DISTRIBUTION THEORY ON P.C.F. FRACTALS

LUKE G. ROGERS AND ROBERT S. STRICHARTZ

Abstract. We construct a theory of distributions in the setting of analysis on post-critically finite self-similar fractals, and on fractafolds and products based on such fractals. The results include basic properties of test functions and distributions, a structure theorem showing that distributions are locally-finite sums of powers of the Laplacian applied to continuous functions, and an analysis of the distributions with point support. Possible future applications to the study of hypoelliptic partial differential operators are suggested.

1. Introduction

The prevalence of fractal-like objects in nature has led both physicists and mathematicians to study dynamic processes on fractals. One rigorous way to do this on post-critically finite (p.c.f.) fractals is by studying differential equations in the natural analytic structure. A brief description of this analytic structure will appear in Section 2 below, but we emphasize that it is intrinsic to the fractal, and is not necessarily related to the analysis on a space in which the fractal may be embedded. For example, the familiar Sierpinski gasket fractal **SG** is often visualized as a subset of $\mathbb{R}^2$, but restricting a smooth function on $\mathbb{R}^2$ to **SG** does not give a smooth function on the fractal [3]. Similarly, we should not expect the solutions of differential equations on fractals to be quite like the solutions of their Euclidean analogues; for example, many fractals have Laplacian eigenfunctions that vanish identically on large open sets [2], whereas eigenfunctions of the Euclidean Laplacian are analytic.

Perhaps the most important tools for studying differential equations in the Euclidean context are Fourier analysis and the theory of distributions. Since the theory of analysis on fractals relies on first constructing a Laplacian operator Δ, it is unsurprising that quite a lot is known about the fractal analogue of Fourier analysis. In some interesting cases the spectrum and eigenfunctions of the Laplacian are known explicitly, and many results about Laplacian eigenfunctions have also been derived by using probability theory to study the heat diffusion on fractals. Fourier-type techniques have also been applied to treat smoothness in the fractal setting: analogues of the Sobolev, Hölder-Zygmund and Besov spaces that are so important in Euclidean analysis of differential equations were introduced and investigated in [34]. Analogues of other basic objects in Euclidean analysis are studied in [19, 4]. By contrast there has not previously been a theory of distributions on fractals, and it is the purpose of the present work to provide one.

It is relatively elementary to define distributions on fractals; as usual they are dual to the space of smooth functions with compact support, where a function u is said to be smooth if $\Delta^k u$ is continuous for all $k \in \mathbb{N}$. The main theorems about distributions are then really theorems about smooth functions, and the key to proving many of them is knowing how

Date: October 20, 2010.
2000 *Mathematics Subject Classification.* Primary 28A80, Secondary 46F05.
Key words and phrases. Analysis on fractals, post-critically finite self-similar fractal, distribution, PDE on fractals.
Second author supported in part by NSF grant DMS 0652440.

to smoothly partition smooth functions. Partitions of unity are used to achieve this in the Euclidean setting, but are not useful on fractals because products of smooth functions are not smooth [3]. (This latter fact also implies that products of functions and distributions are not distributions, so the distributions are not a module over the smooth functions.) Instead we rely on a partitioning theorem for smooth functions proved in [28], see Theorem 2.8 below. Using this partitioning result we are able to prove analogues of the standard structure theorems describing distributions as derivatives of continuous functions (Theorem 5.10), and identifying the positive distributions as positive measures (Theorem 5.6). We can also characterize the distributions of point support as finite linear combinations of certain "derivatives" of Dirac masses that can be explicitly described (Corollaries 6.6 and 6.13), provided we make certain assumptions about the point in question. These assumptions are needed in order to understand the local behavior of smooth functions at the point, and are related to work done in [32, 39, 5, 27, 1]. Unfortunately many of the proofs are quite technical in nature; we have tried to explain in advance the strategies behind the proofs, which are more conceptual.

At the end of this paper we suggest several interesting questions related to the hypoellipticity of differential operators that are natural to consider in the context of distribution theory. It should also be noted that there are a number of results on local solvability of differential equations [36, 25] that could be reformulated in this context. We expect that this work will provide the foundation for many subsequent investigations.

Finally, one would really like to have a theory of tempered distributions on p.c.f. fractal sets, and the authors would be very interested to see such a theory. While it is tempting to believe that this kind of theory could be obtained by simple modifications of the material presented here, we do not believe this to be the case. Indeed, the authors (together with Erin P. J. Pearse) have some partial results toward developing a theory of tempered distributions on the Sierpinski Gasket, but have so far been unable to resolve certain technical details that arise even in this simple case.

2. Setting

We begin by describing the basic elements of analysis on a post-critically finite self-similar set X, as laid out in the monograph of Kigami [13]; in this section all unreferenced results may be found in [13], which also includes proofs and references to the original literature. The reader who prefers to have a concrete example of a p.c.f. set in mind may choose to think of X as the Sierpinski Gasket, in which case a more elementary exposition of the material that follows may be found in [37].

P.C.F. Fractals. Let X be a self-similar subset of $\mathbb{R}^d$ (or more generally a compact metric space). By this we mean that there are contractive similarities $\{F_j\}_{j=1}^N$ of $\mathbb{R}^d$, and X is the unique compact set satisfying $X = \cup_{j=1}^N F_j(X)$. Then X has a natural cell structure in which we associate to a word $w = w_1 w_2 \dots w_m$ of length m the map $F_w = F_{w_1} \circ \cdots \circ F_{w_m}$, and call $F_w(X)$ an m-cell. If w is an infinite word then we let $[w]_m$ be its length m truncation and note that $F_w(X) = \bigcap_m F_{[w]_m}(X)$ is a point in X.

We say $F_j(x)$ is a critical value of $X = \cup_{l=1}^N F_l(X)$ if there is $y \in X$ and $k \neq j$ such that $F_j(x) = F_k(y)$. An infinite word w is critical if $F_w(X)$ is a critical value, and $\tilde{w}$ is post-critical if there is $j \in \{1, \dots, N\}$ such that jw is critical. We always assume that the set of post-critical words is finite, in which case the fractal is said to be *post-critically finite (p.c.f.)*. The *boundary* of X is then defined to be the finite set V_0 consisting of all points $F_{\tilde{w}}(X)$ for which $\tilde{w}$ is post-critical; this set is assumed to contain at least two points. We also let $V_m = \cup_w F_w(V_0)$, where the union is over all words of length m. Points in $V_* = \cup_{m\geq 0} V_m$ that are not in V_0 are called *cell boundary points*, and a key property of p.c.f. fractals is that cells can intersect only at these points. A cell boundary point x is called a *junction point* if it has the property that any sufficiently small neighborhood U of x is disconnected by the removal of x. In this case $U \setminus \{x\}$ has finitely many connected components.

We fix a probability measure μ on X that is self-similar in the sense that there are $\mu_1, \ldots, \mu_N$ such that the cell corresponding to $w = w_1 \ldots w_m$ has measure $\mu(F_w(X)) = \prod_{j=1}^m \mu_{w_j}$. The usual Bernoulli measure in which each $\mu_j = \frac{1}{N}$ is one example.

Dirichlet Form. Our analysis on X will be constructed from an irreducible self-similar Dirichlet form. A closed quadratic form E on $L^2(\mu)$ is called Dirichlet if it has the (Markov) property that if $u \in \mathrm{dom}(E)$ then so is $\tilde{u} = u\chi_{0<u<1} + \chi_{u\geq 1}$ and $E(\tilde{u}, \tilde{u}) \leq E(u, u)$, where χ_A is the characteristic function of A. Self-similarity of E means that there are renormalization factors $r_1, \ldots, r_N$ such that

$$E(u, v) = \sum_{j=1}^{N} r_j^{-1} E(u \circ F_j, v \circ F_j). \tag{2.1}$$

It follows immediately that $E(u, v)$ can also be expressed as the sum over m-words of $r_w^{-1} E(u \circ F_w, v \circ F_w)$ where $r_w = r_{w_1} \cdots r_{w_m}$. In order to use results from [28] we assume that $0 < r_j < 1$ for all j, in which case E is *regular*, meaning that that $C(X) \bigcap \mathrm{dom}(E)$ is dense both in $\mathrm{dom}(E)$ with E-norm and in the space of continuous functions $C(X)$ with supremum norm.

It is far from obvious that interesting fractals should support such Dirichlet forms, but in fact the conditions described so far are satisfied by many p.c.f. self-similar sets that have sufficient symmetry. In particular, if X is a nested fractal in the sense of Lindstrøm [17] then a Dirichlet form of the above type may be constructed using a diffusion or a harmonic structure [15, 7, 30]. Some other approaches may be found in [21, 14, 18, 23, 22, 10, 24].

Harmonic Functions. Given a function on V_0 (usually thought of as an assignment of boundary values) there is a unique continuous function on X that has these boundary values and minimizes the energy. Such functions are called *harmonic*, and form a finite dimensional space containing the constants. It is easy to see that there are *harmonic extension matrices* A_j, $j = 1, \ldots, N$ with the property that if h is harmonic then A_j maps the values of h on V_0 to its values on $F_j(V_0)$. The largest eigenvalue of each A_j is 1, corresponding to the constant functions; it is useful to know that the second eigenvalue has magnitude at most r_j, and is equal to r_j in many interesting examples ([13], Appendix A).

The Laplacian and Normal Derivatives. We can use the energy and measure to define weak Laplacian operators on the fractal by requiring that $E(u, v) = -\int (\Delta u) v \, d\mu$ for all u, v in a suitable subspace of $\mathrm{dom}(E)$. If the subspace consists of those functions that vanish on V_0 (i.e. Dirichlet boundary conditions) then we obtain the Dirichlet Laplacian Δ_D, while if the subspace is all of $\mathrm{dom}(E)$ we obtain the Neumann Laplacian Δ_N. According to the Gauss-Green formula given below the latter is equivalent to requiring that u has vanishing normal derivatives on V_0. In either case $-\Delta$ is a non-negative self-adjoint operator on $L^2(\mu)$ with compact resolvent (see Theorem 2.4.2 of [13]). We will generally use the Dirichlet Laplacian and will denote it by Δ, its eigenvalues by λ_j and the corresponding eigenvectors by ψ_j. When $\Delta u \in C(X)$ we write $u \in \mathrm{dom}(\Delta)$ and think of these as the (continuously) differentiable functions on X. Inductively define $\mathrm{dom}(\Delta^k)$ for each k and then $\mathrm{dom}(\Delta^\infty) = \cap_k \mathrm{dom}(\Delta^k)$. We say f is smooth if $f \in \mathrm{dom}(\Delta^\infty)$. Harmonic functions have zero Laplacian.

By introducing a normal derivative ∂_n at boundary points the defining equation for the Laplacian can be generalized to obtain a Gauss-Green formula $E(u,v) = -\int(\Delta u)v\,d\mu + \sum_{x\in V_0} v(x)\partial_n u(x)$ when $v \in \mathrm{dom}(E)$, as in Theorem 3.7.8 of [13]. This formula may be localized to a cell $F_w(X)$, in which case $\partial_n^w u(q) = \lim_m E(u, v_m)$ at the boundary point $q = \bigcap_m F_{wr_j^m}(X)$; here v_m is identically zero except on $F_{wr_j^m}(X)$, where we require that $v_m \circ F_{wr_j^m}$ be harmonic on X with all boundary values equal to 0 other than $v_m(q) = 1$. The superscript w in $\partial_n^w u(q)$ indicates which cell the normal derivative is taken with respect to, as there is one for each cell that intersects at q. In general the normal derivatives exist once Δu exists as a measure. If $u \in \mathrm{dom}(\Delta)$ then the normal derivatives at a point sum to zero. Conversely, suppose the cells $\{F_{w^{(j)}}(X)\}_j$ partition X and pairwise intersect only at a finite set of points, and u is defined piecewise by functions u_j on $F_{w^j}(X)$ with $u_j \circ F_{w^j} \in \mathrm{dom}(\Delta)$. Then $u \in \mathrm{dom}(\Delta)$ if and only if for every intersection point x of these cells, all $u_j(x)$ are equal, all $\Delta u_j(x)$ are equal, and the normal derivatives of the u_j at x sum to zero. We call these constraints the *matching conditions* for the Laplacian.

Resistance Metric. In addition to the Laplacian and other derviatives, the Dirichlet form also provides us with a metric intrinsic to the fractal. We define the resistance metric R by

$$R(x,y) = \min\{E(u)^{-1} : u \in \mathrm{dom}(E),\ u(x) = 0,\ u(y) = 1\}.$$

In Sections 2.3 and 3.3 of [13] it is proven that under our assumptions this minimum exists and defines a metric, and that the R-topology coincides with the topology induced from the embedding of X into $\mathbb{R}^d$. Of particular importance for us is the fact that continuity may be treated using the resistance metric, for which purpose the following Hölder-$\frac{1}{2}$ estimate is very useful:

$$|u(x) - u(y)|^2 \leq E(u)R(x,y) \leq \|u\|_2\|\Delta u\|_2 R(x,y). \tag{2.2}$$

If $u \in \mathrm{dom}(\Delta)$ vanishes on V_0 then we obtain the first inequality trivially from the definition and the second by applying the Cauchy-Schwarz inequality to $E(u) = -\int u\Delta u\,d\mu$. For general $u \in \mathrm{dom}\,\Delta$ we can simply subtract the harmonic function with the same boundary values and apply the same estimate. In particular, this shows that the L^2 domain of the Laplacian embeds in the continuous functions.

Fractafolds. Since the results in this paper are primarily local in nature, we will be able to work on a connected fractafold based on X with a restricted cellular construction, which we denote by Ω. Some results on fractafolds and their spectra may be found in [33]. As with a manifold based on Euclidean space, a fractafold based on X is just a connected Hausdorff space in which each point has a neighborhood homeomorphic to a neighborhood of a point in X. One way to construct a fractafold is by suitably gluing together copies of X, for example by identifying appropriate boundary points. This leads us to the idea of a cellular construction, which is the analogue of a triangulation of a manifold. A restricted cellular construction consists of a finite or countably infinite collection of copies X_j of X, together with an admissible identification of their boundary points. Admissibility expresses the requirement that the result of the gluing be a fractafold; more precisely, it means that if $\{x_1, \dots, x_J\}$ are identified then there is a junction point $x \in X$ and a neighborhood U of x such that each of the components $U_1, \dots, U_J$ of $U\setminus\{x\}$ is homeomorphic to a neighborhood of the corresponding point x_j in X_j. (Recall that a junction point is a point of $V_* \setminus V_0$ such that for any sufficiently small neighborhood U the set $U \setminus \{x\}$ is disconnected.) We call any such point x a *gluing point*, and make the obvious definition that a neighborhood of x is a union of neighborhoods of x in each of the cells X_j that meet at x in the manner previously described.

It should be noted that the above is not the most general kind of cellular construction (hence the term *restricted* in the definition), because some fractals have non-boundary points (called terminal points) at which cells may be glued (see [33], Section 2). Dealing with such points introduces certain technicalities that, while not insurmountable, cause complications in defining the Green's operator (see below) that we will need for proving Theorem 5.10. It is worth noting that if X has some topological rigidity then all fractafolds have restricted cellular structure. This is true, for example, for fractafolds based on the Sierpinski Gasket ([33] Theorem 2.1).

Thus far our fractafold has only topological structure; however if Ω has a restricted cellular construction then a smooth structure may be introduced in the same manner as it was on X itself, specifically by defining a Dirichlet energy and a measure and thus a weak Laplacian. We can take the energy on Ω to be the sum of the energies on the cells of the cellular construction, and the measure (which is not necessarily finite, but is finite on compacta) to be the sum of the measures on the cells:

$$E(u,v) = \sum_j E_{X_j}(u\big|_{X_j}, v\big|_{X_j}) = \sum_j a_j E_X(u\big|_{X_j} \circ \iota_j, v\big|_{X_j} \circ \iota_j)$$
$$\mu(A) = \sum_j \mu_{X_j}(A \cap X_j) = \sum_j b_j \mu_X(\iota_j^{-1}(A \cap X_j))$$

where $\iota_j : X \to X_j$ is the map from the cellular construction. In the same way that the angle sum at a vertex of a triangulation of a manifold determines the curvature at the vertex, the choice of the weights a_j and b_j amount to a choice of metric on Ω, with $a_j = b_j = 1$ for all j being the flat case (see [33], Section 6). As all of the computations made later in the paper may be made on one cell at a time, we will henceforth suppress the weights a_j and b_j.

Well-known results about p.c.f. fractals imply the existence of a Green's function (for which there is an explicit formula) on finite unions of cells in a fractafold with cellular construction.

Lemma 2.1. *Let $K = \cup_1^J X_j$ be a connected finite union of cells in Ω and such that $K \neq \Omega$. Then there is a Green's operator G_K with the property that if ν is a Radon measure on K (i.e. a Borel measure that is finite on compacta, outer regular on Borel sets and inner regular on open sets), then $G_K\nu$ is continuous, $-\Delta G_K \nu = \nu$ on the interior of K, and $G_K\nu\big|_{\partial K} = 0$. The same conclusion holds in the case $K = \Omega$ under the additional assumption $\int d\nu = 0$.*

Remark 2.2. It is clear that ∂K is a subset of the boundary points of the cells X_j, specifically consisting of those gluing points at which not all glued cells are included in K.

Proof. We recall from Sections 3.6 and 3.7 of [13] that our assumptions on X imply there is a Green's operator G on X with continuous kernel $g(x, y)$, such that $-\Delta G\nu = \nu$ and $G\nu\big|_{\partial X} = 0$ for all Radon measures ν. There is an explicit formula giving $g(x, y)$ as a series.

If G_j is the Green's operator for the cell X_j it is easy to verify that $-\Delta \sum G_j \nu = \nu$, except at the gluing points where the Laplacian can differ from ν by Dirac masses, the size of which may be computed explicitly by summing the normal derivatives of the $G_j\nu$ at the points that are glued. However it is also apparent that by assigning values at each of the gluing points and extending harmonically on the cells we obtain a continuous and piecewise harmonic function, the Laplacian of which is a sum of Dirac masses at the gluing points.

Provided the boundary ∂K is non-empty (which is obvious if $K \neq \Omega$), a linear algebra argument (Lemma 3.5.1 in [13]) shows that for any prescribed set of weights for Dirac masses of the Laplacian at interior gluing points, there is a unique piecewise harmonic function that vanishes on ∂K and has this Laplacian. Subtracting this piecewise harmonic function from $\sum G_j \nu$ gives the required $G_K \nu$.

On the other hand, if ∂K is empty then the kernel of Δ is precisely the constant functions. We can therefore invert $-\Delta$ on the measures that annihilate constants, that is, those for which $\int d\nu = 0$. This can be done explicitly in the same manner as in the previous case, except that the linear algebra step now shows the Laplacians of piecewise harmonics span the space of mean-zero linear combinations of Dirac masses at the gluing points. In this case the choice of piecewise harmonic function is unique up to adding a constant; our convention is to choose this constant so $\int G_K \nu(x) d\mu(x) = 0$. □

As is usual in the classical theory of distributions, we treat distributions on an open domain. The following assumption is used in a number of places in the paper, only some of which are mentioned explicitly.

Assumption 2.3. *Throughout this paper we will assume that Ω has no boundary.*

In some examples it is possible to deal with boundary points by passing to an appropriate cover, but relatively little is known in terms of covering theory for general fractafolds. Elementary examples to keep in mind include non-compact cases like open subsets of $X \setminus V_0$ or the infinite Sierpinski Gasket treated in [38], and compact fractafolds like the double cover of the Sierpinski Gasket **SG**, which consists of two copies of **SG** with each boundary point from one copy identified with exactly one boundary point of the other (see [33] for more details).

Smooth Cutoffs and Partitioning. As mentioned earlier, the structure theorems we shall prove for distributions rest heavily on results from [28]. In what follows we assume that Ω is a fractafold with restricted cellular structure and is based on a fractal X with regular harmonic structure.

Recall that $x \in X$ is a junction point if and only if there is a neighborhood $U \ni x$ such that $U \setminus \{x\}$ is disconnected into a finite number of components U_j. For a smooth function u the quantities $\Delta^k u(x)$ and $\partial_n^j \Delta^k u(x)$ exist for all $k \in \mathbb{N}$; the superscript j on ∂_n^j indicates the normal derivative with respect to the cell U_j. For a fixed j, the two sequences $\Delta^k u(x)$ and $\partial_n^j \Delta^k u(x)$ make up the *jet* of u at x in U_j. The first result we need from [28] is a Borel-type theorem on the existence of smooth functions with prescribed jets.

Theorem 2.4 ([28], Theorem 4.3 and Equation 4.8)**.** *Given values $\rho_0, \rho_1, \dots$ and $\sigma_0, \sigma_1, \dots$ there is a smooth function f on U_j that vanishes in a neighborhood of all boundary points except q, where the jet is given by $\Delta^k f(q) = \rho_k$ and $\partial_n^j \Delta^k f(q) = \sigma_k$ for all k. If we write U_j as $U_j = F_w(X)$ for a word w, and fix a number L of jet terms, then for any $\epsilon > 0$ we may construct f so that for $0 \leq k \leq L$, we have the estimate*

$$\|\Delta^k f\|_\infty \leq C(k)(r_w \mu_w)^{-k} \Big(\sum_{l=0}^{L} r_w^l \mu_w^l |\rho_l| + \sum_{l=0}^{L-1} r_w^{l+1} \mu_w^l |\sigma_l| \Big) + \epsilon \tag{2.3}$$

where $C(k)$ depends only on k and the harmonic structure on X.

Remark 2.5. Of course, it follows immediately that we can construct a smooth function with prescribed jets at each of the boundary points of a cell K and an estimate like (2.3), just by applying the theorem separately to each of the boundary points and summing the result.

Corollary 2.6. *If K is a cell in Ω and U is an open neighborhood of K, then there is a smooth function f such that $f = 1$ on K, $f = 0$ outside U, and $\|f\|_\infty \leq C$, where C is a constant that does not depend on K or U.*

Proof. Let $\{q_j\}$ be the boundary points of K and at each q_j take cells $V_{j,k} \subset U$ such that $\bigcup_k V_{j,k} \cup K$ contains a neighborhood of q_j. By making all of these cells sufficiently small and removing any inside K we may further assume that the $V_{j,k}$ intersect K only at q_j, intersect each other only at q_j and do not intersect $V_{j',k'}$ for any $j' \neq j$.

On each $V_{j,k}$ construct the smooth function $f_{j,k}$ guaranteed by Theorem 2.4 with $f_{j,k} = 1$ at q_j and all other jet terms at q equal to zero, and taking $\epsilon = 1$. Then the piecewise function

$$f(x) = \begin{cases} 1 & \text{for } x \in K \\ f_{j,k} & \text{for } x \in V_{j,k} \\ 0 & \text{otherwise} \end{cases}$$

is equal 1 on K and 0 off U by construction. It is also smooth, simply because the pieces are smooth and the matching conditions for Δ^l apply at each of the boundary points of the $V_{j,k}$ for all l. The bound $\|f\|_\infty \leq C$ independent of K and U now follows from (2.3) because the scale-dependent terms are all raised to the power zero, so are constant. □

A more difficult task than that in Corollary 2.6 is to construct a *positive* bump function that is equal to 1 on K and to zero outside the neighborhood U of K. A result of this type was proven in [28] under certain assumptions on the diffusion Y_t corresponding to the Laplacian. A sufficient assumption is that the heat kernel $p_t(x, y)$ (i.e. the transition density of of Y_t) satisfies an estimate of the form

$$p_t(x, y) \leq \frac{\gamma_1}{t^{\alpha/\beta}} \exp\Big(-\gamma_2\Big(\frac{R(x, y)^\beta}{t}\Big)^{1/(\beta-1)}\Big) \tag{2.4}$$

where α, β, γ_1 and γ_2 are constants. The estimate 2.4 is known to be valid in great generality on p.c.f. fractals (Corollary 1.2 of [9]).

Theorem 2.7 ([28] Corollary 2.9)**.** *Under the assumption* (2.4)*, for a cell K and an open neighborhood $U \supset K$, there is a smooth function f such that $f = 1$ on K, $f = 0$ outside U, and $f(x) \geq 0$ for all x.*

The final theorem from [28] that we will use extensively is concerned with the smooth partitioning of a smooth function.

Theorem 2.8 ([28], Theorem 5.1)**.** *Let $K \subset X$ be compact and fix $\bigcup U_\alpha \supset K$ an open cover. If $f \in$ dom(Δ^∞) then there is a decomposition $f = \sum_1^J f_j$ in which each f_j is in* dom(Δ^∞) *and has support in some U_{α_j}.*

Remark 2.9. Compactness of K is used only to obtain finiteness of the decomposition, and may be omitted for finite covers. An analogous countable (and locally finite) decomposition is then valid in the σ-compact case; in particular it is valid on Ω, because of the existence of a cellular structure.

Remark 2.10. The proof uses a result on the existence of smooth functions with prescribed jet at a point (Theorem 4.3 of [28]) to smoothly join cutoffs to a piece of the original function as in the proof of Theorem 2.6. This is very different from the Euclidean case where one simply multiplies the smooth function by a smooth bump. In particular, the construction of the cutoff depends explicitly on the growth rate of the jet of f at the boundary points under consideration, so for a collection of sets indexed by j, the mapping $f \mapsto \{f_j\}$ to a sequence of smooth functions supported on these sets is nonlinear.

Although the non-linearity will make some later proofs more complicated, this method does provide good estimates. From (2.3) and standard arguments for controlling the normal derivative $\partial_n \Delta^k f$ at a point by the norms $\|\Delta^j\|_\infty$, $j = 0\dots,k+1$, over a neighborhood of the point (like those in Section 6 below) we find that $f \mapsto f_j$ can be arranged to satisfy

$$\|\Delta^k f_j\|_\infty \leq C \sum_{l=0}^{k} \|\Delta^l f\|_\infty \tag{2.5}$$

where C is a constant depending only on k and K.

3. Test Functions

We define test functions on Ω in the usual way, and provide notation for the space of smooth functions on Ω topologized by uniform convergence on compacta.

Definition 3.1. The space of *test functions* $\mathcal{D}(\Omega)$ consists of all $\phi \in \operatorname{dom}(\Delta^\infty)$ such that $\operatorname{Sppt}(\phi)$ is compact. We endow it with the topology in which $\phi_i \to \phi$ iff there is a compact set $K \subset \Omega$ containing the supports of all the ϕ_i, and $\Delta^k\phi_i \to \Delta^k\phi$ uniformly on K for each $k \in \mathbb{N}$. There is a corresponding family of seminorms defined by

$$|\phi|_m = \sup\{|\Delta^k\phi(x)| : x \in \Omega, k \leq m\} \tag{3.1}$$

though it should be noted that the topology on $\mathcal{D}(\Omega)$ is not the usual metric topology produced by this family. For a discussion of the topology on $\mathcal{D}(\Omega)$ and its relation to these seminorms, see Chapter 6 of [29]. At times we may need the space of test functions on Ω having support in some compact set $K \subset \Omega$. By a slight abuse of notation we will write $\mathcal{D}(K)$ for this space.

Definition 3.2. $\mathcal{E}(\Omega) = \operatorname{dom}(\Delta^\infty)$ with the topology $\phi_i \to \phi$ iff for every compact $K \subset \Omega$ we have $\Delta^k\phi_i \to \Delta^k\phi$ uniformly on K for each $k \in \mathbb{N}$. There is a corresponding family of seminorms defined by

$$|\phi|_{m,K} = \sup\{|\Delta^k\phi(x)| : x \in K, k \leq m\}. \tag{3.2}$$

The following result is immediate from Theorem 2.8 and (2.5). It will be used frequently in the results proved below.

Lemma 3.3. *If $\phi \in \mathcal{D}(\Omega_1 \cup \Omega_2)$, then $\phi = \phi_1 + \phi_2$ for some $\phi_j \in \mathcal{D}(\Omega_j)$. For each m there is $C = C(M, \Omega_1, \Omega_2)$ so $|\phi_j|_m \leq C|\phi|_m$, $j = 1, 2$.*

One consequence is that $\mathcal{D}(\Omega)$ is dense in $\mathcal{E}(\Omega)$, because we may fix an increasing compact exhaustion $\cup_j K_j = \Omega$ of our domain and for arbitrary $\phi \in \mathcal{E}(\Omega)$ write $\phi = \phi_j + \tilde{\phi}_j$, where ϕ_j is supported in K_{j+1} and $\tilde{\phi}_j$ is supported in K_j^c, so that $\phi\big|_{K_j} = \phi_j\big|_{K_j}$. The functions ϕ_j are in $\mathcal{D}(\Omega)$ and it is clear that $\Delta^k\phi_j \to \Delta^k\phi$ uniformly on compacta, hence $\phi_j \to \phi$ in $\mathcal{E}(\Omega)$. Another density result that follows from Lemma 3.3 is as follows.

Theorem 3.4. *$\mathcal{D}(\Omega)$ is dense in $C_c(\Omega)$, the space of continuous functions with compact support, with supremum norm.*

Proof. The dual of $C_c(\Omega)$ is the space of Radon measures, so by the Hahn-Banach Theorem, it suffices to show that if such a measure ν satisfies

$$\int \phi \, d\nu = 0, \quad \text{for all } \phi \in \mathcal{D}(\Omega), \tag{3.3}$$

then $\nu \equiv 0$.

Let K be a cell and $\{U_i\}$ a sequence of open sets containing K so that $\nu(U_i \setminus K) \to 0$. Using Corollary 2.6 we see that for each i we can take $\phi_i \in \mathcal{D}(\Omega)$ with $\phi_i \equiv 1$ on K, the bound $\|\phi_i\|_\infty \leq C$ for all i, and $\text{Sppt}(\phi_i) \subset (U_i)$. Then for ν satisfying (3.3) we compute

$$\nu(K) = \left| \int \phi_i \, d\nu - \nu(K) \right| \leq \|\phi_j\|_\infty \nu(U_j \setminus K) \leq C\nu(U_j \setminus K) \to 0.$$

As ν vanishes on all cells it is the zero measure, and the result follows. □

Since Ω is locally compact and Hausdorff, it is a standard result that $C_c(\Omega)$ is supremum-norm dense in $C_0(\Omega)$, where the latter consists of those continuous functions f for which the set $\{x : |f(x)| \geq \epsilon\}$ is compact for all $\epsilon > 0$. Hence $\mathcal{D}(\Omega)$ is also dense in $C_0(\Omega)$.

In the special case where Ω is compact we may also characterize $\mathcal{D}(\Omega) = \mathcal{E}(\Omega)$ by the decay of the Fourier coefficients obtained when ϕ is written with respect to a basis of Laplacian eigenfunctions. This provides an alternate proof of the density of $\mathcal{D}(\Omega)$ in $C_c(\Omega)$, which of course coincides with $C(\Omega)$ in this case.

Theorem 3.5. *If Ω is compact then $\mathcal{D}(\Omega) = \mathcal{E}(\Omega)$ is the space of smooth functions with Fourier coefficients that have faster than polynomial decay, and hence is dense in $C(\Omega)$.*

Proof. Clearly $\phi \in \mathcal{D}(\Omega)$ is in L^2, so can be written $\phi = \sum_{i=0}^\infty a_i \psi_i$, where ψ_i is the Laplacian eigenfunction with eigenvalue $-\lambda_i$. It follows that $(-\Delta)^k \phi = \sum_i a_i \lambda_i^k \psi_i$ with convergence in L^2. Since $\Delta^k \phi$ is in $C(\Omega) \subset L^2$ for all k we see that the sequence a_i must decay faster than any polynomial in the λ_i. Conversely any such sequence gives rise to a function for which every power of the Laplacian is in L^2, whereupon the function is smooth by iteration of (2.2). In addition, any $u \in C(\Omega)$ can be explicitly approximated by functions from $\mathcal{D}(\Omega)$ by taking successive truncations of the Fourier series $u = \sum_{i=0}^\infty a_i \psi_i$. To see this gives convergence in $\mathcal{D}(\Omega)$ write $(-\Delta)^k \sum_{i=j}^\infty a_i \psi_i = \sum_{i=j}^\infty a_i \lambda_i^k \psi_i$ and note this converges to zero in L^2 and therefore almost everywhere. Now from (2.2)

$$\sup_\Omega \Big| \sum_{i=j}^\infty a_i \lambda_i^k \psi_i \Big|^2 \leq C \Big\| \sum_{i=j}^\infty a_i \lambda_i^k \psi_i \Big\|_2 \Big\| \sum_{i=j}^\infty a_i \lambda_i^{k+1} \psi_i \Big\|_2$$

and both terms on the right converge to zero. □

In [34] there is a definition of Sobolev spaces on p.c.f. fractals of the type studied here. These spaces may be defined by applying the Bessel potential $(I - \Delta)^{-s}$ (for the Dirichlet or Neumann Laplacian) or the Riesz potential $(-\Delta)^{-s}$ (for the Dirichlet Laplacian) to L^p functions on the fractal, and adding on an appropriate space of harmonic functions. In particular, the space of L^2 functions with $\Delta^k u \in L^2$ for $0 \leq k \leq m$ may be identified with a particular L^2 Sobolev space ([34] Theorem 3.7). Writing $W^{s,2}$ for the L^2 Sobolev space arising from $(I - \Delta)^{-s}$, we have in consequence of the preceding:

Corollary 3.6. *If Ω is compact then $\mathcal{D}(\Omega) = \cap_{s>0} W^{s,2}$.*

4. Distributions

Definition 4.1. The space of *distributions* on Ω is the dual space $\mathcal{D}'(\Omega)$ of $\mathcal{D}(\Omega)$ with the weak-star topology, so $T_i \to T$ if and only if $T_i\phi \to T\phi$ for all $\phi \in \mathcal{D}(\Omega)$.

As usual, the most familiar examples of distributions are the Radon measures. If $d\nu$ is such a measure then we define T_ν by $T_\nu\phi = \int \phi\, d\nu$. Theorem 3.4 ensures that the mapping $\nu \mapsto T_\nu$ is injective, so we may identify ν and T_ν. One way to obtain further examples is to apply the adjoint of the Laplacian to a distribution, which clearly produces another distribution.

Definition 4.2. If $T \in \mathcal{D}'(\Omega)$ we define $\Delta T \in \mathcal{D}'(\Omega)$ by $(\Delta T)\phi = T(\Delta\phi)$ for all $\phi \in \mathcal{D}(\Omega)$.

It is clear that powers of the Laplacian applied to the Radon measures provide a rich collection of examples of distributions. Later we prove that all distributions arise in essentially this way (Theorem 5.10), but we first need to establish some more elementary properties.

Theorem 4.3. *A linear functional T on $\mathcal{D}(\Omega)$ is a distribution if and only if for each compact $K \subset \Omega$ there are m and M such that*

$$|T\phi| \leq M|\phi|_m \tag{4.1}$$

Proof. It is clear that the existence of such an estimate ensures continuity of T. To prove the converse we assume no such estimate exists, so there is K compact and a sequence ϕ_j such that $|T\phi_j| \geq j|\phi_j|_j$. Then the support of $\tilde{\phi}_j = \phi_j/T\phi_j$ is in K for all j and

$$\left\|\Delta^k\tilde{\phi}_j\right\|_\infty \leq \frac{|\phi_j|_k}{|T\phi_j|} \leq \frac{1}{j} \quad \text{once } j \geq k.$$

Therefore $\tilde{\phi}_j \to 0$ in $\mathcal{D}(\Omega)$ but has $T\tilde{\phi}_j = 1$ for all j, contradicting the continuity of T. □

In the special case that Ω is compact we saw in Theorem 3.5 that $\mathcal{D}(\Omega)$ consists of smooth functions having Fourier coefficients that decay faster than polynomially. This allows us to identify its dual with coefficient sequences having at most polynomial growth. Recall that $\{\psi_i\}$ are the eigenfunctions of the Laplacian.

Lemma 4.4. *If Ω is compact and $T \in \mathcal{D}'(\Omega)$ then the sequence $T\psi_i$ has at most polynomial growth. Conversely, any sequence $\{b_i\}$ of polynomial growth defines a distribution via $\phi = \sum_i a_i\psi_i \mapsto \sum_i a_ib_i$.*

Proof. We saw in Theorem 3.5 that $\sum_{i=0}^j a_i\psi_i$ converges to ϕ in $\mathcal{D}(\Omega)$ if and only if $\{a_i\}$ has faster than polynomial decay in λ_i. It follows that for any $T \in \mathcal{D}'(\Omega)$, $\sum_{i=0}^j a_iT\psi_i = T\sum_{i=0}^j a_i\psi_i \to T\phi$, from which the sequence $T\psi_i$ has at most polynomial growth.

Conversely suppose that $\{b_i\}$ has polynomial growth, $|b_i| \leq C\lambda_i^m$, and consider the map $\phi = \sum_i a_i\psi_i \mapsto \sum_i a_ib_i$. This is a well defined linear map T on functions having Fourier coefficients with faster than polynomial decay, hence on $\mathcal{D}(\Omega)$, with the estimate

$$|T\phi| \leq \sum_i |a_ib_i| \leq C\sum_i |a_i|\lambda_i^m = C\|\Delta^m\phi\|_2 \leq C\|\Delta^m\phi\|_{\sup} \leq C|\phi|_m$$

which shows that T is a distribution. □

In particular, if we identify $\mathcal{D}(\Omega)$ as a subset of $\mathcal{D}'(\Omega)$ by letting $\phi' \in \mathcal{D}(\Omega)$ act on $\mathcal{D}(\Omega)$ via $\phi \mapsto \langle\phi, \phi'\rangle$, where $\langle, \rangle$ is the L^2 inner product, then this implies that the test functions are dense in the distributions when Ω is compact. To see this, let $T \in \mathcal{D}'(\Omega)$, so $\{T\psi_i\}$ has polynomial growth, and define a distribution T_j by

$$T_j\psi_i = \begin{cases} T\psi_i & \text{if } i \leq j \\ 0 & \text{if } i > j. \end{cases}$$

Then T_j is represented by a test function in the manner described above, and writing an arbitrary $\phi \in \mathcal{D}(\Omega)$ as $\phi = \sum_i a_i\psi_i$ we see $(T - T_j)\phi = \sum_{i=j}^{\infty} a_i T\psi_i \to 0$, so that $T_j \to T$ in $\mathcal{D}'(\Omega)$. This is true more generally.

Theorem 4.5. *$\mathcal{D}(\Omega)$ is dense in $\mathcal{D}'(\Omega)$.*

Proof. Let $T \in \mathcal{D}'(\Omega)$. We may assume Ω is non-compact, and since it has restricted cellular construction we can take an increasing exhaustion $\cup K_j$ of Ω by compact sets K_j with the property that K_j is contained in the interior of K_{j+1}, and each K_j is a finite union of cells. Recall $\mathcal{D}(K_j)$ denotes the space of test functions with support in K_j. For each j the action of T on $\mathcal{D}(K_j)$ identifies T as an element of $\mathcal{D}'(K_j)$ so by the preceding there is a sequence $\{T_{j,k}\}_{k=0}^{\infty} \subset \mathcal{D}(K_j)$ for which $T_{j,k} \to T$ in $\mathcal{D}'(K_j)$, and hence in $\mathcal{D}'(K_l)$ for all $l \leq j$.

Now consider the diagonal sequence $T_{j,j}$. For any test function ϕ there is some j_0 such that K_{j_0} contains the support of ϕ, and hence $T_{j,j}\phi$ is defined for $j \geq j_0$ and converges to $T\phi$. So $T_{j,j} \to T$ in $\mathcal{D}'(\Omega)$. Of course, $T_{j,j}$ only corresponds to a test function ϕ_j on K_j, not to an element of $\mathcal{D}(\Omega)$. To remedy this, note that for the test function ϕ_j corresponding to $T_{j,j}$ on K_j we may apply Theorem 2.4 to each of the (finite number of) boundary points of K_j and thereby continue ϕ_j smoothly to a function ϕ'_j on Ω with support in K_{j+1}. Denote by T'_j the distribution corresponding to this new test function ϕ'_j. Since ϕ_j and ϕ'_j coincide on K_j we see that $T_j\phi = T'_j\phi$ for all ϕ having support in K_j. It follows that T'_j converges to T in $\mathcal{D}'(\Omega)$, and since each T'_j corresponds to a test function, the proof is complete. □

Definition 4.6. If $\Omega_1 \subset \Omega$ is open, we say the distribution T *vanishes on* Ω_1 if $T\phi = 0$ for all ϕ supported on Ω_1. This is written $T\big|_{\Omega_1} = 0$,

To make a meaningful definition of the support of a distribution we again need the partitioning property. By Lemma 3.3 we know that any $\phi \in \mathcal{D}(\Omega_1 \cup \Omega_2)$ can be written as $\phi = \phi_1 + \phi_2$ for $\phi_j \in \mathcal{D}(\Omega_j)$. If both $T\phi_1 = 0$ and $T\phi_2 = 0$ it follows that $T\phi = 0$. We record this as a lemma, and note that it establishes the existence of a maximal open set on which T vanishes.

Lemma 4.7. *If $T\big|_{\Omega_1} = 0$ and $T\big|_{\Omega_2} = 0$ then $T\big|_{\Omega_1\cup\Omega_2} = 0$.*

Definition 4.8. The *support* of T is the complement of the maximal open set on which T vanishes, and is denoted $\mathrm{Sppt}(T)$. In the special case where $\mathrm{Sppt}(T)$ is compact we call T a *compactly supported distribution*.

Theorem 4.9. *The space of compactly supported distributions is (naturally isomorphic to) the dual $\mathcal{E}'(\Omega)$ of the smooth functions on Ω.*

Proof. The inclusion $\mathcal{D}(\Omega) \subset \mathcal{E}(\Omega)$ defines a natural map from $\mathcal{E}'(\Omega)$ to $\mathcal{D}'(\Omega)$. We have seen (after Lemma 3.3 above) that $\mathcal{D}(\Omega)$ is dense in $\mathcal{E}(\Omega)$, from which it follows that the kernel of this map is trivial. For convenience we identify $\mathcal{E}'(\Omega)$ with its isomorphic image in $\mathcal{D}'(\Omega)$, so we need only verify it is the space of distributions with compact support.

Fix an increasing sequence of compacta K_j with $\Omega = \cup_j K_j$. If $T \in \mathcal{E}'(\Omega)$ fails to be compactly supported then for each j there is ϕ_j supported in $\Omega \setminus K_j$ such that $T\phi_j \neq 0$, and by renormalizing we may assume $T\phi_j = 1$ for all j. However for any compact K there is j such that $K \subset K_j$ and thus $\phi_l \equiv 0$ on K once $l \geq j$. This implies $\phi_j \to 0$ in $\mathcal{E}(\Omega)$ and $T\phi_j = 1$ for all j, a contradiction.

Conversely, let $T \in \mathcal{D}'(\Omega)$ be supported on the compact $K \subset \Omega$. Fix a strictly larger compact K_1 (so that K is contained in the interior of K_1) and an open neighborhood Ω_1 of K_1. By Remark 2.9 the conclusion of Theorem 2.8 is valid for the cover by $\Omega \setminus K_1$ and Ω_1, even though Ω is noncompact. In particular we can fix a decomposition mapping in which $f \in \mathcal{E}(\Omega)$ is written as $f = f_1 + f_2$, with f_2 supported on $\Omega \setminus K_1$ and therefore $f_1\big|_{K_1} \equiv f\big|_{K_1}$. Now let T_1 on $\mathcal{E}(\Omega)$ be given by $T_1 f = T f_1$. This is well defined, because if $f, g \in \mathcal{E}(\Omega)$ and $f_1 = g_1$, then $f - g \equiv 0$ in a neighborhood of K and the support condition ensures $Tf = Tg$. It is also linear, even though the mapping $f \mapsto f_1$ is nonlinear (see Remark 2.10), because $(f + g)_1 = f_1 + g_1$ on K_1, which contains a neighborhood of K. Lastly, T_1 is continuous, as may be seen from the fact that a sequence $\{\phi_j\} \subset \mathcal{E}(\Omega)$ such that $\Delta^k \phi_j \to 0$ uniformly on compacta will have $\Delta^k(\phi_j)_1 \to 0$ on $K_1 \supset K$, or from (2.5). We conclude that every compactly supported distribution is in $\mathcal{E}'(\Omega)$. □

5. Structure Theory

Definition 5.1. A distribution T has *finite order m* if for each compact K there is $M = M(K)$ such that $|T\phi| \leq M|\phi|_m$ for all $\phi \in \mathcal{D}(K)$.

The following theorem indicates the importance of the finite order distributions.

Theorem 5.2. *Compactly supported distributions have finite order.*

Proof. Let T be a distribution with compact support K and let K_1 be a compact set such that $K \subset \text{int}(K_1)$. By Lemma 3.3 we may decompose any $\phi \in \mathcal{D}(\Omega)$ as $\phi = \phi_1 + \phi_2$ where ϕ_1 is supported on K_1, ϕ_2 is supported in K^c, and $|\phi_j|_m \leq C_m|\phi|_m$ for $j = 1, 2$. Clearly $T\phi = T\phi_1$, but there are m and M so that (4.1) holds on K_1, from which we conclude that

$$|T\phi| = |T\phi_1| \leq M|\phi_1|_m \leq C_m M|\phi|_m.$$ □

It is easy to see that the Radon measures on Ω are examples of distributions of finite order. In fact they have order zero, because the action of ν on $\mathcal{D}(\Omega)$ via $\nu\phi = \int \phi \, d\nu$ trivially satisfies the bound $|\nu\phi| \leq \|\phi\|_\infty = |\phi|_0$. The converse is also true.

Theorem 5.3. *If T is a distribution of order zero then there is a Radon measure ν such that $T\phi = \int \phi \, d\nu$ for all $\phi \in \mathcal{D}(\Omega)$.*

Proof. Let K be compact. Since T has order zero there is $M = M(K)$ so that $|T\phi| \leq M\|\phi\|_\infty$ whenever $\phi \in \mathcal{D}(\Omega)$ has support in K. Theorem 3.4 shows that these functions are dense in $C(K)$, so we may extend T to a bounded linear operator on $C(K)$. Such operators are represented by Radon measures, so there is ν_K with $T\phi = \int \phi \, d\nu_K$ for all test functions ϕ with support in K. Now let $\bigcup K_j$ be a compact exhaustion of Ω and consider the measures ν_{K_j}. These converge weak-star as elements of the dual of $C_c(\Omega)$ to a Radon measure ν, but by construction $\int \phi \, d\nu_{K_j} \to T\phi$ on $\mathcal{D}(\Omega)$ and the result follows. □

Remark 5.4. As written, the preceding proof relies on Theorem 3.4 and hence on the Hahn-Banach theorem. Since each of the K_j is compact, a constructive proof can be obtained by instead using Theorem 3.5.

A well known application of the preceding is obtaining a characterization of the distributions that have positive values on positive test functions. To prove this we need Theorem 2.7, and therefore must make the corresponding assumption (2.4) on the behavior of the heat kernel corresponding to the Laplacian.

Definition 5.5. T is a positive distribution if $T\phi \geq 0$ whenever $\phi \geq 0$ is a positive test function.

Theorem 5.6. *Positive distributions have order zero. If the Laplacian on X is such that* (2.4) *holds and if T is a positive distribution, then there is a positive measure ν such that* $Tf = \int f d\nu$.

Proof. Let K be compact. Using Theorem 2.6 there is $\psi_K \in \mathcal{D}(\Omega)$ such that $\psi \equiv 1$ on K. If $\phi \in C^\infty$ with support in K then the functions $\|\phi\|_\infty \psi_K \pm \phi$ are both positive, whence

$$-\|\phi\|_\infty T\psi_K \leq T\phi \leq \|\phi\|_\infty T\psi_K.$$

We conclude that T has order zero, so by Corollary 5.3 it is represented by integration against a measure ν. If there is a cell K for which $\nu(K) < 0$ then we can take U_j to be a neighborhood of K for which $\nu(U_j \setminus K) < 1/j$ and let f_j be as in Theorem 2.7. It follows that

$$Tf_j = \int f_j \, d\nu \leq \nu(K) + \frac{1}{j}\|f_j\|_\infty$$

and for a sufficiently large j this is negative, in contradiction to the positivity of T. We conclude that $\nu(K) \geq 0$ for all cells K, and therefore that ν is a positive measure. □

We noted at the beginning of Section 4 that the adjoint of the Laplacian maps $\mathcal{D}'(\Omega)$ to itself. In particular, if ν is a Radon measure, hence a distribution of order zero, then for each compact K there is $M(K)$ such that

$$|(\Delta^m \nu)\phi| = |\nu(\Delta^m \phi)| \leq M(K)|\Delta^m \phi|_0 \leq M(K)|\phi|_m$$

so $\Delta^m \nu$ is a distribution of order m. This result has a converse, which we prove using a modification of the Green's function introduced in Lemma 2.1. The basic idea is to produce a Green's operator that inverts the Laplacian on test functions, so that the adjoint of this operator lowers the order of a finite-order distribution. Iterating to produce a distribution of zero order then produces a measure by Theorem 5.3. We need some preliminary results about $\mathcal{D}(K)$, the space of test functions on Ω with support in K, in the case that K is a finite union of cells.

Lemma 5.7. *Let K be a connected finite union of cells in* Ω*. Then* $\mathcal{D}(K)$ *is dense in* $L^2(K)$.

Proof. Given f in $L^2(K)$ we extend to Ω by setting it to be zero outside K. Since K has finite measure, we can find $f_1 \in C(\Omega)$ with $\|f - f_1\|_{L^2(K)} < \epsilon$. Now $\mathcal{D}(\Omega)$ is sup norm dense in $C_c(\Omega)$, so we can find $f_2 \in \mathcal{D}(\Omega)$ with $\|f_1 - f_2\|_\infty < \epsilon\mu(K)^{-1/2}$ and hence $\|f_1 - f_2\|_{L^2(K)} < \epsilon$. Next, consider any boundary point q of K. Take small cells in K that form a neighborhood of q in K, and such that the measure of these cells is less than $\epsilon^2(1 + |f_2(q)|)^{-2}$. Applying Lemma 2.4 we can construct on each of these cells a smooth cutoff function with the property that its jet is equal to the jet of f_2 at q and all terms of the jets at other boundary points of the small cell are zero. Therefore subtracting all of these functions from f_2 does not change the smoothness on K, but makes all of the boundary jets vanish. Extending the resulting function f_3 by setting it to be zero outside K we then have $f_3 \in \mathcal{D}(K)$. Now $\|f_2 - f_3\|_{L^2(K)}$ is bounded by the sum of the L^2 norms of all of these cutoff functions, and by (2.3) with $k = L = 0$ we see that that the L^∞ norm of each of these functions is bounded by $|f_2(q)| + \epsilon$, hence the corresponding L^2 norm is bounded by 2ϵ. This shows that $\|f_2 - f_3\|_{L^2(K)} \leq 2\epsilon MN$, where M is number of points in the boundary of K and N is the maximum number of cells that can intersect at a cell boundary point (which is finite and depends only on the fractal), so the proof is complete. □

Lemma 5.8. *Let K be a connected finite union of cells in Ω. Then $\Delta : \mathcal{D}(K) \to \mathcal{D}(K)$ and its image consists of all test functions that are orthogonal to the harmonic functions on K. Moreover there is a linear operator $\tilde{G}_K : \mathcal{D}(K) \to \mathcal{D}(K)$ such that $-\Delta\tilde{G}_K(\Delta\phi) = \Delta\phi$ for all $\phi \in \mathcal{D}(K)$.*

Proof. If $K \neq \Omega$ then for $\phi \in \mathcal{D}(K)$ the matching conditions for the Laplacian ensure that both ϕ and $\partial_n\phi$ vanish on ∂K because ϕ is identically zero outside K. If $f \in \text{dom}(\Delta)$, then the Gauss-Green formula reduces to

$$\int_K (\Delta\phi) f \, d\mu = \int_K \phi(\Delta f) \, d\mu$$

because there are no non-zero boundary terms. This latter is also true when $K = \Omega$, because of Assumption 2.3 that Ω has no boundary. We conclude that f is orthogonal to the image $\Delta(\mathcal{D}(K))$ if and only if Δf is orthogonal to $\mathcal{D}(K)$. The latter is dense in $L^2(K)$ so the first result is proven.

Let $h_1 \dots, h_{i(K)}$ be an L^2–orthonormal basis for the finite dimensional space of harmonic functions on K. As $\mathcal{D}(K)$ is dense in $L^2(K)$ we can approximate each h_j by a test function $\tilde{h}_j$ sufficiently closely that the Gramm matrix $\langle \tilde{h}_j, h_j \rangle$ of L^2 inner products is invertible, and thereby construct $\phi_1, \dots, \phi_{i(K)}$ in $\mathcal{D}(K)$ such that $\langle \phi_i, h_j \rangle = \delta_{ij}$, where δ_{ij} is Kronecker's delta. Given $\phi \in \mathcal{D}(K)$ we let

$$\tilde{\phi} = \sum_{i=1}^{i(K)} \langle \phi, h_i \rangle \phi_i \tag{5.1}$$

and define

$$\tilde{G}_K\phi(x) = \begin{cases} G_K(\phi - \tilde{\phi})(x) & \text{if } x \in K \\ 0 & \text{if } x \notin K \end{cases}$$

where G_K is the Green's operator defined in Lemma 2.1. It is then clear that for $\psi \in \mathcal{D}(K)$,

$$-\Delta\tilde{G}_K\psi = -\Delta G_K(\psi - \tilde{\psi}) = \psi - \tilde{\psi} \tag{5.2}$$

except perhaps at points of ∂K, where we must first verify that the matching conditions for the Laplacian hold. Since $\tilde{G}_K\psi$ vanishes outside K, the matching conditions require that $\partial_n\tilde{G}_K\psi(q) = 0$ whenever $q \in \partial K$. One way to verify this is from the Gauss-Green formula for a harmonic function h, which yields

$$0 = \langle \psi - \tilde{\psi}, h \rangle = \int_K (-\Delta\tilde{G}_K\psi)h = -\sum_{q \in \partial K} (-\partial_n\tilde{G}_K\psi(q))h(q)$$

from which we see that it suffices to know the solvability of the Dirichlet problem on K, that is, for every assignment of boundary values on ∂K there is a harmonic function h with those boundary values. This latter is true because of Lemma 2.1; for example it may be proven by taking a function that is piecewise harmonic on cells and has the desired boundary data and subtracting the result of applying G_K to its Laplacian (which is simply a sum of Dirac masses at the interior gluing points). We conclude that $\tilde{G}_K\psi \in \mathcal{D}(K)$ and that (5.2) holds everywhere.

Finally, if $\psi = \Delta\phi$ for some $\phi \in \mathcal{D}(K)$, then ψ is orthogonal to the harmonics, so $\tilde{\psi} = 0$ and $-\Delta\tilde{G}_K\psi = \psi$ as desired. □

The adjoint of $\tilde{G}_K$ is defined on distributions by $(\tilde{G}_K T)\phi = T(\tilde{G}_K\phi)$. This operator is really defined on the dual of $\mathcal{D}(K)$, which is a larger space, but we will not make use of this fact.

Theorem 5.9. *If T is a distribution of order $m \geq 1$ then $\tilde{G}_K T$ is a distribution of order $m-1$, and if T is a distribution of order zero then $\tilde{G}_K T$ is integration with respect to a continuous function on K.*

Proof. Let T be a distribution of order $m \geq 1$, so that $\tilde{G}_K T\phi$ is bounded by

$$|\tilde{G}_K\phi|_m = \sup\{\|\Delta^k \tilde{G}_K(\phi)\|_\infty : k \leq m\}.$$

When $k \geq 1$ we have $\Delta^k \tilde{G}_K(\phi) = -\Delta^{k-1}(\phi - \tilde{\phi})$, and when $k = 0$ we see that $\|\tilde{G}_K(\phi)\|_\infty \leq C\|\phi - \tilde{\phi}\|_\infty$ because the operator G_K in Lemma 2.1 is clearly bounded on L^∞. Hence $|\tilde{G}_K\phi|_m \leq C|\phi - \tilde{\phi}|_{m-1} \leq C(m-1, K)|\phi|_{m-1}$, where the latter inequality is from (5.1) with a constant $C(m-1, K)$ that may depend on the set of functions ϕ_j. Thus $\tilde{G}_K T$ has order $m-1$.

If T has order zero then by Theorem 5.3 it is represented by integration against a Radon measure ν. Provided $K \neq \Omega$ we can apply Lemma 2.1 directly to see $\nu = \Delta f$ for some f that is continuous on K and vanishes on ∂K, so can be extended continuously to be zero outside K. This ensures there are no boundary terms when we compute with the Gauss-Green formula:

$$\begin{aligned}
\tilde{G}_K T\phi = T\tilde{G}_K\phi &= \int_K \tilde{G}_K\phi \, d\nu = \int_K (\tilde{G}_K\phi)(\Delta f) \, d\mu \\
&= \int_K (-\Delta\tilde{G}_K\phi) f \, d\mu = \int_K (\phi - \tilde{\phi}) f \, d\mu \\
&= \int_K \Big(\phi - \sum_{i=1}^{i(K)} \langle \phi, h_i \rangle \phi_i\Big) f \, d\mu \\
&= \int_K \phi \Big(f - \sum_{i=1}^{i(K)} \langle \phi_i, \bar{f} \rangle \bar{h}_i\Big) \, d\mu
\end{aligned}$$

and the bracketed term in the last line is continuous because f is continuous and all of the h_i are harmonic.

The argument is slightly different if $\Omega = K$. We instead set $t = \int d\nu / (\int d\mu)$ so $\int d(\nu - t\mu) = 0$, at which point Lemma 2.1 applies to show $\nu - t\mu = \Delta f$ for a continuous f, and we can compute as before

$$\begin{aligned}
\tilde{G}_\Omega T\phi &= \int_\Omega \tilde{G}_\Omega \phi \, d\nu \\
&= t \int_\Omega \tilde{G}_\Omega \phi \, d\mu + \int_\Omega \tilde{G}_\Omega \phi (\Delta f) \, d\mu \\
&= \int_\Omega (\phi - \tilde{\phi}) f \, d\mu \\
&= \int_\Omega \phi \Big(f - \int_\Omega f \, d\mu\Big) \, d\mu
\end{aligned}$$

where we used that $\int_\Omega \tilde{G}_\Omega \phi \, d\mu = 0$ (from the proof of Lemma 2.1) and that the harmonic functions are constants in this case. ◻

We now have all the necessary tools to prove the main structure theorem for distributions.

Theorem 5.10. *Any distribution T may be written as a locally finite sum of the form $T = \sum \Delta^{m_j} \nu_j$ or $T = \sum \Delta^{m_j+1} f_j$, where the ν_j are Radon measures and the f_j are continuous functions with compact support.*

Proof. Suppose first that Ω is non-compact and take $K_1, K_2, \ldots$ a sequence of subsets such that each K_j is a connected finite union of cells, K_j is contained in the interior of K_{j+1}, and $\cup_j K_j = \Omega$. Such a sequence exists because Ω has a restricted cellular construction. It will be convenient to also set $K_0 = \emptyset$. For each j let $\tilde{G}_j = \tilde{G}_{K_j}$ be the operator from Lemma 5.8. The key point of the proof is that for any distribution S, we have $(-\Delta)^m \tilde{G}_j^m S = S$ as elements of $\mathcal{D}'(K_j)$ (though not as elements of $\mathcal{D}'(\Omega)$). This may be verified by direct computation. First notice that if $l \geq 1$ and $\phi \in \mathcal{D}(K_j)$ then $(-\Delta)\tilde{G}_j((-\Delta)^l \phi) = (-\Delta)^l \phi$ by Lemma 5.8, so that $\tilde{G}_j((-\Delta)^l \phi) - (-\Delta)^{l-1}\phi$ is harmonic and vanishes on ∂K_j, so must be identically zero. Thus $\tilde{G}_j((-\Delta)^l \phi) = (-\Delta)^{l-1}\phi$ and inductively $\tilde{G}_j^m((-\Delta)^m \phi) = \phi$. Then for all $\phi \in \mathcal{D}(K_j)$,

$$(-\Delta)^m \tilde{G}_j^m S\phi = \tilde{G}_j^m S\big((-\Delta)^m \phi\big) = S\big(\tilde{G}_j^m (-\Delta)^m \phi\big) = S\phi. \tag{5.3}$$

Fix $T \in \mathcal{D}'(\Omega)$. Inductively suppose that for $i = 0, \ldots, j-1$ we have found m_i and a measure ν_i supported on K_i such that $T - \sum_0^{j-1} \Delta^{m_i}\nu_i$ vanishes on $\mathcal{D}(K_{j-1})$. The base case $j = 0$ is trivial because $K_0 = \emptyset$. Now $T - \sum_0^{j-1} \Delta^{m_i}\nu_i$ is in $\mathcal{D}'(\Omega)$, hence its restriction to $\mathcal{D}(K_j)$ is in $\mathcal{D}'(K_j)$. We call this restriction T_j. As K_j is compact, T_j has finite order m_j. It satisfies $T_j = (-\Delta)^{m_j}\tilde{G}_j^{m_j} T_j$ by the argument already given, and by Theorem 5.9 there is a measure ν_j supported on K_j such that $\nu_j = (-1)^{m_j}\tilde{G}_j^{m_j} T_j$. Therefore $T_j = \Delta^{m_j}\nu_j$ in $\mathcal{D}'(K_j)$, which is equivalent to saying that $T - \sum_0^j \Delta^{m_i}\nu_i$ vanishes on $\mathcal{D}(K_j)$.

It is immediate from the definition that $\sum_j \Delta^{m_j}\nu_j$ is a locally finite sum. If we fix $\phi \in \mathcal{D}(\Omega)$ then there is a j such that $\phi \in \mathcal{D}(K_j)$, whereupon $(T - \sum_{i=1}^l \Delta^{m_i}\nu_i)\phi = 0$ for all $l \geq j$. This proves that $T = \sum_j \Delta^{m_j}\nu_j$.

The proof that $T = \sum_j \Delta^{m_j+1} f_j$ is similar. Obviously we wish to use the latter part of Theorem 5.9 to go from the measure ν_j to a continuous function. The only technicality is that the resulting f_j is continuous on K_j rather than on all of Ω. We fix this at each step of the induction as follows. Suppose we have determined T_j as the restriction of $T - \sum_{i=0}^{j-1} \Delta^{m_i+1} f_i$ and from Theorem 5.9 a function g_j continuous on K_j such that $\Delta^{m_j+1} g_j = T_j$ in $\mathcal{D}'(K_j)$. Let f_j be a continuous extension of g_j to Ω obtained by requiring $f_j = 0$ on ∂K_{j+1} and outside K_{j+1}, and letting it be piecewise harmonic on the cells of the cellular structure on $K_{j+1} \setminus K_j$ (here we use that K_j is in the interior of K_{j+1}). Clearly $\Delta^{m_j+1} f_j = T_j$ in $\mathcal{D}'(K_j)$, because we $f_j = g_j$ on K_j, so $T - \sum_{i=0}^j \Delta^{m_i+1} f_i$ vanishes on $\mathcal{D}(K_j)$ and we may complete the proof as before.

In the case when Ω is compact the proof is somewhat more elementary because we need only a single set $K = \Omega$, but there is a small technical difference due to the fact that the final equality of (5.3) is no longer true. Indeed if we follow the above proof we find that $\tilde{G}_K((-\Delta)^l \phi) - (-\Delta)^{l-1}\phi$ is harmonic by the same reasoning as for the non-compact case, but now the harmonic functions are constants. By definition (see the last line of the proof of Lemma 2.1), the constant part of $\tilde{G}_K((-\Delta)^l \phi)$ is zero, and if $l > 1$ then $(-\Delta)^{l-1}\phi$ is orthogonal to the constants as shown in Lemma 5.8. Thus $\tilde{G}_K((-\Delta)^l \phi) = (-\Delta)^{l-1}\phi$ for all $l > 1$, but $\tilde{G}_K((-\Delta)\phi) = \phi - \tilde{\phi}$, where $\tilde{\phi} = (\mu(K))^{-1} \int \phi \, d\mu$ is constant. The analogue of (5.3) is therefore

$$(-\Delta)^m \tilde{G}_K^m S\phi = S\big(\tilde{G}_j^m (-\Delta)^m \phi\big) = S\phi - \tilde{\phi} S 1$$

where 1 is the constant function.

Applying this to a distribution T of finite order m, we write $T\phi = \tilde{\phi} T 1 + (-\Delta)^m \tilde{G}_K^m T\phi$ and note $(-1)^m \tilde{G}_K^m T$ is a measure ν by Theorem 5.9. Since $\tilde{\phi} T 1$ is the result of applying the distribution $T1\mu/\mu(K)$ to ϕ we conclude $T = (T1/\mu(K))\mu + \Delta^m \nu$. If instead we apply the version with power $m+1$ then we obtain $T\phi = \tilde{\phi} T 1 + (-\Delta)^{m+1}\tilde{G}_K^{m+1} T\phi$. In this expression

$(-1)^{m+1}\tilde{G}_K^{m+1}T$ is a continuous function f, and $\tilde{\phi}T1$ is the result of applying the distribution given by the constant function $T1/\mu(K)$ to ϕ. Hence $T = T1/\mu(K)+\Delta^{m+1}f$. This completes the proof. ◻

6. Distributions supported at a point

A distribution with support a point q is of finite order by Theorem 5.2, and simple modifications of the arguments in Theorem 5.10 show that it is a power of the Laplacian applied to a measure with support in a neighborhood of q. The purpose of this section is to identify it more precisely as a finite sum of certain derivatives of the Dirac mass at q; in general these derivatives are not just powers of the Laplacian, but instead reflect the local structure of harmonic functions at q.

Identification of a distribution T of order m supported at q is achieved by describing a finite number of distributions T_j, $j = 1, \dots, J$ with the following property: if $\phi \in \mathcal{D}(\Omega)$ has $T_j\phi = 0$ for all j then for any $\epsilon > 0$ there is a neighborhood U_ϵ of q and a decomposition $\phi = \phi_q + (\phi - \phi_q)$ into test functions such that $|\phi_q|_m < \epsilon$ and $\phi - \phi_q$ vanishes on U_ϵ. The reason is that then $T\phi = T\phi_q$ because of the support condition, and $|T\phi_q| \leq C\epsilon$, from which we conclude that T vanishes whenever all T_j vanish. It follows from a standard argument (for example, Lemma 3.9 of [29]) that T is a linear combination of the T_j.

The argument described in the previous paragraph motivates us to find conditions on a test function ϕ that ensure we can cut if off outside a small neighborhood of a point q while keeping the norm $|\phi_q|_m$ of the cutoff small. In order to proceed we will need some notation for a neighborhood base of q. If q is a non-junction point then it lies in a single copy of X in the cellular structure, and within this copy there is a unique word w such that $F_w(X) = q$. The cells containing q are then of the form $U_i = F_{[w]_i}(X)$. For junction points the situation is different, as q can be the intersection point of several copies of X, or can be in a single copy but be given by $F_{w_j}(X) = q$ for a finite number of words $w_1, \dots, w_J$. We will not distinguish between these possibilities but will instead make the convention that the distinct words determining q may be used to distinguish copies of X if necessary. With this assumed, let $U_{i,j} = F_{[w_j]_i}(X)$, and $U_i = \cup_j U_{i,j}$.

Fix q and let $G_{i,j}$ denote the Dirichlet Green's operator on the cell $U_{i,j}$, omitting the j index if q is a non-junction point. If $\phi \in \mathcal{D}(\Omega)$ we can then decompose ϕ on $U_{i,j}$ into

$$\phi\big|_{U_{i,j}} = H_{i,j}\phi + G_{i,j}\Delta\phi$$

where $H_{i,j}\phi$ is the (unique) harmonic function on $U_{i,j}$ whose values on $\partial U_{i,j}$ coincide with those of $\phi\big|_{U_{i,j}}$. By induction we obtain

$$\phi\big|_{U_{i,j}} = \sum_{l=0}^{m-1} G_{i,j}^l H_{i,j}\Delta^l\phi \ + G_{i,j}^m\Delta^m\phi\big|_{U_{i,j}} \tag{6.1}$$

and write $h_{i,j}^l = G_{i,j}^l H_{i,j}\Delta^l\phi$.

Lemma 6.1. *In the decomposition* (6.1) *we have at each $x \in U_{i,j}$ and $p \in \partial U_{i,j}$ that $\Delta^k h_{i,j}^l(x) = \partial_n\Delta^k h_{i,j}^l(p) = 0$ if $k > l$, while for $k \leq l$,*

$$\begin{aligned}\left|\Delta^k h_{i,j}^l\right| &\leq c(k,l) r_{[w_j]_i}^{l-k}\mu_{[w_j]_i}^{l-k}\left\|H_{i,j}\Delta^l\phi\right\|_{L^\infty(U_{i,j})}\\ \left|\partial_n\Delta^l h_{i,j}^l(p)\right| &\leq c(k,l) r_{[w_j]_i}^{l-k-1}\mu_{[w_j]_i}^{l-k}\left\|H_{i,j}\Delta^l\phi\right\|_{L^\infty(U_{i,j})}\end{aligned}$$

Proof. The conclusions for the cases $k > l$ are immediate from the fact that $\Delta^l h^l_{i,j}$ is harmonic, while the remaining estimates are derived from the fact that the Laplacian scales by $r_w\mu_w$ on a cell $F_w(X)$ while the normal derivative scales by r_w. □

Our purpose in making the above definitions is that estimates on the functions $h^l_{i,j}$ are precisely what is needed to ensure we can cut off a smooth function in the manner previously described.

Theorem 6.2. *If ϕ is such that $\Delta^m\phi(q) = 0$ and $\|H_{i,j}\Delta^l\phi\|_{L^\infty(U_{i,j})} = o\Big(r_{[w_j]_i}\mu_{[w_j]_i}\Big)^{m-l}$ for $0 \leq l \leq m-1$ as $i \to \infty$, then for all $\epsilon > 0$ there is ϕ_q such that $|\phi_q|_m \leq \epsilon$ and $\phi - \phi_q$ is supported away from q.*

Proof. We begin by constructing a neighborhood of $U_{i,j}$ by adjoining cells at each of the points $p \in \partial U_{i,j}$. At each p we require finitely many such cells, and we choose them so as to intersect $U_{i,j}$ only at p. It will also be convenient to assume that these cells have comparable scale to the $U_{i,j}$, in the sense that they have the form $F_{\tilde{w}}(X)$ for some word with length $i \leq |\tilde{w}| \leq i + i_0$ for some constant i_0. Let K be one of the cells adjoined at a point p, and let n_p be the number of cells adjoined at p. Using Theorem 2.4 we define a smooth function f_K on K with jet $\Delta^k f_K(p) = \Delta^k\phi(p)$ and $\partial^K_n \Delta^k f_K(p) = -(1/n_p)\partial^{U_{i,j}}_n \Delta^k\phi(p)$, and with vanishing jets at the other boundary points of K. Having done this for the set $\mathcal{K}$ of adjoined cells we see from the matching conditions for the Laplacian that

$$\phi_q(x) = \begin{cases} \phi(x) & \text{for } x \in U_i \\ \sum_{K\in\mathcal{K}} f_K & \text{for } x \in \bigcup_{K\in\mathcal{K}} K \\ 0 & \text{otherwise} \end{cases}$$

defines a test function with the property that $\phi - \phi_q = 0$ on U_i.

We must estimate $|\phi_q|_m$. There is an easy estimate for $\Delta^k\phi$ for $k \leq m$ from Lemma 6.1:

$$\begin{aligned} |\Delta^k\phi| &\leq \sum_{l=0}^{m-1} \left|\Delta^k h^l_{i,j}\right| + \left|G^{m-k}_{i,j}\Delta^m\phi\right| \\ &\leq \sum_{l=k}^{m-1} c(k,l) r^{l-k}_{[w_j]_i}\mu^{l-k}_{[w_j]_i} \left\|H_{i,j}\Delta^l\phi\right\|_{L^\infty(U_{i,j})} + c(k,m) r^{m-k}_{[w_j]_i}\mu^{m-k}_{[w_j]_i}\left\|\Delta^m\phi\right\|_{L^\infty(U_{i,j})} \\ &\leq \sum_{l=k}^{m} c(k,l) o\big(r^{m-k}_{[w_j]_i}\mu^{m-k}_{[w_j]_i}\big) \\ &= o\big(r^{m-k}_{[w_j]_i}\mu^{m-k}_{[w_j]_i}\big) \end{aligned} \tag{6.2}$$

where we used $\Delta^m\phi(q) = 0$ to obtain that $\Delta^m\phi(q) = o(1)$ when $i \to \infty$. As a result we have good control of $|\phi_q|_m$ on $U_{i,j}$.

A similar calculation allows us to estimate the size of the normal derivative $|\partial_n \Delta^k \phi(p)|$ at any of the points p where pieces f_K are attached. We compute

$$\begin{aligned}
|\partial_n \Delta^k \phi(p)| &\leq \sum_{l=0}^{m-1} \left|\partial_n \Delta^k h_{i,j}^l\right| + \left|\partial_n G_{i,j}^{m-k} \Delta^m \phi\right| \\
&\leq \sum_{l=k}^{m-1} c(k,l) r_{[w_j]_i}^{l-k-1} \mu_{[w_j]_i}^{l-k} \left\|H_{i,j}\Delta^l \phi\right\|_{L^\infty(U_{i,j})} + c(k,m) r_{[w_j]_i}^{m-k-1} \mu_{[w_j]_i}^{m-k} \left\|\Delta^m \phi\right\|_{L^\infty(U_{i,j})} \\
&\leq \sum_{l=k}^{m} c(k,l) o\big(r_{[w_j]_i}^{m-k-1} \mu_{[w_j]_i}^{m-k}\big) \\
&= o\big(r_{[w_j]_i}^{m-k-1} \mu_{[w_j]_i}^{m-k}\big).
\end{aligned} \tag{6.3}$$

Fix $K \in \mathcal{K}$ and examine f_K. By assumption $K = F_{\tilde{w}}(X)$, so by (2.3) with the fixed number of jet terms m we know

$$\|\Delta^k f_K\|_\infty \leq C(k) \Big(\sum_{k'=0}^{m} r_{\tilde{w}}^{k'-k} \mu_{\tilde{w}}^{k'-k} \left|\Delta^{k'} \phi(p)\right| + \sum_{k'=0}^{m-1} r_{\tilde{w}}^{k'+1-k} \mu_{\tilde{w}}^{k'-k} \left|\partial_n^{U_{i,j}} \Delta^{k'} \phi(p)\right| \Big) + \epsilon \tag{6.4}$$

provided $0 \leq k \leq m$. The terms involving $\left|\Delta^{k'} \phi(p)\right|$ may be replaced by the estimate (6.2). For the terms involving normal derivatives we use that $\partial_n \Delta^{k'} h_{i,j}^{k'}(p) = (1/n_p) \partial_n \Delta^{k'} \phi$ and (6.3). The result is

$$\begin{aligned}
\|\Delta^k f_K\|_\infty &\leq C(k) \Big(\sum_{k'=0}^{m} o\big(r_{\tilde{w}}^{k'-k} \mu_{\tilde{w}}^{k'-k} r_{[w_j]_i}^{m-k'} \mu_{[w_j]_i}^{m-k'}\big) + \sum_{k'=0}^{m-1} o\big(r_{\tilde{w}}^{k'+1-k} \mu_{\tilde{w}}^{k'-k} r_{[w_j]_i}^{m-k'-1} \mu_{[w_j]_i}^{m-k'}\big) \Big) + \epsilon \\
&\leq o\big(r_{[w_j]_i}^{m} r_{\tilde{w}}^{-k} \mu_{[w_j]_i}^{m} \mu_{\tilde{w}}^{-k}\big) \sum_{k'=0}^{m} \Big(\frac{r_{\tilde{w}} \mu_{\tilde{w}}}{r_{[w_j]_i} \mu_{[w_j]_i}} \Big)^{k'} \Big(1 + \frac{r_{\tilde{w}}}{r_{[w_j]_i}} \Big) + \epsilon.
\end{aligned} \tag{6.5}$$

However, $\tilde{w}$ and $[w_j]_i$ have comparable length and are adjacent, so they differ only in the final i_0 letters and therefore the ratios $r_{\tilde{w}} r_{[w_j]_i}^{-1}$ and $\mu_{\tilde{w}} \mu_{[w_j]_i}^{-1}$ are bounded by constants depending only on i_0 and the harmonic structure and measure. It follows that

$$\|\Delta^k f_K\|_\infty \leq C(m,k,r,\mu) o\big(r_{[w_j]_{i-1}} \mu_{[w_j]_{i-1}}\big)^{m-k}$$

and combining this estimate for each $K \in \mathcal{K}$ with (6.2) proves that

$$\|\Delta^k \phi_q\|_{L^\infty} = o\big(r_{[w_j]_{i-1}} \mu_{[w_j]_{i-1}}\big)^{m-k} \quad \text{as } i \to \infty$$

for $0 \leq k \leq m$. In particular we can make $|\phi_q|_m < \epsilon$ by making i sufficiently large. □

Theorem 6.2 suggests that the natural candidates for the distributions supported at q are appropriately scaled limits of the maps $\phi \mapsto H_{i,j} \Delta^l \phi$ as $i \to \infty$. The question of how to take such limits has been considered by a number of authors [16, 39, 26, 27, 1], and is generally quite complicated. At the heart of this complexity is the fact that the local behavior of smooth functions in a neighborhood of a point q depends strongly (in fact almost *entirely*) on the point q rather than the function itself. This property – often called "geography is destiny" – contrasts sharply with the Euclidean situation where neighborhoods of points are analytically indistinguishable. Its immediate implication for the structure of distributions with point support is that the nature of these distributions must depend on the point in question. In order of increasing complexity we consider three cases: cell boundary points, periodic points and a class of measure-theoretically generic points.

Cell Boundary Points. A cell boundary point q is of the form $q = F_{w_j}(X)$ for words $w_1, \dots w_J$, each of which terminates with an infinite repetition of a single letter. The distributions corresponding to approaching q through the sequence $[w_j]_i$ may be understood by examining the eigenstructure of the harmonic extension matrices A_{i_j}, the definition of which appeared in the Harmonic Functions part of Section 2.

For notational convenience we temporarily fix one contraction F, let A be the corresponding harmonic extension matrix, and suppose q is $\cap F^i(X)$. Let r and μ be the resistance and measure scalings of F, and $\gamma_1, \dots, \gamma_n$ be the eigenvalues of A, ordered by decreasing absolute value, with eigenspaces $E_1, \dots, E_n$. Of course $\gamma_1 = 1$ and E_1 is the constant functions. Let $H_i u$ be the harmonic function on $F^i(X)$ that equals u on $\partial F^i(X)$. Fix an inner product and let P_1 be projection onto E_1 and P_s for $s > 1$ be projection onto the orthogonal complement $\tilde{E}_s$ of $E_1 + \cdots + E_{s-1}$ in $E_1 + \cdots + E_s$. In what follows, G is the Dirichlet Green's operator on X and G_i is the same on $F^i(X)$.

Definition 6.3. Inductively define derivatives d_s and differentials D^k, $k \geq 1$ at the point q by setting $D^0 u = 0$, and for each s such that $(r\mu)^k < |\gamma_s| \leq (r\mu)^{k-1}$

$$d_s u = \lim_{i \to \infty} \gamma_s^{-i} P_s H_i (u - G D^{k-1} \Delta u) \tag{6.6}$$

if these limits exist. Note that d_s always exists for harmonic functions as the fact that $P_s A = \gamma_s P_s$ implies the sequence is constant in this case. Provided the necessary $d_s u$ exist we then let

$$D^k u = h + G D^{k-1} \Delta u \tag{6.7}$$

where h is the unique harmonic function on X with $d_s h = d_s u$ for those s with $|\gamma_s| > (r\mu)^k$ and $d_s h = 0$ for all other s. We will also make use of $\bar{D}^k u$, where $\bar{D}^0 = u(q)$ is a constant function and

$$\bar{D}^k u = \bar{h} + G \bar{D}^{k-1} \Delta u$$

where $\bar{h}$ is harmonic on X with $d_s \bar{h} = d_s u$ for those s with $|\gamma_s| \geq (r\mu)^k$ and $d_s \bar{h} = 0$ for all other s.

The reader might find it helpful at this point to refer to Example 6.8 where these derivatives are constructed explicitly for the case of the Sierpinski Gasket.

Lemma 6.4. *For $u \in \operatorname{dom}(\Delta^k)$ and each s with $|\gamma_s| > (r\mu)^k$ the derivative $d_s u$ exists, and*

$$|d_s u| \leq C(k) \sum_{l=0}^{k} \|\Delta^l u\|_\infty. \tag{6.8}$$

The differential satisfies

$$\left\| u - D^k u \right\|_{L^\infty(F^i(X))} \leq C(k) i^k (r\mu)^{ki} \|\Delta^k u\|_\infty, \tag{6.9}$$

and if we further suppose that $\Delta^k u \in \operatorname{dom}(E)$ then

$$\left\| u - \bar{D}^k u \right\|_{L^\infty(F^i(X))} \leq C(k) (r\mu)^{ki} r^{i/2} E^{1/2}(\Delta^k u). \tag{6.10}$$

Proof. The proof is inductive. When $k = 0$ there are no s with $|\gamma_s| > 1 = (r\mu)^0$, so the first statement is vacuous and (6.9) is immediate. Suppose both hold up to $k - 1$.

Write $u - G D^{k-1} \Delta u$ as $H_0 u + G(\Delta u - D^{k-1} \Delta u)$, from which

$$d_s u = d_s H_0 u + \lim_i \gamma_s^{-i} P_s H_i (G \Delta u - G D^{k-1} \Delta u) \tag{6.11}$$

provided the latter limit exists. On the cell $F^i(X)$,

$$G(\Delta u - D^{k-1} \Delta u) = H_i G(\Delta u - D^{k-1} \Delta u) + G_i (\Delta u - D^{k-1} \Delta u)$$

thus

$$H_{i+1}G(\Delta u - D^{k-1}\Delta u) = AH_iG(\Delta u - D^{k-1}\Delta u) + H_{i+1}G_i(\Delta u - D^{k-1}\Delta u).$$

In particular, if we project by P_s then the action of A is multiplication by γ_s (because $P_sA = \gamma_sP_s$). Scaling and (6.9) implies $G_i(\Delta u - D^{k-1}\Delta u)$ is bounded by

$$\Big|G_i(\Delta u - D^{k-1}\Delta u)\Big| \le C(r\mu)^i \Big\|\Delta u - D^{k-1}\Delta u\|_{L^\infty(F^i(X))} \le CC(k-1)i^{k-1}(r\mu)^{ik}\|\Delta^k u\|_\infty \quad (6.12)$$

and the action of H_{i+1} and P_s can only improve this estimate, so

$$\begin{aligned} |\gamma_s|^{-(i+1)}\Big|P_sH_{i+1}G(\Delta u - D^{k-1}\Delta u) - \gamma_sP_sH_iG(\Delta u - D^{k-1}\Delta u)\Big| \\ \le |\gamma_s|^{-(i+1)}\Big|G_i(\Delta u - D^{k-1}\Delta u)\Big| \\ \le CC(k-1)i^{k-1}\Big(\frac{r^k\mu^k}{|\gamma_s|}\Big)^i\|\Delta^k u\|_\infty \end{aligned} \quad (6.13)$$

This shows $\{\gamma_s^{-i}P_sH_iG(\Delta u - D^{k-1}\Delta u)\}$ is Cauchy when $|\gamma_s| > (r\mu)^k$, and that its limit is bounded by $C(k)\|\Delta^k u\|_\infty$. It follows from (6.11) that d_su exists for these values of s, and since $|d_sH_0u| \le \|u\|_\infty$ we also obtain (6.8).

Summing the tail of (6.13) establishes that

$$\Big|d_su - \gamma_s^{-i}P_sH_i(u - GD^{k-1}\Delta u)\Big| \le C(k)\Big(\frac{r^k\mu^k}{|\gamma_s|}\Big)^i\|\Delta^k u\|_\infty.$$

Now let h be the unique harmonic function with $d_sh = d_su$ for those s with $|\gamma_s| > (r\mu)^k$ and $d_sh = 0$ otherwise. Since $\gamma_s^{-i}P_sH_ih = d_sh$ is a constant sequence, we find

$$\Big|P_sH_i(u - h - GD^{k-1}\Delta u)\Big| \le C(k)(r\mu)^{ik}\|\Delta^k u\|_\infty \quad (6.14)$$

for those s with $|\gamma_s| > (r\mu)^k$. Recalling $D^ku = h + GD^{k-1}\Delta u$ from (6.7) write

$$\begin{aligned} (u - D^ku)\Big|_{F^i(X)} &= H_i(u - D^ku) + G_i(\Delta(u - D^ku)) \\ &= H_i(u - h - GD^{k-1}\Delta u) + G_i(\Delta u - D^{k-1}\Delta u). \end{aligned} \quad (6.15)$$

We have estimated $G_i(\Delta u - D^{k-1}\Delta u)$ in (6.12) and the terms $P_sH_i(u - h - GD^{k-1}\Delta u)$ for $|\gamma_s| > (r\mu)^k$ in (6.14). What remains are the terms $P_sH_i(u - h - GD^{k-1}\Delta u)$ for $|\gamma_s| \le (r\mu)^k$. Each of these is obtained as a sum, with

$$\begin{aligned} \Big|P_sH_i(u - h - GD^{k-1}\Delta u)\Big| &= \Big|\sum_{j=0}^{i-1}\gamma_s^{i-j}P_sH_jG_j(u - h - GD^{k-1}\Delta u)\Big| \\ &\le CC(k-1)\sum_{j=0}^{i-1}|\gamma_s|^{i-j}j^{k-1}(r\mu)^{jk}\|\Delta^k u\|_\infty \\ &\le CC(k-1)(r\mu)^{ik}\|\Delta^k u\|_\infty\sum_{j=0}^{i-1}j^{k-1}\Big(\frac{|\gamma_s|}{(r\mu)^k}\Big)^{(i-j)k} \\ &\le C(k)i^k(r\mu)^{ik}\|\Delta^k u\|_\infty \end{aligned} \quad (6.16)$$

because $|\gamma_s| \le (r\mu)^k$. This proves (6.9) for k and completes the induction.

The proof of (6.10) uses essentially the same inductive argument with $\bar{D}$ replacing D and the estimate from (6.10) replacing that from (6.9) throughout. Note that in (6.13) we can have $|\gamma_s| \ge (r\mu)^k$ because there is an additional factor of $r^{i/2}$ so the series still converges geometrically. Also, in (6.16) the working is simplified because for $\bar{D}$ we have these $|\gamma_s| < (r\mu)^k$ and the $r^{i/2}$ term is bounded, so the convergence is geometric here also.

This allows us to remove the polynomial term in i. The base case $k = 0$ is true because of the Hölder estimate (2.2). □

The map d_s takes a smooth function to the orthogonal complement $\tilde{E}_s$ of $E_1 + \cdots + E_{s-1}$ in $E_1 + \cdots + E_s$. We now fix orthonormal bases for each of the $\tilde{E}_s$, and refer to the coordinates of d_s with respect to the basis for $\tilde{E}_s$ as the *components* of d_s; these components have values in $\mathbb{C}$.

Corollary 6.5. *Each component $d_{s,v}$ of a d_s for which $(r\mu)^k < |\gamma_s| \leq (r\mu)^{k-1}$ is a distribution supported at q and of order at most k. If $|\gamma_s| < (r\mu)^{k-1}$ then its order is equal to k, and it is otherwise of order either $k-1$ or k. If $d_{s,v}$ is a component that is a distribution of order k, then $\Delta^l d_{s,v}$ defined by $\Delta^l d_{s,v}\phi = d_{s,v}\Delta^l\phi$ is also supported at q and has order $k + l$.*

Proof. It is apparent from the definition that d_s is linear on $\mathcal{D}(\Omega)$ and that $d_s\phi = 0$ if $\phi \in \mathcal{D}(\Omega)$ is identically zero in a neighborhood of q, so it follows from (6.8) that the components of d_s are distributions of order at most k and are supported at q.

Suppose $|\gamma_s| < (r\mu)^{k-1}$ and let $u_{s,v}$ denote the harmonic function with $d_{s,v} = 1$ and all other d_t equal zero. Then the values of $u_{s,v}$ are $O(\gamma_s) = o(r\mu)^{k-1}$ and $\Delta^l u = 0$ for $l \leq 1$, so Theorem 6.2 implies that for any $\epsilon > 0$ there is a function ψ equal to $u_{s,v}$ in a neighborhood of q but with $|\psi|_{k-1} < \epsilon$. Since $d_{s,v}u_{s,v} = 1$ and $d_{s,v}u_{s,v} = d_{s,v}\psi$ because of the support condition, it cannot be that $d_{s,v}$ is order $k-1$ or less, so it has order k.

In the case $|\gamma_s| = (r\mu)^{k-1} < (r\mu)^{k-2}$ the above argument says that $d_{s,v}$ has order at least $k-1$. Both of the values $k-1$ and k occur in examples. For instance, when $k = 1$, the derivative $d_1 u = u(q)$ corresponding to the constant harmonic functions has order $0 = k-1$. A case where there is a d_s of this type with order k occurs on the Sierpinski Gasket, see Example 6.8 below. This shows that scaling alone cannot identify the order of d_s when $|\gamma_s| = (r\mu)^{k-1}$.

The statement regarding $\Delta^l d_s$ is immediate. □

We now return to using the index j to distinguish the words w_j for which $x = F_{w_j}(X)$, and accordingly denote by d_s^j the derivative d_s corresponding to the approach through cells $F_{[w_j]_i}$.

Theorem 6.6. *Let T be a distribution of order k supported at the cell boundary point q, where $q = F_{w_j}(X)$, $j = 1, \dots, n$. The word w_j terminates with infinite repetition of a letter which, by a suitable relabeling we assume is j. Then T is a finite linear combination of the distributions $\Delta^l d_{s,v}^j$, for which $|\gamma_s| \geq (r_j\mu_j)^{k-l}$. The linear combination runs over all such s, all basis elements v for $\tilde{E}_s$, and all cells $j = 1, \dots, n$ that meet at q.*

Proof. Suppose that $\phi \in \mathcal{D}(\Omega)$ has the property that $\Delta^l d_{s,v}^j\phi = 0$ for all $(r_j\mu_j)^{k-l} \leq |\gamma_s|$. It follows from Definition 6.3 that $\bar{D}^k\phi = 0$ and more generally that $\bar{D}^{k-l}\Delta^l\phi = 0$ for all $l \leq k$.

However, the harmonic part of $H_{i,j}\Delta^l\phi = H_{i,j}\Delta^l(\phi - \bar{D}^k\phi)$ on the cell $U_{i,j}$ of scale i corresponding to the word w_j is bounded by the maximum over the boundary vertices of this cell, so from (6.10):

$$\left\|H_{i,j}\Delta^l\phi\right\|_{L^\infty(U_{i,j})} = o(r_j\mu_j)^{(k-l)i} = o\big(r_{[w_j]_i}\mu_{[w_j]_i}\big)^{k-l}$$

for $0 \leq l \leq k-1$. We also have that $\Delta^k\phi(q) = 0$ because $\Delta^l\phi(q) = \Delta^l d_{1,v}^j\phi = 0$, so Theorem 6.2 shows that for any $\epsilon > 0$ there is $\psi \in \mathcal{D}(\Omega)$ that is equal to $\phi - \bar{D}^k\phi$ in a neighborhood of q and with $|\psi|_k < \epsilon$.

Using the support condition and the fact that T has order k yields

$$|T\phi| = |T\psi| \leq M|\psi|_k < M\epsilon$$

for some fixed M depending only on T, and all $\epsilon > 0$. Thus $T\phi = 0$, and we have shown that the kernel of T contains the intersection of the kernels of the distributions described. By a standard result (e.g. Lemma 3.9 of [29]), T is a linear combination of these distributions. □

Remark 6.7. Since d_1 corresponds to the eigenspace of constants, the distributions $d_1^j \Delta^l$ are independent of j and are simply powers of the Laplacian applied to the Dirac mass at x. It should also be noted that for each j the distribution d_2^j corresponds to the largest eigenvalue less than 1, so gives the normal derivative at x when approaching through the cells $F_{[w_j]i}$, $i \to \infty$. The matching conditions for the Laplacian say that at a junction point these distributions are not linearly independent but instead satisfy $\sum_j d_2^j u = 0$. In consequence we do not need all of the basis elements listed in Theorem 6.6 at such a point.

The linear combination in Theorem 6.6 may include distributions of the form $\Delta^l d_{s,v}^j$ having $|\gamma_s| = (r_j \mu_j)^{k-l}$, and it is a-priori possible for these to be of order $k+1$. If this were to occur then we would have a non-trivial linear combination of these $(k+1)$-order distributions such that the linear combination is of order only k. We do not know of an example in which this occurs, but cannot eliminate it as a possibility because our arguments rely on scaling information.

Example 6.8. The canonical example of a p.c.f. self-similar fractal of the type we are describing is the Sierpinski Gasket **SG** with its usual symmetric harmonic structure (see [37] for details of all results described below). In this case $r = 3/5$ and $\mu = 1/3$, so the Laplacian scales by $r\mu = 1/5$. There are three harmonic extension matrices A_i, $i = 1, 2, 3$ corresponding to the boundary points, which are the vertices of the gasket. Each A_i has eigenvalues $\gamma_1 = 1$, $\gamma_2 = 3/5$ and $\gamma_3 = 1/5$ with one-dimensional eigenspaces. We consider A_1. In the basis where the first coordinate is the value at the boundary point q_1 and the other two are the values at the other boundary points q_2 and q_3, the eigenspace corresponding to the eigenvalue 1 is spanned by $(1, 1, 1)$, the one corresponding to $3/5$ is spanned by $(0, 1, 1)$, and that corresponding to $1/5$ is spanned by $(0, 1, -1)$. Hence the images of the projections P_s are the subspaces spanned by $(1, 1, 1)$, $(2, -1, -1)$ and $(0, 1, -1)$ for $s = 1, 2, 3$ respectively.

Now $D^0 u = 0$, so at the point q fixed by F_1 we have

$$\begin{aligned} d_1 u &= \lim_{i\to\infty} \gamma_1^{-i} P_1 H_i u \\ &= \lim_{i\to\infty} 1^{-i} \frac{1}{3} \langle (1,1,1), H_i u\rangle (1,1,1) \\ &= \lim_{i\to\infty} \frac{1}{3}\big(u(q_1) + u(F_1^i q_2) + u(F_1^i q_3)\big)(1,1,1) \\ &= u(q_1)(1,1,1) \end{aligned}$$

and the corresponding harmonic function is the constant $u(q_1)$. Next we have

$$\begin{aligned} d_2 u &= \lim_{i\to\infty} \gamma_2^{-i} P_2 H_i u \\ &= \lim_{i\to\infty} \Big(\frac{3}{5}\Big)^{-i} \frac{1}{5} \langle (2,-1,-1) H_i u\rangle (2,-1,-1) \\ &= \frac{1}{5} \lim_{i\to\infty} \Big(\frac{5}{3}\Big)^{i} \big(2u(q_1) - u(F_1^i q_2) - u(F_1^i q_3)\big)(2,-1,-1) \\ &= \frac{1}{5} \partial_n u(q_1)(2,-1,-1) \end{aligned}$$

where $\partial_n u(q_1)$ is the usual normal derivative at q_1. The harmonic function with this value of d_2 (and all other $d_s = 0$) is $-(\partial_n u(q_1)/2)(0,1,1)$. This exhausts the projections with $|\gamma_s| > r\mu$. We define $D^1 u = h$ where h is the harmonic function with the boundary values determined by $d_1 u$ and $d_2 u$, namely $u(q_1)(1,1,1) - (\partial_n u(q_1)/2)(0,1,1)$, which is of course the unique harmonic function matching u and $\partial_n u$ at q_1. All computations so far are valid for $u \in \mathrm{dom}(\Delta)$ (and in fact under slightly weaker conditions). If we knew that $u \in \mathrm{dom}(\Delta^2)$ we could take $D^1\Delta u$, which would be the harmonic function matching Δu and $\partial_n \Delta u$ at q_1, while $GD^1\Delta u$ would be the function with zero boundary values and Laplacian equal to this harmonic function. Now this function is symmetric under the (Euclidean) reflection fixing q_1 and exchanging q_2 with q_3, and the image of the projection onto P_3 is antisymmetric under this reflection. Hence $P_3 H_i G D^1 \Delta u = 0$ and therefore

$$\begin{aligned} d_3 u &= \lim_{i\to\infty} \gamma_3^{-i} P_3 H_i (u - GD^1\Delta u) \\ &= \lim_{i\to\infty} \Big(\frac{1}{5}\Big)^{-i} \frac{1}{2} \langle (0,1,-1) H_i u\rangle (0,1,-1) \\ &= \frac{1}{2} \lim_{i\to\infty} 5^i \big(u(F_1^i q_2) - u(F_1^i q_3)\big)(2,-1,-1) \\ &= \frac{1}{2} \partial_T u(q_1)(0,1,-1) \end{aligned}$$

where $\partial_T u(q_1)$ is the usual tangential derivative of u at q_1. It is worth noting that the latter exists as soon as Δu is Hölder continuous, which is certainly the case if $\Delta u \in \mathrm{dom}(E)$.

There are no other eigenvalues of A_1, so we have found all of the derivatives from the above definition. They are simply the function value, normal derivative and tangential derivative at q_1. If u has sufficient smoothness then $D^k u$ will be the harmonic function matching $u(q_1)$, $\partial_n u(q_1)$ and $\partial_T u(q_1)$, plus the Dirichlet Greens operator applied to $D^{k-1}\Delta u$. Thus it approximates u in the sense that $\Delta^j u$ is approximated up to value and normal derivative for $j \le k-1$ and also up to tangential derivative for $j \le k-2$. If we instead consider $\bar{D}^k$ then we also have tangential derivative matching for $j = k-1$.

Now consider the structure of distributions at a general point $q \in V_* \setminus V_0$, where two cells meet. Without loss of generality we suppose the cells are are indexed by $j = 0, 1$. The two corresponding normal derivatives $\partial_N^j u(q)$, $j = 0, 1$ satisfy the single linear relation that they sum to zero, and by Corollary 6.5 they are of order 1. The two tangential derivatives $\partial_T^j u(q)$, $j = 0, 1$ are independent, and it is known that they cannot be controlled by $\|u\|_\infty + \|\Delta u\|_\infty$ (see [37], page 60). They are therefore of order 2. It is also possible to see in this example that any non-trivial linear combination of the $\partial_T^j u(q)$ also has order 2. Writing δ_q for the Dirac mass at q, we conclude from Theorem 6.6 that any distribution T of order k at a cell boundary point of **SG** can be written as a linear combination of the form

$$T = \sum_{l=0}^{k} a_l \Delta^l \delta_q + \sum_{l=0}^{k-1} b_l \Delta^l \partial_N^0 \delta_q + \sum_{l=0}^{k-2} \sum_{j=0,1} c_{l,j} \Delta^l \partial_T^j \delta_q. \tag{6.17}$$

In this setting we can say somewhat more, because Theorem 2 of [20] shows that given three arbitrary sequences of numbers, it is possible to construct a smooth function u on **SG** for which these sequences occur as $\Delta^l \delta_q u$, $\Delta^l \partial_N^0 \delta_q u$ and $\Delta^l \partial_T^j \delta_q u$ respectively. This shows that there are no linear relations between these distributions, meaning that they are a minimal set for the representation in Theorem 6.6. (Note: In that paper these three sequences are referred to as a jet, which is a different definition of jet than used here.)

This example also illustrates the issue described in the proof of Corollary 6.5, namely that there can be a d_s with $|\gamma_s| = (r\mu)^{k-1}$ and yet d_s is order k. In this case we have $\partial_T = d_3$ with $\gamma_3 = 1/5 = r\mu$, so $k = 2$, and d_3 is of order 2.

Periodic and Eventually Periodic Points. Periodic points are those $x = F_w(X)$ for which w is a periodic word, meaning that w is composed of an infinite repetition of a fixed finite word v. Eventually periodic points are those for which the word w is periodic after some finite number of letters. For these points there is a theory similar to that used for cell boundary points; we do not have to consider derivatives corresponding to multiple cells, but instead of looking at the eigenstructure of a matrix A_i we must examine that of A_v, which is a finite composition of the A_i matrices. If γ_s is an eigenvalue of A_v with eigenspace E_s, then we can define the derivative d_s as we did for cell boundary points. It is easy to see that the analogues of Lemma 6.4, Corollary 6.5 and Theorem 6.6 all hold, simply by changing the notation to refer to the infinitely repeated matrix being A_v, the eigenvalues γ_s being those of A_v, and the Laplacian scaling factor to be $r_v\mu_v$ instead of $r_j\mu_j$.

Generic Points. We now consider a point $x = F_w(X)$ that is not in a cell boundary, so $w = w_1 w_2 \dots$ is an infinite word. The behavior of harmonic functions on the cell $F_{[w]_n}(X)$ can be understood by considering the product $A_{[w]_n} = \prod_{j=1}^{n} A_{w_j}$. We need to understand their scaling properties, for which we use the following approach from [39]. Define for each unit vector α the corresponding Lyupunov exponent

$$\log\gamma(\alpha) = \lim \frac{1}{i} \log \|A_{[w]_i}\alpha\| \tag{6.18}$$

if the limit exists. In this definition we may take $\|\cdot\|$ to be any norm on the $\#V_0$-dimensional space containing α; all such norms are equivalent, so γ is unaffected by this choice. We will use the norm corresponding to an inner product and make use of notions of orthogonality in the inner product space.

Let us suppose that these limits exist at x. Observe that if $\gamma(\alpha) > \gamma(\alpha')$ then for a linear combination $\gamma(a\alpha + b\alpha') = \gamma(\alpha)$ if $a \neq 0$. It follows that there are at most $\#V_0$ distinct values $\gamma_1 > \gamma_2 \cdots$ that occur. Corresponding to these is a direct sum decomposition $E_1 \oplus E_2 \oplus \cdots$ with the property that writing $\alpha = \alpha_1 + \alpha_2 + \cdots$ we have $\gamma(\alpha) = \gamma_s$ if and only if $\alpha_1 = \cdots = \alpha_{s-1} = 0$ and $\alpha_s \neq 0$. (Note that the spaces E_j can be more than one-dimensional, as occurs for a power of a fixed matrix with an eigenspace of dimension greater than one.) Since the constant functions are harmonic we actually know that $\gamma_1 = 1$ and E_1 is spanned by $(1, 1, \dots, 1)$. We let P_s be the orthogonal projection onto E_s.

The subspaces E_s provide the natural decomposition of harmonic functions into their scaling components at x. However we cannot expect to directly mimic Definition 6.3 because the estimate (6.18) does not imply the existence of a renormalized limit of the form

$$\lim_{i\to\infty} \gamma_s^{-i} P_s H_i(u - GD^{k-1}\Delta u) \tag{6.19}$$

Indeed it is easy to see that (6.18) does not even imply that $A_{[w]_i}\alpha$ is $O(\gamma(\alpha))^i$.

A natural way to proceed was introduced in [16, 39] and further treated in [27]. Let u be the function we are considering, and $H_i u$ be the harmonic function on $F_{[w]_i}(X)$ with boundary values equal to u on $\partial F_{[w]_i}(X)$ as usual. If we assume that the harmonic scaling matrices A_j are all invertible we can unravel the scaling structure for harmonic functions at x by applying the inverse of $A_{[w]_i}$ to $H_i u$. For later use we record an elementary result about the scaling of the adjoint of $A_{[w]_i}^{-1}$.

Lemma 6.9. *If* $\alpha \in E_s$ *then* $\lim \frac{1}{i} \log\left\|(A_{[w]_i}^{-1})^*\alpha\right\| = -\log\gamma_s$.

Proof. Writing $\langle\cdot,\cdot\rangle$ for the usual inner product,

$$\begin{aligned}\left\|(A_{[w]_i}^{-1})^*\alpha\right\| &= \sup_{\alpha''}\frac{\left|\langle\alpha'',(A_{[w]_i}^{-1})^*\alpha\rangle\right|}{\|\alpha''\|}\\ &= \sup_{\alpha'}\frac{\left|\langle A_{[w]_i}\alpha',(A_{[w]_i}^{-1})^*\alpha\rangle\right|}{\|A_{[w]_i}\alpha'\|}\\ &= \sup_{\alpha'}\frac{\left|\langle\alpha',\alpha\rangle\right|}{\|A_{[w]_i}\alpha'\|}.\end{aligned}$$

Since the logarithm is monotone, this implies

$$\frac{1}{i}\log\left\|(A_{[w]_i}^{-1})^*\alpha\right\| = \sup_{\alpha'}\Bigl(\frac{1}{i}\log\frac{\left|\langle\alpha',\alpha\rangle\right|}{\|\alpha'\|}-\frac{1}{i}\log\frac{\|A_{[w]_i}\alpha'\|}{\|\alpha'\|}\Bigr).$$

This sequence of suprema is certainly bounded below by the sequence with $\alpha'=\alpha$, which has limit $-\log\gamma(\alpha)=-\log\gamma_s$. Hence $\liminf_i\frac{1}{i}\log\left\|(A_{[w]_i}^{-1})^*\alpha\right\|\geq-\log\gamma_s$. Now each of the suprema is achieved at some point α'_i on the unit sphere, and if we take a subsequence on which $\frac{1}{i}\log\left\|(A_{[w]_i}^{-1})^*\alpha\right\|$ converges to its lim sup then the α'_i converge on a subsubsequence to some α'_∞ by compactness. If it were the case that the lim sup were greater than $-\log\gamma_s$ it would easily follow that α'_∞ has zero projection on E_t for $t\leq s$. However it is easy to see that in a neighborhood of such an α'_∞ that $\frac{|\langle\alpha',\alpha\rangle|}{\|A_{[w]_i}\alpha'\|}$ is bounded above by a constant multiple of $\|A_{[w]_i}\alpha\|^{-1}$, which yields the result. □

In order to account for the scaling behavior of the Laplacian, we set

$$\log\beta_w=\lim_{i\to\infty}\frac{1}{i}\log r_{[w]_i}\mu_{[w]_i} \tag{6.20}$$

provided the limit exists.

Definition 6.10. Assume that $x=F_w(X)$ is a point at which the limits in (6.18) and (6.20) exist, and that all A_j are invertible. Inductively define derivatives d_s and differentials D^k, $k\geq1$, at the point x by setting $D^0u=0$, and for each s such that $\beta_w^k<\gamma_s\leq\beta_w^{k-1}$

$$d_su=\lim_{i\to\infty}P_sA_{[w]_i}^{-1}H_i(u-GD^{k-1}\Delta u) \tag{6.21}$$

if these limits exist. Note that d_s always exists for harmonic functions because the sequence is constant in this case. Provided the necessary d_su exist we then let

$$D^ku=h+GD^{k-1}\Delta u \tag{6.22}$$

where h is the unique harmonic function on X with $d_sh=d_su$ for those s with $\gamma_s>\beta_w^k$ and $d_sh=0$ for all other s. We will also make use of $\bar{D}^ku$, where $\bar{D}^0=u(q)$ and

$$\bar{D}^ku=\bar{h}+G\bar{D}^{k-1}\Delta u$$

with $\bar{h}$ harmonic on X with $d_s\bar{h}=d_su$ for those s with $\gamma_s\geq\beta_w^k$ and $d_s\bar{h}=0$ for all other s.

This is essentially a generalization of Definition 6.3, because if $x=F_w(X)$ is a cell boundary point then w ends with infinite repetition of a single letter j, the Lyapunov exponents are the eigenvalues γ_s of A_j, and the action of $A_{[w]_i}^{-1}$ on the corresponding eigenspace is just multiplication by γ_s^{-i}. (The definitions do not precisely coincide because our orthogonalization procedures on the subspaces are different. To see this, consider powers of a fixed 2×2 matrix with non-orthogonal eigenvectors. Then the projections in Definition 6.3 are onto the eigenspace of the larger eigenvalue and its orthogonal complement, while those in Definition 6.10 are onto the eigenspaces of the smaller eigenvalue and its orthogonal complement. Though they are clearly equivalent objects it is perhaps arguable that the former is correctly thought of as a derivative and the latter as a tangent, see Section 4.2 of [39])

The following result may be seen as a generalization of Theorem 1 of [39], see also Theorems 5 and 6 of [27]. It is proved by essentially the same method as Lemma 6.4. At several points in the proof we use the observation that for a positive sequence a_i satisfying $\lim i^{-1}\log a_i = \log a$ and a value $\epsilon > 0$ there is a constant $C(\epsilon)$ so $C(\epsilon)^{-1}e^{-\epsilon i}a^i \leq a_i \leq C(\epsilon)e^{\epsilon i}a^i$.

Lemma 6.11. *Assume that all A_j are invertible, and that $x = F_w(X)$ is a point at which the limits in* (6.18) *and* (6.20) *exist. For $u \in \operatorname{dom}(\Delta^k)$ and each s such that $\gamma_s > \beta_w^k$, the derivative d_s exists, and*

$$|d_s u| \leq C(k)\sum_{l=0}^{k}\|\Delta^l u\|_\infty. \tag{6.23}$$

For all sufficiently small $\epsilon > 0$, the differential satisfies

$$\left\|u - D^k u\right\|_{L^\infty(F_{[w]_i}(X))} \leq C(k,\epsilon)\beta_w^{ik}e^{\epsilon i}\|\Delta^k u\|_\infty. \tag{6.24}$$

If in addition we assume that $\Delta^k u \in \operatorname{dom}(E)$ then

$$\left\|u - \bar{D}^k u\right\|_{L^\infty(F_{[w]_i}(X))} \leq C(k,\epsilon)r_{[w]_i}^{1/2}\beta_w^{ik}e^{\epsilon i}E^{1/2}(\Delta^k u). \tag{6.25}$$

Proof. The proof is inductive. When $k = 0$ there are no s with $\gamma_s > 1 = \beta_w^0$, so the first statement is vacuous and (6.24) is immediate. Suppose both hold up to $k-1$.

Write $u - GD^{k-1}\Delta u$ as $H_0 u + G(\Delta u - D^{k-1}\Delta u)$, so

$$d_s u = d_s H_0 u + \lim_i P_s A_{[w]_i}^{-1} H_i (G\Delta u - GD^{k-1}\Delta u) \tag{6.26}$$

provided the latter limit exists. Writing G_i for the Dirichlet Green's operator on the cell $F_{[w]_i}(X)$, we have on that cell

$$G(\Delta u - D^{k-1}\Delta u) = H_i G(\Delta u - D^{k-1}\Delta u) + G_i(\Delta u - D^{k-1}\Delta u)$$

from which

$$H_{i+1}G(\Delta u - D^{k-1}\Delta u) = A_{w_{i+1}} H_i G(\Delta u - D^{k-1}\Delta u) + H_{i+1}G_i(\Delta u - D^{k-1}\Delta u),$$

therefore

$$\begin{aligned} A_{[w]_{i+1}}^{-1} H_{i+1}(G\Delta u - GD^{k-1}\Delta u) - A_{[w]_i}^{-1} H_i (G\Delta u - GD^{k-1}\Delta u)& \\ = A_{[w]_i}^{-1} H_{i+1} G_i (\Delta u - D^{k-1}\Delta u),& \end{aligned} \tag{6.27}$$

and by substitution into (6.26),

$$d_s u = d_s H_0 u + \sum_0^\infty P_s A_{[w]_i}^{-1} H_{i+1} G_i (\Delta u - D^{k-1}\Delta u) \tag{6.28}$$

provided that the series converges.

Since G_i inverts the Laplacian with Dirichlet boundary conditions on $F_{[w]_i}(X)$, we have for any sufficiently small $\epsilon > 0$ the bound

$$\begin{aligned} \left|G_i(\Delta u - D^{k-1}\Delta u)\right| &\leq C r_{[w]_i}\mu_{[w]_i}\left\|\Delta u - D^{k-1}\Delta u\right\|_{L^\infty(F_{[w]_i(X)})} \\ &\leq C(k-1,\epsilon) r_{[w]_i}\mu_{[w]_i}\beta_w^{(k-1)i}e^{(\epsilon/4)i}\left\|\Delta^k u\right\|_\infty \\ &\leq C(k-1,\epsilon)\beta_w^{ki}e^{(\epsilon/2)i}\left\|\Delta^k u\right\|_\infty \end{aligned} \tag{6.29}$$

because of the inductive hypothesis (6.24) and the Laplacian scaling estimate (6.20). This is also applicable to $H_{i+1}G_i(\Delta u - D^{k-1}\Delta u)$ by the maximum principle. Using Lemma 6.9 to estimate the size of $\|(A^{-1}_{[w]_i})^* P_s\alpha\|$, it follows that for any sufficiently small $\epsilon > 0$, and any vector α,

$$\begin{aligned}\left|\langle P_s A^{-1}_{[w]_i} H_{i+1} G_i(\Delta u - D^{k-1}\Delta u), \alpha\rangle\right| &= \left|\langle H_{i+1} G_i(\Delta u - D^{k-1}\Delta u), (A^{-1}_{[w]_i})^* P_s\alpha\rangle\right| \\ &\leq C(k-1,\epsilon)\beta_w^{ki}\gamma_s^{-i} e^{(3\epsilon/4)i}\left\|\Delta u\right\|_\infty,\end{aligned}$$

This and the assumption $\gamma_s > \beta_w^k$ imply that if $\epsilon > 0$ was chosen small enough then the series in (6.28) converges, and is bounded by $C\|\Delta^k u\|_\infty$. The estimate (6.23) follows because $d_s H_0 u$ is bounded by $C\|u\|_\infty$.

Now $u - D^k u = u - h - GD^{k-1}\Delta u = H_0 u - h + G(\Delta u - D^{k-1}\Delta u)$, where h is the harmonic function with $d_s u = P_s h$ for all s satisfying $\gamma_s > \beta_w^k$ and $P_s h = 0$ otherwise. An expression for h can be obtained by summing (6.28) over these values of s. Comparing it to the expression

$$A^{-1}_{[w]_i} H_i(u - D^k u) = H_0 u - h + \sum_0^{i-1} A^{-1}_{[w]_l} H_{l+1} G_l(\Delta u - D^{k-1}\Delta u)$$

from (6.27), it is apparent that for those s with $\gamma_s > \beta_w^k$ we have

$$P_s A^{-1}_{[w]_i} H_i(u - D^k u) = -\sum_i^\infty P_s A^{-1}_{[w]_l} H_{l+1} G_l(\Delta u - D^{k-1}\Delta u)$$

which we note satisfies for all $\|\alpha\| \leq 1$ and sufficiently small $\epsilon > 0$

$$\begin{aligned}\left|\langle P_s A^{-1}_{[w]_i} H_i(u - D^k u), \alpha\rangle\right| &\leq \sum_i^\infty \left|\langle H_{l+1} G_l(\Delta u - D^{k-1}\Delta u), (A^{-1}_{[w]_l})^* P_s\alpha\rangle\right| \\ &\leq \sum_i^\infty C(k-1,\epsilon)\beta_w^{kl}\gamma_s^{-l} e^{(3\epsilon/4)l}\left\|\Delta^k u\right\|_\infty \\ &\leq C(k-1,\epsilon)\beta_w^{ki}\gamma_s^{-i} e^{(3\epsilon/4)i}\left\|\Delta^k u\right\|_\infty. \end{aligned} \tag{6.30}$$

For those s satisfying $\gamma_s \leq \beta_w^k$ we have instead

$$P_s A^{-1}_{[w]_i} H_i(u - D^k u) = \sum_0^i P_s A^{-1}_{[w]_l} H_{l+1} G_l(\Delta u - D^{k-1}\Delta u).$$

and for all vectors α with $\|\alpha\| \leq 1$,

$$\begin{aligned}\left|\langle P_s A^{-1}_{[w]_i} H_i(u - D^k u), \alpha\rangle\right| &\leq \sum_0^i C(k-1,\epsilon)\beta_w^{kl}\gamma_s^{-l} e^{(3\epsilon/4)l}\left\|\Delta^k u\right\|_\infty \\ &\leq C(k-1,\epsilon)\beta_w^{ki}\gamma_s^{-i} e^{(3\epsilon/4)i}\left\|\Delta^k u\right\|_\infty. \end{aligned} \tag{6.31}$$

Equations (6.30) and (6.31) give the same estimate for each $P_s A^{-1}_{[w]_i} H_i(u - D^k u)$. Mapping forward again by $A_{[w]_i}$ increases each term by a factor at most $C(\epsilon)\gamma_s^i e^{(\epsilon/4)i}$, so summing over all s we finally have

$$\left|H_i(u - D^k u)\right| \leq C\beta_w^{ki} e^{\epsilon i}\left\|\Delta^k u\right\|_\infty$$

for some constant $C = C(k,\epsilon)$. Now the restriction of $(u - D^k u)$ to $F_{[w]_i}(X)$ is

$$(u - D^k u)\Big|_{F[w]_i(X)} = H_i(u - D^k u) + G_i(\Delta(u - D^k u)) = H_i(u - D^k u) + G_i(\Delta u - D^{k-1}\Delta u))$$

the second term of which is bounded by $\beta_w^{ki} e^{\epsilon i} \|\Delta^k u\|_\infty$ from (6.29), and the first term of which we have just estimated in the same way. This establishes (6.24) and completes the induction.

The proof of (6.25) is the same, except that (6.25) is used in place of (6.24) throughout. The validity of the estimate for $k = 0$ is a consequence of the Hölder estimate (2.2). □

As previously, we fix orthonormal bases for the spaces E_s and see that the components of d_s are distributions.

Corollary 6.12. *Suppose that x satisfies the assumptions of Lemma 6.11 and $\beta_w^k < \gamma_s$. Any component $d_{s,v}$ of the derivative d_s is a distribution of order at most k supported at x. If also $\gamma_s < \beta_w^{k-1}$ then $d_{s,v}$ has order equal to k. If $d_{s,v}$ has order k then defining $\Delta^l d_{s,v}$ by $\Delta^l d_{s,v}\phi = d_{s,v}\Delta^l \phi$ yields a distribution supported at x and of order $k + l$.*

Proof. Linearity of $d_{s,v}$ is immediate from Definition 6.10, so it is a distribution of order at most k by (6.23). Again using Definition 6.10 it is apparent that $d_{s,v}\phi = 0$ if $\phi \in \mathcal{D}(\Omega)$ vanishes in a neighborhood of x, so $d_{s,v}$ is supported at x.

To see that $d_{s,v}$ has order at least k, consider the harmonic function h with boundary values equal to the unit vector in the v direction in E_s. Then $H_i h = A_{[w]_i} H_0 h$, so the sequence in (6.21) is constant equal to $H_0 h$, and $d_{s,v}h = 1$. Now for $\epsilon > 0$ so small that $\gamma_s e^{3\epsilon} \leq \beta_w^{k-1}$ we have

$$\|H_i h\|_\infty \leq C(\epsilon)\gamma_s^i e^{\epsilon i} \leq C(\epsilon)\beta_w^{(k-1)i} e^{-2\epsilon i} \leq C(\epsilon)\big(r_{[w]_i}\mu_{[w]_i}\big)^{k-1} e^{-\epsilon i} = o\big(r_{[w]_i}\mu_{[w]_i}\big)^{k-1}$$

and of course $\Delta^l h \equiv 0$ for all $l > 0$, so Theorem 6.2 applies with $m = k - 1$, and there is a test function ϕ such that $\phi = h$ in a neighborhood of x and $|\phi|_{k-1}$ is as small as we desire. Since $d_{s,v}h = d_{s,v}\phi$ by the support condition, $d_{s,v}$ cannot be of order $k - 1$ or less. The final statement of the lemma is obvious. □

Theorem 6.13. *Suppose that all of the matrices A_j are invertible, and that $x = F_w(X)$ is a point at which the limits in* (6.18) *and* (6.20) *exist. Then all distributions of order at most k at x are linear combinations of the distributions $\Delta^l d_{s,v}$, with $\gamma_s \geq \beta_w^{k-l}$.*

Proof. As in the proof of Theorem 6.6, it suffices to show that T vanishes whenever the distributions $\Delta^l d_{s,v}$, with $\gamma_s \geq \beta_w^{k-l}$ vanish.

Suppose $\phi \in \mathcal{D}(\Omega)$ satisfies $\Delta^l d_{s,v}\phi = 0$ for those $\gamma_s \geq \beta_w^{k-l}$. Then the differential $\bar{D}^k\phi$ (which exists by Lemma 6.11) must be zero, as must $\bar{D}^{k-l}\Delta^l\phi$ for each $0 \leq l \leq k$. From (6.25) we then see that for all sufficiently small $\epsilon > 0$,

$$\begin{aligned}\|H_i\Delta^l\phi\|_{L^\infty(F_{[w]_i}(X))} \leq \|\Delta^l\phi\|_{L^\infty(F_{[w]_i}(X))} &\leq C(k,\epsilon) r_{[w]_i}^{1/2}\beta_w^{i(k-l)} e^{\epsilon i} E^{1/2}(\Delta^k u)\\ &\leq C(k,\epsilon) r_{[w]_i}^{1/2}\big(r_{[w]_i}\mu_{[w]_i}\big)^{k-l} e^{2\epsilon i} E^{1/2}(\Delta^k u)\\ &= o\big(r_{[w]_i}\mu_{[w]_i}\big)^{k-l}.\end{aligned}$$

Applying Theorem 6.2 we find that for any $\delta > 0$ there is ψ equal to ϕ in a neighborhood of x and such that $|\psi|_k < \delta$. In particular, since T is order k and supported at x, there is M independent of ϕ such that

$$|T\phi| = |T\psi| \leq M|\psi| \leq \delta$$

and thus $T\phi = 0$. □

In concluding this section it seems appropriate to say something about the set of points x satisfying the conditions in Definition 6.10. The set at which the limit β_w exists has full

μ-measure by the law of large numbers, and in fact

$$\beta_w = \sum_j \mu_j \log r_j \mu_j$$

at μ-a.e. point. The set on which the Lyupunov exponents exist may be treated by the theory of random matrices introduced by Furstenberg and Kesten [8]. In particular, it is possible to make certain assumptions on the matrices A_i that guarantee that this set is also of full μ-measure. This topic is discussed quite thoroughly in the paper [27] of Pelander and Teplyaev, so we will not cover it here. One consequence of their work, however, is that there are conditions that imply the spaces E_s are independent of the choice of point x. For example, if the semigroup generated by the A_i is strongly irreducible and contracting then there is a single vector α_1 such that at μ-almost every x, the space E_1 is spanned by α_1 and has Lyapunov exponent γ_1. If the same strong irreducibility and contraction holds after taking the quotient to remove E_1 then E_2 is also one-dimensional and independent of x on a full measure set. For a fractal where the irreducibility and contraction properties are true for the semigroup generated by the A_i on each of the subspaces found by removing $E_1, E_2, .., E_{s-1}$ in turn, we could conclude that all of the distributions of the form d_s are independent of x on a set of full μ-measure. Hence in this situation any distribution of order m with point support in a fixed set of full μ-measure would be a finite linear combination of distributions $\Delta^l d_s$ for suitable values of l, where the d_s are independent of x. This generic behavior is very different from that seen at cell boundary points and eventually periodic points, where the structure of point-supported distributions can vary substantially from point to point.

7. Distributions on products

In this section we give a theory of distributions on finite products of post-critically finite self-similar fractals, using the analytic theory for such products developed in [35]. This gives genuinely new examples, because products of p.c.f. self-similar sets are not usually themselves p.c.f. Since there is no essential difference between a product $X = X' \times X''$ with two factors and a general finite product, we state our results only for the two factor case.

Following the notational conventions of [35], points are $x = (x', x'')$, functions on X are called u or f, on X' they are u' or f', while on X'' they are u'' or f''. The energies on X' and X'' are E' and E'' and the Laplacians are Δ' and Δ''. They come from a regular harmonic structure as in Section 2 and have energy and measure scaling factors r', μ', r'' and μ''. The corresponding Laplacians Δ' and Δ'' are defined componentwise, so $u \in \text{dom}(\Delta')$ with $\Delta' u = f$ if u and f are continuous on X and have the property that for each fixed $x'' \in X''$ we have $\Delta' u(\cdot, x'') = f(\cdot, x'')$. A similar definition is used for Δ''. By Lemma 11.2 of [35], Δ' and Δ'' commute on $\text{dom}(\Delta') \cap \text{dom}(\Delta'')$.

Definition 7.1. A function u on X is smooth if for all $j, k \in \mathbb{N}$, $(\Delta')^j(\Delta'')^k u$ is a continuous function on X. The definition extends to a finite union of cells in the obvious manner, and u is smooth on a domain in X if it is smooth on every finite union of cells in the domain.

As before, we only wish to treat the case of an open domain.

Assumption 7.2. *In what follows, we assume that Ω is a domain with no boundary.*

We define the test functions $\mathcal{D}(\Omega)$ on a domain Ω to be the smooth functions of compact support with the usual topology and the corresponding seminorms

$$|\phi|_m = \sup\Big\{\big|(\Delta')^j(\Delta'')^k\phi(x)\big| : x \in \Omega,\ j + k \leq m''\Big\}.$$

The distributions $\mathcal{D}'(\Omega)$ form the dual space with weak-star topology. A distribution T has order m if on any compact K there is $M = M(K)$ so that $|T\phi| \leq M|\phi|_m$ for all test functions supported on K

The goal of this section is to provide conditions under which analogues of our main results for distributions on p.c.f. fractals are also valid on products of these fractals. In order to avoid duplicating a great deal of work, we only give details of the proofs where they differ significantly from those for the case of a single p.c.f. fractal. In particular it is fairly easy to verify that all of the results of Section 3 (except Corollary 3.6), Section 4, and Section 5 prior to Theorem 5.6, depend only on the partitioning property of Theorem 2.8 and the estimate (2.5) (either directly or through Lemma 3.3) as well as the fact that for compact Ω there is an orthonormal basis of L^2 consisting of eigenfunctions. The latter is obviously true for the product $X' \times X''$ because it is true for the factors, so the original proofs transfer to the product setting once we know the partitioning property and the corresponding estimate for product spaces. These are proved in Theorem 7.8 and (7.6) below, so all of the aforementioned results are also true for products of p.c.f. fractals with regular harmonic structure, and connected fractafolds with restricted cellular structure based on such products.

Only small changes are needed to obtain analogues of the remaining results from Section 5. The proof of Theorem 5.6 required that on any cell there was a positive smooth function equal to 1 on the cell and vanishing outside a specified neighborhood: such a function may be obtained on the product space as a product of functions of this type on the factors, so the theorem is true for products in which each factor has the estimate (2.4) for the heat kernel. The other results are used to prove the structure theorem (Theorem 5.10). Of these, Lemma 5.8 remains true with the same proof if it is modified to say that Δ' maps $\mathcal{D}(K)$ to itself with image orthogonal to those ϕ having $\Delta'\phi = 0$ and there is $\tilde{G}'_K$ such that $-\Delta'\tilde{G}'_K(\Delta'\phi) = \Delta'\phi$; there is a corresponding result for Δ''. A version of Theorem 5.9 is then true with $\tilde{G}''_K\tilde{G}'_K$ replacing $\tilde{G}_K$ throughout. The original proof shows that for $m \geq 1$, $\tilde{G}''_K\tilde{G}'_K$ takes a distribution of order m to one of order at most $m-1$. To show that $\tilde{G}''_K\tilde{G}'_K$ takes a distribution of order zero to a continuous function it suffices to approximate the corresponding measure ν by a sequence of linear combinations of product measures. Applying $\tilde{G}''_K\tilde{G}'_K$ to a product measure gives a continuous function by the original proof of Theorem 5.9, so applying it to the sequence gives a uniformly convergent sequence of continuous functions whose limit represents the distribution $\tilde{G}''_K\tilde{G}'_K\nu$. The proof of Theorem 5.10 needs no further changes.

At this point we have essentially all of the results of Sections 3, 4, and 5 in the product setting (the only exception is Corollary 3.6). In addition there are some things that can be said about distributions with point support that generalize the results of Section 6. We will return to these after giving the details of the partitioning argument, because some aspects of the procedure for cutting off a smooth function will be important for the proofs.

Partitioning on products. We prove analogues of the partitioning property in Theorem 2.8 and the estimate (2.5) in the product setting. As in the single variable case, the proof relies on a cell-by-cell construction of a smooth function, for which the following matching condition is essential. Note that a cell in X is a product of cells from X' and X'', so has the form $K = F'_{w'}(X') \times F''_{w''}(X'')$, where w' and w'' are finite words. Its boundary consists of faces $\{q'_i\} \times F''_{w''}(X'')$ and $F'_{w'}(X') \times \{q''_j\}$ for $q'_i \in V'_0$ and $q''_j \in V''_0$.

Lemma 7.3. *Suppose the cells $K_1, \dots, K_k$ all contain the face $L = \{q'\} \times F''_{w''}(X'')$, and that the union $\cup_1^k K_l$ contains a neighborhood of every point in L except those of the form*

$(q', F''_{w''}q'')$ with $q'' \in V_0$. If u_j is smooth on K_j for each j, then the piecewise defined function $u = u_j$ on K_j is smooth on $\cup_1^k K_l$ if and only if for each $x'' \in X''$, both
(a) The functions $(\Delta')^l(\Delta'')^m u_j(q', x'')$ are independent of j for each l and m, and
(b) For each x'', $\sum_j (\partial_n')_j (\Delta')^l(\Delta'')^m u_j(q', x'') = 0$, where $(\partial_n')_j$ indicates the normal derivative in the x' variable from within K_j.

Proof. For fixed x'', (b) is the necessary and sufficient matching condition in the first variable for $(\Delta')^l(\Delta'')^m u(\cdot, x'')$ to exist (as a function rather than a measure with atom at q'). Condition (a) is then equivalent to continuity of $(\Delta')^l(\Delta'')^m u$. □

Our construction uses an analogue of the Borel theorem from [28]. That result yields the existence of a smooth function with a prescribed jet at a cell boundary point of a p.c.f. fractal, whereas we need existence of a smooth function with prescribed smooth jet on the face of a cell in the product X.

Theorem 7.4. *Fix a face $\{q'\} \times X''$ and a neighborhood $U \subset X'$ of q'. Given two sequences $\{\rho_l(x'')\}_{l=0}^\infty$ and $\{\sigma_l(x'')\}_{l=0}^\infty$ of functions that are smooth in x'', there is a smooth function u with support in $U \times X''$ such that for each $x'' \in X''$, $(\Delta')^k(\Delta'')^m u(q', x'') = (\Delta'')^m \rho_k(x'')$ and $\partial_n'(\Delta')^k(\Delta'')^m u(q', x'') = (\Delta'')^m \sigma_k(x'')$.*

Proof. The proof is almost the same as that for Theorem 4.3 of [28]. Specifically we form the series

$$u(x', x'') = \sum_l \rho_l(x'') g_{l,m_l}(x') + \sigma_l(x'') f_{l,n_l}(x') \tag{7.1}$$

where the functions g_{l,m_l} and f_{l,n_l} are as defined in that proof, so they satisfy

$$(\Delta')^k g_{l,m_l}(q) = \delta_{kl} \qquad \partial_n'(\Delta')^k g_{l,m_l}(q) = 0$$
$$(\Delta')^k f_{l,n_l}(q) = 0 \qquad \partial_n'(\Delta')^k f_{l,n_l}(q) = \delta_{kl}$$

and have supports in cells of scale m_l and n_l respectively. Convergence of the series (7.1) is achieved by making an appropriate choice of m_l and n_l. In particular, it follows from the cited proof that if $|\rho_l(x'')| \leq R_l$ and $|\sigma_l(x'')| \leq S_l$ for all x'', then one can choose m_l and n_l depending only on R_l and S_l such that for each x'' the series converges to a function that is smooth in x', supported in $U \times X''$, and has $(\Delta')^k u(q, x'') = \rho_k(x'')$ and $\partial_n'(\Delta')^k u(q, x'') = \sigma_k(x'')$.

Now we require convergence not only of the series for $u(x', x'')$, but also that for $(\Delta'')^m u(x', x'')$ for each m, so we must diagonalize. Set

$$R_l = \max_{0 \leq l'' \leq l} \max_{x'' \in X''} \left|(\Delta'')^{l''} \rho_l(x'')\right|$$
$$S_l = \max_{0 \leq l'' \leq l} \max_{x'' \in X''} \left|(\Delta'')^{l''} \sigma_l(x'')\right|$$

which are finite by the assumed smoothness and the compactness of X'', and let m_l and n_l be chosen as described above. For fixed x'', all terms after the m-th in the partial sum

$$(\Delta'')^m \sum_l^L \rho_l(x'') g_{l,m_l}(x') + \sigma_l(x'') f_{l,n_l}(x')$$

have coefficients bounded by R_l and S_l, so the above reasoning implies that the partial sums converge to a function that is smooth in x', and has $(\Delta')^k(\Delta'')^m u(q', x'') = (\Delta'')^m \rho_k(x'')$ and $\partial_n'(\Delta')^k(\Delta'')^m u(q', x'') = (\Delta'')^m \sigma_k(x'')$ for all m.

Finally, it will be useful later to have estimated the contribution of each term to the L^∞ norm of $(\Delta')^k(\Delta'')^m u$. It is convenient to write $w'(m_l)$ and $w'(n_l)$ for the words such that

$F_{w'(m_l)}(X')$ is the support of g_{l,n_l} and $F'_{w'(n_l)}(X)$ is the support of f_{l,n_l}. Note that scaling then implies (see equations 4.4 and 4.5 of [28]) that

$$\left\|(\Delta')^k(\Delta'')^m\rho_l(x'')g_{l,m_l}(x')\right\| \leq c(k,l)\big(r'_{w'(m_l)}\mu'_{w'(m_l)}\big)^{l-k}\left\|(\Delta'')^m\rho_l(x'')\right\|_\infty$$
$$\left\|(\Delta')^k(\Delta'')^m\sigma_l(x'')f_{l,n_l}(x')\right\| \leq c(k,l)\big(r'_{w'(n_l)}\mu'_{w'(n_l)}\big)^{l-k}r'_{w'(n_l)}\left\|(\Delta'')^m\sigma_l(x'')\right\|_\infty$$

and in the construction in [28] it is noted that the contributions of terms with $l > k$ may be made smaller than any prescribed $\epsilon > 0$, so taking ϵ to be a small multiple of $\|\rho_0\|_\infty$ we obtain

$$\begin{aligned}\left\|(\Delta')^k(\Delta'')^m u\right\|_\infty \leq &\sum_{l=0}^{k} c(k,l)\big(r'_{w'(m_l)}\mu'_{w'(m_l)}\big)^{l-k}\left\|(\Delta'')^m\rho_l(x'')\right\|_\infty \\ &+\sum_{l=0}^{k-1} c(k,l)\big(r'_{w'(n_l)}\mu'_{w'(n_l)}\big)^{l-k}r'_{w'(n_l)}\left\|(\Delta'')^m\sigma_l(x'')\right\|_\infty. \qquad (7.2)\end{aligned}$$

□

Remark 7.5. This result may be localized to any cell in X simply by rescaling the desired jet for the cell to obtain a corresponding jet on X, applying the theorem, and then composing the resulting function with the inverse of the map to the cell. It may also be applied to a face in a finite union of cells, so that the face is of the form $\{q'\}\times(\cup_{j=1}^J K''_j)$ with each K''_j a cell in X'', because $\cup_{j=1}^J K''_j$ is compact.

In order to make use of the preceding result we require a small lemma.

Lemma 7.6. *If u is smooth on X and $q' \in V'_0$ then $\partial'_n u(q',x'')$ is smooth with respect to x'' and $(\Delta'')^l\partial'_n u(q',x'') = \partial'_n(\Delta'')^l u(q',x'')$. There is a bound*

$$\left\|\partial'_n(\Delta'')^l u(q',x'')\right\|_\infty \leq C\Big(\left\|(\Delta'')^l u\right\|_\infty + \left\|\Delta'(\Delta'')^l u\right\|_\infty\Big) \qquad (7.3)$$

Proof. For each x'' and each scale m, let $h_m(x',x'')$ be the function that is piecewise harmonic at scale m in the x' variable and coincides with u on $V'_m\times\{x''\}$. Then $h_m(x',x'')$ is smooth in x'', because its values are obtained as uniform limits of linear combinations of the values from $V'_m\times\{x''\}$. Moreover, the normal derivative $\partial'_n h_m(q',x'')$ is a linear combination (with coefficients depending on m) of the differences $(h_m(p'_1,x'') - h_m(p'_2,x''))$, where p'_1 and p'_2 are neighbors of q' at scale m. Thus $\partial'_n h_m(q',x'')$ is smooth in x'' and $(\Delta'')^l\partial'_n h_m(q',x'') = \partial'_n(\Delta'')^l h_m(q',x'')$.

For each fixed x'', we may express $(\Delta'')^l u(x',x'')$ on a cell K'_m of scale m containing q' as the sum of $(\Delta'')^l h_m$ and an integral involving the Dirichlet Green kernel G'_m for Δ' on K'_m. Taking the normal derivative we obtain

$$\partial'_n(\Delta'')^l u(q',x'') = \partial'_n(\Delta'')^l h_m(q',x'') + \int(\Delta'(\Delta'')^l u(y',x''))\partial'_n G'_m(q',y')\,d\mu'(y'). \qquad (7.4)$$

However an easy scaling argument shows that $\partial'_n G'_m(q',y')$ is bounded independent of m and y', so the integral term is bounded by a constant multiple of $\left\|\Delta'(\Delta'')^l u\right\|_\infty\mu'(K'_m)$, independent of m and x''. Since $\mu'(K'_m) \to 0$ as $m \to \infty$ we conclude that $(\Delta'')^l\partial'_n h_m(q',x'')$ converges to $\partial'_n(\Delta'')^l u(q',x'')$ uniformly in x'' for each l. Then (7.3) is obtained by using (7.4) with $m = 0$. □

We may use the preceding results to smoothly cut off a smooth function on a neighborhood of a cell.

Theorem 7.7. *Let u be smooth on a cell $K = F'_{w'}(X') \times F''_{w''}(X'')$, and $U \supset K$ be open. There is a function v such that $v = u$ on K, $v = 0$ on $X \setminus U$ and v is smooth on X. Moreover for each k,*

$$\left\|(\Delta')^k(\Delta'')^m v\right\|_\infty \leq C(k,U)\sum_{l=0}^{k}\sum_{n=0}^{m}\left\|(\Delta')^l(\Delta'')^n u\right\|_{L^\infty(K)}. \tag{7.5}$$

Proof. Let $K' = F'_{w'}(X')$ and $K'' = F''_{w''}(X'')$. Fix a face of K having the form $\{q'\} \times K''$ and let $\rho_k(x'') = (\Delta')^k u(q', x'')$ and $\sigma_k(x'') = \partial'_n(\Delta')^k u(q', x'')$. The functions ρ_k are smooth in x'' by the definition of smoothness of u, and the functions σ_k are smooth in x'' by Lemma 7.6. Now take a finite number of small cells K'_j in X' with the following properties: the intersection $K' \cap K'_j = \{q'\}$ for all j, the intersection $K'_j \cap K'_{\tilde{j}} = \{q'\}$ for all $j \neq \tilde{j}$, the union $K' \cup (\cup_j K'_j)$ contains a neighborhood of q' in X', and $(K' \cup (\cup_j K'_j)) \times K'' \subset U$. Let the number of K'_j be J, and apply Theorem 7.4 to each K'_j to obtain a smooth function u_j that has jets $\rho_k(x'')$ and $(-1/J)\sigma_k(x'')$ at q' and is supported in a neighborhood of q' that is strictly contained in K'_j. By construction, the matching conditions of Lemma 7.3 apply to the functions u on K and u_j on $K'_j \times K''$, so the piecewise defined function is smooth on the union of these cells.

Repeat the previous construction for each of the finite number of faces having the form $\{q'_i\} \times K''$. As these faces are disjoint we may choose the small cells in the construction so that those used for q'_i do not intersect those for q'_j for $j \neq i$. The result is a finite collection of cells $K'_j \times K'' \subset U$ and functions u_j such that the piecewise function u on K and u_j on $K'_j \times K''$ is smooth on the union of the cells, and vanishes identically in a neighborhood of any boundary face of $(K' \cup (\cup K'_j)) \times K''$ that has the form $\{p'\} \times K''$. We call this function v'.

Having treated the vertical faces $\{q'_i\} \times K''$, we then treat the horizontal faces $(K' \cup (\cup K'_j)) \times \{q''\}$ of the new function v' in the same manner. All of the results we needed were valid on faces of finite unions of cells, so the same proof allows us to piecewise extend to a smooth function v on a larger finite union of cells, which we call L, but with the additional condition that v vanishes identically in a neighborhood of each horizontal face of L. Then $L \subset U$ and v vanishes in a neighborhood of all faces of the boundary of L, so Lemma 7.3 ensures that extending v to be identically zero outside L gives a smooth function on X. By construction, $v = u$ on K.

For the estimate (7.5) we note that

$$\left\|(\Delta'')^m \rho_k(x'')\right\|_\infty \leq \left\|(\Delta')^k(\Delta'')^m u\right\|_{L^\infty(K)}$$

by definition, while rescaling (7.3) to the cell K implies that

$$\left\|(\Delta'')^m \sigma_k(x'')\right\|_\infty \leq C\Big((r'_{w'})^{-1}\left\|(\Delta')^k(\Delta'')^m u\right\|_{L^\infty(K)} + \mu'_{w'}\left\|(\Delta')^{k+1}(\Delta'')^m u\right\|_{L^\infty(K)}\Big).$$

Substituting into (7.2) and using $r'_{w'(n_l)} \leq r'_{w'}$ and $\mu'_{w'} < 1$ we have

$$\begin{aligned}\left\|(\Delta')^k(\Delta'')^m\right\|_\infty &\leq \sum_{l=0}^{k} c(k,l)\big(r'_{w'(m_l)}\mu'_{w'(m_l)}\big)^{l-k}\left\|(\Delta')^l(\Delta'')^m u\right\|_{L^\infty(K)}\\ &\leq C(k,U)\sum_{l=0}^{k}\left\|(\Delta')^l(\Delta'')^m u\right\|_{L^\infty(K)}.\end{aligned}$$

This type of estimate deals with all of the vertical faces, and an analogous argument is valid for the horizontal faces, so (7.5) holds. □

Theorem 7.8. *If u is smooth on X and $\cup\Omega_j$ is an open cover of X then there are constants $C(k,m)$ and smooth functions u_j such that u_j is supported on Ω_j, $\sum_j u_j = u$, and*

$$\|(\Delta')^k(\Delta'')^m u_j\|_\infty \leq C(k,m)\sum_{l=0}^{k}\sum_{n=0}^{m}\|(\Delta')^l(\Delta'')^n u\|_\infty. \tag{7.6}$$

Proof. The open cover is finite, say $\{\Omega_j\}_1^J$ because X is compact. Moreover we may partition X into a finite number of cells K_l such that each K_l is contained in some Ω_j. We proceed by induction on l, with the base case being that we apply Theorem 7.7 to u on K_1 to obtain a smooth function v_1 with support in the open Ω_j that contains K_1. At the l-th step we apply Theorem 7.7 to $u-\sum_1^{l-1} v_m$ on K_l to obtain a smooth function v_l with support in the open Ω_j that contains K_l. Note that $u-\sum_1^l v_m$ vanishes on $\cup_{m=1}^l K_m$ so once we have exhausted the cells we have $\sum_l v_l = u$. By construction, each of the v_l is smooth, supported on some Ω_j and satisfies (7.5). Setting u_j to be the sum of those v_l that are supported on Ω_j completes the proof. □

Distributions with point support on products. It is useful to begin with the observation that if $T' \in \mathcal{D}'(X')$ and $T'' \in \mathcal{D}'(X'')$ are distributions on the components of a product space $X'\times X''$ then there is a tensor distribution $T'\times T''$ which is a distribution on the product. This is not entirely immediate, but follows readily from the structure theorem for the component spaces. Specifically, the fact that T' is locally $(-\Delta')^k f$ for a continuous f implies that for a $\phi\in\mathcal{D}(X'\times X'')$ there are k and f such that

$$\begin{aligned}\Delta'' T'\phi(x',x'') &= \Delta''\int_{X'} f(x')(-\Delta')^k\phi(x',x'')\,d\mu'(x')\\ &= \int_{X'} f(x')(-\Delta')^k\Delta''\phi(x',x'')\,d\mu'(x')\\ &= T'\Delta''\phi(x',x'')\end{aligned}$$

where we used that Δ' and Δ'' commute. In particular $T'\phi$ is smooth in the second variable, so $T'\times T''\phi = T''(T'\phi)$ is well defined. Repeating the calculation with T'' in place of Δ'' ensures that $T''(T'\phi) = T'(T''\phi)$, so the order in which the distributions are applied is not important. Linearity of $T'\times T''$ is immediate and it is easy to check the continuity condition that ensures it is a distribution on $X'\times X''$.

In the special case where T' is supported at x' and T'' is supported at x'' it is apparent that $T'\times T''$ is supported at (x',x''), so this construction and the results of Section 6 supply a large number of distributions with point support. In fact we can show that if x' and x'' are either cell boundary points or satisfy the conditions of Theorem 6.13, then the distributions with support at (x',x'') are of this type. As in Section 6, the key is to show that if $\phi\in\mathcal{D}(X'\times X'')$ is annihilated by a sufficiently large collection of tensor distributions at (x',x'') and if $\epsilon>0$ is given, then it is possible to cut off ϕ on a small neighborhood of (x',x'') such that the resulting function has $\left\|(\Delta')^j(\Delta'')^k\phi\right\|_\infty < \epsilon$ for all j and k such that $j+k\leq m$. It follows that all distributions of order at most m and support $x'\times x''$ are linear combinations of the given tensor distributions.

Our main tool is an adaptation of Theorem 6.2.

Theorem 7.9. *Given a test function ϕ and a cell $K = K'\times K''$ with $K' = F'_{w'}(X')$ and $K'' = F''_{w''}(X'')$, there is a test function ψ such that $\psi=\phi$ on K and*

$$\left\|(\Delta')^j(\Delta'')^k\psi\right\|_\infty \leq C(m,n)\sum_{l=0}^{m}\sum_{i=0}^{n}(r'_{w'}\mu'_{w'})^{l-j}(r''_{w''}\mu''_{w''})^{i-k}\left\|(\Delta')^l(\Delta'')^i\phi\right\|_{L^\infty(K)}+\epsilon \tag{7.7}$$

for all $0 \leq j \leq m$ *and* $0 \leq k \leq n$.

Proof. The method for cutting-off a smooth function on a cell has already been described in the proof of Theorem 7.7. Since we cut off first in one variable and then in the other, the estimates from the proof of Theorem 6.2 may be applied directly. Suppose that we cut off in the first variable and then in the second. Taking (6.4) for the Laplacian $(\Delta')^k$ in the first variable on a fixed slice $U' \times \{y''\}$ and substituting from the second lines of both of (6.2) and (6.3), gives

$$\left\|(\Delta')^j \psi\right\|_{L^\infty(U' \times \{y''\})} \leq C(m) \sum_{l=0}^{m} (r'_{w'}\mu'_{w'})^{l-j} \left\|(\Delta')^l \phi\right\|_{L^\infty(K' \times \{y''\})} + \epsilon$$

provided $j \leq m$. In this calculation we used that the harmonic part of a function (which was denoted $H_{i,j}$ in the proof of Theorem 6.2) is bounded by the L^∞ norm of the function because of the maximum principle, and we extracted the scaling factor $r'_{w'}\mu'_{w'}$ of the Laplacian on $K' = F'_{w'}(X')$ using the same argument as in (6.5).

The same estimate is true with the same proof when ψ is replaced by $(\Delta'')^k \psi$ and ϕ by $(\Delta'')^k \phi$. We use this fact when we repeat the estimate in the second variable, because in this case we are cutting off the function that was modified at the first step. A little algebra then produces the desired estimate. ◻

Theorem 7.10. *Let T be a distribution supported at $(x', x'') \in X' \times X''$. Suppose that x' is such that either Theorem 6.6 or Theorem 6.13 may be used to identify the distributions with support at x', and make the same assumption for x''. Then T is a finite linear combination of tensor products $T' \times T''$ where T' is supported at x' and T'' is supported at x''.*

Proof. In light of the preceding discussion and Theorem 7.9, it suffices to show that if the given tensor distributions vanish on a test function ϕ then the right side of (7.7) may be made less than 2ϵ by taking K sufficiently small. The proof of this estimate is elementary: we simply go from (x', x'') to (y', y'') by using two Taylor-like expansions, one in each variable.

Since T has compact support it also has finite order m. It then seems reasonable that each of the terms $T' \times T''$ should be made up of a T' of order $k \leq m$ and a T'' of order at most $m-k$. Unfortunately we cannot prove this in general because our scaling estimates are insufficiently refined, as was explained in Remark 6.7. This result is true if the distributions at x' and x'' are such that none have scaling exactly equal to that of the Laplacian (meaning that if they are as in Theorem 6.5 then there is no γ_s equal to a power of $r\mu$, and if they are as in Corollary 6.12 then there is no γ_s equal to a power of β_w). Given the limitations of our estimates we must instead allow the possibility that T' is order $k+1$ and T'' is order $m-k+1$.

Suppose then that $T' \times T''\phi = 0$ for all T' of order up to $k+1$ and T'' of order up to $m-k+1$. It follows that the differential $(\bar{D}'')^{m-k}$ vanishes on the one-variable smooth function $T'\phi(x', \cdot)$. The same reasoning as was used at the beginning of the proofs of Theorem 6.6 and Theorem 6.13 shows that then $T'\phi = o(r''_{w''}\mu''_{w''})^{m-k}$ on the set $\{x'\} \times K''$, so in particular at (x', y'').

We now wish to repeat the argument to go from (x', y'') to (y', y''). Instead of having vanishing distributions in the first variable at (x', y'') we have only estimates on their size, which we use to estimate the size of $(\bar{D}')^m \phi$. Recall from Definitions 6.3 and 6.10 that the differential $(\bar{D}')^m \phi$ for the second variable on the cell K' consists of a harmonic function with coefficients obtained using distributions of order at most $m+1$, as well as

$G'(\bar{D}')^{m-1}\Delta'\phi$. where G' is the Green's operator for the cell K'. The harmonic function is itself made up of pieces (one for each $k \leq m$) with scaling bounded by $(r'_{w'}\mu'_{w'})^k$ (or an equivalent quantity involving $\beta'_{w'}$) and coefficients obtained using distributions in the first variable with order at most $k+1$. The estimate of the previous paragraph says that these coefficients are $o(r''_{w''}\mu''_{w''})^{m-k}$, so each term of the harmonic functions is $o((r'_{w'}\mu'_{w'})^k(r''_{w''}\mu''_{w''})^{m-k})$ on $K' \times \{y''\}$. A similar argument applies to $G'(\bar{D}')^{m-1}\Delta'\phi$, because the G' produces an extra factor of $r'_{w'}\mu'_{w'}$, and the harmonic piece of $(\bar{D}')^{m-1}\Delta'\phi$ that has scaling $(r'_{w'}\mu'_{w'})^{k-1}$ is obtained via distributions of order at most k applied to $\Delta'\phi$, each of which is a distribution of order $k+1$ applied to ϕ. We may repeat this reasoning inductively across the terms of $(\bar{D}')^m\phi$ to obtain a bound of the form

$$\left|(\bar{D}')^m\phi\right| = o\Big(\sum_{k=0}^{m}(r'_{w'}\mu'_{w'})^k(r''_{w''}\mu''_{w''})^{m-k}\Big). \tag{7.8}$$

on the set $K' \times \{y''\}$. Since we also know (from (6.10) and (6.25)) that

$$\left|\phi - (\bar{D}')^l\phi\right| = o(r'_{w'}\mu'_{w'})^m$$

on $K' \times \{y''\}$ we conclude that the estimate (7.8) is also true for ϕ itself. The point $y'' \in K''$ was arbitrary, so we have

$$\|\phi\|_{L^\infty(K)} = o\Big(\sum_{k=0}^{m}(r'_{w'}\mu'_{w'})^k(r''_{w''}\mu''_{w''})^{m-k}\Big). \tag{7.9}$$

Our working thus far has shown that if $T' \times T''\phi = 0$ for all T' of order up to $k+1$ and T'' of order up to $m-k+1$, then (7.9) holds. However, this assumption obviously implies that $T' \times T''\big((\Delta')^l(\Delta'')^i\phi\big) = 0$ for all T' of order up to $k+1-l$ and T'' of order up to $m-k+1-i$ if $l+i \leq m$ and $0 \leq k \leq (m-i-l)$. Thus (7.9) improves to

$$\left\|(\Delta')^l(\Delta'')^i\phi\right\|_{L^\infty(K)} = o\Big(\sum_{k=0}^{m-i-l}(r'_{w'}\mu'_{w'})^k(r''_{w''}\mu''_{w''})^{m-i-l-k}\Big). \tag{7.10}$$

Substituting into (7.7) for the cutoff of ϕ yields

$$\left\|(\Delta')^j(\Delta'')^k\phi\right\|_\infty \leq \epsilon + o\Big(\sum_{l=0}^{m}\sum_{i=0}^{n}(r'_{w'}\mu'_{w'})^{l-j}(r''_{w''}\mu''_{w''})^{i-k}\sum_{s=0}^{m-i-l}(r'_{w'}\mu'_{w'})^k(r''_{w''}\mu''_{w''})^{m-i-l-k}\Big).$$

The simplest way to complete the argument is to choose K' and K'' such that $r'_{w'}\mu'_{w'}$ and $r''_{w''}\mu''_{w''}$ are comparable, at which point all terms in the sum are bounded. It follows that the sum term is $o(1)$ so can be made less than ϵ by requiring that K' and K'' are also sufficiently small. □

8. Hypoellipticity

An important question in the analysis of PDE is to identify conditions under which a distributional solution of a PDE is actually a smooth function. In Euclidean space, an archetypal example is Weyl's proof that a weak solution of the Laplace equation is actually C^∞. In order to study these questions one uses the notion of hypoellipticity, which we may now define in the setting of fractafolds based on p.c.f. fractals and their products. We will not settle any of the questions about hypoellipticity here, but simply suggest some natural problems for which the distribution theory we have introduced is the correct setting.

We first define the singular support of a distribution, which intuitively corresponds to those points where the distribution is not locally smooth.

Definition 8.1. A distribution T is smooth on the open set $\Omega_1 \subset \Omega$ if there is $u \in \mathcal{E}(\Omega_1)$ such that

$$T\phi = \int u\phi \, d\mu \text{ for all } \phi \in \mathcal{D}(\Omega_1)$$

Using Lemma 3.3 for the case of a single p.c.f. fractal, or the analogous result derived from Theorem 7.8 in the product setting, we see that if T is smooth on Ω_1 and on Ω_2 then it is smooth on $\Omega_1 \cup \Omega_2$, thus there is a maximal open set on which T is smooth.

Definition 8.2. For a distribution T, Let Ω_T be the maximal open set on which T is smooth. The singular support of T is the set

$$\text{SingSppt}(T) = \text{Sppt}(T) \setminus \Omega_T$$

Let P be a polynomial of order k on $\mathbb{R}^m$, so $P(\xi) = \sum_{|\kappa| \leq k} a_\kappa \xi^\kappa$ where $\kappa = \kappa_1 \dots \kappa_m$ is a multi-index, $|\kappa| = \sum \kappa_j$ is its length, and $\xi^\kappa = \prod \xi_j^{\kappa_j}$. Consider the linear differential operator $P(\Delta) = P(\Delta_1, \dots, \Delta_m)$ on a product $\prod X_j$ of p.c.f. self-similar fractals X_j with Laplacians Δ_j. It is clear that for any distribution T we have $\text{SingSppt}(P(\Delta)T) \subseteq \text{SingSppt}(T)$, because when T is represented by $u \in \mathcal{E}(\Omega_T)$ then $P(\Delta)T$ is represented by $P(\Delta)u$. By analogy with the Euclidean case, we define a class of constant coefficient hypoelliptic linear differential operators.

Definition 8.3. $P(\Delta)$ is called hypoelliptic if $\text{SingSppt}(P(\Delta)T) = \text{SingSppt}(T)$ for all $T \in \mathcal{D}'(\Omega)$.

Given the importance of hypoelliptic operators in the analysis of PDE on Euclidean spaces, it is natural to seek conditions that imply hypoellipticity of an operator on a p.c.f. fractal or on products of p.c.f. fractals. We expect that if $P(\Delta)$ is elliptic then it should be hypoelliptic; it also seems possible that the celebrated hypoellipticity criterion of Hörmander [11, Section 11.1] might imply hypoellipticity in the fractal case, though we do not expect conditions of this type to be necessary because of examples like that motivating Conjecture 8.9.

Definition 8.4. For a polynomial $P(\xi) = \sum_{|\kappa| \leq k} a_\kappa \xi^\kappa$, the principal part of P is $P_0 = \sum_{|\kappa|=k} a_\kappa \xi^\kappa$. P is called elliptic if $P_0(\xi) \neq 0$ for $\xi \neq 0$; equivalently P is elliptic if there is $c > 0$ so $\left|P_0(\xi)\right| \geq c|\xi|^k$. We call $P(\Delta)$ elliptic if $P(\xi_1^2, \dots, \xi_m^2)$ is elliptic.

Remark 8.5. The above definition is consistent with the usual one in the case that X is a Euclidean interval rather than a fractal set, but they do not coincide because we are dealing with a smaller class of operators. Specifically, for such X the Laplacian is $\partial^2/\partial x^2$, so our class of operators $\{P(\Delta)\}$ is smaller than the usual collection of constant coefficient linear partial differential operators $P(\partial/\partial x_1, \dots, \partial/\partial x_m)$. Similarly our class of elliptic operators is a strict subset of the classical one.

Conjecture 8.6. *If X is a p.c.f. fractal and $P(\Delta)$ is an elliptic operator on the product space X^m, then $P(\Delta)$ is hypoelliptic.*

In the case that $m = 1$, all operators $P(\Delta)$ are elliptic, and they can all be shown to be hypoelliptic. Indeed, by factoring the polynomial we can reduce to the case of the linear polynomial $\Delta + c$ for some complex constant c. The hypoellipticity of $\Delta + c$ is readily obtained from the fact that on small cells $(\Delta + c)$ has a resolvent kernel that is smooth away from the diagonal, as may be seen by representing the resolvent as an integral with respect to the heat kernel or by applying results from [12].

Conjecture 8.7. *A sufficient condition for the hypoellipticity of* $P(\Delta)$ *is that* $D^\kappa P(\xi)/P(\xi) \to 0$ *as* $\xi \to \infty$ *for any partial derivative* D^κ *with* $|\kappa| > 0$ *(compare to Theorems 11.1.1 and 11.1.3 of* [11]*).*

In the Euclidean setting the above condition is necessary as well as sufficient, but we do not expect this to be the case on fractals. One reason for believing this is that on certain fractals there are a natural class of operators that are not elliptic but that ought to behave somewhat like elliptic operators, and for these operators the condition can fail. In what follows we discuss the idea behind this phenomenon. Note, however, that this discussion is limited to global smoothness expressed using decay of the Fourier coefficients; hypoellipticity is a more subtle question.

Consider the algebraic equation

$$P_0(\lambda_1, \dots, \lambda_m)\hat{u}(\lambda_1, \dots, \lambda_m) = \hat{f}(\lambda_1, \dots, \lambda_m)$$

for any choice of $(\lambda_1, \dots, \lambda_m)$ with each λ_j an eigenvalue of Δ_j. Since all of these λ_j are negative, ellipticity of $P(\Delta)$ says that $|P_0(\lambda_1, \dots, \lambda_m)| \geq c|\lambda_1 + \cdots + \lambda_m|$ for such $(\lambda_1, \dots, \lambda_m)$. Inverting the equation we see this is sufficient to show the Fourier transform $\hat{u}$ has faster decay than $\hat{f}$, so u should be as smooth or smoother than f. However the ellipticity condition should only be necessary for this approach if the points $(\lambda_1, \dots, \lambda_m)$ are dense in the positive orthant $\{\xi : \xi_j \geq 0\}$. In [6] it is shown that this is not the case for the Sierpinski Gasket fractal; specifically it is shown that in the case $m = 2$, the points (λ_1, λ_2) omit an open neighborhood of a ray in the positive orthant. It follows that there is $a > 0$ and $b < 0$ such that $a\Delta_1 + b\Delta_2$ is not elliptic but $-a/b$ lies in the omitted neighborhood, so $|a\lambda_1 + b\lambda_2| \geq c|\lambda_1 + \lambda_2|$ whenever λ_j is an eigenvalue of Δ_j. Note in particular that $a\Delta_1 + b\Delta_2$ does not satisfy the assumption of Conjecture 8.7.

Following [6] we call operators of the above type quasielliptic. Given that quasielliptic operators satisfy elliptic-type estimates on the spectrum, it seems likely that they will have similar smoothness properties to elliptic operators, perhaps including hypoellipticity. L^p estimates for these operators may be found in recent work of Sikora [31].

Definition 8.8. The operator $P(\Delta)$ is quasielliptic if there is $c > 0$ such that $|P_0(\xi)| \geq c|\xi|$ for all $\xi \in \{(\lambda_1, \dots, \lambda_m) : \lambda_j \text{ is an eigenvalue of } \Delta_j\}$.

Conjecture 8.9. *The quasielliptic operators of* [6] *are hypoelliptic.*

References

[1] Carlos Avenancio-Leon and Robert S. Strichartz, *Local behavior of harmonic functions on the Sierpinski gasket*, Illinois J. Math. **51** (2007), no. 4, 1061–1075. MR 2417415

[2] Martin T. Barlow and Jun Kigami, *Localized eigenfunctions of the Laplacian on p.c.f. self-similar sets*, J. London Math. Soc. (2) **56** (1997), no. 2, 320–332. MR 1489140 (99b:35162)

[3] Oren Ben-Bassat, Robert S. Strichartz, and Alexander Teplyaev, *What is not in the domain of the Laplacian on Sierpinski gasket type fractals*, J. Funct. Anal. **166** (1999), no. 2, 197–217. MR 1707752 (2001e:31016)

[4] Nitsan Ben-Gal, Abby Shaw-Krauss, Robert S. Strichartz, and Clint Young, *Calculus on the Sierpinski gasket. II. Point singularities, eigenfunctions, and normal derivatives of the heat kernel*, Trans. Amer. Math. Soc. **358** (2006), no. 9, 3883–3936 (electronic). MR 2219003 (2007h:28009)

[5] ———, *Calculus on the Sierpinski gasket. II. Point singularities, eigenfunctions, and normal derivatives of the heat kernel*, Trans. Amer. Math. Soc. **358** (2006), no. 9, 3883–3936 (electronic). MR 2219003

[6] Brian Bockelman and Robert S. Strichartz, *Partial differential equations on products of Sierpinski gaskets*, Indiana Univ. Math. J. **56** (2007), no. 3, 1361–1375. MR 2333476 (2008h:31012)

[7] Pat J. Fitzsimmons, Ben M. Hambly, and Takashi Kumagai, *Transition density estimates for Brownian motion on affine nested fractals*, Comm. Math. Phys. **165** (1994), no. 3, 595–620. MR 1301625 (95j:60122)

[8] H. Furstenberg and H. Kesten, *Products of random matrices*, Ann. Math. Statist **31** (1960), 457–469. MR 0121828 (22 #12558)

[9] B. M. Hambly and T. Kumagai, *Transition density estimates for diffusion processes on post critically finite self-similar fractals*, Proc. London Math. Soc. (3) **78** (1999), no. 2, 431–458. MR 1665249 (99m:60118)

[10] B. M. Hambly, V. Metz, and A. Teplyaev, *Self-similar energies on post-critically finite self-similar fractals*, J. London Math. Soc. (2) **74** (2006), no. 1, 93–112. MR 2254554 (2007i:31011)

[11] Lars Hörmander, *The analysis of linear partial differential operators. II Differential operators with constant coefficients.*, Grundlehren der Mathematischen Wissenschaften, vol. 257, Springer-Verlag, Berlin, 1983, MR 705278 (85g:35002b)

[12] Marius Ionescu, Erin P. J. Pearse, Luke G. Rogers, Huo-Jun Ruan, and Robert S. Strichartz, *The resolvent kernel for PCF self-similar fractals*, Trans. Amer. Math. Soc. **362** (2010), no. 8, 4451–4479. MR 2608413

[13] Jun Kigami, *Analysis on fractals*, Cambridge Tracts in Mathematics, vol. 143, Cambridge University Press, Cambridge, 2001. MR 1840042 (2002c:28015)

[14] ———, *Harmonic analysis for resistance forms*, J. Funct. Anal. **204** (2003), no. 2, 399–444. MR 2017320 (2004m:31010)

[15] Takashi Kumagai, *Short time asymptotic behaviour and large deviation for Brownian motion on some affine nested fractals*, Publ. Res. Inst. Math. Sci. **33** (1997), no. 2, 223–240. MR 1442498 (98k:60130)

[16] Shigeo Kusuoka, *Dirichlet forms on fractals and products of random matrices*, Publ. Res. Inst. Math. Sci. **25** (1989), no. 4, 659–680. MR 1025071 (91m:60142)

[17] Tom Lindstrøm, *Brownian motion on nested fractals*, Mem. Amer. Math. Soc. **83** (1990), no. 420, iv+128. MR 988082 (90k:60157)

[18] Robert Meyers, Robert S. Strichartz, and Alexander Teplyaev, *Dirichlet forms on the Sierpiński gasket*, Pacific J. Math. **217** (2004), no. 1, 149–174. MR 2105771 (2005k:31028)

[19] Jonathan Needleman, Robert S. Strichartz, Alexander Teplyaev, and Po-Lam Yung, *Calculus on the Sierpinski gasket. I. Polynomials, exponentials and power series*, J. Funct. Anal. **215** (2004), no. 2, 290–340. MR 2150975

[20] Kasso A. Okoudjou, Luke G. Rogers, and Robert S. Strichartz, *Generalized eigenfunctions and a Borel theorem on the Sierpinski gasket*, Canad. Math. Bull. **52** (2009), 105–116.

[21] Roberto Peirone, *Convergence and uniqueness problems for Dirichlet forms on fractals*, Boll. Unione Mat. Ital. Sez. B Artic. Ric. Mat. (8) **3** (2000), no. 2, 431–460. MR 1769995 (2001i:31016)

[22] ———, *Convergence of Dirichlet forms on fractals*, Topics on concentration phenomena and problems with multiple scales, Lect. Notes Unione Mat. Ital., vol. 2, Springer, Berlin, 2006, pp. 139–188. MR 2267882 (2008b:28013)

[23] ———, *Existence of eigenforms on fractals with three vertices*, Proc. Roy. Soc. Edinburgh Sect. A **137** (2007), no. 5, 1073–1080. MR 2359927

[24] ———, *Uniqueness of eigenforms on nicely separated fractals*, Analysis on graphs and its applications, Proceedings of Symposia in Pure Mathematics, vol. 77, Amer. Math. Soc., Providence, 2008, pp. 231–241.

[25] Anders Pelander, *Solvability of differential equations on open subsets of the Sierpiński gasket*, J. Anal. Math. **102** (2007), 359–369. MR 2346562 (2009d:31020)

[26] Anders Pelander and Alexander Teplyaev, *Infinite dimensional i.f.s. and smooth functions on the Sierpiński gasket*, Indiana Univ. Math. J. **56** (2007), no. 3, 1377–1404. MR 2333477

[27] ———, *Products of random matrices and derivatives on p.c.f. fractals*, J. Funct. Anal. **254** (2008), no. 5, 1188–1216. MR 2386935

[28] Luke G. Rogers, Robert S. Strichartz, and Alexander Teplyaev, *Smooth bumps, a Borel theorem and partitions of smooth functions on P.C.F. fractals*, Trans. Amer. Math. Soc. **361** (2009), no. 4, 1765–1790. MR 2465816

[29] Walter Rudin, *Functional analysis*, second ed., International Series in Pure and Applied Mathematics, McGraw-Hill Inc., New York, 1991. MR 1157815 (92k:46001)

[30] C. Sabot, *Existence and uniqueness of diffusions on finitely ramified self-similar fractals*, Ann. Sci. École Norm. Sup. (4) **30** (1997), no. 5, 605–673. MR 1474807 (98h:60118)

[31] Adam Sikora, *Multivariable spectral multipliers and quasielliptic operators*, Indiana Univ. Math. J. **58** (2009), no. 1, 317–334. MR 2504414 (2010d:47068)

[32] Robert S. Strichartz, *Taylor approximations on Sierpinski gasket type fractals*, J. Funct. Anal. **174** (2000), no. 1, 76–127. MR 1761364 (2001i:31018)

[33] ———, *Fractafolds based on the Sierpiński gasket and their spectra*, Trans. Amer. Math. Soc. **355** (2003), no. 10, 4019–4043 (electronic). MR 1990573 (2004b:28013)

[34] ———, *Function spaces on fractals*, J. Funct. Anal. **198** (2003), no. 1, 43–83. MR 1962353 (2003m:46058)

[35] ———, *Analysis on products of fractals*, Trans. Amer. Math. Soc. **357** (2005), no. 2, 571–615 (electronic). MR 2095624 (2005m:31016)

[36] ______, *Solvability for differential equations on fractals*, J. Anal. Math. **96** (2005), 247–267. MR 2177187 (2006j:35092)

[37] ______, *Differential equations on fractals*, Princeton University Press, Princeton, NJ, 2006, A tutorial. MR 2246975 (2007f:35003)

[38] Alexander Teplyaev, *Spectral analysis on infinite Sierpiński gaskets*, J. Funct. Anal. **159** (1998), no. 2, 537–567. MR 1658094 (99j:35153)

[39] ______, *Gradients on fractals*, J. Funct. Anal. **174** (2000), no. 1, 128–154. MR 1761365 (2001h:31012)

University of Connecticut, Storrs, CT 06269-3009 USA
E-mail address: rogers@math.uconn.edu

Cornell University, Ithaca, NY 14850-4201 USA
E-mail address: str@math.cornell.edu