# On the Behavior of Spatial Critical Points under Gaussian Blurring

## A Folklore Theorem and Scale-Space Constraints

Marco Loog[1], Johannes Jisse Duistermaat[2], and Luc M.J. Florack[2]

[1] Image Sciences Institute, University Medical Center
Utrecht, The Netherlands
marco@isi.uu.nl
[2] Department of Mathematics, Utrecht University
Utrecht, The Netherlands

**Abstract.** The main theorem we present is a version of a "Folklore Theorem" from scale-space theory for nonnegative compactly supported functions from $\mathbb{R}^n$ to $\mathbb{R}$. The theorem states that, if we take the scale in scale-space sufficiently large, the Gaussian-blurred function has only one spatial critical extremum, a maximum, and no other critical points.
Two other interesting results concerning nonnegative compactly supported functions, we obtain are:

1. a sharp estimate, in terms of the radius of the support, of the scale after which the set of critical points consists of a single maximum;
2. all critical points reside in the convex closure of the support of the function.

These results show, for example, that all catastrophes take place within a certain compact domain determined by the support of the initial function and the estimate mentioned in 1.
To illustrate that the restriction of nonnegativity and compact support cannot be dropped, we give some examples of functions that fail to satisfy the theorem, when at least one assumption is dropped.

**Keywords and Phrases.** Large-scale behavior, loss of detail, nonnegative function, compact support, spatial critical point, deep structure.

## 1 Introduction

In this paper we discuss a "Folklore Theorem" from scale-space that could be roughly stated as follows: if a function $f : \mathbb{R}^n \to \mathbb{R}$ is blurred sufficiently, then the blurred function has a single critical point, which is an extremum. The intuitive idea behind this theorem is appealing: if scale is taken sufficiently large, every detail of the function will be lost (as scale increases "images become less articulated", one sees the "erosion of structure" as Koenderink calls it [7]), because they are too small with respect to the scale. All that is left is a blurred function virtually indistinguishable from a single Gaussian blob. This theorem, however, is false. *The principal aim of this article* is the formulation of certain

M. Kerckhove (Ed.): Scale-Space 2001, LNCS 2106, pp. 183–192, 2001.

restrictions on the functions considered, that leads to the Folklore Theorem to hold.

As we will see in Subsection 3.1, the falseness is already shown by a simple example like the $n$th order ($n > 0$) derivative of a one-dimensional Gaussian, or an arbitrary periodic function.

The theorem we present (Subsection 3.2) states that for nonnegative functions $f : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ (where $\mathbb{R}_{\geq 0}$ stands for the set of nonnegative real numbers) with compact support, it *does* hold that $f_\sigma$, which is $f$ at scale $\sigma$, has only one critical point for all $\sigma$ larger than a certain scale $\varsigma$. To show that the restriction of nonnegativity or compact support cannot be dropped in general, we discuss in Subsection 3.3 some examples which illustrate this. Especially the example in which we retain the compact support, but drop the nonnegativity is counter-intuitive and an interesting observation.

A further result we present is a sharp estimate of the scale $\varsigma$, in terms of the radius of the support, after which we certainly have a unique maximum. This, together with a result presented in Section 4, which states that all spatial critical points of a blurred function reside in the convex closure of the support of the initial function, gives us a certain bounded domain within scale-space to which attention can be restricted, when for example, tracking down catastrophes [1], finger-prints [9], or other potentially interesting features [3].

Most of the results are presented and proven in quite a formal way. However, the proofs of these results are postponed to Appendix A to keep the text more readable. The next section gives some definitions and notations, which are used in the remainder of the article. Section 5 provides the discussion and the conclusions.

## 2 Definitions and Notations

We start with the formal definitions of the support of a function, and the radius of the support.

**Definition 1.** *The support* supp$(f)$ *of a function* $f : \mathbb{R}^n \to \mathbb{R}$ *is*

$$\mathrm{supp}(f) := \overline{\{x \in \mathbb{R}^n | f(x) \neq 0\}},$$

*where the line indicates that the closure of the set is taken. The radius* $\mathrm{r}(f)$ *of the support of the function* $f$*, is defined as the smallest radius of a ball containing* supp$(f)$.

We also define Gaussian blurring of a function from which scale-space is generated.

**Definition 2.** *Let the Gaussian kernel* $g_\sigma : \mathbb{R}^n \to \mathbb{R}$*, for* $\sigma \in \mathbb{R}_{\geq 0}$*, be defined as:*

$$g_\sigma(x) := (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{\|x\|^2}{2\sigma^2}}.$$

*Furthermore, given a function* $f : \mathbb{R}^n \to \mathbb{R}$*, let the function* $f_\sigma : \mathbb{R}^n \to \mathbb{R}$ *be defined as:*

$$f_\sigma(x) := (f * g_\sigma)(x) = \int_{\mathbb{R}^n} f(y)\, g_\sigma(x-y)\, dy,$$

*to which we refer as the blurred function $f_\sigma$ at a fixed scale $\sigma$. Scale-space is defined as the complete family of functions $f_\sigma$, with $\sigma \in \mathbb{R}_{\geq 0}$.*

# 3 Large-Scale Behavior of Critical Points

This section presents the theorem that states that for nonnegative functions $f$ with compact support, it holds that $f_\sigma$ has only one critical point for all $\sigma$ larger than a certain scale $\varsigma$. However, we first discuss shortly two simple examples, which show that the theorem cannot hold for general (integrable) functions.

## 3.1 Arbitrary Number of Critical Points

*Example 1.* Derivatives of the normalized Gaussian function define an autoconvolution algebra, i.e., if $D_k$ denotes a partial derivative operator with multi-index order, $k = (k_1, \ldots, k_n)$, then $D_k g_\sigma * D_l g_\tau = D_{k+l} g_{\sqrt{\sigma^2+\tau^2}}$. In particular this implies that derivatives of a normalized Gaussian are, topologically speaking, blur invariant, hence their critical points are preserved regardless of the amount of blurring.

*Example 2.* As a second example, note that $\sin(\langle \omega, x \rangle)$ and $\cos(\langle \omega, x \rangle)$, with $\omega, x \in \mathbb{R}^n$ are eigenfunctions under Gaussian blurring with eigenvalue $e^{-\frac{1}{2}\sigma^2 \|\omega\|^2}$. Thus the same conclusion can be drawn regarding their (infinite number of) critical points as in the previous example: the number of critical points remains equal for all scales and is infinite. Cf. [6], for an analysis of the behavior of the extrema at large scale, in which functions are considered that are periodic, band-limited, and one-dimensional.

Clearly, in general, the Folklore Theorem does not hold. Moreover, both examples in this section show that we can construct functions with an arbitrary number, ranging from one to infinite, of extrema at all scales. Hence, we look for restrictions on a function, that do lead to the behavior that only a single large-scale extremum exists beyond scale, say, $\varsigma$. As Example 1 shows, requiring the function to satisfy: $\lim_{\|x\| \to \infty} f(x) = 0$ is not enough; it is still possible that an arbitrary number of critical points coexist to an arbitrarily large scale.

## 3.2 Nonnegative and Compactly Supported Functions

If we restrict the functions to compactly supported functions, it seems reasonable to expect the Folklore Theorem. Because the function is compactly supported, its support is bounded and has a maximal radius, say $r$, so, one expects that for scales that are large enough – proportional to $r$ – there is not much more left of the blurred function then something similar to a Gaussian, in which one can no longer distinguish any kind of detail. However, as we will see in Subsection 3.3, this restriction is still too weak to make the theorem hold. We need an extra restriction, which is the requirement that the function is nonnegative.

**Theorem 1.** *Given a nonnegative function* $f : \mathbb{R}^n \to \mathbb{R}$ *with compact support. If* $\mathrm{r}(f)$ *equals* $r$*, then* $f_\sigma$ *has exactly one critical point for every* $\sigma > r$*. This unique critical point is a maximum.*

Refer to Appendix A for the proof.

*Remark 1.* Without further proof, we mention that the foregoing estimate is sharp, in the sense that given a support with radius $r$, it is possible to construct a nonnegative, compactly supported function $f$ such that $f_\sigma$ has multiple extrema for all $\sigma$ arbitrarily close to $r$. The idea is to construct a nonnegative function $f$ that is compactly supported with $\mathrm{r}(f) = r$, and comes arbitrarily close to the sum of two Dirac functions on $\mathbb{R}^n$ separated by a distance equal to $2r$.

The result stated in Theorem 1, in combination with Remark 1, suggests that we should call $2\sigma$, and not just $\sigma$, the scale of $f_\sigma$ at which $f$ is observed (cf. [4,7,8]).

*Remark 2.* Under much weaker assumptions, one has the following weaker result, which however still might be useful. The proof, of which we do not give the details here, starts with an estimate from below for the quantity $\int f(y)\, g_\sigma(x - y)\, dy$, which appears in the denominator of the mapping $F$ in the proof of Theorem 1. In the estimate, we use the assumption that $x \in B_{\delta\sigma}$.

**Theorem.** *Assume that* $\int_{\mathbb{R}^n} \|y\|^2 |f(y)|\, dy < \infty$ *and* $\int_{\mathbb{R}^n} f(y)\, dy > 0$ *(the case that* $\int_{\mathbb{R}^n} f(y)\, dy < 0$ *can be treated analogously). Let* $B_r$ *denote the ball with radius* $r$ *and center at the origin. Then there exists a* $\delta > 0$ *and a* $\sigma_0 > 0$ *such that for every* $\sigma > \sigma_0$ *the restriction of* $f_\sigma$ *to* $B_{\delta\sigma}$ *has precisely one critical point* $\xi(\sigma)$*, which is a maximum. Furthermore, if* $\sigma \to \infty$*, then* $\xi(\sigma)$ *converges to*

$$\frac{\int_{\mathbb{R}^n} y f(y) dy}{\int_{\mathbb{R}^n} f(y) dy}.$$

In the next subsection it is shown that the restrictions on the function cannot be dropped in general.

### 3.3 Nonnegative or Compactly Supported Functions

Dropping one of the requirements on the function in Theorem 1 can lead to examples of functions for which the theorem does not hold. We present three examples, which illustrate this. The first two are straightforward and are presented without rigorous proof. For the third one, it is demonstrated that the theorem does not hold.

Note that Examples 1 and 2 of Subsection 3.1 illustrate what can happen when both assumptions are dropped. The first and third example we present in this subsection do not require the initial function to be compactly supported. However, what we do require in the third example is that if $\|x\|$ goes to infinity that the value of the function goes to zero. In the second example we drop the nonnegativity assumption.

*Example 3.* As in Example 2, we consider periodic functions, however now we require the functions to be nonnegative. Hence, consider a bounded periodic function, and add a suitable constant so as to make it nonnegative. Note that a constant function is blur-invariant, and that its addition has no affect on critical points. Thus even positive functions may have multiple (in this case infinitely many) critical points that survive at all scales.

*Example 4.* In this example, we show that things can go wrong when dropping the nonnegativity assumption, but not the compactness requirement. We think that this result is quite counter-intuitive, because one would expect that if the support of the initial function is bounded, this restricts the size of the possible details that can be present in a function (e.g. an image) and so there should be a finite scale $\sigma$ for which $f_\sigma$ consists of one Gaussian-like blob. Hence $f_\sigma$ should have a unique extremum.

However, it is quite easy to construct compactly supported functions that do not satisfy Theorem 1. To do so, let $f$ be nonnegative, compactly supported, $n$-times differentiable, and one-dimensional. First, note that all derivatives of $f$ are also compactly supported. Now, because $f$ satisfies the assumptions of the theorem and hence has a single maximum for large scales, the $n$th-order derivatives still has, at least, $n+1$ extrema for large scales.

*Remark 3.* Most images reflect a certain physical spatial and/or temporal measurements, and most of these measured physical quantities are nonnegative, because there is some clear absolute zero. So, there seems to be, at least in practice, hardly any loss of generality in restricting the class of admissible functions to nonnegative functions. However, quit often we do not study the function itself, but for example its first or second order derivative, or its Laplacian, and hence there are multiple critical points at large scales in general.

*Example 5.* A last example we give, is the function, $f := \sum_{k=0}^{\infty} 2^{-k\mu} \delta_{2^k} * g_s$ for some $s > 0$, here $\delta_{2^k} := \delta(\cdot - 2^k)$, which is the Dirac function situated in $2^k$. Interesting about this function $f$ is that its first $m$th-order moments, with $m < \mu$ are finite, i.e., $\int_{\mathbb{R}} x^m f(x)\, dx < \infty$ for all $m < \mu$, and $f$ goes to zero when $|x|$ goes to infinity.

These two conditions do not hold for the function from Example 3, but do hold for the compactly supported functions in Theorem 1. One could think of these conditions as restrictions on the speed of growth of the function, and one might expect that these requirements are sufficient for the theorem to hold. In Appendix A, however, it is demonstrated that this nonnegative function has more than one extremum for large scales $\sigma \in \mathbb{R}_{\geq 0}$.

## 4 Spatial Constraints for Critical Points

In Section 3, we gave an explicit estimate of a scale $\varsigma$ beyond which there is only one spatial maximum for $f_\sigma$ when $\sigma > \varsigma$. This gives a lower bound for the scale beyond which it might be needless to analyze the function $f_\sigma$ any further, because further blurring does not bring further topological changes to the function as it remains a single blob upon succussive blurring.

In this section we present a result, Theorem 2, that also puts spatial constraints on the region of interest of scale-space. Again, it is required for the theorem to hold that the function is nonnegative. Moreover, the result becomes really useful, when the function is also compactly supported, which is exemplified by Corollary 1.

**Theorem 2.** *Let the function $f$ be nonnegative, then every spatial critical point of $f_\sigma$ is in the closure of the convex hull of* $\mathrm{supp}(f)$*.*

Refer to Appendix A for the proof.

As a direct consequence of Theorems 1 and 2, we can, for example, restrict the region in scale-space that would be of interest when we want to track down toppoints (i.e. scale-space catastrophes $(x, \sigma) \in \mathbb{R}^n \times \mathbb{R}_{\geq 0}$ for which

$$\mathrm{grad} f_\sigma(x) = 0 \wedge \det \left( \frac{\partial^2 f_\sigma}{\partial x_j x_k}(x) \right)_{jk} = 0,$$

see [1]) of compactly supported, nonnegative functions $f$.

**Corollary 1.** *For a nonnegative function $f$ with compact support, all toppoints in scale-space are situated in $C \times [0, \mathrm{r}(f)] \subset \mathbb{R}^n \times \mathbb{R}_{\geq 0}$, where $C$ is the convex closure of* $\mathrm{supp}(f)$*.*

## 5 Conclusions and Discussion

This article presented two theorems which, for example, enable us to restrict the interesting region in scale-space, with respect to toppoints.

The first, and main, theorem states that if a nonnegative function with compact support is blurred sufficiently, there is a scale for which any further blurring does not give rise to a topological change: the function is a single blob, and remains a single blob upon succussive blurring. This is a particular version of, what we called a "Folklore Theorem": if a function $f : \mathbb{R}^n \to \mathbb{R}$ is blurred sufficiently, then the blurred function has a single critical point, which is an extremum.

This result might be clear intuitively, because the function has only a bounded domain where one can distinguish details (or structure [7]), and so one expects that if scale is taken large enough every single detail is lost. However, based on the same reasoning, we could argue that the theorem is even true for not necessarily nonnegative functions; a conjecture that has to be rejected as Example 4 shows us. Hence the Folklore Theorem does not hold in all its generality. (Besides Example 4, we discussed some other examples, which show that the Folklore Theorem is not true in general.)

Furthermore, in the same theorem, we give a sharp estimate of the scale beyond which there are no more details distinguishable in the blurred function. This scale is equal to $r$, where $r$ equals $\mathrm{r}(f)$, the radius of the support of the function nonnegative $f$. This, in combination with Remark 1, suggests that we should call $2\sigma$ the scale of $f_\sigma$ at which a function $f$ is observed.

The second theorem presented spatial constraints for the critical points. This result is complementary to Theorem 1, which can be interpreted as imposing scale constraints in scale-space (cf. Corollary 1), but now the constraints are in the spatial domain. Theorem 2 states that, for every nonnegative, compactly supported function $f$ every spatial critical point of $f_\sigma$ is in the closure of the convex hull of the support of the initial function $f$.

We conclude that the scale-space constraints we considered in this article are all based merely on the knowledge that the functions under consideration are nonnegative and compactly supported. Other function classes, for which comparable statements are possible, could be investigated. However, from the examples we give, it is clear that it can be hard to formulate such statements for different or broader classes of functions.

Another interesting direction for subsequent research is to generalize Theorem 1 and give explicit – and sharp – estimates of scales after which there are only 2 critical points, 3 critical points, etc. left, or to generalize Theorem 2 in such a way that one can restrict the "support of the spatial critical points for the function $f_\sigma$" to different domains than the convex closure of supp($f$).

# A Proofs and Demonstrations

This appendix gives the proofs of both Theorems 1 and 2, as well as the proof of the statement in Example 5. We start with a definition of an inner product that is used in the proof of Theorem 1. Note that the domain of integration is omitted in the remainder of this appendix. One should read $\int_{\mathbb{R}^n}$ for $\int$.

**Definition 3.** *Given a nonnegative function $f$, a spatial coordinate $x$ and a scale $\sigma$. Define the inner product $\langle\cdot,\cdot\rangle_x$ for two functions $a : \mathbb{R}^n \to \mathbb{R}$ and $b : \mathbb{R}^n \to \mathbb{R}$ as*

$$\langle a(y), b(y)\rangle_x := \frac{\int a(y)\, b(y)\, f(y)\, g_\sigma(x-y)\, dy}{\int f(y)\, g_\sigma(x-y)\, dy},$$

*furthermore define $\overline{a(y)} := \langle a(y), 1\rangle_x$.*

*Proof.* (Theorem 1) To proof the theorem, we first derive a system of equations that is satisfied by a spatial critical point $\xi \in \mathbb{R}^n$ at a scale $\sigma$. Clearly, for $\xi$ the following $n$ equations hold:

$$\frac{\partial f_\sigma}{\partial x_j}(\xi) = \frac{1}{\sigma^2} \int (\xi_j - y_j)\, f(y)\, g_\sigma(\xi - y)\, dy = 0, \tag{1}$$

for all $j \in \{1, \ldots, n\}$. Hence, if $\xi$ is a critical point:

$$\xi_j \int f(y)\, g_\sigma(\xi - y)\, dy = \int y_j f(y)\, g_\sigma(\xi - y)\, dy,$$

and we conclude that $\xi$ satisfies $\xi = F(\xi; \sigma)$, where every $F_j : \mathbb{R}^n \times \mathbb{R}_{\geq 0} \to \mathbb{R}$ of $F$ is defined as:

$$F_j(x; \sigma) := \frac{\int y_j f(y) g_\sigma(x-y) dy}{\int f(y) g_\sigma(x-y) dy} = \bar{y}_j.$$

Hence, $\xi$ is a fixed point of $F(\cdot;\sigma)$.

$F(\cdot;\sigma)$ is a contraction if the operator norm of the matrix $\Phi(x) := (\frac{\partial F}{\partial x_k}(x;\sigma))_k$ is smaller than 1, i.e., $\|\Phi(x)\|$, for all $x \in \mathbb{R}^n$ [2]. Writing out the entries $\frac{\partial F_j}{\partial x_k}(x;\sigma)$ of $\Phi(x)$, we get:

$$\begin{aligned} \frac{\partial F_j}{\partial x_k}(x;\sigma) &= \frac{1}{\sigma^2}\left(\langle y_j, y_k\rangle_x - \langle y_j, 1\rangle_x \langle y_k, 1\rangle_x\right) \\ &= \frac{1}{\sigma^2}\big\langle y_j - \langle y_j, 1\rangle_x, y_k - \langle y_k, 1\rangle_x\big\rangle_x = \frac{1}{\sigma^2}\langle y_j - \bar{y}_j, y_k - \bar{y}_k\rangle_x. \end{aligned} \tag{2}$$

From Equation (2), it follows that $\Phi(x)$ is a symmetric matrix and so its operator norm equals the maximum of the absolute values of the eigenvalues of $\Phi(x)$.

Rewrite $\Phi(x)$ as follows:

$$\Phi(x) = \frac{1}{\sigma^2}\left(\langle y_j - \bar{y}_j, y_k - \bar{y}_k\rangle_x\right)_{jk} = \frac{\int (y-\bar{y})(y-\bar{y})^{\mathrm{t}} f(y)\, g_\sigma(x-y)\, dy}{\sigma^2 \int f(y)\, g_\sigma(x-y)\, dy},$$

and note that $\Phi(x)$ is positive semi-definite, because the matrix $(y-\bar{y})(y-\bar{y})^{\mathrm{t}}$ is positive semi-definite and $f$ is nonnegative. From this, it follows that the operator norm equals the maximum of the eigenvalues of $\Phi(x)$ [2]. Furthermore, it holds that the trace of the matrix $\Phi(x)$ is greater or equal to this maximum eigenvalue. Let $\lambda$ be this eigenvalue, then the following holds:

$$\begin{aligned} \lambda \le \operatorname{trace}(\Phi(x)) &= \sum_{j=1}^{n} \frac{1}{\sigma^2}\langle y_j - \bar{y}_j, y_j - \bar{y}_j\rangle_x \\ &= \frac{\sum_{j=1}^{n} \int (y_j - \bar{y}_j)^2 f(y)\, g_\sigma(x-y)\, dy}{\sigma^2 \int f(y)\, g_\sigma(x-y)\, dy}. \end{aligned} \tag{3}$$

Furthermore, it is easy to verify that $\int \sum_j (y_j - \gamma_j)^2 f(y)\, g_\sigma(x-y)\, dy = \int \|y - \gamma\|^2 f(y)\, g_\sigma(x-y)\, dy$ attains its minimal value for $\gamma = \bar{y} \in \mathbb{R}^n$. Now, because the radius of the convex closure of $\operatorname{supp}(f)$ equals $r$, there is an $m \in \mathbb{R}^n$ for which $\|y - m\| \le r$ for $y \in \operatorname{supp}(y)$. And so from the last two statements, and Equation (3) we have:

$$\begin{aligned} \lambda \le \operatorname{trace}(\Phi(x)) &= \frac{\sum_{j=1}^{n} \int (y_j - \bar{y}_j)^2 f(y)\, g_\sigma(x-y)\, dy}{\sigma^2 \int f(y)\, g_\sigma(x-y)\, dy} \\ &\le \frac{\int \sum_{j=1}^{n} (y_j - m_j)^2 f(y)\, g_\sigma(x-y)\, dy}{\sigma^2 \int f(y)\, g_\sigma(x-y)\, dy} \\ &\le \frac{\int r^2 f(y)\, g_\sigma(x-y)\, dy}{\sigma^2 \int f(y)\, g_\sigma(x-y)\, dy} \end{aligned}$$

As already stated, the function $F(\cdot;\sigma)$ is a contraction if $\lambda < 1$. The function $F(\cdot;\sigma)$ being a contraction directly implies that a spatial critical point $\xi$ is unique. Hence $f_\sigma$ has a unique spatial critical point if $\frac{r^2}{\sigma^2} < 1$. So for all $\sigma > r$, $f_\sigma$ has a unique critical point, which is a maximum, because $f_\sigma$ is nonnegative and $\lim_{\|x\|\to\infty} f_\sigma(x) = 0$. □

*Proof.* (Theorem 2) Let $a \in \mathbb{R}^n$ and $c \in \mathbb{R}$, such that for every $y \in \mathrm{supp}(f)$, we have: $\langle y, a\rangle \leq c$ (here $\langle \cdot, \cdot \rangle$ is the standard inner product between two vectors). Note that $\langle \cdot, a\rangle \leq c$ defines a half-space in $\mathbb{R}^n$, and so the foregoing states that $\mathrm{supp}(f)$ is situated in this half-space.

Now assume $\xi$ to be a critical point of $f_\varsigma$ ($\varsigma$ is fixed), then the $n$ Equations (1), satisfied by a critical point, imply that

$$\begin{aligned} \langle \xi, a\rangle \int f(y)\, g_\varsigma(\xi - y)\, dy &= \int \langle y, a\rangle f(y)\, g_\varsigma(\xi - y)\, dy \\ &\leq c \int f(y)\, g_\varsigma(\xi - y)\, dy, \end{aligned}$$

Here it is used that $f$ is nonnegative to obtain the inequality. This inequality in turn implies that $\langle \xi, a\rangle \leq c$. Now, the intersection of all closed half-spaces determined by $\langle \cdot, a\rangle \leq c$ and containing $\mathrm{supp}(f)$, equals the closure $C$ of the convex hull of $\mathrm{supp}(f)$ (see [5]). Furthermore, because for a critical point $\xi$ we also have that $\langle \xi, a\rangle \leq c$ for all the former choices of $a \in \mathbb{R}^n$ and $c \in \mathbb{R}$, we conclude that all critical points reside within the same closure $C$. □

*Proof.* (Example 5) Let $\phi := \sum_{k=0}^{\infty} 2^{-k\mu}\delta_{2^k}$. If this distribution has the property of having multiple extrema for all scales larger than a certain scale $\varsigma$, then the function $f := \phi * g_s$ has the same property.

To show that $\phi$ possesses this property, we start by taking the derivative of $\phi_\sigma$, which equals $\phi'_\sigma(x) = \frac{1}{\sigma^2}\sum_{k=0}^{\infty} 2^{-k\mu}(2^k - x)\, g_\sigma(x - 2^k)$, with $\sigma > 0$. Firstly, the function $\phi'_\sigma$ is positive for $x$ smaller than 1. Secondly, for $\sigma > 0$, $\phi'_\sigma(2)$ is negative. To see this, note that for $x = 2$, the first summand equals $2^{-0\mu}(2^0 - x)\, g_\sigma(x - 2^0) = -g_\sigma(1)$, which is negative. Furthermore, for the other summand ($k \geq 1$), which are nonnegative, the following holds:

$$2^{-k\mu}(2^k - 2)\, g_\sigma(2 - 2^k) < 2^{-2k}2^k g_\sigma(2^{k-1}) \leq 2^{-k} g_\sigma(1),$$

hence

$$\sum_{k=1}^{\infty} 2^{-k\mu}(2^k - x)\, g_\sigma(x - 2^k) < \sum_{k=1}^{\infty} 2^{-k} g_\sigma(1) = g_\sigma(1),$$

which shows that the latter statement holds.

Thirdly, for every $\sigma$ there is an $x > 2$ for which $\phi'_\sigma(x)$ is positive. Proof: take $x = 2^p - \sigma$ ($p \in \mathbb{Z}_{>0}$), then the summand of $\phi'_\sigma$, where $k = p$, equals: $2^{-p\mu}\sigma e^{-\frac{1}{2}}$. Now, choose $p$ such that $2^{p-2} > \sigma$ and such that $2^{\frac{p-3}{2}} > \sigma$ – implying that

$2\sigma^2 < 2^{p-2}$, then the following holds:

$$
\begin{aligned}
-\sum_{k=0}^{p-1} 2^{-k\mu}(2^k - 2^p + \sigma)\, g_\sigma(2^p - \sigma - 2^k) \leq \\
\sum_{k=0}^{p-1} 2^{-0\mu} 2^p g_\sigma(2^{p-1} + 2^{p-2} - 2^k) \leq \\
\sum_{k=0}^{p-1} 2^p g_\sigma(2^{p-2}) \leq \\
p\, 2^p\, e^{-\frac{(2^{p-2})^2}{2\sigma^2}} \leq p\, 2^p\, e^{-2^{p-2}}.
\end{aligned}
$$

Because, $e^{-2^{p-2}}$ goes to 0 extremely rapidly when $p$ goes to infinity, it follows that there is a $\rho \in \mathbb{Z}_{>0}$ for which

$$
-\sum_{k=0}^{\rho-1} 2^{-k\mu}(2^k - 2^\rho + \sigma)\, g_\sigma(2^\rho - \sigma - 2^k) \leq \rho\, 2^\rho\, e^{-2^{\rho-2}} \leq 2^{-\rho\mu}\sigma e^{-\frac{1}{2}}.
$$

Hence, taking $x = 2^\rho - \sigma$ gives $\phi'_\sigma(x) > 0$.

Now, from the three foregoing statements, we have that $\phi'_\sigma(0)$, $\phi'_\sigma(2)$, and $\phi'_\sigma(2^\rho - \sigma)$ are positive, negative, and positive, respectively. Finally, because $\lim_{x\to\infty} \phi'_\sigma(x)$ is negative, we conclude that for every scale $\sigma > 0$, $\phi_\sigma$ has at least three extrema. □

## References


1. J. Damon. Local morse theory for solutions to the heat equation and gaussian blurring. *Journal of Differential Equations*, 115(2):368–401, 1995.
2. J. A. Dieudonné. *Foundations of Modern Analysis*, volume 10 of *Pure and Applied Mathematics*. Academic Press, New York . London, 1960.
3. L. M. J. Florack. *Image Structure*, volume 10 of *Computational Imaging and Vision*. Kluwer Academic Publishers, Dordrecht . Boston . London, 1997.
4. L. M. J. Florack, B. M. Ter Haar Romeny, J. J. Koenderink, and M. A. Viergever. Linear scale-space. *Journal of Mathematical Imaging and Vision*, 4(4):325–351, 1994.
5. B. Grünbaum. *Convex Polytopes*, volume 16 of *Pure and Applied Mathematics*. Interscience Publishers, London, 1967.
6. P. Johansen, S. Skelboe, K. Grue, and J. Damgaard Andersen. Representing signals by their toppoints in scale-space. In *Proceedings of the 8th International Conference on Pattern Recognition*, pages 215–217, Paris, 1986.
7. J. J. Koenderink. The structure of images. *Biological Cybernetics*, 50:363–370, 1984.
8. T. Lindeberg and B. M. Ter Haar Romeny. *Linear Scale-Space I: Basic Theory*, volume 1 of *Computational Imaging and Vision*, chapter 1, pages 1–38. Kluwer Academic Publishers, Dordrecht . Boston . London, 1994.
9. A. L. Yuille and T. Poggio. Fingerprint theorems for zero-crossings. *Journal of the Optical Society of America. A, Optics and Image Science*, 2:683–692, 1985.