

**Manuscript version: Author's Accepted Manuscript**

The version presented in WRAP is the author's accepted manuscript and may differ from the published version or Version of Record.

**Persistent WRAP URL:**

<http://wrap.warwick.ac.uk/135640>

**How to cite:**

Please refer to published version for the most recent bibliographic citation information. If a published version is known of, the repository item page linked to above, will contain details on accessing it.

**Copyright and reuse:**

The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions.

Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

**Publisher's statement:**

Please refer to the repository item page, publisher's statement section, for further information.

For more information, please contact the WRAP Team at: [wrap@warwick.ac.uk](mailto:wrap@warwick.ac.uk).

# EQUIVALENCE OF WEAK AND STRONG MODES OF MEASURES ON TOPOLOGICAL VECTOR SPACES

HAN CHENG LIE AND T. J. SULLIVAN

ABSTRACT. A strong mode of a probability measure on a normed space  $X$  can be defined as a point  $u$  such that the mass of the ball centred at  $u$  uniformly dominates the mass of all other balls in the small-radius limit. Helin and Burger weakened this definition by considering only pairwise comparisons with balls whose centres differ by vectors in a dense, proper linear subspace  $E$  of  $X$ , and posed the question of when these two types of modes coincide. We show that, in a more general setting of metrisable vector spaces equipped with measures that are finite on bounded sets, the density of  $E$  and a uniformity condition suffice for the equivalence of these two types of modes. We accomplish this by introducing a new, intermediate type of mode. We also show that these modes can be inequivalent if the uniformity condition fails. Our results shed light on the relationships between among various notions of maximum a posteriori estimator in non-parametric Bayesian inference.

## 1. INTRODUCTION

A probability measure  $\mu$  on a topological vector space  $X$  can be described using summary statistics such as means and covariances or, as in this article, *modes*, meaning points of maximum  $\mu$ -mass in an appropriate sense. There are multiple ways of defining such modes, particularly for infinite-dimensional vector spaces; the *strong mode* of Dashti et al. [9] and the *weak mode* of Helin and Burger [17] describe the notion of a point of maximum probability in a separable Banach space by examining the probabilities of norm balls in the small-radius limit.

In the setting of Bayesian inverse problems, the modes of the posterior are known as *maximum a posteriori estimators* or *MAP points*). MAP points play an important role because of their interpretation as being the most likely solutions of the inverse problem given the observed data [26]. A far-from-exhaustive list of studies that have considered MAP points is [2, 4–7, 11, 14–16, 18–21, 24, 27, 28]. Modes play an important role in other areas of applied mathematics as well. For example, in mathematical models of chemical reactions, rare events of diffusion processes on energy landscapes play an important role; the rare events of interest are typically transitions from one energy well or metastable state to another. The qualitative description of such rare events requires studying the paths that a diffusion process is most likely to take. Studying the modes of the law of the diffusion process on the associated path space is often done with the help of a least-action principle formulated by an Onsager–Machlup functional, and an application of Freidlin–Wentzell theory or large deviations theory [10, 12, 13, 22].

---

*Date:* August 21, 2018.

*2010 Mathematics Subject Classification.* 28C20 (62G35 49Q20 49N45) .

*Key words and phrases.* mode, small ball probabilities, topological vector space, maximum a posteriori estimation, minimum action principle.

Onsager–Machlup functionals — and, more generally, variational approaches — play an important role in Bayesian inverse problems. When  $\mu$  is a Bayesian re-weighting of a Gaussian prior, the strong mode of Dashti et al. (see [9, Definition 3.1] or our Definition 2.2) is the minimiser of an appropriate Onsager–Machlup functional [9, Theorem 3.5]. This functional can be formally regarded as a misfit regularised by the negative logarithm of the prior density. Under the assumptions considered in [17], the weak mode of Helin and Burger (see [17, Definition 4] or our Definition 2.3) corresponds to the points where certain logarithmic derivatives of  $\mu$  vanish; see [17, Theorems 2 and 3]. Recently, Agapiou et al. [1] resolved one of the questions raised by Helin and Burger, by extending the weak mode formalism and variational approach to analyse  $L^1$ -priors on Besov spaces; these priors are perceived as having sparsity-promoting properties advantageous to inverse problems and imaging.

Helin and Burger [17] asked under what circumstances a weak mode is also strong mode. While it followed immediately from the definitions that a strong mode is a weak mode, the converse was far from clear. Their motivation for the question would be to determine whether the weak mode was a genuinely different type of mode. Another motivation for demonstrating the equivalence of strong and weak modes is that, if it were known that every weak mode is a strong mode, then one could apply the variational characterisation of weak modes in terms of logarithmic derivatives in order to obtain the strong modes of Dashti et al. In particular, one could enlarge the available set of tools for identifying strong modes — and thus MAP estimators for Bayesian inverse problems — by drawing upon the powerful theory of differentiable measures [3].

In this work, we formulate the question of when a weak mode is a strong mode in a more general context than that considered in [17]. We present two conditions that jointly suffice for the equivalence of strong and weak modes: a uniformity condition (2.6), and the topological density of the set  $E$  with which one defines the weak mode. We introduce an intermediate type of mode and reduce the original question to the task of identifying sufficient conditions for two simpler equivalence statements, presented as Theorem 2.5 and Theorem 2.6 below. We state the necessary definitions and results in Section 2, and prove Theorems 2.5 and 2.6 in Section 3 and Section 4 respectively. We also illustrate the importance of the uniformity condition (2.6) by giving in Example 4.4 an example of a measure that does not satisfy the uniformity condition and has weak modes but no strong modes.

The significance of Theorems 2.5 and 2.6 for inverse problems is that, at the level of generality in which Helin and Burger formulated their question, weak modes and strong modes are not equivalent. Some of the assumptions therein, e.g. the existence of representatives of Radon–Nikodym derivatives that are continuous on  $X$ , are neither sufficient nor necessary. In addition, our main results show that one can consider weak modes for a much larger class of measures, since we make no assumptions about differentiability of  $\mu$ . When one wishes to compute modes, then such assumptions may be useful for the variational characterisation of weak modes as zeros of logarithmic derivatives, but they are not relevant for the purpose of determining the equivalence of strong and weak modes.

In the context of inverse problems, our results provide theoretical support for certain choices of negative log-likelihood or potential functions, in which the Onsager–Machlup functional assumes the value  $+\infty$  outside a topologically dense subset

of  $X$  [9, Equation (2.1)]. Our results apply not only to normed spaces, but also to metric spaces such as  $F$ -spaces, so that one may apply the MAP approach in function spaces that have desirable sparsity properties, e.g.  $L^p$  for  $0 < p < 1$ . We believe that the study of MAP points on more general spaces than normed vector spaces will prove to be relevant in future research on Bayesian inverse problems, and that our results provide a useful first step in this direction.

In Section 5, we consider some extensions of our results. We show that modes can be defined for complex vector spaces in Section 5.1, and demonstrate that similar equivalence statements hold for the *local* modes defined by Agapiou et al. [1]. In Section 5.3, we show that a probability measure on a topological vector space need not have a mode (Example 5.5), and we show that the set of modes can depend on the choice of the set  $K$  (Example 5.6). The results of the latter section indicate that the study of modes contains subtle surprises.

## 2. MAIN RESULTS

In this paper,  $\mathbb{N} := \{1, 2, \dots\}$ ,  $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$ , and  $\mathbb{K} := \mathbb{R}$  or  $\mathbb{C}$ . We shall work in the following setting:

**Assumption 2.1.**  *$X$  is a first countable, Hausdorff topological vector space over  $\mathbb{K}$ ;  $\mu$  is a non-zero measure on the Borel  $\sigma$ -algebra of  $X$  that is finite on bounded sets; and  $K$  is a bounded, open neighbourhood of the origin.*

In [9, 17],  $X$  is a separable Banach space,  $K$  is its unit norm ball, and  $\mu$  is a Borel probability measure with topological support  $\text{supp}(\mu) = X$ . Here, we do not assume that  $X$  is complete or separable, that  $\text{supp}(\mu) = X$ , or that  $\mu$  is finite on  $X$ . The existence of a set  $K$  satisfying Assumption 2.1 implies that  $X$  is metrisable [23, Theorem 6.2.1].

For any  $\lambda \in \mathbb{K}$ ,  $x \in X$ , and  $E \subset X$ ,  $x + \lambda E := \{x + \lambda y \mid y \in E\}$ . We denote the collection of open neighbourhoods of  $x \in X$  by  $\mathcal{N}(x)$ . For any  $U \in \mathcal{N}(0)$ , let  $U(x, \lambda) := x + \lambda U$ . For  $K$  as in Assumption 2.1, define the evaluation map  $f: X \times \mathbb{R}_{>0} \rightarrow \mathbb{R}_{\geq 0}$  by

$$(2.1) \quad f(x, r) := \mu(K(x, r)) \geq 0.$$

In Section 5.1, we consider the more general case of nonzero (but not necessarily strictly positive) radius parameter  $r$ . This is important in situations when a problem is formulated over a complex topological vector space  $X$ : in such cases, admitting only real, strictly positive values of the radius parameter  $r$  would limit the ensuing analysis to the real restriction of  $X$ . Assumption 2.1 implies that  $f(x, r)$  is finite for all  $(x, r) \in X \times \mathbb{R}_{>0}$ . Below, we reformulate [9, Definition 3.1] of a MAP estimator for  $\mu$ .

**Definition 2.2.** A *strong mode* (or simply *mode*) of  $\mu$  is any  $u \in X$  satisfying

$$(2.2) \quad \lim_{r \downarrow 0} \frac{\sup_{z \in X} f(z, r)}{f(u, r)} = 1.$$

The intuition behind (2.2) is that, if  $u$  is a mode, then, for sufficiently small  $r$ , translating the  $rK$ -neighbourhood of  $u$  does not yield an increase in measure.

Next, we reformulate [17, Definition 4] of a weak MAP estimator for  $\mu$ .

**Definition 2.3.** Let  $E \subset X$ . An  $E$ -weak mode (or simply *weak mode*) of  $\mu$  is any  $u \in \text{supp}(\mu)$  such that

$$(2.3) \quad \lim_{r \downarrow 0} \frac{f(u-v, r)}{f(u, r)} \leq 1, \quad \forall v \in E.$$

A weak mode  $u$  differs from a strong mode in that it considers translations of  $K(u, r)$  only by vectors in  $E$ ; thus every mode is an  $E$ -weak mode for every  $E \subset X$ . We require  $u \in \text{supp}(\mu)$  in order to exclude “0/0 problems”: for example, if  $\mu := \frac{1}{2}(\mu_1 + \mu_2)$ , where  $\mu_1$  (resp.  $\mu_2$ ) is the standard one-dimensional Gaussian measure on  $\{+1\} \times \mathbb{R}$  (resp.  $\{-1\} \times \mathbb{R}$ ), and  $E := \text{span}(e_2)$ , then  $0 \notin \text{supp}(\mu)$  is an  $E$ -weak mode in the 0/0 sense.

We now introduce a new definition.

**Definition 2.4.** Let  $E \subset X$ . An  $E$ -strong mode of  $\mu$  is any  $u \in \text{supp}(\mu)$  such that

$$(2.4) \quad \lim_{r \downarrow 0} \frac{\sup_{z \in u-E} f(z, r)}{f(u, r)} \leq 1.$$

Note that an  $X$ -strong mode is simply a strong mode. By comparing (2.2), (2.3), and (2.4), and observing that  $u - E \subset X$  for any  $u \in X$  and  $E \subset X$ , we have the following chain of implications:

$$(2.5) \quad \text{Strong} \implies E\text{-strong} \implies E\text{-weak}$$

Helin and Burger asked when an  $E$ -weak mode (for a topologically dense subspace  $E$  of some Banach space  $X$ ) is also a strong mode. In this article, we address their question by investigating the converses of the implications in (2.5).

**Theorem 2.5.** *Under Assumption 2.1, if  $E$  is topologically dense in  $X$ , then  $u$  is an  $E$ -strong mode if and only if  $u$  is a strong mode.*

**Theorem 2.6.** *Suppose that Assumption 2.1 holds, and let  $E$  be a nonempty subset. If the pair  $(u, E)$  satisfies the uniformity condition*

$$(2.6) \quad \exists (v^*, r^*) \in E \times (0, 1) \text{ such that, for all } r \in (0, r^*), \\ f(u - v^*, r) = \sup_{z \in u-E} f(z, r),$$

*then  $u$  is an  $E$ -weak mode if and only if  $u$  is an  $E$ -strong mode.*

Note that neither Theorem 2.5 nor Theorem 2.6 make any claim about the existence of modes of any type. In addition, neither of these theorems requires  $E$  to be a subspace of  $X$ , unlike the assumption on [9, p. 2] or [17, Assumption (A1)].

Theorems 2.5 and 2.6 yield a positive answer to Helin and Burger’s question:

**Corollary 2.7.** *Under Assumption 2.1, if  $E$  is topologically dense in  $X$ , and if  $(u, E)$  satisfies the uniformity condition (2.6), then  $u$  is an  $E$ -weak mode if and only if  $u$  is a strong mode.*

On the other hand, *without* the uniformity condition (2.6), Example 4.4 shows that the answer to Helin and Burger’s question is generally *negative*, i.e. there exist finite Borel measures — even on  $X = \mathbb{R}^2$  — with weak modes but no strong modes.

3. THE EQUIVALENCE OF  $E$ -STRONG AND STRONG MODES

In this section, we seek conditions for which (2.4) implies (2.2):

$$\lim_{r \downarrow 0} \frac{\sup_{z \in u-E} f(z, r)}{f(u, r)} \leq 1 \implies \lim_{r \downarrow 0} \frac{\sup_{z \in X} f(z, r)}{f(u, r)} = 1.$$

In Lemma 3.3, we show that the evaluation map  $f$  defined in (2.1) is lower semicontinuous. We use this powerful property to derive the second key result of this section, Proposition 3.4, which shows that topological density of  $E$  suffices for every  $E$ -strong mode to be a strong mode. By (2.5), this yields the conclusion of Theorem 2.5.

We begin this section with a useful definition.

**Definition 3.1.** The *coincident limiting ratios* condition holds for  $x$  and  $E \subset X$  if

$$(CLR) \quad \lim_{r \downarrow 0} \frac{\sup_{z \in x-E} f(z, r)}{f(x, r)} = \lim_{r \downarrow 0} \frac{\sup_{z \in X} f(z, r)}{f(x, r)}.$$

**Lemma 3.2.** *If  $u$  is an  $E$ -strong mode, and if  $u$  and  $E$  satisfy (CLR), then  $u$  is a strong mode.*

*Proof.* By Definition 2.4 and (CLR), the limit in (2.2) is less than or equal to 1. Since the limit cannot be strictly less than 1, the result follows.  $\square$

We now show the lower semicontinuity of the evaluation map  $f(\cdot, r): X \rightarrow \mathbb{R}_{\geq 0}$ , for any  $r > 0$ . We shall use this property to prove Theorem 2.5.

**Lemma 3.3.** *For arbitrary  $r > 0$ ,  $f(\cdot, r)$  is lower semicontinuous on  $X$ .*

*Proof.* Fix an arbitrary  $x \in X$  and an arbitrary sequence  $(x_n)_{n \in \mathbb{N}}$  converging to  $x$ . To prove the desired statement, it suffices to show that

$$(3.1) \quad \forall y \in X, \quad \mathbb{I}_{K(x, r)}(y) \leq \liminf_{n \rightarrow \infty} \mathbb{I}_{K(x_n, r)}(y),$$

by considering separately the cases of when  $y \in K(x, r)$ , when  $y \notin \text{cl}(K(x, r))$ , and when  $y \in \partial K(x, r)$ . Given (3.1), it follows from (2.1) and Fatou's lemma that

$$f(x, r) \leq \int_X \liminf_{n \rightarrow \infty} \mathbb{I}_{K(x_n, r)}(y) \mu(dy) \leq \liminf_{n \rightarrow \infty} \int_X \mathbb{I}_{K(x_n, r)}(y) \mu(dy).$$

Since the rightmost term is equal to  $\liminf_{n \rightarrow \infty} f(x_n, r)$ , it follows that  $f(\cdot, r)$  is lower semicontinuous at  $x$ , for arbitrary  $x \in X$ . It remains only to establish (3.1).

Let  $y \in K(x, r)$  be arbitrary. Since  $y \in K(x, r)$ , and since  $K(x, r)$  is open, there exists a bounded  $U \in \mathcal{N}(0)$  such that, for all  $v \in U$ ,  $y - v \in K(x, r)$ . Since  $y - v \in K(x, r)$  if and only if  $y \in K(x + v, r)$ , it follows that  $z \in (x + U) \implies y \in K(z, r)$ . Since  $x_n \rightarrow x$  as  $n \rightarrow \infty$ , there exists some  $N \in \mathbb{N}$  such that  $x_n \in (x + U)$  for all  $n \geq N$ . Then  $\mathbb{I}_{K(x_n, r)}(y) = 1$  for  $n \geq N$ , and

$$(3.2) \quad \forall y \in K(x, r), \quad \mathbb{I}_{K(x_n, r)}(y) \rightarrow \mathbb{I}_{K(x, r)}(y).$$

Now let  $y \notin \text{cl}(K(x, r))$  be arbitrary. Since  $\text{cl}(K(x, r))$  is closed, there exists a bounded  $U' \in \mathcal{N}(0)$  such that, for all  $v \in U'$ ,  $y - v \notin \text{cl}(K(x, r))$ . Since  $y - v \notin \text{cl}(K(x, r))$  if and only if  $y \notin \text{cl}(K(x + v, r))$ , it follows that  $z \in (x + U') \implies y \notin \text{cl}(K(z, r))$ . Since  $x_n \rightarrow x$ , there exists some  $N \in \mathbb{N}$  such that  $x_n \in x + U'$  for all  $n \geq N$ . This implies that  $\mathbb{I}_{K(x_n, r)}(y) = 0$  for all  $n \geq N$ . Thus,

$$(3.3) \quad \forall y \notin \text{cl}(K(x, r)), \quad \mathbb{I}_{K(x_n, r)}(y) \rightarrow \mathbb{I}_{K(x, r)}(y).$$

Observe that  $\liminf_{n \rightarrow \infty} \mathbb{I}_{K(x_n, r)}(y)$  is either 0 or 1, because  $\mathbb{I}_{K(x_n, r)}(y)$  is either 0 or 1. On the other hand, since  $K(x, r)$  is open,  $\mathbb{I}_{K(x, r)}(y) = 0$  for every  $y \in \partial K(x, r)$ . Thus,  $\mathbb{I}_{K(x, r)} \leq \liminf_{n \rightarrow \infty} \mathbb{I}_{K(x_n, r)}$  on  $\partial K(x, r)$ .  $\square$

We now use lower semicontinuity to show that topological density of  $E$  in  $X$  suffices for any pair  $(x, E)$  to satisfy (CLR).

**Proposition 3.4.** *Suppose that Assumption 2.1 holds. If  $E$  is topologically dense in  $X$ , then  $\sup_{z \in X} f(z, r) = \sup_{z \in x-E} f(z, r)$  for all  $(x, r)$ . In particular, (CLR) holds for any  $x \in X$ .*

*Proof.* The second conclusion follows immediately from the first, so it suffices to prove the first conclusion. Let  $(x, r) \in X \times \mathbb{R}_{>0}$  be arbitrary, and suppose that  $\sup_{z \in X} f(z, r) = +\infty$ . Let  $M \in \mathbb{N}$  be arbitrary. Recall that Lemma 3.3 yields the lower semicontinuity of  $f(\cdot, r)$ . Given lower semicontinuity of  $f(\cdot, r)$ , the set  $\{z \in X \mid f(z, r) > M\}$  is open; the set is also nonempty since  $\sup_{z \in X} f(z, r) = +\infty$ . Since the topological density of  $E$  in  $X$  implies the topological density of  $x - E$  in  $X$ , it follows that there exist  $z \in x - E$  with  $f(z, r) > M$ . Since  $M$  was arbitrary, it follows that  $\sup_{z \in x-E} f(z, r) = +\infty$  as well.

Now suppose that  $\sup_{z \in X} f(z, r)$  is finite, and let  $(x_n)_{n \in \mathbb{N}} \subset X$  be such that  $f(x_n, r) \in \mathbb{R}$  and  $f(x_n, r) \rightarrow \sup_{z \in X} f(z, r)$ . It holds that  $\{z \in X \mid f(z, r) > f(x_n, r)\}$  is open and nonempty. Since  $x - E$  is topologically dense in  $X$ , there exist  $z \in x - E$  such that  $f(z, r) > f(x_n, r)$ . Since  $f(x_n, r) \rightarrow \sup_{z \in X} f(z, r)$ , this implies that  $\sup_{z \in x-E} f(z, r) \geq \sup_{z \in X} f(z, r)$ .  $\square$

Theorem 2.5 follows from the preceding observations.

*Proof of Theorem 2.5.* By (2.5), every strong mode is an  $E$ -strong mode for any subset  $E \subset X$ . For the converse, since  $E$  is topologically dense in  $X$ , Proposition 3.4 implies that (CLR) holds for  $u$  and  $E$ . Applying the hypothesis that  $u$  is an  $E$ -strong mode and Lemma 3.2 completes the proof.  $\square$

For the sake of completeness, we show that, under the hypotheses that  $E$  is topologically dense in  $X$  and  $u$  is an  $E$ -strong mode, it holds that  $u$  is a strong mode if and only if the coincident limiting ratios condition holds for  $u$  and  $E$ . We shall use the following lemma:

**Lemma 3.5.** *Let  $u$  be an  $E$ -strong mode, and suppose that  $u$  and  $E$  satisfy*

$$(3.4) \quad 1 \leq \lim_{r \downarrow 0} \frac{\sup_{z \in u-E} f(z, r)}{f(u, r)}.$$

*If  $u$  is also a mode, then  $u$  satisfies (CLR).*

*Proof.* We prove the contrapositive. Suppose that (CLR) does not hold; then

$$1 \leq \lim_{r \downarrow 0} \frac{\sup_{z \in x-E} f(z, r)}{f(x, r)} < \lim_{r \downarrow 0} \frac{\sup_{z \in X} f(z, r)}{f(x, r)}$$

and by (2.2) it follows that  $u$  is not a mode.  $\square$

To establish a sufficient condition for the hypothesis (3.4) to hold, we use the lower semicontinuity of  $f(\cdot, r)$  for any  $r > 0$ , and the fact that Assumption 2.1 implies  $X$  is a metrisable space.

**Proposition 3.6.** *Let  $u \in \text{supp}(\mu)$  and  $\emptyset \neq E \subset X$ . Suppose Assumption 2.1 holds. If either  $E$  contains the origin or is topologically dense in a neighbourhood of the origin, then  $u$  and  $E$  satisfy (3.4).*

*Proof.* Suppose  $0 \in E$ . Then  $u \in u - E$ , and  $f(u, r) \leq \sup_{z \in u - E} f(z, r)$  for all  $r$ .

Suppose that  $E$  is topologically dense in a neighbourhood  $V$  of the origin. Then  $u - E$  is dense in  $u - V$ , and there exists a sequence of points in  $(u - E) \cap (u - V)$  that converges to  $u$ . Let  $r > 0$  be arbitrary. By Lemma 3.3, it follows that  $f(\cdot, r)$  is lower semicontinuous at  $u$ . Using lower semicontinuity of  $f(\cdot, r)$  at  $u$ , and using the fact that  $X$  is a metrisable space, it follows that  $\liminf_{n \rightarrow \infty} f(u_n, r) \geq f(u, r)$ . Since  $(u_n)_{n \in \mathbb{N}} \subset u - E$ , it follows that  $\sup_{z \in u - E} f(z, r) \geq f(u, r)$ .  $\square$

**Corollary 3.7.** *Let  $E$  be topologically dense in  $X$ , and let  $u \in \text{supp}(\mu)$  be an  $E$ -strong mode. Then  $u$  is a strong mode if and only if (CLR) holds for  $u$  and  $E$ .*

*Proof.* Sufficiency of (CLR) for  $u$  to be a mode follows from Lemma 3.2. Necessity follows from Lemma 3.5, which we may apply given that the topological density of  $E$  ensures that we may apply Proposition 3.6.  $\square$

#### 4. THE EQUIVALENCE OF $E$ -WEAK AND $E$ -STRONG MODES

In this section, we seek conditions for which (2.3) implies (2.4), i.e.

$$\left( \lim_{r \downarrow 0} \frac{f(u - v, r)}{f(u, r)} \leq 1, \quad \forall v \in E \right) \implies \lim_{r \downarrow 0} \frac{\sup_{z \in u - E} f(z, r)}{f(u, r)} \leq 1.$$

A recurring observation in this section is that finding general conditions for the relation above to hold is relatively easier when  $X$  is finite-dimensional, compared to when  $X$  is infinite-dimensional. When  $X$  is infinite-dimensional, one such condition is given by the uniformity condition (2.6), which has the interpretation that there exists some point  $u - v^* \in u - E$  such that, for all sufficiently small  $r$ , the  $\mu$ -measure of  $K(u - v^*, r)$  dominates that of  $K(u - v, r)$  for all  $v \in E$ .

The following proposition provides sufficient conditions for a weak mode to be  $E$ -strong, in the case when  $X$  is a finite-dimensional vector space:

**Proposition 4.1.** *Let  $X = \mathbb{R}^n$ , let  $E$  be a dense proper subset of  $X$ , let  $K(0, 1) = B_p(0, 1)$  be the unit  $\ell_p$ -ball centred at the origin for some  $0 < p \leq \infty$ , and let  $\mu$  admit a continuous density  $g$  with respect to the Lebesgue measure  $\lambda$ . If  $u$  is an  $E$ -weak mode and  $g(u) > 0$ , then  $u$  is a strong mode, and thus also an  $E$ -strong mode.*

**Remark 4.2.** Note that the restriction to  $\ell_p$ -balls is unnecessary. We may choose  $K(0, 1)$  to be any bounded set, since the family of  $r$ -dilates of  $K(0, 1)$  will satisfy the property of bounded eccentricity; this property suffices for the application of the Lebesgue differentiation theorem [25, Chapter 3, Corollary 1.7].

*Proof of Proposition 4.1.* Fix an arbitrary  $v \in E$ . Observe that

$$\lim_{r \downarrow 0} \frac{f(u - v, r)}{f(u, r)} = \lim_{r \downarrow 0} \frac{\mu(B_2(u - v, r))}{\mu(B_2(u, r))} = \lim_{r \downarrow 0} \frac{\mu(B_2(u - v, r))}{\lambda(B_2(u - v, r))} \frac{\lambda(B_2(u - v, r))}{\mu(B_2(u, r))}.$$

Since the Lebesgue measure is translation-invariant, it follows that  $\lambda(B_2(u - v, r)) = \lambda(B_2(u, r))$ . Using this and the product rule for convergent real sequences, we have

$$\lim_{r \downarrow 0} \frac{f(u - v, r)}{f(u, r)} = \lim_{r \downarrow 0} \frac{\mu(B_2(u - v, r))}{\lambda(B_2(u - v, r))} \lim_{r \downarrow 0} \frac{\lambda(B_2(u, r))}{\mu(B_2(u, r))}.$$

Thus, by the quotient rule for convergent sequences, the hypothesis that  $g(u) > 0$ , and the Lebesgue differentiation theorem, we obtain

$$\lim_{r \downarrow 0} \frac{\lambda(B_2(u, r))}{\mu(B_2(u, r))} = \left( \lim_{r \downarrow 0} \frac{\mu(B_2(u, r))}{\lambda(B_2(u, r))} \right)^{-1} = (g(u))^{-1}.$$

Applying the Lebesgue differentiation theorem again and using the hypothesis that  $u$  is an  $E$ -weak mode, we thus obtain

$$1 \geq \lim_{r \downarrow 0} \frac{f(u - v, r)}{f(u, r)} = \frac{g(u - v)}{g(u)}, \quad \forall v \in E.$$

The above implies that

$$(4.1) \quad \sup_{z \in u - E} g(z) \leq g(u).$$

By continuity of  $g$  and density of  $E$  in  $X$ , (4.1) implies the stronger statement that  $\sup_{z \in X} g(z) \leq g(u)$ . Translation invariance of Lebesgue measure yields

$$\sup_{z \in X} f(z, r) \leq \sup_{z \in X} g(z) \lambda(B_2(u', r)),$$

for any  $u' \in X$ . Using the properties of limits, the definition of  $f$ , the Lebesgue differentiation theorem, and (4.1), we have

$$\begin{aligned} \lim_{r \downarrow 0} \frac{\sup_{z \in X} f(z, r)}{f(u, r)} &\leq \sup_{z \in X} g(z) \left( \lim_{r \downarrow 0} \frac{f(u, r)}{\lambda(B_2(u, r))} \right)^{-1} \\ &= \sup_{z \in X} g(z) \left( \lim_{r \downarrow 0} \frac{\mu(B_2(u, r))}{\lambda(B_2(u, r))} \right)^{-1} \\ &= \sup_{z \in X} g(z) (g(u))^{-1} \leq 1. \quad \square \end{aligned}$$

Proposition 4.1 indicates the power of the combination of topological density of  $E$  and a continuous density with respect to a nice reference measure like Lebesgue measure.

The next lemma shows that the uniformity condition (2.6) implies that every  $E$ -weak mode is an  $E$ -strong mode.

**Lemma 4.3.** *Let  $\emptyset \neq E \subset X$ , and suppose  $u$  is an  $E$ -weak mode. If  $(u, E)$  satisfy the uniformity condition (2.6), then  $u$  is an  $E$ -strong mode.*

*Proof.* From the elementary properties of limits, we have

$$\lim_{r \downarrow 0} \frac{\sup_{z \in u - E} f(z, r)}{f(u, r)} - \lim_{r \downarrow 0} \frac{f(u - v^*, r)}{f(u, r)} = \lim_{r \downarrow 0} \frac{\sup_{z \in u - E} f(z, r) - f(u - v^*, r)}{f(u, r)}.$$

The hypothesis (2.6) implies that the right-hand side is zero, since the ratio inside the limit vanishes for all  $r < r^*$ . Thus the left-hand side is zero, and the limits agree. Therefore, given the assumption that  $u$  and  $E$  satisfy (2.3), it follows that  $u$  and  $E$  also satisfy (2.4).  $\square$

The proof of Theorem 2.6 follows.

*Proof of Theorem 2.6.* By the right implication in (2.5), every  $E$ -strong mode  $u$  is also an  $E$ -weak mode, regardless of whether or not the pair  $(u, E)$  satisfies the uniformity condition (2.6). On the other hand, if  $(u, E)$  satisfies the uniformity condition, then Lemma 4.3 yields that any  $E$ -weak mode  $u$  is also  $E$ -strong.  $\square$

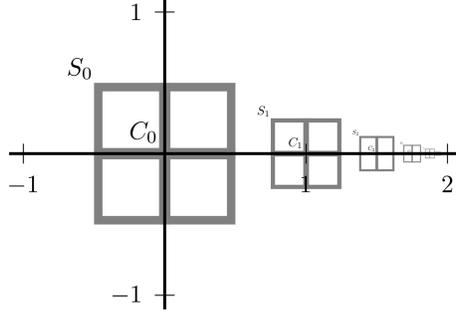


FIGURE 4.1. Illustration of the crossed squares  $C_n \cup S_n$  used in the construction of the measure  $\mu$  on  $\mathbb{R}^2$  in Example 4.4 that has countably infinitely many weak modes but no  $\mathbb{Q}^2$ -strong mode.

The next example involving a geometric sequence of crossed squares demonstrates that it is possible for a vector  $u \in \text{supp}(\mu)$  to satisfy (2.3) but not (2.4), even when  $X$  is a finite-dimensional Banach space and  $\mu$  is a nonatomic, finite measure on  $X$ . The essential idea of the example is to construct a measure such that the ratios of measures of balls decay sufficiently rapidly for (2.3) to hold, but not uniformly rapidly, so that (2.4) does not hold. In addition, although there exist intervals of  $r$ -values in which the supremum  $\sup_{z \in u-E} f(z, r)$  is attained at some element of  $u - E$ , none of these intervals extend to the origin. Hence, there does not exist a  $v \in E$  such that the values of  $\sup_{z \in u-E} f(z, r)$  are dominated by those of  $f(u - v, r)$  for all sufficiently small  $r$ , and the measure  $\mu$  constructed below fails to satisfy the uniformity condition.

**Example 4.4** ( $E$ -weak modes that are not  $E$ -strong). Let  $X = \mathbb{R}^2$ , let  $E = \mathbb{Q}^2$ , let  $K(0, 1) := B_\infty(0, 1)$ , and let  $u$  be the origin of  $X$ . Fix  $\alpha = \frac{7}{8}$ . Define the sequence  $(v_n)_{n \in \mathbb{N}_0} \subset E$  by

$$v_0 = u = (0, 0), \quad v_n := \left( \sum_{m=0}^{n-1} 2^{-m}, 0 \right).$$

For  $i = 1, 2$ , let  $e_i$  denote the  $i^{\text{th}}$  Euclidean basis vector. For  $n \in \mathbb{N}_0$ , let

$$C_n^i := \left\{ v_n + r e_i \mid -\frac{1}{2^{n+1}} < r < \frac{1}{2^{n+1}} \right\}$$

be a line segment of length  $2^{-n}$  centred at  $v_n$  that is parallel to the  $e_i$ -axis. Define the cross-shaped set  $C_n := C_n^1 \cup C_n^2$ , and let  $S_n := \partial B_\infty(v_n, 2^{-(n+1)}) = \bigcup_{i=1}^4 S_n^i$  be the boundary square, where  $(S_n^i)_{i=1}^4$  are the four edges of the square. Thus, as illustrated in Figure 4.1, we have a sequence of crossed squares of the form  $C_n \cup S_n$  centred at  $v_n$ . The side lengths of the crossed squares form a geometric sequence, and the sequence of  $e_1$ -coordinates of the centres give the finite-term approximations of a geometric series converging to 2. Note that, under any  $\ell_p$ -norm on  $X$ , the distance from  $v_n$  to  $v_{n-1}$  is  $2^{-(n-1)}$ , and the distance between the corresponding crossed squares is  $2^{-(n+1)}$ . Therefore,  $\bigcup_{n \in \mathbb{N}_0} (C_n \cup S_n)$  is a disjoint countable union of closed sets that is bounded in  $X$ .

Let  $\mathcal{H}^1$  denote the one-dimensional Hausdorff measure, normalised so that it assigns to any straight line segment its Euclidean length. In particular,  $\mathcal{H}^1(S_n^i) =$

$\mathcal{H}^1(C_n^k)$  for any  $i = 1, \dots, 4$  and  $k = 1, 2$ . Define the measures  $\mu_n^C$ ,  $\mu_n^S$ , and  $\mu_n$  by

$$\begin{aligned}\mu_n^C(B) &:= \mathcal{H}^1(B \cap C_n) = \sum_{i=1}^2 \mathcal{H}^1(B \cap C_n^i), \\ \mu_n^S(B) &:= \frac{\alpha}{2} \mathcal{H}^1(B \cap S_n) = \frac{\alpha}{2} \sum_{i=1}^4 \mathcal{H}^1(B \cap S_n^i), \\ \mu_n(B) &:= \mu_n^C(B) + \mu_n^S(B)\end{aligned}$$

for any Borel set  $B \subseteq \mathbb{R}^2$ , so that  $\text{supp}(\mu_n) = C_n \cup S_n$ . Observe that, since  $\alpha < 1$ ,

$$\mu_n^C(X) = 2^{-(n-1)} > \alpha 2^{-(n-1)} = \mu_n^S(X),$$

and  $\mu_n(X) = (1 + \alpha)2^{-(n-1)}$ . Let  $\mu$  be the measure

$$(4.2) \quad \mu(B) := \sum_{n=0}^{\infty} \mu_n(B) = \sum_{n=0}^{\infty} \mathcal{H}^1(B \cap C_n) + \frac{\alpha}{2} \sum_{n=0}^{\infty} \mathcal{H}^1(B \cap S_n)$$

for any Borel set  $B \subseteq \mathbb{R}^2$ . Observe that  $\mu$  is a finite measure that is not absolutely continuous with respect to Lebesgue measure.

We now show that every element in the set  $(v_n)_{n \in \mathbb{N}_0}$  is an  $E$ -weak mode, i.e. that for every  $n \in \mathbb{N}_0$  we have

$$\lim_{r \downarrow 0} \frac{\mu(B_\infty(v_n - v, r))}{\mu(B_\infty(v_n, r))} \leq 1, \quad \forall v \in E.$$

Note that for every  $n \in \mathbb{N}_0$ , there exists some  $R(n)$  such that  $\mu(B_\infty(v_n, r)) = 4r$  for all  $r < R(n)$ .

Let  $n \in \mathbb{N}_0$  be arbitrary. It suffices to consider  $v \in E$  such that  $v_n - v$  belongs to  $\bigcup_{n \in \mathbb{N}_0} (C_n \cup S_n)$ , since otherwise  $\mu(B_\infty(v_n - v, r)) = 0$  for all sufficiently small  $r$ , and the limit above is zero. If  $v_n - v \in Z$ , then for sufficiently small  $r$  the ratio in the limit is 1. This leaves three cases to consider. Below, we denote the closure of the set  $A$  by  $\text{cl}(A)$ .

*Case 1.* Suppose  $v_n - v \in C_n^i \setminus (v_n)_{n \in \mathbb{N}_0}$  for some  $i \in \{1, 2\}$ . For sufficiently small  $r$ ,  $\mu(B_\infty(v_n - v, r)) = \mu_n^C(B_\infty(v_n - v, r) \cap C_n^i) = 2r$ , so in this case the limit equals  $\frac{1}{2}$ .

*Case 2.* Suppose  $v_n - v \in S_n^j \setminus \text{cl}(C_n)$  for some  $j \in \{1, \dots, 4\}$ . For sufficiently small  $r$ ,  $\mu(B_\infty(v_n - v, r)) = \mu_n^S(B_\infty(v_n - v, r) \cap S_n^j) = 2\alpha r$ , so in this case the limit equals  $\frac{\alpha}{2}$ .

*Case 3.* Suppose  $v_n - v \in S_n^j \cap \text{cl}(C_n)$  for some  $j$ . For sufficiently small  $r$ ,  $\mu(B_\infty(v_n - v, r)) = \mu_n^S(B_\infty(v_n - v, r) \cap S_n^j) + \mu_n^C(B_\infty(v_n - v, r) \cap C_n) = 2\alpha r + r$ , so in this case the limit equals  $\frac{2\alpha+1}{4} < \frac{3}{4}$ .

Since in all the preceding cases the limit in question is less than or equal to 1,  $v_n$  is an  $E$ -weak mode.

Next, we show that  $v_0$  is not an  $E$ -strong mode. Fix an arbitrary  $n \in \mathbb{N}$ . Since the  $\ell_\infty$ -radius of the crossed square  $C_n \cup S_n$  is  $2^{-(n+1)}$  and the distance of  $C_n \cup S_n$  to the nearest of the two adjacent crossed squares is  $2^{-(n+2)}$ , it follows that for arbitrary  $\frac{1}{2^{n+1}} < r < \frac{3}{2^{n+2}}$ , we have

$$\mu(B_\infty(v_n, r)) = \mu_n(X) = (1 + \alpha)2^{-(n-1)} = (1 + \alpha)\frac{4}{2^{n+1}},$$

and for the same values of  $r$  we have

$$\frac{4}{2^{n+1}} < \mu(B_\infty(v_0, r)) = \mu_0^C(B_\infty(v_0, r)) < \frac{6}{2^{n+1}}.$$

Computing the ratio and using that  $\alpha = \frac{7}{8}$  yields

$$\frac{5}{4} = (1 + \alpha) \frac{4}{6} < \frac{\mu(B_\infty(v_n, r))}{\mu(B_\infty(v_0, r))} < 1 + \alpha.$$

Since  $n \in \mathbb{N}$  was arbitrary, it follows that there exists a countable sequence of intervals of values of  $r$  of the form  $\frac{1}{2^{n+1}} < r < \frac{3}{2^{n+2}}$ , such that over each interval the ratio  $\sup_{z \in v_0 - E} \mu(B_\infty(z, r)) / \mu(B_\infty(v_0, r))$  cannot be strictly less than  $\frac{5}{4}$ . Thus

$$\lim_{r \downarrow 0} \frac{\sup_{z \in v_0 - E} f(z, r)}{f(v_0, r)} \geq \frac{5}{4} > 1,$$

and so  $v_0$  is an  $E$ -weak mode, but not an  $E$ -strong mode.

By replacing  $v_0$  in the preceding analysis with  $v_m$  for an arbitrary  $m \in \mathbb{N}$  and making appropriate modifications, it follows that the set of  $E$ -weak modes coincides exactly with the set  $\{v_n\}_{n \in \mathbb{N}_0}$ , and that no element in this set is an  $E$ -strong mode. Thus,  $\mu$  has countably infinitely many weak modes, but no  $E$ -strong mode.

**Remark 4.5.** Example 4.4 yields two important observations. First, a measure may admit a weak mode but not a mode; this follows from (2.5). Second, the example demonstrates that there exists a finite measure  $\mu$  on a vector space  $X$  and some dense subset  $E$  of  $X$  such that some  $u \in \text{supp}(\mu)$  is an  $E$ -weak mode but not an  $E$ -strong mode. In addition, the measure  $\mu$  does not satisfy the uniformity condition. Therefore, the uniformity condition (2.6) cannot be removed in general. This is not surprising, given that (2.4) describes uniform decay of  $f(u - v, r)$  values over all  $v \in E$ , whereas (2.3) allows for non-uniform decay.

In this section, both Proposition 4.1 and Example 4.4 make heavy use of fundamental properties of Lebesgue measure. However, many inverse problems involve infinite-dimensional vector spaces, for which there exists no analogue of Lebesgue measure. While Lemma 4.3 provides a sufficient condition that holds for any  $X$ , and while Example 4.4 shows that this sufficient condition is sharp in full generality, it may be useful to determine alternative sufficient conditions for an  $E$ -weak mode to be  $E$ -strong, when the measure  $\mu$  and  $E$  have sufficient regularity properties, e.g. those assumed by Helin and Burger [17].

Given a  $\sigma$ -finite measure  $\mu$  that satisfies Assumption 2.1, and given an arbitrary vector  $v \in X$ , define the translated measure  $\mu_v$  by

$$(4.3) \quad \mu_v(A) = \mu(A - v),$$

for each Borel set  $A \subseteq X$ . If  $\mu_v$  is absolutely continuous with respect to  $\mu$ , i.e. if  $\mu_v \ll \mu$ , then denote the corresponding equivalence class of Radon–Nikodym derivatives by  $[\frac{d\mu_v}{d\mu}] \in L^1(\mu)$ ;  $\frac{d\mu_v}{d\mu}$  shall denote a representative of this equivalence class. Recalling that  $\mathcal{N}(x)$  denotes the set of open neighbourhoods of  $x \in X$ , let  $C(U)$  denote the set of functions that are continuous on  $U \in \mathcal{N}(x)$ . For any  $x \in X$  and  $U \in \mathcal{N}(x)$ , define the set

$$(4.4) \quad T(x) := \left\{ v \in X \mid \mu_v \ll \mu, \exists \left( U, \frac{d\mu_v}{d\mu} \right) \in \mathcal{N}(x) \times \left[ \frac{d\mu_v}{d\mu} \right] \text{ with } \frac{d\mu_v}{d\mu} \in C(U) \right\}.$$

We have the following result:

**Lemma 4.6.** *Let  $\mu$  be a  $\sigma$ -finite measure on  $X$  and  $u \in \text{supp}(\mu)$ . Suppose that  $T(u)$  is nonempty, and let  $v \in T(u)$ . Then*

$$(4.5) \quad \lim_{r \downarrow 0} \frac{f(u-v, r)}{f(u, r)} = \frac{d\mu_v}{d\mu}(u).$$

The result and the idea of its proof is similar to that of [17, Lemma 2]. In our case, however,  $X$  is not a Banach space, and  $\mu$  need not be finite. Two other key differences are that we do not assume the existence of a representative of  $[\frac{d\mu_v}{d\mu}]$  that is continuous on all of  $X$ , and that we do not assume that the continuous representative attains either its infimum or its supremum on  $U$ .

*Proof of Lemma 4.6.* For arbitrary  $(x, y, r) \in X \times X \times \mathbb{R}_{>0}$ , the notation (2.1) for the evaluation map  $f$  and the notation (4.3) for the translated measure  $\mu_x$  yield

$$f(x-y, r) = \mu(K(x-y, r)) = \mu_y(K(x, r)).$$

Next, notice that if  $\mu_y \ll \mu$ , then for any element  $\frac{d\mu_y}{d\mu}$  of the equivalence class  $[\frac{d\mu_y}{d\mu}] \in L^1(\mu)$ , we have

$$\mu(K(x-y, r)) = \int_{K(x, r)} \frac{d\mu_y}{d\mu}(z) \mu(dz) \leq \left( \sup_{z \in K(x, r)} \frac{d\mu_y}{d\mu}(z) \right) \mu(K(x, r)).$$

Dividing both sides by  $\mu(K(x, r)) = f(x, r)$  and using that  $\mu(K(x-y, r)) = f(x-y, r)$ , we obtain

$$(4.6) \quad \frac{f(x-y, r)}{f(x, r)} \leq \sup_{z \in K(x, r)} \frac{d\mu_y}{d\mu}(z).$$

Modifying the argument for a lower bound yields

$$(4.7) \quad \inf_{z \in K(x, r)} \frac{d\mu_y}{d\mu}(z) \leq \frac{f(x-y, r)}{f(x, r)} \leq \sup_{z \in K(x, r)} \frac{d\mu_y}{d\mu}(z).$$

Note that it is not necessary for either the infimum or the supremum to be attained. In addition, the infimum (supremum) may be  $-\infty$  ( $+\infty$ ).

Suppose that  $u$  and  $v$  satisfy the hypotheses. In particular, since  $u \in \text{supp}(\mu)$  it follows that  $f(u, r) > 0$  for all  $r > 0$ . Since  $v \in T(u)$ , there exists a neighbourhood  $U \in \mathcal{N}(u)$  and a representative  $\frac{d\mu_v}{d\mu}$  of  $[\frac{d\mu_v}{d\mu}] \in L^1(\mu)$  such that  $\frac{d\mu_v}{d\mu}$  is continuous on  $U$ . For the remainder of this proof, we may assume without loss of generality that the radius parameter  $r$  of  $K(u, r)$  is small enough that  $K(u, r) \subset U$ .

To establish finiteness of the supremum in (4.7) for  $x = u$  and sufficiently small values of  $r$ , we use the hypothesis that  $u$  is an  $E$ -weak mode. Suppose that the supremum is not finite for any sufficiently small value of  $r$ . Then for arbitrary  $C > 1$ , there exists a sequence  $(r_n)_{n \in \mathbb{N}}$  in the interval  $(0, r^*)$  that satisfies  $r_n \downarrow 0$  and  $\sup_{z \in K(u, r_n)} \frac{d\mu_v}{d\mu}(z) > C$  for all  $n \in \mathbb{N}$ . By continuity of  $\frac{d\mu_v}{d\mu}$ , there exists some  $M \in \mathbb{N}$  such that  $\frac{d\mu_v}{d\mu} > \frac{C}{2}$  on  $K(u, r_m)$  for all  $m \geq M$ , and thus

$$\mu(K(u-v, r_m)) = \int_{K(u, r_m)} \frac{d\mu_v}{d\mu}(z) \mu(dz) \geq \frac{C}{2} \mu(K(u, r_m)), \quad \forall m \geq M.$$

Taking the limit as  $m \rightarrow \infty$  and letting  $C > 4$  be arbitrary, we obtain a contradiction with (2.3). Thus the supremum is finite for sufficiently small  $r$ .

For the infimum in (4.7), we consider the two possible cases. In the first case, there exists a neighbourhood  $U' \in \mathcal{N}(u)$  such that  $[\frac{d\mu_v}{d\mu}|_{U' \cap U}] = [0] \in L^1(\mu)$ . Since  $\frac{d\mu_v}{d\mu}$  is continuous on  $U' \cap U$ , it follows that  $\frac{d\mu_v}{d\mu}$  vanishes uniformly on  $U' \cap U$ . Thus for all  $r$  sufficiently small such that  $K(u, r) \subset U' \cap U$ , both the infimum and supremum in (4.7) are zero. In the second case, for all  $U' \in \mathcal{N}(u)$  it holds that  $[\frac{d\mu_v}{d\mu}|_{U'}] \neq [0]$ . Therefore, by taking a decreasing sequence of subsets  $(U'_m)_{m \in \mathbb{N}} \subset \mathcal{N}(u)$  such that  $U'_m \subset U$  and  $U'_m \downarrow \{u\}$ , it follows from the continuity of  $\frac{d\mu_v}{d\mu}$  on  $U'_m \cap U$  and  $\mu$ -a.e. non-negativity of  $\frac{d\mu_v}{d\mu}$  that  $\frac{d\mu_v}{d\mu}$  is strictly positive on  $U'_m \cap U$ . Hence  $\frac{d\mu_v}{d\mu}(u) > 0$ . Using the latter observation and continuity, we may further assume that  $\frac{d\mu_v}{d\mu}$  is bounded away from zero on  $K(u, r)$  for all sufficiently small  $r$ .

Let  $(r_n)_{n \in \mathbb{N}}$  be an arbitrary decreasing sequence contained in the interval  $(0, r^*)$ , such that  $r_n \rightarrow 0$ . By the boundedness of  $K$ , by passing to a subsequence<sup>1</sup> if necessary, we may assume that  $(K(u, r_n))_{n \in \mathbb{N}}$  is a strictly decreasing sequence of sets, i.e.  $K(u, r_{n_2}) \subset K(u, r_{n_1})$  whenever  $n_1 < n_2$ . Since  $r_n \rightarrow 0$  and  $u \in K(u, r_n)$  for all  $n \in \mathbb{N}$ , it follows that  $K(u, r_n) \downarrow \{u\}$ . Combining these observations with the definitions of the supremum and infimum, it follows that  $(\sup_{z \in K(u, r_n)} \frac{d\mu_v}{d\mu}(z))_{n \in \mathbb{N}}$  and  $(\inf_{z \in K(u, r_n)} \frac{d\mu_v}{d\mu}(z))_{n \in \mathbb{N}}$  are strictly decreasing and increasing sequences respectively. In addition, the first (resp. second) sequence is bounded from below (resp. above) by  $\frac{d\mu_v}{d\mu}(u)$ . By continuity of  $\frac{d\mu_v}{d\mu}$  on  $K(u, r_n)$  for every  $n \in \mathbb{N}$ , it follows that both sequences converge to  $\frac{d\mu_v}{d\mu}(u)$ . By (4.7), the desired conclusion follows.  $\square$

**Corollary 4.7.** *Let  $\mu$  be a  $\sigma$ -finite measure on  $X$  and  $u \in \text{supp}(\mu)$ . Suppose that  $T(u)$  is nonempty, and let  $E \subset T(u)$  be nonempty. If  $u$  is an  $E$ -weak mode, then  $\sup_{v \in E} \frac{d\mu_v}{d\mu}(u) \leq 1$ .*

*Proof.* By Definition 2.3 of an  $E$ -weak mode and by Lemma 4.6,  $\frac{d\mu_v}{d\mu}(u) \leq 1$  for all  $v \in E$ . By definition of the supremum, the conclusion follows.  $\square$

The following result establishes an important sufficient condition for an  $E$ -weak mode to be  $E$ -strong:

**Proposition 4.8.** *Let  $\mu$  be a  $\sigma$ -finite measure on  $X$  and  $u \in \text{supp}(\mu)$ . Suppose that  $T(u)$  is nonempty, and let  $E \subset T(u)$  be nonempty. Suppose  $u$  is an  $E$ -weak mode. If  $\sup_{v \in E} \frac{d\mu_v}{d\mu}(u)$  is continuous at  $u$ , then  $u$  is an  $E$ -strong mode.*

*Proof.* For arbitrary  $r > 0$ ,  $v \in E \subset T(u)$  and a fixed representative  $\frac{d\mu_v}{d\mu}$ , it follows from (4.6) and the fact that suprema commute that

$$\frac{\sup_{v \in E} f(u - v, r)}{f(u, r)} \leq \sup_{v \in E} \sup_{z \in K(u, r)} \frac{d\mu_v}{d\mu}(z) = \sup_{z \in K(u, r)} \sup_{v \in E} \frac{d\mu_v}{d\mu}(z).$$

Recall that  $E \subset T(u)$  and that  $\frac{d\mu_v}{d\mu}$  is continuous at  $u$ , for all  $v \in T(u)$ . Since the supremum of a collection of functions that are lower semicontinuous at a point is

<sup>1</sup>Passage to a subsequence is not necessary if  $K$  is star-shaped with respect to the origin.

itself lower semicontinuous at the point, it follows that  $\sup_{v \in E} \frac{d\mu_v}{d\mu}$  is lower semicontinuous at  $u$ . Thus, we have

$$(4.8) \quad \sup_{v \in E} \frac{d\mu_v}{d\mu}(u) \leq \lim_{r \downarrow 0} \sup_{z \in K(u, r)} \sup_{v \in E} \frac{d\mu_v}{d\mu}(z),$$

with equality if and only if  $\sup_{v \in E} \frac{d\mu_v}{d\mu}$  is continuous at  $u$ . The conclusion then follows by the hypotheses and Corollary 4.7.  $\square$

In order to show that an  $E$ -weak mode is  $E$ -strong, it suffices to show the continuity of  $\sup_{v \in E} \frac{d\mu_v}{d\mu}$  at  $u$ . Using the fact that every lower semicontinuous function can be obtained as the limit of an increasing sequence of continuous functions, let  $(v_n)_{n \in \mathbb{N}}$  be a sequence in  $E$ . Define a sequence of functions  $(g_m)_{m \in \mathbb{N}}$  by

$$g_m(x) := \max \left\{ \frac{d\mu_{v_n}}{d\mu}(x) \mid n \in \mathbb{N}, n \leq m \right\}, \quad \forall x \in X.$$

Note that for any  $m \in \mathbb{N}$ , there is an open neighbourhood  $U_m \in \mathcal{N}(u)$  such that  $g_m$  is continuous on  $U_m$ ; simply take  $U_m := \bigcap_{n \leq m} U_{v_n}$ , where  $\frac{d\mu_{v_n}}{d\mu}$  is continuous on  $U_{v_n} \in \mathcal{N}(u)$ . Since we will take the limit as  $m \rightarrow \infty$ , we will need that  $\bigcap_{m \in \mathbb{N}} U_{v_n} \in \mathcal{N}(u)$ , i.e. there exists some  $W \in \mathcal{N}(u)$  such that for all  $n \in \mathbb{N}$ ,  $\frac{d\mu_{v_n}}{d\mu}$  is continuous on  $W$ . By choosing the sequence  $(v_n)_{n \in \mathbb{N}} \subset E \subset T(u)$  appropriately, we may assume without loss of generality that  $(g_m)_{m \in \mathbb{N}}$  converges to  $\sup_{v \in E} \frac{d\mu_v}{d\mu}$  on  $W$ .

In order for  $\sup_{v \in E} \frac{d\mu_v}{d\mu}$  to be continuous on  $W$ , we need that  $(g_m)_{m \in \mathbb{N}}$  converges uniformly on  $W$ . The standard result for guaranteeing uniform convergence is the Arzelà–Ascoli theorem. In order to apply the latter, we need that the  $(g_m)_{m \in \mathbb{N}}$  form an equicontinuous and pointwise bounded sequence on a compact Hausdorff space. In other words, there must exist a *compact* neighbourhood  $W$  of  $u$  such that  $\frac{d\mu_{v_n}}{d\mu}$  is continuous on  $W$  for all  $n \in \mathbb{N}$ . Since  $X$  is a Hausdorff topological vector space by Assumption 2.1, the existence of a point  $u$  and a compact neighbourhood  $W \in \mathcal{N}(u)$  imply that  $X$  is in fact locally compact. However, it is known that every locally compact Hausdorff topological vector space over  $\mathbb{R}$  or  $\mathbb{C}$  is finite-dimensional. Thus, for infinite-dimensional  $X$ , the approach described above fails in general.

Another approach is to avoid using the commutativity of suprema, and instead to consider the special case where one can switch the order in which the  $r$ -limit and supremum over  $v \in E$  are taken. In this case,

$$\begin{aligned} \lim_{r \downarrow 0} \frac{\sup_{v \in E} f(u - v, r)}{f(u, r)} &\leq \lim_{r \downarrow 0} \sup_{v \in E} \sup_{z \in K(u, r)} \frac{d\mu_v}{d\mu}(z) \\ &= \sup_{v \in E} \lim_{r \downarrow 0} \sup_{z \in K(u, r)} \frac{d\mu_v}{d\mu}(z) = \sup_{v \in E} \frac{d\mu_v}{d\mu}(u), \end{aligned}$$

where the last equation follows from continuity of  $\frac{d\mu_v}{d\mu}$  for all  $v \in E$ . Note that this case allows for the neighbourhoods of continuity of  $\frac{d\mu_v}{d\mu}$  to differ, which is advantageous. A sufficient condition for the commutativity of the  $r$ -limit and supremum over  $v \in E$  is that, for some  $U^* \in \mathcal{N}(u)$  and  $v^* \in E$ ,

$$(4.9) \quad \frac{d\mu_{v^*}}{d\mu}(z) = \sup_{v \in E} \frac{d\mu_v}{d\mu}(z), \quad \forall z \in U^*.$$

The lemma below describes a special case in which (4.9) is satisfied, without requiring continuity of  $\frac{d\mu_v}{d\mu}$  for any  $v$ . It also shows that (4.9) implies the uniformity condition (2.6).

**Lemma 4.9.** *Let  $\mu$  be a  $\sigma$ -finite measure on  $X$  and  $u \in \text{supp}(\mu)$ . Let  $E$  contain the origin, and suppose that for every  $v \in E$ , it holds that  $\mu_v \ll \mu$ . If there exists some  $U^* \in \mathcal{N}(u)$  such that  $\frac{d\mu_v}{d\mu} \leq 1$  on  $U^*$  for all  $v \in E$ , then  $\sup_{v \in E} \frac{d\mu_v}{d\mu}(u)$  is continuous at  $u$ . In addition, (2.6) holds.*

*Proof.* If  $v^*$  is the origin, then  $\frac{d\mu_{v^*}}{d\mu}$  is constant and equal to 1 on all of  $X$ . By the hypotheses, we have that  $\sup_{v \in E} \frac{d\mu_v}{d\mu}$  is constant, hence continuous, on  $U^*$ .

For the second statement, let  $r^* > 0$  be sufficiently small so that  $K(u, r) \subset U^*$  for all  $r < r^*$ . Since

$$\frac{d\mu_v}{d\mu}(z) \leq \sup_{v \in E} \frac{d\mu_v}{d\mu}(z) = \frac{d\mu_{v^*}}{d\mu}(z) = 1, \quad \forall z \in U^*,$$

it follows that for all  $v \in E$ ,

$$f(u - v, r) = \int_{K(u, r)} \frac{d\mu_v}{d\mu}(z) \mu(dz) \leq \mu(K(u, r)) = f(u, r), \quad \forall r < r^*.$$

Since equality is attained for  $v = v^*$  being the origin, it follows that (2.6) holds.  $\square$

In this section, we saw that when  $X$  is an infinite-dimensional Hausdorff topological vector space, the property that  $\mu$  and  $E$  are such that  $\mu$  has absolutely continuous translates  $\mu_v$  for all  $v \in E$  does not suffice for every  $E$ -weak mode to be  $E$ -strong. The significance of Lemma 4.9 is to provide an additional condition on the Radon–Nikodym derivatives such that the desired equivalence holds, and to emphasise the importance of the uniformity condition (2.6). Indeed, the key observations of this section are that the uniformity condition is a sufficient condition, and that in general it cannot be weakened; see Example 4.4 and Remark 4.5.

## 5. DISCUSSION

In this section,  $\mathbb{K}_\times := \mathbb{K} \setminus \{0\}$ , where  $\mathbb{K} := \mathbb{R}$  or  $\mathbb{C}$ .

**5.1. Arbitrary non-zero scaling factors.** As noted in Section 1, the fact that we restrict the scaling factor  $r$  in (2.1) to  $\mathbb{R}_{>0}$  even when the base field  $\mathbb{K}$  of  $X$  is  $\mathbb{C}$  implies that the preceding results apply only to the real restriction of  $X$ . However, a careful examination of the proofs indicates that the restriction to strictly positive  $r \in \mathbb{R}$  is not necessary, even when  $X$  is a real topological vector space. This is because the set  $K$  in (2.1) is a bounded neighbourhood of the origin, by Assumption 2.1. In particular,  $\rho K$  is a bounded neighbourhood of the origin for any  $\rho \in \mathbb{K}_\times$ , for both  $\mathbb{K} = \mathbb{R}$  and  $\mathbb{K} = \mathbb{C}$ . We therefore define, analogously to (2.1),

$$(5.1) \quad J(x, \rho) := x + \rho K, \quad \phi(x, \rho) := \mu(J(x, \rho)) \geq 0,$$

for  $J \in \mathcal{N}(0)$  and  $(x, \rho) \in X \times \mathbb{K}_\times$ . Let  $|\rho|$  denote the complex modulus of  $\rho \in \mathbb{C}$ . Below, we adapt the definitions from Section 2 to this setting.

**Definition 5.1.** A *strong mode* (or simply *mode*) of  $\mu$  is any  $u \in X$  such that

$$(5.2) \quad \lim_{|\rho| \rightarrow 0} \frac{\sup_{z \in X} \phi(z, \rho)}{\phi(u, \rho)} = 1.$$

An  $E$ -strong mode of  $\mu$  is a point  $u \in \text{supp}(\mu)$  such that

$$(5.3) \quad \lim_{|\rho| \rightarrow 0} \frac{\sup_{z \in u-E} \phi(z, \rho)}{\phi(u, \rho)} \leq 1.$$

An  $E$ -weak mode (or simply weak mode) of  $\mu$  is a point  $u \in \text{supp}(\mu)$  such that

$$(5.4) \quad \lim_{|\rho| \rightarrow 0} \frac{\sup_{z \in u-E} \phi(z, \rho)}{\phi(u, \rho)} \leq 1.$$

From the definitions, we obtain the same chain of implications (2.5) when we allow  $\rho \in \mathbb{K}_\times$  as when we restricted  $\rho$  to  $\mathbb{R}_{>0}$ .

The following result, analogously to Corollary 2.7, answers Helin and Burger's question for vector spaces over  $\mathbb{K}$ :

**Theorem 5.2.** *Under Assumption 2.1, if  $E$  is topologically dense in  $X$ ,  $u \in \text{supp}(\mu)$ , and the pair  $(u, E)$  satisfies*

$$\begin{aligned} \exists(v^*, r^*) \in E \times (0, 1) \text{ such that, for all } \rho \in \{z \in \mathbb{C} \mid 0 < |z| < r^*\}, \\ \phi(u - v^*, \rho) = \sup_{z \in u-E} \phi(z, \rho), \end{aligned}$$

then  $u$  is an  $E$ -weak mode if and only if  $u$  is a strong mode.

The second hypothesis is an extension of the uniformity condition (2.6) to  $\mathbb{K}_\times$ .

*Proof.* The proof of Proposition 3.4 applies for a fixed scaling. Therefore, after modifying these proofs by replacing  $f$ ,  $r$ , and small- $r$  limits with  $\phi$ ,  $\rho$ , and small  $|\rho|$ -limits respectively, we obtain that  $E$ -strong mode of  $\mu$  in the sense of (5.3) also is a strong mode in the sense of (5.2).

To prove the remaining equivalence, suppose that  $u$  is an  $E$ -weak mode, and let  $v^* \in E$  be as in the  $\mathbb{K}_\times$ -uniformity condition. Observe as in the proof of Lemma 4.3 that

$$\lim_{|\rho| \downarrow 0} \frac{\sup_{z \in u-E} \phi(z, \rho)}{\phi(u, \rho)} - \lim_{|\rho| \downarrow 0} \frac{\phi(u - v^*, \rho)}{\phi(u, \rho)} = \lim_{|\rho| \downarrow 0} \frac{\sup_{z \in u-E} \phi(z, \rho) - \phi(u - v^*, \rho)}{\phi(u, \rho)},$$

and that the right-hand side is zero by the  $\mathbb{K}_\times$ -uniformity condition. Thus if  $u$  is an  $E$ -weak mode, it is also an  $E$ -strong mode. Using (2.5) completes the proof.  $\square$

**5.2. Local modes.** In [1, Definition 2.5], one defines a point  $u$  in a Banach space  $X$  to be a local mode and local  $E$ -weak mode by allowing translations to points in a norm ball centred at  $u$  and the intersection of this ball with  $E$ .

**Definition 5.3.** Let  $V \in \mathcal{N}(0)$  be bounded. A point  $u \in X$  is a *local strong mode* of  $\mu$  with respect to  $V$  if

$$(5.5) \quad \lim_{|\rho| \downarrow 0} \frac{\sup_{z \in u-V} \phi(z, \rho)}{\phi(u, \rho)} = 1.$$

A *local  $E$ -weak mode* of  $\mu$  with respect to  $V$  is a point  $u \in \text{supp}(\mu)$  such that

$$(5.6) \quad \lim_{|\rho| \downarrow 0} \frac{\sup_{z \in u-E \cap V} \phi(z, \rho)}{\phi(u, \rho)} \leq 1.$$

For brevity, we omit defining the a local  $E$ -strong mode; note that we have the analogous chain of implications as in (2.5). The global nature of a point  $u$  that satisfies (5.2) and (5.4) is evident by writing the latter in a similar form to (5.5) and (5.6), in which case  $V = X$ .

In [1, Theorem 2.10], it was shown that for any log-concave probability measure  $\mu$  on separable Banach space  $X$ , every local strong mode is a global strong mode, and every local  $E$ -weak mode is also a global  $E$ -weak mode, for the case when  $X$  is a Banach space and  $V$  is the unit norm ball. We now consider the question of when a local  $E$ -weak mode is a local strong mode.

**Theorem 5.4.** *Suppose that Assumption 2.1 holds, and let  $V \in \mathcal{N}(0)$  be bounded. If  $E$  is topologically dense in  $V$ , and if*

$$\exists(v^*, r^*) \in E \cap V \times (0, 1) \text{ such that, for all } \rho \in \{z \in \mathbb{C} \mid 0 < |z| < r^*\},$$

$$\phi(u - v^*, \rho) = \sup_{z \in u - E \cap V} \phi(z, \rho),$$

then  $u$  is a local  $E$ -weak mode if and only if  $u$  is a local strong mode (5.5).

We sketch the proof below.

*Proof.* If, in the proof of Proposition 3.4, we replace  $X$  with  $x - V$ ,  $x - E$  with  $x - E \cap V$ , and  $f(\cdot, r)$  with  $\phi(\cdot, \rho)$ , then the density of  $E$  in  $V$  implies that

$$\sup_{z \in x - V} \phi(z, \rho) = \sup_{z \in x - E \cap V} \phi(z, \rho),$$

for all  $(x, \rho) \in X \times \mathbb{K}_\times$ . Thus, the local coincident limiting ratios condition with respect to  $V$  holds, i.e.

$$\lim_{|\rho| \downarrow 0} \frac{\sup_{z \in x - E \cap V} \phi(z, \rho)}{\phi(x, \rho)} = \lim_{|\rho| \downarrow 0} \frac{\sup_{z \in x - V} \phi(z, \rho)}{\phi(x, \rho)},$$

so local  $E$ -strong modes and local strong modes are equivalent. Using the local uniformity condition hypothesis and elementary properties of limits, we have

$$\lim_{|\rho| \downarrow 0} \frac{\sup_{z \in x - V} \phi(z, \rho)}{\phi(x, \rho)} - \lim_{|\rho| \downarrow 0} \frac{\phi(u - v^*, \rho)}{\phi(u, \rho)} = 0,$$

which completes the proof that every local  $E$ -weak mode is a local strong mode. Since the analogue of (2.5) holds for local modes, the proof is complete.  $\square$

**5.3. Examples concerning the existence and  $K$ -dependence of modes.** We conclude with some further discussion of the existence and  $K$ -dependency of modes. Recall that Example 4.4 gave an example of a measure on  $\mathbb{R}^2$  with weak modes but no strong modes; our next example is of a measure on  $\mathbb{R}$  with no modes at all:

**Example 5.5** (Probability measure without a mode). Fix parameters  $1 < a < b < \infty$  with  $b > 2$ ; write  $A := \sum_{n \in \mathbb{N}} a^{-n} < \infty$ . A good choice, which helps with the intuition, is to take  $a = 2$  and  $b = 3$  or  $4$ . Define  $\mu$  to be the probability measure with Lebesgue density (probability density function)  $\rho: \mathbb{R} \rightarrow \mathbb{R}_{\geq 0}$ ,

$$(5.7) \quad \rho(x) \equiv \frac{d\mu}{dx}(x) := \sum_{n \in \mathbb{N}} \frac{1}{2A} (b/a)^n \mathbb{I}[|x - n| < b^{-n}].$$

As illustrated in Figure 5.1,  $\rho$  is a step function with height  $\frac{1}{2A}(b/a)^n$  on each interval of length  $2b^{-n}$  centred on each natural number  $n$ , so that this interval has mass  $a^{-n}/A$ ; the density function  $\rho$  takes the value 0 outside these intervals. (The restriction that  $b > 2$  ensures that the supports of the indicator functions  $\mathbb{I}[|x - n| < b^{-n}]$  in (5.7) are pairwise disjoint.)

As can be seen from inspection,  $\rho$  is unbounded above and has no global maximum. Thus,  $\mu$  has no mode in the naïve sense of a maximiser of its density function.

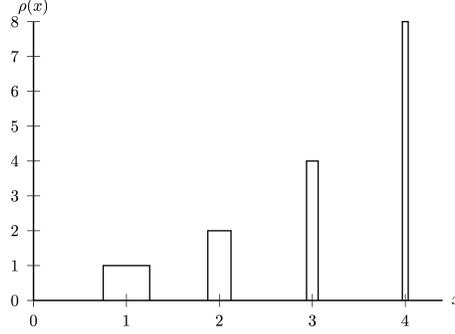


FIGURE 5.1. The probability density function  $\rho$  of a probability measure on the real line with no modes in the naïve, strong, or weak senses, given by (5.7) with parameters  $a = 2$ ,  $b = 4$ . The local uniform distribution centred on each  $n \in \mathbb{N}$  has mass  $2^{-n}$ , with  $\rho$  having magnitude  $2^{n-1}$  on an interval of length  $2^{1-2n}$ .

Furthermore, the small-balls approach does not rectify the situation. For the ball radius  $r_n = b^{-n} > 0$ , the unique maximiser of  $x \mapsto \mu((x - r_n, x + r_n))$  is  $x_n^* = n$  (with maximum value  $a^{-n}/A$ ). Clearly, the sequence  $(x_n^*)_{n \in \mathbb{N}}$  does not converge in  $\mathbb{R}$ , and indeed has no convergent subsequence, so we cannot hope that a mode will arise as an accumulation point of the sequence  $(x_n^*)_{n \in \mathbb{N}}$  of approximate modes, in the style of [9], since this sequence has no such accumulation points.

Indeed, it is possible to directly verify that  $\mu$  has no strong modes in the sense of (2.2). For any proposed strong mode  $u \in \mathbb{R}$ , let  $N := \lceil u \rceil + 1$ . Then, for  $r := r_N = b^{-N}$  and  $x_N^* = N$ ,

$$\frac{\sup_{z \in \mathbb{R}} f(z, r)}{f(u, r)} \geq \frac{f(x_N^*, r)}{f(u, r)} \geq \frac{A^{-1}a^{-N}}{2b^{-N} \cdot (2A)^{-1}(b/a)^{N-1}} = \frac{a^{-N}}{b^{-1}a^{1-N}} = \frac{b}{a} > 1.$$

Hence,  $u$  cannot be a strong mode for  $\mu$ . Similarly,  $u$  is not an  $E$ -weak mode for any dense  $E \subseteq \mathbb{R}$ .

This example can be modified to produce a smooth density  $\rho$  with the same pathology. Also, under the map  $x \mapsto \pi^{-1} \arctan(x)$ , this pathological measure is pushed forward to one with compact support in  $[0, 1]$ ; in this case, the sequence  $(x_n^*)_{n \in \mathbb{N}}$  of approximate modes would converge to 1, which is again not a mode.

We note that Example 5.5 relies essentially on the *unboundedness* of the density; in [8, Example 2.2], another example is given of a probability measure on  $\mathbb{R}$  with no strong mode, which instead relies on *discontinuity* of the density.

In light of Example 5.5, it would be interesting to determine conditions on  $\mu$  that guarantee the existence of a mode, or conditions on  $\mu$  and  $E$  that guarantee the existence of an  $E$ -weak mode.

The definitions of a strong mode in (2.2) and an  $E$ -weak mode in (2.4) require a choice of the reference neighbourhood of the origin  $K$ . The question then arises as to whether, for distinct choices of  $K, K' \in \mathcal{N}(0)$ , a mode defined in terms of  $K$  is also a mode defined in terms of  $K'$ . The following two-dimensional example indicates that this is not always the case:

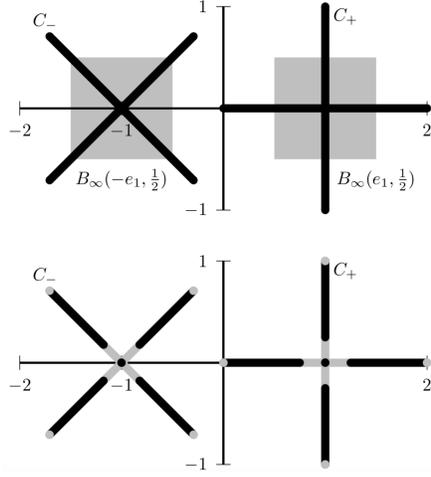


FIGURE 5.2. ABOVE. Uniform measure  $\mu$  on the 1-dimensional set shown in black has  $-e_1$  (resp.  $e_1$ ) as its unique strong mode with respect to  $K = B_\infty$  (resp.  $K = B_1$ ), as can be seen by comparing the length of  $C_- \cap K(-e_1, \frac{1}{2})$  to that of  $C_+ \cap K(e_1, \frac{1}{2})$ . BELOW. The local modes of  $\mu$ , localised using  $V = B_2(0, \frac{1}{4})$ , are the same for both choices of  $K$ , and are highlighted in black, whereas the remainder of  $\text{supp}(\mu)$  is shown in grey.

**Example 5.6** (Modes depend upon  $K$ ). Let  $K = B_p$  be the  $\ell_p$ -norm open unit ball in  $X = \mathbb{R}^2$ ; we will pay particular attention<sup>2</sup> to the cases  $p = 1$  and  $\infty$ . Let  $e_1 := (1, 0) \in \mathbb{R}^2$ ,

$$C_- := -e_1 + \{(x, y) \in \mathbb{R}^2 \mid x^2 = y^2 \leq 1\}$$

$$C_+ := e_1 + \left\{ (x, y) \in \mathbb{R}^2 \mid \begin{array}{l} \text{either } x = 0 \text{ and } |y| \leq 1, \\ \text{or } |x| \leq 1 \text{ and } y = 0 \end{array} \right\},$$

and let  $\mu(E) := \mathcal{H}^1(E \cap (C_- \uplus C_+))$ , where  $\mathcal{H}^1$  again denotes 1-dimensional Hausdorff measure, normalised to give unit measure to any line segment of unit length. Thus,  $\mu$  is uniform measure on the disjoint union  $\text{supp}(\mu) = C_- \uplus C_+$  of two right-angled crosses, namely  $C_+$  centred at  $e_1$  and aligned with the axes and  $C_-$  centred at  $-e_1$  and aligned at  $\pi/4$  to the axes;  $\mu$  has finite total mass 4.

There are two intuitively plausible candidates for modes of this measure  $\mu$ , namely  $u = \pm e_1$ . Exactly *which* is the strong mode depends whether one chooses  $K = B_1$  or  $K = B_\infty$ , as illustrated in Figure 5.2. To be more precise, for  $0 < r < \frac{1}{2}$ ,

$$\begin{aligned} \mu(B_1(-e_1, r)) &= 2\sqrt{2}r, & \mu(B_\infty(-e_1, r)) &= 4\sqrt{2}r, \\ \mu(B_1(e_1, r)) &= 4r, & \mu(B_\infty(e_1, r)) &= 4r, \end{aligned}$$

and, for any  $\tilde{u} \in X \setminus \{\pm e_1\}$  and sufficiently small  $r > 0$ ,

$$\mu(B_1(\tilde{u}, r)) \leq 2r, \quad \mu(B_\infty(\tilde{u}, r)) \leq 2\sqrt{2}r.$$

<sup>2</sup>This is a choice of convenience. In fact, this example exhibits the same behaviour for any two choices of  $0 < p \leq \infty$  on opposite sides of  $p = 2$ .

Thus,  $u \mapsto \mu(B_1(u, r))$  is globally maximised by taking  $u = e_1$ , whereas  $u \mapsto \mu(B_\infty(u, r))$  is globally maximised by taking  $u = -e_1$ . It then follows that  $u = e_1$  is a strong mode of  $\mu$  with respect to  $K = B_1$ , while  $u = -e_1$  is a strong mode of  $\mu$  with respect to  $K = B_\infty$ . These modes are unique since, because  $2\sqrt{2} < 4 < 4\sqrt{2}$ ,  $u = e_1$  is *not* a strong mode of  $\mu$  with respect to  $B_\infty$ , and  $u = -e_1$  is *not* a strong mode of  $\mu$  with respect to  $B_1$ .

The mechanisms underlying this example are twofold: first,  $\mu$  is not absolutely continuous with respect to 2-dimensional Lebesgue measure on  $X$ , and so the Lebesgue differentiation theorem does not apply here; secondly, the anisotropy of the 1-norm and  $\infty$ -norm matches the anisotropy of the  $C_-$  and  $C_+$  components of  $\text{supp}(\mu)$ . This anisotropy is particularly important: both choices of  $u = \pm e_1$  are strong modes of  $\mu$  with respect to  $K = B_2$ , the Euclidean ball.

Note also that, for sufficiently small  $V$ , e.g.  $V = B_2(0, \frac{1}{4})$ ,  $u = e_1$  becomes a *local* strong mode of  $\mu$  with respect to  $B_\infty$ , since this choice of  $V$  means that  $\mu(B_\infty(e_1, r))$  is never compared in (5.5) against any  $\mu(B_\infty(z, r))$  with  $z \in C_-$ . Similarly,  $u = -e_1$  becomes a local strong mode of  $\mu$  with respect to  $B_1$ . Note, however, that the full set of local modes is perhaps counterintuitive: it is, as illustrated in Figure 5.2,

$$(C_- \cap B_2(-e_1, 1) \cap B_2(-e_1, \frac{1}{4})^c) \uplus (C_+ \cap B_2(e_1, 1) \cap B_2(e_1, \frac{1}{4})^c) \uplus \{\pm e_1\},$$

namely those points of  $\text{supp}(\mu)$  that are not at the extremes of the ‘‘arms’’ of the crosses, nor too close to the isolated local modes at  $\pm e_1$ .

Example 5.6 indicates the importance of the choice of the sets  $K$  and  $J$  in the definitions (2.1) and (5.1) of  $f$  and  $\phi$  respectively. In the case when  $(X, \|\cdot\|)$  is a normed vector space, or when  $(X, d)$  is a metric vector space, then it may be natural to choose  $K$  to be the the unit open  $\|\cdot\|$ -norm or  $d$ -metric ball, since these sets are native to the structure of the vector space. If there is no unambiguous choice, then it may be useful to determine whether there exist conditions for which a mode defined in terms of  $K$  is also a mode in terms of  $K'$ .

#### ACKNOWLEDGEMENTS

We thank Gerd Wachsmuth, in particular for suggesting a simpler proof of a previous version of Theorem 2.5 and for pointing out an error in a result concerning weak and  $E$ -strong modes in a previous version of this manuscript. We also thank Sergios Agapiou and Tapio Helin for their feedback on an early version of this manuscript, and the two anonymous peer reviewers for their helpful comments.

The authors acknowledge support provided by the Freie Universität Berlin within the Excellence Initiative of the German Research Foundation (DFG); by DFG grant CRC 1114 ‘‘Scaling Cascades in Complex Systems’’; and by the National Science Foundation (NSF) under grant DMS-1127914 to the Statistical and Applied Mathematical Sciences Institute (SAMSI) and SAMSI’s QMC Working Group II ‘‘Probabilistic Numerics’’. Any opinions, findings, and conclusions or recommendations expressed in this article are those of the authors and do not necessarily reflect the views of the above-named funding agencies and institutions.

#### REFERENCES

1. S. Agapiou, M. Burger, M. Dashti, and T. Helin, *Sparsity-promoting and edge-preserving maximum a posteriori estimators in non-parametric Bayesian inverse problems*, Inverse Problems **34** (2018), no. 4, 045002, 37.

2. A. Alexanderian, N. Petra, G. Stadler, and O. Ghattas, *A fast and scalable method for  $A$ -optimal design of experiments for infinite-dimensional Bayesian nonlinear inverse problems*, SIAM J. Sci. Comput. **38** (2016), no. 1, A243–A272.
3. V. I. Bogachev, *Differentiable Measures and the Malliavin Calculus*, Mathematical Surveys and Monographs, vol. 164, American Mathematical Society, Providence, RI, 2010.
4. T. Bui-Thanh and Q. P. Nguyen, *FEM-based discretization-invariant MCMC methods for PDE-constrained Bayesian inverse problems*, Inv. Probl. Imaging **10** (2016), no. 4, 943–975.
5. M. Burger and F. Lucka, *Maximum a posteriori estimates in linear inverse problems with log-concave priors are proper Bayes estimators*, Inv. Probl. **30** (2014), no. 11, 114004, 21.
6. L. Calatroni, J. C. De Los Reyes, and C.-B. Schönlieb, *Infimal convolution of data discrepancies for mixed noise removal*, SIAM J. Imaging Sci. **10** (2017), no. 3, 1196–1233.
7. D. Calvetti, H. Hakula, S. Pursiainen, and E. Somersalo, *Conditionally Gaussian hypermodels for cerebral source localization*, SIAM J. Imaging Sci. **2** (2009), no. 3, 879–909.
8. C. Clason, T. Helin, R. Kretschmann, and P. Piironen, *Generalized modes in Bayesian inverse problems*, 2018, arXiv:1806.00519.
9. M. Dashti, K. J. H. Law, A. M. Stuart, and J. Voss, *MAP estimators and their consistency in Bayesian nonparametric inverse problems*, Inv. Probl. **29** (2013), no. 9, 095017, 27.
10. A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, second ed., Applications of Mathematics (New York), vol. 38, Springer-Verlag, New York, 1998.
11. M. M. Dunlop and A. M. Stuart, *MAP estimators for piecewise continuous inversion*, Inv. Probl. **32** (2016), no. 10, 105003, 50.
12. W. E, W. Ren, and E. Vanden-Eijnden, *Minimum action method for the study of rare events*, Comm. Pure Appl. Math. **57** (2004), no. 5, 637–656.
13. M. I. Freidlin and A. D. Wentzell, *Random Perturbations of Dynamical Systems*, second ed., Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 260, Springer-Verlag, New York, 1998, Translated from the 1979 Russian original by Joseph Szücs.
14. K. Hämäläinen, A. Kallonen, V. Kolehmainen, M. Lassas, K. Niinimäki, and S. Siltanen, *Sparse tomography*, SIAM J. Sci. Comput. **35** (2013), no. 3, B644–B665.
15. L. Harhanen, N. Hyvönen, H. Majander, and S. Staboulis, *Edge-enhancing reconstruction algorithm for three-dimensional electrical impedance tomography*, SIAM J. Sci. Comput. **37** (2015), no. 1, B60–B78.
16. P. Héas, F. Lavancier, and S. Kadri-Harouna, *Self-similar prior and wavelet bases for hidden incompressible turbulent motion*, SIAM J. Imaging Sci. **7** (2014), no. 2, 1171–1209.
17. T. Helin and M. Burger, *Maximum a posteriori probability estimates in infinite-dimensional Bayesian inverse problems*, Inv. Probl. **31** (2015), no. 8, 085009, 22.
18. T. Helin and M. Lassas, *Hierarchical models in statistical inverse problems and the Mumford–Shah functional*, Inv. Probl. **27** (2011), no. 1, 015008, 32.
19. N. Hyvönen and M. Leinonen, *Stochastic Galerkin finite element method with local conductivity basis for electrical impedance tomography*, SIAM/ASA J. Uncertain. Quantif. **3** (2015), no. 1, 998–1019.
20. A. Krol, S. Li, L. Shen, and Y. Xu, *Preconditioned alternating projection algorithms for maximum a posteriori ECT reconstruction*, Inv. Probl. **28** (2012), no. 11, 115005, 34.
21. J. Liu and X. Zheng, *A block nonlocal TV method for image restoration*, SIAM J. Imaging Sci. **10** (2017), no. 2, 920–941.
22. Y. Lu, A. Stuart, and H. Weber, *Gaussian approximations for transition paths in Brownian dynamics*, SIAM J. Math. Anal. **49** (2017), no. 4, 3005–3047.
23. L. Narici and E. Beckenstein, *Topological Vector Spaces*, second ed., Pure and Applied Mathematics (Boca Raton), vol. 296, CRC Press, Boca Raton, FL, 2011.
24. M. Pereyra, *Maximum-a-posteriori estimation with Bayesian confidence regions*, SIAM J. Imaging Sci. **10** (2017), no. 1, 285–302.
25. Elias M. Stein and Rami Shakarchi, *Real analysis*, Princeton Lectures in Analysis, vol. 3, Princeton University Press, Princeton, NJ, 2005, Measure theory, integration, and Hilbert spaces.
26. A. M. Stuart, *Inverse problems: a Bayesian perspective*, Acta Numer. **19** (2010), 451–559.
27. G. Vaksman, M. Zibulevsky, and M. Elad, *Patch ordering as a regularization for inverse problems in image processing*, SIAM J. Imaging Sci. **9** (2016), no. 1, 287–319.

28. Z. Yao, Z. Hu, and J. Li, *A TV-Gaussian prior for infinite-dimensional Bayesian inverse problems and its numerical implementations*, *Inv. Probl.* **32** (2016), no. 7, 075006, 19.

HAN CHENG LIE, INSTITUT FÜR MATHEMATIK, FREIE UNIVERSITÄT BERLIN, 14195 BERLIN, GERMANY

*Current address:* Zuse Institute Berlin, Takustr. 7, 14195 Berlin, Germany

*E-mail address:* [hlie@math.fu-berlin.de](mailto:hlie@math.fu-berlin.de)

T. J. SULLIVAN, FREIE UNIVERSITÄT BERLIN AND ZUSE INSTITUTES BERLIN, TAKUSTR. 7, 14195 BERLIN, GERMANY

*E-mail address:* [sullivan@zib.de](mailto:sullivan@zib.de)