SCALE–BASED GEOMETRY OF NONDIFFERENTIABLE FUNCTIONS, MEASURES, AND DISTRIBUTIONS

JAMES DAMON

Department of Mathematics
University of North Carolina at Chapel Hill,
Chapel Hill, NC 27599-3250, USA

PRELIMINARY VERSION

Introduction

In this paper, we consider the problem of associating geometric structures to discrete objects arising in various "real world" situations and to the mathematical objects used to model them, so that we may capture geometric properties of these objects. Such "real world" objects typically may be: *discretely defined functions* which take constant values on the cells of a grid, as for example illustrated by 2 or 3–dimensional computer images viewed as intensity-valued functions on a grid; *piecewise linear approximations* to surfaces obtained by sampling at a finite number of points; *discrete measures* obtained from finite data sets of points; *numerical solutions* to partial differential equations, etc. as in Fig. 1.

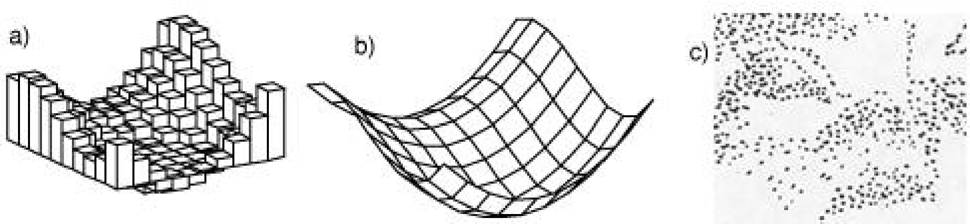

FIGURE 1. a) Discrete Function, b) Piecewise Linear Function, and c) Discrete Point Set

Such objects are nondifferentiable, even noncontinuous, and they are subject to "noise and errors". They are described using mathematical objects typically treated by methods of analysis such as: L^p–functions or special subspaces such as functions in Sobolev spaces; measures, including regular Borel measures, probability measures, or signed measures; and more generally distributions, including tempered distributions and distributions with compact support.

Such objects are usually not candidates for geometric analysis, although we do "see"geometric features in the objects in Fig. 1. However, we do not speak

Partially supported by a grant from the National Science Foundation.

about geometric properties such as curvature properties of L^p functions, nor the shape or other geometric features of measures. This is because traditionally, geometrical methods are applied to differentiable functions and manifolds to obtain differential–geometric properties. Suppose we wish to extract geometric features from a discrete function or measure, or the corresponding mathematical objects. One possible approach would be to first approximate them using differentiable functions or distributions. For discrete objects this would involve some method such as for functions using splines or surface–fitting or for measures, approximations using differentiable distributions. Then, geometric properties could be extracted from the approximations.

There are several problems with this approach. First, it is not clear whether we could expect reasonable geometric structures as we use finer approximations by differentiable functions, etc. Second, by their very nature, the topologies on these spaces essentially disregard differentiability properties; hence, such an approximation process would be unstable for geometric properties that depended upon derivative properties. Thus, it is unclear to what extent any associated geometric properties can be made independent of the approximation. Despite these problems, our goal will be to show how to associate in an intrinsic way geometric structures to both discrete objects and their mathematical representations, whether they be nondifferentiable functions, measures, or distributions.

In fact, our goal will be to associate scale–based geometric structures satisfying the following properties.

Principal Properties for Scale-based Geometry:

(1) *Applicability to Discrete and Nondifferentiable Objects :* It should be possible to associate such geometric structures to both discrete "real world" objects and the mathematical entities modeling them.
(2) *Genericity :* For "almost all" objects, the associated geometric structures should have a structure describable in terms of a simple "catalogue" of possible local properties.
(3) *Stability :* For "almost all" objects, the geometric structure should be structurally stable. This means that under sufficiently small perturbations within the space of objects, the perturbed geometric structures should be "geometrically equivalent" to the initial one.
(4) *Generic Transitions :* There should be a specific list of generic transitions which occur in the geometric structure under deformations.
(5) *Applicability to Associated Objects :* Geometric structures should be obtained, not just for the original objects, but also for associated objects which capture specific geometric feature types appearing in the objects.

It is impossible to satisfy all of these conditions using classical geometric methods (applied to differentiable approximations), especially if we want to satisfy both 1) and 3). We will overcome this intrinsic incompatibility via "scale–based geometry" which is based on two ingredients. The first introduces "scale" as an additional parameter implicit in the analysis of "real world objects"; and the second concentrates on "qualitative geometry".

By "qualitative geometry" of differentiable functions, submanifolds, or distribution functions, we mean differential geometric properties identified not by particular values for invariants or quantities (e.g. curvature), but rather by identifying: regions where the invariants have specific properties; special subsets where various quantities take extremal values; or subsets separating regions where different properties are exhibited. Such subsets can be typically defined by algebraic conditions on derivatives and can be analyzed using singularity theoretic methods.

However, such properties are not intrinsic for nondifferentiable functions. In general, there is even a lack of stability of geometric properties/structure under small perturbations of smooth functions if we only can control them using the C^0 topology. This problem can be overcome through the introduction of scale. Our use of scale contrasts with the standard usage in applied mathematics, where the term "scale"refers to identifying specific scales of interest in a problem for modeling; however, it shares similarities with the notion of scale–invariance properties of equations (see e.g. [Ba]).

We are conerned instead with the interaction of different properties at different and varying scales. This approach to scale follows the idea of introducing scale for problems in computer imaging, due to Witkin via "Gaussian blurring"[Wi] (also see Ijima [WII]). Gaussian blurring is defined by convolution with a Gaussian kernel and introduces an extra variable "scale". It was originally introduced to both overcome problems with small perturbations by "noise" and also to provide a measure of scale for geometric features. Koenderink [K] showed the essential nature of scale for questions involving imaging, and deduced that for scale to satisfy certain intrinsic properties, it must be given by Gaussian blurring. Gaussian blurring leads, via the "semigroup"property, to solutions of evolution equations, especially the heat equation. There has since evolved a large body of results devoted to properties of Gaussian blurring, scale space, and variants such as nonlinear blurring, see e.g. the books [tHR], [Li], [Sp], and [Mo] and many additional references therein.

In this paper, we apply scale in a broader sense than that given by Gaussian blurring, by allowing convolution with any of a large class of "scale–based kernels", which are characterized by their scaling behavior under the action of a scaling group. These include among many others: Gaussian kernels, their derivatives, and certain other kernels obtained by modifying Gaussians. These kernels may depend on multi–parameter scales for tracking several different scales for different features. We consider the properties of scale space resulting from the action of a "Poincaré Scaling group"acting on scale space, with the corresponding behavior of convolutions with the scale– based kernel.

This approach will show that in a very general sense, scale provides a bridge allowing classical geometric methods to be applied to a much larger classes of objects, the tempered distributions, and to specific subspaces which include nondifferentiable L^p-functions, Borel and/or probability measures, and distributions of compact support. Furthermore, we will give criteria ensuring that the application of scale–based geometry to these objects satisfies the genericity and stability properties listed above.

Scale–based geometry involves the interaction of three classes of objects: 1) the scale–based kernels, 2) the geometric structures capturing the various possible geometric properties, and 3) the various subspaces of tempered distributions to which we would like to apply scale–based versions of the properties. First, we give general criteria on the scale–based kernel, the specific scale-based geometric properties, and the subspaces of distributions so that on a given compact region C of scale space, there exists an open dense subset of distributions in the subspace which possess generic scale–based geometric properties. This applies to the full space of tempered distributions (Theorem 5.8) and to subspaces that satisfy a general criterion (Theorem 6.3); in particular this criterion is satisfied by many standard subspaces (Theorem 6.4 and Corollary 6.5), including spaces of L^p-functions, Sobolev spaces, spaces of regular Borel or probability measures, and distributions of compact support.

Second, the local properties of the scale–based geometric structure are given in terms of the local structure of an associated "closed Whitney stratified set"(Theorems 5.10 and 6.3). This representation leads to a specific criteria (Proposition 8.1) for determining whether the scale–based properties agree with those for classical geometry, or how they must be modified. However, the same generic scale–based properties which occur for smooth functions will also occur generically for nondifferentiable functions, measures, etc.

Third, we deduce the structural stability of the associated geometric structures on the compact region C of scale space under all sufficiently small perturbations (Theorems 5.11 and 6.3). This structural stability holds for perturbations within the full space of tempered distributions as well as within specific subspaces. As a consequence, generic nondifferentiable functions, measures, etc, can be approximated by differentiable functions, distributions, etc. in many different ways, yet for any sufficiently close approximation, we obtain the same scale-based geometric properties on the given compact region of scale space.

These results are proven using a "relative transversality theorem", which is an extension of the classical Thom Transversality theorem and applies to subspaces of distributions under convolution [D2]. This requires an introduction of convolution jet space to replace the classical jet space for smooth functions (§4). The structure of this convolution jet space results from the scaling properties of the scale–based kernel under the scaling group. We then prove the genericity and stability results using transversality results involving closed Whitney stratified sets.

Fourth, the method of proof allows us to deduce that even discrete and piecewise linear functions and discrete measures can exhibit geometric properties generically on compact subsets of scale space (Corollaries 7.4 and 7.5). While at first it seems surprising that any discrete object can exhibit geometric properties generically, we give several senses in which "most discrete functions"have this property (Theorems 7.6 and 7.8).

Fifth, we show that the criterion for scale–based kernels which we have mentioned is valid for simple Gaussian scale kernels and more generally the class of "extended Gaussian kernels"obtained from Gaussian kernels by allowing anisotropy, applying derivatives, and allowing multiple independent scales (Proposition 9.4). We extend the results about a basis of solutions for the heat equation as given in [D1] to obtain a basis for convolution jet space for this more general class of kernels (Corollaries 9.7 and 9.9). The criterion for geometric properties is valid for specific geometric properties associated to: classical differential geometry by, e.g. Porteous [Po1], [Po2], Koenderink [K2], Bruce-Giblin et al [BGG], [BGT], Mumford [Mu]; edge–based geometry Canny [Ca], Rieger [Ri1], [Ri2], medial geometry defined via "ridges"by Pizer-Eberly et al [Eb1], [Eb2], [PE], "watershed regions"[NO], and more generally "relative critical sets"(this author [D3], [D4], Miller [Mi], and Keller [Ke]), as well as general higher order differential invariants by ter Haar Romeny et al [RFSV] (see §8). Again, the criterion for subspaces is valid for subspaces of L^p functions, spaces of measures such as positive Borel measures, probability measures, etc, and full spaces of tempered distributions or distributions of compact support.

Sixth, we also allow derived distributions which arise by applying a discrimination criteria (e.g. as in computer imaging identifying characteristic properties of features, textures, etc.). These criteria may be defined by applying "discrimination filters", (distributional) differential operators, or statistical methods yielding functions, measures or more general distributions which identify the "amount of a property" occurring in a given region of space. We may then extract scale–based geometric properties of these discrimination distributions. Because the discrimination process typically involves a specific scale, genericity results can only be expected for regions that are at a sufficiently larger scale. Hence, results are described in terms of genericity of scale–based properties occurring on a given compact region C of scale space.

We give sufficient conditions that the geometric properties of these discrimination distributions are given by generic geometric structures on C for a dense open subset of the space of initial distributions (Proposition 10.2 for general operators, and Theorem 10.5 for linear operators). Thus, we deduce geometric properties of the regions where the discriminated properties occur (§10). We apply the criteria for linear operators to three situations to obtain genericity of scale–based geometry on a given C. These include partial differential operators (Corollary 10.6), a linear operation which can approximate "block functions" (see Fig. 2) on cells of a sufficiently fine mesh (Corollary 10.7), and a specific simple form of texture discrimination which involves integral comparison with fixed "mask functions" on the cells of a sufficiently fine mesh (Corollary 10.9). As a block function is a discrete approximation to a Dirac δ–function, all three of these conditions involve approximating a δ–function by discrimination distributions.

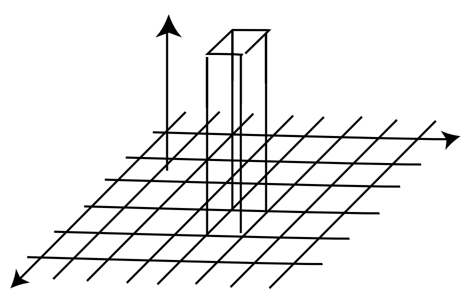

FIGURE 2. A Block function on a square mesh for $\mathbb{R}^2$

If the initial object can be analyzed using a number of different discrimination criteria, these yield discrimination distributions which together define a vector-valued tempered distribution. The scale–based geometric properties of these vector-valued distributions identify how the identified features interact on various regions. Furthermore, we allow different scaling actions for the different distributions, reflecting the fact that the properties are occurring at independent scales. Even with different scale–based kernels and different scaling actions for each component, we still show that scale–based geometric properties of vector–valued tempered distributions satisfy both genericity and stability conditions (§11). Lastly, we combine the results of

§10 and §11 to provide a criterion involving "generic independence of discrimination features" which guarantees the genericity and stability of multi–feature geometry for an open dense set of initial distributions (Proposition 12.2). Using this, we give a sufficient condition, in terms of independence conditions for the mask functions on individual cells, that a simple form of multi–texture discrimination will have generic scale–based geometric properties (Corollary 12.5). This gives evidence that more general multi-texture detection schemes will likewise exhibit generic scale–based properties.

Finally, in Part 4, we consider how these methods apply to families of distributions yielding generic transitions occurring within families. To understand the transitions that can occur in families, we must understand how transversality to closed Whitney stratified sets can fail in families. We apply the notion of $\mathcal{K}_V$–equivalence from singularity theory to analyze the failure of transversality to a subset V (§13). We apply it in several different settings, as the families of convolved functions on scale space may depend on parameters in a number of ways. We can view the scaling parameters themselves as parameters (§14). Alternately, we can view the families as depending on external parameters (§16). Third, we can consider kernels which contain auxillary parameters and hence introduce parameters for the convolution functions on scale space (§15).

CONTENTS

1. From Classical to Scale–based Geometry: An Overview

Classical Geometry. By "classical methods of geometry", we are not referring to geometry from a time period, but rather to the method of obtaining geometric properties by applying differential–geometric methods to differentiable functions or submanifolds. The most obvious example is *classical differential geometry.* Within it we often want to be able to identify regions where specific geometric features are distinguished. For a surface in $\mathbb{R}^3$, distinguished curves such as "crests", "ridges", and parabolic curves capture properties of the surface. These curves have "generic properties"see e.g. Porteous[Po1], [Po2], Koenderink [K2], Bruce-Giblin-Tari [BGT], etc. Very general results on generic geometric properties of differentiable manifolds are given in Wall [Wa]. For functions, there are more possiblilties. If we view the graph of the function, we can again consider its generic differential geometry; however, the dependent variable may not be commensurate with the independent variables, so such differential geometric information will often not have intrinsic meaning.

However, there are alternate ways to extract geometric information directly from the function. These include:

(1) *Edge–Based Geometry* : Regions of relative values for the function are separated by "edges", where the change in values are most significant. Some possibilities include: the "Canny edge"[Ca], and its higher dimensional analogues, see e.g. Rieger [Ri4].
(2) *Level Set Geometry* : The functions level sets themselves have differential geometric properties, such as extremals of curvature, which reveal geometric information about the function. Gauch [Gau], Rieger [Ri3]. Furthermore, there are "level set methods"for allowing regions to evolve as level sets to detect certain geometric regions, Osher–Sethian [OS], [Sn].
(3) *Medial Geometry* : One of the very first geometric structures to be associated to a region with smooth boundary was the Blum medial axis [BN], which associates a skeletal structure capturing the shape of the region. This has been extended to functions using the notion of "ridges", which can be defined in a number of different ways [Eb1]. The height ridges of Pizer and Eberly [PE] exhibit medial properties quite different from those of the medial axis, see e.g. [MPP], [MPL]; however, they can be placed into a larger "relative critical set"structure which explains the properties of height ridges and has properties with some advantages over the Blum medial axis ([D3], [D4], [Mi], [Ke]). Medial properties can also be alternately investigated using optimal parameter height ridges [Fr], [Fu2]. As well, a watershed definition of medialness leads to segmentation into watershed regions [NO].
(4) *Higher Order Differential Invariants* : Using invariant theory, it is possible to determine the expressions involving higher order derivatives which are invariant under translation and rotation,ter Haar Romeny et al [RFSV]. Geometric properties which are invariant under rotations and translations can then be expressed via algebraic conditions involving these differential invariants.

Genericity and Stability via Transversality. The preceding geometric properties can be characterized in terms of algebraic conditions on derivatives (both algebraic equalities and inequalities). Such conditions define for functions defined on $U \subset \mathbb{R}^n$, "semialgebraic"sets W in jet space $J^\ell(U, \mathbb{R})$, which consists of ℓ-jets at points

$(x, y) \in U \times \mathbb{R}$ (i.e. the ℓ–th order Taylor expansions at (x, y)). Such sets W can be decomposed into smooth pieces $\{W_i\}$ which form a "Whitney stratified subset"of jet space (see §5).

For classical geometry, if we were considering all smooth functions $C^\infty(U, \mathbb{R}^p)$, then given $f \in C^\infty(U, \mathbb{R}^p)$, there is the "jet-extension map"$j^\ell(f) : U \to J^\ell(U, \mathbb{R}^p)$ sending $x \mapsto j^\ell(f)(x)$, the ℓ–jet of f at $(x, f(x))$. Then, the standard Thom Transversality Theorem (see e.g. [GG, Thm 4.9]) asserts that the set $\mathcal{W}$ of smooth functions f on U for which $j^\ell(f)$ is transverse to W forms a residual subset (or alternately of the second category) in $C^\infty(U, \mathbb{R}^p)$. Because $C^\infty(U, \mathbb{R}^p)$ is a Baire space, $\mathcal{W}$ is dense. It also follows that the set of smooth functions $\mathcal{W}$ for which $j^\ell(f)$ is transverse to a countable number of smooth submanifolds $\{W_i\}$ is still residual. The standard Thom transversality theorem may be further strengthened when the collection $\{W_i\}$ are the strata of a closed Whitney stratified set W. Then, a strengthened form of the Thom transversality theorem (see [GMc]) asserts that $\mathcal{W}$ is open and dense.

Example 1.1 (Genericity for classical differential–geometric properties)**.** The generic properties of curves in surfaces such as ridges, crest, parabolic lines, etc. have been determined and can be described by singularity classification of height functions and distance–squared functions, see e.g. [Po1], [Po2], [K2], [BGT], and more generally [Wa]. This classification is given by algebraic conditions on partial derivatives.

The generic properties of Canny edges for functions in the plane have been determined by Rieger [Ri1], and given in terms of algebraic conditions on derivatives of the functions. They have several unexpected generic properties, such as genericity of corners, that differ from what one would expect from an "edge". This has been partially extended to higher dimensions in [Ri4].

For the evolution of level surfaces under mean curvature flow, the case of level curves has remarkable properties discovered by Gage, Hamilton [Ga], [GH] and Grayson [Gr]. These hold for all sufficiently differentiable curves. However, already for level surfaces in $\mathbb{R}^3$, the behavior becomes more subtle and is not fully understood. Generic geometric properties of level curves are given in [Ri3].

For the Blum medial axis, there is a large body of work devoted to its computation and properties. Among the properties of the medial axis that have been worked out include: the generic properties in dimensions ≤ 6, Bryzgalova [Brz], Yomdin [Y], Mather [M3], stability [M3], deformation properties in one-parameter families, Bogaevski [Bg], the relation between the geometry of the region boundary and that of the medial axis, Nachman-Pizer [NaP], and the representation of the medial axis as a shock set of an evolution equation for the boundary, Kimia-Tannenbaum-Zucker [KTZ]. The analysis of generic properties was further extended to a related object the symmetry set Bruce-Giblin-Gibson [BGG] [BG].

The generic properties for height ridges follows from the determination of the properties of the more general relative critical set. For relative critical sets, the generic properties, stability, and generic transitions in one and two parameter families have been determined in [D3], [D4], Miller [Mi], Keller [Ke]. The generic properties for optimal parameter height ridges have been determined in the simplest case of 1–dimensional ridges and a single parameter [Mi]. For medial structure associated to watersheds, the generic structure follows from the structure of stable, unstable and connecting manifolds for gradient dynamical ststems due originally to Smale [Sm] (see also [NO] for watersheds in computer imaging).

Introduction of Scale. All of this says nothing about the nondifferentiable objects we considered in the introduction. To consider them, we view them as tempered distributions, and introduce scale by convolving them with a fixed "scaling kernel" $G(x, \lambda)$, for $x \in \mathbb{R}^n$ and scaling parameters $\lambda \in \Lambda \subset \mathbb{R}^k_+$. The simplest such scaling kernel is the Gaussian kernel, with a single scale parameter. However, ultimately, we shall consider several different objects each with independent or interacting scales, leading us to consider multiple kernels with multiple independent scales. Hence, we specifically want to consider kernels with scaling parameters, and ultimately we consider "distributions with parameters". Thus, we extend consideration to "uniformly tempered distributions" to allow both scaling parameters and deformation parameters.

First, we enlarge the Schwartz space to the space of uniformly rapidly decreasing functions. Then, we replace the space of tempered distributions by a larger space of "uniformly tempered distributions" on $\mathbb{R}^n \times \Lambda$. This space allows us to consider tempered distributions depending on scaling parameters, and contains the tempered distributions as a subspace. Then, there is an analogous convolution map from uniformly tempered distributions. Provided G is "uniformly rapidly decreasing", we now obtain smooth functions $u * G \in C^\infty(R^n \times \Lambda, \mathbb{R})$, where the $\lambda \in \Lambda$ measures certain scales on $\mathbb{R}^n$. This convolution map is shown to be continuous for the regular C^∞–topology (Theorem 3.6) with image denoted by $\mathcal{H}_{\Lambda,G}$. Restricted to the subspace of tempered distributions, it yields a subspace $\mathcal{H}_G$.

The subspaces of smooth functions $\mathcal{H}_{\Lambda,G}$ or $\mathcal{H}_G$ form rather small subspaces of $C^\infty(\mathbb{R}^n \times \Lambda, \mathbb{R})$; thus, there is no a priori reason to expect them to satisfy any transversality conditions; they may entirely miss any $\mathcal{W}$ given by the Thom transversality theorem.

To overcome the nonapplicability of the Thom transversality theorem to the subspace of functions $\mathcal{H}_G$ arising as convolutions with G, we will apply instead a "relative transversality theorem" [D2, Thm 1.3] which replaces the jet space $J^\ell(\mathbb{R}^n \times \Lambda, \mathbb{R})$ by the convolution jet spaces $\mathcal{H}^\ell_G$ or $\mathcal{H}^\ell_{\Lambda,G}$ consisting of ℓ–jets of functions in $\mathcal{H}_G$ or $\mathcal{H}_{\Lambda,G}$. To be able to apply this theorem, we must know that the convolution jet spaces have a sufficiently nice structure. For a general linear subspace $\mathcal{H} \subset C^\infty(U, \mathbb{R})$, the associated jet space $\mathcal{H}^\ell$ need not even have fibers of constant dimension. For example, if $\mathcal{H}$ is spanned by a single function f, then if x_0 is a critical point with $f(x_0) = 0$, then the fiber $\mathcal{H}^1_{x_0} = 0$.

To determine the structure of the convolution jet spaces $\mathcal{H}^\ell_G$ and $\mathcal{H}^\ell_{\Lambda,G}$, we use the action of the "Poincaré scaling group" PS, which is formed as a semi–direct product of the scaling group and translations. Provided the scaling kernel G is a "scale–based kernel", which means it satisfies certain scaling conditions, we show that the convolution jet spaces are trivial fiber bundles which are also semialgebraic submanifolds (§4). This allows us to use the relative transversality theorem.

Genericity and Stability for Scale–Based Geometry. For a property $\mathcal{P}$ defined by transversality to a closed Whitney stratified set $W \subset \mathcal{H}^\ell_G$ or $\mathcal{H}^\ell_{\Lambda,G}$, we first apply the relative transversality theorem to conclude that for a compact subset of scale space $C \subset \mathbb{R}^n \times \Lambda$, there is an open dense subset of (uniformly) tempered distributions u whose convolutions satisfy $j^\ell(u * G)$ is transverse to W relative to $\mathcal{H}^\ell_G$, so that with u exhibits $\mathcal{P}$ transversely on C (theorem 5.8). Then, we may apply the Thom isotopy theorem [Th], [M2] to conclude that on the subset C, such a generic distribution exhibits $\mathcal{P}$ on a Whitney stratified subset whose local structure is determined by that of W (Theorem 5.10). Moreover, a further application, together with arguments of Mather, allow us to conclude that the subset where $\mathcal{P}$ is exhibited on C is structurally stable under any sufficiently small perturbation in the space of (uniformly) tempered distributions.

In the preceding, we would like to furthermore know that if, for example, we have a nongeneric positive Borel measure, then we can make an arbitrarily small perturbation to another near positive Borel measure which has generic scale–based geometric properties. We identify conditions for the scale–based kernel G (condition (A)) and for a subspace $\mathcal{T}$ of tempered distributions (condition (B)) which imply that the preceding conclusions (especially the density) hold within the subspace $\mathcal{T}$. This condition (B) is satisfied by subspaces we have already mentioned such as the L^p functions, positive (or signed) regular Borel measures, etc. (Theorems 6.3 and 6.4). Moreover, condition (A) is valid for a large class of "Extended Gaussian Kernels", obtained from Gaussian kernels by allowing anisotropy, multiple independent scales, and applying derivatives (§9). Thus, we are able to conclude that all of the subspaces we have mentioned will generically exhibit properties on scale space for any of the extended Gaussian kernels.

Furthermore, the transversality arguments allow us to directly obtain conclusions for genericity for discrete or piecewise linear functions and discrete measures. Although the discrete functions and measures form a very restrictive class of nondifferentiable functions or measures, we still establish the density of the set of discrete functions and measures which exhibit a scale–based property generically on a compact subset of scale subspace (Corollaries 7.4 and 7.5). Although openness is meaningless for sets of discrete objects, we explain how nongeneric discrete functions and measures can be approximated by generic ones obtained by refining the defining mesh (Theorems 7.6 and 7.8).

To apply these results to specific geometric properties, we must be able to translate classical geometric conditions mentioned above to scale–based conditions. Because the conditions defining a classical geometric property are given by semi–algebraic conditions, the (closure of the) subspace defined by them is a closed semi–algebraic subset W'. The properties of $\mathcal{H}^\ell_G$ imply that its intersection with $\mathcal{H}^\ell_G$ is again a closed semi–algebraic subset W. Then, by a theorem of Lojasiewicz [Lo], W still has a Whitney stratification. Hence, we obtain scale –based versions of the properties by transversality to W relative to $\mathcal{H}^\ell_G$ (§8).

If W intersects $\mathcal{H}^\ell_G$ transversally, then we show by a fiber square argument that the geometric structure and local properties associated to $\mathcal{P}$ for convolutions with G will exhibit exactly the same generic properties as will smooth functions. This is verified e.g. for Canny edges [Ri2] and relative critical sets [D4], [Mi], and [Ke].

If the intersection is not transverse, then convolutions with G will still exhibit generic versions of property $\mathcal{P}$. They may now display different properties from those exhibited by smooth functions. However, the exact form of the generic properties still follow from the local structure of the intersection $W \subset \mathcal{H}^\ell_G$ as a closed Whitney stratified subset (§8).

Scale–based Geometry for Texture/Feature Discrimination. These results can be further extended to associated distributions defined by discrimination filters or statistical methods yielding measures which capture the amount of a property in given regions of space. In contrast with the results concerning subspaces $\mathcal{T}$, the operation Ψ which associates to an initial distribution u an associated discrimination distribution $v = \Psi(u)$ will typically introduce specific scales. Thus, the scales at which we can expect genericity and stability must be sufficiently larger so that sufficiently many distributions can arise as associated objects. Thus, we consider a fixed compact subset of scale space C, and ask whether the associated discrimination distributions will exhibit a given scale–based geometric property generically and stably on C. We give sufficient conditions in terms of C, in both an abstract form (Proposition 10.2) and a concrete one in the case of linear Ψ (Theorem 10.5), which ensure that there will again be a dense open subset of initial distributions having the property that the associated distibutions will exhibit a scale–based property $\mathcal{P}$ generically on C . Because we have established genericity for discrete functions in (§7), we translate Theorem 10.5 into an explicit condition of being able to approximate block functions on sufficiently fine mesh by discrimination distributions. We give a specific simple method for texture discrimination involving integration against given mask functions on cells of a mesh. For this texture discrimination, the criterion yields that if the mesh is sufficiently fine (depending on C), then for an open dense set of initial distributions, their associated discrimination functions will exhibit scale–based geometric structures generically and stably on C (Corollary 10.9).

Multifeature Geometry. Furthermore, we may wish to simultaneously compare the properties of a number of different geometric features/textures. We do so in two steps. First, we consider spaces of p–tuple of (uniformly) tempered distributions, or equivalently $\mathbb{R}^p$–vector valued (uniformly) tempered distributions $u = (u_1, \dots, u_p)$. We wish to allow multiple independent scales, so we consider p–tuples of scale–based multi–kernels $G = (G_1, \dots, G_p)$, where there is a single scaling group $\mathbb{R}^k_+$; however, we allow it to act differently for each kernel G_i. Now there is no single geometric action of the Poincaré scaling group PS on scale space, but rather it acts coordinatewise on the convolved distributions $u * G = (u_1 * G_1, \dots, u_p * G_p)$. We again define the image spaces under convolution $\mathcal{H}^{(p)}_{\Lambda,G}$ (respectively $\mathcal{H}^{(p)}_G$), which are products of the individual $\mathcal{H}_{\Lambda,G_i}$ (respectively $\mathcal{H}_{G_i}$). It follows that the corresponding convolution jet spaces $\mathcal{H}^{\ell,(p)}_{\Lambda,G}$ (respectively $\mathcal{H}^{\ell,(p)}_G$) are fiber products of the convolution jet spaces for the individual kernels. Although there is no geometric action of PS as for the case of a single distribution, nonetheless, the earlier results on the structure of convolution jet spaces from (§4) together with the fiber product representation allow us to conclude that both $\mathcal{H}^{\ell,(p)}_{\Lambda,G}$ and $\mathcal{H}^{\ell,(p)}_G$ possess the same properties as the convolution jet spaces from (§4).

Then, we can apply analogous arguments as in (§5 and §6) to obtain for scale–based geometric properties of vector– valued distributions, genericity via transversality (Theorem 11.3), local generic structure and stability (Theorem 11.8), and corresponding results for subspaces (Theorem 11.10). In particular, it follows that given distinct scale–based geometric properties $\mathcal{P}_i$ defined by transversality to closed Whitney stratified sets $W^{(i)}$for the individual distributions u_i, and a compact subset C of scale–space, there is a dense open subset of vector–valued distributions u

for which the u_i exhibit $\mathcal{P}_i$ generically on C and moreover, the geometric structures $W^{(i)}(u_i * G_i)$ will intersect in general position.

To extend the preceding to multiple features of a single initial distribution, we combine the results of §10 and §11. We obtain a criterion for a collection of p discrimination operators Ψ_i defined on a subspace $\mathcal{T}$ being "generically independent on C". We show (Proposition 12.2) that this generic independence implies there is an open dense subset of initial distributions $u \in \mathcal{T}$ so that the vector–valued distribution $(\Psi_1(u), \dots, \Psi_p(u))$ will exhibit a scale–based geometric property $\mathcal{P}$ generically and stably on C. Again this translates into concrete criteria in terms of approximating a collection of independent block functions. The simple method of texture discrimination can be applied to multiple textures. The criterion implies that if the mask functions on a given cell for the various textures are mutually independent in the sense of Corollary 12.5, then the geometric properties of each texture discrimination function will be exhibited generically on C with the geometric structures in general position.

Generic Transitions in Families. The preceding results do not distinguish the scale parameters, nor consider the effect of varying the extra parameters appearing in kernels, nor allowing the distributions to depend on extra parameters. All of these parametrized versions for scale–based geometry are considered in Part 4. The crucial difference is that now, as parameters vary, the generic behavior may fail for certain parameter values. This is due to the loss of transversality to the closed Whitney stratified set(s). The investigation of this failure and the generic transitions that result from it can be carried out using more refined methods from singularity theory for an equivalence denoted $\mathcal{K}_V$-equivalence, which captures properties of nonlinear (nontransverse) sections of varieties V. We explain the general methods for applying $\mathcal{K}_V$-equivalence, including the use of versal unfoldings to describe generic transitions in §13. We then investigate how these general methods apply in the three situations involving parameters already mentioned in §14, §15, and §16. These include allowing external parameters, and deducing generic transitions in scale space as was done for relative critical sets for one and two parameter families in [D4] and [Ke]. Alternately, generic transitions in scale space can be determined varying scale parameters, as for level surface transitions under Gaussian blurring [D1] or Canny edges of Gaussian blurred functions in [Ri2]. Third, we consider transitions which result by varying extra parameters in kernels, as arise in edge detection using edge kernels, optimal parameter height ridges [Fu2], such as optimal scale ridges [Fr], whose generic properties were determined by Miller [Mi].

Part 1. Genericity and Stability Theorems for (Uniformly) Tempered Distributions

In Part 1, we will follow the outline introduced in §1 for establishing the genericity and stability of scale–based geometric properties. To do so we will begin by investigating properties of convolutions of tempered distributions with kernels which have scaling parameters. This will be carried out in the framework of "uniformly tempered distributions"which can be thought of as "distributions depending on parameters".

We begin in §2 by recalling basic properties of both the Schwartz space $\mathcal{S}_n$ of rapidly decreasing functions and the space of tempered distributions $\mathcal{S}'_n$, including many important subspaces. For nonanalysts, we recall basic properties such as convolutions, distributional derivatives, functions of moderate growth, etc. In §3, for a parameter space $\Lambda \subset \mathbb{R}^k$, we introduce the Schwartz space $\mathcal{S}_{n,\Lambda}$ of uniformly rapidly decreasing functions and the corresponding space of "uniformly tempered distributions"$\mathcal{S}^*_{n,\Lambda}$, which are not distributions in the usual sense. All of our results for these spaces will use strong topologies for these spaces.

We show that tempered distributions naturally and continuously extend to uniformly tempered distributions (Lemma 3.3). We also show that uniformly tempered distributions possess properties analogous to those of tempered distributions. Most importantly, we prove Theorem 3.7, which among other things, proves that convolution with a kernel $G \in \mathcal{S}_{n,\Lambda}$ defines a continuous linear transformation

$$c_{\Lambda,\,G} : \mathcal{S}^*_{n,\Lambda} \to C^\infty(\mathbb{R}^n \times \Lambda)$$

for the strong topology on $\mathcal{S}^*_{n,\Lambda}$ and the regular C^∞–topology on $C^\infty(\mathbb{R}^n \times \Lambda)$. Then, its restriction to $\mathcal{S}'_n$, denoted c_G, likewise defines a continuous transformation.

In §4 we introduce the spaces of G–convolved functions in $C^\infty(\mathbb{R}^n \times \Lambda)$, which are the images $\mathcal{H}_{\Lambda,\,G}$ of $c_{\Lambda,\,G}$ (resp. $\mathcal{H}_G$ of c_G). There are associated convolution jet spaces $\mathcal{H}^\ell_{\Lambda,\,G}$ (resp. $\mathcal{H}^\ell_G$). We explicitly determine the structure of these spaces, in the case there is a scaling action of a (multiplicative) scaling group $\mathbb{R}^k_+$ on $\mathbb{R}^n$. This action extends to an associated Poincaré scaling group PS, which is a semidirect product of $\mathbb{R}^k_+$ and $\mathbb{R}^n$. Provided the kernel G satisfies a scaling condition (Definition 4.1), we deduce from the action of PS that $\mathcal{H}^\ell_{\Lambda,\,G}$ and $\mathcal{H}^\ell_G$ are semialgebraic submanifolds of regular jet space, and trivial fiber bundles kernel over $\mathbb{R}^n \times \Lambda$ (Proposition 4.9).

In §5, we introduce the class of geometric properties defined by transversality to closed Whitney stratified sets. We recall the important properties of Whitney stratified sets, including the Thom Isotopy Theorem, and its consequences for both the structure of Whitney stratified sets and their properties under pullbacks by transverse mappings. Using the continuity of $c_{\Lambda,\,G}$ and c_G, the structure of $\mathcal{H}^\ell_{\Lambda,\,G}$ and $\mathcal{H}^\ell_G$, and the relative transversality theorem in [D2], we prove a transversality theorem using $\mathcal{H}^\ell_{\Lambda,\,G}$ and $\mathcal{H}^\ell_G$ (Theorem 5.8). This theorem combined with the Thom Isotopy Theorem and its consequences allows us to deduce the genericity, local generic structure, and stability for scale–based geometric structures of G–convolved (uniformly) tempered distributions (Theorems 5.10 and 5.11).

2. Tempered Distributions

Distributions in general and tempered distributions in particular are standard objects in analysis, and their treatment can be found in such standard references as Schwartz [Sch], Gelfand–Shilov [GS], Rudin [Ru2], and Hormander [H]. However, because they play such an important role here we briefly mention the key properties that we will use and refer readers to the preceding references for more details.

Distributions are continuous linear functionals on spaces of functions. We use standard notation for derivatives; for a multi-index $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n)$, we let $D_x^{\boldsymbol{\alpha}} = D_{x_1}^{\alpha_1} \cdots D_{x_n}^{\alpha_n}$. Also, we let $|\boldsymbol{\alpha}| = \sum_{i=1}^n \alpha_i$.

Several of the most important standard spaces are : $\mathcal{D} = C_0^\infty(\mathbb{R}^n)$, the C^∞ functions with compact support; $\mathcal{E} = C^\infty(\mathbb{R}^n)$, the C^∞ functions; and $\mathcal{S}_n$, the Schwartz space of rapidly decreasing functions. For $\mathcal{E}$, we have the topology of uniform convergence of derivatives on compact subsets defined by the semi-norms

$$||f||_N = \sup_{||x|| \leq N} \sup_{|\boldsymbol{\alpha}| \leq N} |D_x^{\boldsymbol{\alpha}} f(x)|$$

This defines the regular C^∞–topology and $\mathcal{E}$ is a Frechet space. $\mathcal{D}$ has instead an inductive topology induced from the subspaces $\mathcal{D}_K$ of functions with support in compact K. It is complete but is not even a Baire space. Third, $\mathcal{S}_n$ consists of smooth functions rapidly decreasing at infinity. This means $f \in \mathcal{S}_n$ iff all semi-norms $p_N(f) < \infty$, for $N = 1, 2, \ldots$, where

$$p_N(f) = \sup_{x \in \mathbb{R}^n} \sup_{|\boldsymbol{\alpha}| \leq N} |(1 + ||x||^2)^N D_x^{\boldsymbol{\alpha}} f(x)|$$

The spaces have continuous inclusions $\mathcal{D} \to \mathcal{S} \to \mathcal{E}$, yielding continuous maps of duals $\mathcal{E}' \to \mathcal{S}_n' \to \mathcal{D}'$, where $\mathcal{D}'$ are the distributions, $\mathcal{E}'$ the distributions with compact support, and $\mathcal{S}_n'$ the tempered distributions. Although the inclusions of the spaces of distributions are set theoretic inclusions, they are not inclusions topologically. We are mainly interested in the tempered distributions.

Example 2.1 (Examples of Tempered Distributions).

(1) *Positive Regular Borel measures of moderate growth* μ are positive regular Borel measures on $\mathbb{R}^n$ for which there is an integer $\ell > 0$ so that $\int (1 + ||x||^2)^{-\ell} d\mu < \infty$. As a special case we have the subspace of probability measures $\mathcal{PM}$. They define tempered distributions by $u_\mu(\varphi) = \int \varphi d\mu$. This extends by linearity to signed measures $\mu = \mu_+ - \mu_-$ with μ_+ and μ_- positive regular Borel measures of moderate growth. For fixed ℓ we denote the space of positive (resp. signed) regular Borel measures by $\mathcal{BM}_\ell$ (resp $\mathcal{SBM}_\ell$) using the variation of $(1 + ||x||^2)^{-\ell} d\mu$ as a norm.

(2) For $1 \leq p < \infty$, consider measurable functions f for which there exists a positive integer $\ell > 0$ so that $\int ((1+||x||^2)^{-\ell} f(x))^p dx < \infty$. Then, f defines a tempered distribution by $u_f(\varphi) = \int \varphi(x) \cdot f(x) dx$. In particular, *all L^p functions, polynomials, and measurable functions bounded by polynomials define tempered distributions.* Again for fixed ℓ and p we let L_ℓ^p denote this space, using the L^p-norm of $(1 + ||x||^2)^{-\ell} f(x)$.

(3) *Positive and signed regular Borel measures of compact support* have moderate growth and so define tempered distributions. We add a "c"to denote these subspaces ($\mathcal{BM}_c$ and $\mathcal{SBM}_c$). The probability measures of compact support will be denoted by $\mathcal{PM}_c$.

(4) *Distibutions with compact support* are tempered distibutions (by the above inclusion map).

For these examples, the inclusions into the space of tempered distribtutions with its topology are continuous.

Example 2.2 (Properties of Tempered Distributions). (1) Differentiation by $D_x^{\boldsymbol{\alpha}}$, multiplication by a polynomial, or multiplication by a function $g \in \mathcal{S}_n$ are continuous linear transformations $\mathcal{S}_n \to \mathcal{S}_n$; and induce continuous linear transformations $\mathcal{S}'_n \to \mathcal{S}'_n$.

(2) The derivative $D_x^{\boldsymbol{\alpha}} u$ of a tempered distribution u is defined by composition $D_x^{\boldsymbol{\alpha}} u(g) = (-1)^{|\boldsymbol{\alpha}|} u(D_x^{\boldsymbol{\alpha}}(g))$ with $D_x^{\boldsymbol{\alpha}}$ viewed as a linear transformation of $\mathcal{S}_n$.

(3) The Fourier transform defines a continuous linear transformation from $\mathcal{S}_n$ to itself. Hence, induces a Fourier transform $\mathcal{F}$ on tempered distributions $\mathcal{S}'_n$ by $\mathcal{F}(u)(g) = u(\mathcal{F}(g))$.

Quite generally, smooth functions f are defined to be of *moderate growth* if for each multi-index $\boldsymbol{\alpha}$, there is an integer m and a constant $C > 0$ (which depend on $\boldsymbol{\alpha}$) such that

$$|D_x^{\boldsymbol{\alpha}}(f)(x)| \leq C(1 + \|x\|^2)^m \qquad \text{for all} \quad x \in \mathbb{R}^n \tag{2.1}$$

For continuous functions to be of moderate growth, we just require that (2.1) holds for $\boldsymbol{\alpha} = 0$, i.e. just for the function.

The first importance of functions of moderate growth appears in the characterization of tempered distributions in terms of ordinary operations of integration and differentiation, as given in [Sch, Chap. 7, §3, Thm VI] and [GS, vol. 2, Chap. II, 4.3].

Theorem 2.3. *If $u \in \mathcal{S}'_n$ then there is a continuous function of moderate growth f and a multiindex $\boldsymbol{\alpha}$, such that for $g \in \mathcal{S}_n$, $u(g) = \int f(x) D^{\boldsymbol{\alpha}}(g)(x)\, dx$.*

The set of smooth functions of moderate growth forms an algebra, denoted $\mathcal{O}_M$ (see [Sch, Chap. 7, §5]). Polynomials are the simplest examples of functions of moderate growth. Multiplication by a fixed function $f \in \mathcal{O}_M$ again defines a continuous linear transformation of $\mathcal{S}_n$, so that tempered distributions can be multiplied by functions in $\mathcal{O}_M$ by composition with the associated linear transformation $(f \cdot u(g) = u(f \cdot g))$.

More generally, we say a function $f = (f_1, \dots, f_p) : \mathbb{R}^n \to \mathbb{R}^p$ is of moderate growth if each coordinate function f_i is of moderate growth. We say that a diffeomorphism φ of $\mathbb{R}^n$ is of moderate growth if both functions φ and $\varphi^{-1}$are of moderate growth. We have the following basic consequence for composition of a function in $\mathcal{S}_n$ with a diffeomorphism of moderate growth.

Lemma 2.4. *If $g \in \mathcal{S}_n$ and φ is a diffeomorphism of $\mathbb{R}^n$ of moderate growth, then $g \circ \varphi \in \mathcal{S}_n$ and the map $g \mapsto g \circ \varphi$ is continuous.*

Proof. The argument for this is a standard type argument involving the estimates of derivatives, however, there does not seem to be a standard reference. It will follow from a more general form to be given in Lemma 3.5. □

In the case of ordinary distributions, the pullback of a distribution u by a diffeomorphism φ is defined as follows (see [H, Thm 6.1.2]).

$$\varphi^*(u)(g) = u(|\det(d\varphi^{-1}(y))| \cdot g(\varphi^{-1}(y))) \tag{2.2}$$

If φ is a diffeomorphism of $\mathbb{R}^n$ of moderate growth, then we may consider (2.2) for a tempered distribution u. We see that $\varphi^*(u)$ is given by composition of u with two continuous linear transformations of $\mathcal{S}_n$. The first is given by composition of g with φ^{-1}, and is continuous on $\mathcal{S}_n$ by Lemma 2.4. Also, $|\det(d\varphi^{-1}(y))| \in \mathcal{O}_M$ and multiplication by this function defines the second. Thus, the RHS of (2.2) defines an element of $\mathcal{S}'_n$.

The other operation which plays such an important role for distributions is convolution. Let τ_x denote translation by x on $\mathbb{R}^n$, and let $\check{g}(x) = g(-x)$. Then, convolution of a tempered distribution u and $g \in \mathcal{S}_n$ is defined by $(u * g)(x) = u(\tau_x \check{g}(y))$. Then, important properties of convolution are the following.

Theorem 2.5 (see [Ru2, Thm. 7.19]). *Suppose $g \in \mathcal{S}_n$ and $u \in \mathcal{S}'_n$, then*

(1) *$u * g$ is smooth of moderate growth (and hence $u * g \in \mathcal{S}'_n$);*
(2) $D^\alpha_x(u * g) = D^\alpha_x(u) * g = u * D^\alpha_x(g)$;

The Strong Topology on $\mathcal{S}'_n$. There are two standard topologies on $\mathcal{S}'_n$ the strong and weak* topologies. All of our results will concern the strong topology which we next recall.

Suppose $B \subset \mathcal{S}_n$ is a bounded subset (in the sense of topological vector spaces, so for any open neighborhood V of 0, there is an s_0 such that for $s > s_0$, $B \subset s \cdot V$). If $\varepsilon > 0$, then if B is a closed bounded set,

$$U(B, \varepsilon) = \{u \in \mathcal{S}'_n : |u(g)| < \varepsilon \text{ for all } g \in B\}$$

is a basic open neighborhood of 0 in $\mathcal{S}'_n$. For a sequence $\mathbf{b} = (b_0, b_1, \dots,)$ of positive numbers, $B(\mathbf{b}) = \{g \in \mathcal{S}_n : \rho_N(g) \leq b_N \text{ for all nonnegative integers } N\}$ is a closed bounded set; and for any bounded set B there is a sequence $\mathbf{b}$ such that $B \subseteq B(\mathbf{b})$, so that $U(B(\mathbf{b}), \varepsilon) \subseteq U(B, \varepsilon)$. These neighborhoods define by translation a topology on $\mathcal{S}'_n$.

For this strong topology, the inclusions of all of the important subspaces we have already mentioned are continuous.

3. Schwartz Space of Uniformly Rapidly Decreasing Functions and Uniformly Tempered Distributions

Because we want to consider functions depending on parameters, we extend the notion of rapidly decreasing function and tempered distribution to allow both to depend on parameters. We consider an open subset $\Lambda \subset \mathbb{R}^k$, and smooth functions on $\mathbb{R}^n \times \Lambda$. Eventually we consider functions (and distributions) mapping to $\mathbb{R}^p$ (see §§11 and 12). We let λ denote coordinates for $\mathbb{R}^k$ and distinguish derivatives by subscripts $D^{\boldsymbol{\alpha}}_x$ denoting derivatives with respect to the x–variables, and $D^{\boldsymbol{\alpha}}_\lambda$, derivatives with respect to the λ–variables.

We suppose that we have compact subsets $K_m \subset \Lambda, m = 0, 1, 2, \dots,$ such that $K_m \subset \mathrm{int}(K_{m+1})$, and $\cup K_m = \Lambda$. We define semi-norms for $f \in C^\infty(\mathbb{R}^N \times \Lambda)$.

$$p_{N,\Lambda}(f) = \sup_{x \in \mathbb{R}^n} \sup_{\lambda \in K_N} \sup_{|\boldsymbol{\alpha}|, |\boldsymbol{\beta}| \leq N} |(1 + ||x||^2)^N |D^{\boldsymbol{\alpha}}_x D^{\boldsymbol{\beta}}_\lambda f(x, \lambda)|$$

We say that f is *uniformly rapidly decreasing* if all semi-norms $p_{N,\Lambda}(f) < \infty$. We then define the *Schwartz space of uniformly rapidly decreasing functions* $\mathcal{S}_{n,\Lambda}$ to consist of the uniformly rapidly decreasing functions with a topology given by the collection of semi-norms $p_{N,\Lambda}$, $N \geq 1$. By the same arguments as for Schwartz spaces $\mathcal{S}_n$, the $\mathcal{S}_{n,\Lambda}$ are Frechet spaces.

Example 3.1. Both the standard Gaussian kernel

$$K_t(x) = \frac{1}{(4\pi t)^{\frac{n}{2}}} \exp(\frac{-\|x\|^2}{4t})$$

and, for any symmetric positive definite matrix A, the "anisotropic" Gaussian kernel

$$K_t(x) = \frac{1}{(4\pi t)^{\frac{n}{2}}} (\det(A))^{\frac{1}{2}} \exp(\frac{-<x, Ax>}{4t})$$

belong to $\mathcal{S}_{n,\mathbb{R}_+}$, where $\mathbb{R}_+$ denotes the positive numbers. Thus, any derivative $D_x^{\boldsymbol{\alpha}} D_t^k(K_t(x))$ also belongs to $\mathcal{S}_{n,\mathbb{R}_+}$. The standard and anisotropic Gaussian kernels (along with their derivatives) are the simplest examples of scale–based kernels to be defined in §4.

Let $\pi : \mathbb{R}^n \times \Lambda \to \mathbb{R}^n$ denote projection; and for some $\lambda_0 \in K_0$, we have inclusion $i : \mathbb{R}^n \to \mathbb{R}^n \times \Lambda$ sending $x \mapsto (x, \lambda_0)$. These maps induce an inclusion $\pi^* : \mathcal{S}_n \subset \mathcal{S}_{n,\Lambda}$ (for which $\mathcal{S}_n$ has the subspace topology) and a projection $i^* : \mathcal{S}_{n,\Lambda} \to \mathcal{S}_n$ so that $i^* \circ \pi^* = id$. Thus, $\mathcal{S}_n$ can be viewed as a closed subspace which is a direct summand of $\mathcal{S}_{n,\Lambda}$. However, we shall next see that this by itself does not imply that the space of tempered distributions forms a subspace of the space of uniformly tempered distributions.

Definition 3.2. A *uniformly tempered distribution* u is a continuous linear transformation $u : \mathcal{S}_{n,\Lambda} \to C^\infty(\Lambda)$, where $C^\infty(\Lambda)$ has the regular C^∞–topology.

With this definition, a uniformly tempered distribution is not a distribution in the usual sense; however, it will provide a framework for convolving with parameter dependent distributions, and allow other parameter dependent operations. We denote the space of uniformly tempered distributions by $\mathcal{S}^*_{n,\Lambda}$.

We first show that the tempered distributions can be viewed as uniformly tempered distributions, defining an inclusion map $\mathcal{S}'_n \hookrightarrow \mathcal{S}^*_{n,\Lambda}$ which is continuous. Thus, we must say what the topology is on $\mathcal{S}^*_{n,\Lambda}$.

The Strong Topology on $\mathcal{S}^*_{n,\Lambda}$. To define the strong topology on $\mathcal{S}^*_{n,\Lambda}$, we do so by viewing it as a space of continuous linear transformations between Frechet spaces. To define a basic neighborhood of 0, we let M be a nonnegative integer, $\varepsilon > 0$, and $B \in \mathcal{S}_{n,\Lambda}$, a closed bounded subset. For $\mathcal{S}_{n,\Lambda}$, we modify the definition of $B(\mathbf{b})$ given in §2 as follows. For the sequence of positive numbers $\mathbf{b} = (b_0, b_1, \dots,)$, $B_\Lambda(\mathbf{b})$ consists of those $g \in \mathcal{S}_{n,\Lambda}$, such that $p_{N,\Lambda}(g) \leq b_N$ for all N. Let $V(M, \varepsilon) = \{g \in C^\infty(\Lambda) : \|g\|_M < \varepsilon\}$. Then, a basic open neighborhood of $0 \in \mathcal{S}^*_{n,\Lambda}$ has the form

$$U(M, B, \varepsilon) = \{u \in \mathcal{S}^*_{n,\Lambda} : u(g) \in V(M, \varepsilon) \text{ for all } g \in B\}$$

for B a closed bounded set. If $B \subset B_\Lambda(\mathbf{b})$, then $U(M, B_\Lambda(\mathbf{b}), \varepsilon) \subseteq U(M, B, \varepsilon)$. Again, the topology is obtained by translating these local neighborhoods.

We are now in position to define and prove the continuity of the inclusion.

Lemma 3.3. *A tempered distribution $u \in \mathcal{S}'_n$ extends to a uniformly tempered distribution $\tilde{u}$ by defining $\tilde{u}(g)(\lambda) = u(g(x, \lambda))$, where for fixed $\lambda \in \Lambda$, $g(x, \lambda) \in \mathcal{S}_n$. Moreover, the induced map $\mathcal{S}'_n \hookrightarrow \mathcal{S}^*_{n,\Lambda}$, sending $u \mapsto \tilde{u}$, is continuous.*

Proof. First, by induction on $|\boldsymbol{\alpha}|$, we establish

$$D_\lambda^{\boldsymbol{\alpha}}(\tilde{u}(g))(\lambda) = \tilde{u}(D_\lambda^{\boldsymbol{\alpha}}(g)(x, \lambda)) \tag{3.1}$$

Since $D^{\boldsymbol{\alpha}}_{\lambda}(g) \in \mathcal{S}_{n,\Lambda}$, it suffices to establish (3.1) for a single partial derivative $D_{\lambda_i}(u(g))$. In this case, it is sufficient to show

$$u\left(\frac{g(x,\lambda_0+\delta\lambda_i)-g(x,\lambda_0)}{\delta\lambda_i} - D_{\lambda_i}(g)(x,\lambda_0))\right) \longrightarrow 0 \quad \text{as } \delta\lambda_i \to 0 \tag{3.2}$$

Here $\delta\lambda_i$ denotes an increment for λ_i. We denote the function inside u in (3.2) by $h(x,\lambda_0,\delta\lambda_i)$. We want to show that $h(x,\lambda_0,\delta\lambda_i)$ as a function of x (for fixed λ_0) approaches 0 in $\mathcal{S}_n$ as $\delta\lambda_i \to 0$. We actually show more that $h(x,\lambda_0,\delta\lambda_i)$ as a function of (x,λ_0) approaches 0 in $\mathcal{S}_{n,\Lambda}$ as $\delta\lambda_i \to 0$.

We replace g by $g_1 = D^{\boldsymbol{\alpha}}_x D^{\boldsymbol{\beta}}_{\lambda}(g)$, and denote the corresponding h by h_1. Then, by the remainder for the first order Taylor expansion,

$$D^{\boldsymbol{\alpha}}_x D^{\boldsymbol{\beta}}_{\lambda}(h_1) \;=\; \frac{1}{2}\frac{\partial^2 g_1}{\partial\lambda_i^2}(x,\lambda_1)\cdot(\delta\lambda_i) \tag{3.3}$$

where $\lambda_1 - \lambda_0 = \theta\cdot\delta\lambda_i$ for $0 \le \theta \le 1$ If we choose an N so $B_r(\lambda_0) \subset K_N$, and multiply (3.3) by $(1+\|x\|^2)^N$, then for $|\delta\lambda_i| < r$,

$$\begin{aligned}
&\sup_{x\in\mathbb{R}^n}\ \sup_{|\boldsymbol{\alpha}|,|\boldsymbol{\beta}|\le N}\left((1+\|x\|^2)^N|D^{\boldsymbol{\alpha}}_x D^{\boldsymbol{\beta}}_{\lambda}(h_1)|\right)\\
&\qquad\le\ \sup_{x\in\mathbb{R}^n}\ \sup_{|\boldsymbol{\alpha}|,|\boldsymbol{\beta}|\le N}\left((1+\|x\|^2)^N|D^{\boldsymbol{\alpha}}_x D^{\boldsymbol{\beta}}_{\lambda}(\frac{\partial^2 g}{\partial\lambda_i^2})|\right)\frac{|\delta\lambda_i|}{2}\\
&\qquad\le\ p_{N+2,\,\Lambda}(g)\cdot\frac{|\delta\lambda_i|}{2}
\end{aligned}$$

or

$$p_{N,\Lambda}(h_1) \;\le\; p_{N+2,\,\Lambda}(g)\cdot\frac{|\delta\lambda_i|}{2} \tag{3.4}$$

Thus, for g in a bounded neighborhood, $p_{N,\Lambda}(h_1) \to 0$ as $\delta\lambda_i \to 0$ for all N so $h_1 \to 0$ in $\mathcal{S}_{n,\Lambda}$. However, if we view g and h_1 as functions of x, with λ_0 fixed, the preceding argument also gives

$$p_N(h_1) \;\le\; p_{N+2}(g)\cdot\frac{|\delta\lambda_i|}{2} \tag{3.5}$$

implying $h_1 \to 0$ in $\mathcal{S}_n$, establishing (3.1).

Next, we show that $\tilde{u} : \mathcal{S}_{n,\Lambda} \to C^{\infty}(\Lambda)$ is continuous. Given any positive integer M, from (3.1),

$$\sup_{\lambda\in K_M}\ \sup_{|\boldsymbol{\beta}|\le M} |D^{\boldsymbol{\beta}}_{\lambda}(\tilde{u}(g))| = \sup_{\lambda\in K_M}\ \sup_{|\boldsymbol{\beta}|\le M} |\tilde{u}(D^{\boldsymbol{\beta}}_{\lambda}(g))| \tag{3.6}$$

As $u \in \mathcal{S}'_n$, that there is an integer N and a constant $C > 0$ such that $|u(g)| \le C\cdot p_N(g)$ for all $g \in \mathcal{S}_n$. Thus,

$$\|\tilde{u}(g)\|_M \le \sup_{\lambda\in K_M}\ \sup_{|\boldsymbol{\beta}|\le M} C\cdot p_N(D^{\boldsymbol{\beta}}_{\lambda}(g))$$

(where we view g as a function in $\mathcal{S}_n$ for various values of λ)

$$\le \sup_{x\in\mathbb{R}^n}\ \sup_{\lambda\in K_M}\ \sup_{|\boldsymbol{\alpha}|\le N}\ \sup_{|\boldsymbol{\beta}|\le M} C(1+\|x\|^2)^N)|D^{\boldsymbol{\alpha}}_x D^{\boldsymbol{\beta}}_{\lambda}(g)| \tag{3.7}$$

Thus, for all M,

$$\|\tilde{u}(g)\|_M \ \le\ C \cdot p_{N+M,\Lambda}(g) \tag{3.8}$$

From (3.8) follows the continuity of $\tilde{u}$.

It remains to show that the inclusion $i : \mathcal{S}'_n \hookrightarrow \mathcal{S}^*_{n,\Lambda}$ is continuous. By linearity, it is sufficient to establish continuity at 0, so let $U(M, B(\mathbf{b}), \varepsilon)$ be a basic neighborhood. We seek a neighborhood $U(B(\mathbf{b}'), \varepsilon)$ of 0 in $\mathcal{S}'_n$ which maps by i into $U(M, B(\mathbf{b}), \varepsilon)$. We let $b'_N = b_{N+M}$. If $u \in U(B(\mathbf{b}'), \varepsilon)$, then we wish to show $\|\tilde{u}(g)\|_M < \varepsilon$ for all $g \in \mathcal{S}_{n,\Lambda}$ for which $p_{N,\Lambda}(g) \le b_N$ for all N. First, $\|\tilde{u}(g)\|_M$ is given by the LHS of (3.6). It is sufficient to show that for each fixed $\lambda \in K_M$ and $|\beta| \le M$, $|\tilde{u}(D^{\boldsymbol{\beta}}_\lambda(g))| (= |u(D^{\boldsymbol{\beta}}_\lambda(g))|) < \varepsilon$. Thus, as $u \in U(B(\mathbf{b}'), \varepsilon)$, by (3.6), it is sufficient to show $p_N(D^{\boldsymbol{\beta}}_\lambda(g)) \le b'_N$ for all N and each fixed $\lambda \in K_M$ and $|\beta| \le M$. By (3.7), this is implied by $p_{N+M,\Lambda}(g) \le b'_N$. Since $b'_N = b_{N+M}$ this follows from $p_{N,\Lambda}(g) \le b_N$ for all N, proving continuity of the inclusion. $\square$

The advantage of $\mathcal{S}^*_{n,\Lambda}$ is that we may perform certain basic parameter dependent operations as for tempered distributions. First, the analogue of Example 2.2. Since we will not be concerned with Fourier transforms we concentrate on the other properties

Lemma 3.4 (Properties of Uniformly Tempered Distributions)**.** (1) *Differentiation by $D^\alpha_x D^\beta_\lambda$, multiplication by a polynomial in x with coefficients smooth in λ, or multiplication by a function $g \in \mathcal{S}_{n,\Lambda}$ are continuous linear transformations $\mathcal{S}_{n,\Lambda} \to \mathcal{S}_{n,\Lambda}$; and induce continuous linear transformations $\mathcal{S}^*_{n,\Lambda} \to \mathcal{S}^*_{n,\Lambda}$.*

(2) *The derivative $D^{\boldsymbol{\alpha}}_x u$ of a uniformly tempered distribution u is defined by composition $D^{\boldsymbol{\alpha}}_x u(g) = (-1)^{|\boldsymbol{\alpha}|} u(D^{\boldsymbol{\alpha}}_x g)$, with $D^{\boldsymbol{\alpha}}_x$ viewed as a continuous linear transformation of $\mathcal{S}_{n,\Lambda}$.*

(3) *The derivative $D_{\lambda_i}(u)$ is defined by $D_{\lambda_i}(u)(g) = D_{\lambda_i}(u(g)) - u(D_{\lambda_i}(g))$*

The proofs follow as for tempered distributions, but using uniform bounds on compact subsets of Λ; except to define $D_{\lambda_i}(u)$, we make use of the relation which holds in the case u is a differentiable function

$$D_{\lambda_i} \int u(x,\lambda) g(x,\lambda) dx = \int D_{\lambda_i}(u)(x,\lambda) g(x,\lambda) dx + \int u(x,\lambda) D_{\lambda_i}(g)(x,\lambda) dx$$

In 1) of Lemma 3.4, we may replace polynomials by smooth functions of uniformly moderate growth. By a smooth function $f(x,\lambda)$ being of *uniformly moderate growth*, we mean that it satisfies (2.1) uniformly on compact subsets K_m. Specifically, for each integer N and pair of multi-indices $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, with $|\boldsymbol{\alpha}|, |\boldsymbol{\beta}| \le N$, there is an integer m and a constant $C > 0$ (which depend on N, $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$) such that

$$\sup_{\lambda \in K_N} |D^{\boldsymbol{\alpha}}_x D^{\boldsymbol{\beta}}_\lambda(f)(x,\lambda)| \le C(1 + \|x\|^2)^m \qquad \text{for all} \quad x \in \mathbb{R}^n \tag{3.9}$$

Then, we say a smooth mapping $\varphi : \mathbb{R}^n \times \Lambda \to \mathbb{R}^n \times \Lambda$ of the form $\varphi(x,\lambda) = (\varphi_1(x,\lambda), \dots, \varphi_n(x,\lambda), \varphi'_1(\lambda), \dots, \varphi'_k(\lambda))$ is of uniformly moderate growth if each φ_i is of uniformly moderate growth and φ'_i are of moderate growth. Also, φ will be a diffeomorphism of uniformly moderate growth if both functions φ and φ^{-1} are.

Lemma 3.5. *If $g \in \mathcal{S}_{n,\Lambda}$ and φ is a diffeomorphism of $\mathbb{R}^n \times \Lambda$ of uniformly moderate growth, then $g \circ \varphi \in \mathcal{S}_{n,\Lambda}$ and the map $g \mapsto g \circ \varphi$ is continuous.*

Proof. We denote coordinates $(x', \lambda') = \varphi(x, \lambda)$. First, applying the definition of uniformly moderate growth to φ_i viewed as zero-th derivatives gives

$$\sup_{\lambda \in K_N} \|\varphi(x, \lambda)\| \leq C(1 + \|x\|^2)^m \tag{3.10}$$

or for $\lambda \in K_N$,

$$1 + \|x'\|^2 \leq C'((1 + \|x\|^2)^{m'} \tag{3.11}$$

A similar equation holds in the reverse direction interchanging x and x' for different constants C'' and m''.

Also, applying the definition of moderate growth to $\varphi'_i(\lambda)$ gives

$$\|D^{\boldsymbol{\alpha}}_{\lambda} \varphi'(\lambda)\| \leq C_N \cdot (1 + \|\lambda\|^2)^{m_N} \qquad \text{for all } |\boldsymbol{\alpha}| \leq N \tag{3.12}$$

for appropriate constants $C_N > 0$, m_N, and all $\lambda \in \Lambda$. Then, we can relate λ and λ' by

$$1 + \|\lambda'\|^2 \leq C'((1 + \|\lambda\|^2)^{m'} \tag{3.13}$$

with a similar equation holding in the reverse direction.

Using the formula for the higher derivative of a composition (see e.g. [Rg]), We may write

$$D^{\boldsymbol{\alpha}}(g \circ \varphi)(x) = \sum_{|\boldsymbol{\beta}| = |\boldsymbol{\alpha}|+1} c_{\boldsymbol{\beta}} D^{\beta_0}(g)(\varphi(x)) D^{\beta_1}(\varphi_{i_1})(x) \cdots D^{\beta_r}(\varphi_{i_r}))(x) \tag{3.14}$$

where $\boldsymbol{\beta} = (\beta_0, \ldots, \beta_r)$ with all $|\beta_i| > 0$, and the coefficients $c_{\boldsymbol{\beta}}$ are independent of g and φ. Here, to keep notation within bounds, we denote φ'_i by φ_{n+i} and do not distinguish between the derivatives with respect to the x_i and λ_i

Hence, if $|\boldsymbol{\alpha}| \leq N$, then by the uniform moderate growth of φ, there is a constant C' and a positive integer M so that

$$\sup_{\lambda \in K_N} \sup_{1 \leq i \leq n} \sup_{|\boldsymbol{\alpha}|, |\boldsymbol{\beta}| \leq N} |D^{\boldsymbol{\alpha}}_x D^{\boldsymbol{\beta}}_{\lambda}(\varphi_i)(x)| \quad \leq \quad C'(1 + \|x\|^2)^M \tag{3.15}$$

By (3.12), this equation is also valid for $i > n$. Then, by (3.14) and (3.15) we obtain

$$\begin{aligned}
&(1 + \|x\|^2)^N |D^{\alpha}(g \circ \varphi)(x)| \\
&\quad \leq \sup_{x \in \mathbb{R}^n} \sup_{\lambda \in K_N} \left(\sup_{|\gamma| \leq |\alpha|} (1 + \|x\|^2)^N |D^{\alpha}(g)(\varphi(x, \lambda))| \cdot \right. \\
&\qquad\qquad \left. \sup_{1 \leq i \leq n} \sup_{|\delta| \leq N} |D^{\delta}(\varphi_i)(x, \lambda)|^N \right) \\
&\quad \leq C'' \sup_{x \in \mathbb{R}^n} \sup_{\lambda \in K_N} \sup_{|\gamma| \leq |\alpha|} \left((1 + \|x\|^2)^{N+NM} |D^{\gamma}(g)(x')|\right) \\
&\quad \leq C_3 \sup_{x' \in \mathbb{R}^n} \sup_{\lambda \in K_N} \sup_{|\gamma| \leq |\alpha|} \left((1 + \|x'\|^2)^{m'(N+NM)} |D^{\gamma}(g)(x')|\right) \\
&\quad \leq p_{N'', \Lambda}(g)
\end{aligned} \tag{3.16}$$

where $N'' = m'(N + NM)$.

By (3.16),

$$p_{N, \Lambda}(g \circ \varphi) \leq C_3 \cdot p_{N'', \Lambda}(g) \tag{3.17}$$

This establishes both that $g \circ \varphi \in \mathcal{S}_{n,\Lambda}$ and the map $g \mapsto g \circ \varphi$ is continuous. $\square$

We can now use the analogue of (2.2) to define the pullback of a uniformly tempered distribution u. Let Φ be a diffeomorphism of uniformly moderate growth of the form $(x', \lambda') = \Phi(x, \lambda) = (\varphi(x, \lambda), \varphi'(\lambda))$. For fixed λ, $\varphi(x, \lambda)$ is a diffeomorphism of $\mathbb{R}^n$, whose inverse we denote by φ^{-1}. Then, exactly as for (2.2), we define the pullback so it agrees with the pullback of functions via the change of variables formula for multiple integrals.

Definition 3.6. If $u \in \mathcal{S}^*_{n,\Lambda}$, and $h \in \mathcal{S}_{n,\Lambda}$, then we define the pullback of a uniformly tempered distribution u by the diffeomorphism of uniformly moderate growth Φ by the formula

$$\Phi^*(u)(h) = u(|\det(d\varphi^{-1}(y))| \cdot h \circ \Phi^{-1}(y, \lambda))) \circ \varphi' \tag{3.18}$$

It follows from Lemma 3.5 that the RHS of (3.18) defines a uniformly tempered distribution. In the case $\varphi' = id$, (3.18) has the same form as (2.2).

Properties of Convolution of $\mathcal{S}^*_{n,\Lambda}$ with $\mathcal{S}_{n,\Lambda}$. We consider the convolution of a uniformly tempered distribution $u \in \mathcal{S}^*_{n,\Lambda}$ with a function $g \in \mathcal{S}_{n,\Lambda}$. Recall τ_x denotes translation by x on $\mathbb{R}^n$, and as for the case of convolutions of distributions, we let $\check{g}(x, \lambda) = g(-x, \lambda)$. Then, by Lemma 3.5, convolution is again defined by $(u * g)(x, \lambda) = u(\tau_x \check{g}(y, \lambda))$. The important properties of convolution are the following.

Theorem 3.7. *Suppose $g \in \mathcal{S}_{n,\Lambda}$ and $u \in \mathcal{S}^*_{n,\Lambda}$, then*

(1) *$u * g$ is a smooth function on $\mathbb{R}^n \times \Lambda$ of uniformly moderate growth (and hence $u * g \in \mathcal{S}^*_{n,\Lambda}$);*
(2) *$D^\alpha_x(u * g) = D^\alpha_x(u) * g = u * D^\alpha_x(g)$;*
(3) *However,*
$$D_{\lambda_i}(u * g) = D_{\lambda_i}(u) * g + u * D_{\lambda_i}(g);$$
so only if $u \in \mathcal{S}'_n$ do we have $D_{\lambda_i}(u * g) = u * D_{\lambda_i}(g)$.
(4) *For $G \in \mathcal{S}_{n,\Lambda}$, the associated convolution map*
$$c_{\Lambda,\, G} : \mathcal{S}^*_{n,\Lambda} \to C^\infty(\mathbb{R}^n \times \Lambda) \tag{3.19}$$
which sends $u \mapsto u * G$, is continuous for the regular C^∞ topology on $C^\infty(\mathbb{R}^n \times \Lambda)$.

Proof. By definition $(u*g)(x, \lambda) = u(\tau_x \check{g}(y, \lambda))$. Both τ_x and $y \mapsto -y$ are (trivially) diffeomorphisms of uniformly moderate growth. Thus, viewed as a function of y with parameters (x, λ), $\tau_x \check{g}(y, \lambda)) = g(y - x, \lambda)$ is by Lemma 3.5, a function in $\mathcal{S}_{n,\, \mathbb{R}^n \times \Lambda}$. First, we can apply an extension of Lemma 3.3.

Lemma 3.8. *Let $\Gamma \subset \mathbb{R}^\ell$ be open. A uniform tempered distribution $u \in \mathcal{S}^*_{n,\Lambda}$ extends to a uniformly tempered distribution $\tilde{u} \in \mathcal{S}^*_{n,\, \Lambda \times \Gamma}$ by defining $\tilde{u}(g)(\lambda, \gamma) = u(g(x, \lambda, \gamma))$, where for fixed $\gamma \in \Gamma$, $g(x, \lambda, \gamma) \in \mathcal{S}_{n,\Lambda}$. Moreover,*

$$D^{\boldsymbol{\alpha}}_\gamma(\tilde{u}(g))(\lambda, \gamma) = \tilde{u}(D^{\boldsymbol{\alpha}}_\gamma(g)(x, \lambda, \gamma)) \tag{3.20}$$

The proof of this Lemma is virtually identical to that of Lemma 3.3.

As a consequence of Lemma 3.8, we conclude that $u(\tau_x \check{g}(y, \lambda))$ viewed as function of (x, λ) is in $C^\infty(\mathbb{R}^n \times \Lambda)$. Also, both 2) and 3) follow by (3.20) and the definition in 2) and 3) (of Lemma 3.4) of the derivatives for a uniformly tempered distribution.

It remains to show that $u * g$ has uniformly moderate growth and 4) that convolution defines a continuous map. For these we adapt an argument in [Ru2, Thm 7.19]. The inequality

$$1 + |x+y|^2 \leq 2(1+|x|^2)(1+|y|^2) \tag{3.21}$$

implies

$$p_{N,\Lambda}(\tau_x(h)) \leq 2^N (1+|x|^2)^N p_{N,\Lambda}(h) \quad \text{for all } x \in \mathbb{R}^n \text{ and } h \in \mathcal{S}_{n,\Lambda} \tag{3.22}$$

Now, if $u \in \mathcal{S}^*_{n,\Lambda}$, then given a nonnegative integer M, there is an integer N and a constant $C_{M,N}$ such that $\|u(h)\|_M \leq C_{M,N} \cdot p_{N,\Lambda}(h)$ for all $h \in \mathcal{S}_{n,\Lambda}$. Then, let $|\boldsymbol{\alpha}|, |\boldsymbol{\beta}| \leq M$.

$$D_x^{\boldsymbol{\alpha}} D_\lambda^{\boldsymbol{\beta}}(u*g)(x,\lambda) = D_\lambda^{\boldsymbol{\beta}}(u * D_x^{\boldsymbol{\alpha}}(g))(x,\lambda) = (-1)^{|\boldsymbol{\alpha}|} D_\lambda^{\boldsymbol{\beta}}(u(\tau_x \check{D}_x^{\boldsymbol{\alpha}}(g))(y,\lambda))) \tag{3.23}$$

First, from (3.23) and (3.22),

$$\begin{aligned}
\sup_{\lambda \in K_M} |D_x^{\boldsymbol{\alpha}} D_\lambda^{\boldsymbol{\beta}}(u*g)(x,\lambda)| &= \sup_{\lambda \in K_M} |D_\lambda^{\boldsymbol{\beta}}(u(\tau_x \check{D}_y^{\boldsymbol{\alpha}} g(y,\lambda)))| \\
&\leq \|u(\tau_x(\check{D}_y^{\boldsymbol{\alpha}} g(y,\lambda)))\|_M \\
&\leq 2^N C_{M,N}(1+|x|^2)^N p_{N,\Lambda}(D_y^{\boldsymbol{\alpha}}(g)) \\
&\leq 2^N C_{M,N}(1+|x|^2)^N p_{N+M,\Lambda}(g) \quad \text{for all } x \in \mathbb{R}^n
\end{aligned} \tag{3.24}$$

Thus, by (3.24) $u * g$ has uniformly moderate growth.

For continuity, we consider g in place of G to be consistent with our earlier notation, so we will establish the continuity of $c_{\Lambda,\,g}$. Given a neighborhood $V = V(M,\varepsilon)$, we seek $U = U(M, B(\mathbf{b}), \varepsilon)$ so that $u \in U$ will imply $u * g \in V$. From the first two lines of (3.24), we alternately obtain

$$\sup_{|x| \leq M} \sup_{\lambda \in K_M} |D_x^{\boldsymbol{\alpha}} D_\lambda^{\boldsymbol{\beta}}(u*g)(x,\lambda)| = \sup_{|x| \leq M} \sup_{\lambda \in K_M} |D_\lambda^{\boldsymbol{\beta}} u(\tau_x(\check{D}_y^{\boldsymbol{\alpha}} g(y,\lambda)))| \tag{3.25}$$

As $u \in U$, from (3.25), we can conclude $\|u * g\|_M < \varepsilon$ provided for any fixed $\|x\| \leq M$, $p_{N,\Lambda}(\tau_x(\check{D}_y^{\boldsymbol{\alpha}} g(y,\lambda))) \leq b_N$ for all N. However, by (3.22)

$$\begin{aligned}
\sup_{|x| \leq M} p_{N,\Lambda}(\tau_x(\check{D}_y^{\boldsymbol{\alpha}} g(y,\lambda)) &\leq \sup_{|x| \leq M} 2^N (1+|x|^2)^N p_{N,\Lambda}(\check{D}_y^{\boldsymbol{\alpha}} g(y,\lambda)) \\
&\leq 2^N \cdot r_{M,N} \cdot p_{N,\Lambda}(\check{D}_y^{\boldsymbol{\alpha}} g(y,\lambda))
\end{aligned} \tag{3.26}$$

where $r_{M,N} = (1+M^2)^N$. Also,

$$p_{N,\Lambda}(\check{D}_y^{\boldsymbol{\alpha}} g(y,\lambda)) \quad \leq \quad p_{N+M,\Lambda}(g) \tag{3.27}$$

Thus, if we define $\mathbf{b}$ by $b_N = 2^N \cdot r_{M,N} \cdot p_{N+M,\Lambda}(g)$ for all N, then from (3.26) and (3.27), $\|u * g\|_M < \varepsilon$, proving that $c_{\Lambda,\,g}$ is continuous. □

As a corollary of Theorem 3.7 and the continuity of the inclusion $\mathcal{S}'_n \to \mathcal{S}^*_{n,\Lambda}$ (Lemma 3.3), we obtain the continuity of $c_{\Lambda,\,G}$ restricted to $\mathcal{S}'_n$.

Corollary 3.9. *Suppose $G \in \mathcal{S}_{n,\Lambda}$ and $u \in \mathcal{S}'_n$, then convolution with G, $u \mapsto u*G$ defines a continuous map (for the regular C^∞–topology on $C^\infty(\mathbb{R}^n \times \Lambda)$)*

$$c_G : \mathcal{S}'_n \to C^\infty(\mathbb{R}^n \times \Lambda) \tag{3.28}$$

4. Scale–based Kernels and the Structure of Convolution Jet Space

We introduce scaling properties and generalized scale space. Using the maps $c_{\Lambda,\,G}$ and c_G, we introduce the spaces of G–convolved functions and the corresponding convolution jet spaces. We show that scaling properties for kernels G allow us to determine the structure of the convolution jet spaces.

Scale Space and the Poincaré Scaling Group. We begin by introducing a generalized scale space which allows for multiple independent scales. For this we now concentrate on the special case of $\Lambda = \mathbb{R}^k_+$. We define $\mathbb{R}^k_+ = \mathbb{R}_+ \times \cdots \times \mathbb{R}_+$ with k factors. Then, under coordinate-wise multiplication, $\mathbb{R}^k_+$ is a group. We suppose we are given an action of $\mathbb{R}^k_+$ on $\mathbb{R}^n$ given by an $n \times k$ matrix $A = (a_{ij})$ of real numbers. For $\sigma = (\sigma_1, \dots, \sigma_k) \in \mathbb{R}^k_+$ and $x = (x_1, \dots, x_n)$, then

$$\sigma \cdot x = (\prod_j \sigma_j^{a_{1j}} x_1, \dots, \prod_j \sigma_j^{a_{nj}} x_n) \tag{4.1}$$

We shall on occasion denote this action by the representaton $\rho : \mathbb{R}^k_+ \to GL(\mathbb{R}^n)$. The entry a_{ij} denotes the "weight"of the action of the j–th scale parameter σ_j on x_i. Generally, we will be most interested in the case where the weights are positive rational numbers. We refer to A as the *weight matrix* for $\mathbb{R}^n$.

Combined with the action of $\mathbb{R}^k_+$ on itself by left multiplication, we obtain an action of $\mathbb{R}^k_+$ on $\mathbb{R}^n \times \mathbb{R}^k_+$ by $\sigma \cdot (x, \lambda) = (\sigma \cdot x, \sigma \cdot \lambda)$. We refer to $\mathbb{R}^n \times \mathbb{R}^k_+$ as *(generalized) scale space.*

We denote the diffeomorphism induced by the action of $\sigma \in \mathbb{R}^k_+$ by Ψ_σ. Also, we view $\mathbb{R}^n$ as an additive group and let the translation action of $x \in \mathbb{R}^n$ be denoted by τ_x. Then we can form the scaling analogue of the Poincaré group by the semi–direct product $PS = \mathbb{R}^k_+ \rtimes \mathbb{R}^n$ with multiplication

$$(\sigma_1, x_1) \cdot (\sigma_2, x_2) \quad = \quad (\sigma_1 \cdot \sigma_2, \sigma_1 \cdot x_2 + x_1).$$

Then, the Poincaré scaling group PS acts on scale space by $(\sigma, x) \mapsto \tau_x \circ \Psi_\sigma$. Next, we introduce scaling properties of kernels with respect to the action of PS. We first consider the special case that G is a "scale–based kernel".

Scale–based Kernels.

Definition 4.1. We say that $G \in \mathcal{S}_{n,\mathbb{R}^k_+}$ is a *scale–based kernel* if there is a k–tuple $B = (B_1, \ldots, B_k)$ so that

$$G(\sigma \cdot (x, \lambda)) = \sigma^B G(x, \lambda) \qquad \text{where} \quad \sigma^B = \prod_i \sigma_i^{B_i}. \tag{4.2}$$

We call B the *scale weight* of G.

Example 4.2. For the *standard or anisotropic Gaussian* kernel given in Example 3.1, we substitute $t = \frac{1}{2}\sigma^2$, so $K(x, \sigma) = K_{\frac{1}{2}\sigma^2}(x)$. This gives a scale-based kernel with single scale parameter σ (the standard deviation), with $B = -n$ and all weights $\mathrm{wt}\,(x_i) = 1$.

We extend this a *multi-Gaussian kernel* as follows. Let $n = \sum_{i=1}^r n_i$ for positive integers n_i, and consider the decomposition $\mathbb{R}^n = \oplus_{i=1}^r \mathbb{R}^{n_i}$ where $x = (x^{(1)}, \ldots, x^{(r)})$ with $x^{(i)} = (x_1^{(i)}, \ldots, x_{n_i}^{(i)})$. Then, the multi–Gaussian is given by the product of Gaussian kernels $\prod K(x^{(i)}, \sigma_i)$ with scaling parameters $(\sigma_1, \ldots, \sigma_r)$. This would allow independent scales σ_i on the distinct subspaces $\mathbb{R}^{n_i}$.

The standard *Gabor filters* $G_{\mathbf{t}}(x) = \cos(a\mathbf{u} \cdot x + b) \exp(-\frac{\|x\|^2}{4t})$, where $\mathbf{u}$ is a unit vector, are not scale–based kernels; however, there is a scale–based version $G_t(x) = \cos(a\frac{\mathbf{u} \cdot x}{\sqrt{t}} + b) \exp(-\frac{\|x\|^2}{4t})$ where $t = \frac{1}{2}\sigma^2$ with weight $\mathrm{wt}\,(\sigma) = 1$ and weights of all x_i equal 1.

There are two basic properties of the class of scale–based kernels: preservation under taking derivatives and under pullbacks by group homomorphisms.

Lemma 4.3. *Suppose G is a scale-based kernel with weight matrix A, and scale weight B, then:*

(1) $\dfrac{\partial G}{\partial x_i}$ *is again scale-based for the same weight matrix A, and scale weight $B' = B - A_i$, where $A_i = (a_{i1}, \ldots, a_{ik})$ (the weight vector for x_i).*

(2) $\dfrac{\partial G}{\partial \lambda_i}$ *is again scale-based for the same weight matrix A, and scale weight $B' = B - \varepsilon_i$, where $\varepsilon_i = (0, \ldots, 1, 0, \ldots, 0)$ with 1 in the i–th position.*

Proof. By differentiating (4.2) with respect to x_i, we obtain

$$\frac{\partial G(\sigma \cdot (x, \lambda))}{\partial x_i} = \sigma^B \frac{\partial G(x, \lambda)}{\partial x_i}$$

$$\sigma^{A_i} \frac{\partial G}{\partial x_i}(\sigma \cdot (x, \lambda)) = \sigma^B \frac{\partial G(x, \lambda)}{\partial x_i} \tag{4.3}$$

Dividing (4.3) by t^{A_i} gives i). An analogous proof yields ii). □

Example 4.4. By Example 4.2, the standard and anisotropic Gaussian kernels $K(x, \sigma) = K_{\frac{1}{2}\sigma^2}(x)$, the multi–Gaussian kernels, and the scale–based Gabor filters $G_{\frac{1}{2}\sigma^2}(x)$ are scale-based kernels; hence by Lemma 4.3, all of the derivatives are also scale-based. This includes for example, the median filter

$$M(x, \sigma) \quad = \quad -\sigma \cdot \frac{\partial K(x, \sigma)}{\partial \sigma} \qquad (= -\sigma^2 \cdot \Delta(K(x, \sigma)))$$

and is often referred to as the Mexican hat function (see e.g. [Db]). As well it includes the "edge detection kernels", given by the directional derivatives $\dfrac{\partial K(x,\sigma)}{\partial \mathbf{u}}$, where $\mathbf{u}$ is a unit vector, but $B = -(n+1)$. The one–dimensional profiles of these kernels are shown in Fig. 3

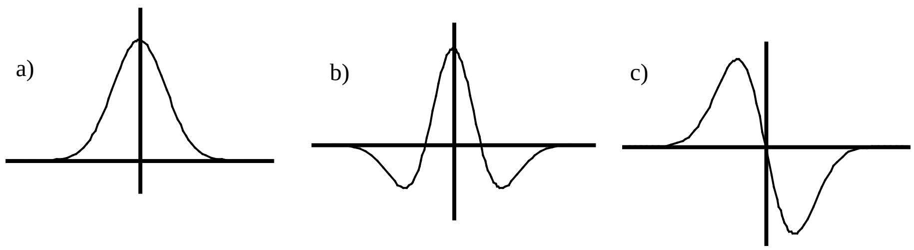

FIGURE 3. Profiles of a) Gaussian Kernel, b) Medial Kernel, and c) Edge Kernel for a fixed σ.

Second, we show the naturality of scale–based kernels under composition with group homomorphisms $\rho' : \mathbb{R}_+^{k'} \to \mathbb{R}_+^k$ of the form

$$\rho'(\sigma'_1, \dots, \sigma'_{k'}) = (\prod_j {\sigma'_j}^{c_{1j}}, \dots, \prod_j {\sigma'_{k'}}^{c_{1j}}) \tag{4.4}$$

We let $C = (c_{ij})$ be the matrix associated to ρ'. Typically in our situation the c_{ij} will be nonnegative rational numbers. Then, given a scale–based kernel G for the scaling group R_+^k via the representation ρ on $\mathbb{R}^n$, we first define an action of $\mathbb{R}_+^{k'}$ on $\mathbb{R}^n$ by $\rho \circ \rho'$. Then, we define the "pullback"of G by ρ' by ${\rho'}^*G(x,\sigma') = G(x, \rho'(\sigma'))$.

Lemma 4.5 (Naturality under pullback). *Suppose G is a scale-based kernel for the scaling group $\mathbb{R}_+^k$ via the representation ρ on $\mathbb{R}^n$, with weight matrix A, and scale weight B. Let $\rho' : \mathbb{R}_+^{k'} \to \mathbb{R}_+^k$ be a homomorphism of the form (4.4) with matrix C. Then, ${\rho'}^*G$ is again scale-based for the scaling group $\mathbb{R}_+^{k'}$ via the representation $\rho \circ \rho'$ on $\mathbb{R}^n$, with weight matrix $A \cdot C$, and scale weight $B \cdot C$.*

Proof. First, ${\rho'}^*G = G \circ (id_{\mathbb{R}^n} \times \rho')$ and ρ' viewed as a function is smooth and has moderate growth. Hence, as $G \in \mathcal{S}^*_{n,\Lambda}$, so also ${\rho'}^*G \in \mathcal{S}^*_{n,\Lambda}$.

Next, we compute

$$\begin{aligned} {\rho'}^*G(\sigma' \cdot x, \sigma' \cdot \lambda') &= G(\rho'(\sigma') \cdot x, \rho'(\sigma' \cdot \lambda')) \\ &= G(\rho'(\sigma') \cdot x, \rho'(\sigma') \cdot \rho'(\lambda')) = (\rho'(\sigma'))^B \cdot G(x, \rho'(\lambda')) \\ &= (\rho'(\sigma'))^B \cdot {\rho'}^*G(x, \lambda') \end{aligned}$$

Then, it is straightforward to see that $(\rho'(\sigma'))^B = (\sigma')^{BC}$, so the scale weight is BC, and that the matrix of $\rho \circ \rho'$ is AC. □

If we pullback a multi–Gaussian kernel, then we obtain a kernel which typically will have the scales mixed so scales affect different corrdinates in different ways. More important is the interaction of scales for multiple kernels which we consider in Part 3.

Example 4.6. In Example 4.4, if we pullback the edge detection kernel $\dfrac{\partial K(x,\sigma)}{\partial \mathbf{u}}$ by the homomorphism $\rho'(\sigma_1, \sigma_2) = \sigma_1 \cdot \sigma_2$, it becomes the kernel $\dfrac{\partial K(x, \sigma_1 \cdot \sigma_2)}{\partial \mathbf{u}}$. Along with the medial kernel $M(x, \sigma_1)$, we have two kernels which are scale–based for different scaling group actions. If we fix σ_2, then we obtain kernels scale–based for a common scale parameter σ_1. This is one of the methods used by Pizer, Eberly, et al. [PEM]. They fixed the ratio σ_2 between the edge detection and the medial scales. However, in reality this ratio can vary for different parts of objects. This variance together with the scale variance of the object is captured by the scale parameters (σ_1, σ_2). In part 3, we show how we may allow multiple independent scales for multiple features with different scaling properties for each object, as illustrated here.

We next give a proposition which relates pullbacks of distributions under the scaling action. We consider $\sigma = \sigma(\lambda) : \Lambda (= \mathbb{R}^k_+) \to \mathbb{R}^k_+$ a smooth function of moderate growth such that $\Psi(x, \lambda) = (\sigma(\lambda) \cdot x, \sigma(\lambda) \cdot \lambda)$ defines a diffeomorphism of $\mathbb{R}^n \times \mathbb{R}^k_+$. We also consider the diffeomorphism of $\mathbb{R}^n \times \mathbb{R}^k_+$: $\psi_\sigma(x, \lambda) = (\sigma(\lambda) \cdot x, \lambda)$ (note that for fixed λ, ψ_σ restricts to a diffeomorphism of $\mathbb{R}^n$ which we still denote by ψ_σ).

Proposition 4.7. *Suppose G is scale–based with weight matrix A and scale weight B. Then, for $u \in \mathcal{S}^*_{n,\Lambda}$*

$$\Psi^*(u * G) = \sigma(\lambda)^{B+|A|}((\psi^*_\sigma u) * G) \tag{4.5}$$

where $|A|$ is the k-tuple given by $|A| = A_1 + \cdots + A_n$.

If we consider the case where σ is constant, then $\Psi = \Psi_\sigma$ defined earlier.

Corollary 4.8. *Suppose G is scale–based with variable weights A and scale weight B. Fix $\sigma \in \mathbb{R}^k_+$. Then, for $u \in \mathcal{S}^*_{n,\Lambda}$*

$$\Psi^*_\sigma(u * G) = \sigma^{B+|A|}((\psi^*_\sigma u) * G) \tag{4.6}$$

Remark . Before proving the proposition, we note that in the special case $u \in \mathcal{S}'_n$ and $B = -|A|$, we obtain the *Scale invariance of distributions under pullback*:

$$(\psi^*_\sigma u) * G \;=\; (u * G) \circ \Psi \;(=\; \Psi^*(u * G)) \tag{4.7}$$

This asserts that if we scale u and convolve then we obtain the same answer as first convolving and then scaling both the space and scale variables.

We observe that Gaussian kernels in Example 4.2 are scale–invariant by Example 4.4. Also, by Proposition 4.7, the medial filter given by the mexican hat function is scale-invariant. In general if we multiply the scale–based kernel G by $\sigma^{-B-|A|}$, then it becomes scale–invariant. Hence, if we multiply the edge filter by σ, it becomes scale invariant; and after multiplying the scale–based version of the Gabor filter by σ^{-n}, it becomes scale–invariant.

Proof. We represent $\Psi = (\psi_1, \psi_2)$, with $x' = \psi_1(x) = \psi_\sigma(x) = \sigma(\lambda) \cdot x$, $(y' = \psi_1(y))$, and $\lambda' = \psi_2(\lambda) = \sigma(\lambda) \cdot \lambda$. Then, we evaluate the RHS of (4.6) applied to $(x, \lambda) = \Psi^{-1}(x', \lambda')$.

$$\begin{aligned}(\psi^*_\sigma u) * G \circ \Psi^{-1}(x', \lambda') &= (\psi^*_\sigma u) * G(x, \lambda) \\ &= (\psi^*_\sigma u)(\tau_x \check{G}(y, \lambda))\end{aligned}$$

where to apply $\psi_\sigma^* u$, we view $\tau_x \check{G}(y,\lambda) = G(x-y,\lambda)$ as a function of (y,λ) for fixed x

$$\begin{aligned} &= (\psi_\sigma^* u)(G(x-y,\lambda)) \\ &= u(|\det(d\psi_\sigma^{-1})| \cdot G(x-\psi_\sigma^{-1}(y'),\lambda)) \end{aligned} \tag{4.8}$$

This last equality follows from the definition of pullback Definition 3.6 and (3.18), as $\psi_\sigma(x,\lambda) = (\sigma\cdot x,\lambda)$.

Since ψ_σ^{-1} is given by the action of σ^{-1}, we compute, recalling $A_i = (a_{i1},\dots,a_{ik})$,

$$|\det(d\psi_\sigma^{-1})| \;=\; \prod_i \sigma^{-A_i} \;=\; \sigma^{-|A|} \tag{4.9}$$

We multiply (4.8) by $\sigma^{B+|A|}$ and apply (4.9). We note that ψ_σ is linear in x and use the definition of multiplication of a uniformly tempered distribution by a function of moderate growth (as $\sigma = \sigma(\lambda)$ is a function of λ).

$$\sigma^{B+|A|}((\psi_\sigma^* u) * G)\circ\Psi^{-1}(x',\lambda') = u(\sigma^{B+|A|}\cdot\sigma^{-|A|}\cdot G(\psi_\sigma^{-1}(x'-y'),\lambda) \tag{4.10}$$

Applying the scale–based property of G to (4.10), we obtain

$$\begin{aligned} \sigma^{B+|A|}((\psi_\sigma^* u) * G)\circ\Psi^{-1}(x',\lambda') &= u(G(\sigma(\lambda)\cdot\psi_\sigma^{-1}(x'-y'),\sigma(\lambda)\cdot\lambda)) \\ &= u(G(x'-y',\lambda')) = u * G(x',\lambda') \end{aligned} \tag{4.11}$$

where in the last equation we used that Ψ is given by the action of $\sigma = \sigma(\lambda)$ (so that $\lambda' = \psi_2(\lambda) = \sigma(\lambda)\cdot\lambda)$. □

Action of the Poincaré Scaling Group. As a result of Corollary 4.8, we can describe the action of the Poincaré scaling group PS on convolutions.

In Theorem 3.7, we defined for a fixed $G \in \mathcal{S}_{n,\Lambda}$ a continuous linear transformations (3.19) sending $u \mapsto u * G$.

$$c_{\Lambda,\,G} : \mathcal{S}^*_{n,\Lambda} \to C^\infty(\mathbb{R}^n \times \Lambda)$$

and the corresponding restriction

$$c_G : \mathcal{S}'_n \to C^\infty(\mathbb{R}^n \times \Lambda)$$

Hence, the image of $c_{\Lambda,\,G}$ is a linear subspace of $C^\infty(\mathbb{R}^n\times\Lambda)$, which we denote by $\mathcal{H}_{\Lambda,\,G}$. This is the space of G–convolved uniformly tempered distibutions. Likewise, the image of c_G is a linear subspace of $C^\infty(\mathbb{R}^n\times\Lambda)$, which we denote by $\mathcal{H}_G$. This is the space of G–convolved tempered distibutions. We refer to these as spaces of G–*convolved distributions.* We define the action of PS as a subgroup of $\mathrm{Diff}(\mathbb{R}^n\times\mathbb{R}^k_+)$ acting on $C^\infty(\mathbb{R}^n\times\mathbb{R}^k_+)$ by $\varphi\cdot f = f\circ\varphi^{-1}$. To determine the action on convolutions, we first recall that convolution commutes with translation,

$$\tau_x^*(u * G) = (\tau_x^* u) * G.$$

Thus, we compute the action of $(\sigma, x) \in PS$ on a convolution by

$$\begin{aligned} (\tau_x\circ\Psi_\sigma)\cdot(u * G) &= ((\tau_x\circ\Psi_\sigma)^{-1})^*(u * G) \\ &= \tau_{-x}^*(\Psi_{\sigma^{-1}}^*(u * G)) \end{aligned}$$

Hence, by applying Corollary 4.8

$$
\begin{aligned}
&= \sigma^{B+|A|}((\tau^*_{-x}(\psi^*_{\sigma^{-1}}u)) * G)\\
&= \sigma^{B+|A|}((((\tau_x \circ \psi_\sigma)^{-1})^* u) * G)
\end{aligned}
\tag{4.12}
$$

The action of elements of PS define diffeomorphisms of (uniformly) moderate growth. By (4.12) together with Lemmas 2.4 and 3.5, we conclude that the action of PS on $C^\infty(\mathbb{R}^n \times \Lambda)$ restricts to an action on $\mathcal{H}_{\Lambda,\,G}$. We also note that as in Corollary 4.8, if $u \in \mathcal{S}'_n$, with fixed $\sigma \in \mathbb{R}^k_+$, then $\psi^*_\sigma u \in \mathcal{S}'_n$. Hence, the action induces an action on both $\mathcal{H}_{\Lambda,\,G}$ and $\mathcal{H}_G$.

We next see the role of scaling properties for determining the structure of jet space for convolved distributions.

Jet Spaces of Convolved Distributions $\mathcal{H}^\ell_{\Lambda,\,G}$ and $\mathcal{H}^\ell_G$. We first construct the associated jet spaces.

$$
\mathcal{H}^\ell_{\Lambda,\,G} \overset{\text{def}}{=} \{j^\ell(f)(x,\lambda) : f = u * G \text{ for } u \in \mathcal{S}^*_{n,\Lambda}\}
\tag{4.13}
$$

Recall from §1, the ℓ–jet $j^\ell(f)(x,\lambda)$ is the ℓ–th order Taylor expansion of f at (x,λ). We denote the set of jets in $\mathcal{H}^\ell_{\Lambda,\,G}$ at the point (x,λ) by $\mathcal{H}^\ell_{\Lambda,\,G\,(x,\lambda)}$. We analogously define $\mathcal{H}^\ell_G$.

We next determine the structure of $\mathcal{H}^\ell_{\Lambda,\,G}$ and $\mathcal{H}^\ell_G$ as fibrations over $\mathbb{R}^n \times \Lambda$ for $\Lambda = \mathbb{R}^k_+$ in the special case that G is scale-based. We first concentrate on $\mathcal{H}^\ell_{\Lambda,\,G}$.

We do so for scale–based kernels by using the consequences of the action of the Poincaré scaling group PS. Its action on $C^\infty(\mathbb{R}^n \times \Lambda)$ induces an action on the jet space $J^\ell(\mathbb{R}^n \times \Lambda, \mathbb{R})$. The induced action of PS on $\mathcal{H}_{\Lambda,\,G}$ and $\mathcal{H}_G$ also restricts to an action on the corresponding jet spaces $\mathcal{H}^\ell_{\Lambda,\,G}$ and $\mathcal{H}^\ell_G$. This allows us to deduce the structure of the convolution jet spaces.

Proposition 4.9. *If G is scale-based then for $\Lambda = \mathbb{R}^k_+$, both $\mathcal{H}^\ell_G$ and $\mathcal{H}^\ell_{\Lambda,\,G}$ are*

(1) *trivial fiber bundles over $\mathbb{R}^n \times \Lambda$, and*
(2) *semi–algebraic submanifolds of $J^\ell(\mathbb{R}^n \times \Lambda, \mathbb{R})$.*

From Proposition 4.9, if $\Lambda \subset \mathbb{R}^k_+$, we may restrict to jet spaces over $\mathbb{R}^n \times \Lambda$ and conclude that $\mathcal{H}^\ell_{\Lambda,\,G}$ is a trivial bundle over $\mathbb{R}^n \times \Lambda$.

Proof. By our above discussion, the action of the Poincaré scaling group PS restricts to an action on the jet spaces $\mathcal{H}^\ell_{\Lambda,\,G}$ and $\mathcal{H}^\ell_G$ commuting with the projection to $\mathbb{R}^n \times \mathbb{R}^k_+$ (here $\Lambda = \mathbb{R}^k_+$). As PS acts transitively on $\mathbb{R}^n \times \mathbb{R}^k_+$, we deduce diffeomorphisms of fiber bundles

$$
\mathcal{H}^\ell_{\Lambda,\,G} \simeq \mathcal{H}^\ell_{\Lambda,\,G,\,(x_0,\lambda_0)} \times (\mathbb{R}^n \times \mathbb{R}^k_+)
\tag{4.14}
$$

and

$$
\mathcal{H}^\ell_G \simeq \mathcal{H}^\ell_{G,\,(x_0,\lambda_0)} \times (\mathbb{R}^n \times \mathbb{R}^k_+)
\tag{4.15}
$$

where the inverse of this diffeomorphism sends

$$
(j^\ell(f)(x_0,\lambda_0), (x,\sigma) \mapsto j^\ell(f \circ (\Psi_\sigma \circ \tau_x)^{-1})(\sigma \cdot x_0 + x, \sigma \cdot \lambda_0)
$$

Also, these equations show that both $\mathcal{H}^\ell_{\Lambda,\,G}$ and $\mathcal{H}^\ell_G$ are the images of the fibers $\mathcal{H}^\ell_{\Lambda,\,G,\,(x_0,\lambda_0)}$, respectively $\mathcal{H}^\ell_{G,\,(x_0,\lambda_0)}$, which are linear subspaces (and hence algebraic submanifolds of $J^\ell(\mathbb{R}^n \times \Lambda, \mathbb{R})$), under the action of the real algebraic group

PS. Hence, the images are semi-algebraic submanifolds by the Tarski–Seidenberg theorem. □

5. Genericity and Stability Theorems for Convolution Operators

We have already established by Theorem 3.7 and Corollary 3.9 the continuity of the convolution maps $c_{\Lambda, G}$ and c_G for $G \in \mathcal{S}_{n,\Lambda}$ with $\Lambda = \mathbb{R}^k_+$. Second, by Proposition 4.9, in the case of a scale–based kernel G we have established that the convolution jet spaces $\mathcal{H}^\ell_G$ and $\mathcal{H}^\ell_{\Lambda, G}$ have the structure of globally trivial fiber bundles. We are now in position to establish theorems regarding both the genericity and stability of scale–based geometric structures. We will do so in terms of transversality to closed Whitney stratified subsets.

Properties of Whitney Stratifications. We recall basic properties of Whitney stratifications. By a *Whitney stratified subset* W of a smooth manifold M, we mean there is a locally finite decomposition of W into disjoint smooth submanifolds $W = \cup_i W_i$, called strata, satisfying: i) (axiom of the frontier) for any two strata W_i and W_j, with closure $\bar{W}_i$, if $\bar{W}_i \cap W_j \neq \emptyset$, then $W_j \subset \bar{W}_i$; and ii) Whitney's conditions a) and b), which give for a pair of strata $W_j \subset \bar{W}_i$, a relation between the limiting tangent spaces and limiting secant lines for W_i and tangent space of W_j (see e.g. [M2] or [Gi]). We also say that $\{W_i\}$ is a Whitney stratification of W.

Several of the key properties of Whitney stratified sets are: i) being a Whitney stratified subset is preserved under diffeomorphisms; ii) orbits of algebraic group actions, where they are locally finite, form a Whitney stratification; iii) more generally, algebraic or semi–algebraic subsets have Whitney stratifications with semi–algebraic strata; and iv) the pullback $f^{-1}(W)$ of a Whitney stratified subset W by a map f transverse to W, i.e. to the strata $\{W_i\}$, is also a Whitney stratified subset with strata $\{f^{-1}(W_i)\}$.

Second, a fundamental property of Whitney stratified sets is given by

Theorem 5.1 (Thom Isotopy Theorem [Th]). *Suppose W is a closed Whitney stratified set and $f : W \to \mathbb{R}$ is a continuous proper mapping such that for each stratum W_i of W $f|W_i \to \mathbb{R}$ is a smooth submersion at each point of W_i. Then, for any $t_1, t_2 \in \mathbb{R}$, $f^{-1}(t_1)$ is homeomorphic to $f^{-1}(t_2)$ by a homeomorphism which is a diffeomorphism on each stratum $f^{-1}(t_i) \cap W_j$.*

First, the Thom Isotopy Theorem has the following consequence for pullbacks of Whitney Stratifications.

Corollary 5.2 ([M2], [Gi]). *Suppose $f_t : M \to N$ is a smooth homotopy of proper smooth maps such that for each $0 \leq t \leq 1$, f_t is transverse to a Whitney stratified set $W \subset N$. Then, $f_0^{-1}(W)$ is homeomorphic to $f_1^{-1}(W)$ by a homeomorphism which is a diffeomorphism on each stratum.*

Second, there are two main consequences concerning the local structure of a Whitney stratified set $W \subset \mathbb{R}^n$.

Corollary 5.3 ([M2], [Gi]). *Let $W \subset \mathbb{R}^n$ be a Whitney stratified set. Suppose W_ℓ is a stratum of W of codimension d, and that $x_i \in W_\ell$ for $i = 1, 2$ belong to the same connected component of W_ℓ. If P_i are d dimensional affine subspaces of $\mathbb{R}^n$ which intersect W_ℓ transversely at x_i, then, there is a local homeomorphism $\varphi : U_1 \cap W, x_1 \xrightarrow{\sim} U_2 \cap W, x_2$, for appropriate neighborhoods $x_i \in U_i \subset P_i$.*

These normal sections $(U \cap W, x)$ determine the normal structure of W near a point $x \in W_\ell$. Second, this local structure can be locally described.

Corollary 5.4 ([M2], [Gi]). *Let P be a normal section through x of a stratum W_ℓ of a Whitney stratified set $W \subset \mathbb{R}^n$. Then, there exists $\epsilon_0 > 0$ such that for $0 < \epsilon < \epsilon_0$, there is a homeomorphism*

$$\bar{B}_\epsilon(x) \cap P \cap W \ \simeq \ c(S_\epsilon(x) \cap P \cap W)$$

where $B_\epsilon(x)$ (resp. $S_\epsilon(x)$) denote the ball (resp. sphere) of radius ϵ; and $c(X)$ denotes the cone on X.

We shall denote this local normal structure of W near x by $W_{loc}(x)$,

Genericity of Properties via Transversality to Whitney Stratifications. We will concentrate on closed Whitney stratified subsets $W \subset J^\ell(U, \mathbb{R})$ with strata $\{W_i\}$ for open subsets $U \subset \mathbb{R}^n \times \Lambda$. For a smooth mapping, $f : U \to \mathbb{R}$, there is the "jet extension map"$j^\ell(f) : U \to J^\ell(U, \mathbb{R})$ sending $x \mapsto j^\ell(f)(x)$. Many singularity theoretic and geometric conditions can be expressed in terms of transversality to orbits of algebraic group actions or more generally to semi–algebraic submanifolds of $J^\ell(U, \mathbb{R})$. Then, by property iv), the pullback of a Whitney stratified subset of $W \subset J^\ell(U, \mathbb{R})$ by the jet extension map $j^\ell(f)$ gives a Whitney stratified set $W(f) = j^\ell(f)^{-1}(W)$ with strata $W_i(f) = j^\ell(f)^{-1}(W_i)$.

We are more specifically concerned with closed Whitney stratified subsets of the convolution jet spaces $\mathcal{H}^\ell_{\Lambda,\, G}$ or $\mathcal{H}^\ell_G$, for G a scale-based kernel. For an element f in $\mathcal{H}_{\Lambda,\, G}(U) (= \mathcal{H}_{\Lambda,\, G\,|U})$,

$$j^\ell(f) : U \longrightarrow \mathcal{H}_{\Lambda,\, G}(U) \hookrightarrow J^\ell(U, \mathbb{R}) \tag{5.1}$$

There is an analogous map as 5.1 for $\mathcal{H}_G$ instead. When the G–convolution jet space is understood, we will abuse notation and denote the composition by $j^\ell(f)$.

Beginning with a closed Whitney stratified set $W \subset J^\ell(U, \mathbb{R})$ with strata $\{W_i\}$, there are two possibilities for obtaining a closed Whitney stratified subset of the convolution jet spaces. We consider this in more detail in §8. For now, we just note that if a convolution jet space, say $\mathcal{H}^\ell_{\Lambda,\, G}$ is transverse to W, then by property iv), the intersection W' is a closed Whitney stratified subset of $\mathcal{H}^\ell_{\Lambda,\, G}$ with strata $\{W'_i = W_i \cap \mathcal{H}^\ell_{\Lambda,\, G}\}$ (with an analogous result for $\mathcal{H}^\ell_G$). Alternately, if $W \subset J^\ell(U, \mathbb{R})$ is a closed semialgebraic subset, but $\mathcal{H}^\ell_{\Lambda,\, G}$ (or $\mathcal{H}^\ell_G$) is not transverse to W then Proposition 4.9 allows us to obtain a natural Whitney stratification for the intersection (see Proposition 8.1). In this second case, transversality to the intersection W' in $\mathcal{H}^\ell_{\Lambda,\, G}$ or $\mathcal{H}^\ell_G$ need not exhibit the same local structure as for W; however, it will still define a geometric structure for G–convolved distributions.

First, we introduce the notion of genericity for convolutions of distributions.

Definition 5.5. We say that a property $\mathcal{P}$ is *generic for G-convolutions of (uniformly) tempered distributions*, if for any compact subset $C \subset \mathbb{R}^n \times \Lambda$ there is an open dense subset of (uniformly) tempered distributions u such that $u * G$ possess property $\mathcal{P}$ at each point of C.

Example 5.6. For $G = K_t$, the Gaussian kernel, K_t–convolutions of tempered distributions are solutions to the heat equation; and the local property $\mathcal{P}$ being *generic for K_t–convolutions of tempered distributions* is equivalent to it being generic for solutions to the heat equation on $\mathbb{R}^n \times \mathbb{R}_{++}$ in the sense of [D5].

Based on the above discussion, we give a restricted definition of genericity which is especially suited for a singularity theoretic approach to scale-based geometric questions expressed in terms of tranversality of jet extension maps.

Definition 5.7. Given a scale-based kernel G, we will say a local geometric property $\mathcal{P}$ is a *tranversally defined scale-based property* if there is a closed Whitney stratified subset $W \subset \mathcal{H}^{\ell}_{\Lambda,\,G}$ with strata W_i, (or respectively $W \subset \mathcal{H}^{\ell}_{G}$) such that $f = u * G$ having property $\mathcal{P}$ at all points of a subset $C \subset \mathbb{R}^n \times \Lambda$ is characterized by $j^{\ell}(f)$ being transverse on C to W (i.e. to the strata W_i), where transversality is relative to $\mathcal{H}^{\ell}_{\Lambda,\,G}$ (respectively $\mathcal{H}^{\ell}_{G}$). Then, we say that f *exhibits property* $\mathcal{P}$ *generically on* C.

Suppose a property $\mathcal{P}$ is defined by transversality to the closed Whitney stratified $W \subset \mathcal{H}^{\ell}_{\Lambda,\,G}$ or $\mathcal{H}^{\ell}_{G}$. If $f = u * G$ is transverse to W, i.e. $j^{\ell}(f)$ is transverse to W, then $j^{\ell}(f)^{-1}(W)$ is a Whitney stratified set with strata $\{j^{\ell\,-1}(W_i)\}$. We refer to $j^{\ell\,-1}(W)$ as the *geometric structure associated to* $\mathcal{P}$. The individual strata will represent different geometric behavior associated to the property $\mathcal{P}$.

Genericity and Stability via Transversality to Whitney Stratifications. We can now give a sufficient condition for genericity for scale-based geometric properties.

Theorem 5.8 (Genericity via Relative Transversality). *1) Suppose W is a closed Whitney stratified subset of $\mathcal{H}^{\ell}_{\Lambda,\,G}$ with strata W_i, then for a compact set $C \subset (\mathbb{R}^n \times \Lambda)$, the set of uniformly tempered distributions*

$$\mathcal{W} = \{u \in \mathcal{S}^*_{n,\Lambda} : j^{\ell}(u * G) \textit{ is transverse on } C \textit{ to all } W_i (\textit{in } \mathcal{H}^{\ell}_{\Lambda,\,G})\} \tag{5.2}$$

is an open dense subset of $\mathcal{S}^*_{n,\Lambda}$.

2) If instead, W is a closed Whitney stratified subset of $\mathcal{H}^{\ell}_{G}$, then

$$\mathcal{W} = \{u \in \mathcal{S}'_{n} : j^{\ell}(u * G) \textit{ is transverse on } C \textit{ to all } W_i (\textit{in } \mathcal{H}^{\ell}_{G})\} \tag{5.3}$$

is an open dense subset of $\mathcal{S}'_n$.

As a corollary of Theorem 5.8, we conclude

Corollary 5.9. *If $\mathcal{P}$ is a transversally defined scale–based property, then it is generic for G-convolutions of (uniformly) tempered distributions.*

Proof. The Thom Transversality Theorem does not apply to the space of functions obtained as G–convolved distributions. We will apply instead a generalization of Thom's theorem, the "relative transversality theorem"[D2, Thm. 1.3], which can be applied to the convolution maps $c_{\Lambda,\,G}$ or c_G. As the proofs for both cases are very similar, we shall concentrate on the case of uniformly tempered distributions. The relative transversality theorem will imply that given a compact subset $C \subset \mathbb{R}^n \times \Lambda$, there is an open dense subset $\mathcal{W} \subset \mathcal{S}^*_{n,\Lambda}$ consisting of u such that $j^{\ell}(u * G)$ is transverse on C to W in $\mathcal{H}^{\ell}_{\Lambda,\,G}$.

To apply the relative transversality theorem, we must show that $\mathcal{S}^*_{n,\Lambda}$ has "smooth image"in $J^{\ell}(\mathbb{R}^n \times \Lambda, \mathbb{R})$ in the sense of [D2, Def. 1.1].

As $c_{\Lambda,\,G}$ is linear, given $(x_0, \lambda_0) \in \mathbb{R}^n \times \Lambda$, there is a finite dimensional linear subspace $\mathcal{V}$ of $\mathcal{S}^*_{n,\Lambda}$ such that the map $\mathcal{V} \to \mathcal{H}^{\ell}_{\Lambda,\,G\ (x_0,\lambda_0)}$ obtained by composing $c_{\Lambda,\,G}$ with $j^{\ell}_{(x_0,\lambda_0)}$ is surjective. Then, it will continue to be surjective for all (x, λ) in a neighborhood U of (x_0, λ_0) and then for all subspaces $u + \mathcal{V}$. Thus, for any $u \in \mathcal{S}^*_{n,\Lambda}$, with $\mathcal{V}' = u + \mathcal{V}$

$$\begin{aligned} \mathcal{V}' \times U &\longrightarrow \mathcal{H}^{\ell}_{\Lambda,\,G} \\ (u', (x,\lambda)) &\mapsto j^{\ell}(u' * G)(x,\lambda) \end{aligned} \tag{5.4}$$

maps surjectively onto $\mathcal{H}^{\ell}_{\Lambda,\,G\,|\mathcal{V}'\times U}$. As $\mathcal{H}^{\ell}_{\Lambda,\,G}$ is a smooth trivial subbundle invariant under translation on $\mathbb{R}^n$ and the scaling action, we obtain the analogue of (5.4) for any point (x', λ') by translation and scaling action. This implies that $c_{\Lambda,\,G}$ has smooth image. Hence, the relative transversality theorem [D2, Thm. 1.3] applies and the set $\mathcal{W}$ is open and dense. An analogous argument works for c_G. □

Generic Structure and Stability for Scale-based Geometric Structures. Given a scale–based kernel G, let $\mathcal{P}$ be a transversally defined scale–based geometric property defined via the closed Whitney stratified subset W, with strata $\{W_i\}$. We can now deduce the generic properties and stability of $W(f)$ for $f = u * G$.

Theorem 5.10 (Scale–based Generic Structure). *Suppose G is a scale–based kernel and that $\mathcal{P}$ is a transversally defined scale–based geometric property. If $f = u * G$ exhibits $\mathcal{P}$ generically on a compact subset C, then at an interior point (x_0, λ_0) of C, there is a strata preserving local homeomorphism*

$$(W(f), (x_0, \lambda_0)) \quad \simeq \quad (W_{loc}(j^{\ell}(f)(x_0, \lambda_0)) \times \mathbb{R}^m, (j^{\ell}(f)(x_0, \lambda_0), 0))$$

where $m = n - \operatorname{codim}(W_i)$ for $j^{\ell}(f)(x_0, \lambda_0) \in W_i$.

Thus, the products of $\{W_{loc}(z)\}$ with appropriate factors $\mathbb{R}^{m_i}$ provide the local models for the generic structure associated to $\mathcal{P}$.

Second, we deduce the stability of $W(f)$. Given $f = u * G$ which exhibits $\mathcal{P}$ generically on a compact set C, we may extend C to a slightly larger compact manifold of dimension $n + k$ with boundaries and corners (in the sense of Mather [M-II]), on which f still exhibits $\mathcal{P}$ generically (by the openness of transversality to closed Whitney stratified sets). We may further arrange that $W(f)$ is transverse to the stratification of the boundary ∂C. Then, stability takes the following form.

Theorem 5.11 (Scale–based Structural Stability). *Suppose G is a scale–based kernel and that $\mathcal{P}$ is a transversally defined scale–based geometric property. If $f = u * G$ exhibits $\mathcal{P}$ generically on a compact subset C (which is an $n + k$ manifold with boundaries and corners) so $W(f)$ intersects ∂C transversely, then, there is an open neighborhood $\mathcal{U}$ of u in $\mathcal{S}^*_{n,\Lambda}$ (resp. $\mathcal{S}'_n$) such that for $u' \in \mathcal{U}$, $f' = u' * G$ exhibits $\mathcal{P}$ generically on C, and there is a strata preserving homeomorphism of C sending $W(f) \cap C$ to $W(f') \cap C$, which is smooth on each stratum.*

Proof of Theorem 5.10. By Corollary 5.3, $(W(f), x) \simeq (W(f)_{loc}(x) \times \mathbb{R}^m, x)$ for appropriate m. Then, the composition of a normal section to $W(f)$ at x composes with $j^{\ell}(f)$ to yield a (nonlinear) normal section to W at $j^{\ell}(f)(x)$. This (nonlinear but transverse) normal section and a normal section are smoothly homotopic while remaining transverse to W at $j^{\ell}(f)(x)$. Hence, by Corollary 5.2 of the Thom Isotopy theorem, $(W(f)_{loc}(x), x) \simeq (W_{loc}(j^{\ell}(f)(x)), j^{\ell}(f)(x))$, yielding the result. □

Proof of Theorem 5.11. Suppose C is a compact manifold with boundaries and corners, and $f = u * G$ exhibits $\mathcal{P}$ generically on C, with $W(f)$ is transverse to ∂C. To be specific we suppose $W \subset \mathcal{H}^{\ell}_{\Lambda,\,G}$. An analogous argument will apply to $\mathcal{H}^{\ell}_{G}$. Then, first there is a neighborhood $\mathcal{W}$ of u consisting of u' for which $f' = u' * G$ satisfies: $j^{\ell}(f')$ is transverse to W on C and $W(f')$ is transverse to ∂C. Then, by results of Mather [M-II] there is a neighborhood $\mathcal{V}$ of f in $C^{\infty}(\mathbb{R}^n \times \Lambda)$ such that for any $f' \in \mathcal{V}$, there is a smooth homotopy f_t between f and f' with $j^{\ell}(f_t)$ transverse to W on C and ∂C. Then, by Corollary 5.2 of the Thom isotopy theorem, $W(f) \cap C$ and $W(f') \cap C$ are homeomorphic by a homeomorphism which is smooth on the strata (and the strata intersected with ∂C). Finally, by the continuity of $c_{\Lambda, G}$ (resp. c_G), there is an open neighborhood $\mathcal{W}'$ of $u \in \mathcal{T}$ which maps into $\mathcal{V}$. Thus, we may replace $\mathcal{W}$ by $\mathcal{W}' \cap \mathcal{W}$ to obtain the desired neighborhood. $\square$

Part 2. Generic Scale–based Geometry for Subspaces of Distributions, and Discrete Functions and Measures

In this part. we extend the genericity and stability results from Part 1 to many important subspaces of distributions. To do so, we must impose conditions on both the scale–based kernel G (condition (A)) and on the subspace $\mathcal{T}$ (condition (B)). Then, we prove (Theorem 6.3) that the distributions in $\mathcal{T}$ will generically exhibit the same scale– based geometric properties as the full space of (uniformly) tempered distributions, including the generic local structure and the stability of the structures. Furthermore, we show condition (B) is satisfied for all of the key subspaces of tempered distributions (Theorem 6.4), with the exception of the spaces of probability measures, for which the results are still shown to hold by (Corollary 6.5).

In §7, we further consider genericity results for discrete and piecewise linear functions and discrete measures. We establish the density of the set of discrete functions and measures which generically exhibit a scale–based geometric property on a compact subset of scale space. Furthermore, we show that for a given compact subset C of scale space, there is a single refinement of the mesh and enlargement of the region, so that any nongeneric discrete function or measure can be approximated by another discrete one with generic properties on C. (Theorems 7.6 and 7.8).

6. Subspaces of Tempered Distributions

Theorems 5.8, 5.10, and 5.11 have established genericity and stability of transversally defined properties $\mathcal{P}$ for spaces of tempered and uniformly tempered distributions. They guarantee that for a compact subset C of scale space, there is an open dense subset of (uniformly) tempered distributions whose G-convolved functions exhibit $\mathcal{P}$ generically and stably on C. We are also interested in various subspaces of distributions such as L^p functions, functions in Sobolev spaces, functions of moderate growth or of compact support, and measures such as positive regular Borel measures, probability measures, etc.

Unless a subspace forms an open subset of the space of (uniformly) tempered distributions, we cannot conclude from Theorem 5.8 that genericity also holds for the subspace. Hence, we can conclude nothing about these particular subspaces of tempered distributions. Nonetheless, we claim that the conclusions of Theorems 5.8, 5.10, and 5.11 are equally valid for all of these subspaces. We give a general criterion which will yield the same consequences and which will apply to all of the subspaces we have mentioned. Second, in §7 we give a formulation of how the discrete objects also have discrete approximations which exhibit these properties generically.

We consider a scaled–based kernel G. We shall either consider $\mathcal{H}_{\Lambda, G}$ or $\mathcal{H}_G$. Hence, we use the notation $(\mathcal{S}^*, \mathcal{H}, \mathcal{H}^\ell)$ to denote either triple $(\mathcal{S}'_n, \mathcal{H}_G, \mathcal{H}^\ell_G)$ or $(\mathcal{S}^*_{n,\Lambda}, \mathcal{H}_{\Lambda, G}, \mathcal{H}^\ell_{\Lambda, G})$.

Definition 6.1. We say that the scaled–based kernel G satisfies condition (A) if there is $(x_0, \lambda_0) \in \mathbb{R}^n \times \Lambda$ and a finite dimensional subspace $\mathcal{V}$ consisting of smooth functions of (uniformly) moderate growth such that the composition is surjective.

$$\begin{array}{ccc} V & \to & \mathcal{H}^\ell_{(x_0,\lambda_0)} \\ h & \mapsto & j^\ell(h * G)(x_0, \lambda_0) \end{array}$$

We shall see in §9 that using the local basis for solutions to the heat equation given in [D1], we can establish condition (A) for the Gaussian kernel and large class of "extended Gaussian kernels"which includes the median kernel, edge kernel as well as many other kernels derived from the Gaussian kernel allowing anisotropy, derivatives, and mixed multiple independent scales.

Along with the condition on G, we also have a condition on the (not necessarily linear) subspace $\mathcal{T} \subset \mathcal{S}^*$. We do not suppose that $\mathcal{T}$ has the subspace topology, but only that the inclusion is continuous.

Definition 6.2. A subspace $\mathcal{T} \subset \mathcal{S}^*$ is said to satisfy condition (B) if

(1) $\mathcal{T}$ is closed under nonnegative linear combinations,i.e. if $u_i \in \mathcal{T}$ then $\sum_i a_i u_i \in \mathcal{T}$ for all $a_i \geq 0$.
(2) $\mathcal{T}$ contains all smooth nonnegative functions of compact support.

We note that all of the subspaces just mentioned as well as many others (with the exception of probability measures) satisfy condition (B). The spaces of probability measures are only closed under convex combinations, and we shall consider separately.

For a subspace $\mathcal{T} \subset \mathcal{S}^*$, we are interested in when a property $\mathcal{P}$ is *generic for G-convolutions of (distributions in) $\mathcal{T}$*. In analogy with Definition 5.5, this shall mean that for any compact subset $C \subset \mathbb{R}^n \times \Lambda$ there is an open dense subset $\mathcal{W}$ of $u \in \mathcal{T}$ such that for all $u \in \mathcal{W}$, $u * G$ possesses property $\mathcal{P}$ at each point of C.

Theorem 6.3. *Suppose G is a scale–based kernel satisfying condition (A); and that $\mathcal{T} \subset \mathcal{S}^*$ is a subspace satisfying condition (B). If $\mathcal{P}$ is a transversally defined scale–based property, then it is generic for G–convolutions of distributions in $\mathcal{T}$.*

Furthermore, for a compact subset C on which $u * G$ exhibits $\mathcal{P}$ generically, the local generic structure of $W(u*G)$ is given on C as in Theorem 5.10 by $\{W_{loc}(z)\}$; and $W(u*G)$ is structurally stable on C for sufficiently small perturbations of $u \in \mathcal{T}$.

Given Theorem 6.3, with the same notation as above, we deduce the following corollary for all reasonable subspaces of (uniformly) tempered distributions.

Theorem 6.4. *All of the following subspaces of distributions satisfy condition (B). Hence, for any scale–based kernel G satisfying condition (A), any transversally defined scale–based property $\mathcal{P}$ is exhibited generically and stably for G–convolutions on compact subsets of scale space.*

(1) *L^p functions and L^p_ℓ–functions for $1 \leq p < \infty$ and (see Example 2.1), functions in Sobolev spaces $H^{\ell,p}$, $1 \leq p < \infty$;*
(2) *positive regular Borel measures of moderate growth $\mathcal{BM}_\ell$, signed regular Borel measures of moderate growth $\mathcal{SBM}_\ell$ (again see Example 2.1);*
(3) *functions in (1) with compact support;*
(4) *C^k functions with compact support, for fixed k, $0 \leq k \leq \infty$;*
(5) *positive and signed regular Borel measures of compact support $\mathcal{BM}_c$ and $\mathcal{SBM}_c$;*
(6) *distibutions with compact support.*

Note: for spaces of measures, we include smooth functions with compact support f as measures $f dx$. The only spaces we have not included are the probability measures. These are taken care of by the following.

Corollary 6.5. *Suppose the scale–based kernel G satisfies condition (A) and $\mathcal{P}$ is a transversally defined scale–based property via the closed Whitney stratified subset*

W. If W is invariant under multiplication by scalars ($z \in W$ implies $cz \in W$ for any $c \in \mathbb{R}_+$), then G–convolutions on the spaces of probability measures $\mathcal{PM}$, or probability measures with compact supports $\mathcal{PM}_c$, also exhibit $\mathcal{P}$ generically (and stably) on compact subsets of scale space.

To prove Theorem 6.3, we again want to use the relative transversality theorem. We give a proposition which exhibits a subspace needed for applying the relative transversality theorem to the convolution map restricted to $\mathcal{T}$.

Proposition 6.6. *Let G be a scale–based kernel. Suppose there is a finite dimensional subspace V' of smooth functions of (uniformly) moderate growth which satisfy condition (A) for G. Then, there is a finite dimensional subspace $\mathcal{V}$ consisting of smooth non–negative functions of compact support such that $\mathcal{V} \to \mathcal{H}^{(\ell)}_{(x_0,\lambda_0)}$ sending $u \mapsto j^{(\ell)}(u * G)(x_0, \lambda_0)$ is surjective.*

Proof of Theorem 6.3. The proof of this Theorem follows the lines of that for Theorem 5.8. Since G satisfies condition (A), Proposition 6.6 provides the linear subspace $\mathcal{V}$ spanned by a finite set of smooth nonnegative functions of compact support, $\{\varphi_i, i = 1, \ldots, m\}$, such that the map $\mathcal{V} \to \mathcal{H}^{(\ell)}_{(x_0,\lambda_0)}$ sending $u \mapsto j^{(\ell)}(u * G)(x_0, \lambda_0)$ is surjective. By openness of surjectivity, there is an open neighborhood U of (x_0, λ_0) such that the corresponding map for $(x', \lambda') \in U$ is surjective. Given (x_1, λ_1), there is an element $(\tau_x, \sigma) \in PS$ so that $(\tau_x, \sigma) \cdot (x_0, \lambda_0) = (x_1, \lambda_1)$. Then, by Corollary 4.8, the subspace $\mathcal{V}_1$ spanned by $\{(\tau_x, \sigma)^*(\varphi_i)\}$ will also map surjectively onto $\mathcal{H}^{(\ell)}_{(x_1,\lambda_1)}$, and hence also for $(x', \lambda') \in U_1$ an open neighborhood of (x_1, λ_1). By assumption 2) of condition (B), $\mathcal{V}_1 \subset \mathcal{T}$. Thus, we conclude that $c_{\Lambda,G}|\mathcal{T}$ has smooth image, and we can directly apply the relative transversality theorem for those $\mathcal{T}$ which are linear subspaces.

For the subspaces only closed under nonnegative linear combinations such as positive Borel measures, we must note that the proof of the relative transversality theorem still works if we find an open convex subspace $\mathcal{V}'$ of $\mathcal{V}$ which satisfies the condition for smooth image in the relative transversality theorem except that the given u belongs to the closure of $\mathcal{V}'$. With this observation, we can apply the relative transversality theorem to the remaining cases. $\square$

Proof of Corollary 6.5. By Proposition 6.6, given $(x_0, \lambda_0) \in \mathbb{R}^n \times \mathbb{R}^k_+$, there is a finite set of smooth nonnegative functions $\{\varphi_i, i = 1, \ldots, m\}$ of (uniformly) moderate growth such that $j^\ell(\varphi_i * G)(x_0, \lambda_0)$ span the linear space $\mathcal{H}^\ell_{(x_0,\lambda_0)}$. The corresponding measures $\varphi_i dx$ define positive regular Borel measures of compact support. We denote the linear subspace spanned by them by $\mathcal{V}$. Given a probability measure $\mu \in \mathcal{PM}$ (respectively $\mathcal{PM}_c$), then $\mathcal{V}_1 = \mu + \mathcal{V} \subset \mathcal{PM}$ (respectively $\mathcal{PM}_c$). Let $\mathcal{V}_2 = \{\mu' \in \mathcal{V}_1 : \int_{\mathbb{R}^n} \mu' = 1\}$. Then, $\mathcal{V}_2$ is an affine subspace of $\mathcal{BM}_b$ (respectively $\mathcal{BM}_c$). Then, we let $\mathcal{V}'$ denote the subspace of $\mathcal{V}_2$ consisting of measures of the form $\mu' = \mu + \sum a_i \varphi_i dx$ with $a_i > 0$. Then, $\mathcal{V}'$ is an open convex subset of $\mathcal{V}_2$ with μ in its closure.

Since the map $\mathcal{V} \to \mathcal{H}^\ell_{(x_0,\lambda_0)}$ sending $\mu' \mapsto j^\ell(\mu' * G)(x_0, \lambda_0)$ is surjective, the jet map $j^\ell(\mu' * G)(x, \lambda)$ will continue to be surjective for (x, λ) in a neighborhood U of (x_0, λ_0). Hence, the map $\mathbb{R}_+ \times \mathcal{V}' \times U \to \mathcal{H}^\ell$ sending $(c, \mu', (x, \lambda)) \mapsto j^\ell(c\mu' * G)(x, \lambda)$ will be a submersion. Hence, it will be transverse to W (in $\mathcal{H}^\ell$). As W is invariant under scalar multiplication, the restriction to $\mathcal{V}' \times U \simeq \{1\} \times \mathcal{V}' \times U$ will be transverse to W. Hence, it follows that $c_{\Lambda,G}$ (respectively c_G) is transverse to W in the sense of [D2, Def 1.4], so we can apply instead the absolute transversality theorem [D2, Thm 1.5] (using $\mathcal{H}^\ell$ in place of $J^\ell(\mathbb{R}^n \times \mathbb{R}^k_+)$) to conclude that the property $\mathcal{P}$ is generic for G–convolutions of probability measures in $\mathcal{PM}$ or $\mathcal{PM}_b$. By the same arguments used to prove Theorems 5.10, and 5.11, we deduce both the generic local structure and stability for $W(\mu * G)$ on any compact subset C on which $\mu * G$ exhibits $\mathcal{P}$ generically. $\square$

Proof of Proposition 6.6. We first fix a family of smooth nondecreasing "step func tions" $\tilde{\chi}_\epsilon$ on $\mathbb{R}$ for which $0 \leq \tilde{\chi}_\epsilon \leq 1$, $\tilde{\chi}_\epsilon(x) = 0$ for $x \leq 0$, and $\tilde{\chi}_\epsilon(x) = 1$ for $x \geq \epsilon$. Then, we define $\chi_R(x) = 1 - \tilde{\chi}_{\frac{1}{R}}(\|x\| - R)$ for $x \in \mathbb{R}^n$. Again, $0 \leq \chi_R \leq 1$, $\chi_R(x) = 1$ for $\|x\| \leq R$, and $\chi_R(x) = 0$ for $\|x\| \geq R + \frac{1}{R}$. Then, we claim

Lemma 6.7. *Let φ be a smooth function of (uniformly) moderate growth with compact support in $\mathbb{R}^k_+$, then $\chi_R\varphi \to \varphi$ in $\mathcal{S}^* = \mathcal{S}^*_{n,\Lambda}$ (respectively $\mathcal{S}'_n$). Specifically, for any neighborhood $\mathcal{U}$ of 0, there is an $R_0 > 0$ so that $\chi_R\varphi - \varphi \in \mathcal{U}$ for all $R \geq R_0$.*

By φ having compact support in $\mathbb{R}^k_+$, we mean there is a compact subset $K \subset \mathbb{R}^k_+$ such that $\operatorname{supp}(\varphi) \subset \mathbb{R}^n \times K$.

Given the Lemma, we proceed with the proof as follows. We give the argument for uniformly tempered distributions which specializes to the case of tempered distributions.

Let $\{\varphi_1, \dots, \varphi_m\}$ be a set of functions of uniformly moderate growth such that

$$\{j^\ell(\varphi_i * G)(x_0, \lambda_0), i = 1, \dots, m\} \text{ span } \mathcal{H}^\ell_{\Lambda,\, G\,(x_0,\lambda_0)}.$$

First, we may assume that each φ_i has compact support in $\mathbb{R}^k_+$. To see this, let $\psi(\lambda)$ be a smooth bump function on $\mathbb{R}^k_+$ with compact support K, such that $\psi \equiv 1$ on a neighborhood of λ_0. Then, as convolution is with respect to x, we easily check

$$j^\ell(\psi(\lambda) \cdot \varphi_i * G)(x_0, \lambda_0) \quad = \quad j^\ell(\varphi_i * G)(x_0, \lambda_0)$$

Thus, we can replace each φ_i by $\psi \cdot \varphi_i$, so we may as well assume each φ_i has compact suport in $\mathbb{R}^k_+$.

Then, define $\Phi \in \hom(\mathbb{R}^m, \mathcal{H}^\ell_{\Lambda,\, G\,(x_0,\lambda_0)})$ by

$$\Phi(u_1, \dots, u_m) = \sum_i u_i j^\ell(c_{\Lambda,\, G}(\varphi_i))(x_0, \lambda_0) \tag{6.1}$$

and an analogous linear transformation Φ_R using instead $\chi_R\varphi_i$ in place of φ_i.

$$\Phi_R(u_1, \dots, u_m) = \sum_i u_i j^\ell(c_{\Lambda,\, G}(\chi_R\varphi_i))(x_0, \lambda_0) \tag{6.2}$$

Then, by the continuity of $c_{\Lambda,\, G}$ and Lemma 6.7, $c_{\Lambda,\, G}(\chi_R\varphi_i)$ approaches $c_{\Lambda,\, G}(\varphi_i)$ as $R \to \infty$. We conclude that $\Phi_R \to \Phi$ in $\hom(\mathbb{R}^m, \mathcal{H}^\ell_{\Lambda,\, G\,(x_0,\lambda_0)})$ as $R \to \infty$. Since Φ is surjective, there is an R_0 so that Φ_R is surjective for $R \geq R_0$. If the $\chi_R\varphi_i$ were nonnegative , then they would be our desired functions. Lastly, we alter them to attain this property as well.

Let $C_{i\,R} = \max_{\|x\| \leq R + \frac{1}{R}} |\varphi_i|$, and let $a_{i\,R} > C_{i\,R}$ denote a number to be chosen. Then, $\varphi'_{i\,R} = \chi_R(\varphi_i + a_{i\,R})$ is smooth, nonnegative, and has compact support. We claim that for each $R \geq R_0$, there are $a_{i\,R}$ so that $\{j^\ell(c_{\Lambda,\, G}(\varphi'_{i\,R}))(x_0, \lambda_0)\}$ span $\mathcal{H}^\ell_{\Lambda,\, G\,(x_0,\lambda_0)}$ as required. In fact,

$$j^\ell(c_{\Lambda,\, G}(\varphi'_{i\,R}))(x_0, \lambda_0) = j^\ell(c_{\Lambda,\, G}(\chi_R\varphi_i))(x_0, \lambda_0) + a_{i\,R}g_R$$

where $g_R = j^\ell(c_{\Lambda,G}(\chi_R))(x_0,\lambda_0)$. For fixed $R \geq R_0$, $\{j^\ell(c_{\Lambda,G}(\chi_R\varphi_i))(x_0,\lambda_0)\}$ span $\mathcal{H}^\ell_{\Lambda,G\,(x_0,\lambda_0)}$; hence, there is a nonempty Zariski open subset of $(a_{1\,R},\ldots,a_{m\,R})$ such that $\{j^\ell(c_{\Lambda,G}(\varphi'_{i\,R}))(x_0,\lambda_0)\}$ span $\mathcal{H}^\ell_{\Lambda,G\,(x_0,\lambda_0)}$. Thus, for each such R we choose $(a_{1\,R},\ldots,a_{m\,R})$ from this Zariski open subset and satisfying $a_{i\,R} > C_{i\,R}$ yielding the desired set of smooth non–negative functions with compact support. □

Finally we prove Lemma 6.7.

Proof of Lemma 6.7. Suppose φ has uniformly moderate growth. Integration $h \mapsto \int \varphi(x,\lambda)h(x,\lambda)dx$ defines a uniformly tempered distribution u_φ. Given a neighborhood $U(M,B(\mathbf{b}),\varepsilon)$ of 0, we will find a constant R_0, an integer M', and a constant $c < \varepsilon/b_{M'}$ such that for $R \geq R_0$,

$$\|(\chi_R - 1)\cdot u_{\varphi\cdot}(h)\|_M \quad \leq \quad c \cdot p_{M',\Lambda}(h) \tag{6.3}$$

This will imply $(\chi_R - 1)\cdot\varphi \in U(M,B(\mathbf{b}),\varepsilon)$ for $R \geq R_0$ as required.

Given $|\boldsymbol{\alpha}| \leq M$, we may expand

$$|D^{\boldsymbol{\alpha}}_\lambda \int_{\mathbb{R}^n} (\chi_R(x) - 1)\varphi(x,\lambda)h(x,\lambda)dx| \;=$$

$$|\int_{\mathbb{R}^n} \sum_{\boldsymbol{\beta}+\boldsymbol{\gamma}=\boldsymbol{\alpha}} c_{\boldsymbol{\beta}\boldsymbol{\gamma}}(\chi_R(x) - 1)D^{\boldsymbol{\beta}}_\lambda\varphi(x,\lambda)\cdot D^{\boldsymbol{\gamma}}_\lambda h(x,\lambda)dx| \tag{6.4}$$

for positive constants $c_{\boldsymbol{\beta}\boldsymbol{\gamma}}$ which are independent of φ and h. Let φ have compact support $K \subset \mathbb{R}^k_+$. By the uniformly moderate growth of φ, there are positive $C_{\boldsymbol{\beta}}$ and $m_{\boldsymbol{\beta}}$ such that

$$|D^{\boldsymbol{\beta}}_\lambda \varphi| \quad \leq \quad C_{\boldsymbol{\beta}}(1 + \|x\|^2)^{m_{\boldsymbol{\beta}}} \quad \text{for all } x \in \mathbb{R}^n \text{ and } \lambda \in K \tag{6.5}$$

Thus, (6.5) holds on $\mathbb{R}^n \times \mathbb{R}^k_+$. Hence,

$$|D^{\boldsymbol{\alpha}}_\lambda \int_{\mathbb{R}^n} (\chi_R - 1)\varphi\cdot h dx| \;\leq\; \sum_{\boldsymbol{\beta}+\boldsymbol{\gamma}=\boldsymbol{\alpha}} c_{\boldsymbol{\beta}\boldsymbol{\gamma}} \int_{\|x\|\geq R} C_{\boldsymbol{\beta}}(1 + \|x\|^2)^{m_{\boldsymbol{\beta}}} \cdot |D^{\boldsymbol{\gamma}}_\lambda h| dx \tag{6.6}$$

Let $M_1 = \max\{M, \max_{|\boldsymbol{\beta}|\leq M}\{m_{\boldsymbol{\beta}}\}\}$. Then, for $\lambda \in K_{M_1+n}$

$$(1 + \|x\|^2)^{m_{\boldsymbol{\beta}}} \cdot |D^{\boldsymbol{\gamma}}_\lambda h| \quad \leq \quad \frac{1}{(1 + \|x\|^2)^n} \cdot p_{M_1+n,\Lambda}(h) \tag{6.7}$$

Let $M' = M_1 + n$ and define

$$C_{M'} \quad = \quad \max_{|\boldsymbol{\alpha}|\leq M}\{\sum_{\boldsymbol{\beta}+\boldsymbol{\gamma}=\boldsymbol{\alpha}} c_{\boldsymbol{\beta}\boldsymbol{\gamma}} C_{\boldsymbol{\beta}}\} \tag{6.8}$$

and

$$\varepsilon(R) \quad = \quad \int_{\|x\|\geq R} \frac{1}{(1 + \|x\|^2)^n} dx \tag{6.9}$$

Then, by taking the supremum of (6.4) over $|\boldsymbol{\alpha}| \leq M$ and $\lambda \in K_{M'}$, and applying (6.5) through (6.9), we obtain

$$\|(\chi_R - 1)\cdot u_{\varphi\cdot}(h)\|_M \quad \leq \quad C_{M'} \cdot \varepsilon(R) \cdot p_{M',\Lambda}(h) \tag{6.10}$$

Since $\varepsilon(R) \to 0$ as $R \to \infty$, we may choose R_0 so that $C_{M'} \cdot \varepsilon(R) < \varepsilon/b_{M'}$ for $R \geq R_0$. Then, (6.10) yields (6.3). □

In §9 we shall give sufficient conditions that the Gaussian kernel and various kernels derived from it satisfy condition (A), so that for these kernels the subspaces we have mentioned do satisfy the conclusions of the theorems.

7. Discrete Functions and Measures

In what follows, we will repeatedly apply Theorem 6.3. Hence, we will frequently make the following "basic scale space assumption".

7.1 (Basic Scale Space Assumption). We assume G is a scale–based kernel satisfying property (A) and that $\mathcal{P}$ is a transversally defined scale–based geometric property defined via the closed Whitney stratified subset W, with strata $\{W_i\}$.

Next, we consider the class of discrete functions or measures, which form a quite specialized class of functions and measures. Because of their special form there is no reason to expect them to exhibit a property $\mathcal{P}$ generically on any compact subset of scale space. Nevertheless, we ask whether: i) any, ii) "many", or iii) "most" discrete functions or measures can exhibit the property $\mathcal{P}$ generically on compact subsets of scale space?

A discrete set of data can define a discrete function or discrete measure in several different ways. Usually, for functions this is done by defining for points of a fixed grid the function values. We take a slightly different approach because we want to obtain functions (and measures) defined on all of $\mathbb{R}^n$. For this we consider a *mesh* $\mathcal{B}$ which consists of a locally finite decomposition of $\mathbb{R}^n$ by regions whose closures are compact with piecewise smooth boundaries, and whose closures only intersect on their boundaries. We refer to the regions as *cells*. We remark at this early point that use the term "mesh" rather than "grid" to emphasize that we are considering functions taking constant values (or piecewise linear values) on the cells rather than just at the points of the grid.

We further suppose we are given a method for refining the mesh by a finer mesh of smaller subregions with the same properties and whose diameters on any compact set can be made arbitrarily small after a finite number of subdivisions. We let $\|\mathcal{B}\|$ denote the maximum diameter of all cells, which we assume is finite.

Example 7.2. A mesh is defined by a grid of cubes defined for $\mathbf{a} = (a_1, \dots, a_n)$, with $a_i \in \mathbb{Z}$, by

$$I_m(\mathbf{a}) = \{x = (x_1, \dots, x_n) \in \mathbb{R}^n : \frac{a_j}{2^m} \leq x_j < \frac{a_j + 1}{2^m}\} \quad \text{for fixed } m.$$

See figure 4. Refinements are by subdivision of the cubes, with m inceasing. Alternately, we can consider a triangulation by simplices, etc where the refinement is by barycentric subdivision.

Then, for us a *discrete function* will be one which is constant on each region of the grid (ultimately, we do not care which values are taken on the various boundary segments), and is zero except on finitely many cells of the mesh. By a *piecewise linear function*, we mean one which is piecewise linear on the closure of each cell and zero except on finitely many cells.

Given a finite set of data points $\mathcal{D} = \{x_i, i = 1, \dots, m\}$, we can associate positive Borel measures in two different ways. One is the associated *counting measure.* Given a Borel set U we let $\mu(U) = card\{x_i : x_i \in U\}$. Alternately, if we also have a mesh, we can define a "box counting measure". We let $c_j = card\{x_i : x_i \in B_j\}$ for each B_j in the mesh; and define $\mu(U) = \sum_j c_j m(U \cap B_j)$, where m denotes

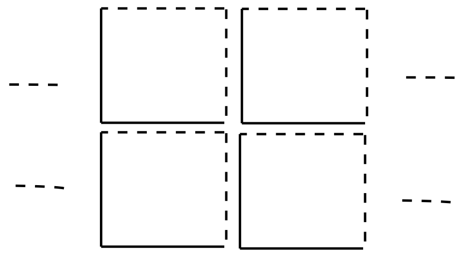

FIGURE 4. One choice of cells for a square mesh for $\mathbb{R}^2$

ordinary Lebesgue measure (see figure 5). By a *discrete simple measure* we shall mean a measure obtained by multiplying a box–counting measure by a constant c. Hence, in particular, we can obtain discrete simple probability measures.

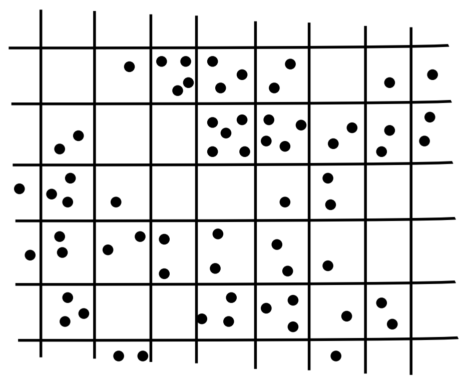

FIGURE 5. Discrete data set on a square mesh for $\mathbb{R}^2$ defines a discrete measure

Remark 7.3. For a fixed data set $\mathcal{D}$, we let μ_0 denote the counting measure and let μ_n denote the box–counting measure obtained using the n–th subdivision of $\mathcal{B}$. Then, it is fairly straightforward to see that $\mu_n \to \mu_0$ as $n \to \infty$ weakly as tempered distributions (in the sense that $\int h\mu_n \to \int h\mu_0$ for all $h \in \mathcal{S}_n$). In what follows, we concentrate on the discrete simple measures obtained from box–counting measures.

We consider discrete functions, piecewise linear functions, or discrete simple measures, and ask whether they can exhibit generic geometric properties in scale space. Despite their crude structure, not only do there exist discrete functions and measures which exhibit a property $\mathcal{P}$ generically on a given compact subset of scale space, but we next show the set of such discrete functions and measures is dense. First, we see as an almost immediate consequence of Theorem 6.4, that we can approximate functions by discrete or piecewise linear functions exhibiting the property $\mathcal{P}$ generically.

Corollary 7.4 (Density of Generic Discrete and Piecewise Linear Approximations)**.** *With the basic scale space assumption, let C be a compact subset of scale space. Given $f_0 \in L^p(\mathbb{R}^n)$ $(1 \leq p < \infty)$, $\varepsilon > 0$ and a mesh, there exists both a discrete function f_1 and a piecewise linear function f_2 (relative to a refinement of the mesh) which exhibit $\mathcal{P}$ generically on C and which approximate f_0 to within ε in $L^p(\mathbb{R}^n)$.*

Proof. The proof follows the lines of a more restrictive result given in [D5, Thm 3.1]. First, by Theorem 6.4, we may approximate f_0 by an $f \in L^p(\mathbb{R}^n)$ which exhibits $\mathcal{P}$ generically on C and is within $\varepsilon/2$ in the L^p–norm. As f exhibits $\mathcal{P}$ generically on C, again by 6.4 there is an $\varepsilon' > 0$, which we may also choose less than $\varepsilon/2$, such that if $f' \in L^p(\mathbb{R}^n)$ is within ε' of f then f' also exhibits $\mathcal{P}$ generically on C.

Then, by e.g. [Ru1, Thm. 3.14],the continuous functions with compact support are dense in $L^p(\mathbb{R}^n)$ for $1 \leq p < \infty$. Thus, we may first approximate f to within $\frac{\varepsilon'}{2}$ in $L^p(\mathbb{R}^n)$ by a continuous function f' with compact support B. Then, B is contained in a compact region Ω consisting of the closure of finite many cells from the mesh. Then, using the uniform continuity of f', by further subdividing Ω we may approximate f' to within $\frac{\varepsilon'}{2}$ in $L^p(\mathbb{R}^n)$ by either a discrete function f_1 or a piecewise linear function f_2 both with support in Ω. These f_1 and f_2 are within ε' of f and hence both exhibit $\mathcal{P}$ generically on C. In addition, they are within $\frac{\varepsilon}{2} + \varepsilon' < \varepsilon$ of f_0, establishing the result. □

Thus, given any compact subset of scale space, the discrete and piecewise linear functions exhibiting $\mathcal{P}$ generically on C is dense in any $L^p(\mathbb{R}^n)$ for $1 \leq p < \infty$. We next prove an analogous result for Borel measures.

Corollary 7.5 (Density of Generic Discrete Measures). *With the basic scale space assumption, suppose we are given a compact subset C of scale–space, $\varepsilon > 0$, and a mesh.*

1) If μ be a positive bounded regular Borel measure on $\mathbb{R}^n$ which is absolutely continuous with respect to Lebesgue measure, then there exists a discrete simple measure ν of compact support (relative to a refinement of the mesh) which exhibits $\mathcal{P}$ generically on C, and which approximates μ to within ε in the space of positive regular Borel measures.

2) Moreover, if the Whitney stratified set W defining $\mathcal{P}$ is invariant under multiplication by positive scalars, then we may also approximate any Borel regular probability measure μ , absolutely continuous with respect to Lebegue measure, by a discrete simple probability measure μ which exhibits $\mathcal{P}$ generically on C.

Proof. By the Radon–Nikodym theorem, we may write $\mu = f_0 dm$, where $f_0 \in L^1(\mathbb{R}^n)$ and dm denotes Lebesgue measure. Then, as μ is positive, we may arrange that $f_0 \geq 0$. By the preceding Corollary 7.4, We may approximate f_0 to within ε by a discrete function f_1 relative to a refinement of the mesh. Also, in using Theorem 6.4 in the proof of Corollary 7.4, we can arrange by the proof of Theorem 6.4 that $f_1 \geq 0$. Moreover, in the proof of Corollary 7.4 we may take f_1 to have nonnegative rational values. Thus, on any cell $f_1 = \frac{c_i}{M}$ for $c_i, M \in \mathbb{Z}$ and nonnegative, and M independent of the cell. Then, $f_1 dm = \frac{1}{M} f_2 dm$ where f_2 takes nonnegative integer values c_i on the cells. Finally, if we pick a data set consisting of c_i points in the i–th cell of this refinement, then $f_1 dm = \frac{1}{M} f_2 dm$ is a discrete simple measure which approximates μ to within ε and exhibits $\mathcal{P}$ generically on C.

If μ is a probability measure, then we first replace ε by $\varepsilon_1 = \frac{\varepsilon}{2}$. Then, there exists a discrete simple measure $f_1 dm = \frac{1}{M} f_2 dm$ which exhibits $\mathcal{P}$ generically on C and approximates μ to within ε_1. Let $c = \int f_1 dm$. Then, $\nu = \frac{1}{c} f_1 dm = \frac{1}{cM} f_2 dm$ is a box-counting probability measure. By the invariance of W under scalar multiplication, ν exhibits $\mathcal{P}$ generically on C. Third, it satisfies

$$\begin{aligned} \|f_0 - \frac{1}{c}f_1\|_{L^1} &\leq \|f_0 - f_1\|_{L^1} + |1 - \frac{1}{c}| \cdot \|f_1\|_{L^1} \\ &< \varepsilon_1 + |1 - \frac{1}{c}| \cdot c \quad \leq \quad 2\varepsilon_1 = \varepsilon \end{aligned} \tag{7.1}$$

Thus, ν is the desired approximation. □

Third, we ask whether "most"discrete functions or measures exhibit $\mathcal{P}$ generically on a given compact subset of scale space C. For a given mesh together with all of its refinements, we can give one form of answer to this question.

Consider any discrete or piecewise linear function f_0 defined on a compact region Ω, which is a finite union of cells from the mesh. Also, we suppose that such f_0 take (possibly a discrete set of) values in a finite interval $[-a, a]$. Then, we show that given $\varepsilon > 0$, by making a single refinement of the mesh and a single enlargement of the region Ω, we may any approximate f_0 to within ε by a discrete (resp. piecewise linear) function exhibiting the property $\mathcal{P}$ generically on C.

Theorem 7.6 (Genericity for Discrete and Piecewise Linear Functions)**.** *With the basic scale space assumption, let C be a compact subset in scale space, and let $\varepsilon > 0$. Given a region Ω contained in a compact subset D (and range of values $[-a, a]$), then there is: a refinement $\mathcal{B}'$ of the mesh, an enlarged region Ω' consisting of the closure of a union of finitely many cells of $\mathcal{B}'$, and a refined set of function values (if finite) such that: every discrete (resp. piecewise linear) function f on Ω (extended to be zero outside Ω) can be approximated by another discrete (resp. piecewise linear) function f_1 on Ω' which exhibits the property $\mathcal{P}$ generically on C. The approximation is within ε in the L^p–norm ($1 \leq p < \infty$) and within ε in the (essential) sup–norm on D.*

Proof. We use condition (A) for G to adapt to our general situation the argument given for a special case in [D5, Thm. 2]. We give the argument for discrete functions with that for piecewise functions virtually identical

First, we can find a finite number of discrete functions $\{f_i\}$ on Ω such that any discrete function on Ω (with values in the range $[-a, a]$) can be approximated to within $\frac{\varepsilon}{2}$ by some f_i. Hence, if we can approximate each f_i by a g_i on a larger region Ω_i to within $\frac{\varepsilon}{2}$ then this approximation also works on the larger $\Omega' = \cup\Omega_i$, so we can approximate any discrete f to within ε on Ω'. Hence, it is enough to approximate a single discrete f_0. We extend this function to be zero outside Ω.

The genericity of the property $\mathcal{P}$ is defined by transversality to a closed Whitney stratified set $W \subset \mathcal{H}^\ell$. Let $F = f_0 * G$. By Proposition 6.6 and the proof of Theorem 6.4, there is a finite dimensional subspace $\mathcal{V}$ spanned by smooth nonnegative functions of compact support $\{\varphi_1, \ldots, \varphi_k\}$ such that the map $f_0 + \mathcal{V} \to \mathcal{H}^\ell_{(x_0,\lambda_0)}$ defined by $f \mapsto j^\ell(f * G)(x_0, \lambda_0)$ is a submersion for all $(x_0, \lambda_0) \in C$. Then, by the parametrized transversality theorem, there is a subset $S \subset \mathcal{V}$ whose complement has measure zero such that $j^\ell(F + f * G)$ is transverse to W on C for all $f \in S$. As the basis for $\mathcal{V}$ has compact support, there is an R_0 so that $\mathrm{supp}(f) \subset B_{R_0}(0)$ for all $f \in \mathcal{V}$. By enlarging R_0 we can assume $B \subset B_{R_0}(0)$. We may choose a compact neighborhood N of 0 in $\mathcal{V}$ such that if $f \in N$, then $|f| < \frac{\varepsilon}{2}$ on $\bar{B}_{R_0}(0)$.

We pick $f' \in S \cap N$ so $j^\ell(F + f' * G)$ is transverse to W on C. Thus, by Theorem 6.4, there is an $\varepsilon' > 0$ so that if g_0 is L^p close to $f_0 + f'$ then $g_0 * G$ exhibits $\mathcal{P}$ generically on C.

Second, we cover $\bar{B}_{R_0}(0)$ by a finite region Ω'_0 which is a union of closures of cells of the mesh. Then, by refining the mesh, we may construct a discrete approximation g_1 to g_0 on Ω'_0 that is within ε_1 in both the (essential) sup–norm and $L^p(\Omega'_0)$–norm. Then, $f_1 = f_0 + g_1$ is a discrete function on Ω'_0 which approximates f_0 to within ε on $B_R(0)$

$$|f_1 - f_0| = |g_1| \leq |g_0 - g_1| + |g_0| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon; \tag{7.2}$$

and approximates $f_0 + f'$ to within ε' in the $L^p(\mathbb{R}^n)$–norm

$$\begin{aligned} \|f_1 - (f_0 + g_0)\|_{L^p} &= \|g_1 - g_0\|_{L^p(\mathbb{R}^n)} \\ &\leq \|g_1 - g_0\|_{L^p(\Omega')} \ < \ \varepsilon' \end{aligned} \tag{7.3}$$

Hence, $f_1 * G$ exhibits property $\mathcal{P}$ generically on C. □

Finally, as a complement to Theorem 7.6, we see that Corollary 7.4 does provide a sup–norm estimate for the retention of a generic property under small perturbations of a discrete function.

Corollary 7.7. *With the basic scale space assumption, suppose the discrete function f_0 defined on Ω exhibits property $\mathcal{P}$ generically on a compact subset C in scale space. Then there is an $\varepsilon > 0$ so that any other discrete function on Ω (allowing further subdivision) which is within ε of f_0 in the (essential) sup–norm exhibits $\mathcal{P}$ generically on C.*

Proof. This is an immediate consequence of Theorem 6.4, which provides an ε' so that if g is within ε' of f_0 in the $L^p(\mathbb{R}^n)$–norm, then g exhibits $\mathcal{P}$ generically on C. However, for any discrete function g, if $f_0 - g$ has sufficiently small (essential) sup–norm ε on the compact Ω, then g is ε' close to f_0 in the $L^p(\mathbb{R}^n)$–norm. Hence, the result follows. □

Refinements of Discrete Simple Measures. The method of proof immediately translates over to the case of discrete simple measures. We define the notion of a refinement of a discrete simple measure. Given a box–counting measure μ defined from a data set $\mathcal{D} = \{x_1, \dots, x_m\}$ and a mesh $\mathcal{B}$, we consider the discrete simple measure $c\mu$. We suppose that each cell of $\mathcal{B}$ has the same volume, and for each refinement $\mathcal{B}'$ of $\mathcal{B}$, there is an integer q so that each cell is decomposed into q cells each of $\frac{1}{q}$ the volume. Then, after k refinements there will be q^k cells from each cell, each of $\frac{1}{q^k}$-th the volume. We replace $\mathcal{D}$ by a data set consisting of $m \cdot q^k$ data points such that if $\mathcal{D}$ has c_j points in cell I_j, then c_j points are placed into each cell of the refinement of I_j, so we obtain a box–counting measure μ' but we will multiply by the constant $\frac{c}{q^k}$. Thus, we refer to $\frac{c}{q^k}\mu'$ as a *refinement of the simple discrete measure* $c\mu$. Then, for any set formed from a union of original cells of $\mathcal{B}$, both measures give the same value.

If $c'\nu$ is a discrete simple measure based on the k-th refinement $\mathcal{B}'$ of $\mathcal{B}$, then we can compare $c'\nu$ and $c\mu$ by

$$\|c\mu - c'\nu\|_{\mathcal{SBM}} \quad = \quad \frac{1}{q^k} \cdot \sum |c \cdot c_j - c' \cdot c'_j|$$

summed over all cells of $\mathcal{B}'$.

Then, we claim we can also uniformly approximate discrete simple measures by others which exhibit $\mathcal{P}$ generically on C. Consider a region Ω contained in a compact subset D, and formed from cells of a mesh $\mathcal{B}$. We consider discrete simple measures μ with support in Ω and values $\nu(I_j) \in [0, a]$ for all cells I_j in $\mathcal{B}$.

Theorem 7.8 (Genericity for Discrete Simple Measures). *With the basic scale space assumption, let C be a compact subset in scale space, and let $\varepsilon > 0$. Then, there is: a refinement $\mathcal{B}'$ of the mesh, an enlarged region Ω' (consisting of the closure of a union of finitely many cells of $\mathcal{B}'$), such that: every discrete simple measure μ defined as above can be approximated by another discrete simple measure ν defined using $\mathcal{B}'$, with support in Ω', and which exhibits the property $\mathcal{P}$ generically on C. The approximation is within ε using $\|\cdot\|_{SB\mathcal{M}}$ such that*

$$\max_j |c \cdot c_j - c' \cdot c'_j| < \varepsilon \quad \textit{on } D$$

Proof. A discrete simple measure $\mu = f_0 \cdot dm$, where dm is Lebesgue measure and f_0 is a discrete function with value $c \cdot c_j$ on the cell I_j. Hence, we can apply the proof of Theorem 7.6, modifying it as in the proof of Corollary 7.5 to obtain an approximation f_1 which is nonnegative and with rational values so that $f_1 \cdot dm$ is the desired discrete simple measure. □

Remark 7.9. We ask in Theorems 7.6 and 7.8, how much larger than Ω does Ω' have to be?

We know by Proposition 6.6 there are smooth nonnegative functions $\{\varphi_1, \ldots, \varphi_m\}$ with support in $\bar{B}_R(0)$ (for some $R > 0$), such that $\{j^\ell(\varphi_i \cdot *G)(x_0, \lambda_0)\}$, for $i = 1, \ldots, m$ span $\mathcal{H}^\ell_{G\,(x_0,\lambda_0)}$. For $(x_0, \lambda_0) = (0, 1)$, where $1 \in \mathbb{R}^k_+$ denotes $(1, 1, \ldots, 1)$, let R_ω denote the infimum of such R. We can express Ω' in terms of Ω and R_ω.

For any compact subset $C_1 \subset \mathbb{R}^n$, there is a finite dimension subspace $\mathcal{V}$ spanned by smooth nonnegative functions with support in $C_1 + \bar{B}_R(0)$ such that the map $\mathcal{V} \to \mathcal{H}^\ell_{(x,1)}$ sending $f \mapsto j^\ell(f * G)(x, 1)$ is a submersion for all $x \in C_1$. Then, for any scale threshold $S = \prod[\varepsilon_i, L_i]$, we can apply the scaling action to obtain functions with support in $C_2 = \cup_{\lambda \in S} \lambda \cdot (C_1 + \bar{B}_R(0))$. This is an allowable Ω' for $\Omega = C_1$.

Part 3. Classical, Scale–based, and Multi-feature Geometries

In Parts 1 and 2, we established theorems which guarantee the genericity and stability of scale–based generic properties $\mathcal{P}$ defined by transversality to closed Whitney stratified subset W of $\mathcal{H}_G^\ell$, where G is a scale–based kernel satisfying condition (A). These properties were shown to furthermore hold for certain subspaces of tempered distributions $\mathcal{T}$ which satisfy condition (B), as well as for generic discrete or piecewise linear functions and discrete measures. To apply these results, we must identify both interesting classes of scale–based kernels which satisfy condition (A), and scale–based properties $\mathcal{P}$.

In Part 3, we provide answers to these questions. At the same time, we expand the original goals of Parts 1 and 2. We provide sufficient conditions that associated discrimination functions, measures, or distributions which determine the amount of a discriminated property in a given region also exhibit generic scale–based properties. We do this for both single and multi–feature discrimination functions and measures.

In §8, we determine when a "classical geometric property"has a scale–based version, and whether this scale–based version has identical properties to those for the classical case. We identify two distinct methods for doing this. One case with scale–parameters treated as distinguished parameters will be treated in §14.

We introduce in §9 a large class of "extended Gaussian kernels"which expand the class of Gaussian kernels by allowing anisotropy, derivatives, and multiple scaling parameters. We establish in Proposition 9.4 that such kernels satisfy condition (A). At the same time, we obtain sufficient conditions (Corollaries 9.6, 9.7, and 9.9) for explicitly verifying the conditions from §8 that a geometric property will have a scale–based version with the same generic properties.

In §10, we consider associated discrimination functions, measures, and distributions. The space of such distributions which arise by applying a filter, a differential operator, or a statistical construction yielding a measure will typically form a proper subspace of tempered distributions. For an operator Ψ, we provide sufficient conditions that the image subspace is sufficiently robust so that such discrimination distributions will exhibit scale–based geometric properties with the same generic properties as all generic distributions. As the discrimination process usually involves a specific scale, genericity results can only be expected for regions that are at a sufficiently larger scale; and hence we restrict to a specific compact region C of scale space. We give sufficient conditions that the discrimination distributions have generic geometric structures on C for a dense open subset of the space of initial distributions (Proposition 10.2). In the case of a linear operation Ψ, this is concretely expressed (in Theorem 10.5) in terms of the scale–based kernel satisfying condition (A), and Ψ satisfying a condition (B_C) (as a substitute for condition (B) for subspaces considered in Part 2). We deduce sufficient conditions in three situations: partial differential operators (Corollary 10.6), a linear operation which can approximate "block functions"(Corollary 10.7), and a simple form of texture discrimination (Corollary 10.9).

In §11 we turn our attention to scale–based properties of vector-valued tempered distributions with values in $\mathbb{R}^p$, which are given by p-tuples of (uniformly) tempered distributions. We allow different scale-based kernels and scaling actions for each component, so the action of the Poincaré scaling group PS does not act via a single geometric action on $\mathbb{R}^n \times \mathbb{R}^k_+$, but rather componentwise. Nonetheless, we show that the convolution jet space has analogous properties to those for single distributions. We deduce that scale–based geometric properties of vector–valued tempered distributions satisfy analogues of the theorems for genericity and stability from §§5, 6, and 7.

Lastly, in §12 we combine the methods of §10 and §11 to analyze multi–feature geometry. If we associate to an initial distribution, multiple distributions defined using a number of different discrimination criteria, we obtain a vector-valued tempered distribution. We give a criteria in terms of "generic independence of discrimination features"for genericity of multi–feature geometry for an open dense set of initial distributions (Proposition 12.2). In particular, we deduce that scale–based geometric properties of these vector-valued distributions will interact generically on C. Using this, we give a sufficient condition, in terms of independence conditions of "mask functions"that a simple form of multi-texture discrimination will have generic scale–based geometric properties (Corollary12.5).

8. Scale-Based Properties derived from Classical Geometric Properties

Suppose we are given a scale–based kernel G which satisfies condition (A). Many of the geometric properties we considered in §1 are "classical geometric properties"defined for smooth (or at least highly differentiable) functions. We describe how we can define associated scale–based properties and to what extent they have the same generic properties as the original classical properties.

Suppose then that the property $\mathcal{P}'$ is a geometric property defined for smooth functions and given by algebraic equalities and inequalities involving the partial derivatives of the functions. These algebraic conditions define a semialgebraic subset whose closure $W' \subset J^\ell(\mathbb{R}^n, \mathbb{R})$ again is a semialgebraic subset. By a theorem of Lojasiewicz [Lo], W' has a Whitney stratification whose strata are again semialgebraic subsets. Thus, the property $\mathcal{P}'$ is defined by transversality to the closed Whitney stratified subset $W' \subset J^\ell(\mathbb{R}^n, \mathbb{R})$, which we henceforth suppose is a closed semialgebraic subset with semialgebraic strata. The individual strata define the various ways that the property $\mathcal{P}'$ can be exhibited.

We identify two approaches for defining an associated scale-based property $\mathcal{P}$.

(1) First, if $\mathcal{P}'$ is defined for smooth functions on $\mathbb{R}^n \times \mathbb{R}^k$, then we can define a corresponding property for convolved functions on scale space $\mathbb{R}^n \times \mathbb{R}^k_+$, without regarding the scale parameters λ as distinguished. This describes how geometric properties propagate throughout physical–scale space.

(2) Second, if instead $\mathcal{P}'$ is defined for smooth functions on $\mathbb{R}^n$, then we can retain the distinguish role of λ, We define a scale–based property for convolved functions on scale space, viewed as unfoldings with the scale parameters λ as unfolding parameters. This allows us to consider how geometric properties of physical space change as scale parameters vary.

In both cases, we ask how exactly the scale-based properties can be determined and how they relate to the original properties $\mathcal{P}'$.

The second case is actually much more subtle mathematically than the first for we are viewing the convolved function $f = u * G$ as an unfolding on the scale parameters λ. Then, we first must determine whether for fixed parameter values, f stably exhibits the property $\mathcal{P}'$. The generic behavior which then occurs typically corresponds to generic properties of a version of $\mathcal{P}$ where the λ are not treated as distinguished parameters. However, there will also be parameters values where $j^\ell_x(f)$ fails to be transverse to W'. Hence, we must also determine how transversality can generically fail in families producing generic transitions. This requires us to consider an equivalence capturing failure of transversality and using the basic theorems of singularity theory as they apply for this equivalence. We postpone considering this case until Part 4, where we more generally allow parameters in the families of distributions, as scaling parameters, and/or as external parameters appearing in the kernel.

Returning to the first case, we provide an answer in terms of the transversality behavior of the stratified set W' relative to $\mathcal{H}^\ell_G$. We first define an associated closed Whitney stratified set in $\mathcal{H}^\ell_G$. Let

$$W \quad \stackrel{\text{def}}{=} \quad W' \cap \mathcal{H}^\ell_G.$$

with $W' \subset J^\ell(\mathbb{R}^n \times \mathbb{R}^k_+, \mathbb{R})$. Then, by Proposition 4.9, $\mathcal{H}^\ell_G$ is a semialgebraic submanifold of $J^\ell(\mathbb{R}^n \times \mathbb{R}^k_+, \mathbb{R})$. Hence, W, being the intersection of semialgebraic subsets, is again a semialgebraic and is closed in $\mathcal{H}^\ell_G$ as W' is closed. We may again apply the theorem of Lojasiewicz [Lo], and conclude that W again has a Whitney stratification with semialgebraic strata.

Proposition 8.1. *In the preceding situation, suppose W' is transverse to $\mathcal{H}^\ell_G$, i.e. the strata $\{W'_i\}$ are transverse. Then, for convolutions $u * G$, the scale–based geometric structure defined by transversality to W in $\mathcal{H}^\ell_G$ has identical generic properties as those exhibited by smooth functions on $\mathbb{R}^n \times \mathbb{R}^k$ which generically exhibit $\mathcal{P}'$.*

If W' is not transverse to $\mathcal{H}^\ell_G$, then W is still a closed Whitney stratified set and still yields a transversally defined property $\mathcal{P}$ exhibited generically and stably on compact subsets of scale space. However, the geometric properties and structure associated to $\mathcal{P}$ will differ from that for $\mathcal{P}'$; although they are determined by the local structure of W.

Proof. Let $f = u * G$. In the case W' is transverse to $\mathcal{H}^\ell_G$, we have the diagram 8.1 which is a fiber square.

$$\begin{array}{ccccc} \mathbb{R}^n \times \mathbb{R}^k_+ & \xrightarrow{j^\ell(f)} & \mathcal{H}^\ell_G & \xrightarrow{i} & J^\ell(\mathbb{R}^n \times \mathbb{R}^k_+, \mathbb{R}) \\ \uparrow & & \uparrow & & \uparrow \\ j^\ell(f)^{-1}(W) & \longrightarrow & W & \xrightarrow{i'} & W' \end{array} \tag{8.1}$$

Hence, $j^\ell(f)$ is transverse to W relative to $\mathcal{H}^\ell_G$ iff $j^\ell(f)$, viewed as a mapping to $J^\ell(\mathbb{R}^n \times \mathbb{R}^k_+, \mathbb{R})$, is transverse to W'. Then, $j^\ell(f)^{-1}(W) = j^\ell(f)^{-1}(W')$. Hence the geometric structures agree and the geometric properties possessed by $u * G$ are those possessed by maps generically exhibiting $\mathcal{P}'$.

In the case W' is not transverse to $\mathcal{H}^\ell_G$, then we can no longer assure a relation between transversality of $j^\ell(f)$ to W in $\mathcal{H}^\ell_G$ and transversality to W'. However, for convolutions $f = u * G$, we do still obtain that transversality of $j^\ell(f)$ to W defines a modified version $\mathcal{P}$ of the property $\mathcal{P}'$. Property $\mathcal{P}$ will be exhibited generically on any compact subset C of scale subspace with associated geometric structure given by $W(f) = j^\ell(f)^{-1}(W)$. By Theorems 5.8, 5.10, and 5.11 we obtain the genericity, local structure, and stability for a corresponding property $\mathcal{P}$ defined by W. $\square$

We briefly indicate to what extent the preceding proposition has been or can be applied to the classical geometric properties identified in §1: 1) generic differential geometry, 2) edge–based geometry, 3) level set geometry, and 4) medial geometry. The principal application has been for medial geometry; however, we also indicate the relevance of Proposition 8.1 for application of the other geometric properties. To consider them we must have them defined for $\mathbb{R}^n \times \mathbb{R}^k$. For example, functions obtained by convolving distributions on $\mathbb{R}^2$ with a kernel with a single scale parameter such as a simple Gaussian, medial or edge kernel as in Example 4.2, will be defined on $2+1$–dimensional scale space.

Example 8.2 (Generic Differential Geometry of Surfaces)**.** As indicated in §1, there has been considerable work on the generic structure of ridges, crests, parabolic curves on surfaces and their relation with the geometry of surfaces. Their generic properties are given by semialgebraic conditions which describe when certain associated geometrically defined functions exhibit specific singularities. Hence, there is a closed Whitney stratified set which transversally defines the generic properties for these sets. As of now, the relation with scale has not been explicitly investigated. For a surface associated to a convolved function, such as a level surface (see Example 8.3) or surface of extremal slope (see Example 8.4), the corresponding curves will have generic properties given by transversality to a corresponding closed Whitney stratified set.

Example 8.3 (Generic Edge–Based Geometry)**.** The generic edge–based geometry of a function on $\mathbb{R}^2$ based on the Canny edge has been determined by Rieger [Ri1]. For higher dimensions, he has determined some of the generic edge–based properties in [Ri4]. The conditions are given by specific algebraic conditions on derivatives. Hence, there is a closed Whitney stratified set describing the generic properties. It has not been yet determined how this behaves relative to any convolution jet space (but see §14).

Example 8.4 (Generic Level Set Geometry)**.** Geometric properties for level sets of functions on $\mathbb{R}^2$. have been identified by [Gau] and the generic properties are given in [Ri3]. They have not been fully determined for functions on higher dimensional spaces, although the results of [BGT] apply off the critical set. The associated generic medial geometry for the collection of level sets can be described through generic properties of an associated "shape map"[D2, §6]. Also, Zakalyukin and Goryunov have obtained generic transitions of medial axis for level surfaces in [GZ]. The behavior relative to scale space has not yet been determined.

The strongest results for convolved functions on scale space have been obtained for medial geometry using relative critical sets. As mentioned in §1, medial properties of convolved functions can be partially understood via "height ridges"of the associated medial function (i.e. cores). The generic properties of these are understood as part of the full relative critical set.

Example 8.5 (Medial Geometry via Relative Critical Sets)**.** One approach to the medial geometry of functions is via the relative critical set [D4], Miller [Mi], and Keller [Ke]. Suppose $f : U \to \mathbb{R}$ is smooth with U an open subset of $\mathbb{R}^n$. At a point $x \in \mathbb{R}^n$, we let $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ denote the eigenvalues of $H(f)(x)$, the Hessian of f at x. Let $\{e_1, \dots, e_n\}$ denote associated unit eigenvalues. If the λ_i are distinct then for a subset $I \subset \{1, \dots, n\}$, we let $C_I(f)$ denote the set of points x where $\nabla f(x)$ is orthogonal to e_i iff $i \in I$. In [D4] and [Mi] were constructed closed Whitney stratified subsets Γ of jet space with the property that if $j^\ell(f)$ is transverse to Γ then the $\{C_I(f)\}$ are smooth strata forming a normal crossings divisor off the partial umbilic set where the λ_i are not all distinct. This divisor is further stratified by the "singular Hessian hypersurface" which generically intersects the strata transversally. Moreover, at least in dimensions ≤ 4, this can be fully extended in an appropriate fashion to the partial umbilic set, so we still have a normal crossings divisor. Certain smooth strata of this divisor form the height ridge set introduced by Pizer-Eberly [PE]. In addition there are strata corresponding to "valleys", as well as certain connecting strata which join and intersect ridges and valleys in precisely specified ways (see also [D7]).

In dimensions > 4, the explicit geometric structure needed to fully determine the relative critical sets on the partial umbilic points has not yet been fully worked out. However, there is partial information given by Miller[Mi] who has determined in terms of the codimension of the strata of the space of symmetric matrices, in which dimensions various partial umbilic behavior will generically first appear.

For the purpose of establishing the stability and genericity without regard to dimension, there is an alternate abstract construction of the closed semi–algebraic Whitney stratified set in dimensions > 4 which is valid for all points, although it does not yield the specific structure on partial umbilic points (see [D7]). Thus, we may apply Theorem 6.4 and Proposition 8.1 to conclude that for a given scale–based kernel G, there is an associated generic relative critical set structure for G–convolved functions. Although we do not know for specific G whether, the generic structure is identical to that for generic smooth functions.

Corollary 8.6. *Given any scale–based kernel G satisfying condition (A) and subspace of tempered distributions $\mathcal{T}$ satisfying condition (B), there is an associated generic relative critical set structure which is exhibited generically and stably for G-convolved distributions from $\mathcal{T}$ on any compact subset of scale space.*

What has to be determined in applying Proposition 8.1 is whether the relative critical set structure for convolved functions has the same generic properties as those for generic smooth functions.

Example 8.7. We consider the case of the "Mexican hat" medial kernel M given in Example 4.4. Convolution with the medial kernel M associates to tempered distributions "medial functions on scale space". In [D4] and [Mi], it is proven that in dimensions $n = 2, 3$ (i.e. $2 + 1$ and $3 + 1$ scale–space), Γ defined in Example 8.5 is transverse to $\mathcal{H}^\ell_M$ (see also §9). Hence, on compact subsets of scale space, these medial functions have relative critical sets exhibiting the same generic properties as smooth functions. In particular, we deduce the explicit generic properties of height ridges of the medial functions (called "cores" by Eberly- Pizer et al. [PE]). For 2 and 3–dimensional grayscale images, the generic properties are given in [D3], [D4], and [Mi], see also [Fu].

9. Extended Gaussian Scale–Based Properties

We described in §8 two methods for defining scale–based geometric properties from classical ones, one of which does not treat λ as distinguished parameter(s). In this case, Proposition 8.1 provides a method for applying the results from Parts 1 and 2 to obtain genericity and stability. We must: i) identify kernels G which satisfy condition (A), and ii) give specific methods for identifying $\mathcal{H}^\ell_G$ or $\mathcal{H}^\ell_{\Lambda,\, G}$, to determine when a closed Whitney stratified set W defining a classical geometric property $\mathcal{P}$ will be transverse to $\mathcal{H}^{\ell}_{G}$ or $\mathcal{H}^{\ell}_{\Lambda,\,G}$. In this section, we give answers to these questions for a general class of "extended Gaussian kernels". This is a natural class of scale–based kernels defined from standard Gaussian kernels by allowing anisotropy, multiple scaling parameters, and the application of certain differential operators. We give a polynomial basis for $\mathcal{H}^{\ell}_{G\,(x_0,\lambda_0)}$ or $\mathcal{H}^{\ell}_{\Lambda,\,G\,(x_0,\lambda_0)}$, and establish that any such kernel satisfies condition (A) (Proposition 9.4). Second, using the polynomial basis, we give sufficient conditions for transversality to $\mathcal{H}^{\ell}_{G}$ or $\mathcal{H}^{\ell}_{\Lambda,\,G}$ (Proposition 9.8 and Corollary 9.9).

A Class of "Extended Gaussian Kernels". We listed in Examples 4.2, 4.4, and 4.6 several types of kernels derived from Gaussian kernels. These include: i) the standard and anisotropic Gaussian kernels; ii) the multi-Gaussian kernel defined for the decomposition $\mathbb{R}^n = \oplus_{i=1}^{r}\mathbb{R}^{n_i}$; and iii) associated kernels such as the medial kernel $M(x,t)$ and the edge- based kernels; and iv) pullbacks of these kernels by homomorphisms.

These belong to a general class defined as follows. Let $n = \sum_{i=1}^{r} n_i$ for positive integers n_i, and consider the decomposition $\mathbb{R}^n = \oplus_{i=1}^{r}\mathbb{R}^{n_i}$ where $x = (x^{(1)},\dots,x^{(r)})$ with $x^{(i)} = (x^{(i)}_1,\dots,x^{(i)}_{n_i})$. We have an action of $\mathbb{R}^r_+$ (with coordinates $(\sigma'_1,\dots,\sigma'_r)$) on $\mathbb{R}^n$ where σ'_i acts trivially on $x^{(j)}$ for $j \neq i$ and by scalar multiplication by σ'_i on $x^{(i)}$, so its weight matrix has i–th row consisting of 0's except for 1 in positions corresponding to $x^{(i)}$.

Let $Q_i(\sigma'_i, D_{x^{(i)}}, D_{\sigma'_i})$ denote a linear differential operator with polynomial coefficients in σ'_i. We assign weights

$$\mathrm{wt}\,(\sigma'_i) = 1, \quad \mathrm{wt}\,(x_i) = 1, \quad wt(\frac{\partial}{\partial\sigma'_i}) = -1, \quad \text{and } \mathrm{wt}\,(\frac{\partial}{\partial x_i}) = -1 \quad \text{for all i.}$$

We say that Q_i is *weighted homogeneous of weighted degree* m if it has the form

$$Q \quad = \quad \sum c_{\boldsymbol{\alpha}\beta i}\cdot {\sigma'_i}^{m_{\boldsymbol{\alpha}\beta i}}\cdot D^{\boldsymbol{\alpha}}_x D^{\beta}_{\sigma'_i},$$

where $m_{\boldsymbol{\alpha}\beta i} = m + \beta + |\boldsymbol{\alpha}|$, and $c_{\boldsymbol{\alpha}\boldsymbol{\beta}i}$ are constants.

Definition 9.1. A kernel K is an *extended Gaussian kernel* if there are: i) a decomposition of $\mathbb{R}^n$ as above; ii) (regular or) anisotropic Gaussian kernels $K(x^{(i)},\sigma'_i) = K_{\frac{1}{2}{\sigma'_i}^2}(x^{(i)})$ for distinct scaling parameters $\sigma' = (\sigma'_1,\dots,\sigma'_r)$; iii) weighted homogeneous differential operators $Q_i((\sigma'_i, D_{x^{(i)}}, D_{\sigma'_i})$ of weighted degrees m_i, $i = 1,\dots,r$; and iv) a homomorphism $\sigma' = \rho'(\sigma) : \mathbb{R}^k_+ \to \mathbb{R}^r_+$ of the form (4.4) such that

$$K(x,\sigma) \quad = \quad \rho'^{\,*}\left(\prod Q_i(K(x^{(i)},\sigma'_i))\right)$$

Such a kernel allows anisotropy, independent scales σ_i, measurement of distinct geometric properties on the distinct subspaces $\mathbb{R}^{n_i}$, as well as mixing of scales. We note that by Lemma 4.3, as (anisotropic) Gaussian kernels are scale–invariant, such an extended Gaussian kernel is scale–based with scale weight $\mathbf{m} = (m_1,\dots,m_r)$. Hence, if we multiply the extended Gaussian kernel by the monomial $\sigma^{-\mathbf{m}} = \prod \sigma_i^{-m_i}$, it becomes scale–invariant.

We next answer when such an extended Gaussian kernel satisfies condition (A), and at the same time explicitly determine $\mathcal{H}^{\ell}_{G}$.

Polynomial Basis for $\mathcal{H}^{\ell}_{G}$. First, for the standard Gaussian kernel $G = K_t(x)$, a polynomial basis for $\mathcal{H}^{\ell}_{G\,(x_0,t_0)}$ for arbitrary (x_0, t_0) can be given using an explicit weighted homogeneous basis for solutions to the heat equation see [Fo] or e.g. [D1, §3]. We denote such a basis for the solutions at 0 of weighted degree m by $\{p_j(x,t)\}$. If we take the union over weights $\leq 2\ell$, then they span a subspace which is mapped by the jet extension map at 0 onto the ℓ-jet space at 0 of solutions to the heat equation. The set of solutions to the heat equation are preserved under translation. Hence, if we translate these polynomials to a point (x_0, t_0), obtaining $\{p_j(x - x_0, t - t_0)\}$, then they again map to a basis for the ℓ-jet space at (x_0, t_0) of solutions to the heat equation. For a solution $p(x,t)$ translated to (x_0, t_0), we substitute $t = \frac{1}{2}\sigma^2$ (and $t_0 = \frac{1}{2}\sigma_0^2$) and expand about the point (x_0, σ_0) in terms of local coordinates $(\bar{x}, \bar{\sigma}) = (x - x_0, \sigma - \sigma_0)$ and obtain $p(\bar{x}, \sigma_0\sigma + \frac{1}{2}\bar{\sigma}^2)$. Then, applying this procedure to $\{p_j(x,t)\}$, and mapping to the jet space at (x_0, σ_0), we obtain a spanning set for $\mathcal{H}^{\ell}_{K\,(x_0,\sigma_0)}$ (and weighted homogeneous solutions of weight $> 2\ell$ will map to 0). Hence, a subset will map to a basis.

Second, consider an anisotropic Gaussian kernel

$$K_t(x) = \frac{1}{(4\pi t)^{\frac{n}{2}}}(\det(A))^{\frac{1}{2}} \exp(\frac{-<x, Ax>}{4t})$$

for a symmetric positive definite matrix A. Let $A = A'^2$ for a symmetric positive definite matrix A'. Then, for a tempered distribution u, $u * K_t$ is a solution to the anisotropic heat equation

$$\frac{\partial f}{\partial t} = Tr(A^{-1}H(f)) \qquad \text{where } H(f) \text{ is the Hessian of } f. \tag{9.1}$$

This equation is weighted homogeneous for weights $\mathrm{wt}\,(x_i) = 1$, all i, and $\mathrm{wt}\,(t) = 2$. Such solutions can be obtained from solutions f_1 to the usual heat equation via $f(x) = f_1(A'x)$. Hence, we obtain a basis for solutions at 0 of weight m by using the basis for the usual heat equation $\{p_j(A'x,t)\}$. We can repeat the preceding argument, taking the union over $m \leq 2\ell$, translating to an arbitrary point (x_0, t_0), substituting $t = \frac{1}{2}\sigma^2$, and expanding about (x_0, σ_0) to obtain a basis for $\mathcal{H}^{\ell}_{K\,(x_0,\sigma_0)}$.

Third, consider a multi-Gaussian kernel

$$K(x,\sigma) = \prod_{i=1}^{r} K_i(x^{(i)}, \sigma_i) \qquad \text{where } K_i(x^{(i)}, \sigma_i) = K_{i\frac{1}{2}\sigma_i^2}(x^{(i)}, t_i)$$

where $K_{it}(x)$ is a (regular or anisotropic) Gaussian kernel. We observe that if $L_i = \dfrac{\partial}{\partial t_i} - \Delta_i$ is the anisotropic heat operator in $(x^{(i)}, t_i)$ as given in (9.1), then $f = u * K(x,\sigma)$ is a solution to the heat equations

$$L_i(f) \;=\; \frac{\partial f}{\partial t_i} - \Delta_i(f) \;=\; 0 \qquad \text{for} \quad i = 1, \dots, r \tag{9.2}$$

Conversely, as the equations (9.2) involve independent coordinates, we claim that $\mathcal{H}^{\ell}_{K_i\,(x_0^{(i)},\sigma_{i0})}$ has a weighted polynomial basis obtained as products of polynomial solutions to the separate heat equations. Specifically, the operators L_i are weighted homogeneous of weighted degree -2. Hence each preserves the weight decomposition on polynomials in (x,t). We let $\mathcal{H}(\ell)$ denote the subspace of polynomials of weight ℓ which are solutions to the equations (9.2). It decomposes

$$\mathcal{H}(\ell) = \oplus_{|\mathbf{k}|=\ell}\mathcal{H}(\mathbf{k}) \tag{9.3}$$

where $\mathbf{k} = (k_1, \ldots, k_r)$ and $\mathcal{H}(\mathbf{k})$ denotes weighted homogeneous solutions to (9.2) which are of weight k_i in the variables $(x^{(i)}, t_i)$. The L_i also preserve the decomposition (9.3). If we let $\mathcal{H}_i(k)$ denotes the subspace of polynomial solutions in $(x^{(i)}, t_i)$ of weight k to the single heat equation $L_i(f) = 0$, then we can decompose $\mathcal{H}(\mathbf{k})$.

Lemma 9.2. *There is a decomposition*

$$\mathcal{H}(\mathbf{k}) \quad \simeq \quad \prod_{i=1}^{r} \mathcal{H}_i(k_i) \tag{9.4}$$

where the inverse isomorphism is given by $(p_1, \ldots, p_r) \mapsto \prod_i p_i$.

Proof. The proof is by induction on r, with the case $r = 1$ trivially true. Let $f \in \mathcal{H}(\mathbf{k})$. First, f is a solution to $L_i(f) = 0$ for $i = 1, \ldots, r$. We first may represent such a weighted solution f of the first equation $L_1(f) = 0$ in the form

$$f \quad = \quad \sum_j g_j(x^{(2)}, \ldots, x^{(r)}, t_2, \ldots, t_r) p_j(x^{(1)}, t_1) \tag{9.5}$$

where the $\{p_j\}$ form a basis of solutions of weight k_1. Then, the remaining operators L_i for $i > 1$ act on the coefficient functions g_j. As f is also a solution to the equations $L_i(f) = 0$ for $i > 1$, these L_i, for $i > 1$, must annihilate the g_j. Hence, by induction each g_j has the required form; hence, so does f. □

Hence combining (9.3) and (9.4) of Lemma 9.2, we obtain for a multi-Gaussian kernel $K(x, \sigma)$ a polynomial basis for $\mathcal{H}^{\ell}_{K\,(x_0, \sigma_0)}$. For a given ℓ, we let

$$\mathcal{E}_\ell \quad = \quad \{\prod_{i=1}^{r} p^i_{j_i}(x^{(i)}, t_i)\} \tag{9.6}$$

where in the collection of products, each p^i_j varies over a polynomial basis for $\oplus_{j=1}^{\ell} \mathcal{H}_i(j)$ given above.

Proposition 9.3. *For a multi-Gaussian kernel $K(x, \sigma)$, $\mathcal{H}^{\ell}_{K\,(x_0, \sigma_0)}$ is spanned by the ℓ-jets obtained from the polynomials in $\mathcal{E}_{2\ell}$ translated to (x_0, t_0) after substituting $t_i = \frac{1}{2}\sigma_i^2$ (and $t_{i0} = \frac{1}{2}\sigma_{i0}^2$), $i = 1, \ldots, r$.*

Proof. Let $K_t(x) = \prod_i K_{i\,t_i}(x^{(i)})$ denote the associated multi–Gaussian kernel, so $K(x, \sigma)$ is obtained from $K_t(x)$ by substituting $t_i = \frac{1}{2}\sigma_i^2$, $i = 1, \ldots, r$. For $f = u * K$, we let $f_1(x, t) = u * K_t$. Then f_1 is a solution to $L_i(f) = 0$ for $i = 1, \ldots, r$. If we translate f_1 to (x_0, t_0) then $f_1(x - x_0, t - t_0)$ is again a solution and conversely. Then, if we decompose the Taylor expansion of f_1 into its weighted homogeneous summands, the translations are the weighted homogeneous summands in terms of $(\bar{x}, \bar{t}) = (x - x_0, t - t_0)$. Then, we obtain f from f_1 by substituting $t_i = \frac{1}{2}\sigma_i^2$ (and $t_{i0} = \frac{1}{2}\sigma_{i0}^2$), $i = 1, \ldots, r$. We see that the terms of weight $> 2\ell$ will map to zero in $\mathcal{H}^{\ell}_{K_i\,(x_0^{(i)}, \sigma_{i0})}$. Hence, the products in $\mathcal{E}_{2\ell}$, after translation and substitution, map to a spanning set. □

We can now draw a conclusion about extended Gaussian kernels in the special case that the homomorphism $\sigma' = \rho'(\sigma) : \mathbb{R}^k_+ \to \mathbb{R}^r_+$ has matrix $C = (c_{ij})$ with entries nonnegative integers such that $\sum_j c_{ij} \geq 1$ for each i.

Proposition 9.4. *For any extended Gaussian kernel $K(x,\sigma)$ with homomorphism ρ' with matrix C as above, there is a polynomial basis for $\mathcal{H}^{\ell}_{K\,(x_0,\sigma_0)}$ for any $(x_0,\sigma_0) \in \mathbb{R}^n \times \mathbb{R}^k_+$. Hence, K satisfies condition (A).*

Proof. The extended Gaussian kernel has the form $K(x,\sigma) = \rho'^*(\tilde{K}(x,\sigma'))$, where

$$\tilde{K}(x,\sigma') \quad = \quad \prod Q_i(\sigma'_i, D_{x^{(i)}}, D_{\sigma'_i})(K_i(x^{(i)},\sigma'_i)) \tag{9.7}$$

If $f = u*K$, then $f = (\prod Q_i(u*K'))_{|\sigma'=\rho'(\sigma)}$, where $K' = \prod_i K_i$. Then, $Q = \prod Q_i$ is a linear operator with polynomial coefficients of order, say M. As $u*K' \in \mathcal{H}_{K'}$, we conclude that

$$\mathcal{H}^{\ell}_{\tilde{K}} = Q^{(\ell+M)}(\mathcal{H}^{\ell+M}_{K'}) \tag{9.8}$$

where

$$Q^{(\ell+M)} : J^{\ell+M}(\mathbb{R}^n \times \mathbb{R}^r_+, \mathbb{R}) \to J^{\ell}(\mathbb{R}^n \times \mathbb{R}^r_+, \mathbb{R}) \tag{9.9}$$

is induced by Q. For any specific (x_0,σ'_0), the induced map on jets over (x_0,σ'_0) is linear. Hence, a polynomial basis for $\mathcal{H}^{\ell+M}_{K'\,(x_0,\sigma'_0)}$ maps under Q to a polynomial spanning set for $\mathcal{H}^{\ell}_{\tilde{K}\,(x_0,\sigma'_0)}$.

Finally, the conditions on ρ imply that composition with $id_{\mathbb{R}^n} \times \rho$ induces

$$\rho^{(\ell)} : J^{\ell}(\mathbb{R}^n \times \mathbb{R}^r_+, \mathbb{R}) \to J^{\ell}(\mathbb{R}^n \times \mathbb{R}^k_+, \mathbb{R}) \tag{9.10}$$

so that $\rho^{(\ell)}$ will send a spanning set of $\mathcal{H}^{\ell}_{\tilde{K}\,(x_0,\sigma'_0)}$ to one for $\mathcal{H}^{\ell}_{K\,(x_0,\sigma_0)}$ As $\rho'(\sigma)$ is polynomial in σ, compostion will take polynomials to polynomials. Also, polynomials have moderate growth, so it follows that a polynomial spanning set for $\mathcal{H}^{\ell+M}_{K'\,(x_0,\sigma_0)}$, given by 9.3 will be mapped by $\rho^{(\ell)} \circ Q^{(\ell+M)}$ to a spanning set for $\mathcal{H}^{\ell}_{K\,(x_0,\sigma_0)}$, so K satisfies condition (A). □

Sufficient Condition for Transversality to $\mathcal{H}^{\ell}_G$. Given a property $\mathcal{P}'$ of smooth functions defined by transversality to a closed Whitney stratified set W', we would like to determine when W' is transverse to $\mathcal{H}^{\ell}_G$.

First, for an important commonly occurring case, we can reduce verifying the transversality of W' to $\mathcal{H}^{\ell}_G$ to the case of a single fiber.

Proposition 9.5. *Suppose that $W' \subset J^{\ell}(\mathbb{R}^n \times \mathbb{R}^r_+, \mathbb{R})$ is invariant under the action of the Poincaré scaling group PS. Let $W_0 = W' \cap J^{\ell}(\mathbb{R}^n \times \mathbb{R}^r_+, \mathbb{R})_{(x_0,\sigma_0)}$. Then, W' is a trivial fiber bundle over $\mathbb{R}^n \times \mathbb{R}^r_+$ with fiber W_0. In order that W' is transverse to $\mathcal{H}^{\ell}_G$ it is necessary and sufficient that W_0 is transverse to $\mathcal{H}^{\ell}_{G\,(x_0,\sigma_0)}$ (in $J^{\ell}(\mathbb{R}^n \times \mathbb{R}^r_+, \mathbb{R})_{(x_0,\sigma_0)}$).*

Proof. W' is invariant under the action of PS, which acts fiberwise on $J^{\ell}(\mathbb{R}^n \times \mathbb{R}^r_+, \mathbb{R})$ and transitively on $\mathbb{R}^n \times \mathbb{R}^r_+$. Hence, the same argument as in Proposition 4.9 implies that W' is a trivial fiber bundle with fiber W_0. As both W' and $\mathcal{H}^{\ell}_G$ are fiber bundles over $\mathbb{R}^n \times \mathbb{R}^r_+$, transversality is equivalent to transversality of the fibers in a given fiber $J^{\ell}(\mathbb{R}^n \times \mathbb{R}^r_+, \mathbb{R})_{(x_0,\sigma_0)}$. Hence, the result follows. □

Suppose that G is an extended Gaussian kernel, and that W' is a closed Whitney stratified set invariant under PS with fiber W_0 over (x_0,σ_0). By Proposition 9.5, there is a set of polynomials which forms a basis for $\mathcal{H}^{\ell}_{G\,(x_0,\sigma_0)}$. Hence,

Corollary 9.6. *A necessary and sufficient condition that W' is transverse to $\mathcal{H}^\ell_G$ is that the polynomial basis spans a subspace transverse to W_0 in $J^\ell(\mathbb{R}^n \times \mathbb{R}^r_+, \mathbb{R})_{(x_0,\sigma_0)}$ at every point of W_0.*

We next give an alternate way to verify Corollary 9.6.

Corollary 9.7. *Suppose $K = \rho'^*(Q(K'))$ is an extended Gaussian kernel with Q of order M and ρ' a homomorphism as in Proposition 9.4, if we have W' invariant under PS then a sufficient condition that W' is tranverse to $\mathcal{H}^\ell_K$ is that the image of the polynomial basis for $\mathcal{H}^{\ell+M}_{K'\,(x_0,\sigma_0)}$ maps under $\rho^{(\ell)} \circ Q^{(\ell+M)}$ to a subspace transverse to W_0.*

We conclude this section with a final proposition that simplifies the verification of Corollary 9.7.

Proposition 9.8. *Suppose that K_t is a multi–Gaussian kernel. Let Q be a composition of first order operators of the form $D^{\boldsymbol{\alpha}}_t \prod_{j=1}^m D_{\mathbf{u}_j}$, where each $D_{\mathbf{u}_j}$ is a derivative with respect to a nonzero vector $\mathbf{u}_j$. Then,*

$$Q^{(\ell+M)} : \mathcal{H}_{K_t}(\ell + M) \to \mathcal{H}_{K_t}(\ell)$$

is surjective, where $M = 2|\boldsymbol{\alpha}| + m$.

Proof. Q commutes with each L_i, so that Q maps $\mathcal{H}_K$ to itself. Then, we prove the result by induction on the weighted degree $2|\boldsymbol{\alpha}| + m$. Thus, it is enough to prove the result for both $Q = D_{\mathbf{u}_i}$ or D_{t_i}. First, either of these only acts on one of the factors in each product term in $\mathcal{E}_\ell$. Hence, it is sufficient to prove the proposition for a single (anisotropic) Gaussian kernel K_t.

First we consider a regular Gaussian kernel. For D_t, the proposition is proven in [D4, Prop. 15.3]. Next, consider $D_{\mathbf{u}_i}$. Since the space of solutions to the (usual) heat equation is invariant under rotation, we may apply a rotation so that $D_{\mathbf{u}_i} = c \cdot D_{x_n}$. Thus, it is enough to prove it for D_{x_n}. Now we can repeat the argument given in [D4, Prop. 15.3], but replacing D_t by D_{x_n}. Specifically, consider the diagram

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \mathcal{H}_k & \xrightarrow{i_k} & W_k & \xrightarrow{L_k} & W_{k-2} & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle \alpha_1} & & \downarrow{\scriptstyle \alpha_2} & & \downarrow{\scriptstyle \alpha_3} & & \\ 0 & \longrightarrow & \mathcal{H}_{k-1} & \xrightarrow{i_{k-1}} & W_{k-1} & \xrightarrow{L_{k-1}} & W_{k-3} & \longrightarrow & 0 \end{array} \tag{9.11}$$

where W_k denotes the space of weighted homogeneous polynomials in (x, t) of weighted degree k and $\mathcal{H}_k$ denotes the weighted subspace space $\mathcal{H}_{K_t}(k)$ of solutions to the heat equation. Also, $L_k = L : W_k \to W_{k-2}$, where L denotes the heat operator, and α_i denote the restriction of D_{x_n} to the respective W_k or $\mathcal{H}_k$ so both α_2 and α_3 are surjective. As L_k is surjective, the rows are exact; hence, we can apply the snake lemma as in [D4, Prop. 15.3] to obtain the exact sequence in (9.12).

$$\begin{aligned} 0 \longrightarrow \ker\alpha_1 \xrightarrow{i_k} \ker\alpha_2 \xrightarrow{L_k} \ker\alpha_3 \\ \longrightarrow \operatorname{coker}\alpha_1 \xrightarrow{i_{k-1}} \operatorname{coker}\alpha_2 \xrightarrow{L_{k-1}} \operatorname{coker}\alpha_3 \longrightarrow 0 \end{aligned} \tag{9.12}$$

As both α_2 and α_3 are surjective, $\operatorname{coker}\alpha_2 = \operatorname{coker}\alpha_3 = 0$. It remains to see that $L_k : \ker(\alpha_2) \to \ker(\alpha_3)$ is surjective. As $\alpha_2 = D_{x_n}|W_k$, and $\alpha_3 = D_{x_n}|W_{k-2}$, $\ker(\alpha_2)$ (resp $\ker(\alpha_3)$) equals polynomials in W_k (resp. W_{k-2}) which only depend on $(x', t) = (x_1, \dots, x_{n-1}, t)$. However, the restriction of L is just the heat operator for $\mathbb{R}^{n-1}$; hence, it is surjective as required, implying the result for the isotropic Gaussian kernels.

Finally, we consider a general anisotropic heat equation as in (9.1). Then any solution has the form $p(A' \cdot x, t)$. First we observe $D_t(p(A' \cdot x, t) = \frac{\partial p}{\partial t}(A' \cdot x, t)$. Hence, as D_t is surjective for the regular heat equation, it is surjective in this case. Second, for $D_{\mathbf{u}_i}$, we let $\varphi(x) = A' \cdot x$. Then,

$$\begin{aligned} D_{\mathbf{u}_i}(p(A' \cdot x, t)) &= D_{\mathbf{u}_i}(p \circ \varphi) \\ &= D_{\mathbf{v}_i}(p)(A' \cdot x, t) \qquad \text{where } \mathbf{v}_i = d\varphi(\mathbf{u}_i) \end{aligned} \tag{9.13}$$

Then, as $D_{\mathbf{v}_i}$ is surjective for solutions to the regular heat equation, as p varies over weighted homogeneous solutions of weight k, we obtain from $\{D_{\mathbf{v}_i}(p)\}$, all possible weighted homogeneous solutions of weight $k-1$. Hence, from $\{D_{\mathbf{v}_i}(p)(A' \cdot x, t)\}$ we obtain all weighted homogeneous solutions to the anisotropic heat equation of weight $k-1$. Hence, the result follows. $\square$

As a corollary, we give a specific spanning set for $\mathcal{H}^{\ell}_{K\,(x_0, \sigma_0)}$.

Corollary 9.9. *Suppose K_t is a multi–Gaussian kernel and Q is an operator of the form $\sigma^{\mathbf{m}} \cdot D_t^{\boldsymbol{\alpha}} \prod_{j=1}^{m} D_{\mathbf{u}_j}$ with $M = 2|\boldsymbol{\alpha}| + m$. Let K be the kernel obtained from $Q(K_t)$ via substitution $t_i = \frac{1}{2}\sigma_i^2$ and $t_{i0} = \frac{1}{2}\sigma_{i0}^2$, $i = 1, \dots, r$. Then, a spanning set for $\mathcal{H}^{\ell}_{K\,(x_0, \sigma_0)}$ is given by*

$$\{(\sigma_0 + \bar{\sigma})^{\mathbf{m}} p(\bar{x}, \sigma_0 \bar{\sigma} + \frac{1}{2}\bar{\sigma}^2)\} \mod m^{\ell+1}_{\bar{x}, \bar{\sigma}}$$

where p ranges over $\mathcal{E}_{2\ell}$, and $m_{\bar{x}, \bar{\sigma}}$ denotes the ideal of smooth germs vanishing when $(\bar{x}, \bar{\sigma}) = (0, 0)$.

Proof. By Proposition 9.8, $Q^{\ell+M}$ is surjective. hence, the result follows by applying Proposition 9.3. $\square$

Example 9.10. The medial kernel is a Gaussian-based kernel obtained by applying the operator $Q = -\sigma \frac{\partial}{\partial \sigma}$ to the standard Gaussian kernel. In [D4] and [Mi] it is shown that for $n = 3, 4$, the associated operator

$$Q^{(\ell)} : J^{\ell}(n, 1)_{(0, \sigma_0)} \to J^{\ell-1}(n, 1)_{(0, \sigma_0)}$$

restricted to the heat equation jet space $\mathcal{H}^{\ell}_{(0, \sigma_0)}$ is transverse to the closed Whitney stratified set Γ defining the relative critical set off strata of codimension $> n$. The arguments used there are special cases for the medial kernel of the general results given in Corollaries 9.7 and 9.9. They yield the genericity and stability results stated in Example 8.7. In light of Theorems 6.3, 6.4 and Corollary 6.5, we conclude that when we convolve L^p functions, regular Borel measures, probability measures, etc., we will obtain the same genericity and stability results.

10. Geometric properties of Discrimination Functions and Measures for Feature Presence

Up to now our attention has focused on scale–based properties of the original (uniformly) tempered distribution. Beginning with a (uniformly) tempered distribution u, suppose we have a method for associating via an operation Ψ another tempered distribution $v = \Psi(u)$ which identifies "how much of a given property"occurs in a given region of physical space. We think of v as being a derived or associated distribution to u. This might be achieved by applying a differential operator, comparing with a texture mask, applying a filter by convolving with an associated kernel, such as e.g. a Gabor filter, which distinguishes certain geometric features. Alternately, a statistical method (as given e.g. in [ZWM]) may be applied to obtain a statistical measure for the amount of a property in a given region. This statistical measure may be in the form of probability or Borel measure. We may now apply scale–based geometry to v to deduce geometric information, in the form of a geometric structure in scale space $W(v * G)$, about the regions where the feature detected by v occurs. This is schematically illustrated in diagram 10.1.

(10.1)

$$\left\{\begin{matrix}\text{original}\\ \text{distrib.}\end{matrix}\right\} \longrightarrow \left\{\begin{matrix}\text{derived}\\ \text{distrib.}\end{matrix}\right\} \xrightarrow{\text{scale}} \left\{\begin{matrix}\text{convolved}\\ \text{distrib.}\end{matrix}\right\} \xrightarrow{\text{geom}} \left\{\begin{matrix}\text{geometric}\\ \text{structure}\end{matrix}\right\}$$

$$u \longrightarrow v = \Psi(u) \longrightarrow v * G \longrightarrow W(v * G)$$

10.1 (Question)**.** To what extent will the generic scale-based properties normally exhibited for spaces of distributions continue to be exhibited generically for the derived or associated distributions v?

To answer this, we assume that the procedure for associating the disribution v to u is given by a continuous operator

$$\Psi : \mathcal{T} \to \mathcal{S}^*_{n,\Lambda} \quad \text{or} \quad \Psi : \mathcal{T} \to \mathcal{S}'_n \tag{10.2}$$

where $\mathcal{T}$ may denote a subspace of $\mathcal{S}^*_{n,\Lambda}$, $\mathcal{S}'_n$, or even $\mathcal{D}'$. If we consider such an operator given e.g. by convolution with a kernel $K \in \mathcal{S}_{n,\Lambda}$ or $\mathcal{S}_n$, then by Theorems 2.5 and 3.7, such operators are continuous linear operators, and have images smooth functions of (uniformly) moderate growth. Such filters could be e.g. (non-scale–based) Gabor filters, edge detection kernels as in Example 4.4, etc. Likewise operators given as linear partial differential operators with constant coefficients defined on $\mathcal{S}^*_{n,\Lambda}$, $\mathcal{S}'_n$, or finite dimensional extensions of these in $\mathcal{D}'$ are also continuous.

Alternately we allow an operator which associates a measure to an original tempered distribution. One basic example occurs for the detection of "textures" in images. As already mentioned, there has been considerable work devoted to characterizing in terms of various associated objects, involving measures, wavelets, etc. (see e.g. [Ju], [ZWM], [Mal]). Rather than identify a texture, we are simply interested in an operation which detects a distinguishing feature of a texture. We will also consider the case that the amount of distinguishing feature appearing in a given region can be expressed by a measure.

We will not concern ourselves here with the details of how such a tempered distribution is constructed, although we do assume that it is given by a continuous operator as in (10.2). If we apply a scale–based geometric property $\mathcal{P}$ to $\Psi(u)$, can we expect that it will be generically exhibited on compact subsets of scale space? Unlike the results of §6, we cannot expect in general that the subspace of derived distributions will satisfy condition (B), especially those defined by detecting specific texture properties which require a specific choice of scale. Instead, we should expect that if the object or region exhibiting the property is large compared to the scale of the detected property then it is possible to have sufficient robustness of the derived distributions to obtain generic properties in the larger scale range. Hence, we shall state all results in terms of a given compact region of scale space C.

In the case that Ψ is nonlinear, there are so many ways to introduce nonlinearities that the best we can do is state a general result that still requires considerable work to verify. It is a corollary of Theorem 6.3, and gives a positive answer in terms of the composition of the map $\Psi|\mathcal{T}$ with c_G (or $c_{\Lambda,\, G}$) having smooth image in the sense of [D2, Def. 1.1], with image equal to $\mathcal{H}^{\ell}_{G}$ (resp. $\mathcal{H}^{\ell}_{\Lambda,\, G}$).

Proposition 10.2. *Suppose G is a scale–based the kernel, $\mathcal{P}$ is a transversely defined scale–based property (via W), and C is a compact subset of scale space. Let Ψ be a continuous operator as in (10.2). Given $\mathcal{T} \subset \mathcal{S}^*_{n,\Lambda}$ or $\mathcal{S}'_n$ (or $\mathcal{D}'$), suppose there is an open subset U of scale space with $C \subset U$ such that the composition of Ψ with c_G (resp. $c_{\Lambda,\, G}$) has smooth image on U, with image equal to $\mathcal{H}^{\ell}_{G}(U)$ (resp. $\mathcal{H}^{\ell}_{\Lambda,\, G}(U)$). Then, there is an open dense subset of $\mathcal{T}$ consisting of u for which the G–convolution $\Psi(u) * G$ will exhibit $\mathcal{P}$ generically on C. The geometric structure has generic local structure $W_{loc}(x)$ for x in the interior of C, and the structure is structurally stable on C.*

Proof. As in the proof of Theorem 6.3, this is a consequence of the relative transversality theorem. $\mathcal{P}$ is defined by transversality to the closed Whitney stratified set $W \subset \mathcal{H}^{\ell}_{G}$ (resp. $\mathcal{H}^{\ell}_{\Lambda,\, G}$). Let $i : \mathcal{T} \hookrightarrow \mathcal{S}'_n$ (or $\mathcal{S}^*_{n,\Lambda}$) denote the inclusion map, and let r denote restriction to U. The composition $r \circ c_G \circ \Psi \circ i$ (or with $c_{\Lambda,\, G}$ in place of c_G) is continuous. The assumption that it has smooth image with image $\mathcal{H}^{\ell}_{G}(U)$ (resp. $\mathcal{H}^{\ell}_{\Lambda,\, G}(U)$) implies by the relative transversality theorem that for the compact subset of scale space $C \subset U$, the set of $u \in \mathcal{T}$ such that $j^{\ell}(\Psi(u) * G)$ is transverse to W relative to $\mathcal{H}^{\ell}_{G}(U)$ (resp. $\mathcal{H}^{\ell}_{\Lambda,\, G}(U)$) is an open dense subset of $\mathcal{T}$. Then, the local generic structure and the stability on compact subsets follow exactly as earlier. □

To use this theorem we must be able to verify the condition of smooth image. This will vary considerably for various nonlinear operators. We limit ourselves here to the linear case. First, we observe a property of scale–based kernels.

Proposition 10.3. *Suppose the scale–based kernel G satisfies condition (A). Let C be a compact subset of scale space.*

(1) *There is a finite dimensional subspace $\mathcal{V}$ of $\mathcal{S}'_n$ (resp. $\mathcal{S}^*_{n,\Lambda}$) consisting of smooth non–negative functions of compact support, so that the map from $\mathcal{V}$ to $\mathcal{H}^{\ell}_{G\,(x_0,\lambda_0)}$ (resp. $\mathcal{H}^{\ell}_{\Lambda,G\,(x_0,\lambda_0)}$) given by $u \mapsto j^{\ell}(u*G)(x_0, \lambda_0)$ is surjective for all $(x_0, \lambda_0) \in C$ (and hence for an open subset U containing C).*

(2) *Alternately, given a mesh $\mathcal{B}$, there is a refinement $\mathcal{B}'$ and a finite dimensional subspace of discrete non–negative functions of compact support defined using $\mathcal{B}'$, which has the same surjectivity properties as in 1).*

(3) *Given bases $\{\varphi_j\}$ for the subspaces in either 1) and 2), there is an $\varepsilon > 0$ so that if a subset of functions $\{\varphi'_j\}$ of compact support with each $\varphi'_j \in \mathcal{S}'_n$ are within ε of φ_j in the (essential) sup norm, then $\{\varphi'_j\}$ also span a subspace satisfying the surjectivity conditions in 1) (and 2)).*

Proof. We first establish the result for smooth non–negative functions of compact support. By Proposition 6.6 and the action of the Poincaré scaling group, given $(x_0, \lambda_0) \in C$ there is a finite set of smooth non–negative functions of compact support $\{\varphi_i, i = 1 \dots r\}$ such that $j^\ell(\varphi_i * G)(x_0, \lambda_0)$ span $\mathcal{H}^\ell_{G\,(x_0,\lambda_0)}$ (resp. $\mathcal{H}^\ell_{\Lambda,G\,(x_0,\lambda_0)}$). Then there is an open neighborhood U_1 of (x_0, λ_0) in scale space such that $j^\ell(\varphi_i * G)(x_0, \lambda_0)$ span $\mathcal{H}^\ell_{G\,(x_1,\lambda_1)}$ (resp. $\mathcal{H}^\ell_{\Lambda,G\,(x_1,\lambda_1)}$) for all $(x_1, \lambda_1) \in U_1$. As C is compact, it can be covered by a finite number of such U_1. The subspace spanned by the $\{\varphi_i\}$ corresponding to the finite cover have the desired property.

To obtain the analogous result for discrete functions, we first use the argument of §7, to approximate $\{\varphi_i, i = 1 \dots r\}$ by discrete functions of compact support $\{\varphi'_i, i = 1 \dots r\}$ defined on a refinement $\mathcal{B}_1$ so that $j^\ell(\varphi'_i * G)(x_0, \lambda_0)$ span $\mathcal{H}^\ell_{G\,(x_0,\lambda_0)}$ (resp. $\mathcal{H}^\ell_{\Lambda,G\,(x_0,\lambda_0)}$). Then, the rest of the argument is the same, where at the last step we must choose a common refinement $\mathcal{B}'$ on which all of the φ'_i are defined.

By the openness of the surjectivity condition and the continuity of c_G or $c_{\Lambda,G}$, we can find an $\varepsilon > 0$ so that if we replace φ_i by another set of functions φ'_i each within ε in the (essential) sup–norm, then this set will still satisfy the surjectivity condition.

□

Given a scale–based kernel G satisfying condition (A), we generally introduce for a linear operator Ψ in (10.2) a substitute for condition (B).

Definition 10.4. Given a compact subset of scale space C, the linear operator Ψ satisfies condition (B_C) if there is a finite dimensional $\mathcal{V} \subset \mathcal{T}$, so that the map from $\mathcal{V}$ to $\mathcal{H}^\ell_{G\,(x_0,\lambda_0)}$ (resp. $\mathcal{H}^\ell_{\Lambda,G\,(x_0,\lambda_0)}$) given by $u \mapsto j^\ell(u * G)(x_0, \lambda_0)$ is surjective for all $(x_0, \lambda_0) \in C$ (and hence for an open subset U containing C).

As a simple consequence, we have.

Theorem 10.5. *Suppose that G is a scale–based kernel satisfying condition (A), C is a compact subset of scale space, and Ψ satisfies condition (B_C). Then, Ψ has smooth image on an open set U containing C. Hence, given any transversally defined scale–based geometric property $\mathcal{P}$, there is an open dense subset of $u \in \mathcal{T}$ such that $\Psi(u) * G$ exhibit $\mathcal{P}$ generically and stably on C.*

Proof. We use the subspace $\mathcal{V}$ to establish the smooth image of Ψ on U just as in the proof of Theorem 5.8. Then, Proposition 10.2 yields the genericity and stability. □

Thus, to deduce genericity and stability results for associated distributions, it is sufficient by Theorem 10.5 to verify that various operators Ψ satisfy condition (B_C). For this, we will apply Proposition 10.3 and show that the finite dimensional subspaces of smooth functions with compact support or discrete functions given there are in the image of Ψ.

Ψ defined by Partial Differential Operators. Let $P(D)$ be a linear partial differential operator with constant coefficients. It defines an operator $\Psi = P(D)$ as in (10.2), (note in the case of $\mathcal{T} \subset \mathcal{D}'$, not all of $\mathcal{D}'$ need be mapped to $\mathcal{S}'_n$). Given a scale–based kernel G satisfying condition (A) and a compact subset C of scale space, we seek conditions ensuring that Ψ satisfies condition (B_C). It is sufficient to show by Proposition 10.3 that image(Ψ) contains the finite dimensional subspace of smooth functions with compact support satisfying the surjectivity condition in that proposition. For this we use results of Malgrange–Ehrenpreis and Nirenberg on solutions to constant coefficient PDE's.

1) First, by the Malgrange–Ehrenpreis Theorem, the equation $P(D)(u) = g$ has a fundamental solution K so $P(D)(K) = \delta$ where δ is the Dirac distribution. If $K \in \mathcal{S}'_n$, then for any $g \in \mathcal{E}'_n$, $g * K \in \mathcal{S}'_n$ and is a solution. Hence, in this case Ψ maps onto $\mathcal{E}'_n$, in particular its image contains all smooth functions with compact support. For example, the fundamental solution for the Laplacian Δ on $\mathbb{R}^n$ is $K(x) = c \cdot \|x\|^{2-n}$ if $n \neq 2$ (or $c \cdot \log(x)$ if $n = 2$). For either of these, $(1 + \|x\|^2)^{-1}K(x) \in L^1(\mathbb{R}^n)$, so K defines a tempered distribution.

2) Second, if the fundamental solution is not in $\mathcal{S}'_n$, then we can still apply a Theorem of Nirenberg (see e.g. [Fo2, Chap. 2]) which asserts that if g is smooth of compact support, then there is a smooth function u which is a solution of $P(D)(u) = g$. Hence, if we allow u to vary over smooth functions, then the image of Ψ contains any given subspace of smooth functions with compact support.

Combining these results with Proposition 10.3 we conclude that Question 10.1 has a positive answer for such operators.

Corollary 10.6. *Suppose G is a scale–based kernel satisfying condition (A), and C is a compact subspace of scale space. If $\Psi = P(D)$, a linear partial differential operator with constant coefficients, then there is a finite dimensional subspace $\mathcal{V}$ of smooth functions (or a finite dimensional subspace of $\mathcal{S}'_n$ if $P(D)$ has a fundamental solution in $\mathcal{S}'_n$) which satisfies condition (B_C). Thus, expanding any subspace of $\mathcal{S}'_n$ by this finite dimensional subspace gives a subspace $\mathcal{T}$ on which Ψ satisfies the conclusions of Theorem 10.5 on generically and stably on C.*

Second we consider conditions using the finite dimensional subspace of discrete functions provided by Proposition 10.3.

Satisfying condition (B_C) by approximating "block functions". Given a scale–based kernel G satisfying condition (A), a compact subset of scale space C, and a mesh $\mathcal{B}$, we may alternately find a finite dimensional subspace of discrete functions of compact support relative to a refinement $\mathcal{B}'$ of $\mathcal{B}$ which satisfy condition (B_C). The support is restricted to a finite subset of cells B_j of $\mathcal{B}'$, and the space is spanned by a finite set $\varphi'_i = \sum a_{ij}\chi_j$, where χ_j is the characteristic function for B_j. By 3) of Proposition 10.3, we can find an $\varepsilon > 0$ so that if we replace φ'_i by another set of functions φ''_i each within ε in, say, the (essential) sup–norm, then this set will still satisfy the surjectivity condition in that proposition.

We refer to the characteristic function χ_j for the cell B_j as a "block function"(see Fig 6).

Suppose that there are m cells B_j whose union form the common support. Also, let $a = \max_{ij} |a_{ij}|$; and finally let $\varepsilon' = \frac{\varepsilon}{am}$. We consider an operator Ψ with the property that for each block function χ_j there is a $u_j \in \mathcal{T}$ such that $\Psi(u_j)$ is within ε' of χ_j. Then, $v_i = \sum a_{ij}u_j$ maps via Ψ to $\varphi''_i = \sum a_{ij}\Psi(u_j)$ which is within ε

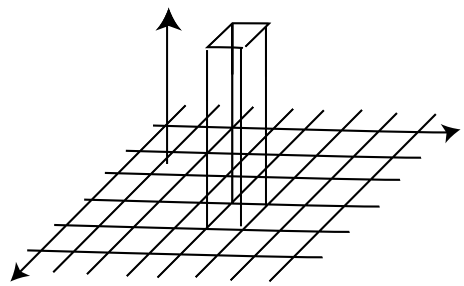

FIGURE 6. A Block function on a square mesh for $\mathbb{R}^2$

of φ_i'. Hence, Ψ satisfies condition (B_C), so Question 10.1 has a positive answer in this case.

Corollary 10.7. *Suppose G is a scale–based kernel satisfying condition (A), and C is a compact subspace of scale space. There is a refinement $\mathcal{B}'$ of the mesh $\mathcal{B}$, a finite subset of cells $\{B_j, j = 1, \dots, m\}$ (of $\mathcal{B}'$) and an $\varepsilon > 0$ so that if Ψ approximates the block functions on the $B_j, j = 1, \dots, m$ to within ε in the (essential) sup–norm, then Ψ satisfies condition (B_C). Thus, Ψ satisfies the conclusions of Theorem 10.5 on generically and stably on C.*

Remark 10.8. We note that in the preceding discussion, we could have viewed the simple measures associated to the φ_i' and instead considered measures approximating them, and obtained an analogous condition.

Simple Texture Discrimination with Generic Geometric Properties. We apply the result of Corollary 10.7 to guarantee genericity and stability of scale–based geometric properties of associated texture discrimination distributions.

Suppose we are considering a scale–based kernel G satisfying condition (A), and that C is a compact subset of scale space. By the discussion in Corollary 10.7, there is a refinement $\mathcal{B}'$ of $\mathcal{B}$ and a finite dimensional subspace spanned by discrete functions of compact support defined using $\mathcal{B}'$, which satisfy the surjectivity condition in Proposition 10.3. To obtain generic scale–based geometric properties for derived distributions defined using a linear operator Ψ, it is sufficient to show we can closely approximate the block functions χ_j to within ε.

This suggests a crude method for detecting texture which guarantees generic scale–based geometric properties on C. On each support cell B_j, we fix a "mask function" K_j with support B_j. We measure the amount of texture in a function g on B_j by the integral $\int g \cdot K_j$, and associate a discrete function with this value on B_j. To extend this process to obtain a continuous linear operator on say $\mathcal{S}_n'$, we choose smooth K_j and multiply by a bump function with support in B_j and closely approximating χ_j. If we still denote this function by K_j, then $K_j \in \mathcal{S}_n'$. Hence, we define Ψ by $\Psi(u) = \sum u(K_j)\chi_j$. Provided $c_j = \int K_j^2 \neq 0$, $\Psi(c_j^{-1} K_j) = \chi_j$. Thus, by Corollary 10.7, Ψ satisfies condition (B_C). Hence, the derived distributions $\Psi(u)$ measuing the texture on the support cells $\cup_j B_j$ will have generic scale–based geometric properties.

We further note that in the preceding argument we can replace $\mathcal{S}_n'$ by any standard subspace $\mathcal{T} \subset \mathcal{S}_n'$ by choosing $u_j \in \mathcal{T}$ with support in B_j such that

$c'_j = u_j(K_j) \neq 0$, so $\Psi(c'^{-1}_j \cdot u_j) = \chi_j$. Also, we can extend texture discrimination to additional cells. This will make no difference in the genericity statements.

Corollary 10.9. *Suppose G is a scale–based kernel satisfying condition (A), $\mathcal{T}$ is a subspace of $\mathcal{S}'_n$ closed under nonnegative linear combinations, and C is a compact subspace of scale space. Consider the refinement $\mathcal{B}'$ of the mesh $\mathcal{B}$, with finitely many support cells $\{B_j, j = 1, \dots, m\}$ and the $\varepsilon > 0$ for the compact subset C given by Proposition 10.3. If there are smooth texture masks K_j and $u_j \in \mathcal{T}$ both with support in $B_j, j = 1, \dots, m$ such that $u_j(K_j) \neq 0$, then the texture discrimination operation*

$$\Psi(u) \quad = \quad \sum_j u(K_j)\chi_j \tag{10.3}$$

satisfies condition (B_C). Thus, by Theorem 10.5 given any transversally defined scale–based geometric property $\mathcal{P}$, there is an open dense subset of $u \in \mathcal{T}$ such that $\Psi(u) * G$ exhibit $\mathcal{P}$ generically and stably on C.

Remark 10.10. Hence, this simple form of texture discrimination given in (10.3) will yield detection distributions with generic scale–based geometric properties on C, including structural stability of the texture–defined geometric structures under sufficiently small perturbations of the initial distributions! Of course, most discrimination criteria will require taking maxima over a series of masks or apply statistical techniques so they will be nonlinear. However, the arguments given here at least provide some reason for optimism in expecting that provided that the texture scale in sufficiently small compared to that of the objects possessing it, the associated scale–based geometric properties will have generic properties.

11. Vector-valued Distributions

In this section we expand our earlier results so they apply to vector–valued distributions with values in $\mathbb{R}^p$. These are equivalent to p–tuples of distributions. For example, these can arise from color images which are 3–dimensional distributions measuring intensity, color, and hue. Also, if we consider a single (uniformly) tempered distribution, but we wish to simultaneously consider several geometric properties, then we will associate a tuple of derived distributions obtained by applying a set of operations.

We begin by considering the geometric properties of a set of independent distributions; and establish the genericity of simultaneous geometric properties.

Vector-valued (Uniformly) Tempered Distributions $\mathcal{S}(\mathbb{R}^n \times \Lambda, \mathbb{R}^p)^*$. We consider generally a p-tuple of uniformly tempered distributions

$$u = (u_1, \dots, u_p) \quad \in \quad \left(\mathcal{S}^*_{n,\Lambda}\right)^p = \mathcal{S}^*_{n,\Lambda} \times \cdots \times \mathcal{S}^*_{n,\Lambda}$$

We denote the Schwarz space of uniformly rapidly decreasing functions mapping to $\mathbb{R}^p$ by $\mathcal{S}(\mathbb{R}^n \times \Lambda, \mathbb{R}^p)$ (abbreviated $\mathcal{S}^{(p)}_{n,\Lambda}$) and the space of uniformly tempered distibutions by $\mathcal{S}(\mathbb{R}^n \times \Lambda, \mathbb{R}^p)^*$ (or more simply by $\mathcal{S}^{*\,(p)}_{n,\Lambda}$). Likewise, we denote p–tuples of tempered distributions by $\mathcal{S}(\mathbb{R}^n, \mathbb{R}^p)'$ (or just $\mathcal{S}'^{(p)}_n$)

Then, given $G = (G_1, \dots, G_p) \in \mathcal{S}(\mathbb{R}^n \times \Lambda, \mathbb{R}^p)$, then for $u \in \mathcal{S}^{*\,(p)}_{n,\Lambda} = (\mathcal{S}(\mathbb{R}^n \times \Lambda, \mathbb{R}^p)^*)$ (or $\mathcal{S}'^{(p)}_n$) we can define $u * G = (u_1 * G_1, \dots, u_p * G_p)$. By Theorem 3.7 applied to each coordinate of $u * G$, we see $u * G \in C^\infty(\mathbb{R}^n \times \Lambda)^p \simeq C^\infty(\mathbb{R}^n \times \Lambda, \mathbb{R}^p)$.

For fixed G, this defines a linear transformation

$$c_{\Lambda, G} : \mathcal{S}(\mathbb{R}^n \times \Lambda, \mathbb{R}^p)^* \to C^\infty(\mathbb{R}^n \times \Lambda, \mathbb{R}^p) \tag{11.1}$$

The restriction of $c_{\Lambda, G}$ to $\mathcal{S}_n^{\prime (p)}$ will again be denoted by c_G. Applying Theorem 3.7, we see these maps $c_{\Lambda, G}$ and c_G are continuous. Analogous to §4, we let $\mathcal{H}^{(p)}_{\Lambda, G}$ (resp. $\mathcal{H}^{(p)}_G$) denote the image of $c_{\Lambda, G}$ (resp. c_G) in $C^\infty(\mathbb{R}^n \times \Lambda, \mathbb{R}^p)$. As in (4.13), we define

$$\mathcal{H}^{\ell,(p)}_{\Lambda, G} \stackrel{\text{def}}{=} \{j^\ell(f)(x, \lambda) : f = u * G \text{ for } u \in \mathcal{S}(\mathbb{R}^n \times \Lambda, \mathbb{R}^p)^*\} \tag{11.2}$$

We have an analogous definition for $\mathcal{H}^{\ell,(p)}_G$.

By the coordinatewise definitions of $c_{\Lambda, G}$ and c_G,

$$\mathcal{H}^{\ell,(p)}_{\Lambda, G} \simeq \prod_{\mathbb{R}^n \times \mathbb{R}^k_+} \mathcal{H}^\ell_{\Lambda, G_i} \tag{11.3}$$

and

$$\mathcal{H}^{\ell,(p)}_{G} \simeq \prod_{\mathbb{R}^n \times \mathbb{R}^k_+} \mathcal{H}^\ell_{G_i} \tag{11.4}$$

where the products on the right are fiber products over $\mathbb{R}^n \times \mathbb{R}^k_+$ for $i = 1, \dots, p$. Then, (11.3) and (11.4) induce diffeomorphisms of the fibers over a point $(x_0, \lambda_0) \in \mathbb{R}^n \times \mathbb{R}^k_+$ which on the RHS are products of fibers of convolution jet spaces.

Definition 11.1. We say a multi–kernel $G = (G_1, \dots, G_p) \in \mathcal{S}_n^{(p)}$ is a *scale–based* multi–kernel if there is a single scaling group $\mathbb{R}^k_+$ such that for each i there is an action of $\mathbb{R}^k_+$ on $\mathbb{R}^n$ for which G_i is scale–based.

Note we do not require that the action be the same for each G_i; in fact, it can be different for each i, allowing us to simultaneously consider different scaling behavior for different distributions.

In this case, we can still deduce the structure of $\mathcal{H}^{\ell,(p)}_G$ and $\mathcal{H}^{\ell,(p)}_{\Lambda, G}$.

Proposition 11.2. *If G is scale-based then for $\Lambda = \mathbb{R}^k_+$, both $\mathcal{H}^{\ell,(p)}_G$ and $\mathcal{H}^{\ell,(p)}_{\Lambda, G}$ are*

1. *trivial fiber bundles over $\mathbb{R}^n \times \Lambda$, and*
2. *semi–algebraic submanifolds of $J^\ell(\mathbb{R}^n \times \Lambda, \mathbb{R}^p)$.*

Proof. We consider the case of $\mathcal{H}^{\ell,(p)}_{\Lambda, G}$, with similar arguments applying to $\mathcal{H}^{\ell,(p)}_G$.

First, by Proposition 4.9 each $\mathcal{H}^\ell_{\mathbb{R}^k_+, G_i}$ is a fiber bundle over $\mathbb{R}^n \times \mathbb{R}^k_+$ with smooth fiber. Hence, the fiber product of the $\mathcal{H}^\ell_{\mathbb{R}^k_+, G_i}$ is again a fiber bundle over $\mathbb{R}^n \times \mathbb{R}^k_+$, with fibers equal to the product of the individual fibers, and hence, smooth. By (11.3), this is $\mathcal{H}^{\ell,(p)}_{\mathbb{R}^k_+, G}$. This still holds for the restriction to any $\mathbb{R}^n \times \Lambda$.

Second, again by Proposition 4.9, $\mathcal{H}^\ell_{\mathbb{R}^k_+, G_i}$ is a semi-algebraic subset of $J^\ell(\mathbb{R}^n \times \mathbb{R}^k_+)$. Let

$$\pi : \prod_{i=1}^p J^\ell(\mathbb{R}^n \times \mathbb{R}^k_+) \to \prod_{i=1}^p (\mathbb{R}^n \times \mathbb{R}^k_+)$$

denote the product of the natural projections. Then by (11.3)

$$\mathcal{H}^{\ell,(p)}_{\mathbb{R}^k_+, G} = \prod_{i=1}^p \mathcal{H}^\ell_{\mathbb{R}^k_+, G_i} \cap \pi^{(-1)}(\Delta(\mathbb{R}^n \times \mathbb{R}^k_+)) \tag{11.5}$$

where $\Delta(\mathbb{R}^n \times \mathbb{R}^k_+)$ denote the diagonal. As π is algebraic, the RHS of (11.5), presents $\mathcal{H}^{\ell,(p)}_{\mathbb{R}^k_+, G}$ as the intersection of two sets each of which is semi–algebraic. Thus, $\mathcal{H}^{\ell,(p)}_{\mathbb{R}^k_+, G}$ is semi–algebraic. □

Contrary to the case of a single distribution, there is a *coordinatewise action of the Poincaré group* PS defined from the single scaling group $\mathbb{R}^k_+$. However, it is not induced geometrically for vector valued distributions as it was in Proposition 4.7.

Then, we still have the analogue of Theorem 5.8.

Theorem 11.3. *Suppose G be a scale–based multi–kernel on $\mathbb{R}^n \times \mathbb{R}^k_+$.*

1) Suppose W is a closed Whitney stratified subset of $\mathcal{H}^{\ell,(p)}_{\Lambda, G}$ with strata W_i, then for a compact set $C \subset (\mathbb{R}^n \times \Lambda)$, the set of uniformly tempered distributions

$$\mathcal{W} = \{u \in \mathcal{S}^{*\,(p)}_{n,\Lambda} : j^\ell(u * G) \textit{ is transverse on } C \textit{ to all } W_i (\textit{in } \mathcal{H}^{\ell,(p)}_{\Lambda, G})\} \tag{11.6}$$

is an open dense subset of $\mathcal{S}^{*\,(p)}_{n,\Lambda}$.

2) If instead, W is a closed Whitney stratified subset of $\mathcal{H}^{\ell,(p)}_{G}$, then

$$\mathcal{W} = \{u \in \mathcal{S}'^{(p)}_{n} : j^\ell(u * G) \textit{ is transverse on } C \textit{ to all } W_i (\textit{in } \mathcal{H}^{\ell,(p)}_{G})\} \tag{11.7}$$

is an open dense subset of $\mathcal{S}'^{(p)}_{n}$.

Proof. By the proof of Theorem 5.8, there is for each factor in (11.3) (resp. (11.4)) a finite dimensional subspace $\mathcal{V}_i$ which maps submersively onto $\mathcal{H}^{\ell}_{\Lambda, G_i\,(x',\lambda')}$ (resp. $\mathcal{H}^{\ell}_{G_i\,(x',\lambda')}$) for all (x', λ') in a neighborhood U of (x_0, λ_0). Then, the product $\mathcal{V} = \prod_i \mathcal{V}_i$ satisfies the condition of smooth image for the relative transversality theorem applied to $c_{\Lambda, G}$ and c_G , implying that $\mathcal{W}$ is open and dense. □

Example 11.4. Let G be a scale–based multi–kernel. Suppose the property $\mathcal{P}_i$ is defined by transversality to a closed Whitney stratified subset $W^{(i)} \subset \mathcal{H}^{\ell}_{\Lambda, G_i}$ (resp. $\mathcal{H}^{\ell}_{G_i}$) with strata $\{W^{(i)}_j\}$. Then, by (11.3), $W = \prod_i W^{(i)}$ is a closed Whitney stratified subset of $\mathcal{H}^{\ell,(p)}_{\Lambda, G}$ (resp. $\mathcal{H}^{\ell,(p)}_{G}$) with strata $\prod_i W^{(i)}_{j_i}$. The simultaneous occurence of the properties $\mathcal{P}_i$ for the $u_i * G_i$ is given by transversality to W.

Hence, we conclude

Corollary 11.5. *Suppose $G = (G_1, \dots, G_p)$ is a scale–based multi–kernel. If $\mathcal{P}_i$ are transversally defined scale–based properties (for the kernel G_i), then for a compact subset C of scale space, there is an open, dense subset of $u = (u_1, \dots, u_p) \in \mathcal{S}^{*\,(p)}_{n,\Lambda}$ (or $\mathcal{S}'^{(p)}_{n}$) such that each $u_i * G$ exhibits $\mathcal{P}_i$ generically on C, and moreover, the Whitney stratified sets $W_i(u_i * G)$ for $i = 1, \dots, p$ intersect tranversally on C.*

Proof. By Theorem 11.3, $j^\ell(u*G)$ is transverse to $W = \prod_i W^{(i)}$, which is equivalent to each $j^\ell(u_i * G_i)$ being transverse to each $W^{(i)}$, and the $W_i(u_i * G) = j^\ell(u_i * G_i)^{-1}(W^{(i)})$ intersecting transversally. □

Corollary 11.6. *Suppose G is a scale–based multi–kernel. If W is tranverse to $\mathcal{H}^{\ell,(p)}_{\Lambda, G}$ (resp. $\mathcal{H}^{\ell,(p)}_{G}$), then on a compact subset C of scale space, there is an open dense subset of $u \in \mathcal{S}^{*\,(p)}_{n,\Lambda}$ (resp. $\mathcal{S}'^{(p)}_{n}$) such that $W(u * G)$ exhibits on C the same generic properties as for generic smooth functions*

Example 11.7. Interesting singularity–theoretic properties of the smooth mapping $f = u{*}G$ can also be investigated for vector–valued distributions u. Such properties are frequently defined by Whitney stratified sets $W \subset J^\ell(\mathbb{R}^n \times \Lambda, \mathbb{R}^p)$. For example, the simplest structure is the singular set of the associated mapping, $\Sigma(f)$, which consists of points $(x, \lambda) \in \mathbb{R}^n \times \Lambda$ such that $\text{rank}(df(x, \lambda)) < r = \min\{n + k, p\}$. $\Sigma(f) = j^1(f)^{-1}(\bar{\Sigma}_1)$, where Σ_1 denotes the 1–jets of corank > 0, i.e. rank $< r$. It identifies the points where the (derivatives of the) coordinate functions of f, $f_i = u_i * G_i$ are dependent.

For the case where the kernels G_i are extended Gaussian kernels for which the differential operators are of the form considered in Proposition 9.8, we have the 1–jets will contain all linear terms in the x_i (modulo higher order terms). For the linear terms involving scale parameters, the precise statement is more subtle. However, in the case there is a single common scale parameter σ for all G_i, we obtain in each component independent linear terms in σ. Thus, we obtain that $\mathcal{H}_G^{(1)} = J^1(\mathbb{R}^n \times \mathbb{R}_+)$. Hence, by Theorem 11.3, the G–convolutions will generically exhibit the Σ_j–singularities for each j. For example, for two distributions, although the individual geometric properties are exhibited generically, there will be a dependence between (the derivatives of) their G_i–convolutions on a curve in scale space.

Just as we deduced consequences of Theorem 5.8 for the genericity and local stability (Theorems 5.10 and 5.11, we can deduce analogous results for $\mathcal{S}^{*\,(p)}_{n,\Lambda}$ and $\mathcal{S}'^{\,(p)}_n$.

Theorem 11.8 (Scale–based Generic Structure and Structural Stability). *Suppose G is a scale–based multi–kernel and that $\mathcal{P}$ scale–based geometric property defined by transversality to a closed Whitney stratified subset W of $\mathcal{H}^{\ell,(p)}_{\Lambda,G}$ (resp. $\mathcal{H}^{\ell,(p)}_G$). Also, suppose $f = u{*}G$ exhibits $\mathcal{P}$ generically on a compact subset C of scale space.*

1) At an interior point (x_0, λ_0) of C, there is a strata preserving local homeomorphism

$$(W(f), (x_0, \lambda_0)) \quad \simeq \quad (W_{loc}(j^\ell(f)(x_0, \lambda_0)) \times \mathbb{R}^m, (j^\ell(f)(x_0, \lambda_0), 0))$$

where $m = n - \text{codim}\,(W_i)$ for $j^\ell(f)(x_0, \lambda_0) \in W_i$.

2) If C is an $n + k$ manifold with boundaries and corners so $W(f)$ intersects ∂C transversely, then, there is an open neighborhood $\mathcal{U}$ of u in $\mathcal{S}^{*\,(p)}_{n,\Lambda}$ (resp. $\mathcal{S}'^{\,(p)}_n$) such that for $u' \in \mathcal{U}$, $f' = u' * G$ exhibits $\mathcal{P}$ generically on C, and there is a strata preserving homeomorphism of C sending $W(f) \cap C$ to $W(f') \cap C$, which is smooth on each stratum.

The proof follows exactly that given for Theorems 5.10 and 5.11 using the Thom Isotopy Theorem.

Furthermore, there is an analogue of Theorem 6.3 for vector–valued distributions. To state it, we extend two definitions. By $G = (G_1, \ldots, G_p)$ satisfying condition (A), we mean that each G_i satisfies condition (A). Likewise, we give an analogue of Definition 6.2

Definition 11.9. A subspace $\mathcal{T} \subset \mathcal{S}'^{\,(p)}_n$ is said to satisfy condition (B) if (as for Definition 6.2): 1) $\mathcal{T}$ is closed under nonnegative linear combinations; and 2) $\mathcal{T}$ contains the space of smooth functions in $C^\infty(\mathbb{R}^n \times \Lambda, \mathbb{R}^p)$ which have compact support and whose coordinate functions are nonnegative.

For example, given subspaces $\mathcal{T}_i \subset \mathcal{S}^*_{n,\Lambda}$ (or $\mathcal{S}'_n$) which satisfy condition (B), then $\mathcal{T} = \mathcal{T}_1 \times \cdots \times \mathcal{T}_p$ also satisfies condition (B). Then, the analogue of Theorem 6.3 is the following.

Theorem 11.10. *Suppose that G satisfies condition (A) and $\mathcal{T}$ satisfies condition (B). Let $\mathcal{P}$ be a scale–based property defined by transversality to a closed Whitney stratified subset $W \subset \mathcal{H}^{\ell,(p)}_{\Lambda, G}$ (resp. $\mathcal{H}^{\ell,(p)}_{G}$). Then, on any compact subset $C \subset \mathbb{R}^n \times \Lambda$, $\mathcal{P}$ is generic for G–convolutions of distributions in $\mathcal{T}$. The local generic structure of $W(u * G)$ is given on C by $\{W_{loc}(z)\}$; and $W(u * G)$ is structurally stable on C under sufficiently small perturbations of $u \in \mathcal{T}$.*

Proof. As each G_i satisfies condition (A), by Proposition 6.6, there is a finite dimensional subspace $\mathcal{V}_i$ spanned by smooth nonnegative functions of compact support whose convolution with G_i composed with $j^\ell(\cdot)(x_0, \lambda_0)$ maps submersively onto $\mathcal{H}^{\ell,(p)}_{\Lambda, G\,(x_0,\lambda_0)}$ (resp. $\mathcal{H}^{\ell,(p)}_{G\,(x_0,\lambda_0)}$). Let $\mathcal{V} = \prod_i \mathcal{V}_i$, which is a subspace of $\mathcal{T}$ by condition 2) of 11.9. Then, we repeat the argument given for Theorem 6.3 to show that $\mathcal{V}$ satisfies the condition of smooth image for the relative transversality theorem. This implies the desired genericity. The remaining results follow as earlier. $\square$

Remark 11.11. Although we do not explicitly state them, there are also analogues of Corollary 6.5 if e.g. $\mathcal{T} = \prod \mathcal{T}_i$ where each $\mathcal{T}_i$ is a space of probability measures, provided that W is invariant under the individual actions of scalar multiplication on each corrdinate. Also, there are analogues of the results of §7 for discrete vector–valued functions and measures.

12. Multi-Feature Geometry

In the preceding section we established that independent tempered distributions can simultaneously generically exhibit distinct geometric properties genericially on a compact subset C of scale space. We now consider a related question for a single tempered distribution. Suppose we associate to a single distribution a finite collection of distributions. For example, we might consider a distinct collection of features/textures which can be discriminated using a collection of filters or statistical measures. This associates a collection of distributions, each capturing the amount of one of these features, textures, etc. present in a given region of space. We then would like to deduce the simultaneous geometric properties of the collection of features.

In this section, we show that given a compact subset C of scale space, provided the collection of features/textures are "sufficiently independent"on a subspace $\mathcal{T} \subset \mathcal{S}^*_{n,\Lambda}$ (or $\mathcal{S}'_n$), then, the collection of geometric properties of features will be simultaneously exhibited generically on C for a dense open subset of (uniformly) tempered distributions in $\mathcal{T}$.

To explain this independence, we consider a collection of continuous operators

$$\Psi_i : \mathcal{T} \to \mathcal{S}^*_{n,\Lambda} \quad \text{or} \quad \Psi_i : \mathcal{T} \to \mathcal{S}'_n \tag{12.1}$$

which discriminate various features or textures, as given in §10. We define an associated operator $\Psi = (\Psi_1, \ldots, \Psi_p)$

$$\Psi : \mathcal{T} \to \mathcal{S}^{*\,(p)}_{n,\Lambda} \quad \text{or} \quad \Psi : \mathcal{T} \to \mathcal{S}'^{\,(p)}_n \tag{12.2}$$

Definition 12.1. Given a scale–based multi-kernel G, and a compact subset of scale space C, we say the *collection of operators* Ψ_i discriminating features, textures, etc. *are generically independent on* C for the subspace $\mathcal{T}$, if there is an open neighborhood U of C so the composition of the associated operator $\Psi = (\Psi_1, \dots, \Psi_p)$ with convolution c_G (resp. $c_{\Lambda, G}$) and restriction to U has smooth image, with image equal to $\mathcal{H}^{\ell,(p)}_G(U)$ (resp. $\mathcal{H}^{\ell,(p)}_{\Lambda, G}(U)$).

We consider properties $\mathcal{P}$ defined for smooth functions in $\mathcal{H}^{(p)}_{\Lambda, G}$ (resp. $\mathcal{H}^{(p)}_G$), defined by transversality to a closed Whitney stratified subset W of $\mathcal{H}^{\ell,(p)}_{\Lambda, G}$ (resp. $\mathcal{H}^{\ell,(p)}_G$). Such a property might, for example, denote the simultaneous occurrence of p distinct properties for the distinct coordinate functions as in Example 11.4. For (uniformly) tempered distributions u, we ask whether for a collection of p features/textures defined by operators Ψ_i,the features will simultaneously exhibit the property $\mathcal{P}$.

Proposition 12.2. *For a scale–based multi-kernel $G = (G_1, \dots, G_p)$, suppose the property $\mathcal{P}$ is a transversally defined scale–based property. Let Ψ_i be continuous operators discriminating features, textures, etc. as in (12.1). If C is a compact subset of scale space, and the Ψ_i are generically independent on C for the subspace $\mathcal{T}$, then there is an open dense subset of $\mathcal{T}$ consisting of u for which the G–convolution $f = \Psi(u) * G$ will exhibit $\mathcal{P}$ generically on C. Furthermore, $W(f)$ has generic local properties given by $\{W_{i\,loc}\}$ and is stable under any sufficiently small perturbation of u in $\mathcal{T}$.*

Proof. This proof virtually repeats the proof of Proposition 10.2, using the relative transversality theorem. If $i : \mathcal{T} \hookrightarrow \mathcal{S}'_n$ (or $\mathcal{S}^*_{n,\Lambda}$) denotes the inclusion map, and r restriction to U, the composition $r \circ c_G \circ \Psi \circ i$ (or with $c_{\Lambda, G}$ in place of c_G) is continuous. The condition that the operators Ψ_i are generically independent asserts that the composition has smooth image with image $\mathcal{H}^{\ell,(p)}_{\Lambda, G}(U)$ (resp. $\mathcal{H}^{\ell,(p)}_G$(U)). The relative transversality theorem implies that for the compact subset $C \subset U$, the set of $u \in \mathcal{T}$ such that $j^\ell(\Psi(u)*G)$ is transverse on C to W relative to $\mathcal{H}^\ell_{\Lambda, G}$ (resp. $\mathcal{H}^\ell_G$) is an open dense subset of $\mathcal{T}$. Then, the local generic structure and the stability on compact subsets follow exactly as earlier. □

Generic Independence of Multifeature Properties. As in §10, we will apply Proposition 12.2 to feature–detection operators Ψ_i which are linear. These occurred in the three situations considered in §10. We examine the form that generic independence takes for each of these classes of operators.

Ψ_i defined by Partial Differential Operators. Suppose each $\Psi_i = P_i(D)$, $i = 1, \dots, p$ are linear differential operators with constant coefficients. By Corollary 10.6, we know that each operator individually will generically yield distributions with generic scale-based geometric properties. THe simultaneous behavior of the set of operators depends very heavily on the operators themselves. In particular, the solution spaces for $P_i(D)(u) = 0$.

Corollary 12.3. *Let G is a scale–based multi kernel satisfying condition (A), and let C be a compact subspace of scale space. Suppose Ψ is an operator as in (12.2) defined via $\Psi_i = P_i(D)$, which are linear partial differential operators with constant coefficients. Let $\mathcal{T}_i = \cap_{j \neq i} \ker(P_i(D))$. If $\Psi_i | \mathcal{T}_i$ satisfies condition (B_C) for each $i = 1, \dots, p$, then the Ψ_i are generically independent on C. Hence, the conclusions* *of Proposition 12.2 apply to yield scale–based genericity and stability for* $\Psi(u) * G$ *for* $u \in \mathcal{T}$.

Generic Independence by approximating "block functions". Second, we consider sufficient conditions for generic independence by approximating "block functions".

Suppose the scale–based multi–kernel G satisfies condition (A). Let C be a compact subset of scale space. Given a mesh $\mathcal{B}$, by Proposition 10.3, we may find for each G_i a finite dimensional subspace $\mathcal{V}_i$ of discrete functions of compact support relative to a refinement $\mathcal{B}_i$ of $\mathcal{B}$ which satisfy condition (B_C). We may then choose a common refinement $\mathcal{B}'$ of the $\mathcal{B}_i$, and assume that each of the finite dimensional subspaces is defined on $\mathcal{B}'$. The combined support is restricted to a finite subset of cells B_j of $\mathcal{B}'$. The space $\mathcal{V}_i$ is spanned by a finite set $\varphi_j^{(i)} = \sum a_{ijk}\chi_k$, where as earlier χ_k is the characteristic function for B_j (or a "block function"). By 3) of Proposition 10.3, we can find a common $\varepsilon > 0$ so that if for any $i = 1, \dots, p$, we replace $\{\varphi_j^{(i)}\}$ by another set of functions $\{\varphi_j^{(i)\,\prime}\}$ each within ε in, say, the (essential) sup–norm, then this set will still satisfy the surjectivity condition in Proposition 10.3.

Then, we can repeat the argument given for Corollary 10.7 for each set of functions spanning each $\mathcal{V}_i$ to conclude that there is an $\varepsilon' > 0$ so that:

if there exist for each support cell B_j, block function χ_j and each $i = 1, \dots, p$, there is a $u_{ij} \in \mathcal{T}$ such that either $\Psi_i(u_{ij})$ is within ε' of χ_j, or $\Psi_i(u_{i'j})$ is within ε' of 0 for $i' \neq i$ in the (essential) sup–norm), then

$v_{ij} = \sum a_{ijk}u_{ik}$ maps via Ψ_i to $\varphi_j^{(i)\,\prime} = \sum a_{ijk}\Psi(u_{ik})$ which is within ε of $\varphi_j^{(i)}$; while $\Psi_i(v_{i'j})$ is within ε of 0 (in the (essential) sup–norm).

We say that the $\Psi = (\Psi_1, \dots, \Psi_p)$ satisfying the above condition *independently approximate block functions* to within ε' on the common support cells B_j.

Corollary 12.4. *Suppose G is a scale–based multi–kernel kernel satisfying condition (A), and C is a compact subspace of scale space. There is a refinement $\mathcal{B}'$ of the mesh $\mathcal{B}$, a finite subset of common support cells $\{B_j, j = 1, \dots, m\}$ (of $\mathcal{B}'$) and an $\varepsilon' > 0$ so that if Ψ independently approximates the block functions on the $B_j, j = 1, \dots, m$ to within ε' in the (essential) sup–norm, then Ψ is generically independent on C, and conclusions of Proposition 12.2 apply to yield scale–based genericity and stability for $\Psi(u) * G$ for $u \in \mathcal{T}$.*

Again we note we could have viewed the simple measures associated to the $\varphi_j^{(i)}$ and instead considered measures approximating them, and obtained an analogous condition.

Generic Geometric Properties from Multiple Texture Discrimination. The preceding Corollary can be applied as in the case of simple texture discrimination, to deduce the genericity of geometric properties associated to multiple texture discrimination.

Again let G be a scale–based multi–kernel satisfying condition (A), and let C be a compact subset of scale space. By the discussion leading up to Corollary 12.4, there is a refinement $\mathcal{B}'$ of $\mathcal{B}$, a finite set of cells of common support $B_j, j = 1, \dots, m$ and $\varepsilon' > 0$ so that if Ψ independently approximates the block functions on the $\{B_j\}$ to within ε' in the (essential) sup–norm, then the Ψ_i are generically independent on C.

Suppose that each Ψ_i provides simple texture discrimination via (10.3). Thus, for each $i = 1, \ldots, p$ and each B_j, there is a smooth mask function K_{ij} with support in B_j. By the mask functions K_{ij} being independent on each B_j, we mean there are tempered distributions u_{ij} with support in B_j so that $(u_{\ell j}(K_{mj})$ is a nonsingular $p \times p$–matrix. This provides a sufficient condition that the simple texture discrimination operators Ψ_i are generically independent.

Corollary 12.5. *Suppose G is a scale–based multi–kernel satisfying condition (A), $\mathcal{T}$ is a linear subspace of $\mathcal{S}'_n$, and C is a compact subspace of scale space. Let $\mathcal{B}'$ be the refinement of the mesh $\mathcal{B}$, with finitely many support cells $\{B_j, j = 1, \ldots, m\}$ and let $\varepsilon' > 0$ be given as above. If each Ψ_i denote simple texture discrimination with a set of mask functions K_{ij} which are independent for each suport cell B_j, then the Ψ_i are generically independent on C. Hence, the conclusions of Proposition 12.2 apply to yield scale–based genericity and stability for $\Psi(u) * G$ for $u \in \mathcal{T}$.*

Proof. To apply Corollary 12.4, we need only demonstrate that Ψ independently approximates the block functions on the $\{B_j\}$. However, if $(u_{\ell j}(K_{mj}))$ is nonsingular, then we can find linear combinations $u'_{\ell j}$ of the $u_{\ell j}$ so that $(u'_{\ell j}(K_{mj}) = \delta_{\ell m}$, giving the desired result. □

Example 12.6. We apply the preceding to the simple detection of horizontal and vertical stripes textures. Suppose that we choose a square mesh with sides of lengths $\frac{1}{N}$. Let H_{ij} be the function which on the cell

$$B_{ij} = \{(x, y) : \frac{i}{N} \leq x < \frac{i+1}{N}, \frac{j}{N} \leq y < \frac{j+1}{N}\}$$

equals $\sin(mN\pi y)$, and 0 on the other cells. We also define V_{ij} to be the function which on the cell B_{ij} equals $\sin(mN\pi x)$, and equals 0 on the other cells. H_{ij} is a mask function for a texture of horizontal stripes on B_{ij} and V_{ij} is a mask function for a texture of vertical stripes on B_{ij} (fig. 7) . Although they are not smooth we can either multiply them by a product of smooth bump functions which approximate the characteristic functions of the intervals. Alternately, we can still apply many classes of tempered distributions directly to them.

FIGURE 7. Striped Textures represented by positive and negative values of H_{ij} and V_{ij} on cell of a square mesh for $\mathbb{R}^2$

It is easily checked that H_{ij} and V_{ij} are orthogonal. Hence, suppose we are interested in medial or edge–based properties of regions exhibiting these textures to some degree. Scale space will introduce a single scale parameter σ for either kernel. A compact region C of scale space implies a bounded region of physical space together with bounds on the scale values. Then, Corollary 12.5 implies that based on these bounds, there is a sufficiently large integer N so that the pairs of simple horizontal and vertical stripe texture discrimination on the square mesh of sides $\frac{1}{N}$, will simultaneously exhibit generic geometric properties for both textures on C, with the standard local generic properties and stability under small perturbations.

Example 12.7. Suppose we consider the textures consisting of 45 and 135–degree line segments "\ "and "/ ".as shown in figure 8 taken from Mumford [Mu, Chap. 4]. These textures define the letter "Z". One way to distinguish the textures is

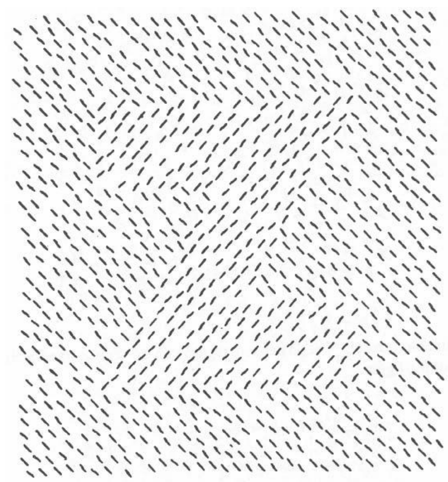

FIGURE 8. Pair of textures defining letter Z taken from Mumford [Mu]

to define measures of the amount of each segment type in an Borel set U, by e.g. assigning the sums of the lengths of the segments intersected with U. Alternately, we could distinguish the textures using the simple method already described, by rotating the square mesh 45–degrees. Then, the mask functions become rotated to new functions H'_{ij} and V'_{ij} as in fig. 9.

FIGURE 9. Striped Textures represented by positive and negative values of H'_{ij} and V'_{ij} on cells of a rotated square mesh for $\mathbb{R}^2$

Provided the frequence m is chosen sufficiently fine, the Mask functions will detect significant proportions of the segments of one type and give essentially zero response to the other type. Thus, by Corollary 12.5, for a fine enough mesh we obtain discrimination functions which will generically have scale–based geometric properties. For example, we illustrate in fig. 10 how the scale-based relative critical set detects the medial properties of the Z and its complement. The scale–based edges will detect the outline of the Z. They will look somewhat as shown. There will be corresponding complementary relative critical set for the 135° discrimination function. For the relative critical set (above a certain threshold) we see both the ridges along the center of the Z, dashed connector sets continuing where the ridge may break, and other connector and valley curves at higher scale where there is an absence of "\ "segments. However, the edges for each discrimination function will be approximately the same in this case. We would obtain approximately the same geometric features if we had used discrimination measures instead.

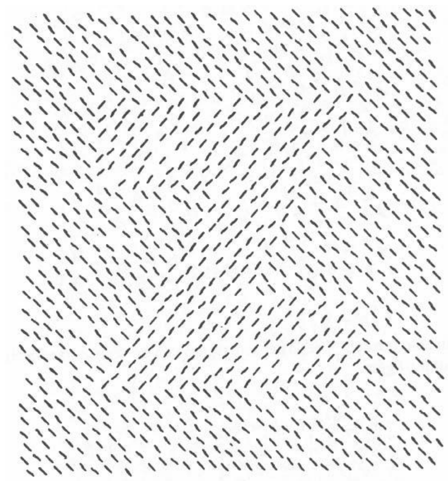

FIGURE 10. Relative Critical Sets and Edges for texture discrimination for the letter Z

Part 4. Scale–based Properties for Parametrized Families (In Preparation)

13. Generic Transitions in Families via $\mathcal{K}_V$-equivalence

14. Scale Parameters as Distinguished Parameters

15. Generic Transitions for Scale-based Kernels with External Parameters

16. Generic Transitions for Geometric Properties for Families of Tempered Distributions

References

[Al] S. Aylward *Continuous Gaussian Mixture modeling*, Information Processing for Medical Imaging, ed. J. Duncan (1997).

[Ba] G.I. Barenblatt *Scaling, self–similarity, and intermediate asymptotics*, Cambridge Texts in Appl. Math., Camb. Univ. Press, 1998.

[Bg] I. Bogaevski *Metamorphoses of singularities of minimum functions and bifucations of shock waves of the Burgers equation with vanishing viscosity*, St. Petersburg (Leningrad) Math. Jour. **1 no. 4** (1990), 807–823.

[BN] H. Blum and R. Nagel *Shape description using weighted symmetric axis features*, Pattern Recognition **10** (1978) 167–180.

[BGG] J.W. Bruce, P.G. Giblin, and C.G. Gibson *Symmetry sets*, Proc. Roy. Soc. Edinburgh **101A** (1983), 163–186

[BG] J. W. Bruce and P. J. Giblin *Growth, motion, and 1–parameter families of symmetry sets*, Proc. Roy. Soc. Edinburgh **104A** (1986), 179–204

[BGT] J.W. Bruce, P.G. Giblin, and F. Tari *Ridges, crests, and subparabolic lines of evolving surfaces*, Int. Jour. Comp. Vision **18(3)** (1996), 195–210.

[Brz] L.N.. Bryzgalova *Singularities of the maximum of a function that depends on the parameters*, Funct. Anal. and Appl. **11** (1977), 49–51.

[Ca] J. Canny *A computational approach to edge detection*, IEEE PAMI **8 (6)** (1987), 679–698.

[D1] J. Damon *Local Morse theory for solutions to the heat equation and Gaussian blurring*, Jour. Diff. Eqtns **115** (1995), 368–401.

[D2] ——— *Generic properties of solutions to partial differential equations*, Arch. Rat. Mech. Anal. **140** (1997), 353–403.

[D3] ——— *Properties of ridges and cores for two dimensional images*, Jour. Math. Imag. and Vision **10** (1999), 163–174.

[D4] ——— *Generic structure of two dimensional images under Gaussian blurring*, SIAM Jour. Appl. Math. **59** (1999), 97–138.

[D5] ——— *Singularities with scale threshold and discrete functions exhibiting generic properties*, Proc. Third Int. Workshop on Real and Complex Singularities, Sao Carlos, M. Ruas, ed., Soc. Bras. Mat., Mat. Contemp. **12** (1997), 45–65.

[D6] ——— *Deformations of sections of singularities and Gorenstein surface singularities*, Amer. J. Math. **109** (1987), 695–722.

[D7] ——— *Generic medial geometry*, in preparation.

[Db] I. Daubechies *Ten Lectures on Wavelets*, CBMS–NSF Conference Series **61** SIAM, 1992.

[Eb1] D. Eberly et al *Ridges for image analysis*, Jour. Math. Imaging and Vision **4** (1994), 351–371.

[Eb2] D. Eberly *Ridges in Image and Data Analysis*, Series in Computational Imaging and Vision Series, Kluwer Academic Publishers, Doordrecht, the Netherlands, 1996.

[Fo] G. B. Folland *Introduction to Partial Differential Equations*, Princeton Math. Notes, Princeton Univ. Press, 1976.

[Fo2] G. B. Folland *Lectures on Partial Differential Equations*, Tata Institute Lect. Notes 70, Tata Institute and Springer Verlag, 1983.

[Fu] J. Furst, J. Miller, R. Keller, S. Pizer *Image Loci are ridges in geometric spaces*, in Scale Space Theory in Computer Vision, ed. B. M. ter Haar Romeny et al, Springer Lecture Notes in Computer Science **1252** (1997), 176–187.

[Fu2] J. Furst, *Height Ridges of Oriented Medialness* Thesis Dept. Comp. Sci. Univ. North Carolina (1998).

[FPM] D. Fritsch, S. Pizer, B. Morse, D. Eberly, A., Liu, *The multiscale medial axis and its applications in image registration*, Pattern Recognition Letters **15** (1994), 445-452.

[Fr] D. Fritsch, *Optimal Scale Height Ridges (?)* Thesis Dept. Comp. Sci. Univ. North Carolina (199?).

[Ga] M. Gage *Curve shortening makes convex curves circular*, Invent. Math. **76** (1984), 357-364

[GH] M. Gage and R. Hamilton *The heat equation shrinking convex plane curves*, Jour. Diff. Geom. **23** (1986), 69–96

[Gau] J. Gauch , Thesis Dept. Comp. Sci. Univ. North Carolina (1988).

[GS] I. M. Gelfand and G. E. Shilov. *Generalized Functions vol I-V*, Academic Press, Orlando, U.S.A., 1968.

[GMc] M. Goresky and R. Macpherson *Stratified Morse Theory*, Ergebnisse der Math. und ihrer Grenzgebiete, Springer Verlag, Berlin–Heidelberg, 1988.

[Gi] C.G. Gibson, et al. *Topological stability of smooth mappings*, Springer Lecture notes in Math. **552**, Springer–Verlag, 1976.

[GG] M. Golubitsky and V. Guillemin *Stable Mappings and their Singularities*, Springer Graduate Texts in Mathematics, Springer-Verlag, Berlin–Heidelberg, 1974.

[GZ] V. Goryunov and V. Zakalyukin , preprint.

[Gr] M. Grayson *The heat equation shrinks embedded curves to round points*, Jour. Diff. Geom. 26 (1987), 285–314

[H] L. Hormander *The Analysis of Linear Partial Differential Operators I*, Grundlehren der Mathematischen Wissenschaften Band **256**, Springer Verlag , Berlin–Heidelberg, 1983.

[Ju] B. Julesz *Visual pattern discrimination*, IRE Transactions of Information Theory, IT **8** (1962), 84-92

[Ke] R. Keller *Generic Transitions of Relative Critical Sets in Parametrized Families with Applications to Image Analysis*, Ph.D. Thesis, Dept. Mathematics, Univ. of North Carolina, 1999.

[K] J.J. Koenderink *The Structure of Images*, Biological Cybernetics **50** (1984), 363–370.

[K2] ——— *Solid Shape*, M.I.T. Press, Cambridge, MA, 1990.

[KTZ] B. B. Kimia, A. Tannenbaum, and S. Zucker *Toward a computational theory of shape: An overview*, O. Faugeras ed., Three Dimensional Computer Vision , M I T Press, 1990.

[Li] T. Lindeberg *Scale Space Theory in Computer Vision*, Series in Engineering and Computer Science, Kluwer Academic Publishers, Doordrecht, the Netherlands, 1994.

[Lo] S. Lojasiewicz, *Ensembles Semi–Analytiques*, Lecture Notes I.H.E.S.,1965.

[Mal] S. Mallat *A Wavelet Tour of Signal Processing*, Academic Press, Boston, MA,1998.

[M] J. Mather *Stability of C^∞-mappings:*

[M-II] ——— *II. Infinitesimal Stability Implies Stability*, Ann. Math. **89** (1969), 254–291.

[M-V] ——— *V. Transversality*, Advances in Math. **37** (1970), 301-336.

[M-VI] ——— *VI. The nice dimensions*, Liverpool Singularities Symposium I, Springer Lecture Notes in Math. **192** (1970), 207–253.

[M2] ——— *Stratifications and mappings*, in Dynamical Systems, M. Peixoto, Editor, (1973) Academic Press, New York

[M3] ——— *How to stratify mappings and jet spaces*, in Singularités d'Applications Differentiables, Plans-sur-Bex, Springer Lecture Notes **535** (1975), 128-176.

[M4] ——— *Distance from a manifold in Euclidean space*, in Proc. Symp. Pure Math. vol **40** Pt 2 (1983), 199–216

[Mi] J. Miller *Relative Critical sets in $\mathbb{R}^n$ and Applications to Image Analysis*, Ph.D. Thesis, Dept. Mathematics, Univ. of North Carolina, 1998.

[Mo] J. M. Morel *Filtrage Itératif des images et équations aux dérivées partielles*, Notes du Cours du Centre Emile Borel **12**, 1998.

[MPP] B. Morse , S. Pizer, et al *Zoom–Invariant Vision of Figural Shape: Effects on cores of Image disturbances*, Comp. Vision and Image Understanding (199?), to appear

[MPL] B. Morse , S. Pizer, and A. Liu *Multiscale medial axis in medical images*, Image and Vision Computing **12** (1993), 112–131.

[Mu] D. Mumford *Pattern Theory and Vision*, Notes du Cours du Centre Emile Borel, 1998.

[NaP] L. R. Nachman and S. Pizer *Three Dimensional shape description using the symmetric axis transform I; Theory*, IEEE PAMI **7(2)** (1985), 187–202

[NO] M. Nielsen and O. F. Olsen *Segmentation via Watersheds*,

[OS] S. Osher and J. A. Sethian *Fronts propagating with curvature dependent speed: Algorithms based on Hamilton-Jacobi formulation*, Jour. Comp. Physics **9** (1988), 12–49

[PE] S. Pizer, D. Eberly et al. *Core–Invariant Vision of Figural Shape: The Mathematics of Cores*, Comp. Vision and Image Understanding **69** (1998), 55–71.

[PEM] S. Pizer, D. Eberly et al. *Axiomatic Zoom–Based Vision*, (199?)

[Po1] I. Porteous *Geometric Differentiation* Camb. Univ. Press, 1994.

[Po2] ——— *Normal singularities of a submanifold*, Jour. Diff. Geom. **5** (1971), 543–564.

[Ri1] J. Rieger *Generic properties of edges and corners on smooth greyvalue surfaces*, Biol. Cybernet. **66** (1992), 497–502.

[Ri2] ——— *Generic evolution of edges on families of diffused greyvalue surfaces*, Jour. Math. Imaging and Vision **5** (1995), 207–217.

[Ri3] ——— *Topographical properties of greyscale images*, Int. Jour. Comp. Imaging **23** (1997), 79–92.

[Ri4] ——— *Hypersurfaces of extremal slope*, (1999), preprint.

[Rg] F. Ronga *A new look at Faa' de Bruno's formula for higher derivatives of composite functions and the expression of some intrinsic derivatives*, Proc. Symp. Pure Math. **vol 40 Pt 2** (1983), 423–431.

[Ru1] W. Rudin *Real and Complex Analysis*, McGraw–Hill, New York, 1966.

[Ru2] ——— *Functional Analysis*, McGraw–Hill, New York, 1973.

[tHR] B. M. ter Haar Romeny (edit.) *Geometry Driven Diffusion in Computer Vision*, Kluwer Academic Publishers, Dordrecht, the Netherlands, 1994.

[RFSV] B. M. ter Haar Romeny, et al. *Higher Order Differential Structure of Images*, Image and Vision Computing **12 (6)** (1994), 317–325

[Sch] L. Schwartz *Théorie des Distributions*, Hermann Publishers, Paris, France, 1966.

[Sn] J. A. Sethian *Level Set Methods*, Monographs Appl. and Comp. Math., Camb. Univ. Press (1996)

[Sm] S. Smale *On Gradient Dynamical Systems* , Annals Math. bf 74 (1961),199–206

[Sp] J. Sporring et al (edit.) *Gaussian Scale Space Theory*, Series in Computational Imaging and Vision Series, Kluwer Academic Publishers, Dordrecht, the Netherlands, 1997.

[Th] R. Thom *Ensembles et morphismes stratifiés*, Bull. Amer. Math. Soc. **75** (1969), 240–284.

[Wa] C. T. C. Wall *Geometric properties of generic differentiable manifolds*, Springer Lect. Notes Math. **592** (1977), 707–774

[WII] J. Weickert, S. Ishikawa, and A. Imiya *On the history of Gaussian Scale Space Axiomatics*, in Gaussian Scale Space Theory, ed. J. Sporring et al, Kluwer, Doordrecht, (1997), 45–58.

[Wi] A. Witkin *Scale Space Filtering*, Proc. Int. Joint Conf. on Artifical Intelligence Karlsruhe (1983), 1019–1021.

[Y] J. Yomdin *On the local structure of the generic central set*, Compositio. Math. **43** (1981), 225–238

[ZWM] S.C. Zhu, Y.N. Wu, and D. Mumford *Minimax Entropy Principle and its Application to Texture Modeling*, Neural Comp. **9** (1997)

Department of Mathematics, University of North Carolina, Chapel Hill, NC 27599-3250, USA