Differentiable families of measures

O.G. Smolyanov
Faculty of Mechanics and Mathematics
Moscow State University
119899 Moscow, Russia

H. v. Weizsäcker
Fachbereich Mathematik
Universität Kaiserslautern
675 Kaiserslautern, Germany

November 11, 1998

Summary. The paper studies differential and related properties of functions of a real variable with values in the space of signed measures. In particular the connections between different definitions of differentiability are described corresponding to different topologies on the measures. Some conditions are given for the equivalence of the measures in the range of such a function. These conditions are in terms of socalled logarithmic derivatives and yield a generalization of the Cameron-Martin-Maruyama-Girsanov formula. Questions of this kind appear both in the theory of differentiable measures on infinite-dimensional spaces and in the theory of statistical experiments.

1 Introduction

This paper describes some general properties of differentiable functions of a real variable having values in the vector space of signed measures on a σ-algebra on some set. The results point out some analogies between two areas where differentiation of such functions has been used. One is the theory of differentiable measures on infinite-dimensional spaces which was started in 1966 by S.V. Fomin and later extended by several authors (cf. e.g. [Ave71],[Sko74]) and which provides a natural framework for Malliavin calculus. The other area is the study of general one parameter families of measures as it is well established in the theory of statistical experiments (cf. the work of J.Hájek and L. LeCam in the early sixties, see [Tor91]).
In particular we describe in a detailed manner the links between differentiability in the sense of Fomin and in the sense of Skorokhod in the more abstract framework and study the differentiability of the Hahn-Jordan decomposition. Even if one is primarily interested in probability measures the derivative is a signed measure and if one wants to consider higher derivatives one necessarily needs also the derivatives of signed measures. Also we prove a formula expressing the relative density (likelihood ratio) of two members of such a family in terms of the logarithmic derivative (derivative of the log-likelihood-ratios). It shows that exponential families are in a certain sense tangential to general differentiable families. Under stronger assumptions this formula was announced (almost without proof) by Yu. Daletskii and G. Sohadze ([Dal88]), extending a result of D. Bell ([Bel85]) on shifts of differentiable measures on Banach spaces. For the family of shifts of Wiener measure this formula specializes to the Cameron-Martin formula in the case of shifts along constant directions and to a variant of the Maruyama-Girsanov formula (cf. [Wei90] and [Buc91]) in the case of shifts along vector fields just because a stochastic integral with respect to Wiener measure coincides with the logarithmic derivative of Wiener measure along the integrand (this applies even in the nonanticipating case, cf. e.g. the references in [Bog90]). It is worth

mentioning that the logarithmic derivative of a measure ν on a Hilbert (or more general) space Q plays a key role for the definition of a quantum theoretic observable in the space $L^2(Q,\nu)$ (cf. [Smo90],[Alb93]). Let us remark also that our theory may be applied to shifts of measures on loop groups (cf. [Mal90]).
The paper is organized as follows: In section 2 we compare various different notions of differentiability for such a one-parameter family of measures. We show that fairly weak assumptions imply the differentiability for the total variation norm at Lebesgue-almost all parameters (theorem2.7). Section 3 introduces the logarithmic derivative and gives the formula mentioned above (Theorem 3.3). Section 4 connects the differentiability of a family of signed measures to the differentiability of the positive parts which one gets by the Hahn-Jordan decomposition.Section 5 introduces the differentiablity in the p-th mean (which is for p=2 essential in the asymptotic theory of statistical experiments). Section 6 applies the previous results in the case where the family is induced by a measurable flow on the underlying space and or more generally by a operator semigroup. Section 7 discusses the behaviour of differentiability under formation of (infinite) product measures. It rephrases some known results in the present language. Finally section 8 reviews the vector space situation, in particular differentiation of a measure along a vector field. For the sake of simplicity the latter is assumed to induce a continuous flow on the underlying space so that the situation is a particular case of section 6. Also the measure approach to infinite dimensional vector calculus from [Smo86] is briefly indicated. Except for this last part we try to give full proofs.

2 Notions of differentiability for curves of abstract measures

Notations. Throughout the following X is a set, $\mathcal{B}$ is a σ-algebra of subsets of X and $M(X)$ is the (vector) space of all signed measures on $\mathcal{B}$. For $\mu \in M(X)$ the measures μ^+ and μ^- are the positive and the negative part of μ in the Hahn-Jordan-decomposition (such that $\mu = \mu^+ - \mu^-$), $|\mu|$ is the measure $\mu^+ + \mu^-$ and $||\mu||$ is the total variation norm of μ. The space of all bounded $\mathcal{B}$-measurable functions is denoted by $\mathcal{B}_b$. The letter C will denote a subspace of $\mathcal{B}_b$ which is normdefining for $\mathcal{M}(X)$, i.e. $||\mu|| = sup\{\int_X \phi\, d\mu : \phi \in C, ||\phi||_\infty \leq 1\}$ where $||\cdot||_\infty$ is the sup-norm. Of course the most important case is $C = C_b(X)$; the space of bounded continuous functions on a completely regular topological space X. In this case always $\mathcal{B} = \mathcal{B}_0(X)$ is the σ-algebra of all Baire subsets of X (i.e. the σ-algebra generated by the elements of $C_b(X)$).
The symbol τ_s denotes the topology of setwise convergence on the space $\mathcal{M}(X)$, the symbol τ_v - the topology of convergence in the total variation norm, the symbol τ_w - the weak topology on $\mathcal{M}(X)$ defined by the duality between $\mathcal{M}(X)$ and $(\mathcal{M}(X),\tau_v)'$ and the symbol τ_C - the weak topology on $\mathcal{M}(X)$ defined by the duality between $\mathcal{M}(X)$ and C .

Definition 2.1: *Let I be an open interval in $\mathbb{R}$. Let τ be a topology on $\mathcal{M}(X)$ which is compatible with the vector space structure. Let $m : t \longmapsto m(t) \equiv \mu_t$ be a map from I to $\mathcal{M}(X)$. Then m (or the family $(\mu_t)_{t\in I}$) is called τ-differentiable at $t_0 \in I$ if m is differentiable at this point as the map from I into $(\mathcal{M}(X),\tau)$, i.e. if there exists the τ−limit $\lim_{\substack{t\longrightarrow 0\\ t\neq 0}} \frac{m(t_0+t)-m(t_0)}{t}$. This limit is called τ-derivative of m at t_0 and is denoted by $m'(t_0)$ or μ'_{t_0}.*

Remark 2.2 1. Note that the symbol for the derivative does not depend on τ; this is motivated by the fact that if a function m is τ_1 - and τ_2-differentiable at t_0 and the topology τ_1

is finer than τ_2 or just every τ_1-convergent sequence is τ_2-convergent, then τ_1-differentiability implies τ_2-differentiability and the derivatives for these two toplogies coincide.
2. Let X be a topological real linear space and let μ_t is of the form $\mu_t(A) = \mu(A+th)$ for some vector $h \in X$ and some fixed measure μ. This measure μ sometimes is called differentiable in the sense of Fomin if the family (μ_t) is τ_s-differentiable whereas differentiability of μ in the sense of Skorokhod means that (μ_t) is $\tau_{C_b(X)}$-differentiable.
3. A classical theorem of Nikodym, see [Dun58], p.70 and p. 121) implies that $(\mathcal{M}(X), \tau_s)$ is sequentially complete and that a sequence in $\mathcal{M}(X)$ is τ_s-convergent iff it is τ_w-convergent; moreover $\tau_s \subset \tau_{\mathcal{B}_b} \subset \tau_w$. Consequently for a map $m : I \longrightarrow \mathcal{M}(X)$ the following are equivalent:
(a) m is τ_w-differentiable at $t_0 \in I$, (b) m is $\tau_{\mathcal{B}_b}$-differentiable at $t_0 \in I$, (c) m is τ_s-differentiable at $t_0 \in I$, (d) for each $A \in \mathcal{B}$ the function $m_A : I \longrightarrow R^1, t \longmapsto m(t)(A) (\equiv \mu_t(A))$ is differentiable at t_0. Moreover, these conditions imply that m is τ_C-differentiable at t_0 since $C \subset \mathcal{B}_b$.
4. Note that in spite of this equivalence the topologies τ_w and $\tau_{\mathcal{B}_b}$ do not coincide. (They do on bounded sets.)

In general, τ_s-differentiability does not imply τ_v-differentiability, and τ_{C_b}-differentiability does not imply τ_s-differentiability, as the following examples show. In later sections we shall give however some sufficient criteria for these implications to hold.

Example 2.3: Let $X = [0,1], I = (-1,1)$ and μ_t be defined by

$$\mu_t(A) = t \int_A \sin\frac{x}{t}\, dx \; for \; t \neq 0 \; and \; \mu_0 = 0.$$

Then $(\mu_t)_{t\in I}$ is τ_s-differentiable at all points of I, but is not τ_v differentiable at $t_0 = 0$. For future reference note that $\mu_0' = 0$.

Example 2.4: Let $X = (-1,1), I = (-1,1)$ and μ_t be defined by

$$\mu_t(A) = \lambda(A \cap (-1,t]) - t\varepsilon_0(A);$$

here λ is the Lebesgue measure on $[-1,1]$ and ε_0 is the Dirac measure at the point 0. Then $(\mu_t)_{t\in I}$ is $\tau_{C_b(\mathbb{R})}$ -differentiable at each point of I with $\mu_t' = \varepsilon_t - \varepsilon_0$; in fact $\frac{d}{dt}\int \phi d\mu_t = \phi(t) - \phi(0)$ for every $\phi \in C_b(\mathbb{R})$. However for $A = [-1,t_0]$ the map $t \mapsto \mu_t(A)$ is not differentiable at t_0, i.e. (μ_t) is τ_s-differentiable nowhere. Note that even $\mu_0' = 0 \ll \mu_0$.

In the following we suppose that the families (μ_t) under consideration are defined on a fixed open interval $I \subset \mathbb{R}$. Here are some general regularity results

Proposition 2.5: *Let (μ_t) be τ_C-differentiable and suppose either that C is complete under the sup-norm or that the map $t \longmapsto ||\mu_t'||$ is bounded from above by a locally integrable function. Then*
a) (μ_t) is τ_v-continuous.
b) There is a probability measure $\nu \in \mathcal{M}(X)$ such that $\mu_t \ll \nu$ for all $t \in I$. For every such ν one can choose the derivatives $f_t = \frac{d\mu_t}{d\nu}$ such that $f_t(x)$ is a $\mathcal{B}(I) \otimes \mathcal{B}$ - measurable function of (t,x).
c) If (μ_t) is τ_s-differentiable then the measure ν in b) dominates even μ_t' for all t.
d) If (μ_t) is twice τ_C-differentiable and C is complete under the sup-norm then (μ_t) is τ_v-differentiable.

Proof: a) Suppose first that C is complete under the sup-norm. Fix t and $\varepsilon > 0$ such that $(t-\varepsilon, t+\varepsilon) \subset I$. The τ_C-differentiability at t implies that the set $A = \{(\mu_{t+h} - \mu_t)/h : 0 < |h| < \varepsilon\}$ is bounded for τ_C. Hence this set is bounded as a subset of the dual of the Banach space $(C, ||\cdot||_\infty)$ by the Banach-Steinhaus theorem. Since C is normdefining for $\mathcal{M}(X)$ this means that the set A is τ_v-bounded. Thus $||\mu_{t+h} - \mu_t|| = O(h)$ as $h \to 0$.
Suppose now that (μ_t) is τ_C-differentiable and $||\mu_t'|| \leq g(t)$ for some locally integrable function g. Then

$$\begin{aligned} ||\mu_{t+h} - \mu_t|| &= sup\{\int \phi \; d(\mu_{t+h} - \mu_t) : \phi \in C, ||\phi||_\infty \leq 1\} \\ &= sup\{\int_t^{t+h} \int \phi \; d\mu_s' \; ds : \phi \in C, ||\phi||_\infty \leq 1\} \\ &\leq \int_t^{t+h} g(s) \; ds \end{aligned}$$

which converges to 0 as $h \to 0$. This proves a).
b). Choose any probability measure ν which dominates μ_t for all rational t, e.g. $\nu = \sum_{i=1}^\infty c_i |\mu_{t_i}|$ where (t_i) is an enumeration of the rational elements of I and (c_i) is chosen so that $\sum_{i=1}^\infty c_i ||\mu_{t_i}|| = 1$. Then ν dominates every μ_t according to a). Then every choice of the densities f_t defines a stochastic process on the probability space $(X, \mathcal{B}, \nu)$ which is continuous in measure. Thus it has a jointly measurable modification (cf. [Nev69], section 3.4) which proves b).
c) Let ν dominate μ_t for all t and let N be a ν-nullset. Then $\mu_t'(N) = \lim_{h\to 0} \frac{1}{h}(\mu_{t+h}(N) - \mu_t(N)) = 0$.
d) If (μ_t) is twice τ_C-differentiable then (μ_t') is τ_v-continuous according to part a). Therefore it suffices to note the foolowing general fact. If m is a map from an interval I into a Banach space E which is differentiable for the topology τ_C of pointwise convergence on C where C is a normdefining subspace of the dual E' and if m' is normcontinuous then m is normdifferentiable. In fact, by the meanvalue theorem

$$\frac{1}{h}(m(t+h) - m(t)) - m'(t) \in \overline{conv}\{m'(\theta) - m'(t) : \theta \in (t-h, t+h)\}$$

. where $\overline{conv}$ denotes the τ_C-closed convex hull. Since C is normdefining we conclude that

$$||\frac{1}{h}(m(t+h) - m(t)) - m'(t)|| \leq sup\{||m'(\theta) - m'(t)|| : \theta \in (t-h, t+h)\}$$

which converges to 0 as $h \to 0$ because m' is normcontinuous. ∎

Remark 2.6 Let us consider the map $m' : t \longmapsto \mu_t' \in \mathcal{M}(X)$. Its range is in the τ-closure of the linear span of the set $\{\mu_t : t \in I\}$, the latter being norm-separable according to proposition 2.5a). If $\tau = \tau_C$ this does not imply that the range of m' is norm-separable as the example 2.4 shows. In the case of τ_s- or equivalently τ_w-differentiability (cf. remark 2.2) this closure is in fact norm-separable and hence m' is not only weakly measurable but also strongly measurable. In particular in this case $||\mu_t'||$ clearly is a measurable function of t. These measurability questions are tacitly surrounded in the proof of the next next result by passing from the space $\mathcal{M}(X)$ to the space $L^1(\nu)$.

Now we come to our first main result:

Theorem 2.7: *Suppose that either*
i) the family $(\mu_t)_{t\in I}$ *is* τ_s*-differentiable or*

ii) the family $(\mu_t)_{t\in I}$ is τ_C -differentiable and $\mu'_t \ll \mu_t$ for Lebesgue almost all $t \in I$.
Assume further that $||\mu'_t|| \leq g(t)$ on $[a,b]$ for some a,b in I and some $g \in L^1[a,b]$. Then
a) For all $t \in [a,b]$

$$\mu_t - \mu_a = \int_a^t \mu'_s \, ds \tag{1}$$

as a $(\mathcal{M}(X), ||.||)$-valued Bochner integral.
b) For Lebesgue almost all points $t \in [a,b]$, $\mu'_t \ll \mu_t$ and the family (μ_t) is τ_v-differentiable at t.
c) There are a probability measure ν and two $\mathcal{B}([a,b]) \otimes \mathcal{B}$-measurable functions f, f' such that $f_t = \frac{d\mu_t}{d\nu}$ for all t, $f'_t = \frac{d\mu'_t}{d\nu}$ for almost all t, and

$$f_t(x) - f_a(x) = \int_a^t f'_s(x) \, ds. \tag{2}$$

holds for all $x \in X$ and all $t \in [a,b]$.

Note that in part b) the absolute continuity statement already has been postulated if only τ_C-differentiability is assumed.

Proof: Choose a measure ν according to proposition 2.5 b) and denote by λ the Lebesgue measure on $[a,b]$. Define a measure m' on $\mathcal{B}(I) \otimes \mathcal{B}$ by $m'(dt,dx) = \mu'_t(dx)dt$. Then $||m'|| = \int_a^b ||\mu'_t||dt < \infty$. A straightforward Fubini argument, using proposition 2.5 c) in the case of τ_s-differentiability, and the additional assumption $\mu'_t \ll \mu_t$ *a.e.* (t) in the case of τ_w-differentiability, shows that $m' \ll \lambda \otimes \nu$. Let $f' \in L^1(\lambda \otimes \nu)$ be a version of $\frac{dm'}{d\lambda \otimes \nu}$. Let f be jointly measurable as in proposition 2.5 b).
Choose $\phi \in C$ in the case of τ_C-differentiability or $\phi = 1_B, B \in \mathcal{B}$ in the case of τ_s-differentiability. Then $t \mapsto \int \phi d\mu_t$ is a differentiable function with integrable derivative over $[a,b]$. So Fubini's theorem and the fundamental theorem of calculus give

$$\begin{aligned}
\int_X \phi(x) \int_a^b f'_t(x) \, dt \; \nu(dx) &= \int_{[a,b]\times X} \phi(x) f'_t(x) \; \lambda \otimes \nu(dt,dx) \\
&= \int_{[a,b]\times X} \phi(x) \; m'(dt,dx) \\
&= \int_a^b \int_X \phi(x) d\mu'_t \; dt \\
&= \int_X \phi(x) \; d(\mu_b - \mu_a) \\
&= \int_X \phi(x)(f_b(x) - f_a(x)) \; \nu(dx).
\end{aligned}$$

for all ϕ as above. Hence (2) holds ν-a.e.. Now we redefine f by $f_t(x) = \int_a^t f'_s(x) \, ds$. Then this changes each section f_t only on a ν-nullset and hence still $f_t = \frac{d\mu_t}{d\nu}$. But with this new f, equation (2) holds everywhere.
It is known that the Banach space $L^1(\lambda \otimes \nu)$ can be identified with the space $L^1_{L^1(\nu)}([a,b])$ of (equivalence classes of) $L^1(\nu)$-valued Bochner integrable functions on $[a,b]$ and thus the equation (2) implies

$$f_t - f_a = \int_a^t f'_s \, ds \tag{3}$$

where the right hand side is a Bochner integral. The Lebesgue differentiation theorem for Bochner integrals (cf. e.g. [Die77],p. 49) yields that for Lebesgue almost all t, $\lim_{h\to 0}\frac{f_{t+h}-f_t}{h} = f'_t$ in the $L^1(\nu)$-norm. This implies that the limit $\lim_{h\to 0}\frac{\mu_{t+h}-\mu_t}{h}$ exists in τ_v and is given by the measure with ν-density f'_t. On the other hand by the weaker differentiability of our assumption this measure must be μ'_t. Thus $f'_t = \frac{d\mu'_t}{d\nu}$ for almost all t and (3) implies the equation (1) .
It remains to prove the statement of absolute continuity in the case of τ_s-differentiability. For this we compare the zero sets of f and f'. For each x the set $\{t : f_t(x) = 0\}$ is contained in the set $\{t : f'_t(x) = 0\}$ up to a Lebesgue nullset: Indeed, for almost every t the function $f_.(x)$ is differentiable at t with derivative $f'_t(x)$ by (2) and Lebesgue differentiation, thus for almost all density points and hence for almost all points t of the zero set of $f_.(x)$ we have $f'_t(x) = 0$. By Fubini's theorem then for Lebesgue almost all t the set $\{x : f_t(x) = 0\}$ is contained in the set $\{x : f'_t(x) = 0\}$ up to a ν-nullset. But this means that for such a t we have $\mu'_t \ll \mu_t$. ∎

In Proposition 5.1 we shall prove a similar result for yet another concept of differentiability for families of nonnegative finite measures, the L^p-differentiability.

3 Logarithmic derivatives

In theorem 2.7b) we saw that for a τ_s-differentiable family (μ_t) most of the measures μ'_t are absolutely continuous with respect to the corresponding μ_t. The following simple but perhaps surprising observation shows that for nonnegative families of measures there is no exceptional t.

Proposition 3.1: *Let (μ_t) be τ_s-differentiable at t_0 . If $\mu_t \geq 0$ for all t in a neighborhood of t_0 then $\mu'_{t_0} \ll \mu_{t_0}$.*

Proof: Let $N \in \mathcal{B}$ be a μ_{t_0}-nullset. Since $\mu_t(N) \geq 0$ for all t near t_0 the value $\mu_{t_0}(N)$ is a local minimum and hence $\mu'_{t_0}(N) = \frac{d}{dt}\mu_t(N)|_{t=t_0} = 0$. ∎

For general signed measures this is not true: For every measure μ^* the family $\mu_t = t\mu^*$ is even τ_v-differentiable at every point t with $\mu'_t = \mu^*$. But for $t_0 = 0$ we have $\mu_0 \equiv 0$ and hence $\mu'_0 \not\ll \mu_0$. In theorem 4.2 below we shall give a general criterion for $\mu'_{t_0} \ll \mu_{t_0}$ in the case of signed measures. First, we introduce the concept of logarithmic derivative.

Definition 3.2 *Let (μ_t) be τ-differentiable at t_0 for one of the topologies τ on $\mathcal{M}(X)$ considered in definition 2.1. Suppose that $\mu'_{t_0} \ll \mu_{t_0}$. Any version of the Radon-Nikodym derivative $\rho_{t_0} = \frac{d\mu'_{t_0}}{d\mu_{t_0}}$ is called logarithmic derivative of (μ_t) at t_0.*

The name is motivated by the following observation. Suppose the family (μ_t) is given by positive densities f_t with respect to some $\nu \in \mathcal{M}(X)$ such that $\frac{\partial}{\partial t}f_t(x)$ is continuous in t and uniformly ν-integrable in x. Then

$$\mu'_t(A) = \int_A \frac{\partial}{\partial t}f_t(x)d\nu = \int_A \frac{\frac{\partial}{\partial t}f_t}{f_t}f_t d\nu = \int_A \frac{\partial}{\partial t}ln f_t(x)d\mu_t(x)$$

and hence ν-a.e.

$$\rho_t(x) = \frac{\partial}{\partial t}ln f_t(x). \tag{4}$$

Actually this equation can be justified even under weaker assumptions. If for some t one has $\mu'_t \ll \mu_t$ then for the density $f'_t = \frac{d\mu'_t}{d\nu}$

$$f_t(x)\rho_t(x) = f'_t(x), \tag{5}$$

holds μ_t-a.e. and even ν-a.e.. But theorem 2.7 part b) shows that under the assumption of that theorem the densities f_t and f'_t can be chosen so that the integral realtion (2) holds, i.e. $f_\cdot(x)$ is absolutely continuous in t with a.e. derivative $f'_\cdot(x)$. Thus in (5) the right hand side can be read as a derivative and therefore (4) holds at least for $\lambda \otimes \nu$-almost all $(t, x) \in [a, b] \times X$ for which $f_t(x) \neq 0$.

The following is the formula mentioned in the introduction. It may be considered as the solution of the measure valued differential equation $\mu' = \rho\mu$. Note that the existence of logarithmic densities, i.e. the condition $\mu'_t \ll \mu_t$ a.e.(t) on $[a, b]$ is automatic in the case of τ_s-differentiability according to theorem 2.7b) whereas in the case of τ_C-differentiability this condition has to be postulated. One way to look at the result is to say that exponential families of probability distributions, i.e. families (μ_t) of the form $\frac{d\mu_t}{d\nu}(x) = c(t)e^{tT(x)}$ are in some sense tangential approximations to more general differentiable families. The formula also shows that under the assumptions of the theorem the Hahn-Jordan decomposition of the measures does not change in the corresponding interval.

Theorem 3.3: *Let $(\mu_t)_{t\in I}$ be differentiable for τ_s or τ_C and suppose that $\int_a^b ||\mu'_t||\ dt < \infty$ for some a, b in I. Let the map $(t, x) \longmapsto \rho_t(x)$ be $\mathcal{B}(I) \otimes \mathcal{B}$ - measurable such that for Lebesgue-almost all t, $\rho_t = \frac{d\mu'_t}{d\mu_t}$. If $\int_a^b |\rho_t(x)|\ dt < \infty$ holds $|\mu_a| + |\mu_b|$-almost everywhere then all measures μ_t, $a \leq t \leq b$ are equivalent and*

$$\frac{d\mu_b}{d\mu_a}(x) = e^{\int_a^b \rho_t(x)dt}. \tag{6}$$

Proof: Instead of a differential equation we consider the corresponding integral equation and employ the following fact: Let $g \in L^1[a, b]$ and $h \in C[a, b]$ satisfy $h(s) = \int_a^s h(t)g(t)dt$ for all $s \in [a, b]$. Then $h(b) = h(a)exp(\int_a^b g(t)dt)$. In fact if $h^*(s) = h(a)exp(\int_a^s g(t)dt)$ then Gronwall's lemma easily implies $h = h^*$.
According to theorem 2.7 c) and Fubini's theorem there are a dominating probability measure ν, and two $\mathcal{B}([a, b]) \otimes \mathcal{B}$-measurable functions f, f' such that for ν-almost all $x \in X$ (5) holds Lebesgue-a.e.. Using (5) in the integral of (2) we get for ν-almost all x and and all t the relation

$$f_t(x) - f_a(x) = \int_a^t f_s(x)\rho_s(x)\ ds. \tag{7}$$

Moreover our assumption implies that for ν-almost every x either $f_a(x) = f_b(x) = 0$ or $\int_a^b |\rho_s(x)|ds < \infty$. Together with our initial remark and (7) this shows that

$$f_b(x) = f_a(x) \cdot exp(\int_a^b \rho_t(x)dt) \tag{8}$$

for ν-almost all x for which either $f_a(x) \neq 0$ or $f_b(x) \neq 0$. In particular f_a and f_b vanish on ν-almost the same points. This implies that the measures μ_a and μ_b are equivalent and

$$\frac{d\mu_b}{d\mu_a}(x) = \frac{f_b}{f_a}(x) = exp(\int_a^b \rho_t(x)dt)\ \ \mu_a - a.e..$$

■

It is not true, however, that the equivalence of the measures μ_t implies the pathwise integrability of the logarithmic density nor is it true that the measures μ_t are always equivalent even if $\mu'_t \ll \mu_t$ for all t.

Example 3.4: Let $\Omega = I = (-1,1)$. Let $f_t(x) = |x-t|^\alpha$ for some positive α. Then $f'_t(x) = sgn(t-x)\alpha|x-t|^{\alpha-1}$ for $t \neq x$. It can be verified that the family $\{\frac{f_{t'}-f_t}{t'-t} : t \neq t'\}$ is uniformly integrable. Thus $\lim_{t'\to t} \frac{f_{t'}-f_t}{t'-t} = f'_t$ in $L^1(\Omega,\lambda)$ and hence $d\mu_t = f_t d\lambda$ defines a τ_v-differentiable family with $d\mu'_t = f'_t d\lambda$. Then $\frac{d\mu'_t}{d\mu_t}(x) = \frac{f'_t}{f_t}(x) = \alpha(t-x)^{-1}$ for $t \neq x$. Thus the logarithmic density is not pathwise integrable but all μ_t are equivalent to Lebesgue measure. On the other hand for $f^<_t(x) = f_t(x)1_{\{t<x\}}$ and $f^>_t(x) = f_t(x)1_{\{t>x\}}$ the resulting families $(\mu^<_t)$ and $(\mu^>_t)$ have supports which decrease and increase strictly as a function of t, respectively. Nevertheless $\mu^{<'}_t \ll \mu^<_t$ and $\mu^{>'}_t \ll \mu^>_t$ as can be seen by direct inspection or by proposition 3.1.

4 Differentiability and the Hahn-Jordan decomposition

Next, we return to the absolute continuity $\mu'_{t_0} \ll \mu_{t_0}$ at a given point t_0. From proposition 3.1 a sufficient condition is the τ_s-differentiability of the positive and the negative parts (μ^+_t) and (μ^-_t) of the family at t_0. We show that this condition is not necessary in the case of τ_s-differentiable (μ_t) but it is necessary in the case of τ_v-differentiable (μ_t).

Example 4.1 We modify the example 2.3 (which was τ_s-differentiable with $\mu'_0 \ll \mu_0$ but not τ_v-differentiable at $t = 0$) in such a way that in addition the positive and negative parts are not $\tau_s - differentiable$ at $t = 0$. For this let g be a bounded function on $I = (-1,1)$, differentiable for $t \neq 0$ but without a limit as $t \to 0$. Let $\mu_0 = 0$ and for $t \neq 0$ let μ_t be the measure with Lebesgue density $f_t(x) = t(sin\frac{x}{t})g(t)$. If (t_n) is a null sequence such that $\lim_n g(t_n) = \alpha$ exists then the sequences $(f^+_{t_n}/t_n)$ and $(f^-_{t_n}/t_n)$ both have the $\sigma(L^1, L^\infty)$-limit $\alpha/2$ and hence their difference (f_{t_n}/t_n) converges weakly to 0 independently of α. Thus (μ_t/t) converges to $\mu_0 = 0$ in τ_s and $\mu'_0 = 0 \ll \mu_0$, whereas each different limit point α of $g(t)$ for $t \to 0$ gives a different limit point of (μ^+_t/t), i.e. (μ^+_t/t) is not differentiable at 0.

Theorem 4.2: *Let $(\mu_t)_{t\in I}$ be a family of elements of $\mathcal{M}(\Omega)$. For every $t_0 \in I$ the following are equivalent:*

1. *(μ_t) is τ_v-differentiable at t_0 and $\mu'_{t_0} \ll \mu_{t_0}$*
2. *the positive and the negative parts (μ^+_t) and (μ^-_t) both are τ_v-differentiable at t_0.*

In this case $(\mu^\pm_t)'(E) = (\mu_t)'(A^\pm_t \cap E)$ where $\Omega = A^+_t \cup A^-_t$ is a Hahn-Jordan decomposition with respect to μ_t.

Proof: 2. $\Longrightarrow$ 1. If (μ^+_t) and (μ^-_t) are τ_v- differentiable at t_0 then $\mu^{+'}_{t_0} \ll \mu^+_{t_0}$ and $\mu^{-'}_t \ll \mu^-_{t_0}$ according to proposition 3.1. This implies the τ_v-differentiability of $\mu_t = \mu^+_t - \mu^-_t$ at t_0 and $\mu'_{t_0} = \mu^{+'}_{t_0} - \mu^{-'}_{t_0} \ll \mu_{t_0}$.
1. $\Longrightarrow$ 2. For simplicity of notation we assume $t_0 = 0$. Let $\Omega = A^+_t \cup A^-_t$ be a Hahn-Jordan-decomposition with respect to μ_t, i.e. $\mu^\pm_t(\mathrm{E}) = \mu_t(\mathrm{A}^\pm_t \cap \mathrm{E})$. Let $B_t = A^+_t \backslash A^+_0$ and $C_t = A^+_0 \backslash A^+_t = A^-_t \backslash A^-_0$. Then

$$\frac{(|\mu_0| + |\mu_t|)(B_t)}{t} = \frac{\mu_t(B_t) - \mu_0(B_t)}{t} \leq ||\frac{\mu_t - \mu_0}{t}||$$

and hence according to the definition of τ_v-differentiability

$$\lim_{t\to 0} \frac{(|\mu_0|+|\mu_t|)(B_t)}{t} \le ||\mu_0'|| \tag{9}$$

We claim that with the help of the assumption $\mu_0' \ll \mu_0$ this implies

$$\lim_{t\to 0} \frac{|\mu_t|(B_t)}{t} = 0. \tag{10}$$

We check this for a sequence (t_n) converging to zero. We may assume $\sum_1^\infty |t_n| < \infty$. Let $D_m = \bigcup_{n=m}^\infty B_{t_n}$. Because of (9) we have $\lim_{m\to\infty} |\mu_0|(D_m) = 0$ and therefore $\mu_0' \ll \mu_0$ implies $|\mu_0'|(D_m) \longrightarrow 0$. For $\varepsilon > 0$ choose m_ε such that $|\mu_0'|(D_{m_\varepsilon}) < \varepsilon$. Write $D = D_{m_\varepsilon}$ and restrict everything to the measurable space $(D, \mathcal{B} \cap D)$. Then $(\tilde\mu_t) = (\mu_t(D \cap \cdot))$ is differentiable at 0 with $\tilde\mu_0' = \mu_0'(D \cap \cdot)$ and $D = (A_t^+ \cap D) \cup (A_t^- \cap D)$ is a Hahn-Jordan decomposition of $\tilde\mu_t$. Thus for $n \ge m$ the set $B_{t_n} = B_{t_n} \cap D = A_{t_n}^+ \cap D \backslash A_0^+ \cap D$ plays the same role for $\tilde\mu_{t_n}$ as for μ_{t_n}. Hence we can apply (9) to $(\tilde\mu_t)$ to get

$$\lim_{n\to\infty} \frac{|\mu_{t_n}|(B_{t_n})}{t_n} \le \lim_{n\to\infty} \frac{(|\tilde\mu_0| + |\tilde\mu_{t_n}|)(B_{t_n})}{t_n} \le ||\tilde\mu_0'|| < \varepsilon.$$

This completes the proof of (10). The set C_t is for $(-\mu_t)$ what B_t is for (μ_t). Hence applying (10) to $(-\mu_t)$ yields

$$\lim_{t\to 0} \frac{|\mu_t|(C_t)}{t} = 0. \tag{11}$$

Now we have

$$\begin{aligned}
\mu_t^+ - \mu_0^+ &= \mu_t(A_t^+ \cap \cdot) - \mu_0(A_0^+ \cap \cdot) \\
&= \mu_t(A_0^+ \cap \cdot) - \mu_0(A_0^+ \cap \cdot) + \mu_t((A_t^+ \backslash A_0^+) \cap \cdot) - \mu_t((A_0^+ \backslash A_t^+) \cap \cdot) \\
&= \mu_t(A_0^+ \cap \cdot) - \mu_0(A_0^+ \cap \cdot) + \mu_t(B_t \cap \cdot) - \mu_t(C_t \cap \cdot)
\end{aligned}$$

and hence because of (10) and (11)

$$\lim_{t\to 0} ||\frac{\mu_t^+ - \mu_0^+}{t} - \mu_0'(A_0^+ \cap \cdot)|| \le \lim_{t\to 0} ||\frac{\mu_t(B_t \cap \cdot)}{t}|| + ||\frac{\mu_t(C_t \cap \cdot)}{t}|| = 0.$$

This shows that (μ_t^+) is differentiable at 0 and $\mu_0^{+\prime} = \mu_0'(A_0^+ \cap \cdot)$. Similary $\mu_0^{-\prime} = \mu_0'(A_0^- \cap \cdot)$. ■

Remark 4.3 D. Fremlin pointed out to us that this result and its proof can be reformulated for Dedekind complete Banach-lattices with an order continuous norm instead of $(\mathcal{M}(\Omega), ||.||)$.

5 Differentiability in the p-th mean

Another context in which the logarithmic derivative is useful is the differentiability in the p-th mean (cf. [Tor91],p. 536 ff. for p=1 and p=2). For the sake of simplicity we assume that all elements of the family (μ_t) are absolutely continuous with respect to some measure ν. In contrast to what is suggested in some statistical texts this is not a real restriction according to the argument in the proof of proposition 2.5b).

Definition 5.1: *Let $p \geq 1$. Suppose that $0 \leq \mu_t \ll \nu$ for some nonnegative measure ν on $\mathcal{B}$ and all $t \in I$. Let $f_t = \frac{d\mu_t}{d\nu}$. Then the family (μ_t) is called L^p-differentiable if the map $F_p : I \longrightarrow L^p(\nu), t \mapsto f_t^{\frac{1}{p}}$ is differentiable for the L^p-norm at all points of I. In the case $p = 2$ the number $||F_2'(t)||_2^2$ is called Fisher-Information of the family (μ_t) at the point t. A family (μ_t) of signed measures is L^p-differentiable if its positive and negative parts have this property.*

According to theorem 4.2 L^1-differentiability is equivalent to τ_v-differentiability with $\mu_t \ll \mu_t$. Part b) of the following proposition is an analogue of theorem 2.7 in this situation.

Proposition 5.2: *a) If for some $p > 1$ the family (μ_t) is L^p-differentiable at $t_0 \in I$ then it is τ_v-differentiable at this point and the logarithmic derivative ρ_{t_0} is in $L^p(\mu_{t_0})$.*
b) Conversely let (μ_t) be τ_C- differentiable with $\mu_t' \ll \mu_t$ for Lebesgue-almost all t such that $\int_a^b ||\rho_t||_{p,\mu_t} dt < \infty$ for all $a, b \in I$. Then (μ_t) is L^p-differentiable at Lebesgue almost all t.
c) Suppose $\mu_t \geq 0$ for all t. Then under the assumption in a) we have $F_p'(t_0) = \frac{1}{p} f_{t_0}' f_{t_0}^{1/p-1} = \frac{1}{p} f_{t_0}^{1/p} \rho_{t_0}$ where $f_t' = \frac{d\mu_t'}{d\nu}$ and under the assuption of b) we have

$$f_b^{1/p} - f_a^{1/p} = \int_a^b \frac{1}{p} f_t^{1/p} \rho_t dt \tag{12}$$

ν-a.e. and as a $L^p(\nu)$-valued Bochner integral.

Proof: Let us first assume that the measures μ_t are nonnegative.
a) Since $f_t = S_p(F_p(t))$ where S_p is as in lemma 5.3 below this lemma shows by the chain rule the first part of a) and the first formula in c). Since $F_p'(t) \in L^p(\nu)$ we have

$$\int |\rho_t|^p d\mu_t = \int |\rho_t|^p f_t \; d\nu = \int |f_t'|^p f_t^{1-p} \; d\nu = \int p^p |F_p'(t)|^p \; d\nu < \infty,$$

i.e. $\rho_t \in L^p(\mu_t)$.
b) Since $||f_t^{1/p} \rho_t||_{p,\nu} = ||\rho_t||_{p,\mu_t}$ our integrability assumption and the measurability considerations in theorem 2.7 imply that $t \mapsto \frac{1}{p} f_t^{1/p} \rho_t$ defines a Bochner integrable map from $[a, b]$ to $L^p(\nu)$. On the other hand by one dimensional change of variable we infer from (7) that

$$f_b^{1/p}(x) - f_a^{1/p}(x) = \int_a^b \frac{1}{p} f_t(x)^{1/p} \rho_t(x) dt$$

ν-a.e.. Then this identity must hold also in the Bochner sense (12). The rest is as in Theorem 2.7 Lebesgue's differentiation theorem for Bochner integrals.
Thus c) is proved and we have to extend a) and b) to the case of signed measures. For a) this is straightforward from the definitions and prop 3.1. For b) we first note that $||\mu_t'|| = ||\rho_t||_{1,\mu_t} \leq ||\mu_t|| \; ||\rho_t||_{p,\mu_t}$ on $[a, b]$. Then Gronwall's lemma shows that $||\mu_t||$ is bounded on this interval. Therefore by this estimate the family (μ_t) satisfies the assumptions of theorem 2.7 and hence it is τ_v-differentiable a.e.. By theorem 4.2 this extends to the negative and positive part of (μ_t) and hence we can apply b) to them to get the desired result. ∎

Lemma 5.3: *Let ν be a measure. For $p \geq 1$ the map $S_p : y \mapsto |y|^p$ from $L^p(\nu)$ to $L^1(\nu)$ is Fréchet-differentiable with $DS_p(y)(z) = p|y|^{p-1}z$.*

Proof: For the reader's convenience we give a proof. We check for a sequence (z_n) converging to 0 in L^p that

$$\int ||y+z_n|^p - |y|^p - p|y|^{p-1}z_n|\; d\nu = o(||z_n||_p) \tag{13}$$

We may assume that $\sum_{n=1}^{\infty} ||z_n||_p < \infty$. Let $z = \sum_{n=1}^{\infty} |z_n|$, this series being convergent in L^p. Using the scalar mean value theorem the integrand on the left hand side of (13) can be estimated from above by

$$|z_n| \sup_{|y^*-y|\leq|z_n|} ||y^*|^{p-1} - |y|^{p-1}| \tag{14}$$

The *sup* is dominated by the function $(2(|y|+|z|))^{p-1}$ which is in $L^{p'}$ where $p' = \frac{p}{p-1}$ is the dual exponent to p. On the other hand the *sup* converges a.e. to 0 since the z_n do so. Thus by dominated convergence this *sup* converges to 0 in $L^{p'}$ and hence Hölder's inequality shows that the 1-norm of (14) is $o(||z_n||_p)$ which proves (13).
■

6 Flows and semigroups

In this section we discuss the special case in which the family (μ_t) is induced by a measurable flow (γ_t) on X, i.e. $\mu_t = \mu_0 \circ \gamma_t^{-1}$, or more generally by a linear operator semigroup $(P_t)_{t\geq 0}$ on $M(X)$. The first result shows that in this situation τ_s-differentiability at some parameter implies continuous τ_v-differentiability at all larger points.

Theorem 6.1: *Let $(P_t : t \geq 0)$ be a semigroup of linear operators in the Banach space $(E, ||.||)$. Suppose that the operator norm $||P_t||$ is bounded on bounded t-intervals. If for $e \in E$ the orbit $m : t \longmapsto P_t e$ is weakly differentiable at some point $t_0 > 0$ then it is continously norm-differentiable. In the case $E = M(\Omega)$ the weak topology can be substituted by τ_s.*

Proof: In view of remark 2.2 2. it suffices to prove the first abstract part of the theorem. From the semigroup property it follows for $t_1 > t_0$ that

$$\lim_{h\to 0}\frac{1}{h}(P_{t_1+h}e - P_{t_1}e) = \lim_{h\to 0} P_{t_1-t_0}\frac{1}{h}(P_{t_0+h}e - P_{t_0}e) = P_{t_1-t_0}m'(t_0)$$

weakly. So the orbit is weakly differentiable at t_1 and $m'(t_1) = P_{t_1-t_0}m'(t_0)$ for all $t_1 \in I$. This implies in particular that m is norm continuous on I. In fact, as in the proof of proposition 2.5 a) $||m(t+h) - m(t)|| = O(h)$ as $h \to 0$ for each $t > t_0$, .
Let E_0 be the closed subspace of E generated by the set $\{P_t x : t \in I\}$. The norm-continuity of m together with the local boundedness of $||P_t||$ implies that the restriction of our semigroup to E_0 is a strongly continuous semigroup. Moreover $\mu'_{t_0} \in E_0$. Thus $t \mapsto \mu'_t = P_{t-t_0}\mu'_{t_0}$ is continuous. Now the argument is finished by the remark in the proof of proposition 2.5d).
■

Except for the continuity of the derivative the result for measures could also be deduced from theorem 2.7 since it is easy to remove the exceptional set of parameters with the help of the semigroup property. However as far as absolute continuity of the derivative is concerned, the exceptional set cannot be removed completely: Let P_t be the rotation by an angle $t \bmod 2\pi$ on $\mathbb{R}^2 = \mathcal{M}(\{1,2\})$. Then for the orbit m of the vector $x = (1,0)$ we have $m(2\pi k) = x$ and $m'(2\pi k) = (0,1)$ for all integers k and hence $m'(t) \not\ll m(t)$ infinitely often. Whether such an example with a Markovian semigroup exists we do not know. If (P_t) is induced by a measurable flow we have indeed

Theorem 6.2: *Let $(\gamma_t : t \in \mathbb{R})$ be a group of bimeasurable bijections of $(\Omega, \mathcal{B})$. Let $\mu \in \mathcal{M}(\Omega)$ and consider $\mu_t(B) = \mu(\gamma_t^{-1}(B))$. Suppose that for some point s, either*

1. *the family (μ_t) is τ_s-differentiable at s or*
2. *the family (μ_t) is τ_C-differentiable at s with $\mu'_s \ll \mu_s$ and $\phi \circ \gamma_t \in C$ for all $t \in \mathbb{R}$ and all $\phi \in C$.*

Then the family itself and its positive and negative parts are τ_v-differentiable on $I = \mathbb{R}$. Moreover $\mu'_t \ll \mu_t$ and $\rho_t(x) = \rho_0(\gamma_{-t}(x))$ μ_t-a.e.for all t. For every b one has the formula

$$\frac{d\mu_b}{d\mu}(x) = e^{\int_0^b \rho_0(\gamma_{-t}(x))dt}. \tag{15}$$

provided the integral in the exponent exists μ-a.e..
The family (μ_t) is L^p-differentiable if and only if $\rho_0 \in L^p(\mu_0)$.

Proof: First let us show that in any of the topologies τ_C, τ_s, τ_v, for all t the derivative μ'_t exists and is equal to $\mu'_s \circ \gamma_{t-s}^{-1}$, whenever μ'_s exists for some s. In fact choose $\phi \in C$ or $\phi = 1_B$. From the group property of the flow it follows that

$$\lim_{h\to 0} \frac{1}{h} \int \phi d(\mu_{t+h} - \mu_t) = \lim_{h\to 0} \frac{1}{h} \int \phi \circ \gamma_{t-s} d(\mu_{s+h} - \mu_s) = \int \phi \circ \gamma_{t-s} d(\mu'_s) = \int \phi d(\mu'_s \circ \gamma_{t-s}^{-1})$$

which implies this assertion, at least for τ_s and τ_C. Moreover the total variation norm is invariant under the flow and hence we can also include τ_v and moreover the norm $||\mu'_t||$ is locally integrable. Also $\mu'_s \ll \mu_s$ for some s implies $\mu'_{s+t} \ll \mu_{s+t}$ for all t. Theorem 2.7 now shows under both of our assumptions that $\mu'_t \ll \mu_t$ for some and hence for all t and that the family is τ_v-differentiable somewhere and hence everywhere. Finally theorem 4.2 then implies the differentiability of the positive and negative part. For the verification of the formula of the logarithmic derivatives let ϕ be a bounded measurable function (resp in C). Then we have

$$\int \phi \; d\mu_t^{'} = \int \phi \circ \gamma_t \; d\mu_0^{'} = \int (\phi \circ \gamma_t)\rho_0 \; d\mu_0 = \int (\phi(\rho_0 \circ \gamma_{-t})) \circ \gamma_t \; d\mu_0 = \int \phi(\rho_0 \circ \gamma_{-t}) \; d\mu_t.$$

The formula (15) now follows from theorem 3.3.
If the family (μ_t) is L^p-differentiable then $\rho_0 \in L^p(\mu_0)$ according to proposition 5.2a). Conversely, if $\rho_0 \in L^p(\mu_0)$ then the same is true for the positive and the negative part of the family. Therefore we may assume $\mu \geq 0$. We have $\rho_t \circ \gamma_t \in L^p(\mu_0)$ then $\rho_t \in L^p(\mu_0 \circ \gamma_t^{-1}) = L^p(\mu_t)$ and $||\rho_t||_{r,\mu_t} = ||\rho_0||_{r,\mu_0}$ for all t. In particular the map $t \longmapsto ||\rho_t||_{r,\mu_t}$ is locally integrable which implies the L^p-differentiability by proposition 5.2b) somewhere and hence everywhere. ■

7 Differentiability of products

For product measures the L^2-differentiability is particularly appropriate as has been noticed already 1962 by J. Hajek [Haj62] and L.LeCam. We refer to [LeC90],chapter 6 for an introduction to local asymptotic normality of statistical experiments. There the results are based on an application of central limit theorems to the logarithmic derivative. Here we collect some elementary statements using Kolmogorov's "Three Series Theorem".Let $(X_n, \mathcal{B}_n)_{n\in\mathbb{N}}$ be a sequence of measurable spaces and let I be a real interval. For each $n \in \mathbb{N}$ let $(\mu_{t,n})_{t\in I}$ be a L^2-differentiable family of probability measures on $(X_n, \mathcal{B}_n)$ and for each $n \in \mathbb{N} \cup \{\infty\}$ let μ_t^n be the n-fold product measure $\bigotimes_{i=1}^n \mu_{t,i}$. Let C be the space of all random variables on X^∞ of the form $\phi = g(x_1, ..., x_n)$ where $n \in \mathbb{N}$ and $g \in (\bigotimes_{i=1}^n \mathcal{B}_i)_b$.

Theorem 7.1: *a) For finite n the family $(\mu_t^n)_{t\in I}$ is L^2-differentiable and its logarithmic derivative ρ_t^n is given by $\rho_t^n(x_1,...,x_n) = \sum_{i=1}^n \rho_{t,i}(x_i)$.*
b) The family $(\mu_t^\infty)_{t\in I}$ is τ_C-differentiable at t if and only if

$$\sum_{i=1}^{\infty} ||\rho_{t,i}||^2_{2,\mu_{t,i}} < \infty.$$

In this case the logarithmic derivative exists and is given by the $L^2(\mu_t^\infty)$-convergent series $\sum_{i=1}^\infty \rho_{t,i}(x_i)$.
c) Suppose $X_n = X$ for all n, $I = \mathbb{R}$ and $\mu_{t,i} = \mu \circ \gamma_{th_i}$ for some sequence (h_i) of real numbers and a fixed group $(\gamma_t)_{t\in\mathbb{R}}$ of bimeasurable bijections of X. Then the following are equivalent:

- *(μ_t^∞) is τ_C-differentiable at some point t.*
- *(μ_t^∞) is L^2-differentiable at all points t.*
- *The sequence (h_i) is square-summable.*

Proof: Let ν_i be a dominating measure for $(\mu_{t,i})$ and let ν^n be the product measure $\bigotimes_{i=1}^n \nu_i$ for $n \in \mathbb{N}\cup\{\infty\}$. For the proof of a) note that for finite n we have $ln\frac{d\mu^n}{d\nu^n}(x_1,...,x_n) = \sum_{i=1}^n lnf_i(x_i)$ and this formula carries over to the logarithmic derivatives.
b) The sum $\sum_{i=1}^n \rho_{t,i}(x_i)$ is a sum of independent random variables with expectation 0 and variances $\sigma_i^2 = ||\rho_{t,i}||^2_{2,\mu_{t,i}}$ with respect to the measures μ_t^n and μ_t^∞. Suppose that the sum in b) is finite. Then the series $\sum_{i=1}^\infty \rho_{t,i}(x_i)$ is converging in $L^2(\mu_t^\infty)$ by the Three Series Theorem and the verification of the τ_C-differentiability with logarithmic derivative $\rho_t^\infty = \sum_{i=1}^\infty \rho_{t,i}(x_i)$ is straightforward. Conversely this differentiability implies that

$$||(\mu_t^\infty)'|| \geq \sup_n \int_{X^n} |\sum_{i=1}^n \rho_{t,i}(x_i)| d\mu_t^n.$$

The sequence $(\sum_{i=1}^n \rho_{t,i}(x_i))$ then is a L^1-bounded martingale and hence it converges μ_t^∞-a.s.. The Three-Series-Theorem then implies the convergence of the series in b).
Part c) follows from Theorem 6.2 and the fact that the logarithmic derivative of the family (μ_{th_i}) at $t = 0$ equals $h_i\rho_0$. Hence in the case of the situation c) the series in b) converges if and only if $\sum_{i=1}^\infty h_i^2 < \infty$. ■

8 Differentiation of measures on vector spaces

Let X be a locally convex space (over the real numbers). By a vector field on X we mean a (continuous) mapping $h : X \longrightarrow X$. Let h be a vector field on X which an associated flow of diffeomorphisms of X exists, i.e. a differentiable [1]mapping $a : \mathbb{R} \times \mathrm{X} \longrightarrow \mathrm{X}$ having the following properties:

$$\begin{aligned} a(t_1 + t_2, x) &= a(t_1, a(t_2, x)); \\ a(0, x) &= x; \\ a'_t(t, x) &= h(a(t, x)) \end{aligned}$$

[1]We call a mapping f from a locally convex space E into another locally convex space differentiable if it is Gateau differentiable and the mapping $(h, x) \longmapsto f'(x)h, E \times E \longrightarrow E$ is continuous on each compact subset.

For simplicity we work in this section in the space $\mathcal{M}_R(X)$ of signed Radon measures. For each $\mu \in \mathcal{M}_R(X)$ and $t \in \mathbb{R}$ let $\mu_t^h \in \mathcal{M}_R(X)$ be the measure defined by $\mu_t^h(B) = \mu(a(t,B))$. The measure $\mu \in M_R(X)$ is called $(\tau-)$differentiable along the vector field h if the family $\{\mu_t^h\}$ is $(\tau-)$differentiable at $t = 0$. In this case by the derivative of μ along the vector field h is the measure from $M_R(X)$, denoted by $d_h\mu$ or $d\mu\, h$ and defined by $d_h\mu = (\mu_0^h)'$. For $\gamma_t(x) = a(-t,x)$ the concepts of the preceding section coincide with those we are considering now. (In particular the norm of the derivative $(\mu_t^h)'$ is independent of t and therefore the assumption of local integrability in theorem 2.7 is automaically fulfilled in these situations.) In particular, if $h(x) = h_0 \in X$ for all $x \in X$ (and $a(t,x) = x + t\,h_0$) then the derivative and differentiability along the vector field h are called derivative (differentiability) along the direction h_0.

In the finite dimensional case $X = \mathbb{R}^n$ a measure μ is $\tau_{C_b(\mathbb{R}^n)}$-differentiable along all directions iff it is absolutely continuous with respect to Lebesgue measure and its partial derivatives in the sense of distributions are given by measures. This means that the density $\frac{d\mu}{d\lambda^n}$ is a "function of bounded variation." ([Zie89],p.43). The fact that $\mu \ll \lambda^n$ can be shown as follows: A straightforward extension of the proof of proposition 2.5a) shows that the map $h \mapsto \mu(\cdot + h) \in \mathcal{M}(\mathbb{R}^n, \tau_v)$ is continuous. This implies $\mu \ll \lambda^n$ by a classical theorem (cf. e.g. [Sak64]). - The measure μ on $\mathbb{R}^n$ is even τ_v-differentiable along all directions iff in addition its partial derivatives are absolutely continuous, i.e. $\frac{d\mu}{d\lambda^n}$ is an "absolutely continuous function."

If $d_h\mu \ll \mu$ then the logarithmic derivative of (μ_t^h) at $t = 0$ is called the logarithmic derivative of μ along the vector field h (direction h_0); we use the symbol $\beta_h(x) = \frac{d(d\mu h)}{d\mu}(x)$[2]

The best known examples in infinite dimensions are the following

Example 8.1: Let $X = \mathbb{R}^\infty$ and let μ be the Radon measure on X defined as follows. Let p be a continuously differentiable strictly positive real even function of real variable such that $\int_{\mathbb{R}} p(t)dt = 1$ and $\int_{\mathbb{R}} \frac{|p'(s)|^2}{p(s)} ds < \infty$, ν be the (Radon) measure on $\mathbb{R}$ with the density $p, (E_j, \mathcal{B}_j, \nu_j)$ $(j \in N)$ be a copy of $(\mathbb{R}, \mathcal{B}_{\mathbb{R}}, \nu)$ and μ be the Lebesgue extension of $\bigotimes_{j=1}^\infty \nu_j$. Finally, let $H = \{(x_j) \in X : \sum_{j=1}^\infty x_j^2 < \infty\}$. Then the measure μ is τ_v-differentiable in the direction $h \in X$ iff $h \in H$. This follows e.g. from theorem 7.1c). For this and related results see e.g. [Haf89].
Let us now take for p the standard normal density. Then $\frac{p'}{p}(s) = -s$ and hence in this case $\beta_h(x) = -\sum_{i=1}^\infty h_i x_i$ which coincides with the scalar product of the Hilbert space H. Let $C_0[0,T]$ be the space of all continuous real functions defined on $[0,T]$ and W be the Wiener measure on $C_0[0,T]$. Then $(C_0[0,T], W)$ is (e.g. via the construction of Wiener measure using the Schauder base with i.i.d. Gaussian coefficients) linearly isomorphic to a measurable (linear) subspace of $(X, \mathcal{B}_X, \mu)$ having measure 1 and the subspace $W_2^1[0,1]$ of $C_0[0,T]$ maps by this isomorphism onto H. Proposition 4.3.1 implies that W is τ_v-differentiable along $h \in C_0[0,T]$ iff $h \in W_2^1[0,1]$. For such a vector h the formula (15) then is the Cameron-Martin formula since $\rho_0(x) = -<h,x>$ and hence $\int_0^1 \rho_0(x - th)\, dt = \int_0^1 -<h, x - th>\, dt = -<h,x> - \frac{1}{2}<h,h>$.

We need also the definition of the derivative and the logarithmic derivative along a vector (sub) space of X which generalizes the derivative along a direction. Thus let H be a locally convex space being a vector subspace of X, let for each $h \in X$ the $\tau-$derivative of $\mu (\in M_R(X))$ along h exist and let the mapping $d\mu : h \longmapsto d\mu h, H \longrightarrow (M_R(X), \tau)$ be continuous (this mapping is automatically linear [Ave71]). In this (and only in this) case

[2] If $h(x) = h_0 \in X$, then $\beta_{h_0}(\ ) \equiv \beta_h(\ )$.

we say that the measure μ is differentiable along H and the mapping $d\mu$ is called the derivative of μ along the space H. If moreover for each $h \in H$ one has $d_h\mu \ll d\mu$ and hence $d_h\mu = \beta_h(\cdot)\mu$, then (any version of) the function $\beta_H : (h,x) \longmapsto \beta_h(x)$ is called logarithmic derivative of μ along H.

We intend to describe the connection between the differentiablity of a measure along a subspace H and along some vector fields. We assume that H is Hilbert space with inner product $(\cdot,\cdot)$ and assume that the topology of X is defined by Hilbert-seminorms. Let h be a vector field in X having the following properties:
(a) $\forall x \in X \;\; h(x) \in H$; (b) $\forall x \in X \;\; h'(x)$ exists and its restriction on H is a trace class operator; (c) the function $x \longmapsto h'(x), X \longrightarrow \mathcal{L}_b(H,H)$ is bounded and continuous; (d) the set $h(X)$ y bounded it X.

Let $C_b^1(X)$ be the vector space of all functions $\varphi \in C_b(X)$ which are differentiable and such that their derivatives are (uniformly) bounded on X as mappings from X into the space X' equipped with the topology of convergence on bounded sets.

Proposition 8.2 *Let $\mu \in M_R(X)$ and let μ be τ_v-differentiable along H. Let there exist a version β_H of logarithmic derivative of μ along H such that for $\mu-$almost all $n \in X$ the functions $h \longmapsto \beta_H(h,x), H \longrightarrow \mathbb{R}$ are linear and continuous and the function $h \longmapsto \beta_H(h,\cdot),\; H \longrightarrow \mathcal{L}_1(X,\mu)$ is continuous. Then μ is τ_v-differentiable along the vector field h and*

$$\beta_h(x) = trh'(x) + \beta_H(h(x),x) \tag{16}$$

Proof:. This result can be found in several places (cf.[Dal85]). The following method of proof seems particularly clear and suitable for the extension discussed below. We use the so called bilinear integrals ([Die77]); but in our situation this is only a trick and all such integrals can be represented as sums of some series of ordinary integrals. For more general situations see below.

By theorem 6.2 it suffices to prove that μ is $\tau_{C_b^1(X)}$- differentiable and (16) holds. Using integration by parts ([Ave71],[Bog90]) the measure μ is $\tau_{C_b^1(X)}$-differentiable along h iff there exists such a measure ν that for all $\varphi \in C_b^1(X)$, $\int_X \varphi' h\,\mu(dx) = -\int_X \varphi\nu(dx)$; in this case $d_h\mu = \nu$.
Now, the formula of vector integration by parts $\int_X (\varphi\otimes h)'\mu(dx) = -\int (\varphi\otimes h)d\mu(dx)$ and Leibniz' formula $(\varphi\otimes h)' = \varphi'\otimes h + \varphi\otimes h'$ yield $\int_X \varphi'\otimes h\otimes\mu(dx) = -\int_X \varphi\otimes h'\otimes\mu(dx) - \int \varphi\otimes h\otimes d\mu(dx)$ for each $\varphi \in C_b^1(X)$. Applying the operation tr to the both sides of the last equality and using that $tr(\varphi'(x)\otimes h(x)) = (\varphi'(x), h(x))$, we can reformulate the preceding criterion of $\tau_{C_b^1(X)}-$differentiability of μ as follows: μ is $\tau_{C_b^1(X)}-$differentiable iff there exists a measure $\nu \in M_R(X)$ such that for each $\varphi \in C_b^1(X)$

$$\int_X \varphi(x)\nu(dx) = \int_X \varphi(x) tr\, h'(x)\mu(dx) + tr\int_X \varphi(x)h(x)\otimes\mu'(dx). \tag{17}$$

But $tr\int\varphi(x)h(x)\otimes\mu'(dx) = \int\varphi(x)(tr\, h(x)\otimes\mu')(dx) = \int\varphi(x)(h(x)\mu'(dx)) = = \int\varphi(x)\beta(h(x),x)\mu(dx)$; hence the right hand side of ?? gives such a measure ν:

$$\nu(= d_h\mu) = tr\, h'\cdot\mu + (h\cdot\mu') = (tr\, h'\beta_H(h(\cdot),(\cdot)))\mu :$$

so $\beta_h(x) = tr\, h'(x)\beta_H(h(x),x)$. ∎

Remark 8.3 Finally let us mention a further extension of these concepts, namely Lie derivatives and the corresponding shifts of so called differential forms of finite codegree.

For each $k, n \in \mathbb{N} \cup \{0\}$ we denote by X_n^k the tensor product of n copies of (a locally convex space) X and k copies of its strong dual endowed with the projective topology (the assumption about the topology on X_n^k is made only for a definitiveness); of course, we set $X_0^0 = \mathbb{R}$. Let us remark that it is possible to define the different sorts of differentiability for families of vectorvalued measures as well as for families of scalar measures.

Let m be a X_n^k−valued Radon measure on X (this means that m is defined on $\mathcal{B}_X$ and whenever $g \in X'$ the composition $g \circ m$ is a (realvalued) Radon measure on X). The Lie t−shift m_{Lt}^h of m along the vector field h is defined as follows:

$$m_{Lt}^h(B) = \int_B \underbrace{h'(x) \otimes \ldots \otimes h'(x)}_{n \text{ times}} \otimes \underbrace{h'^*(x) \otimes \ldots \otimes h'^*(x)}_{k \text{ times}} m_t^h(dx)$$

where m_t^h is defined similarly to μ_t^h above: $m_t^h(B) = m(a(t,B))$ and $\int_B$ is again a bilinear integral. The measure m is said to be $(\tau-)$Lie differentiable along the vector field h if the family (m_{Lt}) is τ−differentiable at $t = 0$, the derivative $m_{L0}^{h'}$ is called the $(\tau-)$Lie derivative of m along the vector field h. Instead of $(m_{L0}^h)'$ we use the symbol $d_h^L m$ (of course $d_h^L m \neq d_h m$ in general). If $k = n = 0$ we return to the case of realvalued measures; in this case the Lie differentiability coincides with the differentiability which was considered above.

An analogue of proposition 4.1 is true in the general case also; in particular under suitable assumptions

$$\begin{aligned} d_h^L m(B) &= d_h m(B) + \int_B (h'^*, m(dx))_1 + \ldots \int_B (h'^*, m(dx))_n \\ &\quad - \int_B (h', m(dx))^1 \ldots - \int_B (h', m(dx))^k. \end{aligned}$$

Here for each $j \in \{1, \ldots, n\}$, $(h'^*, a)_j$ is the element of X_n^k defined so that the mapping $a \longmapsto (h'^*, a)_j$, $X_n^k \longrightarrow X_n^k$ be continuous and for each $c = a_1 \otimes \ldots \otimes a_n \otimes b_1 \otimes \ldots \otimes b_k \in X_n^k$ one has $(h'^*, c)_j = a_1 \otimes \ldots h'^* a_j \otimes \ldots \otimes a_n \otimes b_1 \ldots \otimes b_k$. The definition of $(h, a)_r$, $r \in \{1, \ldots, k\}$ is similar.

A useful particular case of such measures are the measures whose range is in the subspace A^k of the space X^k, generated by (all) antisymmetric tensors from X^k; such measures can be considered as differential forms of codegree k on the space X. ([Smo86])

References

[Alb93] Albeverio, S., Kondratiev, Yu.G., Streit,L. Non Gaussian Analysis. *In preparation*, 1993.

[Ave71] Averbukh, V.I., Smolyanov, O.G. and Fomin, S.V. . *Trudi of Moscow Math. Soc.*, 24:133–174, 1971.

[Bel85] Bell, D. A quasi-invariance theorem for measures in Banach spaces. *Trans. Am. Math. Soc.*, 290:851–855, 1985.

[Bog90] Bogachev, V.I., Smolyanov, O.G. Analytical properties of infinite-dimensional distributions. *Russian Math. Surveys*, 45-3:1–104, 1990.

[Buc91] Buckdahn, R. Anticipative Girsanov transformations. *Probab. Th. Rel. Fields*, 89:211–238, 1991.

[Dal85] Daletski Yu. L. and Maryanin, B.D. Smooth measures on infinite dimensional manifolds. *Dokl. Akad. Nauk SSSR*, 285, No. 6:1287–1300, 1985.

[Dal88] Daletski Yu. L. and Sohadze, G. Absolute Continuity of Smooth Measures. *Functional Analysis and Applications* , 22, N2:77–78, 1988.

[Die77] Diestel, J. and Uhl, J.J., Jr. *Vector measures.* American Math. Soc., Providence, 1977.

[Dun58] Dunford, N. and Schwartz, J.T. *Linear Operators*, volume I: General Theory. Interscience, New York, 1958.

[Haf89] Hafizov, M.U. On the space of differentiability of a product measure. *Vestnik of Moscow University, ser. math. mech.*, 2:81–84, 1989.

[Haj62] Hajek, J. Asymptotically most powerful rank order tests. *Ann. Math. Stat.*, 33:1124–1147, 1962.

[LeC90] LeCam, L. and Yang, G.L. *Asymptotics in Statistics, Some Basic Concepts.* Springer Series in Statistics, Berlin - Heidelberg - New York, 1990.

[Mal90] Malliavin, M.-P. and Malliavin, P. Integration on Loop Groups. I. Quasi Invariant Measures. *Journ. of Funct. Analysis*, 93:207–237, 1990.

[Nev69] J. Neveu. *Mathematische Grundlagen der Wahrscheinlichkeitstheorie.* R. Oldenbourg Verlag, München - Wien, 1969.

[Sak64] Saks, S. *Theory of the Integral, 2nd. ed.* Hafner, Dover Reprint, 1964.

[Sko74] Skorokhod, A.V. *Integration in Hilbert Space.* Springer, Heidelberg-New York, 1974.

[Smo86] Smolyanov, O.G. De Rham currents and the Stokes formula in Hilbert spaces. *Soviet Math. Dokl.*, 33:140–144, 1986.

[Smo90] Smolyanov, O.G. Infinite dimensional distributions and second quantization in string theories. *Proc. V. International Vilnius Conf. on probability and mathematical statistics*, 1990.

[Tor91] E.N. Torgersen. *Comparison of Statistical Experiments.* Cambridge University Press, Cambridge - New York, 1991.

[Wei90] Weizsäcker, H.v. and Winkler, G. *Stochastic Integrals, An Introduction.* Vieweg Advanced Lectures in Mathematics, 1990.

[Zie89] Ziemer,William P. *Weakly Differerentiable Functions.* Springer Graduate Texts, 1989.