

Constructive Disintegration and Conditional Modes

Nathaël Da Costa* Marvin Pförtner* Jon Cockayne†

August 4, 2025

Conditioning, the central operation in Bayesian statistics, is formalised by the notion of disintegration of measures. However, due to the implicit nature of their definition, constructing disintegrations is often difficult. A folklore result in machine learning conflates the construction of a disintegration with the restriction of probability density functions onto the subset of events that are consistent with a given observation. We provide a comprehensive set of mathematical tools which can be used to construct disintegrations and apply these to find densities of disintegrations on differentiable manifolds. Using our results, we provide a disturbingly simple example in which the restricted density and the disintegration density drastically disagree. Motivated by applications in approximate Bayesian inference and Bayesian inverse problems, we further study the modes of disintegrations. We show that the recently introduced notion of a “conditional mode” does not coincide in general with the modes of the conditional measure obtained through disintegration, but rather the modes of the restricted measure. We also discuss the implications of the discrepancy between the two measures in practice, advocating for the utility of both approaches depending on the modelling context.

Contents

1. Introduction	2
1.1. Related Work	6
1.2. Contributions	7
1.3. Structure of the Paper	7
2. Constructive Disintegration	7
2.1. Disintegration Building Blocks	10

*Tübingen AI Center, University of Tübingen. Equal contribution.

†University of Southampton

2.2. Construction of Disintegrations on Manifolds	14
3. Modes of Disintegrations	18
3.1. Restricted Modes Disagree with Modes of Disintegrations	23
3.2. Failure Modes in Normed Spaces	25
4. Discussion	26
References	28
A. Differential Geometry Background	31
A.1. Riemannian Geometry	32
B. Riemannian Volume of Small Balls	34
C. Modes Lemmas	36

1. Introduction

The construction of conditional distributions is central to statistics and probabilistic machine learning. On discrete event spaces, these can often be constructed directly, by applying the definition of conditional probability. However, this construction breaks down in case the event to be conditioned on has probability zero, which is virtually always the case on continuous event spaces. In this case, the construction is more involved, and many texts focus solely on the setting where a Lebesgue density is available or, more generally, where the conditional is dominated by, and thus has a density with-respect-to, the base measure. However, recent works have generated a profusion of examples in which this familiar construction is no longer valid, primarily due to conditionals being defined on a *submanifold* of the support of the measure. Some examples include:

- **Bayesian deep learning** in which reparameterisation invariances often lead to complex multimodalities in posterior measures (e.g. Wiese et al. [2023], Laurent et al. [2024]), potentially mitigated by restricting a prior to a submanifold to limit these invariances,
- **Directional statistics** [Mardia and Jupp, 1999] where learning problems are defined on manifolds such as the Stiefel manifold of orthonormal frames, a submanifold of $\mathbb{R}^{d \times k}$,
- **Probabilistic numerics** [Cockayne et al., 2019] which utilises noise-free observations of an unknown, resulting in posteriors concentrated on submanifolds.

There is therefore a growing need to be able to characterise conditional distributions on a more intrinsic, measure-theoretic level than by using densities.

In modern probability theory, conditioning is formalised as *disintegration*. A disintegration

of a measure μ on $\mathbb{X}$ along a map $h : \mathbb{X} \rightarrow \mathbb{Y}$ is defined as the almost-surely¹ unique family of distributions $\{\mu^h(\cdot | y)\}_{y \in \mathbb{Y}}$ that satisfies the law of total probability. Existence and almost-sure uniqueness of disintegrations is provided by *disintegration theorems* which have been available under mild conditions since at least Dellacherie and Meyer [1978]; see Chang and Pollard [1997] for a more recent exposition. However proofs are typically abstract and, crucially, non-constructive outside of conjugate settings, such as with Gaussian μ and linear h . This paper provides, to our knowledge, the first general set of results for constructing disintegrations.

Transferring intuition from the density setting, a “folklore” result in machine learning conflates disintegration of a distribution with restriction of the distribution to the submanifold on which the conditional is supported. For examples of this see e.g. [Bishop, 2006, Figure 2.18], which describes the von Mises distribution as being obtained by “considering a two-dimensional Gaussian distribution [...] and conditioning on the unit circle”. In fact this is a *restriction* rather than a conditional, and our novel results highlight that these can differ dramatically. In Figure 1 we provide a (disturbingly) simple example related to this distribution in which restriction and disintegration result in radically different distributions. This example takes an ambient standard Gaussian measure on $\mathbb{R}^2$ (Figure 1a) and contrasts restricting (Figure 1c) and disintegrating (Figure 1d) the measure on successively larger ellipses; see Example 2.1 for more detail. The colour maps on contours in Figures 1c and 1d depict appropriate densities for the restriction or disintegration, respectively. Notably, for smaller ellipses, the mass concentrates at different poles between the two approaches.

A special case of our characterisation result on $\mathbb{R}^d$ is reproduced below:

Theorem 1.1. *Let μ be a Borel probability measure on $\mathbb{R}^d$ that has a density $\frac{d\mu}{d\lambda}$ w.r.t. the Lebesgue measure λ on $\mathbb{R}^d$. Let $h : \mathbb{R}^d \rightarrow \mathbb{R}^n$ such that μ -almost all points $x \in \mathbb{R}^d$ are C^1 regular (see Definition A.8). Then there exists a version of the disintegration μ^h with probability densities²*

$$\frac{d\mu^h(\cdot | y)}{d\lambda_{h^{-1}(\{y\})}}(x) \propto \frac{d\mu}{d\lambda}(x) \cdot \left| \det \left(Dh(x)|_{\ker(Dh(x))^\perp} \right) \right|^{-1}$$

for all C^1 regular values $y \in \mathbb{R}^n$, where $\lambda_{h^{-1}(\{y\})}$ is the restricted Lebesgue measure to the fiber $h^{-1}(\{y\})$ (see Definition 2.13).

Theorem 1.1 shows that the disintegration density differs from the restriction density by a corrective factor depending on the Jacobian of h . See Theorem 2.14 for further details. In the case where h is linear, this Jacobian is constant, and we hence recover the classical proportionality formula for densities $d\mu^h(\cdot | y)/d\lambda_{h^{-1}(\{y\})} \propto d\mu/d\lambda$.

To demonstrate the importance of our results we focus on the problem of mode computation, i.e., computation of maximum a-posteriori (MAP) points in Bayesian computation.

¹Specifically, $h_*\mu$ -almost surely; see Definition 2.2 and Theorem 2.4.

²The densities are with respect to the restriction $\lambda_{h^{-1}(\{y\})}$ of the Lebesgue measure to the observation fiber $h^{-1}(\{y\})$ (see Definition 2.13).

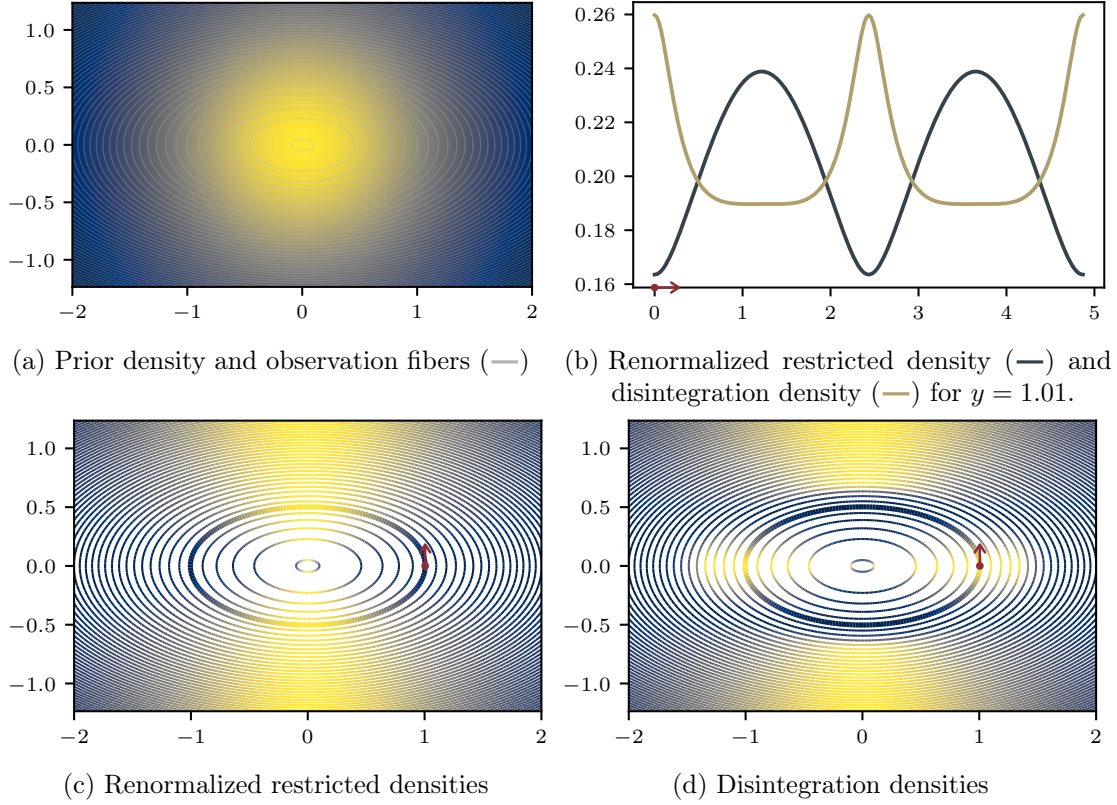


Figure 1: Conditioning a standard Gaussian prior μ on observations made through the quadratic observation operator $h(x) = \frac{x_1^2}{a^2} + \frac{x_2^2}{b^2}$ with $a = 1$ and $b = \frac{1}{2}$. In machine learning folklore, the corresponding conditional distribution is constructed by first restricting the prior’s density $\frac{d\mu}{d\lambda}$ to the observation fiber $h^{-1}(\{y\})$, which is then renormalized w.r.t. the canonical volume measure on the fiber. However, in general, the law of total probability fails to hold for this restricted measure. Our main result, Theorem 2.14, shows that an additional corrective factor of $\|\nabla h(x)\|_2^{-1}$ needs to be multiplied into this restricted density to obtain the density of the true conditional distribution, which is defined implicitly through the notion of a *disintegration*. Intuitively speaking, this is due to the fact that, for the law of total probability to hold, the predictive probability mass has to be distributed between fewer fibers where the fibers lie less dense, which is the case if $\|\nabla h(x)\|_2$ is small (see plots above). The plots show a stark disagreement between the restricted and disintegration densities. Notably the modes of the measures differ maximally. See Example 2.1 for details.

MAP estimation summarises a distribution with a single, representative estimate that can be interpreted informally as a “point of maximal probability”. This is particularly significant in Bayesian inference. In many practical settings sampling from a full posterior is computationally expensive and a point estimate is sufficient for decision-making or

further modelling. MAP estimates are also of central use in many posterior approximation settings, where an approximation to the posterior is often constructed to have the same mode as the true posterior.

Modes are traditionally defined as maximisers of a Lebesgue density, or more generally, as centers that supremise a particular ratio of open metric balls compared to other candidate centers as the radius of the balls tends to zero. The aforementioned difficulties in characterising disintegrations have led some works to propose that MAP points be computed by constraining this maximisation problem to the submanifold on which the conditional is supported—in other words, computing modal points of the restricted measure rather than the disintegrated measure—often resulting in a computationally tractable procedure. We refer to such points as *restricted modes*, and our work highlights that restricted modes can differ drastically from modes of a disintegration. To give some specific examples of this practice:

- Chen et al. [2024, Definition 3] defines a *conditional mode* precisely as above to construct samplers for Gaussian processes conditioned on a particular class of nonlinear, noise-free observations. In fact, this computes a *restricted mode*, potentially meaning that the posterior approximation techniques they propose approximate the restriction rather than the conditional or, in the worst case, do not approximate either distribution.
- Cinquin et al. [2024] propose a Laplace approximation for Bayesian neural networks which computes the MAP estimate of the network weights under a function-space prior that is restricted to the set of functions representable by the neural network.
- Tronarp et al. [2021, Section 3] defines the MAP estimator of a probabilistic ODE solver again as above, resulting in similar potential pitfalls to Chen et al. [2024].

A special case of our novel result Theorem 3.5, as it applies on $\mathbb{R}^d$, is reproduced below, to highlight that modes of a disintegration meaningfully differ from restricted modes in a way that depends on the observation operator h .

Theorem 1.2. *Let μ be a Borel probability measure on $\mathbb{R}^d$ that has a continuous density $\frac{d\mu}{d\lambda}$ w.r.t. the Lebesgue measure λ on $\mathbb{R}^d$. Let $h: \mathbb{R}^d \rightarrow \mathbb{R}^n$ such that μ -almost all points $x \in \mathbb{R}^d$ are C^1 regular (see Definition A.8). Then there is a version of the disintegration μ^h whose weak modes are given by*

$$\begin{aligned} \arg \min_{x \in \mathbb{R}^d} \quad & -\log \frac{d\mu}{d\lambda}(x) + \log \left| \det Dh(x)|_{\ker(Dh(x))^\perp} \right| \\ \text{s.t.} \quad & h(x) = y \end{aligned}$$

for all C^1 regular values $y \in \mathbb{R}^n$.

The proof can be found in Section 3. Restricted modes are obtained by dropping the log det term in Theorem 1.2, and so only coincide with modes of the disintegration in settings where this term is constant, such as for linear h .

Figure 1b graphically illustrates the severe impact this can have by contrasting densities on a particular ellipse in the wider setting of Figure 1. Disintegration and restriction show almost opposite modes in this setting.

1.1. Related Work

Disintegrations Disintegrations have received limited attention in the literature. Tjur [1975] and Tjur [1980, Chapter 9] each developed disintegrations (also referred to as decompositions) and studied their existence, almost-sure uniqueness and continuity properties under an assumption of a continuous observation operator h and completely regular $\mathbb{X}, \mathbb{Y}$. Dellacherie and Meyer [1978] provided existence results under weaker assumptions (only requiring h to be measurable, but with a separability assumption on $\mathbb{Y}$). Chang and Pollard [1997] advocate for disintegrations as the prototype for rigorous conditional distributions, rather than relying on intuition derived from taking limits of conditionals derived in more “well-behaved” settings.

Diaconis et al. [2013] obtain a related formula to our construction of disintegrations, in the special case of Euclidean spaces. Instead of constructing disintegrations from the ground up, they apply the co-area formula. Other recent works include Possobon and Rodrigues [2022], which studies the regularity of the disintegration map as a function of the observation. Cockayne et al. [2019] propose a sampling routine which, under some assumptions, approximately samples a disintegration. However these assumptions are difficult to verify.

Modes Recent years have seen a surge in research into the theoretical definition of modal points of distributions, enabling a departure from the familiar notion of a MAP point as the maximiser of a Lebesgue density. The modern, more general view was initially conceived by Dashti et al. [2013], who proposed to seek points that maximise the limit of a ratio of small ball probabilities with different candidate centers, tractably realised through the connection to an *Onsager-Machlup* functional. They studied the properties of MAP points in separable Hilbert spaces for distributions dominated by a Gaussian reference measure. The MAP estimator of Dashti et al. [2013] was termed a *strong MAP estimator* by Helin and Burger [2015], who introduced the idea of a *weak MAP estimator*, weakening the definition by comparing balls centered at a candidate MAP point to shifted balls, where shifts are limited to a topologically dense subset. Lie and Sullivan [2018] showed that these two definitions are in fact equivalent under an additional uniformity condition on the measure.

In more recent works, Clason et al. [2019] introduced the idea of a *generalised mode* that studies ball ratios in a more general asymptotic sense, studying the limiting ball ratio along sequences of points approaching candidate MAP points. This prompted Klebanov et al. [2025] to note the similarity of the three existing mode definitions and generate and study a family of feasible alternative definitions by constructing a grammar from mathematical symbols in the definitions. Of the 282 “grammatically correct” definitions

generated, ten were deemed meaningful in terms of satisfying some natural properties to expect of such definitions. Ayanbayev et al. [2022] studies Γ -convergence of global weak modes and demonstrates some of the difficulties extending this approach to strong modes. Lambley [2023] studies strong modes of Bayesian posterior measures in Banach spaces, under Gaussian priors, and demonstrates their equivalence with other notions of modes. Specifically studying MAP points in the setting of this paper (i.e. when the support of the conditional measure has no prior mass) has received much less attention. As discussed above both Chen et al. [2024] and Tronarp et al. [2021] construct what we call *restricted modes* in pursuit of the MAP of a disintegration. Chen et al. [2021] and Cinquin et al. [2024] adopt a similar approach.

1.2. Contributions

The main contributions of our work are as follows:

- We prove several intermediate important properties of disintegrations, leading to a formula for the density of a disintegration on a Riemannian manifold $\mathbb{X}$, with respect to the volume measure on the fibers $h^{-1}(\{y\})$.
- We characterise modes of disintegrations and of (Riemannian) restricted measures, highlighting their disagreement in general.
- We illustrate numerically the potentially large extent of disagreement between the two notions of modes on a simple toy problem, that is nevertheless representative of contemporary applications.

1.3. Structure of the Paper

The rest of the paper proceeds as follows. In Section 2 we introduce disintegrations and prove several intermediate results leading to our constructive definition of disintegrations on Riemannian manifolds in Section 2.2. Section 3 focuses on how these ideas pertain to the computation of modes of disintegrations, with Section 3.1 proving disagreement of restricted modes and disintegration modes and Section 3.2 providing a more explicit characterisation of failure modes in $\mathbb{R}^d$. We conclude with some discussion on the implications of these results in Section 4.

2. Constructive Disintegration

We begin by introducing transition and Markov kernels, which are fundamental to the definition of a disintegration. Let $(\mathbb{X}, \mathcal{A}_{\mathbb{X}})$ and $(\mathbb{Y}, \mathcal{A}_{\mathbb{Y}})$ be measurable spaces.

Definition 2.1 (Transition Kernel). A function $\kappa(\cdot \mid \cdot): \mathcal{A}_{\mathbb{X}} \times \mathbb{Y} \rightarrow [0, \infty]$ is called a *transition kernel* from $\mathbb{Y}$ to $\mathbb{X}$ if

- (i) $\kappa(\cdot \mid y)$ is a measure on $(\mathbb{X}, \mathcal{A}_{\mathbb{X}})$ for all $y \in \mathbb{Y}$, and
- (ii) $\kappa(X \mid \cdot)$ is $\mathcal{A}_{\mathbb{Y}}\text{-}\mathcal{B}([0, \infty])$ -measurable for all $X \in \mathcal{A}_{\mathbb{X}}$.

If $\kappa(\cdot \mid y)$ is a probability measure for all $y \in \mathbb{Y}$, then κ is called a *Markov kernel*.

We extend measure-theoretic operations like pushforwards $f_{\star}\kappa$, integrals $\kappa(f, y)$ and $f\kappa$, as well as Radon-Nikodym derivatives $\frac{d\kappa}{d\nu}$ to transition kernels.

Next we introduce disintegrations as Markov kernels satisfying certain regularity properties. To this end, let μ be a measure on $(\mathbb{X}, \mathcal{A}_{\mathbb{X}})$ and $h: \mathbb{X} \rightarrow \mathbb{Y}$ a $\mathcal{A}_{\mathbb{X}}\text{-}\mathcal{A}_{\mathbb{Y}}$ -measurable function.

Definition 2.2 (Disintegration). A *disintegration* of μ with respect to h is a transition kernel $\mu^h(\cdot \mid \cdot)$ from $\mathbb{Y}$ to $\mathbb{X}$ such that

- (i) $\mu^h(\mathbb{X} \setminus h^{-1}(\{y\}) \mid y) = 0$,
- (ii) $\mu(X) = \int_{\mathbb{Y}} \mu^h(X \mid y) h_{\star}\mu(dy)$ for any $X \in \mathcal{A}_{\mathbb{X}}$.

If μ is a probability measure, then the elements of the disintegration are sometimes referred to as *regular conditional probability measures*. In this case, (ii) is equivalent to the law of total probability when conditioning on events of nonzero probability.

Remark 2.1. By the monotone convergence theorem, Definition 2.2(ii) is equivalent to $\mu(f) = \int_{\mathbb{Y}} \mu^h(f \mid y) h_{\star}\mu(dy)$ for any non-negative measurable f .

For the remainder of Section 2, we will always work under the following set of assumptions:

Assumption 2.3.

- (i) $\mathbb{X}$ is a metrizable topological space equipped with its Borel σ -algebra, i.e. $\mathcal{A}_{\mathbb{X}} = \mathcal{B}(\mathbb{X})$.
- (ii) $\mathcal{A}_{\mathbb{Y}}$ is countably generated and contains all singleton sets $\{y\}$.
- (iii) μ is a σ -finite Radon measure on $\mathbb{X}$ and its pushforward $h_{\star}\mu$ is also σ -finite.

Remark 2.2. The assumption that $h_{\star}\mu$ is a σ -finite measure is non-trivial; for example the projection of the Lebesgue measure on $\mathbb{R}^2$ onto a coordinate axis is not σ -finite.

Under Assumption 2.3, one can prove existence and uniqueness (up to $h_{\star}\mu$ null sets) of the disintegration. As discussed in Section 1, various forms of this result exist under different and potentially more general assumptions. The version we introduce is from Chang and Pollard [1997].

Theorem 2.4 (Disintegration Theorem; see Chang and Pollard, 1997, Theorem 1). *There exists a $h_{\star}\mu$ -almost everywhere uniquely defined disintegration μ^h , i.e., the function $y \mapsto \mu^h(\cdot \mid y)$ is $h_{\star}\mu$ -almost everywhere uniquely determined.*

Remark 2.3. Chang and Pollard [1997] allow for disintegrating μ with respect to a different measure than $h_\star\mu$. For simplicity and clarity, we focus in this paper exclusively on the standard case of disintegrating μ with respect to its pushforward, though our results can be extended to the more general case.

When thinking of conditioning, most practitioners will first think of Bayes' rule. To argue that disintegrations provide a rigorous way of conditioning measures more generally, it is crucial to show that, under some assumptions, Bayes' theorem can be recovered from Definition 2.2. This is the focus of our first result; a statement of Bayes' theorem purely in terms of disintegrations.

In the next theorem we think of $\mathbb{X} = \mathbb{W} \times \mathbb{Z}$, where $\mathbb{W}$ should be thought as the parameter space, $\mathbb{Z}$ the data space, and μ the joint (probability) measure over $\mathbb{W} \times \mathbb{Z}$. Hence $\pi_{\mathbb{W}\star}\mu$ is the prior, $\pi_{\mathbb{W}\star}(\mu^{\pi_{\mathbb{Z}}}(\cdot | z))$ is the posterior given data $z \in \mathbb{Z}$, and $\pi_{\mathbb{Z}\star}(\mu^{\pi_{\mathbb{W}}}(\cdot | w))$ is the likelihood at parameter $w \in \mathbb{W}$.

The above exposition makes it clear that we must disintegrate μ with respect to both the coordinate projections $\pi_{\mathbb{Z}}$ and $\pi_{\mathbb{W}}$. Hence, we require $\mathbb{W}$ and $\mathbb{Z}$ to each serve the role of both $\mathbb{X}$ and $\mathbb{Y}$ in Definition 2.2, and hence to satisfy the assumptions on both spaces from (Assumption 2.3). We phrase these succinctly in the following theorem:

Theorem 2.5 (Bayes' Theorem). *Suppose that both $\mathbb{W}$ and $\mathbb{Z}$ are metrizable separable topological spaces equipped with their Borel σ -algebra, that $\mathbb{X} = \mathbb{W} \times \mathbb{Z}$, and that μ is a finite Radon measure on $\mathbb{X}$. Further suppose that there is a σ -finite measure ν on $\mathbb{Z}$ such that $\pi_{\mathbb{Z}\star}(\mu^{\pi_{\mathbb{W}}}(\cdot | w)) \ll \nu$ for $\pi_{\mathbb{W}\star}\mu$ -almost all w . Then*

$$\frac{d(\pi_{\mathbb{W}\star}(\mu^{\pi_{\mathbb{Z}}}(\cdot | z)))}{d(\pi_{\mathbb{W}\star}\mu)}(w) = \frac{d(\pi_{\mathbb{Z}\star}(\mu^{\pi_{\mathbb{W}}}(\cdot | w)))/d\nu(z)}{d(\pi_{\mathbb{Z}\star}\mu)/d\nu(z)}$$

where the right-hand-side is well-defined for $\pi_{\mathbb{Z}\star}\mu$ -almost every z .

Proof. We will show

$$\mu^{\pi_{\mathbb{Z}}}(\mathrm{d}z' \times \mathrm{d}w' | z) = \frac{d(\pi_{\mathbb{Z}\star}(\mu^{\pi_{\mathbb{W}}}(\cdot | w')))/d\nu(z)}{d(\pi_{\mathbb{Z}\star}\mu)/d\nu(z)} \pi_{\mathbb{W}\star}\mu(\mathrm{d}w') \delta_z(\mathrm{d}z')$$

where δ_z is the Dirac measure at z . Clearly, the right-hand-side is supported on $\pi_{\mathbb{Z}}^{-1}(\{z\}) = \{z\} \times \mathbb{W}$, so it satisfies Definition 2.2(i). Now for $Z \subset \mathbb{Z}$ and $W \subset \mathbb{W}$ measurable, define

$$\tilde{Z} := Z \cap \{z \in \mathbb{Z} : d(\pi_{\mathbb{Z}\star}\mu)/d\nu(z) \neq 0\}.$$

Now ν is absolutely continuous with respect to $\pi_{\mathbb{Z}\star}\mu$ when restricted to $\tilde{Z}$ with

$$\frac{d\nu}{d(\pi_{\mathbb{Z}\star}\mu)}(z) = \left(\frac{d(\pi_{\mathbb{Z}\star}\mu)}{d\nu}(z) \right)^{-1}.$$

Then

$$\int_{\mathbb{Z}} \left(\int_W \frac{d(\pi_{\mathbb{Z}\star}(\mu^{\pi_{\mathbb{W}}}(\cdot | w)))/d\nu(z)}{d(\pi_{\mathbb{Z}\star}\mu)/d\nu(z)} \pi_{\mathbb{W}\star}\mu(\mathrm{d}w) \right) \cdot \delta_z(Z) \pi_{\mathbb{Z}\star}\mu(\mathrm{d}z)$$

$$\begin{aligned}
&= \int_{\tilde{Z}} \int_W \frac{d(\pi_{\mathbb{Z}\star}(\mu^{\pi_{\mathbb{W}}}(\cdot | w))) / d\nu(z)}{d(\pi_{\mathbb{Z}\star}\mu) / d\nu(z)} \pi_{\mathbb{W}\star}\mu(dw) \pi_{\mathbb{Z}\star}\mu(dz) \\
&= \int_W \int_{\tilde{Z}} \frac{d(\pi_{\mathbb{Z}\star}(\mu^{\pi_{\mathbb{W}}}(\cdot | w))) / d\nu(z)}{d(\pi_{\mathbb{Z}\star}\mu) / d\nu(z)} \pi_{\mathbb{Z}\star}\mu(dz) \pi_{\mathbb{W}\star}\mu(dw) \\
&= \int_W \int_{\tilde{Z}} \frac{d(\pi_{\mathbb{Z}\star}(\mu^{\pi_{\mathbb{W}}}(\cdot | w)))}{d\nu} (z) \frac{d\nu}{d(\pi_{\mathbb{Z}\star}\mu)} (z) \pi_{\mathbb{Z}\star}\mu(dz) \pi_{\mathbb{W}\star}\mu(dw) \\
&= \int_W \int_{\tilde{Z}} \frac{d(\pi_{\mathbb{Z}\star}(\mu^{\pi_{\mathbb{W}}}(\cdot | w)))}{d\nu} (z) \nu(dz) \pi_{\mathbb{W}\star}\mu(dw) \\
&= \int_W \pi_{\mathbb{Z}\star}(\mu^{\pi_{\mathbb{W}}}(\cdot | w))(\tilde{Z}) \pi_{\mathbb{W}\star}\mu(dw) \\
&= \int_W \mu^{\pi_{\mathbb{W}}}(\mathbb{W} \times \tilde{Z} | w) \pi_{\mathbb{W}\star}\mu(dw) \\
&= \int_{\mathbb{W}} \mu^{\pi_{\mathbb{W}}}(W \times \tilde{Z} | w) \pi_{\mathbb{W}\star}\mu(dw) \\
&= \mu(W \times \tilde{Z}) \\
&= \mu(W \times Z),
\end{aligned}$$

where the first and last equality hold because $\pi_{\mathbb{Z}\star}\mu(Z \setminus \tilde{Z}) = 0$, and second equality holds by Fubini's theorem.

The sets of the form $W \times Z \in \mathcal{A}_{\mathbb{X}}$ are stable under finite intersections and generate $\mathcal{A}_{\mathbb{X}} = \mathcal{A}_{\mathbb{W}} \otimes \mathcal{A}_{\mathbb{Z}}$, so we have shown that our expression satisfies Definition 2.2(ii), and hence is indeed the disintegration $\mu^{\pi_{\mathbb{Z}}}(\cdot | z)$. The theorem then follows by projecting onto $\mathbb{W}$. $\square$

2.1. Disintegration Building Blocks

Because of the almost everywhere nature of Theorem 2.4, all equalities in this section are meant $h_{\star}\mu$ -almost everywhere.

In this subsection we establish six scaffolding results about disintegrations that are of independent interest. We will later demonstrate in Section 2.2 how they can be combined to give, under certain regularity assumptions, an explicit formula for disintegrations on manifolds in Theorem 2.14. We anticipate that other results further characterising disintegrations can be obtained from these lemmas.

The results are, we believe, quite intuitive and straightforward to explain narratively, so we summarise them briefly here. Lemma 2.6 provides the most elementary building block, showing that disintegrations under the identity map yield Dirac measures. Lemma 2.7 gives an expression for the disintegration of a product measure with respect to a coordinate map in terms of the product of the disintegration. Lemma 2.8 shows that the disintegration of the pushforward is the pushforward of the disintegration. Lemma 2.9 shows that the disintegration of a measure dominated by some base measure is dominated by the disintegration of the base measure, and its density is a renormalization of the prior

density. Lemma 2.10 shows that the disintegration of a measure restricted to some set is the renormalization of the restriction of the disintegration of the measure. Finally, Proposition 2.11 shows that disintegrations are invariant to composition of the observation map with a bijection.

Lemma 2.6 (Disintegration w.r.t. the Identity). *If $\mathbb{Y} = \mathbb{X}$ and $h: \mathbb{X} \rightarrow \mathbb{Y}$ is the identity map, then*

$$\mu^h(\cdot \mid y) = \delta_y$$

the Dirac measure at y .

Proof. $\mu^h(\cdot \mid y)$ is supported on $h^{-1}(\{y\}) = \{y\}$. Since $\mu^h(\cdot \mid y)$ is a probability measure, we necessarily have $\mu^h(\cdot \mid y) = \delta_y$. We check: for $X \subset \mathbb{X}$ measurable,

$$\int_{\mathbb{Y}} \delta_y(X) h_{\star} \mu(dy) = \int_{\mathbb{X}} \mathbb{1}_X(x) \mu(dx) = \int_X \mu(dx) = \mu(X)$$

as required. $\square$

In the next lemma $\mathbb{X}_1, \mathbb{X}_2$ are under the same assumptions as $\mathbb{X}$ (Assumption 2.3(i)), and μ_1, μ_2 are under the same assumptions as μ (Assumption 2.3(iii)) but on $\mathbb{X}_1, \mathbb{X}_2$ respectively.

Lemma 2.7 (Disintegration of Product Measure). *Suppose $\mathbb{X} = \mathbb{X}_1 \times \mathbb{X}_2$, $\mu = \mu_1 \times \mu_2$, and that μ_2 is finite. Let $h_1: \mathbb{X}_1 \rightarrow \mathbb{Y}$, and $h = h_1 \circ \pi_1: \mathbb{X} \rightarrow \mathbb{Y}$, $(x_1, x_2) \mapsto h_1(x_1)$. Then, for $y \in \mathbb{Y}$,*

$$\mu_2(\mathbb{X}_2) \mu^h(\cdot \mid y) = \mu_1^{h_1}(\cdot \mid y) \times \mu_2.$$

Proof. Property (i) in Definition 2.2 for μ^h follows from the corresponding property of $\mu_1^{h_1}$:

$$(\mu_1^{h_1}(\cdot \mid y) \times \mu_2)(\mathbb{X} \setminus h^{-1}(\{y\})) = \mu_1^{h_1}(\mathbb{X}_1 \setminus h_1^{-1}(\{y\}) \mid y) \cdot \mu_2(\mathbb{X}_2) = 0.$$

Now for $X_1 \subset \mathbb{X}_1$, $X_2 \subset \mathbb{X}_2$ measurable

$$\begin{aligned} \int_{\mathbb{Y}} (\mu_1^{h_1}(\cdot \mid y) \times \mu_2)(X_1 \times X_2) h_{\star} \mu(dy) &= \int_{\mathbb{X}_1} \mu_1^{h_1}(X_1 \mid y) \cdot \mu_2(X_2) h_{\star} \mu(dy) \\ &= \mu_2(X_2) \int_{\mathbb{X}_1} \mu_1^{h_1}(X_1 \mid y) h_{\star} \mu(dy) \\ &= \mu_2(\mathbb{X}_2) \mu_2(X_2) \int_{\mathbb{X}_1} \mu_1^{h_1}(X_1 \mid y) h_{1\star} \mu_1(dy) \\ &= \mu_2(\mathbb{X}_2) \mu_1(X_1) \mu_2(X_2) \\ &= \mu_2(\mathbb{X}_2) \mu(X_1 \times X_2). \end{aligned}$$

where we used $h_{\star} \mu = \mu_2(\mathbb{X}_2) h_{1\star} \mu_1$. The sets of the form $X_1 \times X_2 \in \mathcal{A}_{\mathbb{X}}$ are stable under finite intersections and generate $\mathcal{A}_{\mathbb{X}} = \mathcal{A}_{\mathbb{X}_1} \otimes \mathcal{A}_{\mathbb{X}_2}$, so this is enough to conclude the proof. $\square$

Remark 2.4. It is important that μ_2 is a finite measure in order for $\pi_{1\star}\mu$ and hence $h_\star\mu$ to be σ -finite.

In the following lemma $\tilde{\mathbb{X}}$ is under the same assumptions as $\mathbb{X}$ (Assumption 2.3(i)).

Lemma 2.8 (Disintegration of Pushforward Measure). *Let $f: \mathbb{X} \rightarrow \tilde{\mathbb{X}}$, $g: \tilde{\mathbb{X}} \rightarrow \mathbb{Y}$ be measurable maps such that $h = g \circ f$. Then, for $y \in \mathbb{Y}$,*

$$(f_\star\mu)^g(\cdot \mid y) = f_\star(\mu^h(\cdot \mid y)).$$

Proof. First checking (i) in Definition 2.2, we have

$$\begin{aligned} f_\star(\mu^h(\cdot \mid y))(\tilde{\mathbb{X}} \setminus g^{-1}(\{y\})) &= \mu^h\left(f^{-1}\left(\tilde{\mathbb{X}} \setminus g^{-1}(\{y\})\right) \mid y\right) \\ &= \mu^h(\mathbb{X} \setminus (f^{-1} \circ g^{-1})(\{y\}) \mid y) \\ &= \mu^h(\mathbb{X} \setminus h^{-1}(\{y\}) \mid y) \\ &= 0. \end{aligned}$$

Moreover, for $X \subset \mathbb{X}$ measurable,

$$\int_{\mathbb{Y}} f_\star(\mu^h(\cdot \mid y))(X) g_\star(f_\star\mu)(dy) = \int_{\mathbb{Y}} \mu^h(f^{-1}(X) \mid y) h_\star\mu(dy) = \mu(f^{-1}(X)) = f_\star\mu(X).$$

□

In the next lemma ν is under the same assumptions as μ (Assumption 2.3(iii)).

Lemma 2.9 (Disintegration of Dominated Measure). *Suppose $\mu \ll \nu$. Then for $y \in \mathbb{Y}$, $\mu^h(\cdot \mid y)$ is absolutely continuous with respect to $\nu^h(\cdot \mid y)$ and*

$$\frac{d\mu^h(\cdot \mid y)}{d\nu^h(\cdot \mid y)} = \frac{d\mu/d\nu}{\nu^h(d\mu/d\nu \mid y)}$$

where $\nu^h(d\mu/d\nu \mid y) > 0$ $h_\star\mu$ -almost everywhere.

Proof. Property (i) in Definition 2.2 for μ^h follows from the corresponding property of ν^h . It remains to show property (ii). Note that for $Y \subset \mathbb{Y}$ measurable

$$\begin{aligned} \int_Y \nu^h\left(\frac{d\mu}{d\nu} \mid y\right) h_\star\nu(dy) &= \int_{\mathbb{Y}} \mathbb{1}_Y(y) \cdot \nu^h\left(\frac{d\mu}{d\nu} \mid y\right) h_\star\nu(dy) \\ &= \int_{\mathbb{Y}} \nu^h\left(\mathbb{1}_{h^{-1}(Y)} \cdot \frac{d\mu}{d\nu} \mid y\right) h_\star\nu(dy) \\ &= \nu\left(\mathbb{1}_{h^{-1}(Y)} \cdot \frac{d\mu}{d\nu}\right) \\ &= \mu(h^{-1}(Y)) \end{aligned}$$

$$= h_\star \mu(Y),$$

so

$$\frac{d(h_\star \mu)}{d(h_\star \nu)}(y) = \nu^h \left(\frac{d\mu}{d\nu} \middle| y \right).$$

Now letting $Y := \{y \in \mathbb{Y} : \nu^h(d\mu/d\nu \mid y) = 0\}$, we have

$$h_\star \mu(Y) = \int_Y \nu^h \left(\frac{d\mu}{d\nu} \middle| y \right) h_\star \nu(dy) = 0$$

so $\nu^h(d\mu/d\nu \mid y) > 0$ $h_\star \mu$ -almost everywhere.

Thus, for $X \subset \mathbb{X}$ measurable,

$$\begin{aligned} \int_{\mathbb{Y}} \nu^h \left(\frac{d\mu}{d\nu} \middle| y \right) \mu^h(X \mid y) h_\star \nu(dy) &= \int_{\mathbb{Y}} \mu^h(X \mid y) h_\star \mu(dy) \\ &= \mu(X) \\ &= \nu \left(\mathbb{1}_X \cdot \frac{d\mu}{d\nu} \right) \\ &= \int_{\mathbb{Y}} \nu^h \left(\mathbb{1}_X \cdot \frac{d\mu}{d\nu} \middle| y \right) h_\star \nu(dy) \\ &= \int_{\mathbb{Y}} \int_X \frac{d\mu}{d\nu}(x) \nu^h(dx \mid y) h_\star \nu(dy) \end{aligned}$$

and hence

$$\nu^h \left(\frac{d\mu}{d\nu} \middle| y \right) \mu^h(\cdot \mid y) = \frac{d\mu}{d\nu} \nu^h(\cdot \mid y)$$

is the unique disintegration of μ with respect to $h_\star \nu$ (see Chang and Pollard [1997, Theorem 1]). $\square$

Lemma 2.10 (Disintegration of Restricted Measure). *Let $X \subset \mathbb{X}$ be measurable. Then, for $y \in \mathbb{Y}$,*

$$\mu^h(\cdot \mid y) \Big|_X = \mu^h(X \mid y) (\mu|_X)^h(\cdot \mid y).$$

Proof. Let ν be the measure on $\mathbb{X}$ defined by $\nu := \mu(\cdot \cap X)$. Then $\nu \ll \mu$ with $\frac{d\nu}{d\mu} = \mathbb{1}_X$. Thus, applying Lemma 2.9, we obtain for any measurable $X' \subset \mathbb{X}$

$$\mu^h(X \mid y) \nu^h(X' \mid y) = \mu^h(\mathbb{1}_X \mid y) \nu^h(X' \mid y) = \int_{X'} \mathbb{1}_X(x) \mu^h(dx \mid y) = \mu^h(X' \cap X \mid y).$$

The result then follows by restricting the measures to X and noting that since $\nu|_X = \mu|_X$ we have $\nu^h(\cdot \mid y)|_X = (\nu|_X)^h(\cdot \mid y) = (\mu|_X)^h(\cdot \mid y)$. $\square$

We have now established all the lemmas required for our main result (Theorem 2.14) in the following subsection. We will nevertheless prove an additional result that will be of interest in the upcoming discussion. Here $\tilde{\mathbb{Y}}$ is under the same assumptions as $\mathbb{Y}$ (Assumption 2.3(ii)).

Proposition 2.11 (Disintegration w.r.t. Equivalent Observations). *Let $f: \mathbb{Y} \rightarrow \tilde{\mathbb{Y}}$ be an isomorphism of measure spaces. Then*

$$\mu^h(\cdot | y) = \mu^{f \circ h}(\cdot | f(y)).$$

Proof. Property (i) in Definition 2.2 follows from the fact that

$$h^{-1}(\{y\}) = (f \circ h)^{-1}(\{f(y)\}).$$

Now for $X \subset \mathbb{X}$ measurable, by (measure-theoretic) change of variables,

$$\int_{\mathbb{Y}} \mu^{f \circ h}(X | f(y)) h_{\star} \mu(dy) = \int_{\tilde{\mathbb{Y}}} \mu^{f \circ h}(X | \tilde{y}) (f \circ h)_{\star} (d\tilde{y}) = \mu(X).$$

□

2.2. Construction of Disintegrations on Manifolds

We begin by stating some additional assumptions on manifold structure of $\mathbb{X}$ that are required for our main result in Section 2, Theorem 2.14. For readers unfamiliar with the differential geometry used in this section, Appendix A provides a concise summary of the relevant definitions and results.

Assumption 2.12.

- (i) Let $\mathbb{X}$ be a d -dimensional C^k manifold. Assume that μ is absolutely continuous with respect to a volume measure on $\mathbb{X}$.
- (ii) Let $\mathbb{Y} = \mathbb{R}^n$ for some $n \in \mathbb{N}$, and $h: \mathbb{X} \rightarrow \mathbb{Y}$ such that the points in $\mathbb{X}$ are μ -almost everywhere C^k regular w.r.t. h .

A C^k regular point w.r.t. h is a point $x \in \mathbb{X}$ around which h is C^k and the derivative of h is surjective (Definition A.8). By the preimage theorem (Theorem A.9), Assumption 2.12 ensures in particular that $h^{-1}(\{y\})$ is a $(d - n)$ -dimensional C^k submanifold of $\mathbb{X}$ for $h_{\star} \mu$ -almost every y . We will use this to construct the disintegration $h_{\star} \mu$ -almost everywhere, on these manifolds.

There is no canonical way to restrict an ambient measure or volume form from $\mathbb{X}$ to the generally lower dimensional fiber $h^{-1}(\{y\})$ directly³. However, if we equip $\mathbb{X}$ with a Riemannian metric then we obtain a Riemannian volume measure $\omega_{\mathbb{X}}$, and there *is* a canonical restriction of this metric. The restricted Riemannian metric on $h^{-1}(\{y\})$ then induces a Riemannian volume measure $\omega_{h^{-1}(\{y\})}$ on $h^{-1}(\{y\})$. This construction allows us more generally to restrict the measure μ , as we will see in the next definition.

³Even when the pushforward of the volume measure $h_{\star} \omega_{\mathbb{X}}$ is σ -finite and we can disintegrate $\omega_{\mathbb{X}}$, such disintegration cannot be interpreted as a canonical *restriction* to the fiber, since we will see the disintegration depends on h through more than just the fiber $h^{-1}(\{y\})$ (Remark 2.8).

Remark 2.5. In Assumption 2.12(i), by a volume measure we mean a measure induced by a d -form (Definition A.6). All such volume measures are absolutely continuous with respect to any Riemannian volume measure $\omega_{\mathbb{X}}$ (Definition A.12), thus we could equivalently assume in Assumption 2.12(i) that μ is absolutely continuous with respect to $\omega_{\mathbb{X}}$.

Definition 2.13 (Riemannian Restricted Measure). Suppose $\mathbb{X}$ is a C^k Riemannian manifold, $\tilde{\mathbb{X}}$ a C^k ($k \geq 1$) Riemannian submanifold of $\mathbb{X}$ and μ is absolutely continuous with respect to $\omega_{\mathbb{X}}$. Then the *Riemannian restricted measure* $\mu_{\tilde{\mathbb{X}}}$ is the measure given by $\mu_{\tilde{\mathbb{X}}} := \frac{d\mu}{d\omega_{\mathbb{X}}} \omega_{\tilde{\mathbb{X}}}$.

Remark 2.6. Note the different notation used for a measure that has been restricted in a *measure-theoretic* sense, as in Lemma 2.10, i.e. $\mu|_{\tilde{\mathbb{X}}}$, compared to the *Riemannian* restricted measure from Definition 2.13, i.e. $\mu_{\tilde{\mathbb{X}}}$. The two are generally distinct. In particular, whenever $\dim \tilde{\mathbb{X}} < \dim \mathbb{X}$, if μ is absolutely continuous with respect to a volume measure on $\mathbb{X}$ then $\mu|_{\tilde{\mathbb{X}}}$ is a zero measure, so not a useful notion of restriction for the purposes of subsequent results.

For $y \in \mathbb{Y}$, $\mu^h(\cdot | y)$ is a probability measure supported on $h^{-1}(\{y\})$. This could lead one to ask whether disintegrations $\mu^h(\cdot | y)$ correspond to renormalizations of Riemannian restricted measures $\mu_{h^{-1}(\{y\})}$, under Assumption 2.12. Perhaps surprisingly they do not. Riemannian restricted measures depend on the choice of Riemannian metric while, by definition, disintegrations do not.

Next we state and prove Theorem 2.14, our main result for this section. The theorem gives an explicit expression for disintegrations under Assumption 2.12, showing that they differ from renormalized Riemannian restricted measures by a determinant-like term. This term quantifies how dense fibers lie locally along directions orthogonal to the fiber. Intuitively speaking, regions in which the fibers lie less dense need to be equipped with higher disintegration density for the law of total probability (Definition 2.2(ii)) to hold. For more concrete, visual intuition, see Figure 1 and Example 2.1.

Theorem 2.14. *Under Assumption 2.12 with $k = 1$, equip $\mathbb{X}$ with a Riemannian metric and denote by $\omega_{\mathbb{X}}$ the corresponding volume measure on $\mathbb{X}$. Then $\mu^h(\cdot | y)$ is absolutely continuous with respect to the Riemannian volume measure $\omega_{h^{-1}(\{y\})}$ on the Riemannian submanifold $h^{-1}(\{y\})$ and*

$$\frac{d\mu^h(\cdot | y)}{d\omega_{h^{-1}(\{y\})}}(x) \propto \frac{d\mu}{d\omega_{\mathbb{X}}}(x) \cdot \left| \omega_{\ker(Dh(x))^\perp}(x) [\nabla h(x)] \right|^{-1}.$$

Proof. Let $y \in \mathbb{R}^n$ a C^1 regular value of h and $x \in h^{-1}(\{y\})$. By the implicit function theorem (Theorem A.9) there are bounded neighbourhoods U of x in $\mathbb{X}$, V of y in $\mathbb{Y} = \mathbb{R}^n$, W of 0 in $\mathbb{R}^{d-n}$ and a C^1 diffeomorphism $\varphi: U \rightarrow V \times W$ such that $\pi_V \circ \varphi = h$, where $\pi_V: V \times W \rightarrow V$ is the projection. Let $\lambda_{V \times W}$, λ_V , λ_W be the Lebesgue measures on $V \times W$, V and W respectively. Then, by Equation (A.4), the pushforward of the

Riemannian volume measure on U through the chart is given by⁴

$$\varphi_*\omega_U = |\omega_U[\nabla\varphi \circ \varphi^{-1}]|^{-1}\lambda_{V \times W} = |\omega_U[\nabla\varphi \circ \varphi^{-1}]|^{-1}(\lambda_V \times \lambda_W).$$

So, disintegrating this measure

$$(\varphi_*\omega_U)^{\pi_V}(\cdot | y) \propto |\omega_U[\nabla\varphi \circ \varphi^{-1}]|^{-1}\lambda_{V \times W}^{\pi_V}(\cdot | y) \quad (\text{Lemma 2.9})$$

$$\propto |\omega_U[\nabla\varphi \circ \varphi^{-1}]|^{-1}(\lambda_V^{\text{id}_V}(\cdot | y) \times \lambda_W) \quad (\text{Lemma 2.7})$$

$$= |\omega_U[\nabla\varphi \circ \varphi^{-1}]|^{-1}(\delta_y \times \lambda_W) \quad (\text{Lemma 2.6})$$

hence

$$\mu|_U^h(\cdot | y) \propto \frac{d\mu}{d\omega_U}\omega_U^h(\cdot | y) \quad (\text{Lemma 2.9})$$

$$= \frac{d\mu}{d\omega_U}\varphi_*^{-1}((\varphi_*\omega_U)^{\pi_V}(\cdot | y)) \quad (\text{Lemma 2.8})$$

$$\propto \frac{d\mu}{d\omega_U}|\omega_U[\nabla\varphi]|^{-1}\varphi_*^{-1}(\delta_y \times \lambda_W).$$

We also have, by Equation (A.4), the pushforward Riemannian volume measure on submanifold $h^{-1}(\{y\}) \cap U$ of U is given by

$$\varphi_*\omega_{h^{-1}(\{y\}) \cap U} = |\omega_{h^{-1}(\{y\}) \cap U}[\nabla\varphi|_{\ker(Dh)} \circ \varphi^{-1}]|^{-1}(\delta_y \times \lambda_W)$$

where we used that $T_x h^{-1}(\{y\}) \cap U = \ker(Dh(x))$. So

$$\omega_{h^{-1}(\{y\}) \cap U} = |\omega_{h^{-1}(\{y\}) \cap U}[\nabla\varphi|_{\ker(Dh)}]|^{-1}\varphi_*^{-1}(\delta_y \times \lambda_W).$$

Thus we see that $\mu|_U^h(\cdot | y) \ll \omega_{h^{-1}(\{y\}) \cap U}$ and, by Equation (A.3),

$$\begin{aligned} \frac{d\mu|_U^h(\cdot | y)}{d\omega_{h^{-1}(\{y\}) \cap U}} &\propto \frac{d\mu}{d\omega_U} \left| \frac{\omega_{h^{-1}(\{y\}) \cap U}[\nabla\varphi|_{\ker(Dh)}]}{\omega_U[\nabla\varphi]} \right| \\ &= \frac{d\mu}{d\omega_U} \left| \frac{\omega_{h^{-1}(\{y\}) \cap U}[\nabla\varphi|_{\ker(Dh)}]}{\omega_{h^{-1}(\{y\}) \cap U}[\nabla\varphi|_{\ker(Dh)}] \cdot \omega_{\ker(Dh)^\perp}[\nabla\varphi|_{\ker(Dh)^\perp}]} \right| \\ &= \frac{d\mu}{d\omega_U} \left| \omega_{\ker(Dh)^\perp}[\nabla\varphi|_{\ker(Dh)^\perp}] \right|^{-1}. \end{aligned}$$

$x \in h^{-1}(\{y\})$ was arbitrary, and U was chosen as a neighbourhood of x . Moreover $\omega_{h^{-1}(\{y\})}|_U = \omega_{h^{-1}(\{y\}) \cap U}$, and by Lemma 2.10 $\mu^h(\cdot | y)|_U \propto \mu|_U^h(\cdot | y)$. Thus, patching together open sets U we get

$$\frac{d\mu^h(\cdot | y)}{d\omega_{h^{-1}(\{y\})}} \propto \frac{d\mu}{d\omega_{\mathbb{X}}} \cdot \left| \omega_{\ker(Dh)^\perp}[\nabla h] \right|^{-1}.$$

□

⁴We abstract away the manifold variable in the notation of this proof, e.g. $\omega_U[\nabla\varphi \circ \varphi^{-1}]$ stands for the function $z \mapsto \omega_U(z)[\nabla\varphi(\varphi^{-1}(z))]$

In the case $\mathbb{X} = \mathbb{R}^d$, with the standard Euclidean metric we get $\omega_{\mathbb{R}^d} = \lambda$ the Lebesgue measure, and the restricted volume form term involves the determinant of the Jacobian of h , restricted to the orthogonal complement of its kernel. This observation yields the special case $\mathbb{X} = \mathbb{R}^d$ result from the introduction (Theorem 1.1).

Remark 2.7. In Theorem 2.14, the disintegration does not depend on the choice of Riemannian metric. Instead, the Riemannian metric provides a way to constructively represent the disintegration. We endow $\mathbb{X}$ with a Riemannian geometry only to provide a reference measure, $\omega_{h^{-1}(\{y\})}$, with-respect-to which we can then construct a density for the disintegration. Note that Riemannian metric can always be constructed on any C^k manifold with $k \geq 1$ (Remark A.1).

Particularly pertinent for Section 3 is that if $\mathbb{X}$ has a metric structure, that structure need not have any relationship with the Riemannian distance function implied by the Riemannian metric chosen.

Remark 2.8. Theorem 2.14 shows that, for fixed $y \in \mathbb{Y}$, $\mu^h(\cdot | y)$ does not merely depend on h through $h^{-1}(\{y\})$, but also on the behaviour of h in a neighbourhood of y , through its gradient. However, comparing this with Proposition 2.11, we see that it is not the *values* of h that affect the disintegration; Proposition 2.11 rather shows that $\mu^h(\cdot | y)$ only depends on h through the family of fibers $\{h^{-1}(\{y'\})\}_{y' \in \mathbb{Y}}$ with distinguished element $h^{-1}(\{y\})$.

Example 2.1. Consider a standard Gaussian prior $\mu = \mathcal{N}(0, I)$ on $\mathbb{X} = \mathbb{R}^2$. Suppose we wish to condition on an exact observation $y \in \mathbb{Y} = \mathbb{R}_{\geq 0}$ made through the quadratic observation operator $h(x) = \frac{x_1^2}{a^2} + \frac{x_2^2}{b^2}$ with $a = 1$ and $b = \frac{1}{2}$. In this case the observation fibers $h^{-1}(\{y\})$ are centered, axis-aligned ellipses with width $2a\sqrt{y}$ and height $2b\sqrt{y}$. The above is visualized in Figure 1a. In the Figure, we choose a uniformly spaced grid of observations $y \in \mathbb{Y}$ to visualize different regular conditional distributions simultaneously. In this setting, Assumption 2.12 is fulfilled (for $k = \infty$), which means that we can construct both the restricted measures $\mu_{h^{-1}(\{y\})}$ from Definition 2.13 as well as the version of the disintegration from Theorem 2.14 (or, more specifically, Theorem 1.1), which we visualize in Figures 1c and 1d, respectively. In these plots, every fiber uses its own independent color map, which means that colors are not comparable between fibers. We can observe a stark contrast between the restriction densities and the disintegration densities along the fibers closest to the origin.

This is due to the corrective term involving the volume form in Theorem 2.14: Here, by Theorem 1.1, this term simplifies to $\|\nabla h(x)\|_2^{-1}$. We have

$$\|\nabla h(x)\|_2^{-1} = \frac{1}{2} \sqrt{\frac{x_1^2}{a^4} + \frac{x_2^2}{b^4}}^{-1} = \frac{a}{2} \sqrt{\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2}} \cdot \frac{a^2}{b^2}^{-1} \leq \frac{a}{2} \sqrt{\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2}}^{-1} = \frac{a}{2\sqrt{y}},$$

since $\frac{a^2}{b^2} \geq 1$. By an analogous argument, $\|\nabla h(x)\|_2^{-1} \geq \frac{b}{2\sqrt{y}}$, where equalities are attained at the intersections with the respective half-axes of the ellipse. This means that the

corrective term is larger if y is small, i.e., for fibers that are closer to the origin. Moreover, the corrective term varies in strength by a factor of $\frac{a}{b}$ between the poles of the ellipse.

Since the observations in Figure 1 are uniformly spaced, the relative magnitude of this term can also be read off the plot from the “density” of the fibers in the normal direction, as formalised by the inverse-magnitude of the gradient in that direction. Intuitively speaking, the corrective factor accounts for the fact that the law of total probability encoded in Definition 2.2(ii) still requires the predictive probability mass to be distributed locally between fibers, even if those fibers lie less dense. This explains the discrepancy between the restriction and disintegration densities. For a more direct comparison of the densities, we also single out an observation at $y = 1.01$ and parameterize the corresponding elliptical fiber curve by arc length starting at the point and velocity vector marked in red in Figures 1c and 1d. Figure 1b shows the densities along this fiber curve. Note that, due to the arclength parameterization, no additional change of variables is needed. Remarkably, this panel also shows that the modes of the restriction and disintegration densities are maximally distant.

3. Modes of Disintegrations

For probability measures μ on Euclidean spaces that have a density $\frac{d\mu}{d\lambda}$ with respect to the Lebesgue measure λ , the modes are defined as the maximizers of said density. However, neither the restricted measures nor the elements of the disintegrations constructed in Section 2.2 have densities with respect to λ in general, as they are supported on lower dimensional fibers. Luckily, there are generalizations of this notion of mode, which measure the local accumulation of probability mass by comparing the probability of metric balls in the small noise limit. We first recap the relevant definitions of and key results about these generalized notions of modes below.

For the remainder of the article, $\mathbb{X}$ will be a metric space and μ a locally finite measure on the Borel σ -algebra of $\mathbb{X}$.

Definition 3.1 (Modes). A point $x^* \in \text{supp}(\mu) \subset \mathbb{X}$ is called a

(i) *weak mode* of μ if

$$\limsup_{r \downarrow 0} \frac{\mu(B_r(x))}{\mu(B_r(x^*))} \leq 1$$

for all $x \in \mathbb{X}$, and a

(ii) *strong mode* of μ if

$$\limsup_{r \downarrow 0} \frac{\sup_{x \in \mathbb{X}} \mu(B_r(x))}{\mu(B_r(x^*))} \leq 1.$$

The definitions above are due to Dashti et al. [2013, Definition 3.1], Helin and Burger [2015, Definition 4], and, in the form presented here, Ayanbayev et al. [2022, Definition 3.6 and 3.7]. However, all of the above only define modes of probability measures, so we

restate definitions and results here (often without proof) if they also hold in our more general setting of locally finite measures.

The following result motivates the naming of weak and strong modes.

Proposition 3.2 (Ayanbayev et al., 2022, Proposition 3.9). *Every strong mode $x \in \mathbb{X}$ of μ is a weak mode of μ .*

Just as for modes defined via Lebesgue densities, weak modes can be characterized as the solutions of optimization problems. However the objective function of these optimization problems needs to be defined only using properties of the measure μ . These objective functions are given by Onsager-Machlup functionals, as in the next definition.

Definition 3.3 (Onsager-Machlup Functional; see Ayanbayev et al., 2022, Definition 3.1). Let $E_\mu \subset \text{supp}(\mu) \subset \mathbb{X}$ be non-empty. A function $I_\mu: E_\mu \rightarrow \mathbb{R}$ is called an *Onsager-Machlup (OM) functional* for μ if

$$\lim_{r \downarrow 0} \frac{\mu(B_r(x_1))}{\mu(B_r(x_2))} = \exp(I_\mu(x_2) - I_\mu(x_1))$$

for all $x_1, x_2 \in E_\mu$. An OM functional is called *exhaustive* if

$$\lim_{r \downarrow 0} \frac{\mu(B_r(x_1))}{\mu(B_r(x_2))} = 0$$

for all $x_1 \in \mathbb{X} \setminus E_\mu$ and $x_2 \in E_\mu$.

As noted by Ayanbayev et al. [2022], the exhaustiveness property ensures that the domain of definition of the OM function is in some sense “maximal”. Lemma C.1 is sometimes useful to verify exhaustiveness of an OM functional.

The weak modes of the measure μ can be characterized as the solutions to a minimization problem whose objective function is given by the OM functional. Intuitively speaking, the OM functional takes the role of a negative log Legesgue density.

Theorem 3.4 (Ayanbayev et al., 2022, Proposition 4.1). *If $I_\mu: E_\mu \rightarrow \mathbb{R}$ is an exhaustive OM functional for μ , then $x^\star$ is a weak mode of μ if and only if*

$$x^\star \in \arg \min_{x \in E_\mu} I_\mu(x).$$

We will now make use of Theorem 2.14 to characterize the modes of the disintegration μ^h by constructing exhaustive OM functionals.

Remark 3.1. Our characterisation is complicated by the fact that a disintegration is only $h_\star\mu$ -almost-surely unique; in other words, there are (typically infinitely-many) *versions* of a disintegration that differ on $h_\star\mu$ -null sets. To simplify terminology we focus on

the specific version of the disintegration constructed in Theorem 2.14, and say that a function is an (exhaustive) OM functional for the disintegration $\mu^h(\cdot | y)$ if it is an (exhaustive) OM functional for the version from Theorem 2.14, which is well-defined under Assumption 2.12 if y is a regular value of h . Note that, under this assumption, the non-regular values of h form a $h_\star\mu$ -null set. This ensures that the functions given below are still (exhaustive) OM functionals for $h_\star\mu$ -almost all $y \in \mathbb{Y}$ in all other versions of the disintegration.

The general mode theory presented thus far only required $\mathbb{X}$ to be a metric space. However, we will now work under Assumption 2.12 in order for Theorem 2.14 to hold. This assumes in particular that $\mathbb{X}$ is a manifold, and we implicitly require that the manifold topology coincides with the metric space topology. As mentioned in Remark 2.7, we do not generally require the Riemannian distance function (see Theorem A.14) to coincide with the metric space structure. This gives us the flexibility to study *non-Riemannian* metric space structures, for example the one induced by ℓ_p norms in Section 3.2. We will later study the special case where the metric space structure coincides with the Riemannian distance function (Proposition 3.9 and Corollary 3.10).

Theorem 3.5. *Under Assumption 2.12 with $k = 1$, equip $\mathbb{X}$ with a Riemannian metric. Assume that $y \in \mathbb{Y}$ is a regular value and that the Riemannian restricted measure $\mu_{h^{-1}(\{y\})}$ admits an OM functional $I_{\mu_{h^{-1}(\{y\})}}: E_{\mu_{h^{-1}(\{y\})}} \rightarrow \mathbb{R}$. Then the disintegration $\mu^h(\cdot | y)$ admits an OM functional $I_{\mu^h(\cdot | y)}: E_{\mu^h(\cdot | y)} \rightarrow \mathbb{R}$, with $E_{\mu^h(\cdot | y)} := E_{\mu_{h^{-1}(\{y\})}}$, and*

$$I_{\mu^h(\cdot | y)}(x) := I_{\mu_{h^{-1}(\{y\})}}(x) + \log \left| \omega_{\ker(Dh(x))^\perp}(x) [\nabla h(x)] \right|.$$

Furthermore if $I_{\mu_{h^{-1}(\{y\})}}$ is exhaustive then $I_{\mu^h(\cdot | y)}$ is exhaustive.

Proof. By Theorem 2.14, the disintegration $\mu^h(\cdot | y)$ can be constructed as

$$\mu^h(dx | y) \propto \left| \omega_{\ker(Dh(x))^\perp}(x) [\nabla h(x)] \right|^{-1} \mu_{h^{-1}(\{y\})}(dx).$$

The function

$$\mathbb{X} \rightarrow \mathbb{R}_{>0}, x \mapsto \left| \omega_{\ker(Dh(x))^\perp}(x) [\nabla h(x)] \right|^{-1}$$

is continuous. Hence, the result follows by Lemma C.2. $\square$

One might wonder whether the OM functional of the Riemannian restricted measure, $I_{\mu_{h^{-1}(\{y\})}}(x)$, is given by $-\log \frac{d\mu}{d\omega_{\mathbb{X}}}$. This fails to hold for general metrics on $\mathbb{X}$ and, perhaps surprisingly, even for non-Euclidean normed spaces $\mathbb{X}$, as we will show in Section 3.2. However, in the following we show that this is indeed the case if the metric on $\mathbb{X}$ is induced by the Riemannian metric used to define $\omega_{\mathbb{X}}$ and under mild regularity assumptions on $\mathbb{X}$, $\frac{d\mu}{d\omega_{\mathbb{X}}}$ and h .

Note that since volume measures on manifolds are locally finite, we can discuss OM functionals $I_{\omega_{\mathbb{X}}}$ and $I_{\omega_{h^{-1}(\{y\})}}$ for the Riemannian volume measures $\omega_{\mathbb{X}}$ and $\omega_{h^{-1}(\{y\})}$ respectively.

Assumption 3.6. μ is absolutely continuous with respect to a volume measure on the manifold $\mathbb{X}$, with continuous Radon-Nikodym derivative.

When $\mathbb{X}$ is equipped with a Riemannian metric, Assumption 3.6 implies that μ is absolutely continuous with respect to $\omega_{\mathbb{X}}$ (Remark 2.5) and $\frac{d\mu}{d\omega_{\mathbb{X}}}$ is continuous.

Corollary 3.7. *Under Assumption 2.12 with $k = 1$ and Assumption 3.6, assume that $y \in \mathbb{Y}$ is a regular value and $\omega_{h^{-1}(\{y\})}$ admits an OM functional $I_{\omega_{h^{-1}(\{y\})}} : E_{\omega_{h^{-1}(\{y\})}} \rightarrow \mathbb{R}$. Then the disintegration $\mu^h(\cdot | y)$ admits an OM functional $I_{\mu^h(\cdot | y)} : E_{\mu^h(\cdot | y)} \rightarrow \mathbb{R}$ with $E_{\mu^h(\cdot | y)} := E_{\omega_{h^{-1}(\{y\})}} \cap \text{supp}^\circ(\frac{d\mu}{d\omega_{\mathbb{X}}})^5$ and*

$$I_{\mu^h(\cdot | y)}(x) := I_{\omega_{h^{-1}(\{y\})}}(x) - \log \frac{d\mu}{d\omega_{\mathbb{X}}}(x) + \log \left| \omega_{\ker(\text{D}h(x))^\perp}(x) [\nabla h(x)] \right|.$$

If, additionally, $I_{\omega_{h^{-1}(\{y\})}}$ is exhaustive then $I_{\mu^h(\cdot | y)}$ is exhaustive.

Proof. The Radon-Nikodym derivative $\frac{d\mu}{d\omega_{\mathbb{X}}}$ is continuous and hence, by Lemma C.2,

$$I_{\mu_{h^{-1}(\{y\})}} = I_{\omega_{h^{-1}(\{y\})}} - \log \frac{d\mu}{d\omega_{\mathbb{X}}}$$

is an OM functional for $\mu_{h^{-1}(\{y\})}$ and $I_{\mu_{h^{-1}(\{y\})}}$ is exhaustive if $I_{\omega_{h^{-1}(\{y\})}}$ is. The claim then follows by Theorem 3.5. $\square$

Next we establish that, when the metric on $\mathbb{X}$ is compatible with the Riemannian geometry, the OM functional is constant.

Assumption 3.8.

- (i) $\mathbb{X}$ is a d -dimensional C^2 Riemannian manifold which, as a metric space, is equipped with the Riemannian distance metric.
- (ii) $\mathbb{Y} = \mathbb{R}^n$ for some $n \in \mathbb{N}$, and $h : \mathbb{X} \rightarrow \mathbb{Y}$ such that the points in $\mathbb{X}$ are μ -almost everywhere C^2 regular w.r.t. h .

Proposition 3.9 (Riemannian Onsager-Machlup Functionals of Volume Measures).

- (i) *Under Assumption 3.8(i), any constant function $I_{\omega_{\mathbb{X}}} : \mathbb{X} \rightarrow \mathbb{R}$ is an exhaustive OM functional for the Riemannian volume measure $\omega_{\mathbb{X}}$, and,*
- (ii) *under Assumption 3.8, if $y \in \mathbb{Y}$ is a regular value, any constant function $I_{\omega_{h^{-1}(\{y\})}} : E_{\omega_{h^{-1}(\{y\})}} \rightarrow \mathbb{R}$ with $E_{\omega_{h^{-1}(\{y\})}} = h^{-1}(\{y\}) \subset \mathbb{X}$ is an exhaustive OM functional for $\omega_{h^{-1}(\{y\})}$.*

⁵We define $\text{supp}^\circ(f) := \{x \in \mathbb{X} : f(x) > 0\}$.

Proof. Note that by Proposition B.1, the Riemannian volume of Riemannian balls is to first order

$$\omega_{\mathbb{X}}(B_r^{\mathbb{X}}(x)) = V_d r^d + o(r^d)$$

as $r \downarrow 0$, where V_d is the volume of the d -dimensional Euclidean ball. Thus we have for $x_1, x_2 \in \mathbb{X}$,

$$\frac{\mu(B_r(x_1))}{\mu(B_r(x_2))} = \frac{V_d r^d + o(r^d)}{V_d r^d + o(r^d)} = \frac{V_d + o(1)}{V_d + o(1)} \xrightarrow{r \downarrow 0} 1$$

which concludes the proof of (i).

For (ii), if $x \notin h^{-1}(\{y\})$ then $\omega_{h^{-1}(\{y\})}(B_r^{\mathbb{X}}(x)) = 0$ for all r small enough. If $x \in h^{-1}(\{y\})$, we have by Proposition B.2

$$\omega_{h^{-1}(\{y\})}(B_r^{\mathbb{X}}(x)) = V_{d-n} r^{d-n} + o(r^{d-n})$$

as $r \downarrow 0$, and the proof of (ii) follows as for (i). $\square$

Note that combining Assumption 3.6 and Assumption 3.8 amounts to Assumption 2.12 with additional regularity requirements. These results lead to a concrete characterization of OM functionals of disintegrations:

Corollary 3.10. *Under Assumptions 3.6 and 3.8, additionally assume that $y \in \mathbb{Y}$ is a regular value. Then $\mu^h(\cdot | y)$ admits an exhaustive OM functional $I_{\mu^h(\cdot | y)}: E_{\mu^h(\cdot | y)} \rightarrow \mathbb{R}$ with $E_{\mu^h(\cdot | y)} := h^{-1}(\{y\}) \cap \text{supp}^\circ(\frac{d\mu}{d\omega_{\mathbb{X}}})$ and*

$$I_{\mu^h(\cdot | y)}(x) := -\log \frac{d\mu}{d\omega_{\mathbb{X}}}(x) + \log \left| \omega_{\ker(Dh(x))^\perp}(x) [\nabla h(x)] \right|.$$

Proof. This follows directly from Corollary 3.7 and Proposition 3.9(ii). $\square$

By Theorem 3.4, this leads to a characterization of weak modes of the disintegration $\mu^h(\cdot | y)$ as the solutions of the constrained optimization problem

$$\begin{aligned} \min_{x \in \mathbb{X}} \quad & -\log \frac{d\mu}{d\omega_{\mathbb{X}}}(x) + \log \left| \omega_{\ker(Dh(x))^\perp}(x) [\nabla h(x)] \right|, \\ \text{s.t.} \quad & h(x) = y. \end{aligned}$$

Theorem 1.2 in the introduction is a simplified version of this result in $\mathbb{R}^d$.

This result shows that the maximizers of the disintegration densities in Figure 1b are indeed the weak modes of the disintegration when equipping the ambient space with the Euclidean metric.

3.1. Restricted Modes Disagree with Modes of Disintegrations

Chen et al. [2024, Definition 3] define a variant of a mode on a fibre by restricting the search space for a strong mode from Definition 3.1 to the fibre. Below we generalize their definition to the setting of metric spaces and provide a weak version of the definition in order to study their relationship to weak and strong modes of disintegrations. We will demonstrate that in general these modes differ from the modes of disintegrations and, for this reason, refer to these modes as “restricted modes” rather than conditional modes as in Chen et al. [2024].

Definition 3.11 (Restricted Modes). Let $y \in \text{supp}(h_\star \mu)$. A point $x^\star \in h^{-1}(\{y\}) \cap \text{supp}(\mu) \subset \mathbb{X}$ is called a

(i) *weak $h^{-1}(\{y\})$ -restricted mode* of μ if

$$\limsup_{r \downarrow 0} \frac{\mu(B_r(x))}{\mu(B_r(x^\star))} \leq 1$$

for all $x \in h^{-1}(\{y\})$, and a

(ii) *strong $h^{-1}(\{y\})$ -restricted mode* of μ if

$$\limsup_{r \downarrow 0} \frac{\sup_{x \in h^{-1}(\{y\})} \mu(B_r(x))}{\mu(B_r(x^\star))} \leq 1.$$

As for regular weak and strong modes, we start by justifying the “weak” and “strong” attributes in the naming of restricted modes.

Proposition 3.12. *Every strong $h^{-1}(\{y\})$ -restricted mode $x \in h^{-1}(\{y\})$ is also a weak $h^{-1}(\{y\})$ -restricted mode.*

Proof. We adapt the proof of Lemma 3.9 in Ayanbayev et al. [2022].

Let $x^\star \in h^{-1}(\{y\}) \cap \text{supp}(\mu)$ be a strong $h^{-1}(\{y\})$ -restricted mode of μ . For all $x \in h^{-1}(\{y\})$, we have

$$\limsup_{r \downarrow 0} \frac{\mu(B_r(x))}{\mu(B_r(x^\star))} \leq \limsup_{r \downarrow 0} \frac{\sup_{x' \in h^{-1}(\{y\})} \mu(B_r(x'))}{\mu(B_r(x^\star))} \leq 1,$$

since $x^\star$ is a strong restricted mode of μ . Hence, $x^\star$ is a weak restricted mode of μ . $\square$

Just as for weak modes, weak restricted modes can be characterized as the minimizers of exhaustive OM functionals. However, for weak restricted modes, we can slightly weaken the notion of exhaustiveness:

Definition 3.13. An Onsager-Machlup functional $I_\mu: E_\mu \rightarrow \mathbb{R}$ for μ is called *$h^{-1}(\{y\})$ -exhaustive* if

$$\lim_{r \downarrow 0} \frac{\mu(B_r(x_1))}{\mu(B_r(x_2))} = 0$$

for all $x_2 \in E_\mu \cap h^{-1}(\{y\})$ and $x_1 \in h^{-1}(\{y\}) \setminus E_\mu$.

We next prove the natural result that weak $h^{-1}(\{y\})$ -restricted modes are minimisers of $h^{-1}(\{y\})$ -exhaustive OM functionals.

Proposition 3.14. *Let $I_\mu: E_\mu \rightarrow \mathbb{R}$ be an $h^{-1}(\{y\})$ -exhaustive Onsager-Machlup functional for μ . Then x^* is a weak $h^{-1}(\{y\})$ -restricted mode of μ if and only if*

$$x^* \in \arg \min_{x \in h^{-1}(\{y\})} I_\mu(x).$$

Proof. We adapt the proof of Proposition 4.1 in Ayanbayev et al. [2022].

Let $x \in h^{-1}(\{y\}) \cap E_\mu$. Then, by the definition of an $h^{-1}(\{y\})$ -exhaustive OM functional, we have

$$\limsup_{r \downarrow 0} \frac{\mu(B_r(x'))}{\mu(B_r(x))} = \lim_{r \downarrow 0} \frac{\mu(B_r(x'))}{\mu(B_r(x))} = \begin{cases} \exp(I_\mu(x) - I_\mu(x')) & \text{if } x' \in E_\mu, \\ 0 & \text{otherwise,} \end{cases}$$

for all $x' \in h^{-1}(\{y\})$. Hence, x is a weak $h^{-1}(\{y\})$ -restricted mode of μ if and only if $\exp(I_\mu(x) - I_\mu(x')) \leq 1$ for all $x' \in h^{-1}(\{y\})$, or, equivalently, $I_\mu(x) \leq I_\mu(x')$ for all $x' \in h^{-1}(\{y\})$.

Now let $x \in h^{-1}(\{y\}) \setminus E_\mu$. Since $I_\mu(x) = +\infty$, it remains to show that x is not a weak $h^{-1}(\{y\})$ -restricted mode of μ . Suppose for the sake of contradiction that this is the case. By definition, we then have $x \in \text{supp}(\mu)$. Pick $x' \in h^{-1}(\{y\}) \cap E_\mu \subset \text{supp}(\mu)$. Then $\frac{\mu(B_r(x'))}{\mu(B_r(x))} > 0$ for every $r > 0$, and hence

$$1 \geq \limsup_{r \downarrow 0} \frac{\mu(B_r(x'))}{\mu(B_r(x))} = \left(\liminf_{r \downarrow 0} \frac{\mu(B_r(x))}{\mu(B_r(x'))} \right)^{-1},$$

which is equivalent to

$$\liminf_{r \downarrow 0} \frac{\mu(B_r(x))}{\mu(B_r(x'))} \geq 1.$$

This contradicts the $h^{-1}(\{y\})$ -exhaustiveness of the OM functional. We conclude that $h^{-1}(\{y\}) \setminus E_\mu$ does not contain any weak $h^{-1}(\{y\})$ -restricted modes of μ . $\square$

Using some of the assumptions and results from the previous sections, we can build $h^{-1}(\{y\})$ -exhaustive OM functionals for μ :

Corollary 3.15. *Under Assumptions 3.6 and 3.8, if $y \in \mathbb{Y}$ is a regular value, the function $-\log \frac{d\mu}{d\omega_{\mathbb{X}}} \big|_{h^{-1}(\{y\})}$ is*

- (i) *an $h^{-1}(\{y\})$ -exhaustive OM functional for μ , and,*
- (ii) *an exhaustive OM functional for $\mu_{h^{-1}(\{y\})}$.*

Proof. Since $-\log \frac{d\mu}{d\omega_{\mathbb{X}}}$ is continuous, (i) follows from Proposition 3.9(i), and (ii) follows from Proposition 3.9(ii), by the same argument as in the proof of Theorem 3.5. $\square$

Corollary 3.15 together with Theorem 3.4 and Proposition 3.14 imply that, under Assumptions 3.6 and 3.8, restricted modes correspond to modes of the Riemannian restricted measure, *not* of the disintegration. This explains our choice to rename this notion of mode from “conditional mode” to “restricted mode”.

In Figure 1 and Example 2.1 we have seen that weak restricted modes and weak modes of the disintegration can disagree catastrophically, even in finite-dimensional Hilbert spaces.

3.2. Failure Modes in Normed Spaces

Many of the results above apply when the metric on $\mathbb{X}$ coincides with the Riemannian distance function. If the underlying space $\mathbb{X}$ is a vector space, then this is the case when $\mathbb{X}$ is equipped with an inner product that generates a norm and a metric on the space. However, especially in probabilistic numerical methods, it is often more natural to work instead in a normed space whose norm is not generated by an inner product. For instance, when approximating strong solutions to partial differential equations, it is common to choose $\mathbb{X}$ to be a Hölder space with the appropriate norm.

It is now natural to ask whether a change in norm affects the results on weak modes of disintegrations and restricted measures above. Below we will show that this is indeed the case. It is well-known that Onsager-Machlup functionals have a strong dependence on the metric used to define them [see e.g., Ayanbayev et al., 2022, Example B.4]. However, perhaps counterintuitively, in our setting the differing behaviour of the weak modes under different norms only arises after disintegrating or restricting, leaving the prior’s modes unaffected.

In this subsection $\mathbb{X} = \mathbb{R}^d$ is a normed space. Write λ for the Lebesgue measure on $\mathbb{R}^d$. The standard Euclidean inner product on $\mathbb{R}^d$ can be viewed as a Riemannian metric, with Riemannian volume form λ . Thus, under Assumption 2.12, we have a canonical restriction $\lambda_{h^{-1}(\{y\})}$ of λ to the fiber $h^{-1}(\{y\}) \subset \mathbb{R}^d$. We will start by providing a version of Proposition 3.9 in normed spaces.

Proposition 3.16 (Onsager-Machlup Functionals of the Lebesgue Measure in Normed Spaces).

- (i) Any constant function $I_\lambda: \mathbb{R}^d \rightarrow \mathbb{R}$ is an exhaustive OM functional for λ , and,
- (ii) under Assumption 2.12 with $k = 2$, if $y \in \mathbb{Y}$ is a regular value then $\lambda_{h^{-1}(\{y\})}$ admits an exhaustive OM functional

$$I_{\lambda_{h^{-1}(\{y\})}}: E_{\lambda_{h^{-1}(\{y\})}} \rightarrow \mathbb{R}, \quad x \mapsto -\log V_{d-n}(x)$$

where $E_{\lambda_{h^{-1}(\{y\})}} = h^{-1}(\{y\}) \subset \mathbb{X}$ and $V_{d-n}(x)$ is the $(d-n)$ -dimensional volume of $B_1^{\mathbb{X}}(0) \cap \ker(Dh(x))$ in $\ker(Dh(x)) \subset \mathbb{R}^d$.

Proof. (i) is a consequence of the translation invariance of λ and of the metric (induced by the norm) on $\mathbb{R}^d$.

For (ii), if $x \notin h^{-1}(\{y\})$ then $\lambda_{h^{-1}(\{y\})}(B_r^{\mathbb{R}^d}(x)) = 0$ for all r small enough. If $x \in h^{-1}(\{y\})$, we have by Proposition B.3

$$\lambda_{h^{-1}(\{y\})}(B_r^{\mathbb{R}^d}(x)) = V_{d-n}(x)r^{d-n} + o(r^{d-n})$$

as $r \downarrow 0$, and the proof of (ii) follows as the proof of Proposition 3.9. $\square$

Proposition 3.16(i) implies that the OM functional of μ , and hence its modes, do not depend on the norm on $\mathbb{R}^d$. However, surprisingly, Proposition 3.16(ii) implies that the OM functional and the modes of $\mu^h(\cdot | y)$ does depend on the norm. Intuitively, this is because the balls $B_r^{\mathbb{X}}(0)$ are not necessarily isotropic, and so the OM functional depends on the orientation of the fibre with respect to the balls.

Combining Proposition 3.16 with Corollary 3.7, we obtain the explicit characterization of the OM functional of the disintegration in normed spaces:

Corollary 3.17. *Let $\mathbb{X} = \mathbb{R}^d$ a normed space, and suppose Assumption 2.12(ii) holds with $k = 2$, as well as Assumption 3.6. If $y \in \mathbb{Y}$ is a regular value, then $\mu^h(\cdot | y)$ admits an exhaustive OM functional $I_{\mu^h(\cdot | y)}: E_{\mu^h(\cdot | y)} \rightarrow \mathbb{R}$ with $E_{\mu^h(\cdot | y)} := h^{-1}(\{y\}) \cap \text{supp}^\circ(\frac{d\mu}{d\lambda})$ and*

$$I_{\mu^h(\cdot | y)}(x) := -\log V_{d-n}(x) - \log \frac{d\mu}{d\lambda}(x) + \log \left| \det Dh(x)|_{\ker(Dh(x))^\perp} \right|,$$

where $V_{d-n}(0)$ is the $(d-n)$ -dimensional volume of $B_1^{\mathbb{X}}(0) \cap \ker(Dh(x))$ in $\ker(Dh(x)) \subset \mathbb{R}^d$.

The counterintuitive behaviour of the OM functionals with respect to a change in norm is illustrated in Figure 2. Here varying the value of p among ℓ_p norms on $\mathbb{R}^2$ leads to different locations and numbers of minimizers, and minima of certain OM functionals become local maxima in others.

4. Discussion

This work has pointed out multiple troubling discrepancies between disintegrations and restrictions of measures, with a particular focus on their modes. Whether restricted modes or disintegration modes are the object of interest is problem-specific. The dividing line between the two approaches is whether the law of total probability is obeyed. Consequently, the paradigm selected hinges on the relevance of said law in a given application. If all potential outcomes y of the predictive distribution $h_*\mu$ are equally valid measurements, the disintegration approach seems more natural. For instance, this is the case in most of Bayesian statistics.

However for some downstream applications, only one “observation” is relevant or even semantically meaningful, and therefore a restriction rather than a disintegration may be preferred. This is the case for the Laplace approximation under function-space priors proposed by Cinquin et al. [2024], and one could argue also for certain probabilistic

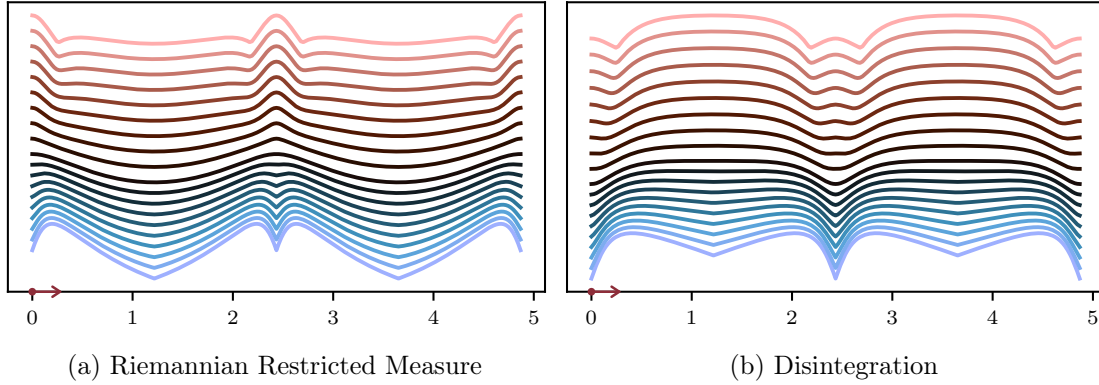


Figure 2: Onsager-Machlup functionals of the restriction and disintegration in Example 2.1 and Figure 1b with respect to different ℓ_p norms. The bottom (blue) lines correspond to $p = 1$, the central (black) lines correspond to $p = 2$, and the top (salmon) lines correspond to $p = \infty$. The remaining values interpolate atan-uniformly between these values. Each of these OM functionals can be shifted by an arbitrary constant, so we spread them vertically for visualization purposes. The change in norm leads to OM functionals with different minimizers, i.e., different weak modes of the distributions. For instance, for the disintegration, the ℓ_2 OM functional has two local minima, which correspond to local maxima of the ℓ_∞ OM functional. The latter admits four other local minima.

numerical methods, particularly those based on solution of ODEs (e.g. in the formalism of Tronarp et al. [2019]) for which one specifies a residual, and it is not clear that setting said residual to any other value than 0 is meaningful. In such cases there is no clear justification for enforcing the law of total probability, as marginalising over all possible observations in $\text{supp}(h_\star\mu)$ gives weight to these irrelevant observations. This suggests that, in such cases, a potential generalisation of the restriction operation to more general settings (such as infinite-dimensional vector spaces) is more appropriate than disintegration.

Acknowledgments

The authors thank Philipp Hennig and Bálint Mucsányi for helpful discussions and feedback on this work. MP and NDC gratefully acknowledge financial support by the European Research Council through ERC StG Action 757275 / PANAMA and ERC CoG Action 101123955 / ANUBIS; the DFG Cluster of Excellence “Machine Learning - New Perspectives for Science”, EXC 2064/1, project number 390727645; the German Federal Ministry of Education and Research (BMBF) through the Tübingen AI Center (FKZ: 01IS18039A); the DFG SPP 2298 (Project HE 7114/5-1); and the Carl Zeiss Foundation (project "Certification and Foundations of Safe Machine Learning Systems in Healthcare"); as well as funds from the Ministry of Science, Research and Arts of the State of Baden-Württemberg. NDC acknowledges the support of the Fonds National de

la Recherche, Luxembourg. The authors thank the International Max Planck Research School for Intelligent Systems (IMPRS-IS) for supporting MP and NDC. JC is supported by EPSRC grant EP/Y001028/1.

References

- Birzhan Ayanbayev, Ilja Klebanov, Han Cheng Lie, and T. J. Sullivan. Γ -convergence of Onsager-Machlup functionals. Part I: With applications to maximum a posteriori estimation in Bayesian inverse problems. *Inverse Problems*, 38(2), 2022. doi:10.1088/1361-6420/ac3f81.
- Christopher M. Bishop. *Pattern recognition and machine learning*. Information science and statistics. Springer, New York, 2006. ISBN 978-0-387-31073-2.
- J. T. Chang and D. Pollard. Conditioning as disintegration. *Statistica Neerlandica*, 51(3): 287–317, November 1997. ISSN 0039-0402, 1467-9574. doi:10.1111/1467-9574.00056. URL <https://onlinelibrary.wiley.com/doi/10.1111/1467-9574.00056>.
- Yifan Chen, Bamdad Hosseini, Houman Owhadi, and Andrew M. Stuart. Solving and learning nonlinear PDEs with Gaussian processes. *Journal of Computational Physics*, 447:110668, December 2021. ISSN 0021-9991. doi:10.1016/j.jcp.2021.110668. URL <https://linkinghub.elsevier.com/retrieve/pii/S0021999121005635>. Publisher: Elsevier BV.
- Yifan Chen, Bamdad Hosseini, Houman Owhadi, and Andrew M. Stuart. Gaussian measures conditioned on nonlinear observations: Consistency, MAP estimators, and simulation, 2024.
- Tristan Cinquin, Marvin Pförtner, Vincent Fortuin, Philipp Hennig, and Robert Bamler. FSP-Laplace: Function-space priors for the Laplace approximation in Bayesian deep learning. In A. Globerson, L. Mackey, D. Belgrave, A. Fan, U. Paquet, J. Tomczak, and C. Zhang, editors, *Advances in Neural Information Processing Systems*, volume 37, pages 13897–13926. Curran Associates, Inc., 2024. doi:10.48550/arxiv.2407.13711.
- Christian Clason, Tapio Helin, Remo Kretschmann, and Petteri Piiroinen. Generalized Modes in Bayesian Inverse Problems. *SIAM/ASA Journal on Uncertainty Quantification*, 7(2):652–684, January 2019. ISSN 2166-2525. doi:10.1137/18m1191804. URL <https://epubs.siam.org/doi/10.1137/18M1191804>. Publisher: Society for Industrial & Applied Mathematics (SIAM).
- Jon Cockayne, Chris J. Oates, T. J. Sullivan, and Mark Girolami. Bayesian Probabilistic Numerical Methods. *SIAM Review*, 61(3):756–789, January 2019. ISSN 0036-1445, 1095-7200. doi:10.1137/17M1139357. URL <https://epubs.siam.org/doi/10.1137/17M1139357>.
- M. Dashti, K. J. H. Law, A. M. Stuart, and J. Voss. MAP estimators and their consistency in Bayesian nonparametric inverse problems. *Inverse Problems*, 29(9), 2013. doi:10.1088/0266-5611/29/9/095017.

- Claude Dellacherie and Paul André Meyer. *Probabilities and potential*. Number 29 in North-Holland mathematics studies. Hermann [u.a.], Paris, 1978. ISBN 978-2-7056-5857-1 978-0-7204-0701-3.
- Persi Diaconis, Susan Holmes, and Mehrdad Shahshahani. Sampling from a Manifold. In *Advances in Modern Statistical Theory and Applications: A Festschrift in honor of Morris L. Eaton*, volume 10, pages 102–126. Institute of Mathematical Statistics, January 2013. doi:10.1214/12-IMSCOLL1006. URL <https://projecteuclid.org/ebooks/institute-of-mathematical-statistics-collections/Advances-in-Modern-Statistical-Theory-and-Applications--A-Festschrift/chapter/Sampling-from-a-Manifold/10.1214/12-IMSCOLL1006>.
- Sylvestre Gallot, Dominique Hulin, and Jacques Lafontaine. *Riemannian Geometry*. Universitext. Springer, Berlin, Heidelberg, 2004. ISBN 978-3-540-20493-0 978-3-642-18855-8. doi:10.1007/978-3-642-18855-8. URL <http://link.springer.com/10.1007/978-3-642-18855-8>.
- Tapio Helin and Martin Burger. Maximum a posteriori probability estimates in infinite-dimensional Bayesian inverse problems. *Inverse Problems*, 31(8), 2015. doi:10.1088/0266-5611/31/8/085009.
- Ilya Klebanov, Hefin Lambley, and T. J. Sullivan. Classification of small-ball modes and maximum a posteriori estimators, March 2025. URL <http://arxiv.org/abs/2306.16278>. arXiv:2306.16278 [math].
- Hefin Lambley. Strong maximum a posteriori estimation in Banach spaces with Gaussian priors. *Inverse Problems*, 39(12), 2023. doi:10.1088/1361-6420/ad07a4.
- Olivier Laurent, Emanuel Aldea, and Gianni Franchi. A symmetry-aware exploration of Bayesian neural network posteriors. In *ICLR*, 2024. URL <https://openreview.net/forum?id=FOSBQuXgAq>.
- Han Cheng Lie and T. J. Sullivan. Equivalence of weak and strong modes of measures on topological vector spaces. *Inverse Problems*, 34(11), 2018. doi:10.1088/1361-6420/aadef2.
- Kanti V. Mardia and Peter E. Jupp. *Directional Statistics*. Wiley, January 1999. ISBN 9780470316979. doi:10.1002/9780470316979. URL <http://dx.doi.org/10.1002/9780470316979>.
- Renata Possobon and Christian S. Rodrigues. Geometric properties of disintegration of measures, February 2022. URL <https://ui.adsabs.harvard.edu/abs/2022arXiv220204511P>. ADS Bibcode: 2022arXiv220204511P.
- Tue Tjur. *A Constructive Definition of Conditional Distributions*. Institute of Mathematical Statistics, University of Copenhagen, 1975. Google-Books-ID: CeP7XwAACAAJ.
- Tue Tjur. *Probability based on Radon measures*. Wiley series in probability and mathematical statistics. J. Wiley, Chichester [Eng.]; New York, 1980. ISBN 978-0-471-27824-5.
- Filip Tronarp, Hans Kersting, Simo Särkkä, and Philipp Hennig. Probabilistic so-

lutions to ordinary differential equations as nonlinear Bayesian filtering: a new perspective. *Statistics and Computing*, 29(6):1297–1315, September 2019. ISSN 1573-1375. doi:10.1007/s11222-019-09900-1. URL <http://dx.doi.org/10.1007/s11222-019-09900-1>.

Filip Tronarp, Simo Särkkä, and Philipp Hennig. Bayesian ODE solvers: the maximum a posteriori estimate. *Statistics and Computing*, 31(3), May 2021. ISSN 0960-3174, 1573-1375. doi:10.1007/s11222-021-09993-7. URL <https://link.springer.com/10.1007/s11222-021-09993-7>. Publisher: Springer Science and Business Media LLC.

Jonas Gregor Wiese, Lisa Wimmer, Theodore Papamarkou, Bernd Bischl, Stephan Günnemann, and David Rügamer. Towards efficient mcmc sampling in Bayesian neural networks by exploiting symmetry. In *Machine Learning and Knowledge Discovery in Databases: Research Track*, pages 459–474, Cham, 2023. Springer Nature Switzerland. ISBN 978-3-031-43412-9.

Appendix

A. Differential Geometry Background

Here we include the necessary background in differential and Riemannian geometry. For a more complete exposition to this background, we refer to Gallot et al. [2004].

Let $\mathbb{X}$ be a d -dimensional C^k manifold, with $k \geq 1$.

Definition A.1 (Vector Bundle). For $l \in \mathbb{N}_0$, a l -vector bundle over $\mathbb{X}$ is a C^k -manifold $\mathbb{E}$ with a C^k map $\pi: \mathbb{E} \rightarrow \mathbb{X}$ such that for each $x \in \mathbb{X}$, $\pi^{-1}(\{x\}) = E_x$ is a l -dimensional real vector space, and there exists an open neighbourhood U of x in $\mathbb{X}$ and a diffeomorphism $\varphi: \pi^{-1}(U) \rightarrow U \times \mathbb{R}^l$ such that the following diagram commutes

$$\begin{array}{ccc} \pi^{-1}(U) & \xrightarrow{\varphi} & U \times \mathbb{R}^l \\ \pi \downarrow & \swarrow \pi_U & \\ U & & \end{array}$$

where $\pi_U: U \times \mathbb{R}^l \rightarrow U$ is the projection onto U , and $\varphi|_{E_x}: E_x \rightarrow \{x\} \times \mathbb{R}^l$ is a linear isomorphism.

A vector bundle should be thought of as attaching a vector space at each point of the manifold. For instance the tangent bundle $T\mathbb{X}$ and the cotangent bundle $T^*\mathbb{X}$ are vector bundles over $\mathbb{X}$.

Definition A.2 (Section). A *section* s of a vector bundle $\mathbb{E}$ over $\mathbb{X}$, is a C^k map $s: \mathbb{X} \rightarrow \mathbb{E}$ such that $\pi \circ s = \text{id}_{\mathbb{X}}$. We write $\Gamma(\mathbb{E})$ for the space of sections over $\mathbb{E}$.

This means $s(x)$ is a ‘vector’ at x . For example, given $f: \mathbb{X} \rightarrow \mathbb{R}$ we have $Df \in \Gamma(T^*\mathbb{X})$.

Definition A.3 (Tensor). For $(p, q) \in \mathbb{N}_0^2$, a (p, q) -tensor $\alpha(x)$ at $x \in \mathbb{X}$ is a multilinear map

$$\alpha: \underbrace{T_x^*\mathbb{X} \times \cdots \times T_x^*\mathbb{X}}_{p \text{ times}} \times \underbrace{T_x\mathbb{X} \times \cdots \times T_x\mathbb{X}}_{q \text{ times}} \rightarrow \mathbb{R},$$

i.e. for each $1 \leq j \leq p$ and $w_1, \dots, w_p, w'_j \in T_x^*\mathbb{X}$, $v_1, \dots, v_q \in T_x\mathbb{X}$, $a_j, b_j \in \mathbb{R}$,

$$\alpha(x)[w_1, \dots, a_j w_j + b_j w'_j, \dots, v_q] = a_j \alpha(x)[w_1, \dots, w_j, \dots, v_q] + b_j \alpha(x)[w_1, \dots, w'_j, \dots, v_q]$$

and similarly in the v_j arguments. Note that we can also view a (p, q) tensor as an element $\alpha(x) \in \underbrace{T_x\mathbb{X} \otimes \cdots \otimes T_x\mathbb{X}}_{p \text{ times}} \otimes \underbrace{T_x^*\mathbb{X} \otimes \cdots \otimes T_x^*\mathbb{X}}_{q \text{ times}} =: (T_x\mathbb{X})^{\otimes p} \otimes (T_x^*\mathbb{X})^{\otimes q}$.

This provides a natural definition of a (p, q) -tensor field α as an element $\alpha \in \Gamma((T_x\mathbb{X})^{\otimes p} \otimes (T_x^*\mathbb{X})^{\otimes q})$. For $x \in \mathbb{X}$, $\alpha(x)$ is thought of as a multilinear map as above.

Definition A.4 (Pullback). If $\mathbb{Y}$ is another C^k manifold and $f: \mathbb{X} \rightarrow \mathbb{Y}$ is a C^k map, we can *pullback* a $(0, q)$ -tensor $\alpha(f(x))$ at $f(x) \in \mathbb{Y}$ to obtain a $(0, q)$ -tensor $f^*\alpha(x)$ at $x \in \mathbb{X}$ given by

$$f^*\alpha(x)[v_1, \dots, v_q] := \alpha(f(x))[Df(x)[v_1], \dots, Df(x)[v_q]]$$

for $v_1, \dots, v_q \in T_x\mathbb{X}$.

If f is injective then using the above construction we have a well-defined notion of a pullback of a $(0, q)$ -tensor field α on $\mathbb{Y}$ to a $(0, q)$ -tensor field $f^*\alpha$ on $\mathbb{X}$.

Definition A.5 (Differential Form). For $q \in \mathbb{N}_0$, a q -form is a $(0, q)$ -tensor field which is antisymmetric, i.e. for each $1 \leq j < q$, $x \in \mathbb{X}$ and $v_1, \dots, v_q \in T_x\mathbb{X}$,

$$\omega(x)[v_1, \dots, v_j, v_{j+1}, \dots, v_q] = -\omega(x)[v_1, \dots, v_{j+1}, v_j, \dots, v_q].$$

Definition A.6 (Volume Measure). Any d -form ω defines a *volume measure* on $\mathbb{X}$ which, by abuse of notation, we also denote by ω . This is the unique measure such that, for any chart $\varphi: U \rightarrow V$, where $U \subset \mathbb{X}$, $V \subset \mathbb{R}^d$, $\varphi_*(\omega|_U) \ll \lambda_V$ the Lebesgue measure on V , with

$$\frac{d\varphi_*(\omega|_U)}{d\lambda_V}(z) = |\varphi^{-1*}\omega(z)[e_1, \dots, e_d]| \quad (\text{A.1})$$

where $e_1, \dots, e_d \in \mathbb{R}^d$ are the standard unit vectors.

Definition A.7 (Orientability). $\mathbb{X}$ is *orientable* if it admits a nowhere vanishing d -form.

Definition A.8 (Regular Point/Value). Let $f: \mathbb{X} \rightarrow \mathbb{Y}$ be a set-theoretic map between manifolds. $x \in \mathbb{X}$ is a *regular point* of f if f is C^k in a neighbourhood of x and $Df(x): T_x\mathbb{X} \rightarrow T_{f(x)}\mathbb{Y}$ is surjective. $y \in \mathbb{Y}$ is a *regular value* of f if each $x \in f^{-1}(\{y\})$ is a regular point.

Theorem A.9 (Implicit Function/Preimage Theorem). *Let $f: \mathbb{X} \rightarrow \mathbb{Y}$ be a set-theoretic map between manifolds. Write $n := \dim \mathbb{Y}$. If $x \in \mathbb{X}$ is a regular point of f , then there exists neighbourhoods U of x in $\mathbb{X}$, V of y in $\mathbb{Y}$, $W \subset \mathbb{R}^{d-n}$ and a C^k diffeomorphism $\varphi: U \rightarrow V \times W$ such that $\pi_V \circ \varphi = f$, where $\pi_V: V \times W \rightarrow V$ is the projection. Moreover, if $y \in \mathbb{Y}$ is a regular value of f , then $f^{-1}(\{y\})$ is a C^k submanifold of $\mathbb{X}$.*

A.1. Riemannian Geometry

Definition A.10 (Riemannian Metric). A *Riemannian metric* g is a $(0, 2)$ -tensor field which is positive definite, i.e. for each $x \in \mathbb{X}$ and each $v \in T_x\mathbb{X} \setminus \{0\}$, $g(x)[v, v] > 0$. It defines an inner product, and hence a norm $\|\cdot\|$, on each tangent space.

Remark A.1. A Riemannian metric can be constructed on any C^k manifold with $k \geq 1$. This can be done by choosing an inner product in each coordinate patch and smoothly “sticking” them together with a partition of unity.

In the remainder of this appendix, $\mathbb{X}$ is a d -dimensional C^k Riemannian manifold equipped with a Riemannian metric g .

Definition A.11 (Gradient). For a C^k map $f: \mathbb{X} \rightarrow \mathbb{R}^n$, we define its *gradient* $\nabla f \in \Gamma((T\mathbb{X})^n)$ as, for each $1 \leq j \leq n$ and $x \in \mathbb{X}$, $\nabla f(x)_j$ is the unique vector in $T_x\mathbb{X}$ such that

$$g(\nabla f(x)_j, \cdot) = \pi_j \circ Df(x)$$

where $\pi_j: \mathbb{R}^n \rightarrow \mathbb{R}$ is the projection onto the j^{th} coordinate.

Definition A.12 (Riemannian Volume Form). If $\mathbb{X}$ is orientable then it has, up-to-sign, a canonical volume form $\omega_{\mathbb{X}}$ called the *Riemannian volume form*. It is the unique, up-to-sign, d -form such that for each $x \in \mathbb{X}$ and orthonormal basis $v_1, \dots, v_d \in T_x\mathbb{X}$,

$$|\omega_{\mathbb{X}}(x)[v_1, \dots, v_d]| = 1. \quad (\text{A.2})$$

Without assuming orientability, $\mathbb{X}$ still admits, up-to-sign, a canonical volume form in each coordinate patch. Thus the absolute value of the Riemannian volume form is still well-defined globally. Abusing notation, we write it $|\omega_{\mathbb{X}}|$. Moreover, the induced *Riemannian volume measure* $\omega_{\mathbb{X}}$ (see Definition A.6) is also well-defined in general.

If $\mathbb{Y}$ and $\mathbb{Z}$ are Riemannian submanifolds of $\mathbb{X}$, $x \in \mathbb{Y} \cap \mathbb{Z}$ and $(T_x\mathbb{Y})^\perp = T_x\mathbb{Z}$, then for $v_1, \dots, v_n \in T_x\mathbb{Y}$, $v_{n+1}, \dots, v_d \in T_x\mathbb{Z}$, we have

$$|\omega_{\mathbb{X}}(x)[v_1, \dots, v_d]| = |\omega_{\mathbb{Y}}(x)[v_1, \dots, v_n] \cdot \omega_{\mathbb{Z}}(x)[\pi_{\mathbb{Z}}(v_{n+1}), \dots, \pi_{\mathbb{Z}}(v_d)]| \quad (\text{A.3})$$

where $\pi_{\mathbb{Z}}: T_x\mathbb{X} \rightarrow T_x\mathbb{Z}$ is the orthogonal projection.

Moreover, given a chart $\varphi: U \rightarrow V$, where $U \subset \mathbb{X}$, $V \subset \mathbb{R}^d$, for each $x \in U$ choose an orthonormal basis $v_1(x), \dots, v_d(x) \in T_x\mathbb{X}$. Write $M(x)$ for the matrix representation of $D\varphi(x)$ with respect to the basis $v_1(x), \dots, v_d(x)$ and the standard basis of $\mathbb{R}^d$, $e_1, \dots, e_d$. Then the Riemannian volume measure pushed through the chart satisfies by Equation (A.1)

$$\begin{aligned} \frac{d\varphi_*\omega_U}{d\lambda_V}(z) &= |\varphi^{-1*}\omega_U(z)[e_1, \dots, e_m]| \\ &= |\omega_U(\varphi^{-1}(z))[D\varphi^{-1}(z)[e_1], \dots, D\varphi^{-1}(z)[e_m]]| \\ &= |\det(M(\varphi^{-1}(z))^{-1})| \\ &= \left| \det\left(M(\varphi^{-1}(z))^\top\right)^{-1} \right| \\ &= |\omega_U(\varphi^{-1}(z))[\nabla\varphi(\varphi^{-1}(z))]|^{-1}. \end{aligned} \quad (\text{A.4})$$

Definition A.13 (Riemannian Tangent Volume). For each $x \in \mathbb{X}$, $T_x\mathbb{X}$ has a canonical *Riemannian tangent volume measure* λ , given by identifying an orthonormal basis of $T_x\mathbb{X}$ with the unit vectors of $\mathbb{R}^d$ to obtain an isomorphism $T_x\mathbb{X} \cong \mathbb{R}^d$, and taking the Lebesgue measure on $\mathbb{R}^d$.

Theorem A.14 (Distance Function). $\mathbb{X}$ is a metric space when equipped with the distance function

$$d^{\mathbb{X}}(x, x') = \inf\{L(\gamma) : \gamma : [0, 1] \rightarrow \mathbb{X} \text{ a } C^k \text{ curve s.t. } \gamma(0) = x, \gamma(1) = x'\}$$

for $x, x' \in \mathbb{X}$, where

$$L(\gamma) = \int_0^1 \|\gamma'(t)\| dt.$$

Theorem A.15 (Exponential Map). Assuming $k \geq 2$, for each $x \in \mathbb{X}$ there is a C^{k-1} map $\exp_x^{\mathbb{X}} : V \rightarrow \mathbb{X}$ defined in a neighbourhood $V \subset T_x \mathbb{X}$ of 0 such that $D \exp_x^{\mathbb{X}}(0) : T_x \mathbb{X} \rightarrow T_x \mathbb{X}$ is the identity and $d^{\mathbb{X}}(x, \exp_x^{\mathbb{X}} v) = \|v\|$ for all $v \in V$. $\exp_x^{\mathbb{X}}$ is called the (Riemannian) exponential map at $x \in \mathbb{X}$.

B. Riemannian Volume of Small Balls

In this appendix $\mathbb{X}$ is a d -dimensional C^2 Riemannian manifold equipped with a Riemannian metric g . Write the metric ball $B_r^{\mathbb{X}}(x) = \{x' \in \mathbb{X} : d(x, x') < r\}$. We study the asymptotics of the volume of balls $\omega_{\mathbb{X}}(B_r^{\mathbb{X}}(x))$ and $\omega_{\tilde{\mathbb{X}}}(B_r^{\mathbb{X}}(x))$ as $r \downarrow 0$, for $\tilde{\mathbb{X}}$ a submanifold of $\mathbb{X}$.

Proposition B.1 (Volume of Riemannian Balls). We have for $x \in \mathbb{X}$

$$\omega_{\mathbb{X}}(B_r^{\mathbb{X}}(x)) = V_d r^d + o(r^d)$$

as $r \downarrow 0$, where V_d is the volume of the d -dimensional unit Euclidean ball.

Proof. $\exp_x^{\mathbb{X}}$ is a diffeomorphism from a neighbourhood $V \subset T_x \mathbb{X}$ of 0 to a neighbourhood $U \subset \mathbb{X}$ of x in $\mathbb{X}$. So we can invert it on U , yielding a chart $\exp_x^{\mathbb{X}}|_V^{-1} : U \rightarrow V$. By Equation (A.1), for r small enough such that $B_r^{\mathbb{X}}(0) \subset V$ and $v_1, \dots, v_d \in T_x \mathbb{X}$ an orthonormal basis,

$$\begin{aligned} \omega_{\mathbb{X}}(B_r(x)) &= \int_{B_r(x)} \left| \exp_x^{\mathbb{X}}|_V^* \omega_{\mathbb{X}}(v)[v_1, \dots, v_d] \right| \lambda(dv) \\ &= \left| \omega_{\mathbb{X}}(x)[D \exp_x^{\mathbb{X}}(x)[v_1], \dots, D \exp_x^{\mathbb{X}}(x)[v_d]] \right| V_d r^d + o(r^d) \\ &= \underbrace{|\omega_{\mathbb{X}}(x)[v_1, \dots, v_d]|}_{=1 \text{ (Equation (A.2))}} V_d r^d + o(r^d) \\ &= V_d r^d + o(r^d) \end{aligned}$$

as $r \downarrow 0$, where λ is the Riemannian tangent volume measure (Definition A.13). $\square$

Proposition B.2 (Volume of Ambient Riemannian Balls). *Let $\tilde{\mathbb{X}}$ be a $\tilde{d}$ -dimensional C^2 Riemannian submanifold of $\mathbb{X}$. Then for $x \in \tilde{\mathbb{X}}$,*

$$\omega_{\tilde{\mathbb{X}}}(B_r^{\mathbb{X}}(x)) = V_{\tilde{d}} r^{\tilde{d}} + o(r^{\tilde{d}})$$

as $r \downarrow 0$, where $V_{\tilde{d}}$ is the volume of the $\tilde{d}$ -dimensional unit Euclidean ball.

Proof. $\exp_x^{\tilde{\mathbb{X}}}$ is a diffeomorphism from a neighbourhood $V \subset T_x \tilde{\mathbb{X}}$ of 0 to a neighbourhood $U \subset \tilde{\mathbb{X}}$ of x in $\tilde{\mathbb{X}}$. So we can invert it on U .

For $r > 0$, it is clear that $B_r^{\tilde{\mathbb{X}}}(x) \subset B_r^{\mathbb{X}}(x)$. Now the Riemannian distance function $d^{\tilde{\mathbb{X}}}(x, \cdot)$ to x on $\tilde{\mathbb{X}}$ is given for $x' \in U$ by

$$d^{\tilde{\mathbb{X}}}(x, x') = \left\| \exp_x^{\tilde{\mathbb{X}}} |_V^{-1}(x') \right\|.$$

Since the derivative of the exponential map at the origin of the tangent space is the identity, we have by Taylor's theorem, for $\exp_x^{\tilde{\mathbb{X}}}(v) \in U$

$$d^{\tilde{\mathbb{X}}}(x, \exp_x^{\tilde{\mathbb{X}}}(v)) = \left\| \exp_x^{\tilde{\mathbb{X}}} |_V^{-1}(\exp_x^{\tilde{\mathbb{X}}}(v)) \right\| = \|v + o(\|v\|)\| = \|v\| + o(\|v\|)$$

as $v \rightarrow 0$. So $B_r^{\tilde{\mathbb{X}}}(x) \subset B_{r+o(r)}^{\mathbb{X}}(x)$ as $r \downarrow 0$. Now by Proposition B.1

$$\omega_{\tilde{\mathbb{X}}}(B_r^{\tilde{\mathbb{X}}}(x)) = V_{\tilde{d}} r^{\tilde{d}} + o(r^{\tilde{d}})$$

as $r \downarrow 0$. Hence

$$\begin{aligned} V_{\tilde{d}} r^{\tilde{d}} + o(r^{\tilde{d}}) &= \omega_{\tilde{\mathbb{X}}}(B_r^{\tilde{\mathbb{X}}}(x)) \\ &\leq \omega_{\tilde{\mathbb{X}}}(B_r^{\mathbb{X}}(x)) \\ &\leq \omega_{\tilde{\mathbb{X}}}(B_{r+o(r)}^{\tilde{\mathbb{X}}}(x)) \\ &= V_{\tilde{d}} (r + o(r))^{\tilde{d}} + o(r^{\tilde{d}}) \\ &= V_{\tilde{d}} r^{\tilde{d}} + o(r^{\tilde{d}}) \end{aligned}$$

as $r \downarrow 0$, i.e.

$$\omega_{\tilde{\mathbb{X}}}(B_r^{\mathbb{X}}(x)) = V_{\tilde{d}} r^{\tilde{d}} + o(r^{\tilde{d}})$$

as $r \downarrow 0$. □

The next proposition is of a similar flavour to Proposition B.2, but the ambient space is a normed space.

Proposition B.3 (Volume of Ambient Normed Balls). *Let $\mathbb{R}^d$ be a normed space and $\tilde{\mathbb{X}} \subset \mathbb{R}^d$ a $\tilde{d}$ -dimensional embedded C^2 manifold equipped with some Riemannian metric. Then for $x \in \tilde{\mathbb{X}}$,*

$$\omega_{\tilde{\mathbb{X}}}(B_r^{\mathbb{R}^d}(x)) = V_{\tilde{d}}(x) r^{\tilde{d}} + o(r^{\tilde{d}})$$

as $r \downarrow 0$ where $V_{\tilde{d}}(x)$ is the Riemannian tangent volume of $B_1^{\mathbb{R}^d}(0) \cap T_x \tilde{\mathbb{X}}$, which in general depends on x , where we view $T_x \tilde{\mathbb{X}}$ as a linear subspace of $\mathbb{R}^d$.

Proof. We write $\|\cdot\|_{\mathbb{R}^d}$ for the ambient norm in $\mathbb{R}^d \supset T_x \tilde{\mathbb{X}}$. $\exp_x^{\tilde{\mathbb{X}}}$ is a diffeomorphism from a neighbourhood $V \subset T_x \tilde{\mathbb{X}}$ of 0 to a neighbourhood $U \subset \tilde{\mathbb{X}}$ of x in $\tilde{\mathbb{X}}$. So we can invert it on U . Then for $x' \in B_r^{\mathbb{R}^d}(0) \cap U \subset \tilde{\mathbb{X}}$ we have by Taylor's theorem

$$\left\| \exp_x^{\tilde{\mathbb{X}}} |V^{-1}(x') \right\|_{\mathbb{R}^d} = \|x' - x + o(\|x' - x\|_{\mathbb{R}^d})\|_{\mathbb{R}^d} = \|x' - x\|_{\mathbb{R}^d} + o(\|x' - x\|_{\mathbb{R}^d})$$

as $x' \rightarrow x$. Hence

$$\exp_x^{\tilde{\mathbb{X}}}(B_{r-o(r)}^{\mathbb{R}^d}(0) \cap T_x \tilde{\mathbb{X}}) \subset B_r^{\mathbb{R}^d}(x) \cap \tilde{\mathbb{X}} \subset \exp_x^{\tilde{\mathbb{X}}}(B_{r+o(r)}^{\mathbb{R}^d}(0) \cap T_x \tilde{\mathbb{X}})$$

as $r \downarrow 0$. So similarly to the proof of Proposition B.1, $B_r^{\mathbb{R}^d}(x) \cap \tilde{\mathbb{X}} \subset \exp_x^{\tilde{\mathbb{X}}}(B_{r+o(r)}^{\mathbb{R}^d}(0) \cap T_x \tilde{\mathbb{X}})$ implies

$$\begin{aligned} \omega_{\tilde{\mathbb{X}}}(B_r^{\mathbb{R}^d}(x)) &\leq \int_{B_{r+o(r)}^{\mathbb{R}^d}(0) \cap T_x \tilde{\mathbb{X}}} \left| \exp_x^{\tilde{\mathbb{X}}} |V^* \omega_{\mathbb{X}}(v)[v_1, \dots, v_d] \right| \lambda(dv) \\ &= V_{\tilde{d}}(x)(r + o(r))^{\tilde{d}} + o(r^{\tilde{d}}) \\ &= V_{\tilde{d}}(x)r^{\tilde{d}} + o(r^{\tilde{d}}) \end{aligned}$$

as $r \downarrow 0$. Similarly the inclusion $\exp_x^{\tilde{\mathbb{X}}}(B_{r-o(r)}^{\mathbb{R}^d}(0) \cap T_x \tilde{\mathbb{X}}) \subset B_r^{\mathbb{R}^d}(x) \cap \tilde{\mathbb{X}}$ yields the reverse inequality

$$\omega_{\tilde{\mathbb{X}}}(B_r^{\mathbb{R}^d}(x)) \geq V_{\tilde{d}}(x)r^{\tilde{d}} + o(r^{\tilde{d}})$$

as $r \downarrow 0$, which concludes the proof. $\square$

C. Modes Lemmas

Lemma C.1 (Ayanbayev et al., 2022, Lemma B.1(a)). *An OM functional $I_\mu: E_\mu \rightarrow \mathbb{R}$ is exhaustive if and only if there is $x_2 \in E_\mu$ such that*

$$\lim_{r \downarrow 0} \frac{\mu(B_r(x_1))}{\mu(B_r(x_2))} = 0$$

for all $x_1 \in \mathbb{X} \setminus E_\mu$.

Lemma C.2. *Let $\mathbb{X}$ be locally compact and let $f: \mathbb{X} \rightarrow \mathbb{R}_{\geq 0}$ be continuous. Define the measure $\nu(dx) := f(x)\mu(dx)$. If $I_\mu: E_\mu \rightarrow \mathbb{R}$ is an OM functional for μ such that E_μ and $\text{supp}^\circ(f) := \{x \in \mathbb{X}: f(x) > 0\}$ have a non-empty intersection, then*

$$I_\nu: E_\nu \rightarrow \mathbb{R}, x \mapsto I_\mu(x) - \log f(x)$$

with $E_\nu := E_\mu \cap \text{supp}^\circ(f)$ is an OM functional for ν . If I_μ is exhaustive, then I_ν is exhaustive as well.

Proof. Since f is continuous and $\mathbb{X}$ is locally compact, f is locally uniformly continuous by the Heine-Cantor theorem. With this, the statement can be proven analogously to Lemma B.8 of Ayanbayev et al. [2022]. $\square$