Self-similarity of solutions to integral and differential equations with respect to a fractal measure

H. Kunze[*], D. La Torre[†], F. Mendivil[‡], E.R. Vrscay[§]

September 20, 2018

Abstract

In this paper we study solutions of a variation of a classical integral equation (based on the Picard operator) in which Lebesgue measure is replaced by a self-similar measure μ. Our main interest is in the fractal nature of the solutions and we use Iterated Function Systems (IFS) tools to investigate the behaviour and self-similarity of these solutions. Both the integral and differential form of the equation are discussed since each brings useful insights. Several convergence results are provided along with illustrative examples that show the applications of the theory when the underlying fractal object is the celebrated Cantor set. Additionally we show that the solution to our integral equation inherits self-similarity from the defining measure μ.

1 Introduction

We are interested in studying the integral equation

$$f(x) = y_0 + \int_0^x \phi(t, f(t))\ d\mu(t) \tag{1.1}$$

where μ is a Borel probability measure on $[0, 1]$. As usual both the integral and differential forms of this equation are useful and provide insights. Our primary interest is the case where μ is a self-similar measure and thus usually singular

[*]Department of Mathematics and Statistics, University of Guelph, Canada. Email:hkunze@uoguelph.ca

[†]University of Milan, Department of Economics, Management and Quantitative Methods, Milan, Italy and Dubai Business School, University of Dubai, Dubai, UAE. Email: davide.latorre@unimi.it, dlatorre@ud.ac.ae

[‡]Department of Mathematics and Statistics, Acadia University, Wolfville, Nova Scotia, Canada. Email: franklin.mendivil@acadiau.ca

[§]Department of Applied Mathematics, University of Waterloo, Ontario, Canada. Email: ervrscay@uwaterloo.ca

(in our examples μ is supported on a Cantor-type set). The solution of (1.1) will have many features in common with the Cantor Ternary function in these cases. We state existence and uniqueness results and also some stability results. Our primary use of the stability results is to explore the solution to (1.1) via a sequence of approximations μ_n to the measure μ; the approximating sequence μ_n is generated by the action of an Iterated Function System with probabilities (IFSP). In addition, we prove that under mild conditions the solution will inherit self-similarity from μ.

There has been considerable previous work on extensions of calculus to several different settings. One of the most popular has been calculus on so-called *time scales* (the original paper is [1], for a more recent book see [2]). In essence a time scale is a closed subset of $\mathbb{R}$, but this does not do the subject justice. There are many similarities between dynamical equations on time scales and what we do but strong differences as well. Using the time scale setting would require us to work with questions of the convergence of sequences of time scales and solutions to dynamical equations on this sequence. Our method is much simpler and more natural for our setting of IFS fractals.

Measure Differential Equations (MDEs) (the original paper is [3], for a more recent book see [4]) is another generalization which encompasses equation (1.1) and its solutions. In fact, our equation is a special case of an MDE. We avoid the general technicalities of MDEs in this paper by concentrating on the simple case a compact domain and bounded positive measure μ. Unlike most work on MDEs, we are specifically interested in singular measures and the resulting fractal behaviour of the solutions.

Another related area is that of Stieltjes derivatives, as nicely explained in [5]. Our problem and definitions also fit exactly into this setting and our definition of a derivative in (2.4) and (2.5) are equivalent to the Stieltjes derivatives. We claim no novelty in this definition (a form of which already appeared in a paper by Young in 1917) but believe that our general results on convergence of solutions on approximations to self-similar sets (and measures) are novel as is our careful examination of these solutions on fractal sets. The paper [6] is the paper most closely related to ours and gives formulas for solutions to first order dynamical equations on the classical 1/3-Cantor set.

This paper is organized as follows. Section 2 gives the basic framework along with existence and uniqueness results, stability results and a discussion of an equivalent formulation in terms of a generalized differential-type equation. There are two purposes for presenting the material in this section (other than setting out our notation). The first purpose is to provide the stability results in order to give a methodology for studying solutions to (1.1) and to justify and motivate our examples in Section 4. The second reason is to provide the reader with a complete and simple presentation of a framework. Certainly equation (1.1) can be studied within the context of Time-Scales or Measure Differential Equations, but giving a complete explanation of these frameworks is unnecessary for our purposes.

Section 3 provides a brief background on IFS and IFSP, but only what is necessary for the rest of the paper. In Section 4 we discuss several examples in

detail in order to illuminate the differences between solutions of classical ODEs and (1.1). Section 5 provides an alternative viewpoint which shows how one may use a change-of-variable to transform (1.1) into a more standard problem. In this section we also briefly indicate how the theory easily extends to higher-order equations. In Section 6 we prove the self-similarity of the solution to (1.1) under quite mild assumptions on the IFSP which generates μ. Finally, the paper closes in Section 7 with some comments and suggestions for future work.

2 Framework

In this section we give the definitions and results which are necessary for our presentation. As mentioned in the Introduction, equation (1.1) could also be cast in the framework of Time-Scales or MDEs, but we choose a simpler and more streamlined framework. Our discussion in this section is valid for a generic measure μ, but our focus in the rest of the paper is on self-similar measures.

As our basic framework, take a Borel probability measure μ on $[0,1]$, a bounded function $\phi : [0,1] \times \mathbb{R} \to \mathbb{R}$, and $y_0 \in \mathbb{R}$ and consider the integral transform T given by

$$T(f)(x) = y_0 + \int_0^x \phi(t, f(t)) \, d\mu(t). \tag{2.2}$$

It is clear that $T(f)$ is bounded if ϕ is bounded. Next, notice that for $0 \leq x < y \leq 1$ we have

$$|T(f)(x) - T(f)(y)| \leq \int_x^y |\phi(t, f(t))| \, d\mu(t) \leq \|\phi\|_\infty \mu([x,y])$$

and thus T maps continuous functions to continuous functions if μ is non-atomic and ϕ is bounded.

Similarly,

$$\begin{aligned} |T(f)(x) - T(g)(x)| &\leq \int_0^x |\phi(t, f(t)) - \phi(t, g(t))| \, d\mu(t) \\ &\leq \mu([0,1]) K_\phi \|f - g\|_\infty \end{aligned} \tag{2.3}$$

where K_ϕ is such that $|\phi(t,x) - \phi(t,y)| \leq K_\phi |x - y|$ for all $t \in [0,1]$ (so is a uniform Lipschitz constant for ϕ in its second argument). Thus if ϕ is Lipschitz then T is Lipschitz as a map from bounded functions to bounded functions, $T : B[0,1] \to B[0,1]$ (recall that $B(X)$ is a Banach space when using the supremum norm).

As mentioned in the Introduction, our primary interest will be in examining the solutions to the integral equation $f(x) = T(f)(x)$ or

$$f(x) = y_0 + \int_0^x \phi(t, f(t)) \, d\mu(t).$$

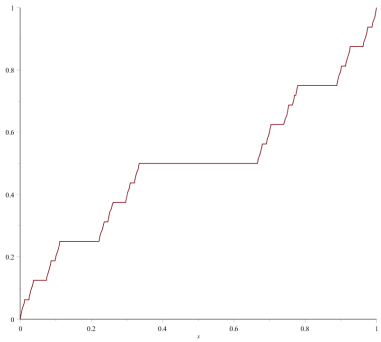

Figure 1: "Devil Staircase": cumulative density function for the uniform measure on the classical 1/3-Cantor set

We are particularly interested in the situation where μ is fractal. For us this means that either the support of μ is a Cantor-type set or the support of μ is $[0,1]$ but its distribution function is fractal-like (see Figure 1 for an example).

First we give some useful notation. For any $0 < \alpha \leq 1$, we use the notation $T|_{[0,\alpha]}$ for the operator T defined by (2.2) but with the restriction $0 \leq x \leq \alpha$. It will become necessary to consider a sequence of operators as defined in (2.2) where both ϕ and μ vary. In these cases we will use the notation $T_{\zeta,\theta}$ for the operator

$$T_{\zeta,\theta}(f)(x) = y_0 + \int_0^x \zeta(t, f(t))\ d\theta(t).$$

It is worth noting that if $f : [0,1] \to \mathbb{R}$ is a fixed point of T then $f|_{[0,\alpha]}$ is a fixed point of $T|_{[0,\alpha]}$. In addition, suppose that for $\alpha > 0$ we have that $T|_{[0,\alpha]}$ has a unique fixed point for each given y_0. Then the usual "local-to-global" pasting technique will give a unique fixed point $f : [0,1] \to \mathbb{R}$ to T for each given y_0.

Recall that $\phi : [0,1] \times \mathbb{R} \to \mathbb{R}$ is *locally Lipschitz in its second argument* if for any compact $K \subset [0,1] \times \mathbb{R}$ there is a $k > 0$ such that $|\phi(t,s) - \phi(t,s')| \leq k|s - s'|$ for all $(t,s),(t,s') \in K$. Notice that if ϕ has a continuous first partial derivative with respect to its second argument then it is locally Lipschitz in its second argument.

We comment that it is possible to remove the condition that ϕ be bounded at the cost of possibly only having local solutions to (1.1).

Proposition 2.1. *Suppose that ϕ is bounded, continuous, and locally Lipschitz in its second argument.*

1. *If μ is non-atomic then (1.1) has a unique solution on $[0,1]$.*
2. *If μ has no atoms of size larger than $\alpha \geq 0$, then (1.1) has a unique solution on $[0,1]$ for any ϕ with uniform Lipschitz constant $K < 1/\alpha$.*

Proof. 1. A straightforward modification of the classical proof will work. More specifically, in [7, Thm 3.1, chapter 2] we simply use $I_\alpha(t_0) = \{t \le t_0 : \mu([t_0, t]) \le \alpha\} \cup \{t \ge t_0 : \mu([t_0, t]) \le \alpha\}$ rather than $I_\alpha(t_0) = [t_0 - \alpha, t_0 + \alpha]$ and the rest of the details are the same. Since μ is non-atomic $I_\alpha(t_0)$ is a closed interval of positive length for any $\alpha > 0$. The "local-to-global" extension works the same and yields a solution defined on all of $[0, 1]$.

2. The assumptions on μ mean that we can choose $a > 0$ so that $\mu([0, a]) \le \alpha$ and thus (2.3) implies that $T : B[0, a] \to B[0, a]$ is contractive and so there is unique solution to (1.1) on $[0, a]$. We continue by repeating the argument with some $b > a$ so that $\mu([a, b]) \le \alpha$ with appropriate initial condition and pasting the solutions together.

□

We will use a sequence of approximations to (1.1) in order to understand the features of a solution. Proposition 2.3 shows that this process works.

Lemma 2.2. *Suppose that $\mu_n \to \mu$ weak* and $\phi_n \to \phi$ uniformly. Then $T_{\phi_n,\mu_n}(f) \to T_{\phi,\mu}(f)$ uniformly on $[0, 1]$. In particular $T_{\phi_n,\mu_n}|_{[0,\alpha]}(f) \to T_{\phi,\mu}|_{[0,\alpha]}(f)$ uniformly on $[0, \alpha]$ for any $0 < \alpha \le 1$.*

Proof. We see that

$$\left| \int_0^1 \phi_n(t, f(t)) d\mu_n(t) - \int_0^1 \phi(t, f(t)) d\mu(t) \right| \le$$
$$\int_0^1 |\phi_n(t, f(t)) - \phi(t, f(t))| d\mu_n(t) + \int_0^1 |\phi(t, f(t))| \, d|\mu_n - \mu|(t).$$

Thus for any x we have

$$|T_{\phi_n,\mu_n}(f)(x) - T_{\phi,\mu}(f)(x)| \le \int_0^1 |\phi_n(t, f(t)) - \phi(t, f(t))| d\mu_n(t) + \int_0^1 |\phi(t, f(t))| \, d|\mu_n - \mu|(t).$$

The first term goes to zero since $\phi_n \to \phi$ uniformly and all μ_n are probability measures. The second term goes to zero since $t \mapsto \phi(t, f(t))$ is continuous and $\mu_n \to \mu$ weak*. □

Proposition 2.3. *Suppose that $\mu_n \to \mu$ weak*, $|\phi(s, t)| \le M$ and $|\phi_n(s, t)| \le M$ for all $(s, t) \in [0, 1] \times \mathbb{R}$ and n, $\phi_n \to \phi$ uniformly on compact subsets of $[0, 1] \times \mathbb{R}$, and ϕ and ϕ_n are all locally Lipschitz with respect to their second argument. Let f_n be the fixed point of T_{ϕ_n,μ_n} and f be the fixed point of $T_{\phi,\mu}$. Then $f_n \to f$ uniformly on $[0, 1]$.*

Proof. The assumptions and definition (2.2) imply that there is a $B > 0$ so that $|f_n(x)| \le B$ and $|f(x)| \le B$ for all $x \in [0, 1]$ and n. Let k be a Lipschitz factor for ϕ and all ϕ_n on the compact set $[0, 1] \times [-B, B]$. By standard arguments (mentioned above), there are $\alpha > 0$ and $0 \le \lambda < 1$ so that $T_{\phi,\mu}|_{[0,\alpha]}$

and all $T_{\phi_n,\mu_n}|_{[0,\alpha]}$ are λ-contractions. However, then by the standard abstract argument (for example, [7, Cor 3.1, chapter 1]) and Lemma 2.2 we have that $f_n|_{[0,\alpha]} \to f|_{[0,\alpha]}$. This along with the "local-to-global" pasting finishes the proof. □

Under slightly different assumptions we can show that the limit of solutions is also a solution.

Proposition 2.4. *Suppose that $\mu_n \to \mu$ weak* and $\phi_n \to \phi$ uniformly on $[0,1] \times \mathbb{R}$ where ϕ_n, ϕ are all locally Lipschitz in their second variable. Furthermore, suppose that $f_n \in C[0,1]$ is a solution to (1.1) using μ_n and ϕ_n and $f_n \to f$ uniformly on $[0,1]$. Then f is a solution to (1.1) using μ and ϕ.*

Proof. By the assumptions on f_n and f there is an $M > 0$ so that $|f_n(x)| \leq M$ for all $x \in [0,1]$ and n. By the assumption of locally Lipschitz there is a $K > 0$ so that $|\phi_m(t,s_1) - \phi_m(t,s_2)| \leq K|s_1 - s_2|$ for all $s_1, s_2 \in [-M, M]$.

Thus we have

$$\begin{aligned}|\phi_m(t,f_m(t)) - \phi(t,f(t))| &\leq |\phi_m(t,f_m(t)) - \phi_m(t,f(t))| + |\phi_m(t,f(t)) - \phi(t,f(t))| \\ &\leq K|f_m(t) - f(t)| + |\phi_m(t,f(t)) - \phi(t,f(t))|\end{aligned}$$

and so $\phi_m(t,f_m(t)) \to \phi(t,f(t))$ uniformly in t over $[0,1]$.

Next we compute that

$$\begin{aligned}|f(x) - \int_0^x \phi(t,f(t))\ d\mu| &\leq |f(x) - f_m(x)| + \left|\int_0^x \phi_m(t,f_m(t))\ d\mu_m - \int_0^x \phi(t,f(t))\ d\mu\right| \\ &\leq |f(x) - f_m(x)| + \int_0^x |\phi_m(t,f_m(t)) - \phi(t,f(t))| d\mu_m \\ &\quad + \int_0^x |\phi(t,f(t))|\ d|\mu - \mu_m|.\end{aligned}$$

The first term can be made arbitrarily small since $f_m \to f$ uniformly. The second term can be made arbitrarily small since $\phi_m(t,f_m(t)) \to \phi(t,f(t))$ uniformly and each μ_m is a probability measure. Finally, the third term can be made arbitrarily small since $\mu_m \to \mu$ in the weak* topology and $t \mapsto \phi(t,f(t))$ is continuous. Thus we have that f is a solution as desired. □

Since we are primarily interested in μ which are the invariant measures for an IFS with probabilities, we have a natural means of constructing the approximating sequence μ_n. In fact, we can choose each μ_n to be absolutely continuous with a piece-wise constant density function. In this case, we can solve for f_n by using classical ODE techniques on each piece and splicing the local solutions together to obtain f_n.

For the remainder of the paper we will assume that μ is non-atomic and that ϕ is bounded, continuous and locally Lipschitz. Thus we have $T : C[0,1] \to C[0,1]$. In addition, the support of μ is a perfect subset of $[0,1]$ (it is the support of μ which could be regarded as a "time scale" if we chose to use that viewpoint).

Equivalent ODE formulation

As usual we can try to formulate (1.1) in differential form. Because μ may not be Lebesgue measure the result is different than in the classical case but leads to a Stieltjes-type derivative (see [5] for a very nice overview of results and comparison with classical calculus). We give some basic derivations just for illustration; for a careful development the reader is invited to see [1, 8, 5, 9].

Intuitively we have

$$T(f)(x+h) - T(f)(x) = \int_x^{x+h} \phi(t, f(t))\ d\mu(t) = \phi(\zeta_h, f(\zeta_h))\mu([x, x+h])$$

for some $\zeta_h \in (x, x+h)$ and thus

$$\frac{T(f)(x+h) - T(f)(x)}{\mu([x, x+h])} = \phi(\zeta_h, f(\zeta_h)) \to \phi(x, f(x)) \text{ as } h \to 0^+$$

if ϕ and f are both continuous at x. We have to be careful at points x which are boundary points of the support of μ (since then the denominator might be zero for all $h > 0$), but fortunately there are only countably many such points and the set of such points has μ measure zero. Motivated by this we define the *μ-derivatives* of a function G to be

$$D_\mu^+(G)(x) := \lim_{h \to 0^+} \frac{G(x+h) - G(x)}{\mu([x, x+h])} \tag{2.4}$$

and

$$D_\mu^-(G)(x) := \lim_{h \to 0^+} \frac{G(x) - G(x-h)}{\mu([x-h, x])} \tag{2.5}$$

when these limits exists. Whenever the two limits exist and are equal we label their common value $D_\mu(G)(x)$ and say that G *is μ-differentiable at x*. Many of the classical properties for differentiable functions hold (in an appropriately amended form) for μ-differentiable functions (see [5] for a nice discussion of this). For instance, monotone functions are μ-differentiable μ-almost everywhere.

Remark 2.5. Letting $\theta(x) = \mu([0, x])$ be the distribution function for μ, we wish to point out to the reader that our $D_\mu(G)$ corresponds with G'_θ from [5].

The standard properties of the derivative hold for D_μ (with a few necessary changes, see [5] for details). In particular, a version of the Fundamental Theorem of Calculus holds which gives that the integral equation (1.1) becomes the μ-differential initial value problem

$$D_\mu(f)(x) = \phi(x, f(x)), \quad f(0) = y_0. \tag{2.6}$$

We must be careful in how we interpret (2.6), in particular $D_\mu(f)(x)$. Clearly if $\mu([x-\epsilon, x+\epsilon]) = 0$ for some $\epsilon > 0$ then $D_\mu(f)$ will not exist at x for any non-constant f. However, such x only occur in the complementary "gaps" of *support*(μ) and thus they play no role in the action of T. In addition, if

$\mu([x-\epsilon, x+\epsilon]) > 0$ but $\mu([x, x+\epsilon]) = 0$ for some $\epsilon > 0$ (so that x is a left endpoint of a "gap"), then only $D_\mu^{-1}(f)$ is well-defined at x. Clearly a parallel statement holds for right endpoints of "gaps". The set of all these endpoints is countable and, since we assume that μ is non-atomic, has μ-measure zero. Thus again it plays a negligible role in the action of T. Annoyingly they do form a set which is dense in $support(\mu)$.

A useful way to think about (2.6) is imagining that vector field ϕ being "turned off" in the "gaps" and turned back on on $support(\mu)$. Of course, when μ is supported on a Cantor-type set it is turned on for the briefest of moments. One could also imagine that all the action only occurs on $support(\mu)$ and the "gaps" are used only to visualize the resulting function (since it is difficult to plot a function whose domain is a Cantor set!)

3 Constructing self-similar measures

In this section we briefly review the construction of self-similar probability measures, since our main interest is in such measures. We use the convenient machinery of *Iterated Function Systems* (IFS) and an IFS with probabilities (IFSP). This construction was developed in [10] (also see [11, 12] for more details).

An IFS on a complete metric space X is a finite set of contractive self-maps $w_i : X \to X$. The *attractor* of the IFS $\{w_i\}$ is the unique non-empty compact set $A \subset X$ which satisfies the self-tiling condition

$$A = \bigcup_i w_i(A). \tag{3.7}$$

This set is usually constructed as the fixed-point of the contractive set-valued mapping $\hat{w}$ given by

$$\hat{w}(B) = \bigcup_i w_i(B).$$

This $\hat{w}$ is contractive in the *Hausdorff distance* on the space of all non-empty compact subsets of X. Thus $\hat{w}^n(B) \to A$ for any appropriate initial set B.

A self-similar measure is also constructed as the fixed point of a contraction on the space $\mathcal{P}(X)$ of Borel probability measures on X. We use $\mathrm{Lip}_1(X)$ to denote the set of real functions on X with Lipschitz factor at most one.

Definition 3.1. Let $\mu, \nu \in \mathcal{P}(X)$. Then the *Monge-Kantorovich distance* between μ and ν is given by

$$d_{MK}(\mu, \nu) = \sup \left\{ \int_X f(x)\ d(\mu - \nu)(x) : f \in \mathrm{Lip}_1(X) \right\}.$$

This metric gives the weak* topology on probability measures (see [12] for more details).

An IFS with probabilities is an IFS $\{w_i\}$ along with an associated set of probabilities p_i. The *invariant measure* μ of an IFSP satisfies the self-similarity

condition

$$\mu(B) = \sum_i p_i \mu(w_i^{-1}(B)), \tag{3.8}$$

where B is an arbitrary Borel set. The measure μ is the fixed point of the so-called *Markov* operator $M : \mathcal{P}(X) \to \mathcal{P}(X)$ given by

$$M\mu(B) = \sum_i p_i \mu(w_i^{-1}(B)). \tag{3.9}$$

In particular, given an arbitrary initial measure $\mu_0 \in \mathcal{P}(X)$, the sequence $\mu_{n+1} = M\mu_n$ converges geometrically fast (in the d_{MK} metric) to the invariant measure μ.

It is worth noting that often M preserves absolute continuity of the measure (and singularity as well); this is true for all of our examples.

As a first simple example, take $X = [0, 1]$, $w_1(x) = x/2$, $w_2(x) = x/2 + 1/2$ and $p_1 = p_2 = 1/2$. Then it is not hard to see that the invariant measure for this IFSP is Lebesgue measure on $[0, 1]$.

As another more interesting example we again take $X = [0, 1]$ but $w_1(x) = x/3$ and $w_2(x) = x/3 + 2/3$ with $p_1 = p_2 = 1/2$. This time we obtain μ as the "uniform" measure on the classical 1/3-Cantor set. Figure 2 illustrates the iteration of M starting with μ_0 being Lebesgue measure on $[0, 1]$. The top row shows the density functions for μ_0, μ_1 and μ_5 while the bottom row shows the (cumulative) distribution functions for μ_0, μ_1, and μ_5. For those readers who have not seen IFSP before it is worth examining this figure as it shows both the construction of the Cantor set (as the support of the limiting measure) and also the redistribution of mass as given by (3.9). The singular nature of the limiting measure is evident from the fact that the "columns" in the densities get taller and thinner.

The top rows of Figures 4, 5, and 6 give the sequence of densities for some other IFSP operators.

From Figure 2 it is clear that the distribution function of μ also has self-similarity. In this case, it is self-similar under the IFSM (an IFS on Maps, see [12]) given by

$$\mathcal{T}(g)(x) = \begin{cases} g(3x)/2, & \text{if } 0 \le x \le 1/3, \\ 1/2, & \text{if } 1/3 < x < 2/3, \\ g(3x-2)/2 + 1/2, & \text{if } 2/3 \le x \le 1. \end{cases} \tag{3.10}$$

As a final comment for this section, it is important to note that the invariant measure μ is non-atomic as long as none of the maps w_i have a finite set as their range.

4 Example DE solutions using fractal measures

In this section we give some examples to illustrate the important behaviour of the solutions to (1.1).

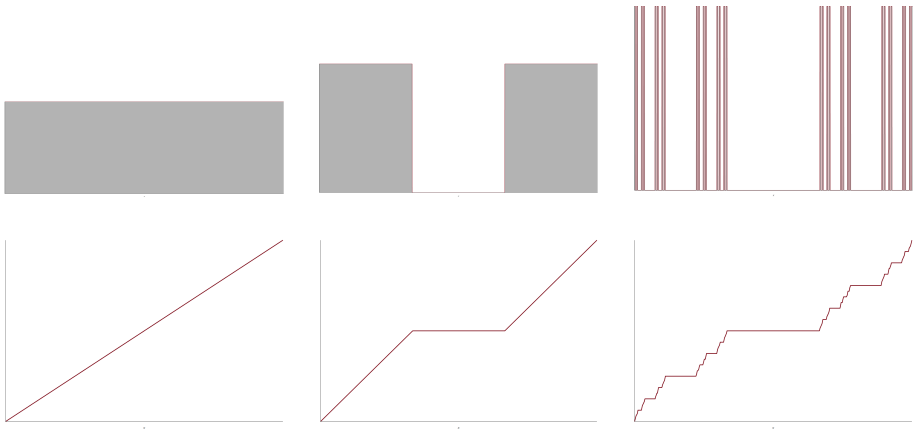

Figure 2: Iteration for "Uniform" Cantor measure: Top row: μ_0, μ_1, μ_5. Bottom row: their distribution functions.

Before we start with the examples, we wish to point out that if $\phi(t,s) = 1$ with $y_0 = 0$ we get the DE $D_\mu(f) = 1$ whose solution is simply the cumulative distribution for μ.

Example 4.1. For our first example we take μ to be the invariant measure for the IFS with probabilities $w_0(x) = x/3$, $w_1(x) = x/3 + 2/3$, $p_0 = p_1 = 1/2$. In this case, the invariant measure μ is a "uniform" measure supported on the classical 1/3-Cantor set. We use Proposition 2.3 to illustrate the solution to

$$D_\mu(f) = f, \quad f(0) = 1. \tag{4.11}$$

In the classical case we expect the solution $f(x) = e^x$ and the result here is a modification of this. In fact, it is a type of "restriction" of the exponential to the Cantor set.

For μ_0 we take Lebesgue measure on $[0,1]$. Then we take the sequence of approximating measures μ_n where $\mu_{n+1} = M\mu_n$, with M the Markov operator for the IFSP (3.9)

$$M\nu = \frac{1}{2}\nu \circ w_0^{-1} + \frac{1}{2}\nu \circ w_1^{-1}.$$

We know that $\mu_n \to \mu$ in the Monge-Kantorovich metric and thus the sequence of solutions f_n converge uniformly to the solution f of (4.11). In Figure 3 the top row illustrates the densities for μ_0, μ_1 and μ_5 (which is visually "close" to μ) and the second row illustrates the solutions f_0, f_1 and f_5 (which is visually close to f). The first thing one notices about f_5 (and, thus f) is that it looks very similar to the graph of the "Devil's staircase" function (shown in Figure 1). This is no surprise since f can only change on the 1/3-Cantor set and is increasing.

Clearly $f_0(x) = e^x$, the solution to the classical problem.

To understand $f_1(x)$, we see that the density for μ_1 is constant and equal to 3/2 over the two intervals $[0,1/3]$ and $[2/3,1]$ and zero on the interval $(1/3,2/3)$.

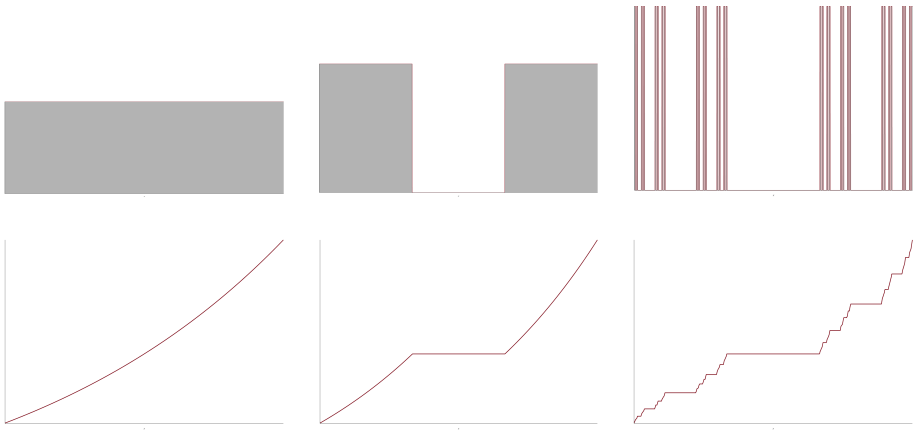

Figure 3: Example 4.1: Top row: μ_0, μ_1, μ_5. Bottom row: f_0, f_1, f_5

Thus the "vector field" is constant over two intervals and "turned off" in the middle. This results in a constant exponential growth over $[0, 1/3]$ and $[2/3, 1]$ and no change over $(1/3, 2/3)$. Said another way, the equation

$$D_{\mu_1} f = f, \quad f(0) = 1, \quad 0 \le x \le 1,$$

results in the three "local" problems

$$\begin{aligned} f' &= \frac{3}{2} f, \quad f(0) = 1, \quad 0 \le x \le 1/3, \\ f' &= 0, \quad f(1/3) = \alpha, \quad 1/3 < x < 2/3, \\ f' &= \frac{3}{2} f, \quad f(2/3) = \beta, \quad 2/3 \le x \le 1, \end{aligned}$$

where α and β are chosen to enforce continuity. Thus,

$$f_1(x) = \begin{cases} e^{3x/2}, & \text{if } 0 \le x \le 1/3, \\ e^{1/2}, & \text{if } 1/3 < x < 2/3, \\ e^{3/2x - 1/2}, & \text{if } 2/3 \le x \le 1. \end{cases}$$

In a similar way the solution $f_2(x)$ is given by

$$f_2(x) = \begin{cases} e^{9x/4}, & \text{if } 0 \le x \le 1/9, \\ e^{1/4}, & \text{if } 1/9 < x < 2/9, \\ e^{9x/4 - 1/4}, & \text{if } 2/9 \le x \le 1/3, \\ e^{1/2}, & \text{if } 1/3 < x < 2/3, \\ e^{9x/4 - 1}, & \text{if } 2/3 \le x \le 7/9, \\ e^{3/4}, & \text{if } 7/9 < x < 8/9, \\ e^{9x/4 - 5/4}, & \text{if } 8/9 \le x \le 1. \end{cases}$$

Notice that $f_0(x) = f_1(x)$ for $x = 0, 1$. Similarly, $f_1(x) = f_2(x)$ for $x = 0, 1$ and $x \in [1/3, 2/3]$. In fact, $f_n(x) = f_{n+1}(x)$ for $x = 0, 1$ and all the points where f_n is constant. Thus moving from f_n to f_{n+1} is a process of refinement; from this we can see the uniform convergence $f_n \to f$.

Compare the exponent of $9x/4$ in f_2 with the exponent of $3x/2$ in f_1 to the exponent x in f_0. The growth rate has to increase when we move from f_0 to f_1 to f_5 since the intervals over which f_n is changing are getting smaller but we have the interpolation property which was just mentioned.

For this example, we have $f_{n+1} = \mathcal{T} f_n$ where

$$\mathcal{T}(g)(x) = \begin{cases} \sqrt{g(3x)}, & \text{if } 0 \le x \le 1/3, \\ e^{1/2}, & \text{if } 1/3 < x < 2/3, \\ e^{1/2}\sqrt{g(3x-2)}, & \text{if } 2/3 \le x \le 1, \end{cases} \tag{4.12}$$

and thus f is the fixed-point of a fractal-type transform on functions. Of course this is not surprising since μ is self-similar.

Example 4.2. For our second example we keep everything the same as the first example but only modify the probabilities to be $p_0 = 1/5$ and $p_1 = 4/5$. Figure 4 illustrates the densities for μ_0, μ_1, and μ_5, and the solutions f_0, f_1 and f_5. It is instructive to compare these to the corresponding parts of Figure 3.

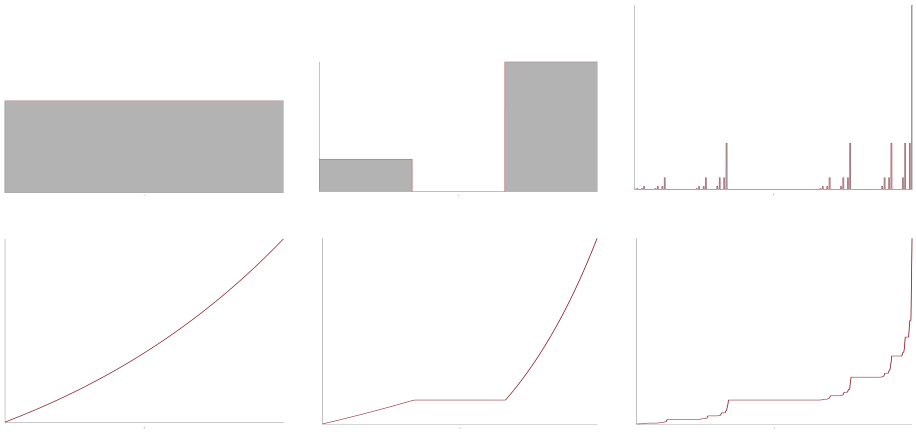

Figure 4: Example 4.2: Top row: μ_0, μ_1, μ_5. Bottom row: f_0, f_1, f_5

In this case we have

$$f_1(x) = \begin{cases} e^{3x/5}, & \text{if } 0 \le x \le 1/3, \\ e^{1/5}, & \text{if } 1/3 < x < 2/3, \\ e^{(12x-7)/5}, & \text{if } 2/3 \le x \le 1. \end{cases} \quad f_2(x) = \begin{cases} e^{9x/25}, & \text{if } 0 \le x \le 1/9, \\ e^{1/25}, & \text{if } 1/9 < x < 2/9, \\ e^{(36x-7)/25}, & \text{if } 2/9 \le x \le 1/3, \\ e^{1/5}, & \text{if } 1/3 < x < 2/3, \\ e^{(36x-19)/25}, & \text{if } 2/3 \le x \le 7/9, \\ e^{9/25}, & \text{if } 7/9 < x < 8/9, \\ e^{(144x-119)/25}, & \text{if } 8/9 \le x \le 1. \end{cases}$$

Notice that the growth rate varies across the intervals as a result of the two different probabilities. The self-similar relationship $f_{n+1} = \mathcal{T} f_n$ is given by the operator

$$\mathcal{T}(g)(x) = \begin{cases} (g(3x))^{1/5}, & \text{if } 0 \le x \le 1/3, \\ e^{1/5}, & \text{if } 1/3 < x < 2/3, \\ e^{1/5}(g(3x-2))^{4/5}, & \text{if } 2/3 \le x \le 1. \end{cases} \tag{4.13}$$

Example 4.3. Our third example illustrates the effects of different geometric contraction factors along with different probabilities but keeps the same function ϕ as before. This time we use the IFS maps $w_0(x) = 2x/5$ and $w_1(x) = x/3+2/3$ along with $p_0 = 4/5$ and $p_1 = 1/5$.

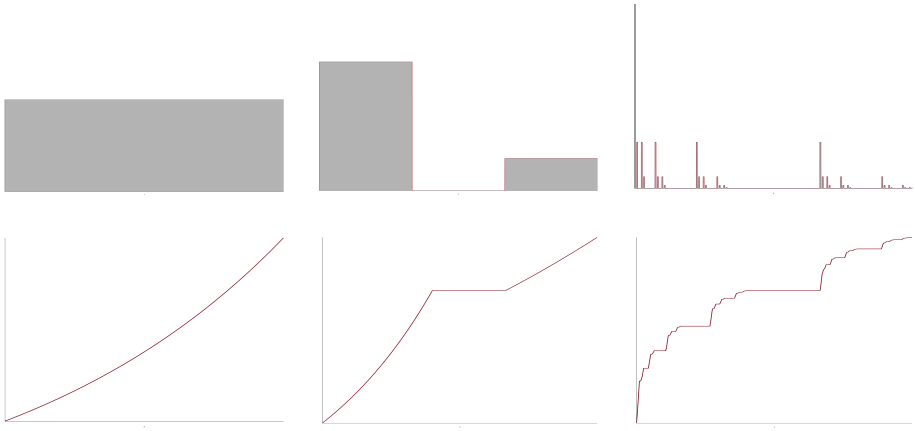

Figure 5: Example 4.3: Top row: μ_0, μ_1, μ_5. Bottom row: f_0, f_1, f_5

In this case we have

$$f_1(x) = \begin{cases} e^{2x}, & \text{if } 0 \le x \le 2/5, \\ e^{4/5}, & \text{if } 2/5 < x < 2/3, \\ e^{(3x+2)/5}, & \text{if } 2/3 \le x \le 1. \end{cases} \quad f_2(x) = \begin{cases} e^{4x}, & \text{if } 0 \le x \le 4/25, \\ e^{16/25}, & \text{if } 4/25 < x < 4/15, \\ e^{(6x+8)/25}, & \text{if } 4/15 \le x \le 2/5, \\ e^{4/5}, & \text{if } 1/3 < x < 2/3, \\ e^{6x/5}, & \text{if } 2/3 \le x \le 4/5, \\ e^{24/25}, & \text{if } 4/5 < x < 8/9, \\ e^{(9x+16)/25}, & \text{if } 8/9 \le x \le 1. \end{cases}$$

The self-similar relationship $f_{n+1} = \mathcal{T} f_n$ is given by the operator

$$\mathcal{T}(g)(x) = \begin{cases} (g(5x/2))^{4/5}, & \text{if } 0 \le x \le 2/5, \\ e^{4/5}, & \text{if } 2/5 < x < 2/3, \\ e^{4/5}(g(3x-2))^{1/5}, & \text{if } 2/3 \le x \le 1. \end{cases} \tag{4.14}$$

Example 4.4. Our next example is one where μ is fully supported but uses the same function ϕ as before. This time we use the IFS maps $w_0(x) = x/2$, $w_1(x) = x/2 + 1/2$ along with the probabilities $p_0 = 4/5$ and $p_1 = 1/5$.

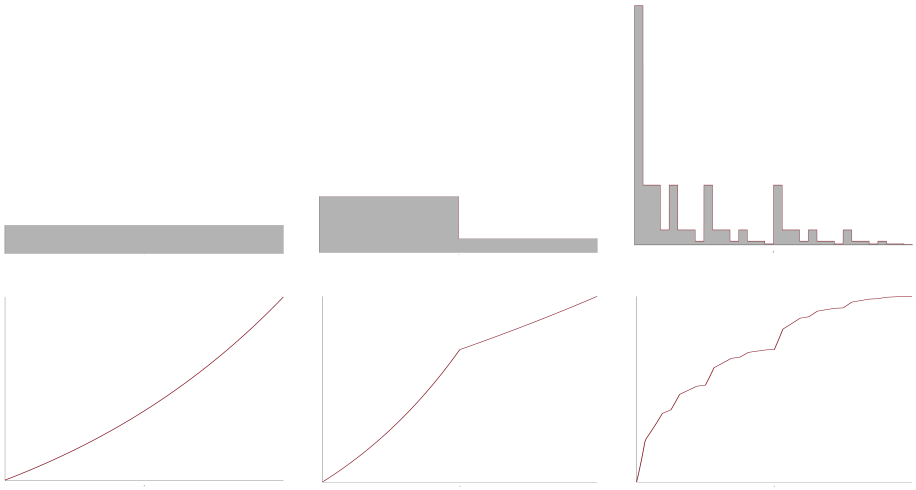

Figure 6: Example 4.4: Top row: μ_0, μ_1, μ_5. Bottom row: f_0, f_1, f_5

The self-similar relationship $f_{n+1} = \mathcal{T} f_n$ is given by the operator

$$\mathcal{T}(g)(x) = \begin{cases} (g(2x))^{4/5}, & \text{if } 0 \le x \le 1/2, \\ e^{4/5}(g(2x-1))^{1/5}, & \text{if } 1/2 \le x \le 1. \end{cases} \tag{4.15}$$

Example 4.5. Our last example in this section will take μ to be the invariant measure for the IFS with probabilities $w_0(x) = x/3$, $w_1(x) = x/3 + 2/3$, $p_0 =$

$p_1 = 1/2$ (so μ is the "uniform" measure on the 1/3-Cantor set) but this time we choose $\phi(t,s) = s - s^2$ and $y_0 = 2$, so we are solving the problem

$$D_\mu(f) = f - f^2, \quad f(0) = 2 \tag{4.16}$$

whose classical solution on $[0,1]$ is $\psi(t) = 1/(1 - e^{-t}/2)$. The plots of the μ_i and f_i are shown in Figure 6. The self-similar relationship $f_{n+1} = \mathcal{T} f_n$ is given by the operator

$$\mathcal{T}(g)(x) = \begin{cases} \Psi_0(g(3x)), & \text{if } 0 \le x \le 1/3, \\ \frac{2}{2-e^{-1/2}}, & \text{if } 1/3 < x < 2/3, \\ \Psi_1(g(3x-2)), & \text{if } 2/3 \le x \le 1 \end{cases} \tag{4.17}$$

where

$$\Psi_0(y) = \frac{2}{2 - \sqrt{2 - 2/y}} \quad \text{and} \quad \Psi_1(y) = \frac{2}{2 - e^{-1/2}\sqrt{2 - 2/y}}.$$

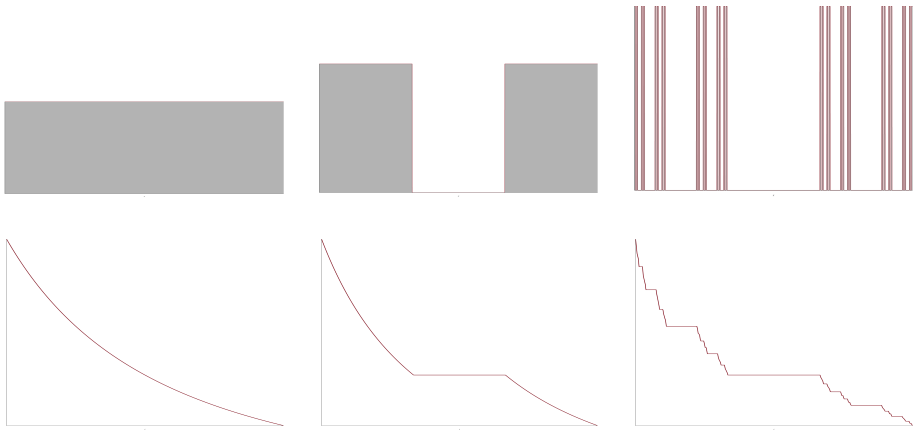

Figure 7: Example 4.5: Top row: μ_0, μ_1, μ_5. Bottom row: f_0, f_1, f_5

5 Lebesgue measure as a prototype for an arbitrary non-atomic measure

In this section we show how we can transform equation (1.1) into a more standard type of integral equation involving Lebesgue measure. This is accomplished by using a change-of-variable which is defined using the measure μ.

We use λ to denote Lebesgue measure on $[0,1]$. As above, let $\phi : [0,1] \times \mathbb{R} \to \mathbb{R}$ and μ be given. Define $\theta : [0,1] \to [0,1]$ by $\theta(x) = \mu([0,x])$, so that θ is

surjective, non-decreasing, and continuous (since μ has no atoms). We define θ^{-1} by

$$\theta^{-1}(y) := \inf\{x : \theta(x) = y\}.$$

Notice that if θ is invertible this is simply the inverse but in general we have $\theta^{-1}(\theta(x)) \leq x$ and $\theta(\theta^{-1}(y)) = y$. Since $\mu([\theta^{-1}(\theta(x)), x]) = 0$ if and only if $\theta^{-1}(\theta(x)) < x$ we know that $\theta^{-1}(\theta(x)) = x$ for any $x \in \mathit{support}(\mu)$ which is not the right-hand endpoint of some "gap", an interval in $[0,1] \setminus \mathit{support}(\mu)$. This makes it clear that $\mathit{range}(\theta^{-1})$ is equal to $\mathit{support}(\mu)$ without these right-hand endpoints.

What motivates our choice of θ is that $\mu = \lambda \circ \theta$ (this is the fact that drives the well-known sampling method of *cdf inversion*). To see this just note that for $0 \leq a < b \leq 1$ we have $\lambda(\theta([a,b])) = \lambda([\theta(a), \theta(b)]) = \theta(b) - \theta(a) = \mu([a,b])$.

Notice that θ is a quotient map $\mathit{support}(\mu) \to [0,1]$. Recall that a quotient map is a continuous surjection which maps closed sets to closed sets. θ is continuous since μ is non-atomic, is surjective since μ is a probability measure, and is a closed map since $[0,1]$ is compact. The idea is that each closure of a gap in $[0,1] \setminus \mathit{support}(\mu)$ is collapsed to a point – the gaps are "closed". The topology of the image is just that of $[0,1]$, all the components of $\mathit{support}(\mu)$ are "glued together". However, the geometry of the identification is given by the way that the mass of μ is distributed on its support.

Now suppose that ψ is a solution to

$$\psi(z) = y_0 + \int_0^z \phi(\theta^{-1}(t), \psi(t))\ d\lambda(t). \tag{5.18}$$

Then we claim that $f = \psi \circ \theta$ is a solution to (1.1). To see this we compute

$$\begin{aligned} f(x) = \psi(\theta(x)) &= y_0 + \int_0^{\theta(x)} \phi(\theta^{-1}(t), \psi(t))\ d\lambda(t) \\ &= y_0 + \int_0^x \phi(\theta^{-1}(\theta(\tau)), \psi(\theta(\tau)))\ d(\lambda \circ \theta)(\tau) \\ &= y_0 + \int_0^x \phi(\theta^{-1}(\theta(\tau)), f(\tau))\ d\mu(\tau) \\ &= y_0 + \int_0^x \phi(\tau, f(\tau))\ d\mu(\tau). \end{aligned}$$

(where we used the substitution $t = \theta(\tau)$ and the fact that $\mu(\{x : \theta^{-1}(\theta(x)) = x\}) = 1$.

Higher order equations

The same process can be used for higher order equations. For example, take ψ a solution to

$$y_0 + \int_{x=0}^y \int_{t=0}^x g(\psi(t))\ d\lambda(t)\ d\lambda(x)$$

(so to $y'' = g(y)$ with $y(0) = y_0$ and $y'(0) = 0$) and define $f(x) = \psi(\theta(x))$. For notational convenience let

$$G(x) = \int_{t=0}^{x} g(\psi(t))\ d\lambda(t).$$

Then

$$\begin{aligned} f(y) = \psi(\theta(y)) &= y_0 + \int_0^{\theta(y)} G(x)\ d\lambda(x) \\ &= y_0 + \int_0^{y} G(\theta(\tau))\ d(\lambda \circ \theta)(\tau) \\ &= y_0 + \int_0^{y} \left\{ \int_{t=0}^{t=\theta(\tau)} g(\psi(t))\ d\lambda(t) \right\} d\mu(\tau) \\ &= y_0 + \int_0^{y} \left\{ \int_{s=0}^{s=\tau} g(\psi(\theta(s)))\ d(\lambda \circ \theta) \right\} d\mu(\tau) \\ &= y_0 + \int_0^{y} \int_0^{\tau} g(f(s))\ d\mu(s)\ d\mu(\tau), \end{aligned}$$

just as before. In this way, we can transform a solution to

$$y'' = g(y),\ y(0) = y_0,\ y'(0) = 0$$

into a solution for

$$D_\mu^2(f) = g(f),\ f(0) = y_0,\ D_\mu f(0) = 0$$

in the same way as we did for first-order equations.

Conversion of μ-derivative to a "standard" derivative

In fact, this all leads to considering the possibility of relating the μ-derivative to a "normal" derivative. To this end, let μ be a non-atomic probability measure on $[0,1]$, $C = \mathit{support}(\mu)$ and $f : [0,1] \to \mathbb{R}$. It is clear that the limit defining $D_\mu f(x)$ does not exist for any x in a complementary gap $[0,1] \setminus C$ and only exists as a one-sided limit for those $x \in C$ which are endpoints of some complementary gap.

As before let $\theta(x) = \mu([0,x])$ and $\theta^{-1}(y) := \inf\{x : \theta(x) = y\}$.

Lemma 5.1. *Let $x, x_n \in C$ with x not an endpoint of a complementary gap. Then $x_n \to x$ iff $\theta(x_n) \to \theta(x)$.*

Proof. One direction is clear since θ is continuous. Thus suppose that $y_n := \theta(x_n) \to \theta(x) =: y$ and for the sake of obtaining a contradiction suppose that $x_n \not\to x$. Then there is some $\epsilon > 0$ and a subsequence x_{n_k} with $|x_{n_k} - x| \geq \epsilon$ but $y_{n_k} \to y$. We must have that either $x_{n_k} < x$ infinitely often or $x < x_{n_k}$

infinitely often. With no loss of generality we assume that $x_{n_k} < x$ for all k and thus $|y_{n_k} - y| = \mu((x_{n_k}, x]) \to 0$. Since x is not an endpoint of a complementary gap there is some $z \in C$ with $z < x$ and $x - z < \epsilon/2$. But this means that

$$|y_{n_k} - y| = \mu((x_{n_k}, x]) \geq \mu([z, x]) > 0$$

for all k, since μ is non-atomic. This contradicts $y_n \to y$. □

Let $f : [0,1] \to \mathbb{R}$ be given and define $F : [0,1] \to \mathbb{R}$ by $F(y) = f(\theta^{-1}(y))$.

Proposition 5.2. *Let $x \in support(\mu)$ not be an endpoint of a complementary gap and define $y = \theta(x)$. Then*

$$D_\mu f(x) = \lim_{\eta \to x, \eta \in support(\mu)} \frac{f(\eta) - f(x)}{\theta(\eta) - \theta(x)} = \lim_{\gamma \to y} \frac{F(\gamma) - F(y)}{\gamma - y} = F'(y).$$

Proof. This is simply an application of the change-of-variable $y = \theta(x)$ along with the definitions and Lemma 5.1. □

Thus any statement involving the μ-derivative on $support(\mu) \setminus \{$ endpoints $\}$ can be transferred to an equivalent statement on $[0,1]$ involving the "standard" derivative. This is accomplished by the change-of-variable θ which "uniformizes" the length scale.

6 Self-similarity of solutions

In the examples in Section 4 we saw that all of the solutions possessed self-similarity. In this short section we will prove that this is a general feature, at least in a slightly restricted class of problems. It is likely that this is true for a wider range of problems, but we leave this investigation for future work.

Assume that either $\phi(t,s) < 0$ or $\phi(t,s) > 0$ for all $s \in \mathbb{R}$ and $t \in [0,1]$. The main purpose of this assumption is for the classical solution to be invertible since it is then either strictly decreasing or strictly increasing.

In addition, take an injective IFS w_i, $i = 0, 1, \ldots, N$, on $[0,1]$ with $w_i(x) < w_j(y)$ for all $x, y \in [0,1]$ whenever $i < j$. This puts an order on the ranges $w_i([0,1])$ and ensures that they are disjoint. Assume that $0 \in w_0([0,1])$ and $1 \in w_N([0,1])$ and let G_i be the open interval (or "gap") between $w_{i-1}([0,1])$ and $w_i([0,1])$. As usual we also have a set of probabilities $p_0, p_1, \ldots, p_N$ which we assume are all strictly positive. Let $\theta(x) = \mu([0,x])$ be the distribution function for μ, the invariant measure for this IFSP.

Let $\psi : [0,1] \to \mathbb{R}$ be the solution to the classical problem

$$dy/dt = \phi(\theta^{-1}(t), y(t)), \quad y(0) = y_0.$$

Define the maps Ψ_i, for $i = 0, 1, \ldots, N$ by

$$\Psi_i(y) = \psi(p_0 + p_1 + \cdots p_{i-1} + p_i \psi^{-1}(y)\,). \tag{6.19}$$

Proposition 6.1. *Let $\{w_i, p_i\}$ be the IFSP discussed above and μ be its invariant measure and let ϕ satisfy the conditions above. Then the solution f to (1.1) for this choice of μ and ϕ is self-similar under the operator $\mathcal{T}$ defined by*

$$\mathcal{T}(g)(x) = \begin{cases} \Psi_i(g(w_i^{-1}(x))), & \text{if } x \in w_i([0,1]), \\ \psi(p_0 + p_1 + \cdots + p_i), & \text{if } x \in G_i. \end{cases}$$

Proof. We first note that θ is self-similar under the operator $\mathcal{J}$ given by

$$\mathcal{J}(g)(x) = \begin{cases} p_0 + p_1 + \cdots p_{i-1} + p_i g(w_i^{-1}(x))), & \text{if } x \in w_i([0,1]), \\ p_0 + p_1 + \cdots + p_i, & \text{if } x \in G_i. \end{cases}$$

Next we note that the solution f can be written as $f = \psi \circ \theta$, as explained in Section 5. Finally, we see that

$$\begin{aligned} \Psi_i(f(w_i^{-1})) &= \psi(p_i \psi^{-1}(\psi(\theta(w_i^{-1}))) + p_0 + \cdots + p_{i-1}) \\ &= \psi(p_i \theta(w_i^{-1}(x)) + p_0 + \cdots + p_{i-1}) = \psi(\theta(x)) \\ &= f(x) \end{aligned}$$

for $x \in w_i([0,1])$. For $x \in G_i$ we have

$$f(x) = \psi(\theta(x)) = \psi(p_0 + p_1 + \cdots + p_{i-1})$$

and thus $f = \mathcal{T} f$, as desired. □

7 Comments and future directions

The change-of-variable we give in Section 5 is simple because the measure μ is non-atomic and thus the cumulative density function is a continuous; invariant measures of IFSP are non-atomic if the IFS maps w_i are reasonable (in our examples they are all injective). We also wish to point out that the equivalent differential equation after the change-of-variable may have a discontinuous vector field, particularly in the non-autonomous case.

Although our examples in Section 4 are all of autonomous equations there is no problem with non-autonomous equations. Of course generating the solutions to the approximations using μ_n is a bit more involved, but the issues are computational and not theoretical. However, for non-autonomous equations the change-of-variable is a bit more involved since it involves a change of time in the vector field as well.

It would be interesting to compare the behaviour of solutions of higher-order μ-differential equations for self-similar μ; this study would be a relatively straightforward extension of our results. Certainly one way to do this is via a system of equations and thus vector-valued integral and μ-differential equations. There should be no essential difficulty in doing this as the estimates would be the same. However, the nature of the oscillations might be interesting, especially if one could get self-similar oscillations down the scales. It would also be interesting to examine fractional differential equations in this context.

Another direction for future work is examining the self-similarity of the solutions to (1.1) in the case that the IFS is overlapping. This introduces substantial complications which can be seen even in the simple case of $\phi(t, s) = 1$ where the solution is the cumulative density function for μ.

Acknowledgments

The third author (FM) was partially supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) in the form of a Discovery Grant (238549).

References

[1] S. Hilger, Analysis on measure chains – a unified approach to continuous and discrete calculus, *Results Math.* **18**, (1990), 18-56.

[2] M. Bohner and A. Peterson, *Dynamic equations on time scales*, Birhäuser, Boston, 2001.

[3] W.W. Schmaedeke, Optimal control theory for nonlinear vector differential equations containing measures, *SIAM J. Control* **3**, ((1965) no 2, 231-280.

[4] B. Brogliato, *Nonsmooth Mechanics. Models, Dynamics and Control, 3rd edition*, Springer-Verlag, Switzerland, 2016.

[5] R. L. Pouso, A. Rodríguez, A new unification of continuous, discrete, and impulsive calculus through Stieltjes derivatives, *Real Anal. Exchange* **40** (2014/15), No. 2, 319-353.

[6] S. Petrovic, Dynamic equations on the Cantor set, *PanAmer. Math. J.* **13** (2003), No. 4, 1-18.

[7] J. Hale, *Ordinary Differential Equations*, Dover, Mineola, NY, 1997.

[8] Monteiro, Giselle and Satco, Bianca, Distributional, differential and integral problems: equivalence and existence results, *Electon. J. Qual. Theory Differ. Equ.* (2017), No. 7, 1-26.

[9] A. Slavík, Dynamic equations on time scales and generalized ordinary differential equations, *J. Math. Anal. Appl.* **385** (2012), 534-550.

[10] J. Hutchinson, Fractals and self-similarity, Indiana Univ. J. Math., 30, 713-747 (1981).

[11] M.F. Barnsley, *Fractals Everywhere*, Academic Press, New York (1989).

[12] H. Kunze, D. La Torre, F. Mendivil and E.R. Vrscay, *Fractal-based Methods in Analysis*, Springer, New York (2012).