at most $m-1$, and $\langle p,\mu\rangle = 0$ for all μ in the 'observational' sub-space M, then $p=0$. This is equivalent to saying that 0 is the only element common to both $\mathbb{P}_{m-1}$ and M°, the orthogonal in $H^{m+s}_{loc}(\mathbb{R}^n)$ of M. This is the 'natural end condition' seen earlier, for the unique solution of the spline interpolation problem.

$\mathbb{P}_{m-1}$ is a linear vector space of dimension $q=\binom{m+n-1}{n}$. Let $\{p_j\}_{j=1,..q}$ be a basis for $\mathbb{P}_{m-1}$, and let $\{\mu_i\}_{i=1,\ldots,N}$ be linearly independent observation functionals in $H^{-m-s}_{\bar{\Omega}}(\mathbb{R}^n)$, Ω bounded. M is the span of $\{\mu_i\}$. The condition $\langle p,\mu\rangle = 0$ for all $\mu \in M$ means $p=0$, becomes

$$\text{(column) rank} \begin{pmatrix} \langle p_1,\mu_1\rangle & \cdots & \langle p_q,\mu_1\rangle \\ \vdots & & \\ \langle p_1,\mu_N\rangle & \cdots & \langle p_q,\mu_N\rangle \end{pmatrix}_{q,N} = q = \dim \mathbb{P}_{m-1},$$

which is to say that $\{\mu_i\}_{i=1,\ldots,N}$ are $\mathbb{P}_{m-1}$-*unisolvent*. The linear independence of the observation functionals μ_i means that the matrix

$$\begin{pmatrix} \langle \mu_1 * K_\lambda,\mu_1\rangle & \cdots & \langle \mu_N * K_\lambda,\mu_1\rangle \\ \vdots & & \\ \langle \mu_1 * K_\lambda,\mu_N\rangle & \cdots & \langle \mu_N * K_\lambda,\mu_N\rangle \end{pmatrix}_{N,N}$$

has rank N.

Theorem: Let $\{\mu_i\}_{i=1,\ldots,N}$ be a set of $\mathbb{P}_{m-1}$-unisolvent linearly independent bounded linear functionals in $H^{-m-s}_{comp}(\mathbb{R}^n)$; let $f \in H^{m+s}(\Omega)$, then there is a unique $s \in D^{-m}\tilde{H}^s(\mathbb{R}^n)$ of the form

$$s = \nu * K_{2m+2s-n} + p, \tag{4.63}$$

where $\nu \in M \cap \mathbb{P}^\circ_{m-1}$, $M = \text{span}(\{\mu_i\})$, $p \in \mathbb{P}_{m-1}$, and $K_{2m+2s-n}$ is given by (4.61), satisfying

$$\langle s,\mu_i\rangle = \langle f,\mu_i\rangle, \quad i = 1,\ldots,N, \tag{4.64}$$

with minimum semi-norm $\| s \|_{D^{-m}\tilde{H}^s(\mathbb{R}^n)}$.

Proof: This is a restatement of the theorem with which the previous section started out, with the kernel associated with $D^{-m}\tilde{H}^{s}(\mathbb{R}^n)$ given its explicit form.

For brevity, put $\lambda = 2m+2s-n$; the requirements of the theorem put two conditions on

$$s = \sum_{i=1}^{N} a_i \mu_i * K_\lambda + \sum_{j=1}^{q} b_j p_j; \tag{4.65}$$

firstly, that s interpolates f at μ_i:

$$a_1 \langle \mu_1 * K_\lambda, \mu_i \rangle + \ldots + a_N \langle \mu_N * K_\lambda, \mu_i \rangle + b_1 \langle p_1, \mu_i \rangle + \ldots + b_q \langle p_q, \mu_i \rangle = \langle f, \mu_i \rangle, \tag{4.66}$$

for $i = 1, \ldots, N$, and secondly, that s satisfies the end conditions for a natural spline, that ν be in the orthogonal to $\mathbb{P}_{m-1}$, namely that

$$a_1 \langle p_j, \mu_1 \rangle + \ldots + a_N \langle p_j, \mu_N \rangle = 0, \quad j = 1, \ldots, q \tag{4.67}$$

or, expressing (4.66) and (4.67) in matrix notation,

$$\begin{aligned} Aa + B^t b &= c, \\ Ba &= 0, \end{aligned} \tag{4.68}$$

where

$$A = \begin{pmatrix} \langle \mu_1 * K_\lambda, \mu_1 \rangle & \ldots & \langle \mu_N * K_\lambda, \mu_1 \rangle \\ \vdots & & \\ \langle \mu_1 * K_\lambda, \mu_N \rangle & \ldots & \langle \mu_N * K_\lambda, \mu_N \rangle \end{pmatrix}_{N,N}$$

$$B = \begin{pmatrix} \langle p_1, \mu_1 \rangle & \cdots & \langle p_1, \mu_N \rangle \\ \vdots & & \\ \langle p_q, \mu_1 \rangle & \cdots & \langle p_q, \mu_N \rangle \end{pmatrix}_{q,N}$$

$$a = (a_1, \ldots, a_N)^t \in \mathbb{R}^N,$$
$$b = (b_1, \ldots, b_q)^t \in \mathbb{R}^q,$$
$$c = (c_1, \ldots, c_N)^t \in \mathbb{R}^N,$$

and

$$c_i = \langle f, \mu_i \rangle_{(H_{loc}^{m+s}, H_{comp}^{-m-s})}.$$

Since the matrices A and B have full rank, the equation system (4.68) can be solved

$$\begin{aligned} a &= A^{-1}c - A^{-1}B^t b, \\ b &= (BA^{-1}B^t)^{-1}BA^{-1}c. \end{aligned} \tag{4.69}$$

In the special case where μ_i are evaluation functionals, i.e. $\mu_i = \delta_{\alpha_i}$, $\alpha_i \in \mathbb{R}^n$, $i = 1, \ldots, N$,

$$\delta_{\alpha_i} * K_\lambda = \tau_{\alpha_i} K_\lambda,$$

or

$$\delta_{\alpha_i} * K_\lambda(x) = K_\lambda(x - \alpha_i), \tag{4.70}$$

and the solution of the spline interpolation problem is even simpler. The N points α_i need only be distinct for δ_{α_i} to be linearly independent, and need not to lie in a hyperplane of $\mathbb{P}_{m-1}$ to be $\mathbb{P}_{m-1}$-unisolvent.

Theorem: Let $(c_1,\dots,c_N)^t \in \mathbb{R}^N$, and $\lambda = 2m+2s-n$. Then there is only one function of the form

$$s(x) = \sum_{i=1}^{N} a_i K_\lambda(x-\alpha_i) + p(x), \tag{4.71}$$

with $p \in \mathbb{P}_{m-1}$, and

$$\sum_{j=1}^{N} a_j q(\alpha_j) = 0$$

for all $q \in \mathbb{P}_{m-1}$, and satisfying

$$s(\alpha_i) = c_i, \quad i = 1,\dots,N.$$

Moreover, if f is another function in $D^{-m}\tilde{H}^s(\mathbb{R}^n)$ interpolating c, then

$$\| s \|_{D^{-m}\tilde{H}^s(\mathbb{R}^n)} \leq \| f \|_{D^{-m}\tilde{H}^s(\mathbb{R}^n)}.$$

4.9 Examples of Splines on $\mathbb{R}^n$

The long preparation of the last sections now pays off - several concrete examples can now be given of rotation-invariant natural splines on $\mathbb{R}^n$. Throughout this section $\{\alpha_i\}_{i=1,\dots,N}$ will be a set of distinct $\mathbb{P}_{m-1}$-unisolvent points in $\mathbb{R}^n$, and $\{c_i\}_{i=1,\dots,N}$ an arbitrary vector in $\mathbb{R}^N$. Except when discussing point mass splines, the condition $-m-\frac{n}{2} < s < \frac{n}{2}$ will hold.

4.9.1 Thin-plate splines: Suppose $n = 2$; for $s = 0$ the kernel function (4.61) is logarithmic. Thus for $m = 2$ spline functions (4.71) take on

the form

$$s(x) = \sum_{i=1}^{N} a_i |x-\alpha_i|^2 \log_e |x-\alpha_i| + b_0 + b\cdot x, \qquad (4.72)$$

where $b, x, \alpha_i \in \mathbb{R}^2$, with

$$\sum_{i=1}^{N} a_i = 0 \quad \text{and} \quad \sum_{i=1}^{N} a_i \alpha_i = 0 \quad (\text{in } \mathbb{R}^2).$$

The semi-norm minimised is

$$\begin{aligned} \| f \|^2_{D^2\tilde{H}^0(\mathbb{R}^2)} &= \int_{\mathbb{R}^2} |(D^2 f(y))\hat{}|^2 dm_n(y) \\ &= \int_{\mathbb{R}^2} |D^2 f(y)|^2 dm_n(y) \\ &= \int_{\mathbb{R}^n} \{(\frac{\partial^2}{\partial y_1^2} f(y))^2 + 2(\frac{\partial^2}{\partial y_1 \partial y_2} f(y))^2 + (\frac{\partial^2}{\partial y_2^2} f(y))^2\} dm_n(y). \end{aligned} \qquad (4.73)$$

The Laplacian of (4.72) is

$$\nabla^2 s(x) = \sum_{i=1}^{N} 4a_i(1 + \log_e |x-\alpha_i|),$$

which is harmonic (except at α_i), and $s(x)$ is therefore biharmonic; it is a fundamental solution of

$$(D^{4\cdot 0} + 2D^{2\cdot 2} + D^{0\cdot 4})e = \delta.$$

Interpolating thin-plate splines can be interpreted as giving the shapes of infinite elastic membranes deformed by vertical point loads (or supports) at α_i, $i = 1, \ldots, N$, constraining the membrane to pass through

(α_i, c_i) in $\mathbb{R}^3$. (4.73) can be interpreted as the bending energy of the membrane. Thin-plate splines are thus $\mathbb{R}^2$ analogues of cubic splines interpreted as the shapes of physical splines (thin elastic beams) constrained to pass through given points in $\mathbb{R}^2$. Thin-plate splines thus have an attractive physical meaning, are easy to compute, but have not been widely applied to geodetic problems; one application has been by *Sandwell (1987)*, to interpolate satellite altimeter geoid heights.

4.9.2 Polynomial splines: When $s = \frac{n-1}{2}$, $\lambda = 2m + 2s - n = 2m - 1$, and there is precisely one function of the form

$$s(x) = \sum_{i=1}^{N} a_i |x - \alpha_i|^{2m-1} + p(x) \qquad (4.74)$$

satisfying $\sum_{i=1}^{N} a_i q(\alpha_i) = 0$ for all $q \in \mathbb{P}_{m-1}$ and interpolating c. Any other function interpolating c has greater norm. *Duchon (1977)* calls these *pseudo-polynomial splines;* they are generalisations to $\mathbb{R}^n$ of classical polynomial splines on $\mathbb{R}$.

For $m = 1$ the splines take on the form

$$s(x) = \sum_{i=1}^{N} a_i |x - \alpha_i| + b, \qquad (4.75)$$

where b is a constant, and $\sum_{i=1}^{N} a_i = 0$. $s \in C(\mathbb{R}^n)$, and minimises

$$\|f\|^2_{D^{-1}\tilde{H}^{\frac{n-1}{2}}(\mathbb{R}^n)} = \int_{\mathbb{R}^n} |y|^{n-1} |(Df(y))\hat{}|^2 dm_n(y).$$

The functions (4.75) are called *multi-conic functions*. Similar functions have found wide acceptance in geodetic and surveying

applications, and we turn to these in a moment, but note here, that for $n = 3$

$$\nabla^2 s(x) = \sum_{i=1}^{N} a_i \frac{1}{|x - \alpha_i|}, \tag{4.76}$$

and hence that s is biharmonic (except at α_i).

For $n = 1$, s is a piecewise linear function, constant for $x < \min_i \alpha_i$, and $x > \max_i \alpha_i$. The semi-norm minimised is

$$\| f \|^2_{D^{-1}\tilde{H}^0(\mathbb{R})} = \int_{\mathbb{R}} \left| \frac{\partial}{\partial y} f(y) \right|^2 dm_1(y), \tag{4.77}$$

so that s is the continuous function, of least overall slope, that interpolates the given data.

For $m = 2$, interpolating splines are of the form

$$s(x) = \sum_{i=1}^{N} a_i |x - \alpha_i|^3 + b_0 + \sum_{j=1}^{n} b_j x_j \tag{4.78}$$

(with $x = (x_1, \ldots, x_n)^t$), subject to

$$\sum_{i=1}^{N} a_i = 0, \quad \text{and} \quad \sum_{i=1}^{N} a_i \alpha_i = 0$$

(with the points $\{\alpha_i\}_{i=1,\ldots,N}$ not confined to a hyperplane in $\mathbb{R}^n$). (4.78) represents *pseudo-cubic splines;* the semi-norm minimised is

$$\| f \|^2_{D^{-2}\tilde{H}^{\frac{n-1}{2}}(\mathbb{R}^n)} = \int_{\mathbb{R}^n} |y|^{n-1} |(D^2 f(y))\hat{}|^2 dm_n(y). \tag{4.79}$$

For $n = 1$ we get the familiar *cubic spline*

$$s(x) = \sum_{i=1}^{N} a_i |x - \alpha_i|^3 + b_0 + b_1 x. \tag{4.80}$$

s is a polynomial of degree 3 on the intervals between successive 'knots' α_i, is twice differentiable at the knots, and, in consequence of the natural end conditions, is a polynomial of degree 1 for $x \leq \min_i \alpha_i$ and $x > \max_i \alpha_i$. The semi-norm minimised is

$$\| f \|^2_{D^2 \tilde{H}^0(\mathbb{R}^n)} = \int_{\mathbb{R}} \left| \frac{\partial^2}{\partial y^2} f(y) \right|^2 dm_1(y), \tag{4.81}$$

so that the cubic spline can be interpreted, loosely speaking, as minimising curvature overall.

Polynomial splines in one dimension have long been popular devices for interpolation, smoothing and approximation of linear functionals - in geodetic contexts as well, for example, in representing empirical covariance functions. A systematic development of polynomial splines *sui generis* may be found in *De Boor and Lynch (1966)*, or *Böhmer (1974)*.

It has been noted that for $s = \frac{n-1}{2}$ and $m = 2$, $H^{m+s}_{loc}(\mathbb{R}^n) \subset C^1(\mathbb{R}^n)$, i.e. that (4.78) has continuous derivatives. It is thus possible to construct interpolating pseudo-cubic splines subject to *Hermite conditions*, i.e. satisfying prescribed point values, and prescribed gradients at these points. In one dimension this corresponds to piecewise cubic Hermite interpolation (described, for example, by *Böhmer (1974), p 111)*. Thus for $n = 2$, $m = 2$ and $s = \frac{1}{2}$, $2m + 2s - n = 3$,

$$D^{\alpha}(\delta_x * |y|^3) = 3(x-y)^{\alpha}|x-y|$$

for $|\alpha| = 1$, i.e. $\alpha = (1,0)$ or $(0,1)$, and the Hermite interpolating pseudo-cubic spline takes the form

$$s(x) = \sum_{i=1}^{N} a_i |x-\alpha_i|^3 + b_0 + \sum_{j=1}^{n} b_j x_j + 3\sum_{i=1}^{N} c_k (x-\alpha_k)|x-\alpha_k|, \quad (4.82)$$

subject to the natural conditions

$$\sum_{i=1}^{N} a_i = 0, \quad \text{and} \quad \sum_{i=1}^{N} a_i \alpha_i + \sum_{k=1}^{N} c_k = 0 \ (\text{in } \mathbb{R}^2).$$

A geodetic application for splines subject to gradient interpolatory conditions is the construction of an astro-geodetic geoid fitting irregularly spaced astro-geodetic deflections of the vertical, or, of greater topical interest, in the construction of a geoid interpolating both satellite-derived geoid heights, and astro-geodetic deflections. Chapter 6 sees such an application, to the calculation of a combined Doppler and astro-geodetic geoid for southern Africa.

4.9.3 Hardy's multiquadric functions: The multi-conic splines (4.75) are special cases of *Hardy's multiquadric functions*

$$s(x) = \sum_{i=1}^{N} a_i \{|x-\alpha_i|^2 + \varepsilon^2\}^{1/2}, \quad (4.83)$$

where ε^2 is an empirically derived smoothing parameter, whose purpose is to achieve a regularisation of the discontinuities multi-conic splines have in their derivatives at the observation points, or 'knots'.

The question of the regularisation of the splines found in this chapter will be taken up briefly in Section 4.12, but this may be the appropriate place to say a few more words about Hardy's functions.

Since R.L. Hardy introduced his multiquadric technique in 1971 *(Hardy, 1971)*, it has found wide acceptance for tasks requiring surface-fitting to irregularly spaced data, both in geodetic applications (for example, in modelling the geopotential, *Hardy and Nelson (1986)*, and in the computation of gravimetric terrain corrections, *Krohn (1976)*), and elsewhere (for example, here in South Africa, for rainfall analysis, *Adamson (1978)*). The efficacy of multiquadric interpolation *vis a vis* other techniques has been convincingly demonstrated by comparative numerical investigations of the type conducted by *Hein and Lenze (1979)*. Hardy supported the introduction of multiquadric functions on heuristic grounds. We have seen that in the simplest case, of multi-conic functions, the technique is well founded, and corresponds to a multivariate generalisation of the simple method of piecewise linear interpolation. It has only recently been established that the multiquadric technique in general is also well posed *(Micchelli, 1986)*.

Hardy (1977) investigated the relation between his multiquadric kernel function and the covariance functions used in the theory of stochastic processes. This is an appropriate thing to do, for covariance functions are reproducing kernel functions, and minimum variance unbiased prediction yields spline functions - *Kimeldorf and Wahba (1970)* point to the close connexion between spline theory and the theory of stochastic processes. Hardy came to the conclusion that the multi-conic kernel

function - the distance function $|x|$ in $\mathbb{R}^n$ - is not a covariance function, i.e. not a reproducing kernel function. Indeed it is not; it is a kernel mapping (a linear map between a dual space E' and a topological vector space E, in a sense made precise in Section 3.2), rather than a reproducing kernel function (a positive bilinear form on E x E). It is not difficult to obtain a reproducing kernel function for multi-conic splines; for example, for n = 2, i.e. for $D^{-1}\tilde{H}^{1/2}(\mathbb{R}^2)/\mathbb{P}_0$. Anticipating a result of Section 4.10, it is

$$K_\alpha(x,y) = -\{|x-y| - |x-\alpha| - |y-\alpha|\}, \qquad (4.92)$$

$x,y \in \mathbb{R}^2$, α any 'node' or 'knot' in $\{\alpha_i\}_{i=1,\ldots,N}$.

4.9.4 Point-mass splines: Here we choose n = 3, s = 1 and m = 0, so that $m+s > -\frac{n}{2}$, but $m+s \not> \frac{n}{2}$. $\lambda = 2m+2s-n = -1$.

$$K_\lambda(x) = \frac{1}{|x|}, \quad x \in \mathbb{R}^3,$$

is a fundamental solution of $\Delta e = \delta$, and $\delta_\alpha * \frac{1}{|x|}$ is the Newtonian potential due to a point mass placed at $\alpha \in \mathbb{R}^3$; with $\mu \in H^{-1}_{comp}(\mathbb{R}^3)$ a finite mass distribution, $\mu * \frac{1}{|x|}$ is the corresponding potential (cf. *Zadro, 1984*). $\tilde{H}^1(\mathbb{R}^3)$ is the space of *Newtonian potentials of finite energy* described by *Schwartz (1964), p 233*, with norm

$$\| f \|^2_{\tilde{H}^1(\mathbb{R}^3)} = \int_{\mathbb{R}^3} |y|^2 |f(y)\hat{}|^2 dm_n(y),$$

alternatively, with norm

$$\| f \|^2_{D^{-1}\tilde{H}^0(\mathbb{R}^3)} = \int_{\mathbb{R}^3} |D^1 f(y)|^2 dm_n(y)$$

(bearing in mind that a constant harmonic function is zero), so that point-mass splines may be interpreted as minimising overall, the gradient of the potential. But $\delta * \frac{1}{|x|} \notin \tilde{H}^1(\mathbb{R}^3)$, for

$$|y|^2 |(\delta * \frac{1}{|y|})\hat{}|^2 = \frac{1}{|y|^2}$$

is not integrable over $\mathbb{R}^3$. Thus point evaluation functionals (and evaluation of derivatives) cannot be admitted as observations in the construction of splines minimising the norm $\| \cdot \|_{\tilde{H}^1(\mathbb{R}^3)}$: the mass points cannot coincide with the positions of point observations, and some form of regularisation is necessary, for example, by 'burying' the mass points or 'knots'. One possible form of regularisation is introduced in Section 4.12; another possibility is to replace the kernel mapping

$$\delta_\alpha \mapsto \delta_\alpha * \frac{1}{|x|} = \frac{1}{|x - \alpha|}$$

by the mapping

$$\delta_\alpha \mapsto \delta_{k\alpha} * \frac{1}{|x|} = \frac{1}{|x - k\alpha|},$$

where the constant k, $0 < k < 1$, is an appropriately chosen 'shrinking' or regularising parameter. (One could conceivably go even further, and partition the observations among a number of k_i taken from the interval $]0,1[$ and chosen to correspond with known levels of density contrasts within the earth). Point-mass representations derived from measurement

of potential will be of the form

$$s(x) = \sum_{i=1}^{N} a_i \frac{1}{|x - k\alpha_i|}, \tag{4.84}$$

and representations of the disturbing potential as derived from gravity anomalies will be of the form

$$s(x) = \sum_{i=1}^{N} a_i \{ \frac{|x|^2 - x \cdot k\alpha_i}{|x||x - k\alpha_i|^3} - \frac{2}{|x||x - k\alpha_i|} \} \tag{4.85}$$

(where the usual spherical approximation has been made in dealing with the disturbing potential, cf. *Reilly and Hebrechtsmeier, 1978).* For the solution of (4.84) or (4.85) it is enough that the observation points α_i, $i = 1,\ldots,N$, be distinct in each case (cf. *Stromeyer and Ballani (1984)* on this point).

There is a close relation between spherical harmonics and point masses; the relation

$$\frac{1}{|x - y|} = \frac{1}{|x|} \sum_{n=0}^{\infty} (\frac{|y|}{|x|})^n P_n(\frac{x \cdot y}{|x||y|}),$$

for $x,y \in \mathbb{R}^3$, and with $P_n(t)$ Legendre polynomials, carries with it the suggestion that the following chapter is an appropriate place to continue the discussion on point-mass splines.

In $\mathbb{R}^2$ the logarithmic potential plays the same role as the potential of a point mass in $\mathbb{R}^3$.

4.9.5 Summary: The table that follows summarises the spline functions

encountered in the the last few pages.

Table 4.1: Summary of the more significant spline functions on $\mathbb{R}^n$, for $n = 1,2,3$.

Name	m	s	$Pf\|x\|^{-2m-2s}$	K_λ	sub-space	continuity
$\mathbb{R}^1$: linear	1	0	$\|x\|^{-2}$	$\|\xi\|$	$D^{-1}\tilde{H}^0$	C^0
cubic	2	0	$\|x\|^{-4}$	$\|\xi\|^3$	$D^{-2}\tilde{H}^0$	C^1
quintic	3	0	$\|x\|^{-6}$	$\|\xi\|^5$	$D^{-3}\tilde{H}^0$	C^2
$\mathbb{R}^2$: logarithmic	0	1	$\|x\|^{-2}$	$\log\|\xi\|$	$\tilde{H}^1$	-
thin-plate	2	0	$\|x\|^{-4}$	$\|\xi\|^2\log\|\xi\|$	$D^{-2}\tilde{H}^0$	C^0
multi-conic	1	$\frac{1}{2}$	$\|x\|^{-3}$	$\|\xi\|$	$D^{-1}\tilde{H}^{\frac{1}{2}}$	C^0
pseudo-cubic	2	$\frac{1}{2}$	$\|x\|^{-5}$	$\|\xi\|^3$	$D^{-2}\tilde{H}^{\frac{1}{2}}$	C^1
pseudo-quintic	3	$\frac{1}{2}$	$\|x\|^{-7}$	$\|\xi\|^5$	$D^{-3}\tilde{H}^{\frac{1}{2}}$	C^2
$\mathbb{R}^3$: point-mass	0	1	$\|x\|^{-2}$	$\|\xi\|^{-1}$	$\tilde{H}^1$	-
multi-conic	2	0	$\|x\|^{-4}$	$\|\xi\|$	$D^{-2}\tilde{H}^0$	C^0
pseudo-cubic	2	1	$\|x\|^{-6}$	$\|\xi\|^3$	$D^{-2}\tilde{H}^1$	C^1
pseudo-quintic	3	1	$\|x\|^{-8}$	$\|\xi\|^5$	$D^{-3}\tilde{H}^1$	C^2

4.10 Reproducing Kernels

It was noted that multi-conic splines in $\mathbb{R}^3$, and thin-plate splines in $\mathbb{R}^2$, were biharmonic (except at the knots). More generally, for $k \in \mathbb{Z}^+$,

$$e(x) = \begin{cases} (-1)^k c_1 |x|^{2k-n} \log_e |x|, & k \geq \frac{n}{2} \text{ and } n \text{ even}; \\ (-1)^k c_3 |x|^{2k-n}, & \text{otherwise} \end{cases} \tag{4.86}$$

$(x \neq 0)$ is the fundamental solution of the k-iterated distributional Laplacian Δ^k in $\mathbb{R}^n$, i.e. $\Delta^k e = \delta$ *(Meinguet, 1983)*. *Meinguet (1979c, 1983)* uses (4.86) as the starting point for the development of the theory of interpolating spline functions, which he calls *surface splines*, and in which reproducing kernel functions play an important part, in contrast to the path followed by *Duchon (1977)*, through a direct construction of the kernel mapping using Fourier transforms and convolution. The contents of this section tie in with the work of *Meinguet (1979a, 1979b, 1979c, 1979d, 1981, 1983)*, and also with *Schempp and Tippenhauer (1974)*.

In requiring the observation functionals $\{\mu_i\}_{i=1,\ldots,N}$ to be $\mathbb{P}_{m-1}$-unisolvent, we require that a subset of size $q = \binom{m+n-1}{n}$ can be chosen - call these $\{\mu_i\}_{i=1,\ldots,q}$ - such that

$$\operatorname{rank} \begin{pmatrix} \langle p_1, \mu_1 \rangle & \cdots & \langle p_q, \mu_1 \rangle \\ \vdots & & \\ \langle p_1, \mu_q \rangle & \cdots & \langle p_q, \mu_q \rangle \end{pmatrix} = q = \dim \mathbb{P}_{m-1},$$

where $\{p_i\}_{i=1,\ldots,q}$ form a basis for $\mathbb{P}_{m-1}$, i.e. it is possible to find a unique interpolant in $\mathbb{P}_{m-1}$ to the functionals $\{\mu_i\}_{i=1,\ldots,q}$. Thus in particular, there is a unique $\tilde{p}_j \in \mathbb{P}_{m-1}$ solving the interpolation problem

$$\langle \tilde{p}_j, \mu_i \rangle_{(H^{m+s}_{loc}, H^{-m-s}_{comp})} = \delta_{ij}, \quad i,j = 1,\ldots,q \tag{4.87}$$

where δ_{ij} is the Kronecker symbol. $\{\tilde{p}_j\}_{j=1,\dots,q}$ are linearly independent and form an orthonormal dual basis to $\{\mu_i\}_{i=1,\dots,q}$. Let $f \in H^{m+s}(\Omega)$, Ω bounded. Its $\mathbb{P}_{m-1}$-interpolant on $\{\mu_i\}_{i=1,\dots,q}$ is given by the interpolation formula (in Lagrange form)

$$p_f = \sum_{i=1}^{q} \langle f, \mu_i \rangle \tilde{p}_i. \tag{4.88}$$

The spline solution is the unique element in $D^{-m}\tilde{H}^s(\mathbb{R}^n)$, interpolating f on M with minimum norm, and is of the form $\theta\mu + p$, $\mu \in M \cap \mathbb{P}^{\circ}_{m-1}$ and $p \in \mathbb{P}_{m-1}$.

$$D^{-m}\tilde{H}^s(\mathbb{R}^n) = D^{-m}\tilde{H}^s(\mathbb{R}^n)/\mathbb{P}_{m-1} + \mathbb{P}_{m-1},$$

where the first space is a Hilbert space, and a Hilbert sub-space of $H^{m+s}_{loc}(\mathbb{R}^n)$. For $m+s > \frac{n}{2}$ it is a space of continuous functions, and as was shown in Section 3.4, a reproducing kernel function can be associated with $D^{-m}\tilde{H}^s(\mathbb{R}^n)$. We will now work towards an explicit identification of the reproducing kernel.

Lemma: The map

$$\theta_p : H_{\overline{\Omega}}^{-m-s}(\mathbb{R}^n) \to \mathbb{P}_{m-1}$$

through

$$\theta_p : \mu \mapsto \sum_{i=1}^{q} \langle p_i, \mu \rangle \tilde{p}_i \tag{4.89}$$

is a kernel mapping for $\mathbb{P}_{m-1}$ (with inner product deriving from the dual orthonormal basis $\{\tilde{p}_i\}_{i=1,\dots,q}$) as a Hilbert sub-space of $H^{m+s}_{loc}(\mathbb{R}^n)$.

Proof: Indeed

$$\theta_p : \mu_i \mapsto \tilde{p}_i, \quad i = 1, \ldots, q,$$

and

$$(\tilde{p}_i, \tilde{p}_j)_{\mathbb{P}_{m-1}} = (\theta_p \mu_i, \tilde{p}_j)_{\mathbb{P}_{m-1}} = \langle \tilde{p}_j, \mu_i \rangle = \delta_{ij}$$

(which also confirms that the Lagrange polynomials $\tilde{p}_i$ do form an orthonormal basis for $\mathbb{P}_{m-1}$). Let $\mu \in H^{-m-s}_{comp}(\mathbb{R}^n)$ and $p \in \mathbb{P}_{m-1}$, then

$$\begin{aligned}(\theta_p \mu, p)_{\mathbb{P}_{m-1}} &= \Big(\sum_{i=1}^{q} \langle \tilde{p}_i, \mu \rangle \tilde{p}_i, \sum_{j=1}^{q} \langle p, \mu_i \rangle \tilde{p}_j \Big) \\ &= \sum_{i=1}^{q} \langle \tilde{p}_i, \mu \rangle \langle p, \mu_i \rangle \\ &= \Big\langle \sum_{i=1}^{q} \langle p, \mu_i \rangle \tilde{p}_i, \mu \Big\rangle \\ &= \langle p, \mu \rangle_{(H^{m+s}_{loc}, H^{-m-s}_{comp})},\end{aligned}$$

so that θ_p is indeed a kernel for $\mathbb{P}_{m-1}$.

In consequence $\mu \mapsto \mu - \sum_{i=1}^{q} \langle \tilde{p}_i, \mu \rangle \mu_i$ is a projection of $H^{-m-s}_{comp}(\mathbb{R}^n)$ onto $\mathbb{P}^{\circ}_{m-1}$, and we have

Theorem: For $\mu \in H^{-m-s}_{comp}(\mathbb{R}^n)$ define

$$\theta_0 = (\theta - \theta_p)\mu, \tag{4.90}$$

where θ is the kernel given by (4.57), then θ_0 is a kernel for

$D^{-m}\tilde{H}^s(\mathbb{R}^n)/\mathbb{P}_{m-1}$. Thus the map ($\lambda = 2m + 2s - n$)

$$\mu \mapsto c\mu * K_\lambda - c \sum_{i=1}^{q} \langle \tilde{p}_i, \mu \rangle \mu_i * K_\lambda,$$

with c the appropriate constant from (4.61), is a kernel for the Hilbert sub-space $D^{-m}\tilde{H}^s(\mathbb{R}^n)/\mathbb{P}_{m-1}$. Associated with it is a reproducing kernel function; if $\mu_i = \delta_{\alpha_i}$

$$K_0(x,y) = \langle \theta_0 \delta_y, \delta_x \rangle$$

$$= \langle cK_\lambda(\cdot - y) - c \sum_{i=1}^{q} \tilde{p}_i(y) K_\lambda(\cdot - \alpha_i), \delta_x - \sum_{j=1}^{q} \tilde{p}_j(x) \delta_{\alpha_i} \rangle$$

$$= c\{K_\lambda(x-y) - \sum_{j=1}^{q} \tilde{p}_j(y) K_\lambda(x - \alpha_j) - \sum_{i=1}^{q} \tilde{p}_i(x) K_\lambda(\alpha_i - y)$$

$$+ \sum_{i=1}^{q} \sum_{j=1}^{q} \tilde{p}_i(x) \tilde{p}_j(y) K_\lambda(\alpha_i - \alpha_j)\}. \qquad (4.91)$$

Example: A reproducing kernel for $D^{-1}\tilde{H}^{1/2}(\mathbb{R}^2)/\mathbb{P}_0$, the space of multiconic functions on $\mathbb{R}^2$: $\mathbb{P}_0$ contains only constant functions; the Lagrange basis is the constant function $p_1 = 1$. By (4.60)

$$c = -2^{-2} \frac{\Gamma(-\frac{1}{2})}{\Gamma(\frac{3}{2})} = -1.$$

Thus from (4.91)

$$K_0(x,y) = -\{|x-y| - |x-\alpha_1| - |\alpha_1 - y|\} \qquad (4.92)$$

is a reproducing kernel for $D^{-1}\tilde{H}^{1/2}(\mathbb{R}^n)/\mathbb{P}_0$. We confirm, that for

$f \in D^{-1}\tilde{H}^{1/2}(\mathbb{R}^2)$

$$f(x) = f(\alpha_1) + (K_0(x,\cdot), f(\cdot))_{D^{-1}\tilde{H}^{1/2}(\mathbb{R}^2)}.$$

$$(K_0(x,y), f(y)) = \int_{\mathbb{R}^2} |y| (D^1 K_0(x,y)\hat{}(D^1 f(y))\hat{}\, dm_2(y)$$

$$= \int_{\mathbb{R}^2} |y|^3 \hat{K}_0(x,y)\hat{f}(y) dm_2(y).$$

Now

$$|x-y|\hat{} = (\tau_{-x}|y|)\hat{} = e_x|y|\hat{} = -e_x \frac{1}{|y|^3},$$

thus

$$K_0(x,y) = -e_x \frac{1}{|y|^3} - |x-\alpha_1|\hat{1} + e_{\alpha_1}\frac{1}{|y|^3},$$

so with

$$-\int_{\mathbb{R}^2} \{-e_x \hat{f}(y) + e_{\alpha_1}\hat{f}(y)\} dm_2(y) = f(x) - f(\alpha_1),$$

and

$$|x-\alpha_1| \int_{\mathbb{R}^2} |y|^3 \hat{1}\hat{f}(y) dm_2(y) = 0,$$

the result follows. It should be noted that $K_0(\alpha_1,\alpha_1) = 0$.

Using reproducing kernel functions, spline functions are represented as

$$s(x) = \sum_{j=1}^{q} c_j \tilde{p}_j(x) + \sum_{i=1}^{N-q} a_i \langle K_0(x,\cdot), \mu_i \rangle, \tag{4.93}$$

where the first q functionals are chosen from $\{\mu_i\}_{i=1,\ldots,N}$ to form a dual basis for $\mathbb{P}_{m-1}$. The coefficients c_j, $j = 1,\ldots,q$ are simply the observation values $c_j = \langle f, \mu_j \rangle$ (cf. (4.88)); the remaining coefficients a_i, $i = 1,\ldots,N-q$, are obtained by solving the linear equations arising from the remaining observations:

$$\sum_{i=1}^{N-q} a_i \langle\langle K_0(\cdot,\cdot),\mu_i\rangle,\mu_j\rangle = c_j - \sum_{i=1}^{q} c_i \langle \tilde{p}_i,\mu_i\rangle, \quad j = 1,\ldots,N-q. \tag{4.94}$$

The coefficient matrix made up of the elements $\langle\langle K_0(\cdot,\cdot),\mu_i\rangle,\mu_j\rangle$, $i,j = 1,\ldots,N-q$, coming from the $N-q$ independent functionals μ_i, is a Gram matrix (made up of inner products), and thus symmetric and positive-definite. There are efficient and numerically stable computation techniques for such linear systems; *Meinguet (1983)* makes a plea for using the Cholesky decomposition. On the other hand, in the equations (4.68) the diagonal elements are zero, and so the coefficient matrix is not positive-definite. (It is conditionally positive-definite with respect to B, i.e. $\varepsilon a^t A a > 0$ if $Ba = 0$, with $a \neq 0$, where ε is either +1 or -1, *Dyn et al (1986)*). The numerical difficulties in solving these equations have been the concern of *Dyn and Levin (1983)*, and *Dyn et al (1986)*.

4.11 Smoothing Splines

The results of the previous sections for natural rotation-invariant interpolating splines extend readily to natural smoothing splines. As before, Ω is a bounded domain, and $-m - \frac{n}{2} < s < \frac{n}{2}$.

Theorem: Let $\varepsilon > 0$; let $\{\mu_i\}_{i=1,\ldots,N}$ be a linearly independent $\mathbb{P}_{m-1}$-unisolvent set in $H_{comp}^{-m-s}(\mathbb{R}^n)$. Let $f \in H^{m+s}(\Omega)$. There is a unique $s \in D^{-m}\tilde{H}^s(\mathbb{R}^n)$ of the form $\mu * K_{2m+2s-n} + p$, with $\mu \in M \cap \mathbb{P}^\circ_{m-1}$ (and $M = \text{span}\mu_i$) minimising the quadratic form

$$\| s \|^2_{D^{-m}\tilde{H}^s(\mathbb{R}^n)} + \varepsilon \sum_{i=1}^{N} |\langle s - f,\mu_i\rangle|^2. \tag{4.95}$$

The required function is of the form

$$s = \sum_{i=1}^{N} a_i \mu_i * K_\lambda + \sum_{i=1}^{q} b_i p_i \tag{4.96}$$

(where $q = \binom{m+n-1}{n} = \dim \mathbb{P}_{m-1}$, and $\{p_i\}_{i=1,\dots,q}$ is a basis for $\mathbb{P}_{m-1}$), and is obtained by solving the linear equations

$$\sum_{i=1}^{N} a_i \langle \mu_i * K_\lambda, \mu_j \rangle + \frac{1}{\varepsilon} a_j + \sum_{i=1}^{q} b_i \langle p_i, \mu_j \rangle = \langle f, \mu_j \rangle, \quad j = 1,\dots,N, \tag{4.97}$$

subject to the natural spline conditions

$$\sum_{j=1}^{N} a_i \langle p_j, \mu_i \rangle = 0, \quad j = 1,\dots,q,$$

or, in matrix notation (cf. (4.68))

$$\begin{aligned} (A + \tfrac{1}{\varepsilon} I)a + B^t b &= c, \\ Ba &= 0. \end{aligned} \tag{4.98}$$

(here $1/\varepsilon$ plays the role of the 'ridging parameter' described by *Marquardt (1970)*).

Proof: Suppose s_0 is a solution to the homogenous equation system corresponding to (4.98), i.e. $c = 0$,

$$\langle s_0, \mu_i \rangle + \frac{1}{\varepsilon} a_i = 0,$$

so

$$\sum_{i=1}^{N} \langle s_0, \mu_i \rangle^2 + \frac{1}{\varepsilon} \sum_{i=1}^{N} a_i \langle s_0, \mu_i \rangle = 0,$$

$$\sum_{i=1}^{N} <s_0,\mu_i>^2 + \frac{1}{\varepsilon} <\sum_{i=1}^{N} a_i\mu_i * K_\lambda, \sum_{j=1}^{N} a_j\mu_j> = 0,$$

or

$$\sum_{i=1}^{N} <s_0,\mu_i>^2 + \frac{1}{\varepsilon} \| s_0 \|^2_{D^{-m}\tilde{H}^s(\mathbb{R}^n)} = 0. \tag{4.99}$$

Here we have made use of the fact that $\sum_{i=1}^{N} a_i <p_j,\mu_i> = 0$, and that $\mathbb{P}_{m-1}$ is in the nullspace of the semi-norm. We infer from (4.99) that $s_0 \in \mathbb{P}_{m-1}$, and that $<s_0,\mu_i> = 0$, $i = 1,\ldots,N$, i.e. $s_0 = 0$. The homogenous equations corresponding to (4.98) admit only the trivial solution; in consequence (4.98) has a unique solution.

If s is a smoothing spline, it satisfies (2.14), which equation here takes the form

$$(s,x)_{D^{-m}\tilde{H}^s(\mathbb{R}^n)} + \varepsilon \sum_{i=1}^{N} <s - x,\mu_i><x,\mu_i> = 0, \tag{4.100}$$

for all $x \in D^{-m}\tilde{H}^s(\mathbb{R}^n)$. We use this fact to check that the solution to (4.98) minimises (4.95). If s is the solution to (4.98), and g is any other element of $D^{-m}\tilde{H}^s(\mathbb{R}^n)$, then (with $g = s - (s - g)$)

$$\| g \|^2_{D^{-m}\tilde{H}^s(\mathbb{R}^n)} + \varepsilon \sum_{i=1}^{N} <s - g,\mu_i>^2 = \| s \|^2 + \varepsilon \sum_{i=1}^{N} <s - f,\mu_i>^2$$

$$+ \| s - g \|^2 + \varepsilon \sum_{i=1}^{N} <s - g,\mu_i>^2 - 2(s,s - g)$$

$$- 2\varepsilon \sum_{i=1}^{N} <s - f,\mu_i><s - g,\mu_i>.$$

The last two terms are zero by (4.100), and thus

$$\| s \|^2 + \varepsilon \sum_{i=1}^{N} <s - f,\mu_i>^2 \leq \| g \|^2 + \varepsilon \sum_{i=1}^{N} <s - g,\mu_i>^2,$$

What this regularisation achieves is the replacement of the kernel function $K_\lambda = F(|x|)$ as given by (4.62) by $\overline{K}_\lambda = F((|x|^2+\varepsilon^2)^{1/2})$. The best known example is

$$\overline{K}_\lambda = \{|x|^2 + \varepsilon^2\}^{1/2},$$

which forms the basis of Hardy's multiquadric technique. This form of regularisation can be called *Hardy regularisation*, as the functions $\overline{K}_\lambda$ have come to be referred to as *Hardy-type functions (Dyn et al, 1986)*. These are C^∞ functions; they have proven to be most suitable functions for interpolation and approximation *(ibid.)*, but the well-posedness of interpolation with Hardy-type functions has only recently been established, by *Micchelli (1986)*.

The Fourier transform of $\{|x|^2+\varepsilon^2\}^\beta$, β neither an integer nor of the form $-\frac{n}{2}-k$, $k \in \mathbb{Z}^+$, is

$$c|\xi|^{-\beta-\frac{n}{2}} K_{\beta+\frac{n}{2}}(\varepsilon|\xi|) \qquad (4.101)$$

(Jones (1982), p 339; Gel'fand and Shilov (1964), p 288, where expressions may also be found for the constant c). $K_{\beta+\frac{n}{2}}$ is the modified Bessel function of third kind *(Abromowitz and Segun, 1965, p 444)*.

The kernel function and the Fourier transform from which it derives of multi-conic splines in $\mathbb{R}^2$ are $|x|$ and $|\xi|^{-3}$, respectively (vide Table 4.1). For multiquadric functions the regularised analogues are $\{|x|^2+\varepsilon^2\}^{1/2}$ and $|\xi|^{-3/2}K_{3/2}(\varepsilon|\xi|)$. Figure 4.1 shows the relative

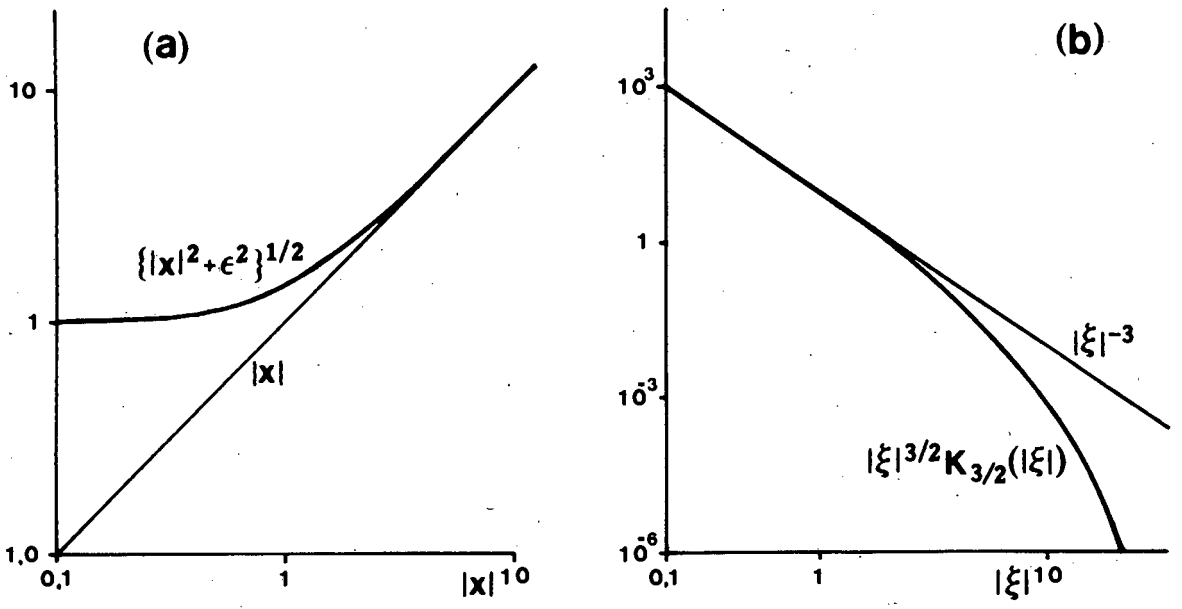


Figure 4.1: (a) Kernel function, and (b) its Fourier transform (up to a constant coefficient) for multi-conic and multiquadric splines on $\mathbb{R}^2$, with $\varepsilon = 1$. (Values for $K_{3/2}|\xi|$ taken from *Abramowitz and Stegun (1965), Tables 10.8 and 10.10).*

behaviour of the multi-conic kernel function and its Hardy-type equivalent. It can be seen from the figure that the effect of the regularisation is to make the Fourier transform of the kernel function fall off considerably more rapidly. It can also be seen that, as ε increases, the Fourier transform falls off even more rapidly, or conversely, that the spline representation becomes smoother. *Dyn et al (1986)* point to the effect of this type of regularisation as 'of the type of a low-pass filter', leading to surface interpolation by radial basis functions containing only the smooth components of the surface splines (4.74).

The question that remains open is what are the general properties of the spaces to which these regularised splines belong, i.e. the properties

of the spaces (cf. (4.37))

$$\{u \in \Phi_n' : \int_{\mathbb{R}^n} \{(1+|y|^2)^{\beta}\}\hat{}\,|\hat{u}(y)|^2 dm_n(y) < \infty\}.$$

For the spaces $H^s(\mathbb{R}^n)$ themselves there is a ready answer: as *Micchelli (1986)* indicates, $c|\xi|^{-\beta-\frac{n}{2}}K_{\beta+\frac{n}{2}}(|\xi|)$ - being the Fourier transform of $(|x|^2+1)^{\beta}$ - is a kernel function for the Sobolev space $H^s(\mathbb{R}^n)$ of (4.37), in the same way as $K_{2m+2s-n}$, (4.62), is a kernel for the Beppo-Levi space $D^{-m}\tilde{H}^s(\mathbb{R}^n)$ of (4.50); but (4.101) is unattractive as a basis function for surface interpolation.

4.13 Approximation of Linear Functionals

In this section a spline approximation satisfying the exactness condition (2.29) is sought within the framework set up earlier in Sections 4.7 and 4.8 for spline interpolation on $\mathbb{R}^n$. As in those sections $-\frac{n}{2} < m+s < \frac{n}{2}+m$; Ω is a bounded domain in $\mathbb{R}^n$; $\{\mu_i\}_{i=1,\ldots,N}$ are $\mathbb{P}_{m-1}$-unisolvent linearly independent bounded linear functionals in $H_{\overline{\Omega}}^{-m-s}(\mathbb{R}^n)$, and they span a linear sub-space M. $\{p_k\}_{k=1,\ldots,q}$ is a basis for $\mathbb{P}_{m-1}$ (of dimension $q = \binom{m+n-1}{n}$); $\mathbb{P}_{m-1}$ is the nullspace of $D^{-m}\tilde{H}^s(\mathbb{R}^n)$.

Theorem: Let $\ell \in H_{\Omega}^{-m-s}(\mathbb{R}^n)$. There is a unique element

$$\mu = \sum_{i=1}^{N} a_i \mu_i \tag{4.102}$$

in M satisfying the exactness condition

$$\langle s, \mu \rangle = \langle s, \ell \rangle$$

for any $s \in \theta(M \cap \mathbb{P}^{\circ}_{m-1}) + \mathbb{P}_{m-1}$, where θ is the kernel mapping (4.57).

Proof: The approximation is to be exact on $\mathbb{P}_{m-1}$, i.e. it must satisfy

$$\sum_{i=1}^{N} a_i \langle p_k, \mu_i \rangle = \langle p_k, \ell \rangle, \quad k = 1, \ldots, q, \tag{4.103}$$

and it must also be made exact on $\theta(M \cap \mathbb{P}^{\circ}_{m-1})$. Let $\{\tilde{\mu}_k\}_{k=1,\ldots,q}$ be a Lagrange dual basis in M of $\{p_k\}_{k=1,\ldots,q}$ spanning $\mathbb{P}_{m-1}$, i.e.

$$\langle p_i, \tilde{\mu}_j \rangle = \delta_{ij}, \quad i,j = 1, \ldots, q \tag{4.104}$$

- a dual version of (4.87), and possible for the same reason: the $\mathbb{P}_{m-1}$-unisolvency of the μ_k. Let

$$\mu_j^* = \mu_j - \sum_{k=1}^{q} \langle p_k, \mu_j \rangle \tilde{\mu}_k, \quad j = 1, \ldots, N; \tag{4.105}$$

then $\mu_j^* \in M \cap \mathbb{P}^{\circ}_{m-1}$. $\sum_{i=1}^{N} a_i \mu_i$ is to be exact on each $\theta\mu_j^*$, i.e.

$$\begin{aligned} \sum_{i=1}^{N} a_i \langle \mu_j * K_\lambda, \mu_i \rangle - \sum_{k=1}^{q} \sum_{i=1}^{N} a_i \langle \tilde{\mu}_k * K_\lambda, \mu_i \rangle \langle p_k, \mu_j \rangle \\ = \langle \mu_j * K_\lambda, \ell \rangle - \sum_{k=1}^{q} \langle \tilde{\mu}_k * K_\lambda, \ell \rangle \langle p_k, \mu_j \rangle, \quad j = 1, \ldots, N, \end{aligned} \tag{4.106}$$

with K_λ the kernel function (4.62). Putting (4.103) and (4.106) together, the equations to be satisfied if the exactness condition is to hold, are

$$\sum_{i=1}^{N} a_i \langle \mu_j * K_\lambda, \mu_i \rangle + \sum_{k=1}^{q} b_k \langle p_k, \mu_j \rangle = \langle \mu_j * K_\lambda, \ell \rangle, \quad j = 1, \ldots, N$$

$$\sum_{i=1}^{N} a_i \langle p_k, \mu_i \rangle = \langle p_k, \ell \rangle, \quad k = 1, \ldots, q, \tag{4.107}$$

where the b_k are the (as yet undetermined) coefficients

$$b_k = -\sum_{i=1}^{N} a_i \langle \tilde{\mu}_k * K_\lambda, \mu_i \rangle + \langle \tilde{\mu}_k * K_\lambda, \ell \rangle. \tag{4.108}$$

(4.107) can be written in matrix form as

$$\begin{aligned} Aa + B^t b &= c \\ Ba \qquad &= d \end{aligned} \tag{4.109}$$

where A,B,a,b are precisely as on pp 86,87 for the interpolation problem, and

$$c_j = \langle \mu_j * K_\lambda, \ell \rangle, \quad j = 1, \ldots, N;$$

$$d_k = \langle p_k, \ell \rangle, \quad k = 1, \ldots, q.$$

Since (4.68), the spline interpolation analogue to (4.109), has a unique solution, so too does (4.109), yielding the unique approximation in M satisfying the exactness condition.

Freeden (1981a) argues along different lines to obtain the equivalent result for harmonic splines. His treatment of the minimum properties - μ has smaller semi-norm than any exact non-element of M, and best approximates ℓ among non-exact elements of M - can be transposed to the present case without much difficulty.

If the reproducing kernel functions of Section 4.10 are applied to the situation here the coefficients b_k would not appear in (4.107), since $\tilde{\mu}_k$, $k = 1,\ldots,q$, would be known explicitly, and the first N equations will reduce to $N - q$ equations involving only linear combinations of $\mu_j * K_\lambda$, $j = 1,\ldots,N - q$, with the μ_j spanning $M \cap \mathbb{P}^\circ_{m-1}$.

Chapter 6 sees an application of the approximation technique of this section, in the thin-plate quadrature of geodetic integrals.

CHAPTER 5

PERIODIC, SPHERICAL AND HARMONIC SPLINES

5.1 Introduction

It is not only the generalised spline theory of Chapters 2 and 3 that will be useful in looking at the geodetically significant cases of splines on the circle, on the sphere, and splines harmonic outside the sphere; the same arguments that were used in the previous chapter for splines on $\mathbb{R}^n$, to find specific kernel mappings from the dual space to which the observations belong to the space in which the interpolating/smoothing spline function is sought, are readily transposed to the new contexts of the present chapter.

Periodic splines are considered since the theory is easily developed and pre-figures much of the spherical spline theory, which is the subject of Section 5.4, following Section 5.3, where some basic facts about spherical harmonics are recalled. Section 5.5 shows how spherical splines readily extend to splines harmonic outside a sphere. It is important for geodetic applications that these harmonic splines may be regularised, by 'burying' the knots, so that the continuity of the splines at the observation points is unimpaired. This technique of regularisation is the Bjerhammar regularisation familiar in other geodetic situations. By comparing the spectra of different spherical splines it is possible to recommend splines most likely to be effective for gravity field approximation.

This chapter places the work of *Freeden (1981a, 1981b, 1984)* and *Wahba*

(1981, 1984) in the setting of the general spline theory adopted in earlier chapters.

5.2 Periodic Splines

Since the treatment of periodic splines serves mainly to set the scene for spherical splines, only splines on the unit circle Π will be considered and splines on tori Π^n will be neglected for $n > 1$. The weighting function used in Section 4.5 will also be neglected, i.e. $s = 0$. We will thus be looking at periodic analogues of univariate polynomial splines of the kinds considered by *Schoenberg (1964)* and *Sünkel (1984a, 1984b)*.

Let $\mathbb{D}(\Pi)$ be the space of functions ϕ on Π whose 2π-periodic extensions to $\mathbb{R}$ are in $C^\infty(\mathbb{R})$, and suppose $\mathbb{D}(\Pi)$ is given the Schwartz topology, i.e. the Frechet topology deriving from the norms

$$\max_{|\alpha| \le m} \sup_{x \in \mathbb{R}} |(D^\alpha \tilde{\phi})(x)|, \tag{5.1}$$

for all $m \in \mathbb{Z}^+$, where $\tilde{\phi}$ is the periodic extension of ϕ, cf. (4.26) *(Rudin, 1973, p 190)*. $\mathbb{D}'(\Pi)$ the space of continuous linear functionals on $\mathbb{D}(\Pi)$, is the space of distributions on Π. All such distributions are tempered. The Fourier transform on $\mathbb{D}'(\Pi)$ is defined as before:

$$\hat{u}\phi = u\hat{\phi}, \quad \phi \in \mathbb{D}(\Pi). \tag{5.2}$$

Since ϕ is periodic, i.e. $\tau_{-2\pi}\tilde{\phi} = \tilde{\phi}$, we have $\hat{\phi} = e_{2\pi}\hat{\phi}$, and the support of

$\hat{\phi}$ is contained in the zeroes of $e^{2\pi i\xi} - 1$, i.e. the support of $\hat{\phi}$ is $\mathbb{Z} = \{0,\pm1,\pm2,\dots\}$, and

$$\hat{\phi}(k) = c_k = \frac{1}{2\pi}\int_0^{2\pi} e_{-k}(x)\phi(x)\,dm(x) \tag{5.3}$$

are the Fourier coefficients of $\hat{\phi}$. The Fourier coefficients of $\hat{u}$ are

$$\hat{u}(k) = ue_{-k}. \tag{5.4}$$

The inverse Fourier transform of (5.3) is

$$\phi(x) = \sum_{k\in\mathbb{Z}} c_k e_k(x), \tag{5.5}$$

and the inverse transform of (5.4) is

$$u = \sum_{k\in\mathbb{Z}} \hat{u}(k)e_k. \tag{5.6}$$

e_k are eigenfunctions of the differential operator D^2, corresponding to eigenvalues k^2:

$$D^2e_k = -k^2e_k.$$

It can be seen from (5.3) that

$$\delta\hat{\phi} = \hat{\phi}(0) = c_0 = \frac{1}{2\pi}\int_0^{2\pi} \phi\,dm = 1\cdot\phi$$

for all ϕ in $\mathcal{D}(\Pi)$, that $\hat{\delta} = 1$, and thus (formally)

$$\hat{1}(x) = \delta(x) = 1 + 2\sum_{k=1}^{\infty} \cos kx. \tag{5.7}$$

The space $\mathbb{H}^0(\Pi)$ is defined to be

$$\mathbb{H}^0(\Pi) = \mathbb{L}^2(\Pi) = \{u \in \mathbb{D}'(\Pi): \frac{1}{2\pi}\int_{\Pi} |u(x)|^2 dm(x) = \sum_{k \in \mathbb{Z}} |u(k)|^2 < \infty\}. \tag{5.8}$$

It is a Hilbert space with inner product

$$(u,v)_{\Pi} = \frac{1}{2\pi}\int_0^{2\pi} u(x)\overline{v(x)}dm(x).$$

$\mathbb{H}^m(\Pi)$ is the space

$$\mathbb{H}^m(\Pi) = \{u \in \mathbb{D}'(\Pi): D^m u \in \mathbb{H}^0(\Pi)\}; \tag{5.9}$$

it is a semi-Hilbert space with semi-norm

$$\| u \|^2_{\Pi,m} = \frac{1}{2\pi}\int_0^{2\pi} |D^m u(x)|^2 dm(x).$$

Its nullspace is $\mathbb{P}_0$, the space of constant functions. Further, $\mathbb{H}^0(\Pi)$ is a dense Hilbert sub-space of $\mathbb{D}'(\Pi)$, and $\mathbb{H}^m(\Pi)$ is a dense semi-Hilbert sub-space. Sobolev's embedding theorem applies equally to the present case: $\mathbb{H}^m(\Pi) \subset C^k(\Pi)$ for $m < k + 1/2$.

Theorem: The map $\theta:(\mathbb{D}')' \cap \mathbb{P}_0^{\circ} \to \mathbb{D}'$ through

$$\theta: \mu \mapsto \mu^*(|k|^{-2m})\hat{} \tag{5.10}$$

is a kernel mapping of $\mathbb{H}^m(\Pi)$ as a semi-Hilbert sub-space of $\mathbb{D}'(\Pi)$.

is the kernel associated with $\mathbb{H}^m(\Pi)$. For $m=1$, K_{2m} is continuous; for $m=2$ it has continuous derivatives, etc.

Example: The kernel mapping (5.10) is used to verify that $\langle w,\delta\rangle = w(0)$, for $w \in \mathbb{H}^m(\Pi)$:

$$\langle w,\delta\rangle = (w,\delta * K_{2m})_{\Pi,m} = \sum_{k\in\mathbb{Z}} |k|^{2m}\hat{w}(k)|k|^{-2m} = \sum_{k\in\mathbb{Z}} \hat{w}(k) = w(0).$$

It is possible to obtain closed expressions for $K_{2m}(x)$. We have *(Spiegel, 1968, p 135)*

$$K_2(x) = \frac{1}{6}\pi^2 - \frac{1}{2}\pi x + \frac{1}{4}x^2, \tag{5.12}$$

and it is clear from (5.11) that K_{2m} is orthogonal to $\mathbb{P}_0$, that DK_{2m} is likewise orthogonal, and that $D^2K_{2m} = -K_{2m-2}$. Hence K_{2m}, $m=2,3,\ldots$, can be found by repeated integration, using the orthogonality condition to determine the constant of integration. In this way, for example,

$$K_3(x) = \sum_{k=1}^{\infty} \frac{\sin kx}{k^3} = \frac{1}{6}\pi^2 x - \frac{1}{4}\pi x^2 + \frac{1}{12}x^3; \tag{5.13}$$

$$K_4(x) = \sum_{k=1}^{\infty} \frac{\cos kx}{k^4} = -\frac{1}{90}\pi^4 + \frac{1}{12}\pi^2 x^2 - \frac{1}{12}\pi x^3 + \frac{1}{48}x^4. \tag{5.14}$$

Interpolation by periodic splines: Let $\{\mu_1,\mu_2,\ldots,\mu_N\}$ be a set of independent bounded linear functionals on $\mathbb{D}'(\Pi)$ which resolve constant functions. Let $f \in \mathbb{D}'(\Pi)$; there exists a unique $s \in \mathbb{H}^m(\Pi)$ of the form

$$s = v * K_{2m} + p_0,$$

where $\nu \in \text{span}(\{\mu_i\}) \cap \mathbb{P}_0^\circ$, $p \in \mathbb{P}_0$, satisfying

$$<s,\mu_i> = <f,\mu_i>, \quad i = 1,\ldots,N,$$

with minimum semi-norm $\| s \|^2_{\Pi,m} = \frac{1}{2\pi}\int_0^\pi |D^m s(x)|^2 dm(x)$.

The required spline is the function

$$s = \sum_{i=1}^{N} a_i \mu_i * K_{2m} + p_0 \tag{5.15}$$

satisfying the interpolatory conditions

$$\sum_{i=1}^{N} a_i <\mu_i * K_{2m}, \mu_j> + b<1,\mu_j> = <f,\mu_j>,$$

for $j = 1,\ldots,N$, and the natural end condition

$$\sum_{i=1}^{N} a_i <1,\mu_i> = 0,$$

and whose coefficients are obtained by solving the linear equations

$$Aa + B^t b = c$$

$$Ba = 0$$

as in (4.68).

A reproducing kernel function corresponding to $K_{2m}(x)$ is

$$K_{2m,0}(x,y) = K_{2m}(x-y) - K_{2m}(x-\alpha_1) - K_{2m}(y-\alpha_1), \tag{5.16}$$

where we have taken $\mu_1 = \delta_{\alpha_1}$. This can be verified by direct computation, by checking that, for $f \in \mathbb{H}^m(\Pi)$,

$$f(x) = f(\alpha_1) + (K_{2m,0}(x,y), f(y))_{\Pi,m}.$$

First note that for g a periodic function, $g(y-x) = e_1(x)g(y)$; thus considering $K_{2m,0}$ a function of y,

$$(K_{2m,0}(x,y))\hat{}(k) = \{e_1(x)\frac{1}{k^{2m}} - K_{2m}(x-\alpha_1)\hat{1} - e_1(\alpha_1)\frac{1}{k^{2m}}\},$$

and

$$(K_{2m,0}(x,y), f(y))_{\Pi,m} = \sum_{k \in \mathbb{Z}} \{e_1(x)\hat{f}(k) - e_1(\alpha_1)\hat{f}(k)\} = f(x) - f(\alpha_1).$$

Trigonometric splines: The periodic splines above come from differential operators D^m inducing semi-norms with nullspace $\mathbb{P}_0$ in each case. Trigonometric splines are obtained from differential operators of degree m, and having nullspace trigonometric polynomials of degree m-1 or less. For example, the operator

$$A_m = D \prod_{i=1}^{m-1} |D^2 + i|^2 = D(D^2+1)(D^2+4)\ldots\ (D^2 + (m-1)^2) \quad (5.17)$$

is associated with the kernel mapping

$$\theta: \mu \mapsto \mu^* \sum_{k=m}^{\infty} \frac{1}{2}\Big(\prod_{j=1}^{m-1} \frac{1}{(k^2 - j^2)^2}\Big)\cos kx, \quad (5.18)$$

for which a closed expression can also be found *(Schoenberg, 1964; Sünkel, 1984a)*.

Much more can be said about periodic and trigonometric splines;

besides the references cited above, *Schumaker (1981)* gives an extensive treatment. Periodic splines share many features with spherical splines; it is to these that we now turn.

5.3 Spherical Harmonics

In considering spherical harmonics, attention will be restricted to the practically significant case of harmonics in $\mathbb{R}^3$, and spherical harmonics in Euclidean spaces of higher dimension will be ignored.

Let Ω be the unit sphere in $\mathbb{R}^3$, with its centre at the origin, and let E be the unbounded region exterior to Ω. A Laplace spherical harmonic of degree n is the restriction $s_n = H_n|_\Omega$ to Ω of an harmonic homogenous polynomial H_n of degree n; we have

$$H_n(x) = |x|^{-(n+1)} s_n(\xi) \tag{5.19}$$

for $x \in E$, $x = |x|\xi$, $\xi \in \Omega$. S_n, the space of spherical harmonics of degree n, is a linear space of dimension $2n+1$ *(Stein and Weiss, 1971, p 139; Jones, 1985, p 2)*. Spherical harmonics of different degree are orthogonal with respect to the $L^2(\Omega)$ inner product, i.e. for $s_n \in S_n$, $s_m \in S_m$, $n \neq m$,

$$(s_n, s_m)_\Omega = \int_\Omega s_n(\xi) s_m(\xi) d\omega(\xi) = 0. \tag{5.20}$$

S_n contains a basis set $\{s_{n,1}, s_{n,2}, \ldots, s_{n,2n+1}\}$ orthonormal with respect to the $L^2(\Omega)$ inner product. The orthogonal direct sum

$$\bigoplus_{n=0}^{\infty} S_n = S_0 \oplus S_1 \oplus \ldots$$

is uniformly dense in $C(\Omega)$, the space of continuous functions on Ω, and dense in the Hilbert space $L^2(\Omega)$ of Lebesgue square-integrable functions on Ω *(Stein and Weiss, 1971, p 141)*. Hence the set $\{s_{n,j}\}_{\substack{n=0,1,\ldots \\ j=1,2,\ldots,2n+1}}$ is a complete orthonormal family in $L^2(\Omega)$: any $f \in L^2(\Omega)$ may be approximated (in the mean) by its Fourier representation,

$$\lim_{N\to\infty} \| f - \sum_{n=0}^{N} \sum_{j=1}^{2n+1} (f, s_{n,j})_\Omega s_{n,j} \|_\Omega = 0. \tag{5.21}$$

The addition theorem of spherical harmonics is also required: for $\xi, \eta \in \Omega$

$$\sum_{j=1}^{2n+1} s_{n,j}(\xi) s_{n,j}(\eta) = \frac{2n+1}{4\pi} P_n(\xi \cdot \eta), \tag{5.22}$$

where P_n is the Legendre polynomial of degree n, and $\xi \cdot \eta$ is the Euclidean inner product in $\mathbb{R}^3$ *(Müller, 1966, p 10)*. (5.22) is a reproducing kernel function for the space S_n. The space $\mathbb{P}_m = \bigoplus_{n=0}^{m} S_n$ of spherical harmonics of degree m or less, has dimension $(m+1)^2$.

In the same way as was done in defining distributions on $\mathbb{R}^n$ and periodic distributions, $\mathbb{D}(\Omega)$ is the space of infinitely differentiable functions on Ω given the canonical Schwartz topology, and $\mathbb{D}'(\Omega)$ the space of distributions on Ω, i.e. the space of continuous linear functionals on $\mathbb{D}$ *(Freeden, 1984)*.

As differential operator we use the surface gradient

$$\nabla_* = \underline{e}_\theta \frac{\partial}{\partial\theta} + \underline{e}_\lambda \frac{1}{\sin\theta} \frac{\partial}{\partial\lambda}; \tag{5.23}$$

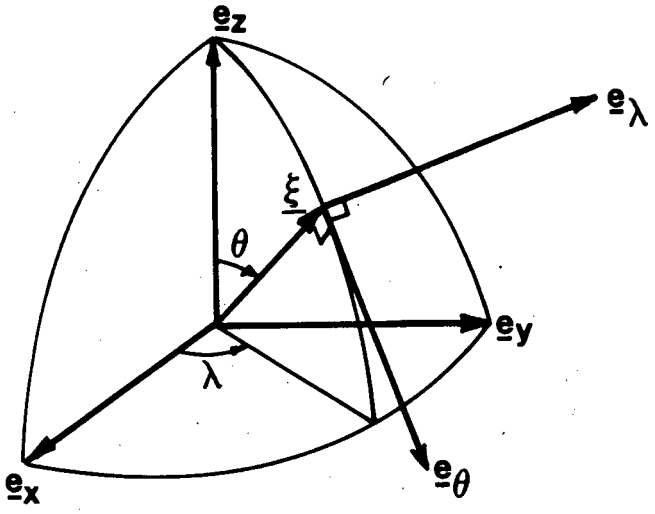


Figure 5.1: Co-ordinates on the unit sphere.

$$\nabla_*^2 = \nabla_*\nabla_* = \Delta^* = \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}(\sin\theta\frac{\partial}{\partial\theta}) + \frac{1}{\sin^2\theta}\frac{\partial}{\partial\lambda^2}. \tag{5.24}$$

This, the restriction to Ω of the Laplace operator in $\mathbb{R}^3$, is the *Laplace-Beltrami operator*. Spherical harmonics $s_{n,j}$ of degree n are eigenfunctions of Δ^*, corresponding to the eigenvalue $\lambda_n = n(n+1)$,

$$\Delta^* s_{n,j} = -n(n+1)s_{n,j}.$$

$H^m(\Omega)$ is defined to be

$$H^m(\Omega) = \{u \in \mathbb{D}'(\Omega): \ |\nabla_*^m u| \in L^2(\Omega)\}, \tag{5.25}$$

where, for m even,

$$\nabla_*^m = \Delta_{m/2}^*,$$

and for m odd,

$$|\nabla_*^m u|^2 = \{\frac{\partial}{\partial\theta}(\nabla_*^{m-1}u)\}^2 + \{\frac{1}{\sin\theta}\frac{\partial}{\partial\lambda}(\nabla_*^{m-1}u)\}^2$$

(Wahba, 1981, 1984), in conformity with the convention of Section 4.6. $H^m(\Omega)$ is a semi-Hilbert space with semi-norm

$$\|u\|^2_{\Omega,m} = \int_\Omega |\nabla_*^m u|^2 d\omega. \tag{5.26}$$

In the same way as was the case for periodic distributions, much of the preparation of Chapter 4 - translation (in the present context, rotation), convolution, Fourier transform -becomes much simpler on the sphere. Let $\phi \in C(\Omega)$, $u \in \mathbb{D}'(\Omega)$, then their Fourier transforms - spherical harmonic coefficients - are

$$\begin{aligned}\hat{\phi}(n,j) &= a_{n,j} = (\phi, s_{n,j})_\Omega, \\ \hat{u}(n,j) &= b_{n,j} = (u, s_{n,j})_\Omega\end{aligned} \tag{5.27}$$

(with $\sum_{n,j} a_{n,j}$ convergent), and

$$\begin{aligned}\phi &= \sum_{n,j} a_{n,j} s_{n,j}, \\ u &= \sum_{n,j} b_{n,j} s_{n,j}\end{aligned} \tag{5.28}$$

(where $\sum_{n,j} = \sum_{\substack{n=0,1,\ldots \\ j=1,2,\ldots,2n+1}}$). In particular, for the Dirac distribution,

$$\begin{aligned}\hat{\delta}(n,j) &= s_{n,j}(0), \\ \hat{\delta}_\xi(n,j) &= s_{n,j}(\xi), \quad \xi \in \Omega,\end{aligned} \tag{5.29}$$

and

$$\delta = \sum_{n,j} s_{n,j}(0)s_{n,j},$$

$$\delta_\xi = \sum_{n,j} s_{n,j}(\xi)s_{n,j}. \tag{5.30}$$

Also

$$\tau_\xi\phi(\eta) = \delta_\xi * \phi(\eta) = \phi(\eta - \xi),$$

$$\tau_\xi u = \delta_\xi * u = \sum_{n,j} b_{n,j}s_{n,j}(\xi) \tag{5.31}$$

$(\xi,\eta \in \Omega)$ in conformity with the results of Section 4.2. Here the definition of *spherical convolution* has been anticipated:

$$u*\phi(\xi) = u(\delta_\xi * \phi) = \sum_{n,j} a_{n,j}b_{n,j}s_{n,j}(\xi), \tag{5.32}$$

which when u is integrable, coincides with the usual definition

$$u*\phi(\xi) = \int_\Omega u(\eta)\phi(\xi - \eta)d\omega(\eta)$$

$$= \int_\Omega \Big(\sum_{n,j} b_{n,j}s_{n,j}(\eta)\Big)\Big(\sum_{n,j} a_{n,j}s_{n,j}(\xi)s_{n,j}(\eta)\Big)d\omega(\eta).$$

A final point: from the relation

$$\Delta_m^* s_{n,j} = -\{n(n+1)\}^m s_{n,j}$$

it follows that

$$(\Delta_m^* u)\hat{}(n,j) = -\{n(n+1)\}^m \hat{u}(n,j). \tag{5.33}$$

5.4 Spherical Splines

We are now in a position to apply the fundamental theorem of Section 4.7 to the present situation of finding spline functions that interpolate (or smooth) data on a sphere. Throughout this section $\lambda_n = n(n+1)$.

Theorem: The map $\theta:(\mathbb{D}')' \cap \mathbb{P}_0^\circ \to \mathbb{D}'$ through

$$\theta: \mu \mapsto \mu * \{\lambda_n^{-m}\}\hat{} \tag{5.34}$$

is a kernel mapping of $H^m(\Omega)$ as a semi-Hilbert sub-space of $\mathbb{D}'(\Omega)$.

Proof: Let $v = \mu * (\lambda_n^{-m})$, then $\hat{v} = \hat{\mu}\lambda_n^{-m}$. Suppose $u \in L\mu$, where

$$L: (\mathbb{D}')' \cap \mathbb{P}_0^\circ \to \mathbb{D}'/\mathbb{P}_0$$

is the kernel associated with $H^m/\mathbb{P}_0$, then

$$\langle \hat{\phi}, \hat{\mu} \rangle = \sum_{n,j} \lambda_n^m \hat{u}(n,j)\hat{\phi}(n,j)$$

for all $\phi \in \mathbb{D}$, and so $\hat{\mu} = \lambda_n^m \hat{u}$, and

$$\lambda_n^m(\hat{v} - \hat{u}) = \lambda_n^m \hat{w} = 0,$$

i.e. $\Delta_*^m w = 0$, or $w \in \mathbb{P}_0$, hence $v \in L\mu$ as was required.

Thus we need the function having λ_n^{-m} as its Fourier transform:

$$K_m(\xi\cdot\eta) = \sum_{n,j} \lambda_n^{-m} s_{n,j}(\xi) s_{n,j}(\eta) = \sum_{n=1}^{\infty} \frac{2n+1}{4\pi} \lambda_n^{-m} P_n(\xi\cdot\eta) \qquad (5.35)$$

($\xi,\eta \in \Omega$) has the required Fourier transform.

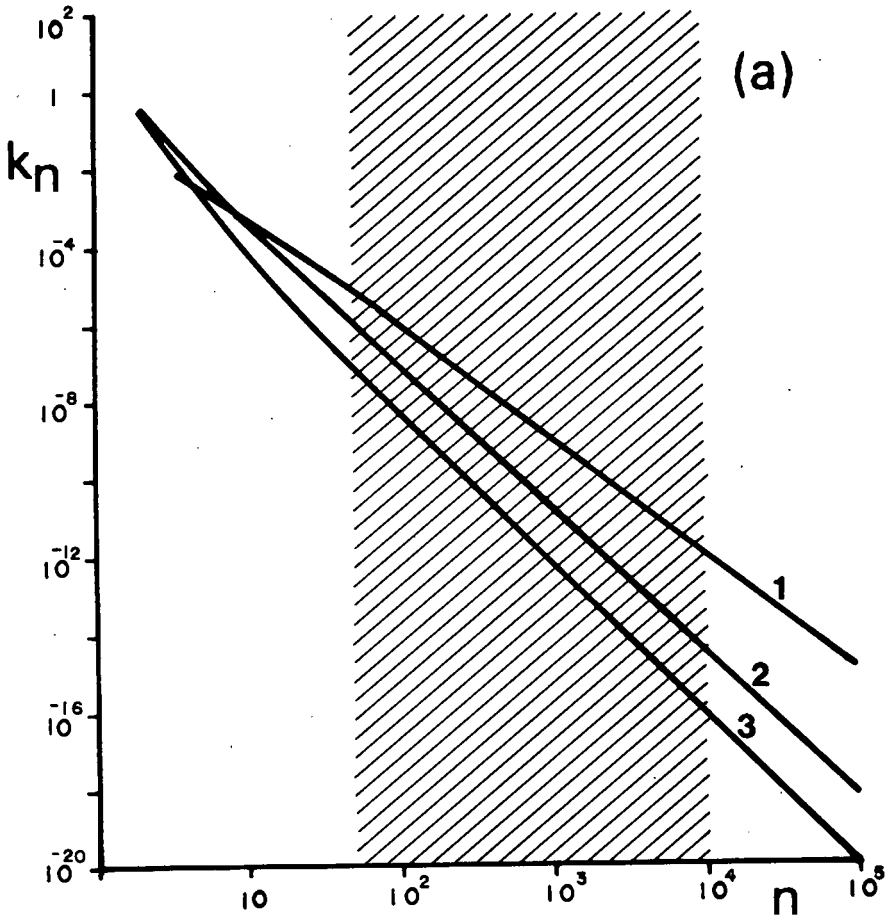
When $m = 0$, K_0 is called the *Szegö kernel*, more on which below.

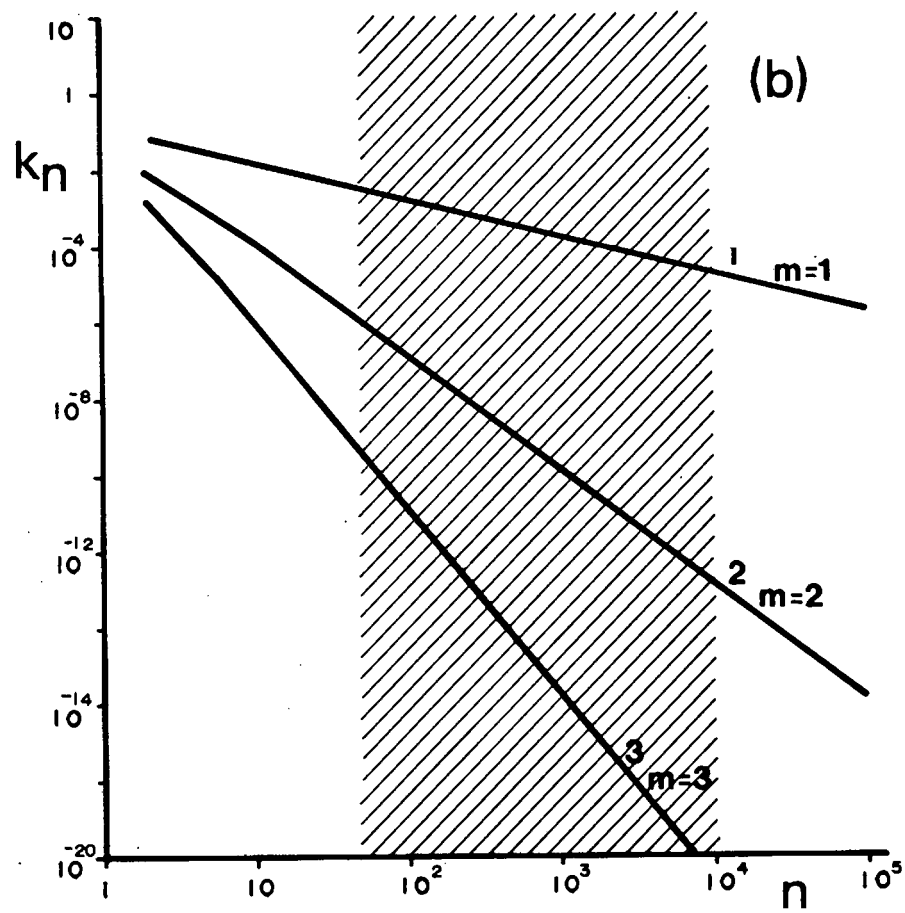
The continuity of K_m, and hence also of the splines that are combinations of K_m, depends on m. For $m = 0$, or 1, $K_m(t)$ is not continuous at $t = 1$; it is continuous for $m = 2$, has continuous derivatives for $m = 3$, and more generally, has continuous derivatives of order $m - 2$.

The realistic approximation of the earth's gravity field places a considerable constraint on the choice of m. The degree variances of the gravity field

$$k_n = \sum_{j=0}^{2n+1} a_{n,j}^2, \qquad (5.36)$$

Figure 5.2 (following page): (a) Disturbing potential degree variances, neglecting a constant variance factor, for
1. Tscherning and Rapp's 1974 degree variance model, $k_n = \{(n-1)(n-2)(n+24)\}^{-1}$;
2. *Schwarz's (1984)* model in mountainous areas, $k_n = (n-1)^{-2} n^{-1,67}$;
3. *Schwarz's (1984)* model in ordinary terrain, $k_n = (n-1)^{-2} n^{-2,06}$.
(b) Spherical spline degree variances, $k_{n,m} = \frac{2n+1}{4\pi}\{n(n+1)\}^{-m}$ for $m = 1,2,3$.


(a)
k_n
10^{2}
1
10^{-4}
10^{-8}
10^{-12}
10^{-16}
10^{-20}
10
10^{2}
10^{3}
10^{4}
10^{5}
n
1
2
3


(b)
k_n
10
1
10^{-4}
10^{-8}
10^{-12}
10^{-14}
10^{-20}
1
10
10^{2}
10^{3}
10^{4}
10^{5}
n
1 m=1
2 m=2
3 m=3

where $a_{n,j}$ are the (normalised) spherical harmonic coefficients of the disturbing potential *(Heiskanen and Moritz, 1967, p 259)*, fall off asymptotically as n^{-3} according to Tscherning and Rapp's widely used model, and fall of as n^{-4} according to Kaula's Rule, for which *Schwarz (1984)* finds support also in the medium frequency range (n = 200 to 1000). This restricts choice of m to m = 2 (with perhaps also m = 3), and thus restricts the degree of continuity that can be expected of the spline representation, as also the data types that can contribute to the representation - no gravity anomalies, for example, if m = 2. Fortunately the process of *Bjerhammar regularisation* provides a ready means of overcoming this difficulty, and allows a more flexible choice of kernel function. This is the subject of the next section; for the moment the discussion of spherical splines is continued by showing how interpolating spherical splines may be constructed.

Let $\{\mu_1, \ldots, \mu_N\}$ be a set of independent bounded linear functionals on $H^m(\Omega)$, and which resolve constant functions. Let $f \in \mathbb{D}'$; there is exactly one element $s \in H^m(\Omega)$, of the form

$$s = \nu * K_m + p_0,$$

where $\nu \in \text{span}(\{\mu_i\}) \cap \mathbb{P}_0^\circ$, $p_0 \in \mathbb{P}_0$, satisfying

$$\langle s, \mu_i \rangle = \langle f, \mu_i \rangle, \quad i = 1, \ldots, N,$$

with minimum semi-norm $\| s \|_\Omega^2 = \int_\Omega |\nabla_*^m s(\xi)|^2 d\omega(\xi)$. The required spline is the function

$$s = \sum_{i=1}^{N} a_i \mu_i * K_m + p_0 \tag{5.37}$$

satisfying the interpolatory conditions

$$\sum_{i=1}^{N} a_i \langle \mu_i * K_m, \mu_j \rangle + b\langle 1, \mu_j \rangle = \langle f, \mu_j \rangle, \quad j = 1, \ldots, N, \tag{5.38}$$

and the natural spline orthogonality condition

$$\sum_{i=1}^{N} a_i \langle 1, \mu_i \rangle = 0, \tag{5.39}$$

and whose coefficients a_i, $i = 1, \ldots, N$, and b are obtained by solving the linear equations

$$\begin{aligned} Aa + B^t b &= c \\ Ba &= 0 \end{aligned} \tag{5.40}$$

as in (4.68).

A reproducing kernel function corresponding to $K_m(\xi \cdot \eta)$ is

$$K_{m,0}(\xi, \eta) = K_m(\xi \cdot \eta) - K_m(\xi \cdot \alpha_1) - K_m(\eta \cdot \alpha_1), \tag{5.41}$$

where we have taken $\mu_1 = \delta_{\alpha_1}$, $\alpha_1 \in \Omega$. This can be verified by checking, that for $f \in H^m(\Omega)$,

$$f(\xi) = f(\alpha_1) + (K_{m,0}(\xi, \eta), f(\eta))_{\Omega, m}.$$

Taking $K_{m,0}(\xi, \eta)$ as a function in η,

$$(K_{m,0}(\xi,\eta))\hat{}(n,j) = \frac{1}{\lambda_n^m} s_{n,j}(\xi) - K_m(\xi\cdot\alpha_1)\hat{1} - \frac{1}{\lambda_n^m} s_{n,j}(\alpha_1),$$

so that

$$(K_{m,0}(\xi,\eta),f(\eta))_{\Omega,m} = \sum_{n,j} \{f(n,j)s_{n,j}(\xi) - f(n,j)s_{n,j}(\alpha_1)\}$$

$$= f(\xi) - f(\alpha_1),$$

since the constant function 1 is in the nullspace of the semi-inner product.

All the splines above - linear combinations of the isotropic functions $K_m(\xi\cdot\eta)$ - share $\mathbb{P}_0$ as the null-space of the spline space inner product. *Freeden (1981a, 1981b)* derives a class of spherical splines with null-space $\mathbb{P}_m$, and kernel functions

$$\sum_{n=m+1}^{\infty} \frac{2n+1}{4\pi} \frac{1}{\{(\lambda_0-\lambda_n)(\lambda_1-\lambda_n)\ldots(\lambda_m-\lambda_n)\}^2} P_n(\xi\cdot\eta),$$

where $\lambda_n = n(n+1)$ as before. For $m = 0$, we have the kernel function K_2; for $m = 1$ the degree variances fall off eventually as n^{-7}, rather too fast for gravity field approximation. 'Covariance' functions corresponding to these splines imply an unrealistic degree of correlation within the gravity field.

Wahba (1984) has used the splines above to derive *Vector splines*, and used these for wind field modelling. A geodetic application would be for modelling horizontal gravity gradients or deflections of the vertical.

5.5 Spherical Harmonic Splines

The spherical splines on the unit sphere extend readily to a sphere of radius R; the same kernel (5.35) can be used, and the semi-norm that is minimised is

$$\| u \|^2_{R\Omega,m} = R^2 \int_{\Omega} |\nabla_*^m u|^2 d\omega. \tag{5.42}$$

Spherical harmonic splines are obtained by applying the definition (5.19) to the kernel formula (5.35): for $x_1 = r_1\xi$, $x_2 = r_2\eta$; $r_1 = |x_1|$, $r_2 = |x_2|$; $r_1, r_2 \geq R$; $\xi, \eta \in \Omega$,

$$K_m(x_1,x_2) = \sum_{n=1}^{\infty} \frac{2n+1}{4\pi} \lambda_n^{-m} \left(\frac{R^2}{r_1 r_2}\right)^{n+1} P_n(\xi\cdot\eta). \tag{5.43}$$

Sometimes the data are given on a sphere, $r_1 = r_2 = R$, so that

$$K_m(\xi\cdot\eta) = \sum_{n=1}^{\infty} \frac{2n+1}{4\pi} \lambda_n^{-m} P_n(\xi\cdot\eta). \tag{5.44}$$

For $m = 0,1$, $K_m(t)$ is not continuous for $t = 1$, i.e. for $\psi = 0$ - the splines (5.37) can not be used for point evaluations; for $m = 2$, derivatives are not continuous at the knots - (5.37) can not then be used for derivative data, e.g. gravity anomalies or deflections of the vertical when approximating the disturbing potential. This difficulty can be overcome by *Bjerhammar regularisation;* by introducing a *Bjerhammar sphere* of radius $R_B < R$, (5.44) becomes *(Moritz, 1980, p 181)*

$$K_{m,B}(\xi\cdot\eta) = \sum_{n=1}^{\infty} \frac{2n+1}{4\pi} \lambda_n^{-m} \left(\frac{R_B}{R}\right)^{2n+2} P_n(\xi\cdot\eta); \tag{5.45}$$

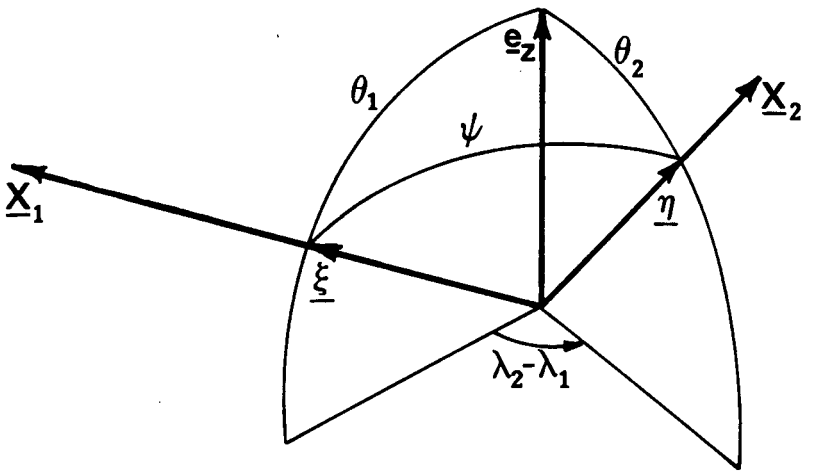


Figure 5.3: Spherical distance on the unit sphere; $\cos\psi = t = \xi \cdot \eta$.

similarly (5.43) becomes

$$K_{m,B}(x_1,x_2) = \sum_{n=1}^{\infty} \frac{2n+1}{4\pi} \lambda_n^{-m} \left(\frac{R_B^2}{r_1 r_2}\right)^{n+1} P_n(\xi \cdot \eta); \qquad (5.46)$$

both these are uniformly convergent for all values of m. The semi-norm remains (5.42). By this device discontinuities at the knots have been placed on the sphere $R_B\Omega$, at a depth $R - R_B$ below the data points on the sphere $R\Omega$; alternatively the knots have been buried at a depth $r_i - R_B$ below the data points x_i. These splines, being harmonic outside $R_B\Omega$, have derivatives of all orders on R.

Some care must be taken in choosing an appropriate radius for the Bjerhammar sphere. This will depend on the choice of spline, i.e. on m, on the data types used for the spline interpolation (or smoothing), and on the spatial distribution of these data. The numerical examples given by *Lelgemann (1981)* show quite graphically how interpolatory behaviour depends on choice of R_B: make it too close to R, and the

interpolant tends to 'local support' at the knots only; make it too small, and $K_{m,B}$ becomes almost a constant function, and linear combinations of splines find it increasingly difficult to accommodate the data, i.e. the linear equations (5.40) become increasingly ill conditioned. For reasonable results in approximating the gravity field the degree variances of (5.45) (or (5.46))

$$k_{n,B} = \frac{2n+1}{4\pi} \lambda_n^{-m} \left(\frac{R_B}{R}\right)^{2n+2}$$

should conform somewhat with the degree variances of the gravity field as contained within the data. For local gravity field approximation, with data spaced at intervals of a few minutes of arc over an area a few degrees square - when medium and high frequency features of the gravity field are being modelled - the degree variances of the disturbing potential fall off roughly as the third or fourth power of n *(Schwarz, 1984)*. Among kernel functions (5.44), $m = 2$ is an appropriate choice.

Figure 5.4(a) shows degree variances for $m = 0$, and $\frac{R_B}{R} = 0{,}97$; 0,99; 0,995 and 0,999. Their relative behaviour should be compared with Figure 5.2(a); none seems to be suited for approximating the disturbing potential. In Figure 5.4(b), showing degree variances for $m = 1$ with $\frac{R_B}{R} = 0{,}99$; 0,995 and 0,999, ratios intermediate to 0,995 and 0,999 appear reasonable. The regularised splines best suited to the gravity field seem to be spherical splines $m = 2$, with $\frac{R_B}{R}$ close to 0,9999. Figure 5.4(c) shows the degree variances for these splines with $\frac{R_B}{R} = 0{,}995$; 0,999 and 0,9999.

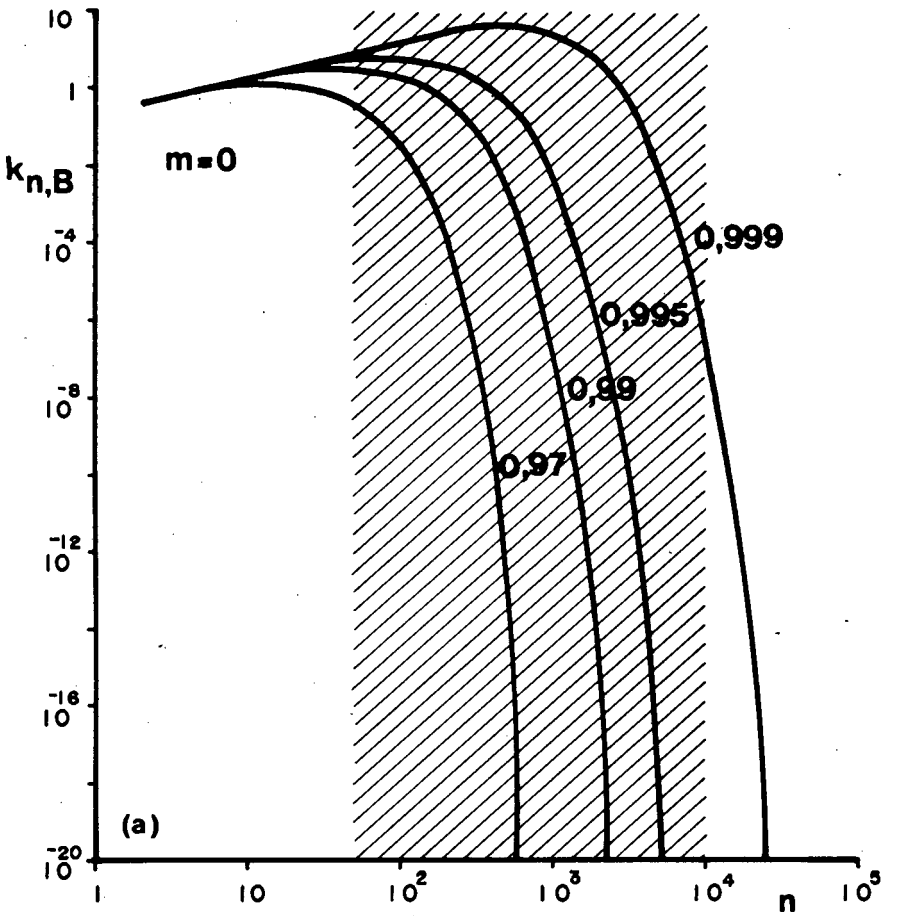
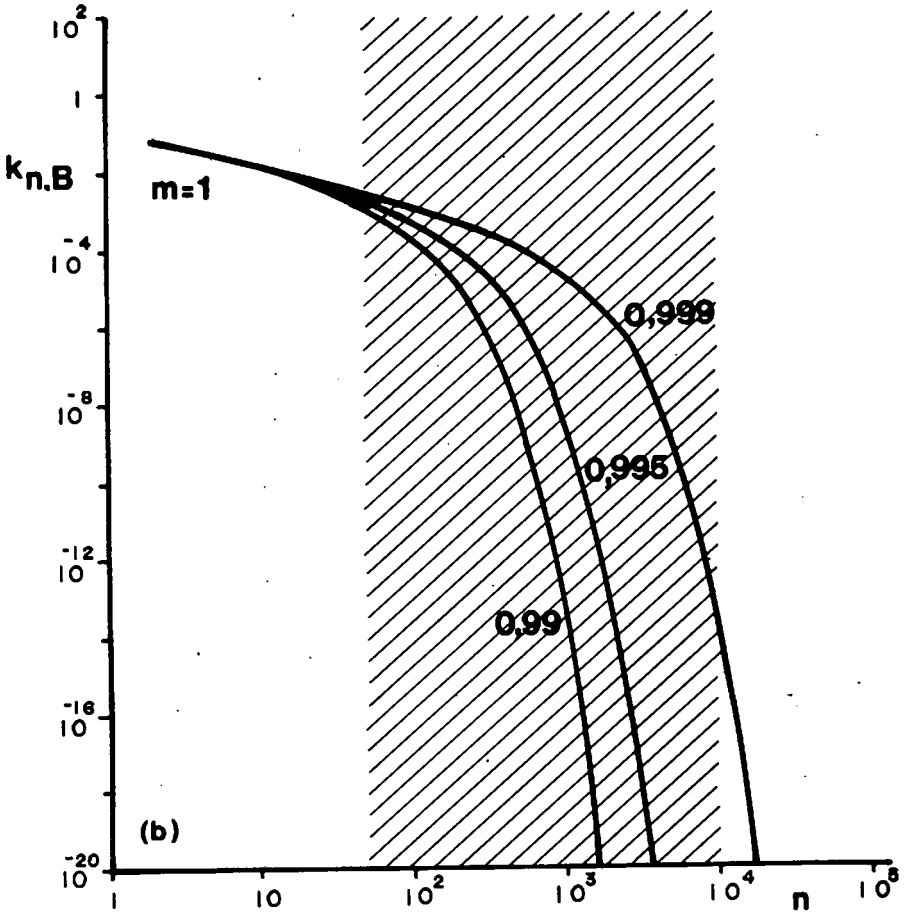


Figure 5.4: Degree variances of regularised spherical splines for (a) m = 0; (following page) (b) m = 1; (c) m = 2, for a variety of ratios R_B/R. Suitable ratios give slopes approximating those of Figure 5.2(a). The shaded area indicates the approximate spectral range of interest for local gravity field approximation.

Lelgemann's (1981) investigation led to an empirical formula for choosing the ratio $\frac{R_B}{R}$:

$$\frac{R_B}{R} = 1 - (2 + \frac{o(k_n)}{2})\psi_m,$$

where $o(k_n)$ is the order of increase of the degree variances k_n, and


10^{2}
1
$k_{n,B}$
m=1
10^{-4}
10^{-8}
10^{-12}
10^{-16}
10^{-20}
0,999
0,995
0,99
(b)
1
10
10^{2}
10^{3}
10^{4}
10^{5}
n

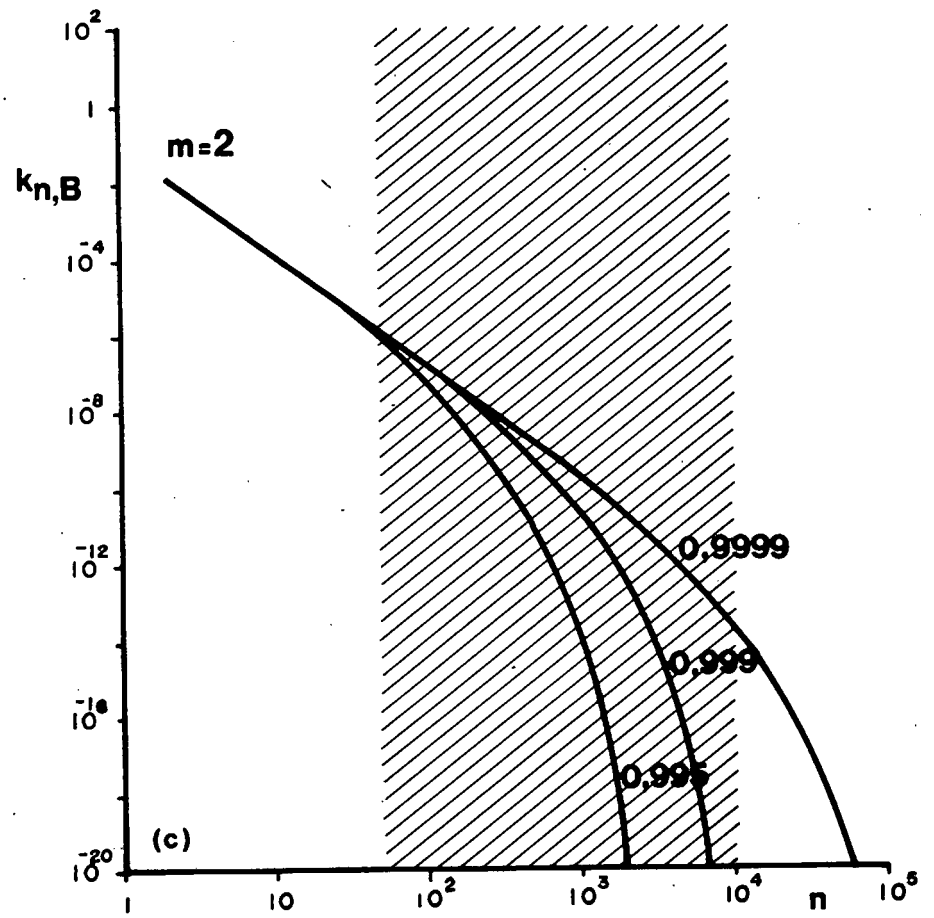

10^{2}
1
m=2
$k_{n,B}$
10^{-4}
10^{-8}
10^{-12}
10^{-16}
10^{-20}
0,9999
0,999
0,995
(c)
1
10
10^{2}
10^{3}
10^{4}
10^{5}
n

Table 5.1: Ratios R_B/R of radius of Bjerhammar sphere to radius of spherical earth as given by *Lelgemann's (1981)* formula, for m = 0,1,2, and mean spacing ψ_m between data points 30', 15', 5'.

m	$o(k_n)$	30'	15'	5'
0	1	0,978	0,989	0,996
1	-1	0,987	0,994	0,998
2	-3	0,996	0,998	0,9993

ψ_m is the mean distance between data points. Table 5.1 gives a selection of results using this formula for the splines above; there seems to be good agreement with ratios inferred from Figure 5.4.

The *Szegö kernel* is obtained when m = 0 in (5.43)or (5.46). The infinite series can be expressed in closed form *(Lelgemann, 1979)*

$$K_0(x_1,x_2) = \sum_{n=0}^{\infty} \frac{2n+1}{4\pi}\left(\frac{R^2}{r_1 r_2}\right)^{n+1} P_n(\xi\cdot\eta) = \frac{r_1^2 r_2^2 - R^4}{R^2 L^3}, \qquad (5.47)$$

or, on introducing a Bjerhammar sphere of radius R_B,

$$K_{0,B}(x_1,x_2) = \frac{r_1^2 r_2^2 - R_B^4}{R_B^2 L^3}, \qquad (5.48)$$

with

$$L^2 = \frac{r_1^2 r_2^2}{R_B^4} - \frac{2 r_1 r_2}{R_B^2}(\xi\cdot\eta) + 1; \qquad (5.49)$$

in the case of observations on a sphere of radius R

$$K_{0,B}(\xi\cdot\eta) = \frac{R^4 - R_B^4}{R_B^2 L^3}, \tag{5.50}$$

$$L^2 = \frac{R^4}{R_B^4} - 2\frac{R^2}{R_B^2}(\xi\cdot\eta) + 1. \tag{5.51}$$

Point-mass splines: There are other choices of degree variances also giving kernels that can be expressed in closed form - these are the kernel functions of *analytical collocation (Moritz, 1980; Lelgemann, 1979, 1980, 1981), Tscherning, 1986).* Only one example will be considered here; with

$$k_n = \frac{2}{2n+1}$$

the *Krarup kernel* is obtained

$$K_{K,B}(x_1,x_2) = \frac{2}{L}, \tag{5.52}$$

with L given by (5.51) for data on a spherical earth, or more generally by (5.49). Further, since

$$L^2 = \frac{R^4}{R_B^4}\left(1 - 2\frac{R_B^2}{R^2}\cos\psi + \frac{R_B^4}{R^4}\right) = \frac{R^4}{R_B^4}\,\ell^2,$$

or

$$L^2 = \frac{r_1^2 r_2^2}{R_B^4}\left(1 - 2\frac{R_B^2}{r_1 r_2}\cos\psi + \frac{R_B^4}{r_1^2 r_2^2}\right) = \frac{r_1^2 r_2^2}{R_B^4}\,\ell^2,$$

the distance ℓ (as shown in Figure 5.5) in the second case is the distance to a point mass proportional to r_2/R_B^2 buried relatively

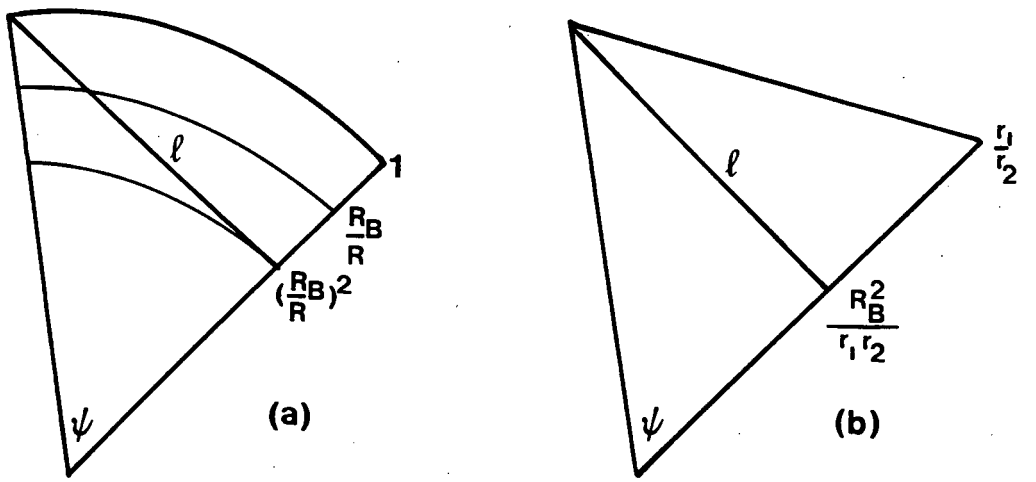


Figure 5.5: Position of mass sources for point-mass splines (a) on a spherical earth; and (b) in the general case.

R_B^2/r_2^2 below the observation point x_2 (cf. *Tscherning, 1986, p 17*); in the first case all point masses are buried relatively R_B/R below the Bjerhammar sphere, or $(R_B/R)^2$ below the surface of a spherical earth. It is thus appropriate to refer to (5.52) as *point-mass splines*.

CHAPTER 6

SPLINE APPROXIMATION OF THE LOCAL GRAVITY FIELD

6.1 Introduction

The application of the theory of spline approximation to the problem of finding local approximations to the earth's gravity field is now relatively straightforward, at least for the simple, essentially two-dimensional models to be considered in this chapter. Computationally, little more is involved than the solution of a set of linear equations. Splines on $\mathbb{R}^2$, and the spherical splines of the previous chapter, will be seen to give results very much the same as for least squares collocation.

Among the applications to be looked at is geoid interpolation. The need to determine orthometric height differences from GPS-derived ellipsoidal heights has made local geoid determination a topical problem. Relative geoid heights may be obtained from gravity anomalies *(Kearsley, 1986; Schwarz et al, 1987)*, but there is also scope for interpolation, either directly, or with supplementary gravity data, of geoid heights determined from comparison of GPS ellipsoidal heights, and spirit-levelled orthometric or normal heights, in the way suggested by *King et al (1985), p 96*. Direct interpolation is the subject of the following section, comparing the efficacy for geoid interpolation of a number of spline techniques. Section 6.4 shows that the inclusion of relatively few gravity anomalies may improve materially the interpolation of geoid heights. Section 6.3 is given over to the more traditional task of interpolating free-air gravity anomalies; since gravity

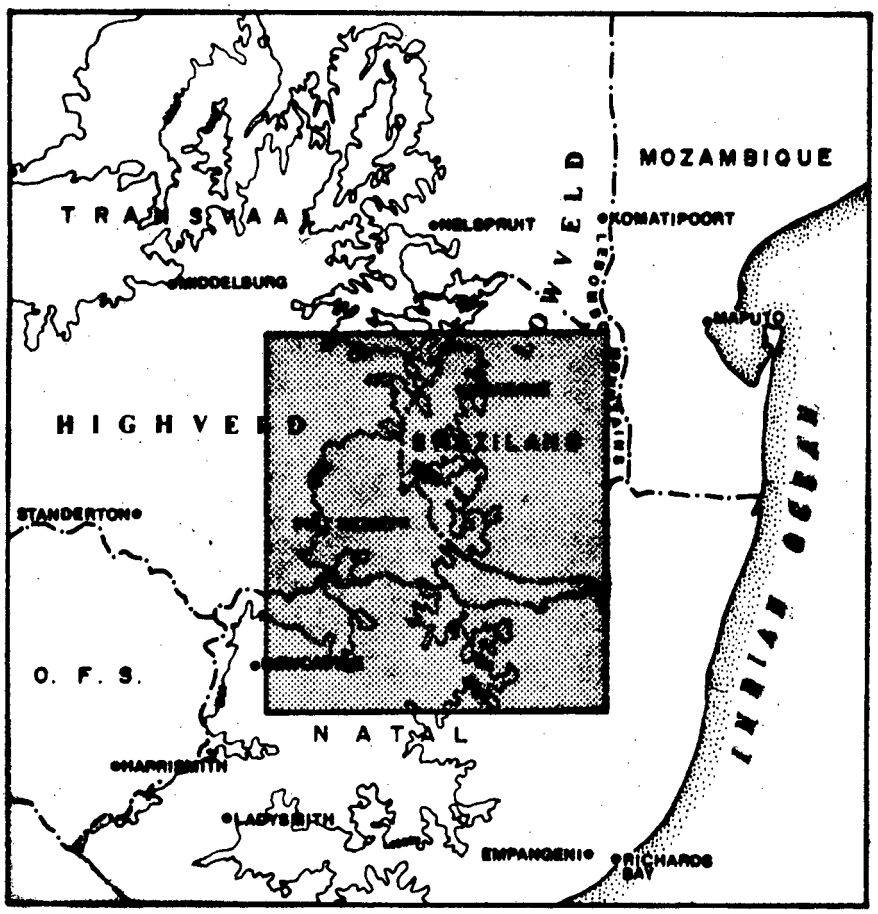
anomalies show more variation, this may be a more demanding test.

The study area for these interpolation tests is an area 2°x2° in extent in northern Natal, south-eastern Transvaal and Swaziland (Figure 6.1(a)). In the west the area straddles the southern African escarpment, and along its eastern edge are the Lebombo moumtains. This mountain range, running along the Swaziland-Mozambique border, though no higher than 700m, is gravimetrically very sharply defined (Figure 6.1(c)). The area between, in the Lowveld of Swaziland and eastern Transvaal, is a complex mix of volcanic rocks and metamorphosed sediments of the Precambrian and Mesozoic periods, with gravity anomalies little correlated with elevation, and somewhat difficult to model *(Merry, 1980)*.

The geoid heights for this area are those of the UCT86 geoid *(van Gysen and Merry, 1987)*. This geoid combines the OSU81 geopotential model and free-air gravity anomalies for southern Africa using a spectral weighting technique, with an a priori spectral weighting function. The known medium-wavelength errors in the geoid *(Merry and van Gysen, 1987)* should not much impair the use of the geoid heights shown in Figure 6.1(b).

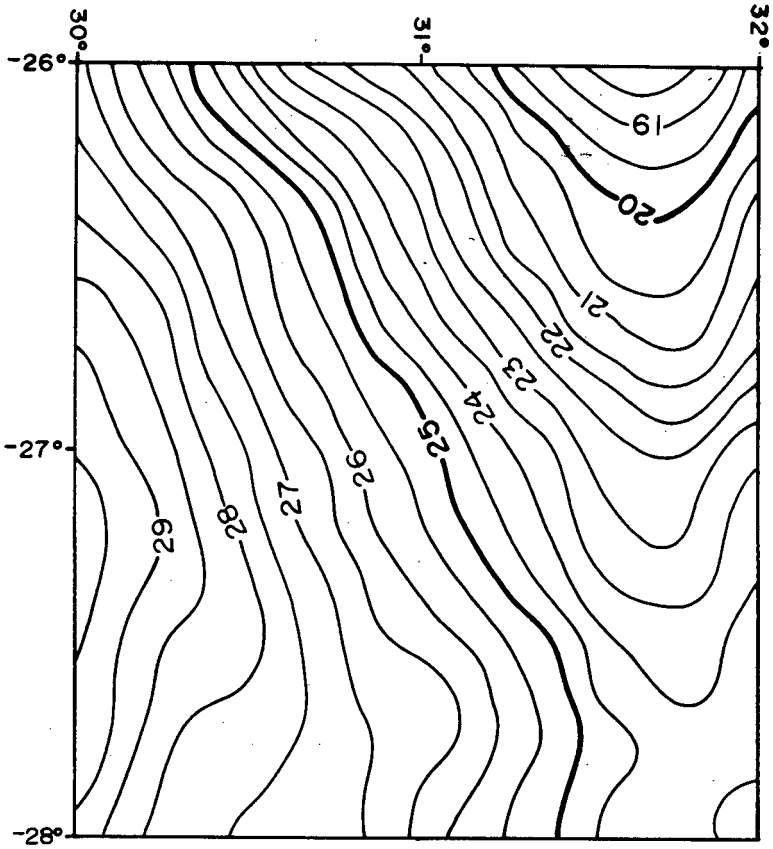
Section 6.5 deals with geoid interpolation of a different kind; a pseudo-cubic geoid for southern Africa is given, which interpolates both Doppler-derived geoid heights and astro-geodetic deflections of the vertical.

The final case study looks at spline approximation of linear functionals - at the thin-plate spline quadrature of several geodetic integrals. The contribution, to a planar approximation, of a circular innermost zone, is evaluated for Stokes's formula, for the formulae of Vening Meinesz, and the gravity anomaly gradient $\frac{\partial \Delta g}{\partial h}$ used in Molodensky's solution of the geodetic boundary value problem.

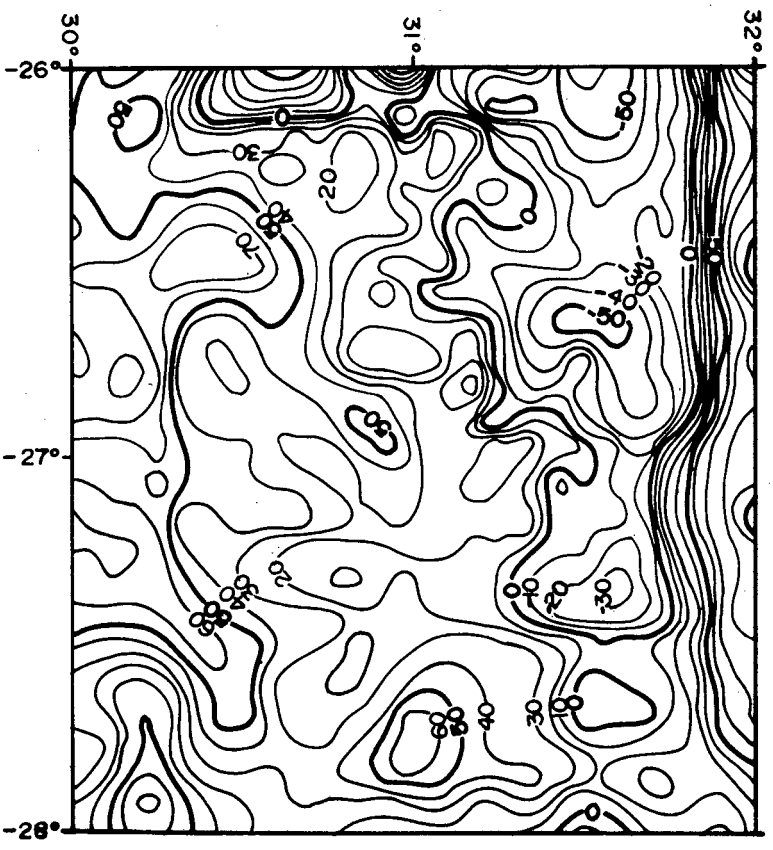


(a)

Figure 6.1: (a) Locality of the 2°x2° study area for the geoid and gravity anomaly interpolation of Sections 6.2-6.4; (b) (overleaf) quasi-geoid heights in the study area (contour interval 1m); (c) free-air gravity anomalies in the study area (contour interval 10 milligal).



(b)



(c)

6.2 Spline Geoid Interpolation

In this first test different spline interpolants of the geoid are compared with least squares collocation. For rotation-invariant splines on Euclidean spaces - those taken from Chapter 4 - the interpolation takes place in $\mathbb{R}^2$, with a *plate carree* projection of geodetic positions,

$$x = (x_1, x_2) = (R\phi, R\lambda\cos\phi_0) \tag{6.1}$$

(where ϕ_0 is an average latitude, and R the radius of a spherical earth). Geoid heights are supposed given at N points $\alpha_i = (\alpha_{1i}, \alpha_{2i})$.

The following splines are considered:

(i) *Multi-conic splines*, of the form

$$s(x) = \sum_{i=1}^{N} a_i |x - \alpha_i| + b \tag{4.75}$$

interpolating the given geoid heights, subject to the end condition

$$\sum_{i=1}^{N} a_i = 0.$$

(ii) *Hardy's multi-quadric functions*, of the form

$$s(x) = \sum_{i=1}^{N} a_i \{|x - \alpha_i|^2 + \varepsilon^2\}^{1/2}, \tag{4.83}$$

with $\varepsilon^2 = 0{,}086$, a value loosely based on the recommendation of *Hardy (1977)*. The results turn out to be relatively insensitive to

the exact choice of ε.

(iii) *Thin-plate splines,*

$$s(x) = \sum_{i=1}^{N} a_i |x-\alpha_i|^2 \log_e |x-\alpha_i| + b_0 + b\cdot x, \tag{4.72}$$

satisfying the interpolatory constraints, and the end conditions

$$\sum_{i=1}^{N} a_i = 0; \quad \sum_{i=1}^{N} a_i \alpha_i = 0 \ (\text{in } \mathbb{R}^2).$$

(iv) *Pseudo-cubic splines,* of the form

$$s(x) = \sum_{i=1}^{N} a_i |x-\alpha_i|^3 + b_0 + b\cdot x, \tag{4.78}$$

interpolating the data, and satisfying the same end conditions as the thin-plate splines, since both have $\mathbb{P}_1$ as nullspace.

(v) *Pseudo-quintic splines,* of the form

$$s(x) = \sum_{i=1}^{N} a_i |x-\alpha_i|^5 + b_0 + b_1 x_1 + b_2 x_2 + b_3 x_1^2 + b_4 x_2^2 + b_5 x_1 x_2, \tag{6.2}$$

with end conditions

$$\sum_{i=1}^{N} a_i = 0,$$

$$\sum_{i=1}^{N} a_i \alpha_{1i} = 0,$$

$$\sum_{i=1}^{N} a_i \alpha_{2i} = 0,$$

$$\sum_{i=1}^{N} a_i \alpha_{1i}^2 = 0,$$

$$\sum_{i=1}^{N} a_i \alpha_{2i}^2 = 0,$$

$$\sum_{i=1}^{N} a_i \alpha_{1i} \alpha_{2i} = 0,$$

or

$$\sum_{i=1}^{N} a_i = 0,$$

$$\sum_{i=1}^{N} a_i \alpha_i = 0,$$

$$\sum_{i=1}^{N} a_i \alpha_i \alpha_i^t = 0.$$

(vi) *Spherical splines*, $m = 2$,

$$s(x) = \sum_{i=1}^{N} a_i K_{2,B}(\xi \cdot \eta_i) + b \tag{6.3}$$

($x = R\xi$, $\alpha_i = R\eta_i$) derived from (5.37), with

$$K_{2,B}(\xi \cdot \eta_i) = \sum_{n=2}^{\infty} \frac{2n+1}{4\pi} \{n(n+1)\}^{-2} \left(\frac{R_B}{R}\right)^{2n+2} P_n(\xi \cdot \eta_i), \tag{6.4}$$

and subject to

$$\sum_{i=1}^{N} a_i = 0.$$

The ratio R_B/R was chosen as 0,9999 on the basis of the discusssion

in Section 5.5

(vii) *Least squares collocating splines,*

$$s(x) = \sum_{i=1}^{N} a_i K_{col,B}(\xi \cdot \eta_i), \tag{6.5}$$

with

$$K_{col,B}(\xi \cdot \eta_i) = \sum_{n=3}^{\infty} \{(n-1)(n-2)(n+24)\}^{-1} \left(\frac{R_B}{R}\right)^{2n+2} P_n(\xi \cdot \eta_i), \tag{6.6}$$

where the degree variances are proportional to the values of Tscherning and Rapp's 1974 model for global geopotential degree variances *(Moritz, 1980, p 190)*, but with $R_B/R = 0,9999$ rather than 0,9996.

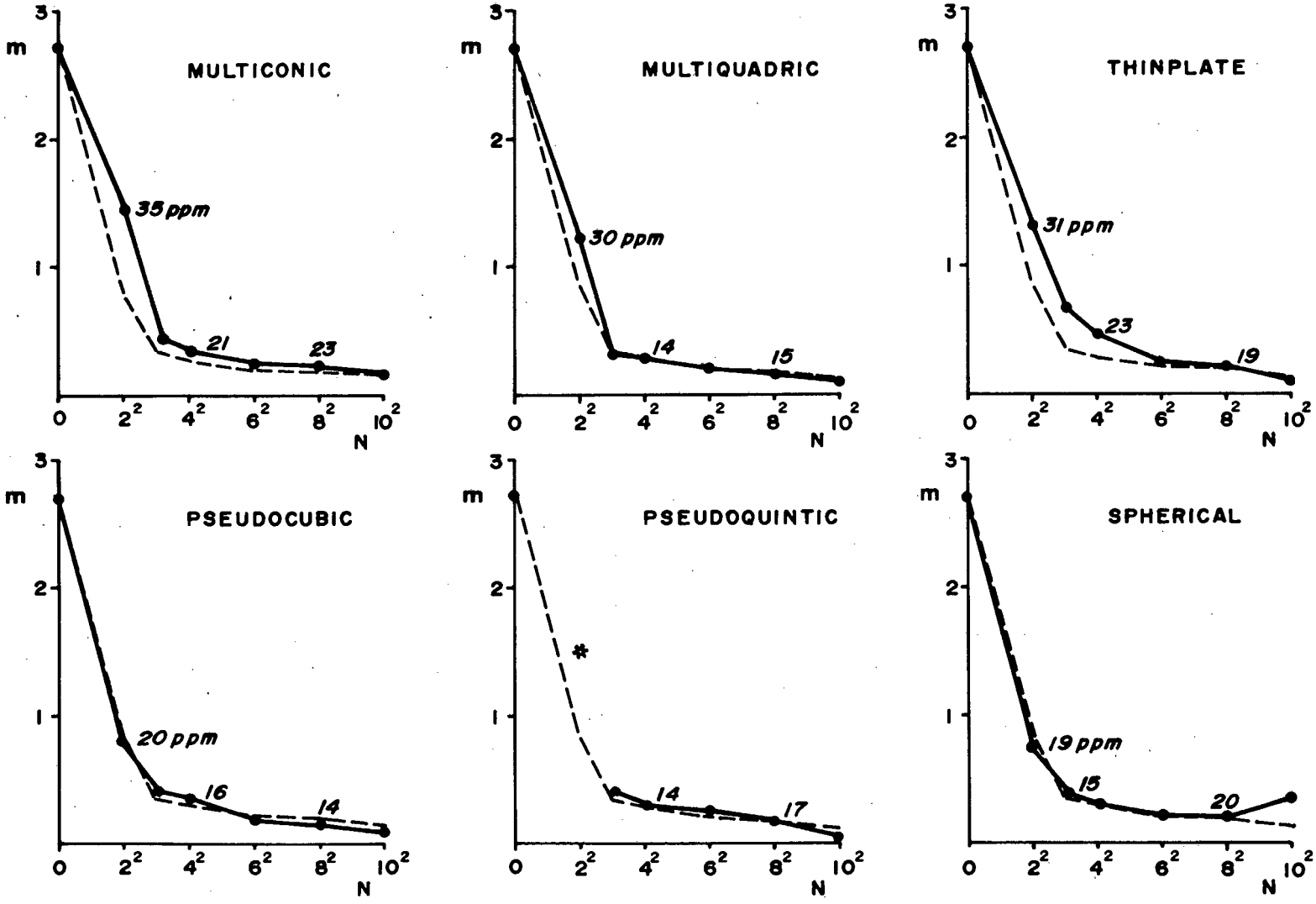
Given geoid heights (on a 5'x5' grid) in the area of Figure 6.1 were uniformly randomly sampled for N = 2, 16, 36, 64 and 100, with the sampling being done using the NAG routine G05EJE *(Numerical Algorithms Group, 1983)*. For each sample the spline representation was found by solving the linear equations (4.68) or (5.40) (using the NAG Crout factorisation routine F04ARE), rather than the positive-definite equations (4.94) consequent to using a reproducing kernel function (4.91) or (5.41).

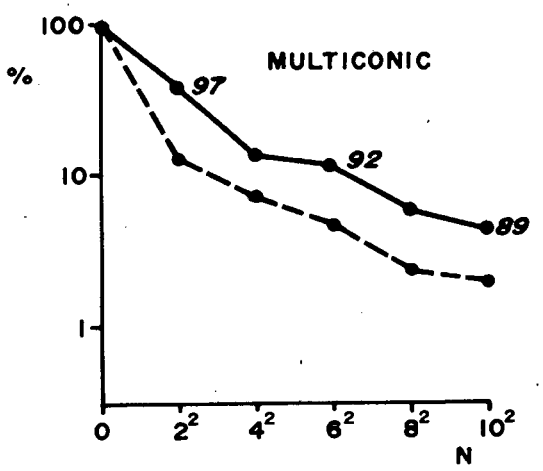
The fidelity of the representation is measured by the root-mean-square (rms) discrepancy against all 5' geoid heights in the 2°x2° sampling area. Figure 6.2(a) shows the rms error (in metres) for each of the six splines (i) to (vi) above, against the corresponding rms error for least squares collocation (vii) (shown by a dashed line). Shown

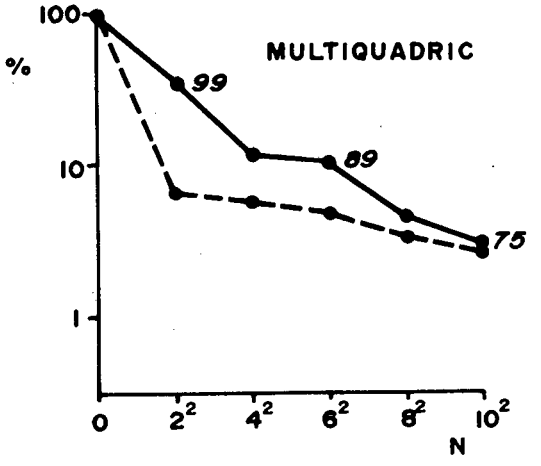
in the figure for N = 0 is the rms variation of the geoid.

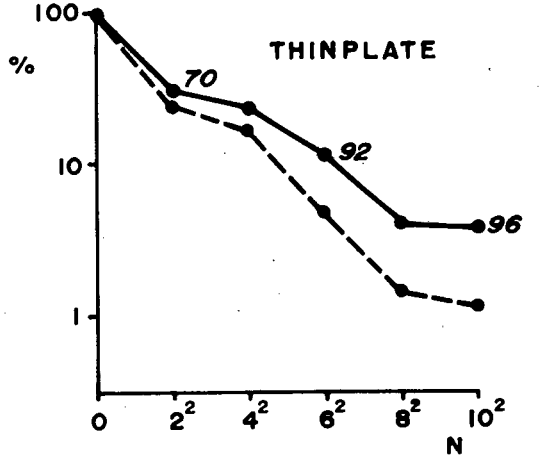
Each technique compares favourably with the collocation solution, the more so as the sample size increases. For N = 4 there is no pseudo-quintic solution as the nullspace alone requires six coefficients. For N = 100 the spherical spline begins to diverge, indicating perhaps, the incipient instability in the linear equations (5.40) *Dyn et al (1986)* warn against, and aggravated by the interpolation of $K_{2,B}$ in (6.4) from a look-up table (with the interpolation being done with a forward difference formula, rather than using a cubic spline, say - contrary to the spirit of the rest of my work).

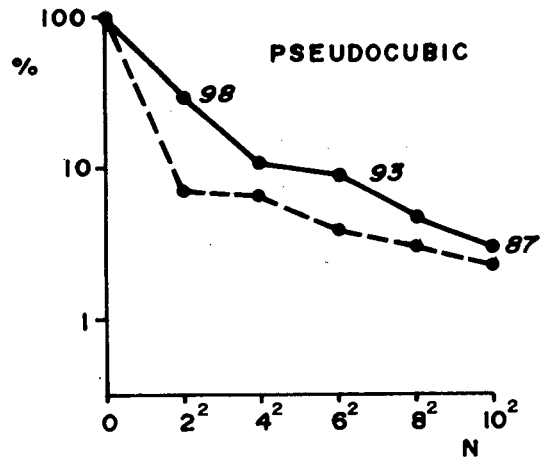
Figure 6.2: (following pages) (a) Rms error (in metres) in representing geoid heights in the area of Figure 6.1, for (i) multi-conic splines, (ii) multi-quadric functions, (iii) thin-plate splines, (iv) pseudo-cubic splines, (v) pseudo-quintic splines, and (vi) spherical splines, m=2. The dashed lines indicate the corresponding rms errors for least squares collocation. Figures in italics give the relative representation error (relative to the closest interpolation point) in parts per million;
(b) predictive power for the same splines, of geoid heights (solid lines), and of least squares predicted heights (dashed lines). Figures in italics give percentage error correlation between spline and least squares representations.

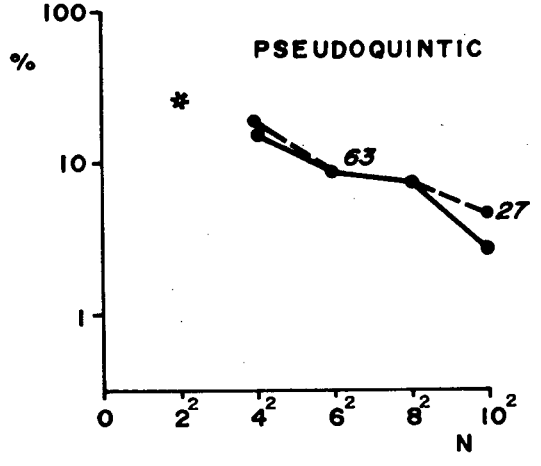

MULTICONIC
35 ppm
21
23
MULTIQUADRIC
30 ppm
14
15
THINPLATE
31 ppm
23
19
PSEUDOCUBIC
20 ppm
16
14
PSEUDOQUINTIC
*
14
17
SPHERICAL
19 ppm
15
20
m
N
0
2^2
4^2
6^2
8^2
10^2
1
2
3


MULTICONIC
100
%
10
1
97
92
89
0
2^2
4^2
6^2
8^2
10^2
N


MULTIQUADRIC
100
%
10
1
99
89
75
0
2^2
4^2
6^2
8^2
10^2
N


THINPLATE
100
%
10
1
70
92
96
0
2^2
4^2
6^2
8^2
10^2
N


PSEUDOCUBIC
100
%
10
1
98
93
87
0
2^2
4^2
6^2
8^2
10^2
N


PSEUDOQUINTIC
100
%
10
1
*
63
27
0
2^2
4^2
6^2
8^2
10^2
N

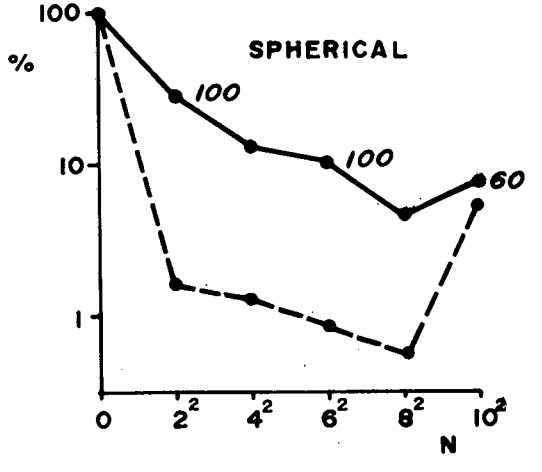

SPHERICAL
100
%
10
1
100
100
60
0
2^2
4^2
6^2
8^2
10^2
N

Figure 6.2(a) also gives an average value for the representation error (error in geoid height relative to distance to the closest interpolation point) expressed in parts per million, indicating that all representations seem to converge at a rate that is approximately linear in relation to the distance between data points.

The suggestion that various splines may be as effective overall as a least squares collocating spline is pursued further in Figure 6.2(b). The solid lines indicate the 'power of prediction' for each technique, i.e. the ratio, expressed as a percentage, between the rms variation in the signal (geoid heights), and the rms error in prediction. Each technique does fairly well. The dashed lines show the power with which each spline predicts least squares collocated heights. It is interesting to see that, with the exception of pseudo-quintic splines, each spline represents the least squares predicted geoid better than the actual geoid, with the $m = 2$ spherical splines doing best (except for $N = 100$), and thin-plate splines also doing well. The figures in italics give the percentage correlation between errors in spline geoid heights and corresponding errors in collocated heights, again pointing to close agreement between each spline technique and least squares collocation (with pseudo-quintic splines excepted).

The reason for this agreement is not far to seek. The geopotential degree variances (6.6) fall off in the main as n^{-3} (and eventually ever faster as the term $(R_B/R)^{2n+2}$ comes to dominate), as do the spherical spline degree variances (6.4). *Schwarz (1984)* indicates a fall-off closer to n^{-4}, and in areas of little topography the

geopotential degree variances fall off even faster, e.g. as $n^{-5,6}$ in southern Ohio *(Arabelos and Tziavos, 1987)*. It can be seen from Table 4.1 (p 98) that, in $\mathbb{R}^2$, the Fourier transform of the multi-conic kernel mapping is proportional to $|x|^{-3}$, for thin-plate splines it is $|x|^{-4}$, for pseudo-cubic splines $|x|^{-5}$, and for pseudo-quintic splines $|x|^{-7}$. The efficacy of thin-plate splines in representing the geoid is thus not surprising, nor that multi-conic (and hence also multi-quadric functions) and pseudo-cubic splines are not much less effective.

6.3 Spline Interpolation of Free-Air Gravity Anomalies

The free-air gravity anomalies of Figure 6.1(c) reflect the complex geology of the region. They are little correlated with elevation, for the rms variation of the anomalies is 30,6 milligal, and the rms error in predicting gravity from linear correlation with elevation is 24,5 milligal. This is thus not an area in which the gravity anomaly field is easily modelled; with a sample of 144 anomalies, gravity anomalies can be modelled with an rms error little better than 7,5 milligal. Representation of this highly variable data should be a good indicator of the relative performance of different splines.

As in the previous section, the interpolation by the splines of Chapter 4 takes place in the plane. The usual technique is followed of removing the topography to get Bouguer gravity anomalies, interpolating these anomalies, and restoring the topography (despite the weak 60% correlation of gravity anomalies with elevation).

The following splines are used, interpolating given gravity anomalies at N points $\alpha_i = (\alpha_{1i}, \alpha_{2i}) \in \mathbb{R}^2$:

(i) *multi-conic splines* (4.75);

(ii) *multi-quadric functions* (4.83), with the same regularising value $\varepsilon^2 = 0{,}086$ as in the previous section;

(iii) *thin-plate splines* (4.72);

(iv) *pseudo-cubic splines* (4.78);

(v) *spherical splines*, with $m = 1$, now with $\alpha_i \in \Omega$, $i = 1, \ldots, N$, and $x = R\xi$, $\alpha_i = R\eta_i$,

$$s(x) = \sum_{i=1}^{N} a_i K_{1,B}(\xi \cdot \eta_i) + b. \tag{6.7}$$

The kernel function is

$$K_{1,B}(\xi \cdot \eta_i) = \sum_{n=2}^{\infty} \frac{2n+1}{4\pi} \{n(n+1)\}^{-1} \left(\frac{R_B}{R}\right)^{2n+2} P_n(\xi \cdot \eta_i). \tag{6.8}$$

If degree variances of the disturbing potential fall off as n^{-3}, those for the corresponding gravity anomalies fall off as n^{-1}, hence the choice of $m = 1$, whereas in the previous section $m = 2$ (and hence also the need for regularisation). The Bjerhammar sphere that seems to work best is at a relative depth $R_B/R = 0{,}99$, in conformity with what *Lelgemann's (1981)* rule would suggest.

If $m = 2$ spherical splines show some numerical instability with increasing numbers of data, this is even more likely for spherical splines with $m = 1$, since the function $K_{1,B}$ is more sharply peaked. A small degree of smoothing is thus introduced, with the 'ridging' parameter of Section 4.11 given the value $\frac{1}{\varepsilon} = 0{,}005$.

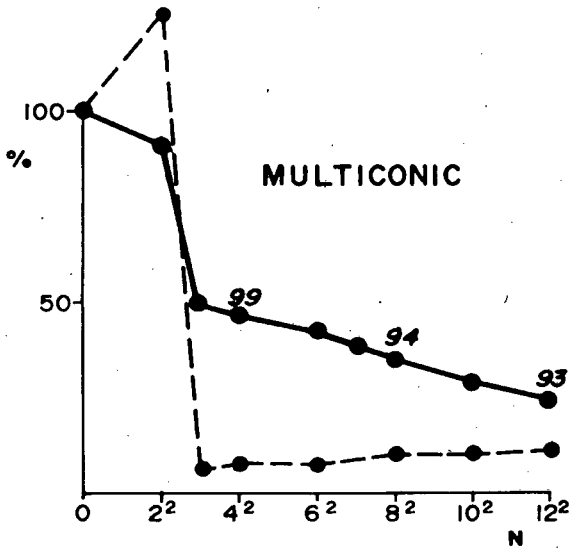
(vi) *Least squares collocating splines* (6.5), with degree variances

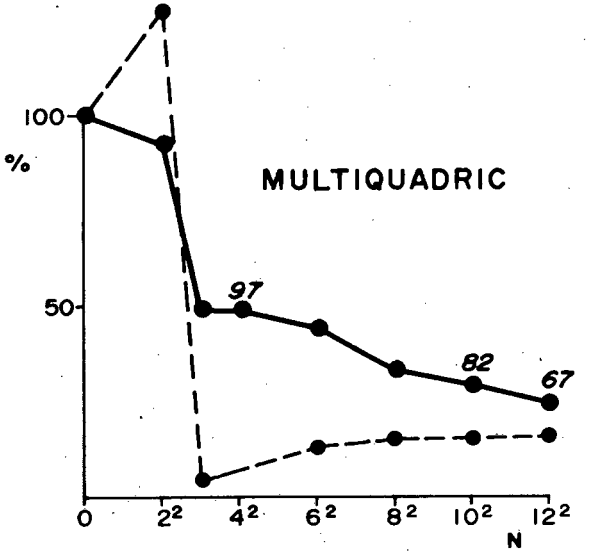
$$K_{col,B}(\xi \cdot \eta_i) = \sum_{n=3}^{\infty} \frac{(n-1)}{(n-2)(n+24)} \left(\frac{R_B}{R}\right)^{2n+2} P_n(\xi \cdot \eta_i), \qquad (6.9)$$

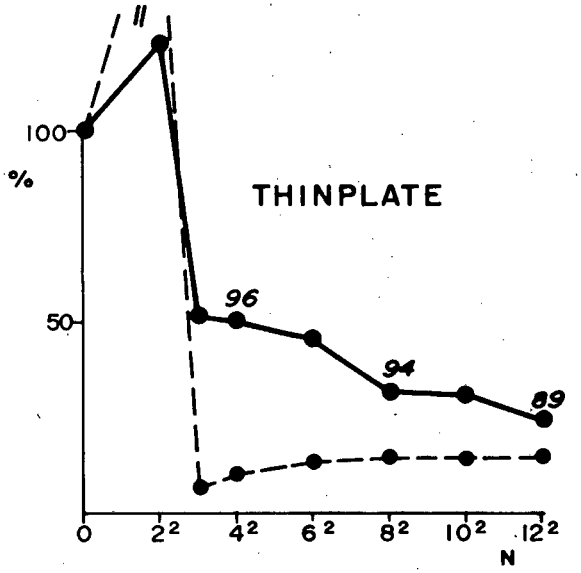
and $R_B/R = 0{,}99$, as in (v).

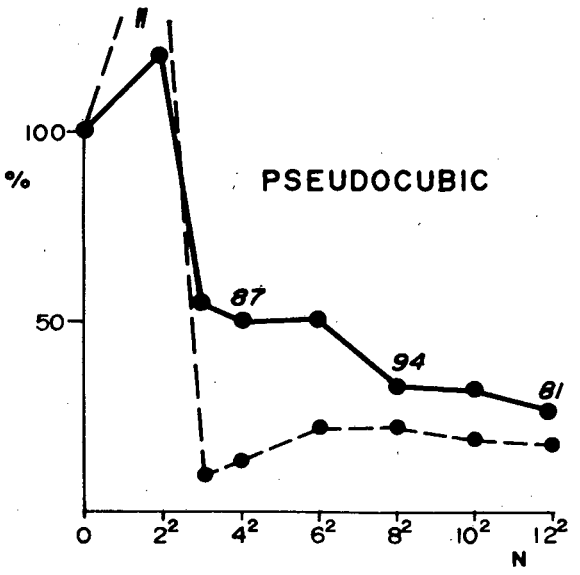
Pseudo-quintic splines are not included; they perform appreciably more poorly in relation to other spline functions than for geoid interpolation. The procedure of the last section - randomly sampling from a 5'x5' data field - is again followed, with $N = 4$, 9, 16, 36, 64, 100 and 144. Figure 6.3 shows how the results compare. With some fluctuations for small samples at $N = 4$ and 9, the different spline techniques seem to give very similar results. The improvement as the

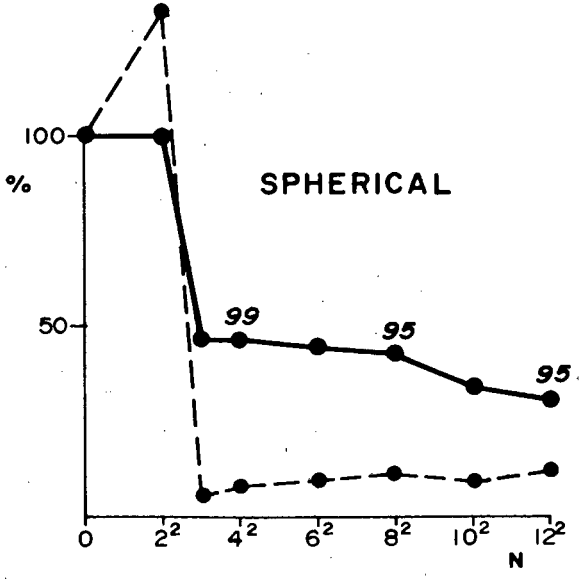
Figure 6.3: (following page) Predictive power (expressed as a percentage) for the indicated spline functions, of free-air gravity anomalies (solid lines), and of least squares predicted anomalies (dashed lines). The figures in italics give the percentage error correlation between spline and least squares representations.

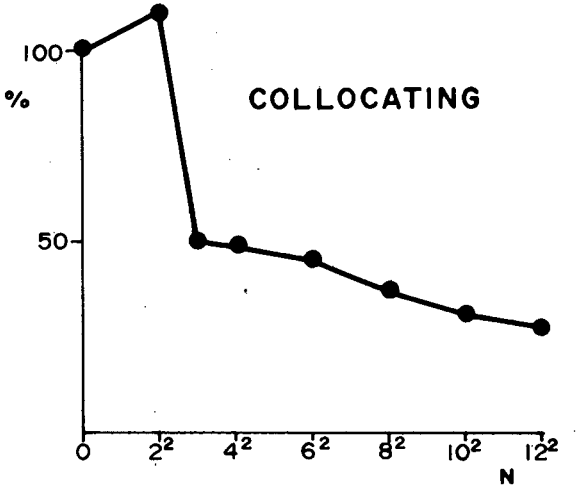

MULTICONIC
100
%
50
99
94
93
0
2^2
4^2
6^2
8^2
10^2
12^2
N


MULTIQUADRIC
100
%
50
97
82
67
0
2^2
4^2
6^2
8^2
10^2
12^2
N


THINPLATE
100
%
50
96
94
89
0
2^2
4^2
6^2
8^2
10^2
12^2
N


PSEUDOCUBIC
100
%
50
87
94
81
0
2^2
4^2
6^2
8^2
10^2
12^2
N


SPHERICAL
100
%
50
99
95
95
0
2^2
4^2
6^2
8^2
10^2
12^2
N


COLLOCATING
100
%
50
0
2^2
4^2
6^2
8^2
10^2
12^2
N

number of data points increases is slower than with geoid heights. In each case spline interpolation agrees better with the least squares predicted anomalies than with the actual anomalies. This is especially evident with the multi-conic and spherical splines. The figures in italics in the figure giving percentage error correlation between spline and least squares errors in representation further strengthen the conclusion that the splines (i) to (v) above interpolate gravity anomalies in very much the same way as least squares collocation.

6.4 Spline Geoid Interpolation with Gravity Anomalies

Spline functions need not interpolate data of one type only. In geoid interpolation - for example, interpolating geoid heights determined from comparison of GPS-derived ellipsoidal and spirit-levelled orthometric or normal heights - gravity anomalies may contribute significantly to improving the representation. This will be indicated at the hand of three examples, involving multiquadric and spherical splines, and least squares collocation.

(i) *Multiquadric functions:* As in Section 6.2, $x = (x_1,x_2) \in \mathbb{R}^2$; geoid heights N_i are given at points $\alpha_i = (\alpha_{1i},\alpha_{2i})$, $i = 1,\ldots,N$. A further M gravity anomalies Δg_j are given at the points $\beta_j = (\beta_{1j},\beta_{2j})$. The two-dimensional approach of the previous sections is continued in representing the geoid as

$$s(x) = \sum_{i=1}^{N} a_i d_\varepsilon(x,\alpha_i) + \frac{1}{\gamma}\sum_{j=1}^{M} b_j\{\varepsilon d_\varepsilon(x,\beta_j)^{-1} - \frac{2}{R_B} d_\varepsilon(x,\beta_j)\}, \qquad (6.10)$$

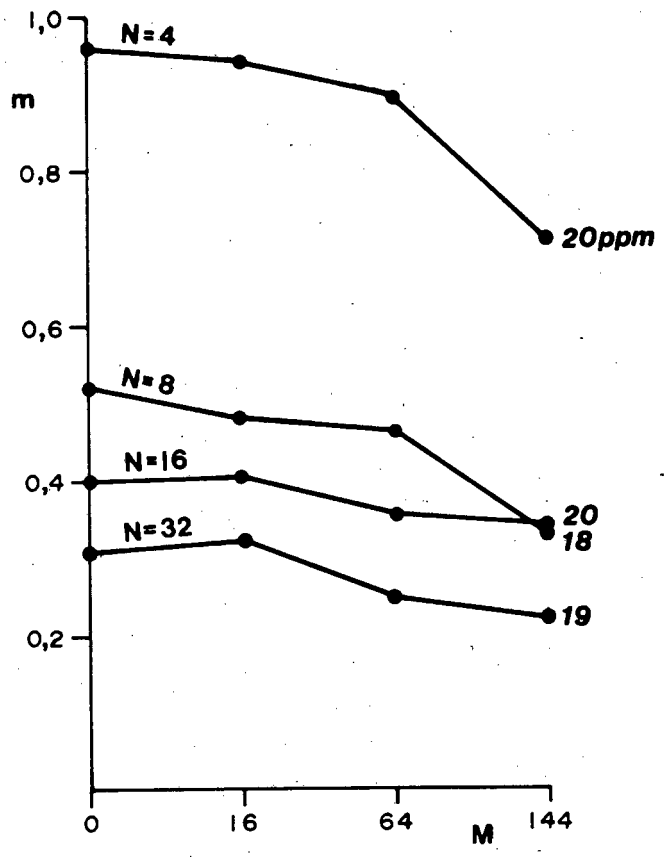


Figure 6.4: Rms errors (in metres) in multiquadric representations of the geoid in Figure 6.1(b), interpolating N geoid heights and smoothing M gravity anomalies. Figures in italics give the relative representation error (relative to the closest interpolation point).

were drawn from a wider area; these are thus uniformly sampled from a 4°x4° area centred on the 2°x2° one. A small degree of smoothing is necessary to obviate the type of divergence seen in Figure 6.2. Only gravity anomalies are smoothed, with the 'ridging' parameter $\frac{1}{\sigma}$ given a value of 0,05. Changing this value by several hundred percent leads to little change in the numerical results. The results are more sensitive to the value of the regularising parameter - a value of ϵ corresponding to $R_B/R = 0,998$ seems to work best, although the results do not change much for a 50% variation in ϵ.

Adding gravity anomalies leads to a modest improvement in the representation of the geoid, and more so for small numbers of given geoid heights (N = 4 and 8). That there should be any improvement at all is surprising. When introducing observations other than function evaluations - in this case, introducing gravity anomaly evaluations as well - it becomes apparent that multiquadric functions (as also the splines in Table 4.1) are not the most natural building blocks for geodetic models. Moreover, it is only through regularisation that gravity anomalies are allowed as data at all. In this regard spherical splines are better suited for representing the geoid using geoid and gravity anomaly data.

(ii) *Spherical splines:* The m = 2 spline interpolating both geoid heights at points $\alpha_i = R\eta_i$, $\eta_i \in \Omega$, $i = 1,\ldots,N$, on a spherical earth of radius R, and gravity anomalies at points $\beta_j = R\zeta_j$, $\zeta_j \in \Omega$, $j = 1,\ldots,M$, has the form, for $x = R\xi$,

$$s(x) = \sum_{i=1}^{N} a_i K_{2,B}(\xi \cdot \eta_i) + \frac{1}{\gamma R_B} \sum_{j=1}^{M} b_j \Delta_g K_{2,B}(\xi \cdot \zeta_j) + c_0, \qquad (6.14)$$

where

$$\Delta_g k_{n,B} = (n-1) k_{n,B}, \quad n = 1,2,\ldots \qquad (6.15)$$

relates (except for the otherwise usual factor R) the 'degree variances' for the spherical spline representation of the disturbing potential to the 'degree cross-covariances' between disturbing potential and gravity anomalies. The coefficients of (6.14) are found from the data by solving the linear equations

$$\begin{aligned} &\sum_{i=1}^{N} a_i K_{2,B}(\eta_k \cdot \eta_i) + \frac{1}{\gamma R_B} \sum_{j=1}^{M} b_j \Delta_g K_{2,B}(\eta_k \cdot \zeta_j) + c_0 = N_k, \\ &\qquad\qquad k = 1,\ldots,N, \\ &\frac{1}{\gamma R} \sum_{i=1}^{N} a_i \Delta_g K_{2,B}(\zeta_\ell \cdot \eta_i) + \frac{1}{\gamma^2 R R_B} \sum_{j=1}^{M} b_j \{\Delta_g^2 K_{2,B}(\zeta_\ell \cdot \zeta_j) + \frac{1}{\sigma} \delta_{\ell j}\} - \frac{2}{\gamma R} c_0 = \Delta g_\ell, \\ &\qquad\qquad \ell = 1,\ldots,M, \end{aligned} \qquad (6.16)$$

subject to the end condition

$$\sum_{i=1}^{N} a_i - \frac{2}{\gamma R} \sum_{j=1}^{M} b_j = 0.$$

Representations of the geoid in Figure 6.1(b), with rms errors as given in Figure 6.5, are obtained in the same way as for multiquadric functions, with N geoid heights sampled from the 2°x2° study area, and M gravity anomalies from a larger 4°x4° area. The gravity anomalies are smoothed, with the ridge parameter set to 0,3, though there is a 50% latitude in choosing this figure in which range results change only

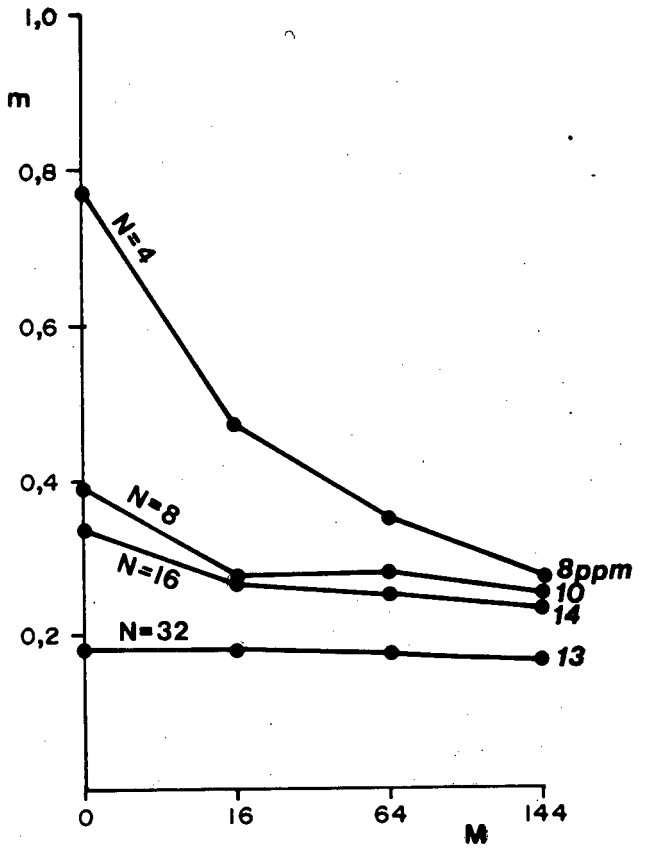


Figure 6.5: Rms errors (in metres) in m = 2 spherical spline representations of the geoid in Figure 6.1(b), interpolating N geoid heights and smoothing M gravity anomalies. ($R_B/R = 0,995$, $1/\sigma = 0,3$). Figures in italics give the relative representation error (in parts per million).

slightly. The ratio R_B/R was set to 0,995.

Including gravity anomaly data leads to an improved spherical spline representation of the geoid, particularly for small numbers of given geoid heights; for larger numbers the improvement is small. Spherical splines perform almost as well as least squares collocation:

(iii) *Least squares collocation:* Here the representation is

$$s(x) = \sum_{i=1}^{N} a_i K_{col,B}(\xi \cdot \eta_i) + \frac{1}{\gamma R_B} \sum_{j=1}^{M} b_j \Delta_g K_{col,B}(\xi \cdot \zeta_j), \tag{6.17}$$

with $K_{col,B}$ given by (6.6), and Δ_g as in (6.15). The coefficients of the representation are obtained from

$$\begin{aligned} &\sum_{i=1}^{N} a_i K_{col,B}(\eta_k \cdot \eta_i) + \frac{1}{\gamma R_B} \sum_{j=1}^{M} b_j \Delta_g K_{col,B}(\eta_k \cdot \zeta_j) = N_k, \\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad k = 1, \ldots, N, \\ &\frac{1}{\gamma R} \sum_{i=1}^{N} a_i \Delta_g K_{col,B}(\zeta_\ell \cdot \eta_i) + \frac{1}{\gamma^2 R R_B} \sum_{j=1}^{M} b_j \{\Delta_g^2 K_{col,B}(\zeta_\ell \cdot \zeta_j) + \frac{1}{\sigma}\delta_{\ell j}\} = \Delta g_\ell, \\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \ell = 1, \ldots, M. \end{aligned} \tag{6.18}$$

In interpolating the geoid of Figure 6.1(b) - in the same way as for spherical splines above - the depth of the Bjerhammar sphere was set at $R_B/R = 0,995$, and the ridge parameter, smoothing gravity anomalies only, at $\frac{1}{\sigma} = 0,2$. Figure 6.6 shows the rms errors in representing the geoid. There is a substantial improvement when N is small, and the relative improvement (not surprisingly) is less as N increases.

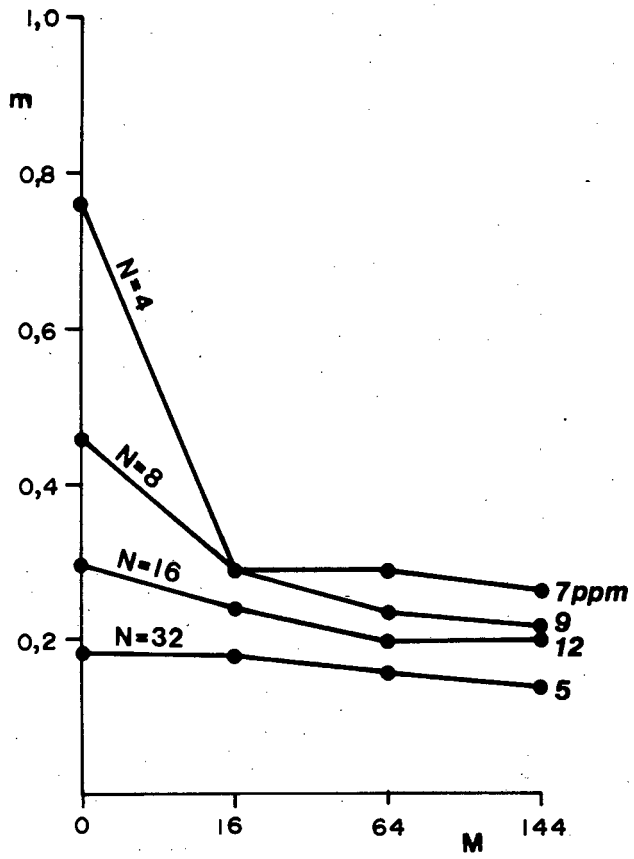


Figure 6.6: Rms errors (in metres) in least squares collocation representations of the geoid in Figure 6.1(b), interpolating N geoid heights and smoothing M gravity anomalies. ($R_B/R = 0,995$, $1/\sigma = 0,2$). Figures in italics give the relative representation error (in parts per million).

To summarise: the inclusion of a small number of gravity anomalies may contribute significantly to the interpolation of geoid heights. Multiquadric functions were able to represent the geoid in the study area with a relative accuracy of about 2 cm per kilometre from the nearest interpolation point, irrespective of the number of geoid points. Least squares collocation does twice as well, at about 1 cm per kilometre, and spherical splines with $m = 2$ only slightly worse.

6.5 Combined Doppler and Astro-Geodetic Geoid for Southern Africa

Using the splines of Chapter 4 for interpolating mixed data is on firmer ground when the data are essentially geometrical. This is the case for geoid representations with data in the form of geoid heights and astro-geodetic deflections of the vertical. This section describes a geoid interpolating these two data types. The geoid heights are obtained as differences between ellipsoidal heights derived from Doppler co-ordinates and levelled orthometric heights, and astro-geodetic deflections of the vertical are used as measures of geoid gradient. Pseudo-cubic splines are used to interpolate these data, for they are sufficiently continuous to be able to accommodate derivatives.

As in the previous sections we continue to work in the plane, with $x = (x_1, x_2) = (R\phi, R\lambda\cos\phi_0)$, (6.1). This introduces no error, other than that the semi-norm (4.79) measuring overall smoothness refers to the plane projection, rather than to a spherical earth.

Geoid gradients in north-south and east-west directions are related to the components ξ, η of the deflection of the vertical by *(Heiskanen and*

Moritz, 1967, p 112)

$$\frac{dN}{ds_\phi} = \frac{1}{R}\frac{dN}{d\phi} = -\xi = -(\Phi - \phi),$$
$$\frac{dN}{ds_\lambda} = \frac{1}{R\cos\phi}\frac{dN}{d\lambda} = -\eta = -(\Lambda - \lambda)\cos\phi. \qquad (6.19)$$

In the projection plane the geoid gradients are

$$\mu = \frac{dN}{dx} = \frac{dN}{ds_\phi}\frac{ds_\phi}{dx} = -\xi,$$
$$\nu = \frac{dN}{dy} = \frac{dN}{ds_\lambda}\frac{ds_\lambda}{dy} = -\eta\frac{\cos\phi}{\cos\phi_0}. \qquad (6.20)$$

The geoid heights $N = h - H$ remain unchanged in the projection.

Let geoid heights N_i be given at N points $\alpha_i = (\alpha_{1i}, \alpha_{2i}) \in \mathbb{R}^2$, and pairs of geoid gradients (μ_j, ν_j) be given at M points $\beta_j = (\beta_{1j}, \beta_{2j})$. The pseudo-cubic spline interpolating these data has the form (with $x = (x_1, x_2)$)

$$s(x) = \sum_{i=1}^{N} a_i |x - \alpha_i|^3 + \sum_{j=1}^{M} 3b_j(x_1 - \beta_{1j})|x - \beta_j| + \sum_{j=1}^{M} 3c_j(x_2 - \beta_{2j})|x - \beta_j|$$
$$+ d_0 + d_1 x_1 + d_2 x_2, \qquad (6.21)$$

where the coefficients $\{a_i\}_{i=1,\ldots,N}$, $\{(b_j, c_j)\}_{j=1,\ldots,M}$, d_0, d_1, d_2 are obtained by solving the linear equations

(following page)

$$s(\alpha_k) = \sum_{i=1}^{N} a_i|\alpha_k - \alpha_i|^3 + \sum_{j=1}^{M} 3b_j(\alpha_{1k} - \beta_{1j})|\alpha_k - \beta_j| +$$

$$+ \sum_{j=1}^{M} 3c_j(\alpha_{2k} - \beta_{2j})|\alpha_k - \beta_j| + d_0 + d_1\alpha_{1k} + d_2\alpha_{2k} = N_k,$$

$$k = 1, \ldots, N,$$

$$\frac{\partial}{\partial x_1} s(\beta_\ell) = \sum_{i=1}^{N} 3a_i(\beta_{1\ell} - \alpha_{1i})|\beta_\ell - \alpha_i|$$

$$+ \sum_{j=1}^{M} 3b_j\{2(\beta_{1\ell} - \beta_{1j})^2 + (\beta_{2\ell} - \beta_{2j})^2\}|\beta_\ell - \beta_j|^{-1}$$

$$+ \sum_{j=1}^{M} 3c_j(\beta_{1\ell} - \beta_{1j})(\beta_{2\ell} - \beta_{2j})|\beta_\ell - \beta_j|^{-1} + d_1 = \mu_\ell,$$

$$\ell = 1, \ldots, M, \qquad (6.22)$$

$$\frac{\partial}{\partial x_2} s(\beta_\ell) = \sum_{i=1}^{N} 3a_i(\beta_{2\ell} - \alpha_{2i})|\beta_\ell - \alpha_i|$$

$$+ \sum_{j=1}^{M} 3b_j(\beta_{1\ell} - \beta_{1j})(\beta_{2\ell} - \beta_{2j})|\beta_\ell - \beta_j|^{-1}$$

$$+ \sum_{j=1}^{M} 3c_j\{(\beta_{1\ell} - \beta_{1j})^2 + 2(\beta_{2\ell} - \beta_{2j})^2\}|\beta_\ell - \beta_j|^{-1} + d_2 = \nu_\ell,$$

$$\ell = 1, \ldots, M,$$

subject to the natural spline end conditions

$$\sum_{i=1}^{N} a_i = 0,$$

$$\sum_{i=1}^{N} a_i\alpha_{1i} + \sum_{j=1}^{M} b_j = 0,$$

$$\sum_{i=1}^{N} a_i \alpha_{2i} + \sum_{j=1}^{M} c_j = 0.$$

The southern African data available for the construction of a pseudo-cubic spline geoid of the form (6.21) are shown in Figure 6.7. The data comprise

(i) 92 pairs of vertical deflection components throughout South Africa observed since 1965 by the Chief Directorate of Surveys and Mapping in support of their geodetic traversing programme *(D.P.M. Rousseau, personal communication)*;

(ii) 9 pairs of deflection components along the South African section of the 30°E meridian arc *(Union of South Africa, 1954)*;

(iii) 47 pairs of deflection components along the parallel 30°S observed in the years 1960 to 1969 by B.M. Jones and P.V. Angus-Leppan *(Jones, 1970)*;

(iv) Five Doppler points in South Africa and one in Namibia that are South Africa's part of the ADOS African Doppler Survey;

(v) a further 21 points in South Africa and the southern part of Namibia determined by Doppler translocation, and tied to the South African ADOS points *(D.P.M. Rousseau, personal communication)*;

(vi) 9 ADOS Doppler points in Botswana, 4 in Swaziland, and 6 in Zimbabwe *(Kumar, 1983)*. Though the points in Zimbabwe and four of the points in Botswana fall outside the study area, they have nevertheless been included in the solution since there are few data along the northern margin. The five ADOS points in Lesotho are not included since (trigonometrically levelled) elevations there are inconsonant by 1,5-3,0 m with those in South Africa.

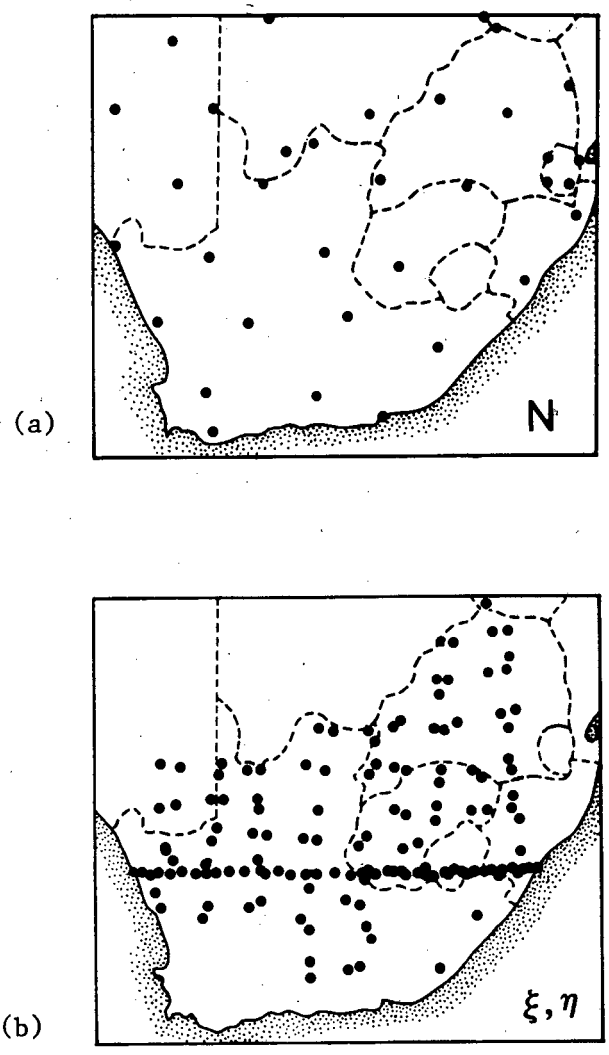


Figure 6.7: Distribution of (a) Doppler-derived geoid heights, and (b) astro-geodetic deflections of the vertical, contributing to the pseudo-cubic spline geoid shown in Figure 6.8.

The astro-geodetic deflections in (6.19) refer to the geoid, while the astronomic latitude and longitude Φ,Λ are observed at ground level, and should be reduced to the geoid. In the absence of real information about plumbline curvature, and in view of the conventional ellipsoidal corrections being of little value *(Heiskanen and Moritz, 1967, p 196)*,

this reduction is neglected.

Further, the deflections ξ,η in (6.19) refer to the geodetic datum used for the horizontal geodetic network; this is the Cape Datum, i.e. the Modified Clarke 1880 Ellipsoid ($a = 6\,378\,249{,}145$ m, $f = 1/293{,}466$) whose position and orientation is defined at the datum point Buffelsfontein. The Doppler co-ordinates refer to the NSWC 9Z-2 reference system, and were transformed to the Cape Datum using the transformation parameters recommended by *Merry (1985):*

$$\begin{aligned} X_0 &= -133{,}4 \text{ m}, \\ Y_0 &= -108{,}8 \text{ m}, \\ Z_0 &= -297{,}9 \text{ m}, \\ \varepsilon_Z &= 0''{,}8. \end{aligned} \tag{6.22}$$

One further point to note is that published elevations in South Africa, though nominally 'heights above sea level', are not properly orthometric heights, but conform more closely to a system of normal heights. Consequently the geoid heights calculated as $N = h - H$ are closer to being height anomalies or quasi-geoid heights.

Figure 6.8 shows the pseudo-cubic spline interpolating the given geoid heights and astro-geodetic deflections of the vertical.

The value of this geoid can be assessed by comparing it with an independent gravimetric solution. Figure 6.9 shows the UCT86 geoid *(van Gysen and Merry, 1987)* transformed to the Cape Datum, and Figure 6.10

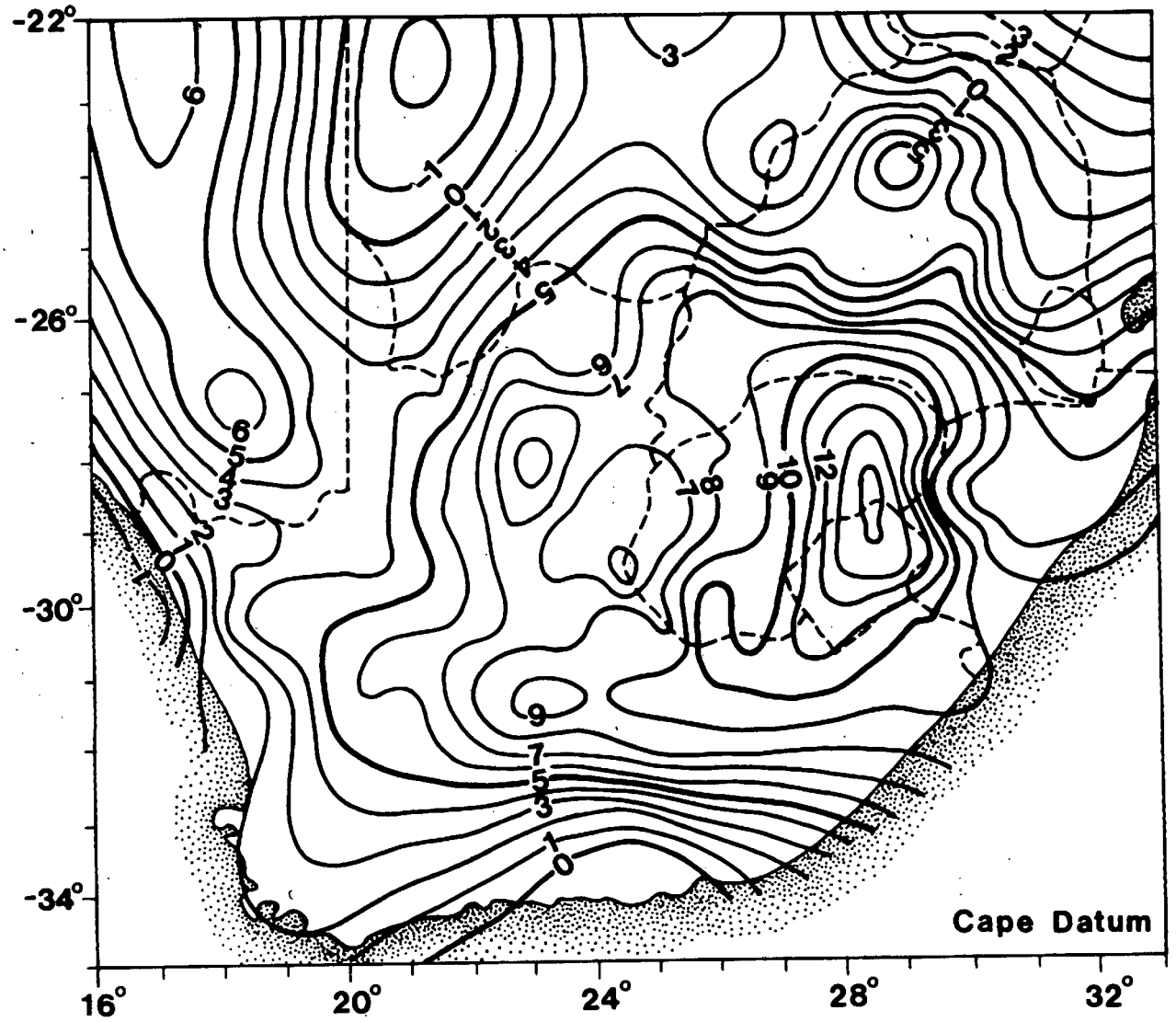


Figure 6.8: Pseudo-cubic spline geoid (on the Cape Datum) interpolating Doppler-derived geoid heights and astro-geodetic deflections of the vertical. Contour interval 1 m.

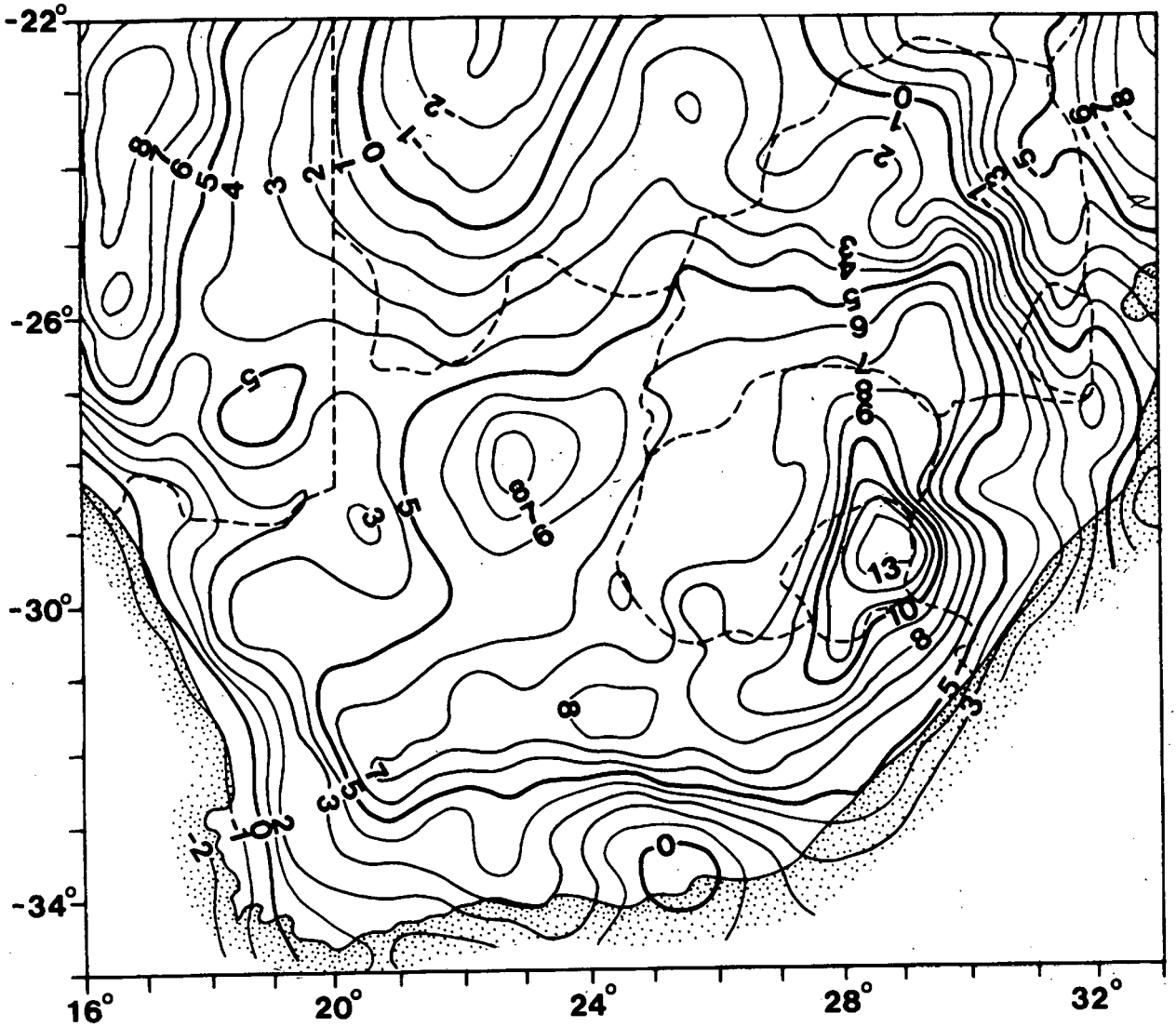


Figure 6.9: UCT86 geoid for southern Africa, transformed to the Cape Datum. Contour interval 1 m.

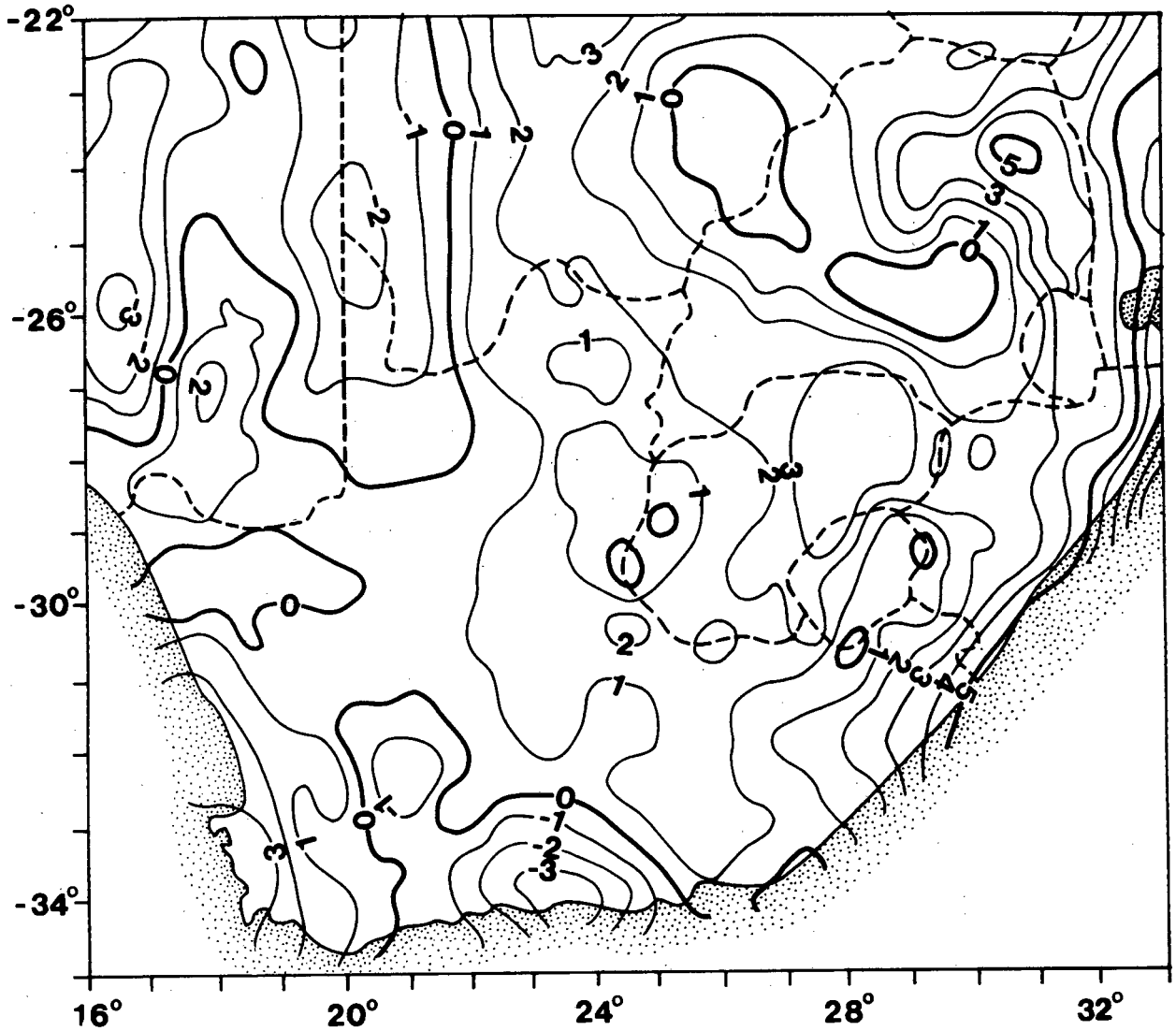


Figure 6.10: Error in the pseudo-cubic spline Doppler/astro-geodetic geoid relative to the UCT86 geoid. Contour interval 1 m.

the difference between the two solutions. There are significant discrepancies between the Doppler/astro-geodetic and gravimetric geoids. The differences could be due to

(i) errors in the gravimetric geoid;

(ii) errors in the Doppler co-ordinates, and in the levelled heights of the Doppler stations;

(iii) errors in the observed astro-geodetic deflections;

(iv) shortcomings in the model, e.g. in using topographic rather than reduced deflections, and in not accounting for the local variability of the vertical deflections; and

(v) shortcomings in the distribution of data.

The consequences of errors in the deflections, and shortcomings in modelling these, are moderated somewhat by the Doppler-derived geoid heights counteracting the introduction of erroneous tilts. There is little that can be done about the distribution of the data available. Where there are no data, or where the data are inadequate for modelling the geoid, there is little success, as in Mozambique and in the Transvaal Lowveld. Nonetheless these areas cannot entirely account for the rms discrepancy between the two solutions of 2,2 m. The gravimetric geoid has a precision considered to be somewhat better than 1,0 m; much of this error is at medium wavelengths *(Merry and van Gysen, 1987)*. The Doppler co-ordinates also have a precision better than 1,0 m. The rms error between gravimetric and Doppler-derived geoid heights at the Doppler points themselves is almost 2,0 m. So, apart from errors at the Doppler points (and due, in considerable measure it would appear, to uncertainty in the elevations of the Doppler stations), the rms

error in the spline geoid is close to 1,0 m.

The ratio between the rms discrepancy between the two solutions, and the rms variation in the geoid itself, is 60%; discounting errors at the Doppler points, the power of the spline solution is 27%. The spline solution is thus modestly successful. Where there are abundant data, as in the central parts of South Africa straddling 30°S, the spline representation is good; it is inadequate where there are few data, as in the north-eastern and eastern Transvaal.

6.6 Thin-Plate Spline Quadrature of Geodetic Integrals

The spline technique can be used for more than finding functions that interpolate or smooth data. As is indicated in Section 4.13 splines may also be put to use in the approximation of linear functionals. Among the functionals of geodetic interest are Stokes's integral and the formulae of Vening Meinesz *(Heiskanen and Moritz, 1967, p 96, p 113)*, and the vertical derivative operator L_1 that appears in the solution of Molodensky's problem by analytical continuation *(Moritz, 1980, pp 377-388)*. This section presents thin-plate spline quadrature rules for the numerical evaluation of these integrals, to a planar approximation, and within a circular innermost zone (in which the integrals are most sensitive to the data). The spline approximations are exact (within the planar approximation) for constant and linear functions, and for thin-plate splines interpolating the data.

The three integrals to be considered are given to the planar approximation as follows *(Heiskanen and Moritz, 1967, p 121; Sideris*

and Schwarz, 1986):

(i) *Stokes's formula:* With $x = (x_1, x_2) \in \mathbb{R}^2$, $y = (y_1, y_2) \in \mathbb{R}^2$, the n'th term in the Molodensky series for the height anomaly is

$$\zeta_n(x) = \frac{1}{\gamma_0} \iint_{\mathbb{R}^2} \frac{g_n(y)}{|x - y|} \, dm_2(y), \tag{6.23}$$

where $dm_2(y) = \frac{1}{2\pi} dm(y)$, (4.19); γ_0 is an average value of gravity; $g_0 = \Delta g$, and g_n, $n = 1,2,\ldots$ are the gravity anomaly correction terms in the Molodensky series, given by (6.26) and (6.27).

(ii) *Formulae of Vening Meinesz:* The n'th terms in Molodensky's solution for the components of the deflection of the vertical at ground level are

$$\begin{pmatrix} \xi_n(x) \\ \eta_n(x) \end{pmatrix} = \frac{1}{\gamma_0} \iint_{\mathbb{R}^2} \frac{g_n(y)}{|x - y|^3} \begin{pmatrix} x_1 - y_1 \\ x_2 - y_2 \end{pmatrix} dm_2(y), \tag{6.24}$$

(iii) *L_1 operator:* Provided f is continuous at $x \in \mathbb{R}^2$,

$$L_1 f(x) = \iint_{\mathbb{R}^2} \frac{f(y) - f(x)}{|x - y|^3} \, dm_2(y). \tag{6.25}$$

L_1 is applied recursively in calculating the terms g_n *(Moritz, 1980, p 386):*

$$L_n = \frac{1}{n} L_1 L_{n-1}, \tag{6.26}$$

$$g_n(y) = -\sum_{i=1}^{n} \{H(y)-H(x)\}^i L_i g_{n-i}(y) \tag{6.27}$$

where H is normal height.

The contribution to each integral of a circular disk of radius r_0, centred on $x \in \mathbb{R}^2$, is

$$\zeta_{n,I}(x) = \frac{r_0}{\gamma_0} \iint_{|x-y|\leq 1} \frac{g_n(y)}{|x-y|}\, dm_2(y), \tag{6.28}$$

$$\begin{pmatrix} \xi_{n,I}(x) \\ \eta_{n,I}(x) \end{pmatrix} = \frac{1}{\gamma_0} \iint_{|x-y|\leq 1} \frac{g_n(y)}{|x-y|^3} \begin{pmatrix} x_1-y_1 \\ x_2-y_2 \end{pmatrix} dm_2(y), \tag{6.29}$$

$$L_{1,I}f(x) = \frac{1}{r_0} \iint_{|x-y|\leq 1} \frac{f(y)-f(x)}{|x-y|^3} dm_2(y). \tag{6.30}$$

The second subscript I indicates 'contribution of innermost zone'. This is the area in which each integral is most sensitive to the data. The linear scale has been changed to give the disk unit radius.

Let $\alpha_i = (\alpha_{1i}, \alpha_{2i})$, $i = 1, \ldots, N$, be distinct points in $\mathbb{R}^2$, not all along a single straight line, at which are given the data in the form of function evaluations δ_{α_i}, i.e. gravity anomalies Δg are given, or the quantity g_n derived from these anomalies. The α_i need not fall within the disk, but if the quadrature is to reflect realistically the sensitivity of each integral to the data, many of the data points should have $|\alpha_i - x| \leq 1$. It will be assumed below, in (6.40), (6.45) and (6.51), that all $|\alpha_i - x| \leq 1$.

Quadrature rules approximating the integrals (6.28)-(6.30) are weighted sums of function evaluations at the nodal points α_i:

$$\zeta_{n,I,s}(x) = \frac{r_0}{\gamma_{0}} \sum_{i=1}^{N} a_i g_n(\alpha_i), \tag{6.31}$$

$$\begin{pmatrix} \xi_{n,I,s}(x) \\ \eta_{n,I,s}(x) \end{pmatrix} = \frac{1}{\gamma_{0}} \sum_{i=1}^{N} \begin{pmatrix} b_{1i} \\ b_{2i} \end{pmatrix} g_n(\alpha_i), \tag{6.32}$$

$$L_{1,I,s} f(x) = \frac{1}{r_{0}} \sum_{i=1}^{N} c_i \{f(\alpha_i) - f(x)\}. \tag{6.33}$$

The additional subscript s indicates 'spline approximation'. The approximation is to be exact for constant and linear functions within the nullspace of the semi-norm (4.73), and exact for thin-plate splines (4.72) interpolating the data.

Section 4.13 shows how the coefficients of such approximations may be found, as the solutions of linear equations of the form (4.109). The coefficient matrices A and B are precisely those of the thin-plate spline interpolation problem. Only the right-hand sides of (4.109) need still to be determined. The elements still required are

$$d_j = \langle \delta_{\alpha_j} * K_2, \ell \rangle, \quad j = 1, \ldots, N, \tag{6.34}$$

(where $K_2 = |y|^2 \log_e |y|$, and ℓ is the desired integral) ensuring the approximation is exact for interpolating splines; and

$$e_1 = \langle 1, \ell \rangle,$$

(continued)

$$e_2 = \langle y_1 - x_1, \ell \rangle, \qquad (6.35)$$

$$e_3 = \langle y_2 - x_2, \ell \rangle,$$

ensuring the approximation is exact on the nullspace $\mathbb{P}_1$.

As in (4.109) the coefficients $\{a_i\}$, $i = 1, \ldots, N$ of the approximation of (6.31), for example, and $\{h_k\}$, $k = 1,2,3$, of the spline nullspace, are obtained from the solution of

$$\begin{aligned} Aa + B^t h &= d \\ Ba \qquad\quad &= e \end{aligned} \qquad (6.36)$$

with

$$A = \begin{pmatrix} 0 & |\alpha_1-\alpha_2|^2 \log_e|\alpha_1-\alpha_2| & \cdots & |\alpha_1-\alpha_N|^2 \log_e|\alpha_1-\alpha_N| \\ |\alpha_2-\alpha_1|^2 \log_e|\alpha_2-\alpha_1| & 0 & & |\alpha_2-\alpha_N|^2 \log_e|\alpha_2-\alpha_N| \\ \vdots & & & \\ |\alpha_N-\alpha_1|^2 \log_e|\alpha_N-\alpha_1| & |\alpha_N-\alpha_2|^2 \log_e|\alpha_N-\alpha_2| & \cdots & 0 \end{pmatrix}_{N,N}$$

$$B = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \alpha_{11}-x_1 & \alpha_{12}-x_1 & & \alpha_{1N}-x_1 \\ \alpha_{21}-x_2 & \alpha_{22}-x_2 & & \alpha_{2N}-x_2 \end{pmatrix}_{3,N}. \qquad (6.37)$$

6.6.1 Stokes's formula: Here ℓ is the integral (6.28), and

$$d_j(x) = \iint_{|x-y| \leq 1} \frac{|y-\alpha_j|^2 \log_e|y-\alpha_j|}{|x-y|} \, dm_2(y). \qquad (6.38)$$

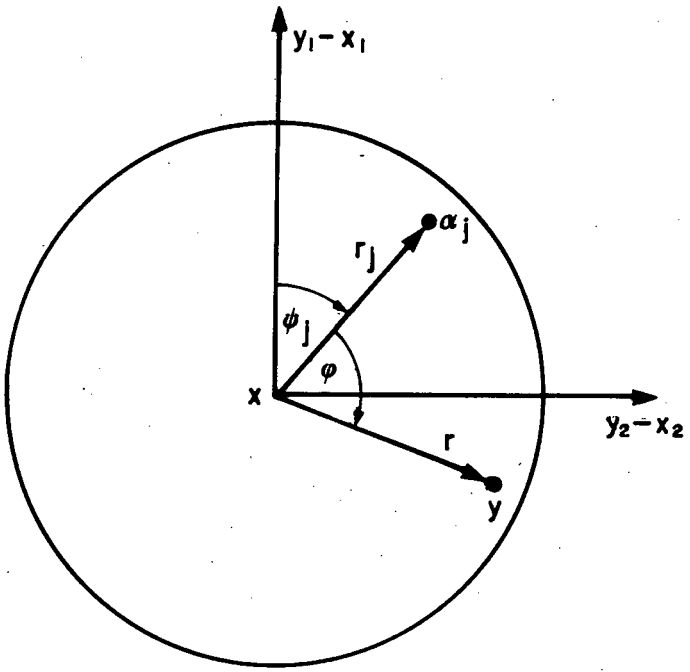


Figure 6.11: Relation on the unit disk, between spline knot α_j and integration variable y.

Putting $r = |x - y|$, and $r_j = |x - \alpha_j|$, as shown in Figure 6.11, we have

$$|y - \alpha_j|^2 = r^2 + r_j^2 - 2rr_j\cos\phi,$$

and

$$\frac{1}{2\pi}\int_0^{2\pi} |y - \alpha_j|^2 \log_e |y - \alpha_j| \, d\phi = (r^2 + r_j^2) \begin{pmatrix} \log_e r \\ \log_e r_j \end{pmatrix} + \begin{pmatrix} r_j^2 \\ r^2 \end{pmatrix} \qquad (6.39)$$

for $\begin{pmatrix} r \geq r_j \\ r < r_j \end{pmatrix}$ *(Spiegel, 1968, equations (15.106) and (14.288))*, and consequently that

$$d_j(x) = \int_0^{r_j} \{(r^2 + r_j^2)\log_e r_j + r^2\}dr$$
$$+ \int_{r_j}^{1} \{(r^2 + r_j^2)\log_e r + r_j^2\}dr = -\frac{1}{9} + \frac{4}{9}r_j^3. \qquad (6.40)$$

In addition,

$$e_1(x) = \iint_{|x-y|\leq 1} \frac{1}{|x-y|}\, dm_2(y) = 1,$$
$$e_2(x) = \iint_{|x-y|\leq 1} \frac{y_1 - x_1}{|x-y|}\, dm_2(y) = 0, \qquad (6.41)$$
$$e_3(x) = \iint_{|x-y|\leq 1} \frac{y_2 - x_2}{|x-y|}\, dm_2(y) = 0.$$

All the elements for the solution of (6.36) are in place. The distribution of the data evaluation points on the disk may be quite arbitrary; they need only be $\mathbb{P}_1$-unisolvent. However, in the numerical examples to follow, data will be presumed on a uniform rectangular grid - in practice the most likely situation.

Figure 6.12(a) shows the quadrature weights a_i, $i = 1,\ldots,5$, for a grid spacing of 1, Figure 6.12(b) the 13 coefficients for a spacing of 1/2, and Figure 6.12(c) the 29 coefficients for a spacing of 1/3. In each case the result is still to be multiplied by r_0/γ_0 (leaving to the user the choice of units for g_n and γ_0).

It is a consequence of the first end condition in (6.41) that the sum of the weights is 1 for each quadrature rule. This satisfies the result of approximating the innermost zone by a constant function *(Heiskanen*

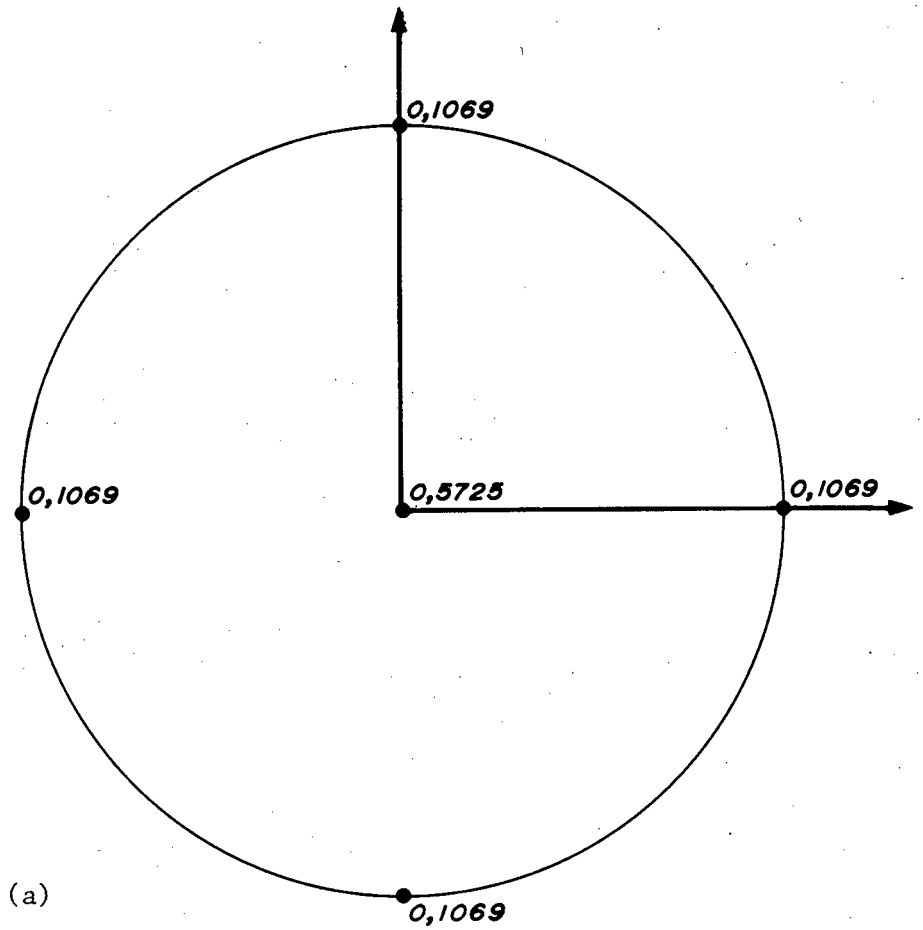
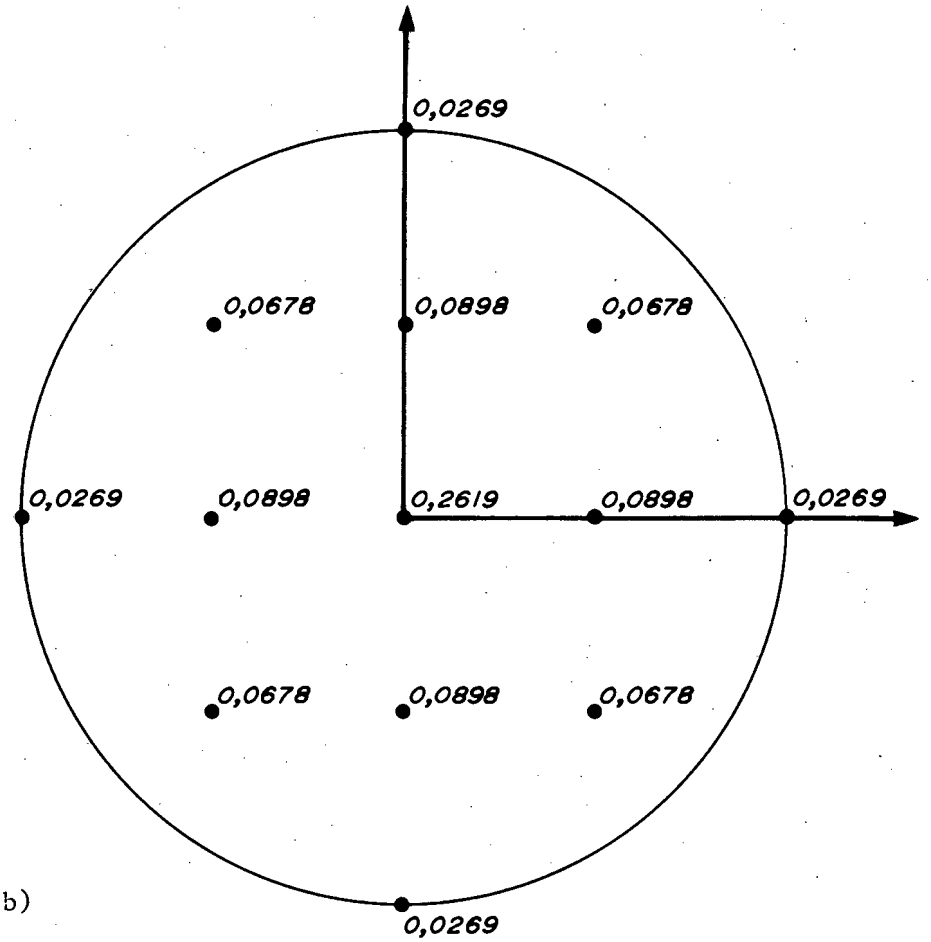
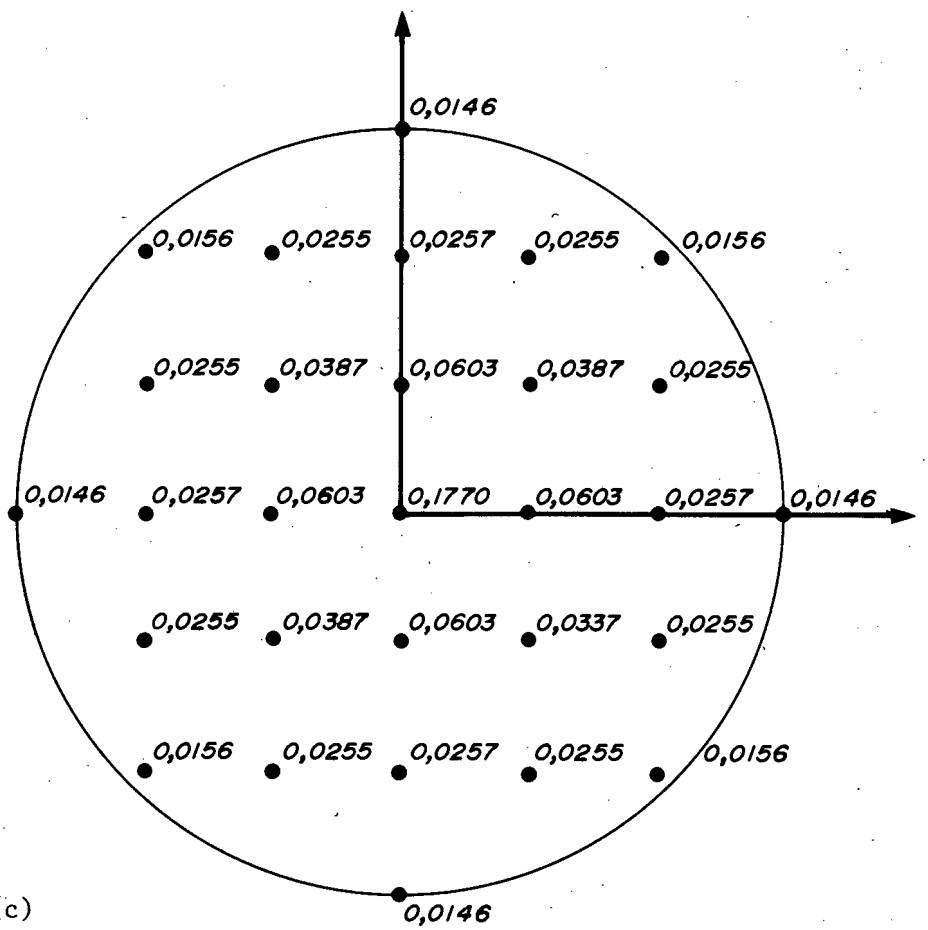


Figure 6.12: Quadrature weights for the thin-plate spline approximation $\zeta_{n,I,s}$ of Stokes's integral $\zeta_{n,I}$ over a unit disk, for
(a) grid interval 1 ($m = 1$, $N = 5$),
(b) grid interval 1/2 ($m = 2$, $N = 13$),
(c) grid interval 1/3 ($m = 3$, $N = 29$).
The weights need to be multiplied by r_0/γ_0.


0,0269
0,0678
0,0898
0,0678
0,0269
0,0898
0,2619
0,0898
0,0269
0,0678
0,0898
0,0678
0,0269

(b)


0,0146
0,0156
0,0255
0,0257
0,0255
0,0156
0,0255
0,0387
0,0603
0,0387
0,0255
0,0146
0,0257
0,0603
0,1770
0,0603
0,0257
0,0146
0,0255
0,0387
0,0603
0,0337
0,0255
0,0156
0,0255
0,0257
0,0255
0,0156
0,0146

(c)

and Moritz, 1967, p 122]. For the three examples given in Figure 6.12 the quadrature coefficients are positive, which is a satisfactory feature.

How well do the quadrature rules (6.31) approximate the integral (6.28)? Three simple examples will give some idea of the effectiveness of thin-plate spline quadrature of Stokes's integral; $r = |x - y|$, and the factor r_0/γ_0 is neglected.

$$\text{(i)} \qquad g_n(y) = r^2; \qquad \zeta_{n,I}(x) = \frac{1}{3};$$

$$\text{(ii)} \qquad g_n(y) = \begin{pmatrix} r - \frac{1}{4} \\ 0 \end{pmatrix} \text{ for } \begin{pmatrix} r \geq \frac{1}{4} \\ r < 1/4 \end{pmatrix}; \quad \zeta_{n,I}(x) = \frac{9}{32}; \qquad (6.42)$$

$$\text{(iii)} \qquad g_n(y) = \frac{r^2}{(2 + \frac{y_2 - x_2}{r})^2}; \quad \zeta_{n,I}(x) = \frac{2}{3^{5/2}}.$$

Table 6.1 shows the results of approximating (6.42) with different numbers of knots; m in the table is the number of intervals into which the unit radius is divided, and N is the number of knots. Spline quadrature appears quite effective for the numerical evaluation of the innermost zone of Stokes's integral.

6.6.2 Formulae of Vening Meinesz: The first formula of the pair (6.29), for the north-south deflection component ξ, will be considered here. Quadrature rules for the east-west component η may be obtained by rotating the results for ξ by 90°. The Vening Meinesz integral applied as ℓ to the thin-plate kernel function gives, for (6.34),

Table 6.1: Thin-plate spline approximations $\zeta_{n,I,s}$ to the three sample integrals in (6.42).

	m	N	$\zeta_{n,I}$	$\zeta_{n,I,s}$	relative error %
(i)	1	5	0,3333	0,4275	28,24
	2	13		0,3331	0,07
	3	29		0,3336	0,09
	4	49		0,3326	0,21
	6	113		0,3332	0,04
	8	197		0,3332	0,05
(ii)	1	5	0,2813	0,3206	14,00
	2	13		0,2945	4,73
	3	29		0,2849	1,31
	4	49		0,2777	1,28
	6	113		0,2809	0,11
	8	197		0,2809	0,11
(iii)	1	5	0,1283	0,1722	34,20
	2	13		0,1294	0,82
	3	29		0,1286	0,24
	4	49		0,1281	0,13
	6	113		0,1282	0,04
	8	197		0,1282	0,05

$$d_j(x) = \iint_{|x-y|\leq 1} \frac{|y-\alpha_j|^2 \log_e|y-\alpha_j|}{|x-y|^3}(x_1 - y_1)dm_2(y). \qquad (6.43)$$

Putting $r = |x-y|$ and $r_j = |x-\alpha_j|$, and with ψ_j as shown in Figure 6.11,

$$-\frac{(x_1 - y_1)}{|x-y|} = \cos(\phi+\psi_j) = \cos\phi\cos\psi_j - \sin\phi\sin\psi_j.$$

Now

$$\frac{1}{2\pi}\int_0^{2\pi} |y-\alpha_j|^2 \log_e|y-\alpha_j| \sin\phi d\phi$$

$$= \frac{1}{4\pi}\int_0^{2\pi} (r^2 + r_j^2 - 2rr_j\cos\phi)\log_e(r^2 + r_j^2 - 2rr_j\cos\phi)\sin\phi d\phi$$

$$= 0,$$

and

$$\frac{1}{2\pi}\int_0^{2\pi} |y-\alpha_j|^2 \log_e|y-\alpha_j| \cos\phi d\phi$$

$$= -rr_j \begin{pmatrix} \log_e r \\ \log_e r_j \end{pmatrix} - \frac{1}{2}rr_j - \frac{1}{4}\begin{pmatrix} r_j^3/r \\ r^3/r_j \end{pmatrix} \qquad (6.44)$$

for $\begin{pmatrix} r \geq r_j \\ r < r_j \end{pmatrix}$ (with the help of *Spiegel (1968), equations (14.285) and (14.288)*). Thus

$$d_j(x) = \int_0^{r_j} \{-r_j \log_e r_j - \frac{1}{2}r_j - \frac{1}{4}\frac{r^2}{r_j}\}\cos\psi_j dr$$

$$+ \int_{r_j}^{1} \{-r_j \log_e r - \frac{1}{2}r_j - \frac{1}{4}\frac{r_j^3}{r^2}\}\cos\psi_j dr$$

$$= \{-\frac{1}{2}r_j + \frac{4}{3}r_j^2 - \frac{1}{4}r_j^3\}\cos\psi_j. \qquad (6.45)$$

Further

$$e_1(x) = \iint_{|x-y|\leq 1} \frac{x_1 - y_1}{|x-y|^3} dm_2(y) = 0,$$

(continued)

$$e_3(x) = -\iint_{|x-y|\leq 1} \frac{(x_1 - y_1)(x_2 - y_2)}{|x-y|^3}\, dm_2(y) = 0, \qquad (6.46)$$

$$e_2(x) = -\iint_{|x-y|\leq 1} \frac{(x_1 - y_1)^2}{|x-y|^3}\, dm_2(y) = -\frac{1}{2}.$$

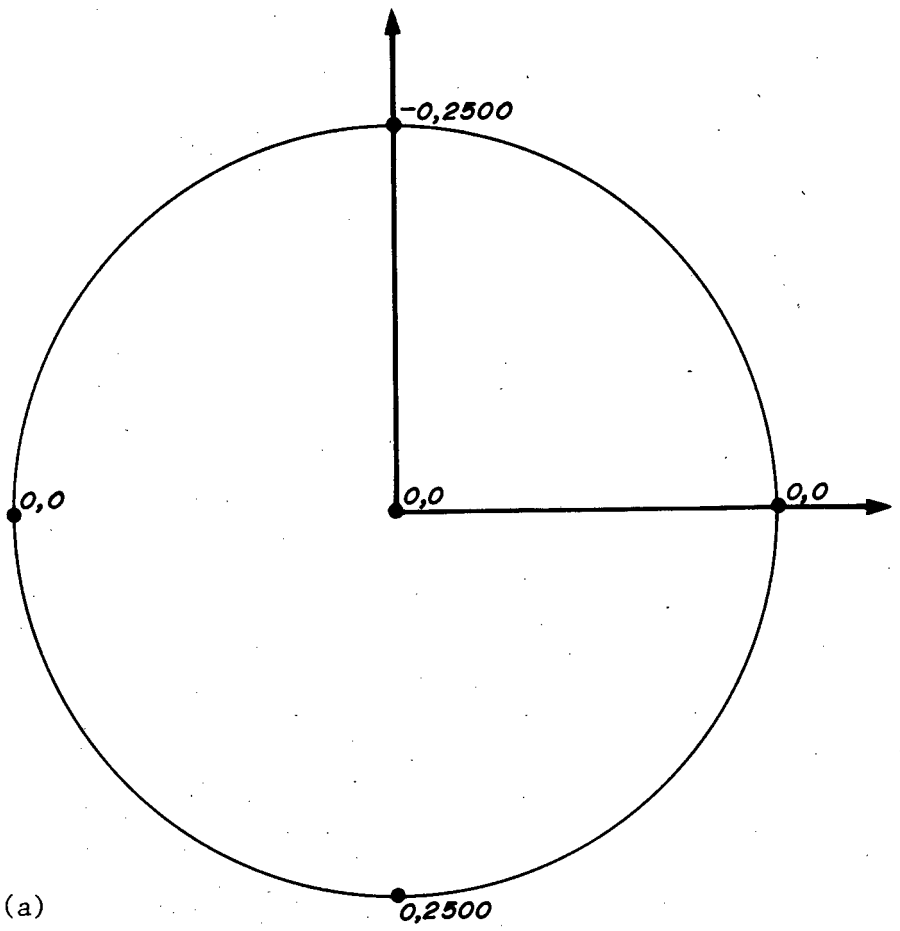
The solution for the quadrature coefficients $\{b_{1i}\}$, $i = 1,\ldots,N$, in (6.32) may now be obtained from (6.36). Figure 6.13(a) shows these coefficients for a grid spacing of 1, and Figure 6.13(b) for a spacing of 1/2. The coefficients satisfy the linear function approximation for the innermost zone given by *Heiskanen and Moritz (1967), p 122.*

The sign change for the coefficients at the points $\alpha = (1,0)$ and $\alpha = (-1,0)$ means the nodes for the 13-point rule in Figure 6.13(b) are not chosen satisfactorily. For a given number of knots it may be better to select more grid points closer to the computation point x. One possibility is to choose the knots as

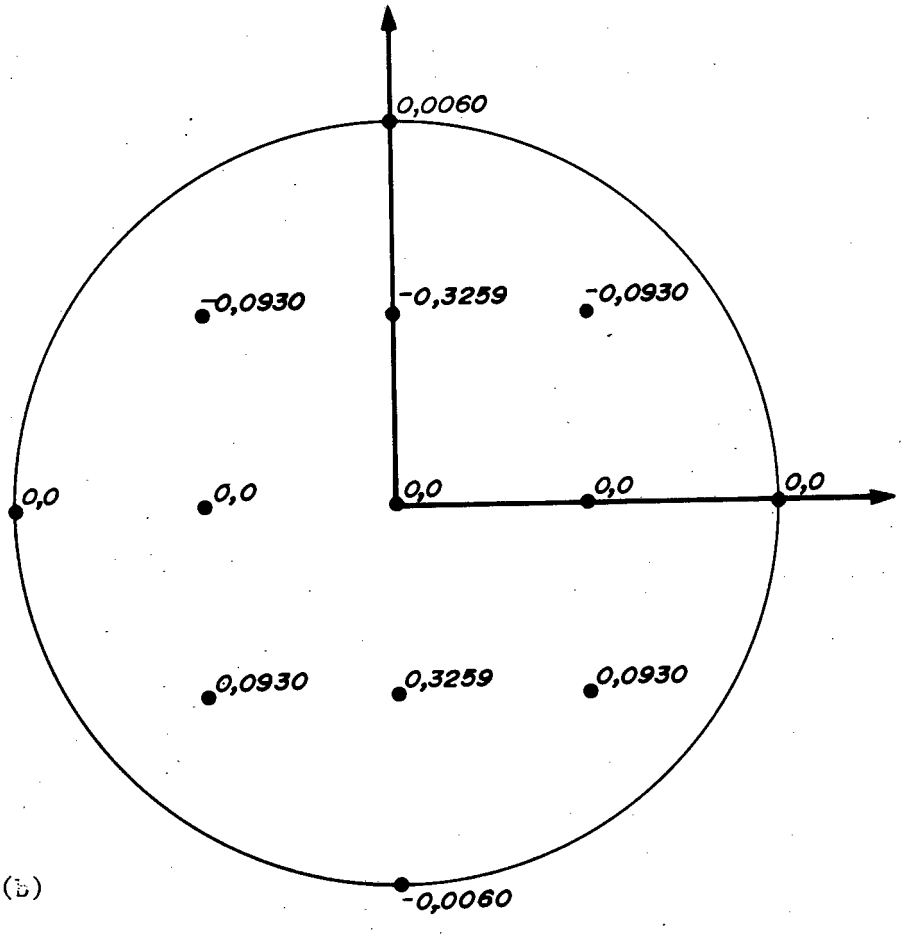
$$\alpha = (\alpha_1,\alpha_2) = (\{\tfrac{k}{m}\}^2,\{\tfrac{\ell}{m}\}^2),\ k,\ell = 0,\ldots,m,\ |\alpha| \leq 1. \qquad (6.47)$$

Figure 6.14 shows the quadrature weights for this form of 'quadratic sampling' from a more dense grid, with $m = 2$ (and $N = 13$).

Figure 6.13: (Following page). Quadrature weights for the thin-plate spline approximation $\xi_{n,I,s}$ of Vening Meinesz's formula for $\xi_{n,I}$ over a unit disk, for
(a) a grid interval of 1 ($m = 1$, $N = 5$), and
(b) a grid interval of 1/2 ($m = 2$, $N = 13$).
The weights need to be multiplied by $1/\gamma_0$.


-0,2500
0,0
0,0
0,0
0,2500

(a)


0,0060
-0,0930
-0,3259
-0,0930
0,0
0,0
0,0
0,0
0,0930
0,3259
0,0930
-0,0060

(b)

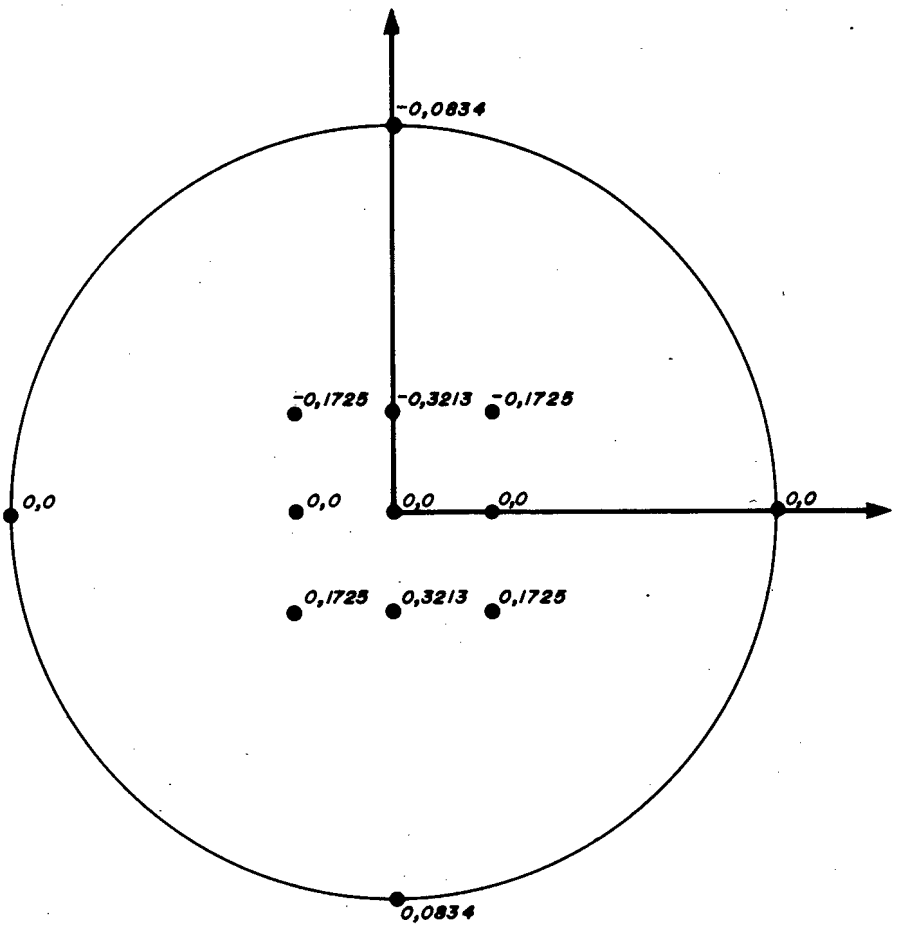


Figure 6.14: Quadrature weights for the thin-plate spline approximation $\xi_{n,I,s}$ of Vening Meinesz's formula for $\xi_{n,I}$ over a unit disk, for a grid interval of 1/4 with 'quadratic sampling' (m = 2, N = 13).

Thin-plate spline quadrature of Vening Meinesz's formula, with quadratic sampling from a more closely spaced data grid, is tested on three simple examples:

$$\text{(i)} \quad g_n(y) = (x_1 - y_1)\{1 - |x-y|^2\}; \qquad \xi_{n,I}(x) = \frac{1}{3};$$

$$\text{(ii)} \quad g_n(y) = (x_1 - y_1)^3 |x-y|^{-1}; \qquad \xi_{n,I}(x) = \frac{3}{16}; \tag{6.48}$$

$$\text{(iii)} \quad g_n(y) = \begin{pmatrix} (x_1 - y_1)(1 - \frac{1}{4}|x-y|^{-1}) \\ 0 \end{pmatrix}; \quad \xi_{n,I}(x) = \frac{3}{8} + \frac{1}{8}\log_e \frac{1}{4}$$

$$\text{for} \begin{pmatrix} |x-y| \geq 1/4 \\ |x-y| < 1/4 \end{pmatrix}.$$

Table 6.2 shows that the quadrature is effective for approximating these integrals.

6.6.3 L_1 operator: The integral (6.30) is a linear functional on the difference $f(y) - f(x)$. Constant functions are excluded from the integrands $f(y) - f(x)$ of (6.30), and from the spline approximation of (6.30) as well. Since constant functions fall within the nullspace of thin-plate splines, the technique of Section 4.10 may be used to exclude any constant function from linear combinations of the kernel $|y|^2 \log_e |y|$. (4.91) gives the projection removing any dependence on constant functions. Under this projection, with a constant zero evaluated at $x = 0$, the kernel function $|y - \alpha_i|^2 \log_e |y - \alpha_i|$ is transformed to something proportional to

$$|y - \alpha_i|^2 \log_e |y - \alpha_i| - |y-x|^2 \log_e |y-x| - |\alpha_i - x| \log_e |\alpha_i - x|. \tag{6.49}$$

Table 6.2: Thin-plate spline approximations $\xi_{n,I,s}$ to the three sample integrals in (6.48), with the quadrature nodes selected as in (6.47).

	m	N	$\xi_{n,I}$	$\xi_{n,I,s}$	relative error %
(i)	1	−5	0,3333	0,0000	100,00
	2	13		0,3015	9,54
	3	29		0,3304	0,89
	5	85		0,3334	0,01
	7	169		0,3342	0,25
(ii)	1	5	0,1875	0,5000	166,67
	2	13		0,2375	26,68
	3	29		0,2002	6,76
	5	85		0,1886	0,57
	7	169		0,1877	0,11
(iii)	1	5	0,2017	0,3750	85,91
	2	13		0,1757	12,91
	3	29		0,2080	3,13
	5	85		0,2029	0,61
	7	169		0,2022	0,24

The proportionality constant, $c = 2^3 \frac{\Gamma(2)}{\Gamma(-1)}$, can be neglected, and (6.49) used as kernel function in forming the linear equations (6.36). The nullspace of this kernel mapping now contains only functions linear in $y_1 - x_1$ and $y_2 - x_2$. The right-hand side of (6.36) requires

$$d_j(x) = \iint_{|x-y|\leq 1} \frac{|y-\alpha_j|^2 \log_e|y-\alpha_j| - |y-x|\log_e|y-x| - |\alpha_i - x|\log_e|\alpha_i - x|}{|x-y|^3} dm_2(y). \tag{6.50}$$

It follows from (6.39) that, with $r = |x-y|$, $r_j = |x-\alpha_j|$,

$$\frac{1}{2\pi}\int_0^{2\pi} \{|y-\alpha_j|^2 \log_e|y-\alpha_j| - r^2\log_e r - r_j^2 \log_e r_j\} d\phi$$

$$= (r^2 + r_j^2)\begin{pmatrix} \log_e r \\ \log_e r_j \end{pmatrix} + \begin{pmatrix} r_j^2 \\ r^2 \end{pmatrix} - r^2\log_e r - r_j^2 \log_e r_j$$

for $\begin{pmatrix} r \geq r_j \\ r < r_j \end{pmatrix}$, and consequently that

$$d_j(x) = \int_0^{r_j} \{\log_e r_j - \log_e r + 1\} dr + \int_{r_j}^1 \{\frac{r_j^2}{r^2}\log_e r - \frac{r_j^2}{r^2}\log_e r_j + \frac{r_j^2}{r^2}\} dr$$

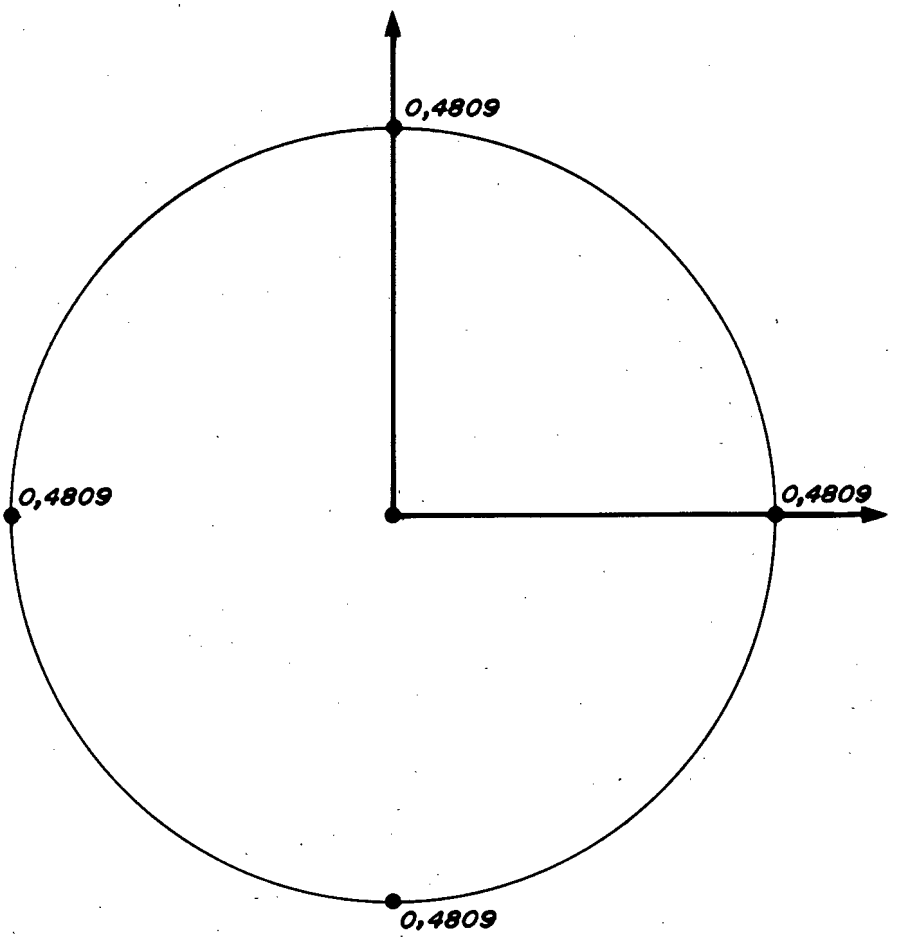
$$= 4r_j - 2r_j^2 + r_j^2 \log_e r_j. \tag{6.51}$$

In addition,

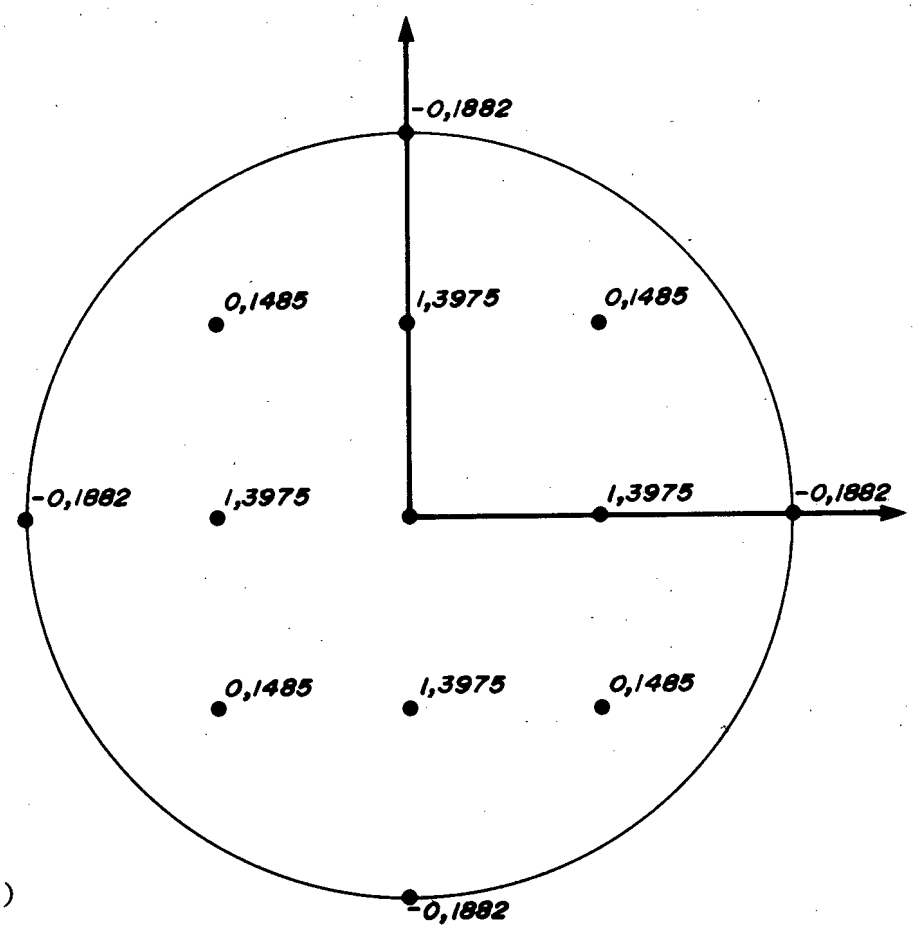
$$e_2(x) = \iint_{|x-y|\leq 1} \frac{y_1 - x_1}{|x-y|^3} dm_2(y) = 0,$$

$$e_3(x) = \iint_{|x-y|\leq 1} \frac{y_2 - x_2}{|x-y|^3} dm_2(y) = 0, \tag{6.52}$$

and e_1 does not appear.


0,4809
0,4809
0,4809
0,4809

(a)


-0,1882
0,1485
1,3975
0,1485
-0,1882
1,3975
1,3975
-0,1882
0,1485
1,3975
0,1485
-0,1882

(b)

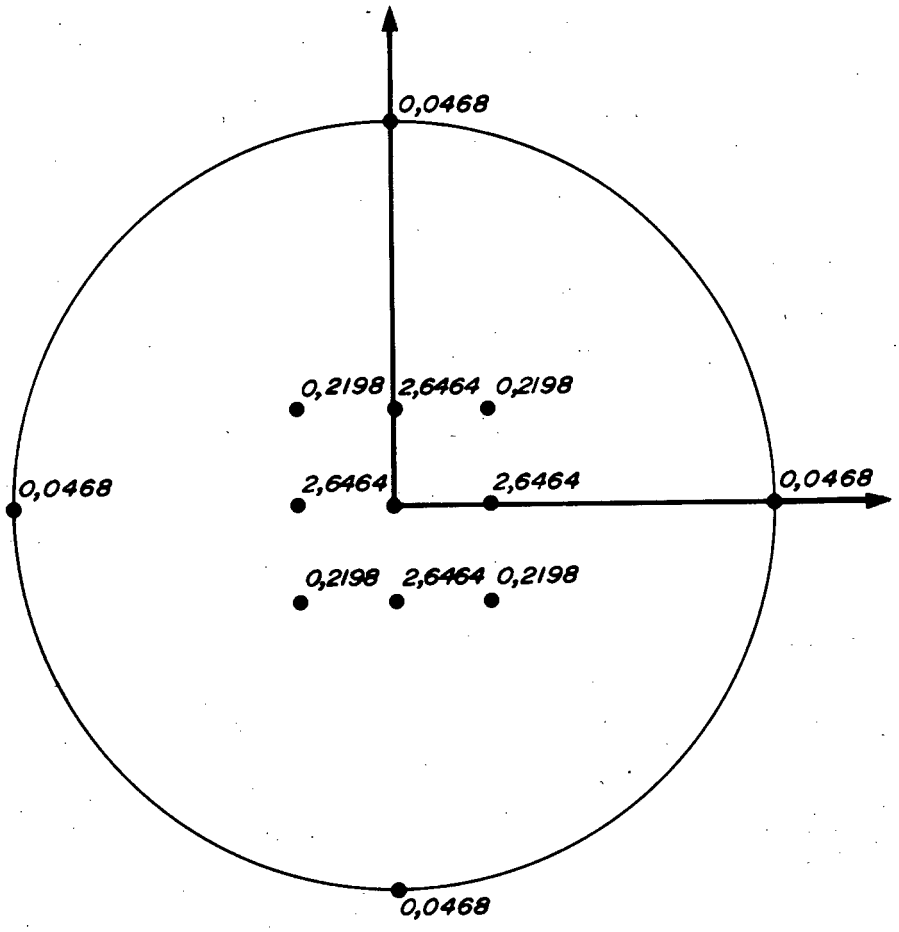


Figure 6.16: Quadrature weights for the thin-plate spline approximation $L_{1,I,s}$ of the L_1 gradient operator, over a unit disk, for a grid interval of 1/4 with 'quadratic sampling' (m = 2, N = 12).

(iii) $f(y) = \dfrac{|x-y|^2}{\left(2+\dfrac{y_2-x_2}{|x-y|}\right)^2}$; $L_{1,I}(x) = \dfrac{2}{3^{3/2}}$.

Table 6.3 shows the results for the quadrature of these integrals.

Gravimetric terrain corrections: The terrain correction, also to a planar approximation, is given by

$$C(x) = \frac{B}{2}\iint_{\mathbb{R}^2} \frac{\{h(x)-h(y)\}^2}{|x-y|^3} dm_2(y), \tag{6.54}$$

where $B = 2\pi\kappa\rho$ is the Bouguer gradient (κ is the gravitational constant, and ρ the uniform density of the topographic masses). The contribution to C of the topography over a unit disk of radius r_0, centred on $x \in \mathbb{R}^2$, is

$$C_I(x) = \frac{B}{2r_0}\iint_{|x-y|\leq 1} \frac{\{h(x)-h(y)\}^2}{|x-y|^3} dm_2(y). \tag{6.55}$$

The quadrature weights for the L_1 operator may be used for the terrain correction C_I also, provided evaluation at the quadrature nodes is $\{h(\alpha)-h(x)\}^2$, and the result is multiplied by $B/2r_0$.

Table 6.3: Thin-plate spline approximations $L_{1,I,s}$ to the three sample integrals in (6.52), with the quadrature nodes selected as in (6.47).

	m	N	$L_{1,I}$	$L_{1,I,s}$	relative error %
(i)	1	4	0,6667	0,0000	100,00
	2	12		0,7164	7,47
	3	28		0,6251	6,23
	5	84		0,6649	0,27
	7	168		0,6683	0,25
(ii)	1	4	0,6363	1,4427	126,73
	2	12		0,2315	63,61
	3	28		0,5961	6,31
	5	84		0,6534	2,68
	7	168		0,6357	0,10
(iii)	1	4	0,3849	0,7748	101,29
	2	12		0,3823	0,68
	3	28		0,3947	2,55
	5	84		0,3866	0,43
	7	168		0,3854	0,12

CHAPTER 7

CONCLUSION

7.1 Summary

Orthogonal projection in Hilbert space forms the basis of many techniques of approximation. In the spline method the projection takes place in a Hilbert space with the 'initial' inner product, (2.18), deriving from an observation mapping F and a 'co-observation' mapping U. Figure 2.2 illustrates the geometry of the projection. Relations between F and U determine the existence and uniqueness of interpolating splines, (2.3) with (2.4), and smoothing splines, (2.5); the requirement of F and U is that they have closed nullspaces and be collectively injective.

If Chapter 2, characterising spline interpolation as orthogonal projection, deals with the geometry of spline approximation, the following chapter can be said to deal with the algebra. There the first steps are taken towards constructing spline elements by drawing on the other important feature of Hilbert spaces, the Riesz mapping associating elements of the Hilbert space with elements of its dual space, (3.4) *et seq.* The concepts taken from *Schwartz (1964, 1973)* relating to the kernel associated with a Hilbert sub-space of a larger locally convex topological vector space, are readily extended to kernels associated with semi-Hilbert sub-spaces, (3.7). Once the associated kernel is identified, the spline interpolation problem may be solved in the way shown by *Duchon (1977)*, here Section 3.3. Chapter 3 is rounded out by a brief treatment of the problem of

spline approximation of linear functionals, of reproducing kernel functions (in contrast to the work of *Meinguet (1979a, 1979d, 1981)* on surface splines, and of *Freeden (1981a, 1984a)* on spherical splines, these are not central to the approach I have adopted), and, very briefly indeed, of the connections with Green's functions.

The purpose of Chapter 4 is to identify kernel mappings associated with spaces of distributions on $\mathbb{R}^n$. I was unfamiliar with the theory of distributions and with the theory of Sobolev spaces, and hence the introductory sections to the chapter setting out concepts and definitions. The Sobolev spaces most commonly encountered in the theory of differential equations are $H^s(\mathbb{R}^n)$; $s \in \mathbb{R}$, (4.37), which are Hilbert spaces with norm

$$\|u\|^2_{H^s(\mathbb{R}^n)} = \int_{\mathbb{R}^n} (1 + |y|^2)^s |\hat{u}(y)|^2 dm_n(y). \tag{4.39}$$

The spaces more relevant to our purposes are $\tilde{H}^s(\mathbb{R}^n)$, for $s < \frac{n}{2}$, (4.42), which are also Hilbert spaces, with norm

$$\|u\|^2_{\tilde{H}^s(\mathbb{R}^n)} = \int_{\mathbb{R}^n} |y|^{2s} |\hat{u}(y)|^2 dm_n(y). \tag{4.43}$$

I have relied on *Duchon (1977)* in showing that $\tilde{H}^s(\mathbb{R}^n)$ is a Hilbert sub-space of the space Φ'_n of tempered distributions on $\mathbb{R}^n$, and that for $-\frac{n}{2} < s < \frac{n}{2}$, and Ω an open bounded set in $\mathbb{R}^n$,

$$\tilde{H}^s(\Omega) = H^s(\Omega). \tag{4.48}$$

Section 4.6 introduces the Beppo-Levi spaces $D^{-m}\tilde{H}^{s}(\mathbb{R}^n)$, all whose m'th derivatives are in $\tilde{H}^{s}(\mathbb{R}^n)$, $s < \frac{n}{2}$, (4.50). These are the semi-Hilbert sub-spaces (with semi-norm (4.51)) in which spline solutions will eventually be found. On bounded sets Ω, with $-m - \frac{n}{2} < s < \frac{n}{2}$, we have the equality

$$(D^{-m}\tilde{H}^{s})(\Omega) = H^{m+s}(\Omega), \tag{4.52}$$

making $D^{-m}\tilde{H}^{s}(\mathbb{R}^n)$ a dense semi-Hilbert sub-space of the Frechet space $H^{m+s}_{loc}(\mathbb{R}^n)$. This is the result required in Section 4.7 for the construction of the kernel mapping associated with the Beppo-Levi space $D^{-m}\tilde{H}^{s}(\mathbb{R}^n)$. The theorem taken from *Duchon (1977)* giving the kernel as

$$\begin{aligned} &\theta\colon H^{-m-s}_{comp} \cap \mathbb{P}^{\circ}_{m-1} \to D^{-m}\tilde{H}^{s}(\mathbb{R}^n), \\ &\theta\colon \mu \mapsto \mu * K_{2m+2s-n}, \end{aligned} \tag{4.57}$$

where

$$K_{\lambda}(x) = \begin{cases} |x|^{\lambda}\log_e|x|, & \lambda \text{ even, non-negative,} \\ |x|^{\lambda} & \text{otherwise,} \end{cases} \tag{4.62}$$

is the most important outcome of the whole chapter. The Fourier transform plays a significant part in obtaining this result.

The rest of Chapter 4 is relatively straightforward. Section 4.8 shows how the interpolating rotation-invariant 'surface splines'

semi-Hilbert sub-spaces of distributions on $\mathbb{R}^n$ - is extended to splines on the circle and on the sphere, there are immediate savings, since the Fourier transform of a periodic or spherical dustribution takes on a particularly simple form (equations (5.3) and (5.27), respectively).

Sünkel (1984a, 1984b) has given a nice explanation of periodic splines. I have done little more than show that the kernel associated with $H^m(\Pi)$, distributions on the unit circle Π whose m'th derivatives are square integrable, is the mapping

$$\begin{aligned} &\theta\colon (\mathbb{D}'(\Pi))' \cap \mathbb{P}_0^\circ \to \mathbb{D}'(\Pi), \\ &\theta\colon \mu \mapsto \mu * K_{2m}, \end{aligned} \tag{5.10}$$

with

$$K_{2m}(x) = \sum_{k=1}^{\infty} \frac{\cos kx}{k^{2m}}. \tag{5.11}$$

In the same way, there is no more comprehensive treatment of spherical and harmonic splines than the work of *Freeden (1981a, 1981b, 1982, 1984)*. I have merely indicated that the kernel associated with $H^m(\Omega)$, the space of distributions on the unit sphere Ω whose m'th surface gradients ∇_* are square integrable, is the mapping

$$\begin{aligned} &\theta\colon (\mathbb{D}'(\Omega))' \cap \mathbb{P}_0^\circ \to \mathbb{D}'(\Omega), \\ &\theta\colon \mu \mapsto \mu * K_m, \end{aligned} \tag{5.34}$$

with

$$K_m(\cos\psi) = \sum_{n=1}^{\infty} \frac{2n+1}{4\pi} \{n(n+1)\}^{-m} P_n(\cos\psi). \tag{5.35}$$

For the rest, Chapter 5 places in a spline context the choice of 'degree variances' (5.36), regularisation by introducing a Bjerhammar sphere, and the Szegö and Krarup kernels from the theory of analytical collocation. Analytical collocation is the same as spline interpolation using a kernel associated with a Hilbert (as opposed to semi-Hilbert) sub-space.

Chapter 6 gives some numerical examples of the use of spline functions for gravity field approximation. For straightforward interpolation of a single data type, e.g. geoid heights or gravity anomalies, there is little to choose between many surface splines, or spherical splines, and least squares collocation. Spline interpolants lie close to least squares interpolants. When it comes to interpolating different data types, with observation functionals other than point evaluation, surface splines do appreciably worse, at least on the evidence of multiquadric interpolation of geoid heights and gravity anomalies in Section 6.4. It is surprising this interpolation worked at all, for the multiquadric functions have little to do with any geodetic reality. Splines are better suited to representations based on mixed data, when the data are related geometrically, as in Section 6.5 giving the example of a pseudo-cubic geoid for southern Africa interpolating Doppler-derived geoid heights and astro-geodetic deflections of the vertical.

The final numerical study shows how thin-plate spline approximations may be found for three standard geodetic integrals: Stokes's and Vening Meinesz's formulae, and the L_1 vertical gradient operator

Tscherning, C.C. (1981): Comparison of some methods for the detailed representation of the earth's gravity field, Reviews of Geophysics and Space Physics, 19(1), pp 213-221.

Tscherning, C.C. (1984): Local approximation of the gravity potential by least squares collocation, Lecture to the International Summer School on Local Gravity Field Approximation, Beijing, 78 pp.

Tscherning, C.C. (1985): Current problems in gravity field approximation, First Hotine-Marussi Symposium, Rome, 21 pp.

Tscherning, C.C. (1986): Functional methods for gravity field approximation, in Suenkel, H. (ed.), Mathematical and Numerical Techniques in Physical Geodesy, Lecture Notes in Earth Sciences 7, Springer-Verlag, Berlin, pp 3-47.

Union of South Africa (1954): Observations of the South African section of the Arc of the 30th Meridian, Department of Lands, Trigonometrical Survey, Mowbray.

Van Gysen, H. and C.L. Merry (1987): The Gravity Field in Southern Africa, University of Cape Town, Department of Surveying, Technical Report TR-3, 42 pp.

Wahba, G. (1981): Spline interpolation and smoothing on the sphere, SIAM Journal on Scientific and Statistical Computing, 2(1), pp 5-16.

Wahba, G. (1984): Surface fitting with scattered noisy data on Euclidean D-space and on the sphere, Rocky Mountain Journal of Mathematics, 14(1), pp 281-299.

Wang, Y.M. (1987): Numerical aspects of the solution of Molodensky's problem by analytical continuation, Manuscripta Geodaetica, 12(4), pp 290-295.

Weidman, J. (1980): Linear Operators in Hilbert Spaces, Graduate Texts in Mathematics 68, Springer-Verlag, New York, 402 pp.

Zadro, M. (1984): Spectral properties of the Newtonian potential field and their application to the interpolation of gravity anomalies, Geophysical Journal of the Royal Astronomical Society, 79, pp 489-493.

that appears in the series solution of Molodensky's problem. These spline-based quadrature rules take the numerical evaluation of geodetic integrals away from the strait-jacket of data on a fixed rectangular or circular grid, in contrast, say, to recent work by *Wang (1987)* on the bicubic spline evaluation of the L_1 gradient operator, where data are still restricted to a grid.

7.2 Splines and Gravity Field Approximation

Tscherning (1981) compares different techniques for obtaining detailed representations of the earth's gravity field. The spline technique is part of the 'operational approach' to gravity field approximation. A change in solution space is required, away from the disturbing potential T harmonic outside the surface of the earth (and regular at infinity), to a space of functions harmonic outside a sphere contained within the earth, or to a space from which the disturbing potential is even further removed in the case of surface splines.

The quadrature formulae of Section 6.6 may be regarded as a contribution to the alternative 'model approach'.

Interpolating splines are part of the technique of 'pure collocation' or 'exact collocation', and smoothing splines part of 'minimum norm collocation'. On the other hand, collocation is usually thought of as involving spherical kernel functions (such as (5.43) or (5.46)), so apart from spherical splines, splines may also form part of the technique of least squares approximation, albeit with as many

coefficients as observations. The close relation of the data points to the spline knots is an important feature of spline theory.

Tscherning *(ibid.)* assesses features of different techniques under a number of headings; the same headings are used here:

Theoretical properties:

(i) *Representation of the disturbing potential is harmonic and regular at infinity:* Spherical splines are harmonic outside the sphere with radius R_B; regularity, the absence of the first degree harmonic, may be achieved by omitting the $n = 2$ term from (5.43), and from its derivatives, (5.44) - (5.46). This is done in (6.4) and (6.8).

(ii) *Set of harmonicity correct:* No.

(iii) *Parameter estimation possible:* Yes, but this is not something I have considered.

(iv) *Regularity at infinity not needed:* Yes.

(v) *Spherical approximation not needed:* Yes, this is possible, but a spherical approximation has been used throughout.

(vi) *Ultimate precision possible* (i.e. convergence to the 'true' disturbing potential T): This can be made possible, but not without estimating improved observation point positions, thus

not with the two-dimensional approach of Chapter 6.

(vii) *Reduced data not needed:* Again, this need not be necessary, but I have used throughout data reduced to the plane or sphere.

Flexibility and reliability:

(viii) *All functionals of the disturbing potential can be evaluated:* Yes, in the case of three-dimensional spline representations. In the case of two-dimensional splines, the types of functionals that may be evaluated is limited.

(ix) *All data types may be used:* Yes, in the case of three-dimensional splines. In two dimensions the data need to be reduced to the plane or sphere; moreover, we see in Section 6.4 that not all data types need be accommodated equally comfortably.

(x) *Data quality may be taken into account:* Yes, the simple 'ridging' technique of data smoothing may be extended to reflect observation variances.

(xi) *Error estimates computable:* It should be possible to derive internal estimates of the error of representation on the basis of data type and data distribution (though one difficulty consequent to changing solution space is to obtain some estimate for $\|T\|$). I have done no work in this direction, relying in Chapter 6 on external comparisons.

(xii) *Global data coverage not required:* Yes, as is true for all operational methods.

(xiii) *Stable in clusters:* Experience shows that the tendency to numerical instability in the linear equations (4.68), and the analogous equations for spherical splines, is exacerbated by data clustering. This may be an incentive to make more use of reproducing kernel functions, so that the equation system to be solved is at least positive-definite.

(xiv) *New data introduced easily:* Yes, the solution of the linear equations (4.68) can be adapted to accommodate additional data without having to solve the augmented equations *de novo*. When the values of existing data are changed, a technique such as Crout factorisation requires no more than a fresh back-substitution through the reduced equations.

(xv) *Behaviour at data holes known:* Yes. Splines are meant to be smooth; they minimise gradient overall (in the case of multi-conic and $m = 1$ spherical splines), or curvature overall (in the case of thin-plate splines in two dimensions), or something analogous to these. Thus, where there are no data, as in data holes or beyond the limits of the data, the overall behaviour is known. (Outside the limits of the data the natural spline end conditions come into play, directing the representation eventually into a polynomial belonging to the spline nullspace, and making the gradient or curvature eventually zero). This does not mean all splines

The spline method is a very general approximation technique with many areas of application. It is not an exclusively geodetic technique; even so, its theoretical properties in relation to gravity field modelling ((i)-(vii) above) mirror those of other operational methods of approximation falling under the headings least squares approximation or minimum norm collocation that have found application in geodetic problems. Minimum norm collocation is a special case of spline smoothing in which the 'co-observation' mapping U is one-to-one.

Spherical splines come closer to representing geodetic reality than the surface splines of Chapter 4, but as can be seen in Sections 6.2 - 6.5, this does not detract from the efficacy of these functions in two-dimensional representations of the gravity field in the form of geoid heights or gravity anomalies.

In properties of flexibility and reliability ((viii)-(xv) above) the spline method again shows itself very much part of the operational approach. It shares with other operational techniques the advantages that all or most functionals of the disturbing potential can be evaluated, that all data types can be permitted, and it is possible to take data quality into account, solutions may be up-dated when new data are introduced, and there is good behaviour where there are gaps in the data or beyond the limits of the data. A disadvantage is a tendency to numerical instability at close clusters of data.

An example of the spline evaluation of a disturbing potential functional is the thin-plate spline approximation of geodetic integrals.

In computational properties ((xvi)-(xxii) above) the spline method shares the best features of both least squares approximation and minimum norm collocation techniques: it can be easily and automatically implemented, with an automatic selection of base functions. There is one base function μ_i*K for each observation functional μ_i, where K is the kernel function appropriate to the type of spline. To these base functions must be added a basis for the spline nullspace, and the base functions between them must satisfy the natural spline end conditions A disadvantage shared with all operational methods is that the coefficients of the representation can only be obtained by solving a system of linear equations.

In summary, splines are part of the operational approach to gravity field approximation; particularly attractive are their computational properties in easy and automatic implementation. A further attraction, not shared with all operational methods, is that the semi-norm being minimised is explicitly given. These are semi-norms that have a clear intuitive appeal, minimising curvature or gradient or something similar. There may be some arbitrariness in selecting smoothing or regularising parameters, though the results are not particularly sensitive to the precise way these are chosen.

7.3 Final Remarks

The distinctive features of a spline theory in a Hilbert space setting are that spline interpolation and smoothing is a process of orthogonal projection (for which I have used the theory of *Delvos and Schempp (1970, 1972)*), and that the Riesz theorem makes it possible to identify

elements of the Hilbert space with elements of the dual space. For the construction of the Riesz mapping I have used the kernel theory of *Schwartz (1964, 1973)*. *Duchon (1977)* shows how to construct kernels associated with semi-Hilbert spaces of distributions on $\mathbb{R}^n$. I have followed his construction, and show how the same construction yields kernels associated with spaces of periodic and spherical distributions, and in turn, the periodic splines of *Sünkel (1984a, 1984b)*, and spherical splines of *Freeden (1981a, 1981b)*.

All these splines share the same pleasing features: they are rotation-invariant, translation-invariant and scale-invariant (for splines on $\mathbb{R}^n$ this last property follows from the remark on p 60). In consequence, unlike bicubic splines or the tensor-product spherical splines of *Dierckx (1984)*, none of the splines I have considered depends on the location, orientation or scale of the Euclidean or spherical co-ordinate system in which the data points are placed.

Duchon's (1977) construction yields whole families of spline functions: multi-conic, pseudo-cubic, thin-plate, point-mass, etc. It is especially gratifying to see that Hardy's multiquadric functions, so widely used without being fully elucidated, are nothing else than regularised multi-conic splines and that these are a multivariate generalisation of linear interpolation. There need be little wonder that Hardy's technique has proved itself so effective.

Turning to numerical applications, it is not surprising that a smooth surface such as the geoid may be effectively interpolated with spline

spline functions. Surface and spherical splines are equally capable of interpolating geoid heights in the 2°x2° study area with a relative precision of 2cm per kilometre to the nearest interpolation point. Adding gravity data reduces this to about 1cm per kilometre for least squares interpolation with Tscherning and Rapp's 1974 degree variances, or to 1,4cm per kilometre for m = 2 spherical splines, irrespective of the number of geoid height points. Relative precisions of this kind may be sufficient for some less demanding applications of orthometric levelling with GPS.

With highly variable data, as gravity anomalies are in the study area, splines do no better and no worse than least squares interpolation.

A further attraction of many splines is the particularly simple form of the kernel function. This makes certain splines especially suited to particular applications involving mixed data, or approximation of functionals. I have shown how pseudo-cubic splines may be used to interpolate both geoid heights and geoid gradients. I show that thin-plate splines may be used to approximate the innermost zone contribution to Stokes's and Vening Meinesz's formulae, and the L_1 gradient operator. The approximations are exact for interpolating splines in the area where the results are most sensitive to the data; these data may be arbitrarily distributed. The technique nicely complements other circular evaluation techniques such as the ring integration of *Kearsley (1986)*. A disadvantage as always, is repeated solution of sets of linear equations (unless the data are in a regular pattern), although the equation systems are easy to set up.

What I hope to have achieved is to show that a single approach, giving the kernel mapping associated with each type of spline, demonstrates the common construction of splines as diverse as thin-plate, multi-conic, pseudo-cubic, spherical or point-mass splines, and that there is sufficient choice to select splines suitable for any application in gravity field modelling ranging from various styles of geoid interpolation to the approximation of geodetic integrals.

REFERENCES

Abramowitz, M. and I.A. Stegun (eds.) (1965): Handbook of Mathematical Functions, Dover, New York, 1046 pp.

Adams, R.A. (1975): Sobolev Spaces, Academic Press, New York, 368 pp.

Adamson, P.T. (1978): The analysis of areal rainfall using multi-quadric functions, Republic of South Africa, Department of Water Affairs, Division of Hydrology, Technical Report TR82, 8 pp.

Anselone, P.M. and P.J. Laurent (1968): A general method for construction of interpolating and smoothing spline functions, Numerische Mathematik, 12, pp 66-82.

Arabelos, D. and I.N. Tziavos (1987): A comparison of two methods to compute degree variances of the gravity field, Manuscripta Geodaetica, 12(4), pp 223-237.

Aronszajn, N. (1950): Theory of reproducing kernels, Transactions of the American Mathematical Society, 68, pp 337-404.

Atteia, M. (1965a): Generalisation de la definition et des proprietes des "spline fonctions", Comptes Rendus de l'Academie des Sciences, Paris, 260, pp 3550-3553.

Atteia, M. (1965b): "Spline-fonctions" generalisees, Comptes Rendus de l'Academie des Sciences, Paris, 261, pp 2149-2152.

Aubin, J.P. (1968): Interpolation et approximation optimales et "Spline-functions", Journal of Mathematical Analysis and Applications, 24, pp 1-24.

Aubin, J.P. (1979): Applied Functional Analysis, John Wiley and Sons, New York, 423 pp.

Boehmer, K. (1974): Spline-Funktionen: Theorie und Anwendungen, B.G. Teubner, Stuttgart, 340 pp.

Constantinescu, F. (1980): Distributions and their Applications in Physics, International Series in Natural Philosophy 100, Pergamon, Oxford, 148 pp.

Cox, D.D. (1984): Multivariate smoothing splines, SIAM Journal on Numerical Analysis, 21(4), pp 789-813.

De Boor, C. (1978): A Practical Guide to Splines, Applied Mathematical Sciences 27, Springer-Verlag, New York, 392 pp.

De Boor, C. and R.E. Lynch (1966): On splines and their minimum properties, Journal of Mathematical Mechanics, 15, pp 953-969.

Delvos, F.J., W. Schaeffer and W. Schempp (1976): Convergence of abstract splines, in Schaback, R. and K. Scherer (eds.), Approximation Theory, Lecture Notes in Mathematics 556, Springer-Verlag, Berlin, pp 155-166.

Delvos, F.J. and W. Schempp (1970): On spline systems, Monatshefte fuer Mathematik, 74(5), pp 399-409.

Delvos, F.J. and W. Schempp (1972): On spline systems: L_m-splines, Mathematische Zeitschrift, 126, pp 154-170.

Delvos, F.J. and W. Schempp (1975): Sard's method and the theory of spline systems, Journal of Approximation Theory, 14, pp 230-243.

Delvos, F.J. and W. Schempp (1976): An extension of Sard's method, in Boehmer, K., G. Meinardus and W. Schempp (eds.), Spline Functions, Lecture Notes in Mathematics 501, Springer-Verlag, Berlin, pp 80-91.

Deny, J. and J.L. Lions (1955): Les espaces du type Beppo Levi, Annales de l'Institut Fourier, 5, pp 305-370.

Dierckx, P. (1984): Algorithms for smoothing data on the sphere with tensor product splines, Computing, 32, pp 319-342.

Donoghue, W.F. (1969): Distributions and Fourier Transforms, Academic Press, New York, 315 pp.

Duchon, J. (1977): Splines minimizing rotation-invariant semi-norms in Sobolev spaces, in Schempp, W. and K. Zeller (eds.), Constructive Theory of Functions of Several Variables, Lecture Notes in Mathematics 571, Springer-Verlag, Berlin, pp 85-100.

Dyn, N. and D. Levin (1983): Iterative solution of systems originating from integral equations and surface interpolation, SIAM Journal on Numerical Analysis, 20, pp 377-390.

Dyn, N., D. Levin and S. Rippa (1986): Numerical procedures for surface fitting of scattered data by radial functions, SIAM Journal on Scientific and Statistical Computing, 7(2), pp 639-659.

Dym, H., and H.P. McKean (1972): Fourier Series and Integrals, Academic Press, New York, 295 pp.

Freeden, W. (1981a): On approximation by harmonic splines, Manuscripta Geodaetica, 6(2), pp 193-244.

Freeden, W. (1981b): On spherical spline interpolation and approxima-

tion, Mathematical Methods in the Applied Sciences, 3, pp 551-575.

Freeden, W. (1982): Interpolation and approximation by harmonic spline functions - theoretical and computational aspects, Bollettino di Geodesia e Scienze Affini, 41(1), pp 105-120.

Freeden, W. (1984a): Spherical spline interpolation - basic theory and computational aspects, Journal of Computational and Applied Mathematics, 11, pp 367-375.

Freeden, W. (1984b): Ein Konvergenzsatz in sphaerischer Spline-Interpolation, Zeitschrift fuer Vermessungswesen, 109(11), pp 569-576.

Gel'fand, I.M. and G.E. Shilov (1964): Generalized Functions, Volume 1, Academic Press, New York, 423 pp.

Hardy, R.L. (1971): Multi-quadric equations of topography and other irregular surfaces, Journal Geophysical Research, 76, pp 1905-1915.

Hardy, R.L. (1977): Least squares prediction, Photogrammetric Engineering and Remote Sensing, 43(4), pp 475-492.

Hardy, R.L. (1983): The biharmonic potential and its geodetic applications, Technical Papers of the 43rd Annual Meeting of the American Congress on Surveying and Mapping, Washington, pp 530-539.

Hardy, R.L. and S.A. Nelson (1986): A multiquadric-biharmonic representation and approximation of disturbing potential, Geophysical Research Letters, 13(1), pp 18-21.

Hayes, J.G. and J. Halliday (1974): The least squares fitting of cubic spline surfaces to general data sets, Journal of the Institute of Mathematics and its Applications, 14, pp 89-103.

Hein, G. and K. Lenze (1979): Zur Genauigkeit und Wirtschaftlichkeit verschiedener Interpolations- und Praediktionsmethoden, Zeitschrift fuer Vermessungswesen, 104(11), pp 492-505.

Heiskanen, W.A., and H. Moritz (1967): Physical Geodesy, W.H. Freeman, San Francisco, 364 pp.

Jarchow, P. (1981): Locally Convex Spaces, B.G. Teubner, Stuttgart, 548 pp.

Jekeli, C. (1983): Achievable accuracies for gravity vector estimation from airborne gravity gradiometry, Proceedings of the International Association of Geodesy Symposia, 18th IUGG General Assembly, Hamburg, Vol 1, pp 577-594.

Jones, B.M. (1970): Deviation of the vertical in South Africa along the parallel 30°S, Proceedings of the Fourth South African Survey

Conference, Durban, 18 pp.

Jones, D.S. (1982): The Theory of Generalized Functions, 2nd Edition, Cambridge University Press, Cambridge, 539 pp.

Jones, M.N. (1985): Spherical Harmonics and Tensors for Classical Field Theory, Research Studies Press, Letchworth, 230 pp.

Kearsley, A.H.W. (1986): The determination of precise geoid height differences using ring integration, Bollettino di Geodesia e Scienze Affini, 45(2), pp 151-174.

Kimeldorf, G.S. and G. Wahba (1970): Spline functions and stochastic processes, Sankya, Series A, pp 173-180.

King, R.W., E.G. Masters, C. Rizos, A. Stolz and J. Collins (1985): Surveying with GPS, University of New South Wales, School of Surveying, Monograph No 9, 128 pp.

Koethe, G. (1979): Topological Vector Spaces II, Grundlehren der mathematischen Wissenschaften 237, Springer-Verlag, New York, 331 pp.

Krarup, T. (1969): A contribution to the mathematical foundation of physical geodesy, Geodaetisk Institut, Copenhagen, Meddelelse No 44, 50 pp.

Krohn, D. (1976): Gravity terrain corrections using multiquadric equations, Geophysics, 41(2), pp 266-275.

Kumar, M. (1983): Status report of the African Doppler Survey (ADOS) as of 15 June 1983, International Coordination of Space Techniques for Geodesy and Geodynamics, Bulletin, 6, pp 47-102.

Lauritzen, S.L. (1973): The probabilistic background of some statistical methods in physical geodesy, Geodaetisk Institut, Copenhagen, Meddelelse No 48, 96 pp.

Lelgemann, D. (1979): Analytical collocation with kernel functions, Bulletin Geodesique, 53, pp 273-289.

Lelgemann, D. (1980): Kollokation und analytische Splinefunktionen, Zeitschrift fuer Vermessungswesen, 105(10), pp 466-479.

Lelgemann, D. (1981): On numerical properties of interpolation with harmonic kernel functions, Manuscripta Geodaetica, 5, pp 157-191.

Lions, J.L. and E. Magenes (1972): Non-Homogenous Boundary Value Problems and Applications, Vol. 1, Grundlehren der mathematischen Wissenschaften 181, Springer-Verlag, Berlin, 357 pp.

Marquardt, D.W. (1970): Generalized inverses, ridge regression, biased linear estimation, and nonlinear estimation, Technometrics, 12(3),

pp 591-612.

Meinguet, J. (1979a): Basic mathematical aspects of surface spline interpolation, in Haemmerlin, G. (ed.), Numerische Integration, International Series of Numerical Mathematics 45, Birkhaeuser Verlag, Basle, pp 211-220.

Meinguet, J. (1979b): A convolution approach to multivariate representation formulas, in Schempp, W. and K. Zeller (eds.), Multivariate Approximation Theory, International Series of Numerical Mathematics 51, Birkhaeuser Verlag, Basle, pp 198-210.

Meinguet, J. (1979c): An intrinsic approach to multivariate spline interpolation at arbitrary points, in Sahney, B.N. (ed.), Polynomial and Spline Approximation, D. Reidel, Dordrecht, pp 163-190.

Meinguet, J. (1979d): Multivariate interpolation at arbitrary points made simple, Journal of Applied Mathematics and Physics (ZAMP), 30, pp 292-304.

Meinguet, J. (1981): From Dirac distributions to multivariate representation formulas, in Ziegler, Z. (ed.), Approximation Theory and Applications, Academic Press, New York, pp 225-248.

Meinguet, J. (1983): Surface spline interpolation: basic theory and computational aspects, Catholic University of Louvain, Institute for Pure and Applied Mathematics, Seminaire de mathematique, Report No 35, 15 pp.

Merry, C.L. (1977): Gravity and the South African height system, South African Survey Journal, 16(1), pp 44-53.

Merry, C.L. (1980): A practical comparison of some methods of predicting point gravity anomalies, Manuscripta Geodaetica, 5, pp 299-314.

Merry, C.L. (1985): Transformation parameters for the Cape Datum, South African Survey Journal, to appear.

Merry, C.L. and H. van Gysen (1987): A regional geoid for southern Africa, 19th General Assembly of the International Union of Geodesy and Geophysics, Vancouver, 34 pp.

Meschkowski, H. (1962): Hilbertsche Raeume mit Kernfunktion, Grundlehren der mathematischen Wissenschaften 112, Springer-Verlag, Berlin, 256 pp.

Micchelli, C.A. (1986): Interpolation of scattered data: distance matrices and conditionally positive definite functions, Construc-

tive Approximation, 2, pp 11-22.

Moritz, H. (1980): Advanced Physical geodesy, Herbert Wichmann, Karlsruhe, 500 pp.

Mueller, C. (1966): Spherical Harmonics, Lecture Notes in Mathematics 17, Springer-Verlag, Berlin, 45 pp.

Numerical Algorithms Group (1983): NAG FORTRAN Library Manual, Mark 10, Oxford.

Petersen, B.E. (1983): Introduction to the Fourier Transform and Pseudo-Differential Operators, Pitman, Boston, 356 pp.

Pryce, J.D. (1973): Basic Methods of Linear Functional Analysis, Hutchinson, London, 320 pp.

Rapp, R.H. (1986): The gravity field of the earth and prospects for improvement of its determination, in Anderson, A.J. and A. Cazenave (eds.), Space Geodesy and Geodynamics, Academic Press, London, pp 55-74.

Reilly, J.P. and E.H. Hebrechtsmeier (1978): A systematic approach to modeling the geopotential with point mass anomalies, Journal of Geophyisical Research, 83(B2), pp 841-844.

Riesz, F. and B. Sz.-Nagy (1955): Functional Analysis, Frederick Ungar, New York, 468 pp.

Rudin, W. (1973): Functional Analysis, McGraw-Hill, New York, 397 pp.

Sandwell, D.T. (1987): Biharmonic spline interpolation of GEOS-3 and SEASAT altimeter data, Geophysical Research Letters, 14(2), pp 139-142.

Sard, A. (1967): Optimal approximation, Journal of Functional Analysis, 1, pp 222-244.

Schempp, W. and U. Tippenhauer (1974): Reprokerne zu Spline-Grundraeumen, Mathematische Zeitschrift, 136, pp 357-369.

Schoenberg, I.J. (1964): On trigonometric spline interpolation, Journal of Mathematics and Mechanics, 13(5), pp 795-825.

Schumaker, L.L. (1981): Spline Functions: Basic Theory, Wiley, New York, 553 pp.

Schwartz, L. (1950): Theorie des distributions, Vol. 1, Actualites Scientifiques et Industrielles 1091, Hermann, Paris, 148 pp.

Schwartz, L. (1951): Theorie des distributions, Vol. 2, Actualites Scientifiques et Industrielles 1122, Hermann, Paris, 169 pp.

Schwartz, L. (1964): Sous-espaces hilbertiens d'espaces vectoriels

topoligiques et noyaux associes (noyaux reproduisants), Journal D'Analyse Mathematique, 13, pp 115-256.

Schwartz, L. (1968): Applications of Distributions to the Elementary Particles in Quantum Mechanics, Gordon and Breach, New York, 134 pp.

Schwartz, L. (1973): Radon Measures on Arbitrary Topological Spaces and Cylindrical Measures, Oxford University Press, Oxford, 393 pp.

Schwarz, K.P. (1982): On the unexpected usefulness of theory in geodesy, Proceedings of the Canadian Institute of Surveying Centennial Convention, Ottawa, April 1982, pp 453-466.

Schwarz, K.P. (1984): Data types and their spectral properties, Lecture to the International Summer School on Local Gravity Field Approximation, Beijing, 65 pp.

Schwarz, K.P., M.G. Sideris and R. Forsberg (1987): Orthometric heights without levelling, ASCE Journal of Surveying Engineering, 113(1), pp 28-40.

Sideris, M.G. and K.P. Schwarz (1986): Solving Molodensky's series by fast Fourier transform techniques, Bulletin Geodesique, 60, pp 51-63.

Spiegel, M.R. (1968): Mathematical Handbook of Formulas and Tables, Schaum's Outline Series, McGraw-Hill, New York, 271 pp.

Stein, E.M, and G. Weiss (1971): Introduction to Fourier analysis on Euclidean spaces, Princeton University Press, Princeton, 297 pp.

Sternberg, W.J. (1944): The Theory of Potential and Spherical Harmonics, Mathematical Expositions No. 3, University of Toronto Press, Toronto, 312 pp.

Stromeyer, D. and L. Ballani (1984): Uniqueness of the inverse gravimetric problem for point mass models, Manuscripta Geodaetica, 9(1-2), pp 125-136.

Suenkel, H. (1984a): Splines: their equivalence to collocation, Ohio State University, Dept. of Geodetic Science and Surveying, Report No. 357, 55 pp.

Suenkel, H. (1984b): New numerical techniques (splines, FFT, point masses), Lecture to the International Summer School on Local Gravity Field Approximation, Beijing, 41 pp.

Treves, F. (1967): Topological Vector Spaces, Distributions and Kernels, Academic Press, New York, 565 pp.