

Operator Differentiable Functions

Dedicated with respect and affection to Shôichirô Sakai

By

Gert K. PEDERSEN*

Abstract

We study the Banach $*$ -algebra $C_{\text{op}}^1(I)$ of C^1 -functions f on the compact interval I such that the corresponding Hilbert space operator function $T \rightarrow f(T)$, for $T = T^*$ and $\text{sp}(T) \subset I$, is Fréchet differentiable. If $f(x) = \int e^{itx} \widehat{f}(t) dt$ we know that the differential is given by the formula

$$df_T(S) = \int_{-\infty}^{\infty} \int_0^1 U_s S U_{(1-s)t} ds \widehat{f}'(t) dt,$$

where $U_t = \exp(itT)$. Functions of this type are dense in $C_{\text{op}}^1(I)$, and $C^2(I) \subset C_{\text{op}}^1(I) \subset C^1(I)$, so several classical results can be deduced. In particular we show that if $T \in \mathfrak{D}(\delta)$, where δ is the generator of a one-parameter group of $*$ -automorphisms of a C^* -algebra \mathfrak{A} (or just a closed $*$ -derivation in \mathfrak{A}), then $f(T) \in \mathfrak{D}(\delta)$ for every f in $C_{\text{op}}^1(I)$, where $\text{sp}(T) \subset I$, and

$$\delta(f(T)) = df_T(\delta(T)).$$

§ 1. Introduction

If \mathfrak{A} is a unital C^* -algebra of operators on a Hilbert space \mathcal{H} , and T is a normal element in \mathfrak{A} , then by the spectral theorem we have a $*$ -isomorphism $f \rightarrow f(T)$ from $C(\text{sp}(T))$ onto the C^* -subalgebra $C^*(T)$ of \mathfrak{A} generated by T and $\mathbf{1}$.

Communicated by T. Kawai, August 16, 1999.

1991 Mathematics Subject Classification: Primary 46L05; Secondary 47A60, 46C15.

Key words and phrases: C^* -algebra, Fréchet derivative, operator differentiability, commutator equations.

Supported in part by SNF, Denmark.

* Department of Mathematics, University of Copenhagen, Denmark.

Suppose instead we fix f and vary T . For this to make sense we assume that f is defined on some region Ω in \mathbb{C} (respectively \mathbb{R}), and we let \mathfrak{A}_n^Ω (respectively \mathfrak{A}_{sa}^Ω) denote the set of normal elements in \mathfrak{A} with spectra in Ω . Then the map

$$T \rightarrow f(T)$$

defines an operator function from \mathfrak{A}_n^Ω into $\mathfrak{A}_n (= \mathfrak{A}_n^{\mathbb{C}})$. It is evident from the Stone-Weierstrass theorem that this function is continuous, but not obvious how to estimate the rate of convergence. For example, we might wish to determine a constant γ – depending on f – such that

$$\|f(S) - f(T)\| \leq \gamma \|S - T\|.$$

Similarly we might wish to estimate the norms of commutators by a constant δ :

$$\|[S, f(T)]\| \leq \delta \| [S, T] \|.$$

Clearly the uniform bound $\|f\|_\infty$ of f is not particularly relevant for these problems since it depends on the pointwise behaviour, whereas $f(T)$ may involve all values of f simultaneously. The best known estimates, $\gamma = \delta = \|\widehat{f}'\|_1$ (valid for real functions and self-adjoint operators), indicate that the problem is non-trivial. We shall see that differentiability conditions on the operator function hold the key to the problems above, and that there is a fascinating – and not widely recognized – theory of differential operators on C^* -algebras associated with the spectral operator functions.

Recall that if \mathfrak{X} and \mathfrak{Y} are Banach spaces a function $f: \mathfrak{D} \rightarrow \mathfrak{Y}$ on a subset \mathfrak{D} of \mathfrak{X} is *Fréchet differentiable* if for each T in \mathfrak{D} there is a bounded linear operator df_T in $\mathcal{B}(\mathfrak{X}, \mathfrak{Y})$ – *the differential* – such that

$$\lim_{S \rightarrow 0} \|S\|^{-1} (f(T+S) - f(T) - df_T(S)) = 0.$$

If df_T depends continuously on T we say that f is *continuously Fréchet differentiable*. Standard references are [13] and [19].

Let now \mathcal{H} be the separable, infinite-dimensional Hilbert space, and denote by $\mathcal{B}(\mathcal{H})_{sa}^I$ the set of self-adjoint operators T on \mathcal{H} with spectra in I , where I is some interval (open or closed, bounded or unbounded – we shall need all possibilities). We say that a function f in $C(I)$ is *operator differentiable* if the operator function

$$f: \mathcal{B}(\mathcal{H})_{\text{sa}}^I \rightarrow \mathcal{B}(\mathcal{H}) \text{ given by } T \rightarrow f(T)$$

is Fréchet differentiable. If the operator function is continuously Fréchet differentiable we say that f is *continuously operator differentiable*. For reasons that will become clear later (Theorem 2.7) we shall denote by $C_{\text{op}}^1(I)$ the class (actually the Banach $*$ -algebra) of operator differentiable functions on I . Although the differential operator df_T is by definition in $\mathcal{B}(\mathcal{B}(\mathcal{H})_{\text{sa}}, \mathcal{B}(\mathcal{H}))$ (and actually in $\mathcal{B}(\mathcal{B}(\mathcal{H})_{\text{sa}})$ if f is real) we shall tacitly extend it to an operator in $\mathcal{B}(\mathcal{B}(\mathcal{H}))$ by standard complexification.

Remark 1.1. Since every separable C^* -algebra \mathfrak{A} can be faithfully represented on \mathcal{H} , the definition of operator differentiability implies that the operator function $T \rightarrow f(T)$ from $\mathfrak{A}_{\text{sa}}^I$ into \mathfrak{A} is Fréchet differentiable with a differential in $\mathcal{B}(\mathfrak{A})$. Conversely, if f satisfies this condition for every separable C^* -algebra, then f is actually Fréchet differentiable on any C^* -algebra, in particular on $\mathcal{B}(\mathcal{H})$, because the existence and continuity of the differential only involves a separable family of elements.

Remark 1.2. Evidently the definition of operator differentiability is patterned after Löwner's definition of operator monotone function [20] and the later theory by Benda and Sherman [5] and Davis [12] of operator convex functions. The differential is implicit in Löwner's treatment, and explicitly used in the version of operator monotonicity and convexity given in [16]. There are some striking differences though: To verify that a function is operator monotone or convex it suffices to check this on all finite-dimensional algebras \mathbb{M}_n . Moreover, the resulting class of functions is highly restricted, see e.g. [16]. By contrast, operator differentiability is a truly infinite-dimensional phenomenon, and the class of (continuously) operator differentiable functions is pleasantly large, cf. Theorem 3.2.

Remark 1.3. Since most of the classical Banach spaces are complex or have natural complexifications, the concept of Fréchet differentiability is usually absorbed in the theory of functions of (one or several) complex variables, cf. [19]. Thus, if we define a function f in $C(\Omega)$, $\Omega \subset \mathbb{C}$, to be operator differentiable on $\mathcal{B}(\mathcal{H})_n^\Omega$ in the obvious way (although the definition may not be very convincing, since the linear structure of $\mathcal{B}(\mathcal{H})_n^\Omega$ is fragile), then necessarily f must be holomorphic in Ω .

Presumably the set $\mathcal{B}(\mathcal{H})_{\text{sa}}$ of self-adjoint operators on Hilbert space

is the most important real Banach space whose complexification, $\mathcal{B}(\mathcal{H})$, is radically different from the original space. This may explain why the theory of operator differentiability – defined for functions of a real variable – has little connection with analytic function theory; even though $\mathcal{B}(\mathcal{H})$ is a complex Banach algebra. Going one step further, the author would maintain that C^* -algebra theory is basically a real theory – concerning the ordered vector spaces \mathfrak{A}_{sa} – whereas the truly complex operator algebras are the non self-adjoint operator algebras.

The contents of this paper is an expanded version of a lecture delivered at the Yamanakoko Conference in January 1998 in honour of S. Sakai.

§ 2. Necessary Conditions

Proposition 2.1 (Widom [33]). *If f in $C(I)$ is operator differentiable then f belongs to $C^1(I)$.*

Proof. Restricting if necessary to a smaller interval we may assume that I is bounded. Put $\mathfrak{A} = C_b(I)$, and let $T = \text{id}$ (regarded as a multiplication operator) and $S = \mathbf{1}$. By assumption there is a bounded, linear operator $df_{\text{id}} : \mathfrak{A} \rightarrow \mathfrak{A}$ such that

$$f(\text{id} + \varepsilon \mathbf{1}) = f(\text{id}) + df_{\text{id}}(\varepsilon \mathbf{1}) + \varepsilon R_\varepsilon,$$

where $R_\varepsilon \rightarrow 0$ as $\varepsilon \rightarrow 0$. Since $f(T) = f \circ T$ for every T in $C_b(I)_{\text{sa}}$, this shows that f is differentiable on I with

$$f' = df_{\text{id}}(\mathbf{1}) \in \mathfrak{A} = C_b(I).$$

Q.E.D.

Proposition 2.2. *If $f \in C_{\text{op}}^1(I)$, and S, T are commuting elements in $\mathcal{B}(\mathcal{H})_{\text{sa}}'$, then*

$$df_T(S) = f'(T)S.$$

Proof. We may assume that S and T are elements in a unital, commutative C^* -algebra $\mathfrak{A} = C(X)$. Thus, $f(T)(x) = f(T(x))$ for every x in X , so that $f(T) = f \circ T$. By assumption we have a differential df_T in $\mathcal{B}(\mathfrak{A})$ such that

$$f(T + \varepsilon S)(x) = f(T)(x) + \varepsilon df_T(S)(x) + \varepsilon R_\varepsilon(x),$$

where $R_\varepsilon \rightarrow 0$ in \mathfrak{A} as $\varepsilon \rightarrow 0$. Since f is differentiable by Proposition 2.1 it follows that

$$df_T(S)(x) = f'(T(x))S(x),$$

as desired.

Q.E.D.

Proposition 2.3. *If $f, g \in C_{\text{op}}^1(I)$, then $fg \in C_{\text{op}}^1(I)$, with*

$$d(fg)_T(S) = df_T(S)g(T) + f(T)dg_T(S).$$

Proof. As in highschool we write

$$\begin{aligned} & \varepsilon^{-1}(fg(T + \varepsilon S) - fg(T)) \\ &= \varepsilon^{-1}(f(T + \varepsilon S) - f(T))g(T + \varepsilon S) + \varepsilon^{-1}f(T)(g(T + \varepsilon S) - g(T)) \\ & \rightarrow df_T(S)g(T) + f(T)dg_T(S), \end{aligned}$$

as desired.

Q.E.D.

Remark 2.4. Note that even though $fg = gf$, the formula in Proposition 2.3 for the differential $d(fg)_T$ has a non-commutative form. However, the identification $d(fg)_T(S) = d(gf)_T(S)$ gives us the following commutator equation

$$[df_T(S), g(T)] = [dg_T(S), f(T)].$$

In particular, taking $g = \text{id}$, we see that

$$(*) \quad [df_T(S), T] = [S, f(T)].$$

Remark 2.5. It follows from Proposition 2.3 that $C_{\text{op}}^1(I)$ is an algebra of functions, and actually a $*$ -algebra since (for $S = S^*$)

$$d\overline{f_T}(S) = (df_T(S))^*.$$

We also see that if $f \in C_{\text{op}}^1(J)$ and $g \in C_{\text{op}}^1(I)$, such that $g(I) \subset J$, then $f \circ g \in C_{\text{op}}^1(I)$ with

$$d(f \circ g)_T(S) \in df_{g(T)}(dg_T(S)).$$

The proof of the chain rule is a straightforward computation, valid for composition of Fréchet differentiable functions in general, cf. [13, 3.14].

Theorem 2.6. *Each function f in $C_{\text{op}}^1(I)$ is continuously operator differentiable, and the norm of the differential is bounded on compact subsets of I .*

Proof. We first show that for fixed S the function $T \rightarrow df_T(S)$ is continuous. For this, take a C^* -algebra \mathfrak{A} (e.g. $\mathfrak{B}(\mathcal{H})$) and form the C^* -algebra $\mathfrak{B} = \mathfrak{A} \otimes \ell^\infty$ of bounded sequences from \mathfrak{A} , together with its closed $*$ -subalgebra $\mathfrak{B}_c = \mathfrak{A} \otimes c$ of convergent sequences.

If (T_n) is a sequence in $\mathfrak{A}_{\text{sa}}^I$ converging to T_∞ in $\mathfrak{A}_{\text{sa}}^I$, put $T = (T_n)$ in $(\mathfrak{B}_c)_{\text{sa}}^I$ and identify S with the constant sequence (S) . Since f is operator differentiable on $(\mathfrak{B}_c)_{\text{sa}}^I$ there is a differential operator df_T . It is easy to verify – using sequences (S_n) with a single non-zero term – that df_T has the form (df_{T_n}) , where each element in the sequence is the differential of f on $\mathfrak{A}_{\text{sa}}^I$ at T_n . Thus,

$$(df_{T_n}(S)) = df_T(S) \in \mathfrak{A} \otimes c,$$

which shows that $df_{T_n}(S) \rightarrow df_{T_\infty}(S)$.

If $T \rightarrow df_T$ was not continuous (from $\mathfrak{A}_{\text{sa}}^I$ into $\mathfrak{B}(\mathfrak{A})$) we could find a sequence (T_n) converging to T_∞ in $\mathfrak{A}_{\text{sa}}^I$, and $\varepsilon > 0$ such that $\|df_{T_\infty} - df_{T_n}\| > \varepsilon$ for all n . There is then a sequence (S_n) in \mathfrak{A}_{sa} with $\|S_n\| = 1$, such that

$$\|df_{T_\infty}(S_n) - df_{T_n}(S_n)\| > \varepsilon$$

for all n . Define elements S, R and R_n in \mathfrak{B}_{sa} by $S = (S_k)$, $R = (T_\infty)$, and

$$(R_n)_k = T_\infty \text{ for } 1 \leq k < n, \quad (R_n)_k = T_k \text{ for } n \leq k.$$

Evidently, then, $R_n \rightarrow R$ inside $\mathfrak{B}_{\text{sa}}^I$, so by the first part of the proof

$$df_{R_n}(S) \rightarrow df_R(S).$$

However,

$$\begin{aligned} \|df_{R_n}(S) - df_R(S)\| &= \sup_k \|df_{(R_n)_k}(S_k) - df_{T_\infty}(S_k)\| \\ &= \sup_k \|df_{T_k}(S_k) - df_{T_\infty}(S_k)\| > \varepsilon, \end{aligned}$$

a contradiction.

If I is compact, set

$$\gamma = \sup \{\|df_T\| \mid T \in \mathfrak{A}_{\text{sa}}^I\}.$$

There is then a sequence (T_n) – necessarily bounded – such that $\|df_{T_n}\| \rightarrow$

γ . Regarding (T_n) as an element T of \mathfrak{B}_{sa}^I we see that $df_T = (df_{T_n})$, whence

$$\gamma = \lim \|df_{T_n}\| = \|df_T\| < \infty.$$

Q.E.D.

The next result have appeared before in weaker forms, for example as Proposition 3.4 in [16] and as Theorem 2.1 in [17].

Theorem 2.7. *If $f \in C_{op}^1(I)$, and if $t \rightarrow A(t)$ is a C^1 -curve from $[0,1]$ into $\mathfrak{B}(\mathcal{H})_{sa}^I$ then*

$$f(A(1)) - f(A(0)) = \int_0^1 df_{A(t)}(A'(t)) dt.$$

Proof. It is easy to verify, cf. Remark 2.5, that the composed map $t \rightarrow f(A(t))$ is Fréchet differentiable with differential $df_{A(t)}(A'(t)) -$ identifying $\mathfrak{B}(\mathcal{H})$ with the algebra $\mathfrak{B}(\mathbb{R}, \mathfrak{B}(\mathcal{H}))$. We subdivide the unit interval, $0 = t_1 < t_2 < \dots < t_n = 1$, and write

$$\begin{aligned} f(A(1)) - f(A(0)) &= \sum_{k=1}^{n-1} f(A(t_{k+1})) - f(A(t_k)) \\ &= \sum_{k=1}^{n-1} df_{A(t_k)}(A'(t_k)) (t_{k+1} - t_k) + R(t_k) (t_{k+1} - t_k), \end{aligned}$$

where $R(t_k) \rightarrow 0$ as $t_{k+1} - t_k \rightarrow 0$. Since the function $t \rightarrow df_{A(t)}(A'(t))$ is continuous by Theorem 2.6, we obtain the difference $f(A(1)) - f(A(0))$ as the limit of a Riemann sum, and the assertion follows. Q.E.D.

Corollary 2.8. *If $f \in C_{op}^1(I)$ for some compact interval I , and $A, B \in \mathfrak{B}(\mathcal{H})_{sa}^I$, then*

$$\|f(B) - f(A)\| \leq \gamma \|B - A\|,$$

where

$$\gamma = \sup \{ \|df_T\| \mid T \in \mathfrak{B}(\mathcal{H})_{sa}^I \}.$$

Proof. With $A(t) = tB + (1-t)A$ we have

$$f(B) - f(A) = \int_0^1 df_{tB+(1-t)A}(B-A),$$

by Theorem 2.7, and the inequality follows.

Q.E.D.

Proposition 2.9. *Let I be a compact interval, and for each f in $C_{\text{op}}^1(I)$ define*

$$\|f\|_d = \|f\|_{\infty} + \sup\{\|df_T\| \mid T \in \mathcal{B}(\mathcal{H})_{\text{sa}}^I\}.$$

Then $C_{\text{op}}^1(I)$ is a Banach space in d -norm.

Proof. If $(f_n) \in C_{\text{op}}^1(I)$, and $f_n \rightarrow f$ in d -norm, then $f \in C(I)$. In fact, $f \in C^1(I)$, because $\|f'\|_{\infty} \leq \|f\|_d$, cf. Proposition 2.2. For fixed elements S, T in $\mathcal{B}(\mathcal{H})_{\text{sa}}^I$ and $0 \leq t \leq 1$ define D_t in $\mathcal{B}(\mathcal{B}(\mathcal{H})_{\text{sa}}, \mathcal{B}(\mathcal{H}))$ as

$$D_t = \lim_n df_{n(1-t)T+tS}$$

Then the function $t \rightarrow D_t$ is continuous as a uniform limit of continuous functions, cf. Theorem 2.6. By Theorem 2.7 we have

$$f_n(T + \varepsilon(S - T)) - f_n(T) = \varepsilon \int_0^1 df_{nT + \varepsilon(S - T)}(S - T) dt,$$

whence in the limit

$$f(T + \varepsilon(S - T)) - f(T) = \varepsilon \int_0^1 D_{t\varepsilon}(S - T) dt.$$

It follows that

$$\|f(T - \varepsilon(S - T)) - f(T) - \varepsilon D_0(S - T)\| \leq \varepsilon \|S - T\| \int_0^1 \|D_{t\varepsilon} - D_0\| dt,$$

so by Lebesgue's theorem f is operator differentiable at T with $df_T = D_0$, i.e. $f \in C_{\text{op}}^1(I)$. Q.E.D.

Proposition 2.10. *If $\{f_x \mid x \in X\}$ is a family of functions in $C_{\text{op}}^1(I)$, depending continuously on the parameter x in a compact Hausdorff space X , and if the function $x \rightarrow df_x$ belongs to $C(X, \mathcal{B}(\mathcal{B}(\mathcal{H})))$, then for each Borel measure μ on X we have that $\int f_x d\mu(x) \in C_{\text{op}}^1(I)$ with*

$$d\left(\int f_x d\mu(x)\right) = \int df_x d\mu(x).$$

Proof. The existence of $f = \int f_x d\mu(x)$ in $C^1(I)$ and of $D = \int df_x d\mu(x)$ in the algebra $C(\mathcal{B}(\mathcal{H})_{\text{sa}}^I, \mathcal{B}(\mathcal{B}(\mathcal{H})))$ is assured by our assumptions. To prove that f is operator differentiable take T in $\mathcal{B}(\mathcal{H})_{\text{sa}}^I$ and S in $\mathcal{B}(\mathcal{H})$ and compute

$$\begin{aligned} & \| \epsilon^{-1}(f(T + \epsilon S) - f(T)) - D_T(S) \| \\ & \leq \int \| (\epsilon^{-1}(f_x(T + \epsilon S) - f_x(T)) - (df_x)_T(S)) d\mu(x) \| \\ & \leq \int \| (df_x)_T(S) + \epsilon^{-1}o_x(\epsilon) - (df_x)_T(S) \| \\ & = \int \| \epsilon^{-1}o_x(\epsilon) \| d\mu(x) \rightarrow 0 \end{aligned}$$

as $\epsilon \rightarrow 0$, by Lebesgue's dominated convergence theorem. Q.E.D.

Corollary 2.11. *If $f \in C_{op}^1([a, b])$ and $g \in C([-\epsilon, \epsilon])$, then for the convolution product we have $f \times g \in C_{op}^1([a + \epsilon, b - \epsilon])$ with*

$$d(f \times g) = \int_{-\epsilon}^{\epsilon} df_{(T-x)}g(x)dx.$$

Proposition 2.12. *For each compact interval I in \mathbb{R} the C^∞ -functions in $C_{op}^1(I)$ are dense in $C_{op}^1(I)$ with respect to the d -norm.*

Proof. If $I = [a, b]$ and f in $C_{op}^1(I)$ are given, together with $\epsilon > 0$, we may as well assume that $f \in C_{op}^1([a - \epsilon, b + \epsilon])$, replacing otherwise f by $f \circ \theta$, where θ is an affine transformation of $[a - \epsilon, b + \epsilon]$ onto $[a, b]$. Because evidently $f \circ \theta \in C_{op}^1([a - \epsilon, b + \epsilon])$, and $\|(f \circ \theta) - f\|_d \rightarrow 0$ as $\epsilon \rightarrow 0$.

Choose a positive C^∞ -function g with support in $[-\epsilon, \epsilon]$ and integral 1. By Corollary 2.11 the convolution product $f \times g$ belongs to $C_{op}^1(I)$.

As $\epsilon \rightarrow 0$ the functions g form an approximate unit for $L^1(\mathbb{R})$. Consequently $\|f \times g - f\|_\infty \rightarrow 0$, and $\|d(f \times g)_T - df_T\| \rightarrow 0$ for each T in $\mathcal{B}(\mathcal{H})_{sa}^I$. But that, as we have seen before, implies that $\|f \times g - f\|_d \rightarrow 0$. Indeed, if $\|(f \times g_n) - f\|_d > \epsilon$ for some sequence (g_n) , where $\text{supp}(g_n) \subset [-\delta_n, \delta_n]$ for each n , and $\delta_n \rightarrow 0$, then $\|d(f \times g_n)_{T_n} - df_{T_n}\| > \epsilon$ for some sequence (T_n) – necessarily bounded – in $\mathcal{B}(\mathcal{H})_{sa}^I$. Passing to $\mathcal{B}(\mathcal{H}) \otimes \ell^\infty$, and noting that f is still operator differentiable we let $T = (T_n)$ in $(\mathcal{B}(\mathcal{H}) \otimes \ell^\infty)_{sa}^I$ and have

$$\|d(f \times g_n)_T - df_T\| = \sup \|d(f \times g_n)_{T_k} - df_{T_k}\| > \epsilon,$$

in contradiction with our previous result. Thus, $\|f \times g_n - f\|_d \rightarrow 0$, as desired. Q.E.D.

§ 3. Sufficient Conditions

For any interval I , let $C_F^n(I)$ denote the space of functions f on I representable in the form

$$(*) \quad f(x) = \int_{-\infty}^{\infty} e^{itx} d\mu(t) \quad x \in I$$

for some finite, signed measure μ (depending on f) for which $\int_{-\infty}^{\infty} |t|^n d|\mu|(t) < \infty$. Note that if $f \in L^1(\mathbb{R})$ then $\mu = \widehat{f}$, so that the condition $(*)$ means that \widehat{f} has moments of n 'th order. Evidently,

$$C_F^n(I) \subset C^n(I)$$

for each n . In the converse direction we have the following:

Proposition 3.1. *If I is a closed interval and $f \in C^{n+1}(I)$ such that $f^{(n+1)} \in L^2(I)$, then $f \in C_F^n(I)$. In particular, if I is compact we have*

$$C^{n+1}(I) \subset C_F^n(I).$$

Proof. If I is (semi-) bounded we can extend f to a function \tilde{f} in $C^{n+1}(\mathbb{R})$ such that $\tilde{f}^{(n+1)} \in L^2(\mathbb{R})$. Replacing f with \tilde{f} we may assume that $I = \mathbb{R}$.

We have

$$(it)^k \widehat{f}(t) = (2\pi)^{-1} \int_{-\infty}^{\infty} e^{-itx} f^{(k)}(x) dx$$

for $0 \leq k \leq n+1$. Since $f^{(n+1)} \in L^2(\mathbb{R})$, it follows from the Plancherel theorem that $(id)^{n+1} \widehat{f} \in L^2(\mathbb{R})$. But then $(1 + |id|)^{-1} (id)^{n+1} \widehat{f} \in L^1(\mathbb{R})$, whence $(id)^n \widehat{f} \in L^1(\mathbb{R})$, as desired. Q.E.D.

Theorem 3.2. (cf. [33], [17, 1.5] & [15, 2.8]) *Each function f in $C_F^1(\mathbb{R})$ is operator differentiable on \mathbb{R} ; and if $f(x) = \int_{-\infty}^{\infty} e^{itx} d\mu(t)$, then*

$$df_T(S) = \int_{-\infty}^{\infty} \int_0^1 e^{itsT} Se^{it(1-s)T} ds it d\mu(t)$$

for all S, T in $\mathcal{B}(\mathcal{H})_{sa}$.

Proof. As in [17] we use Dyson's formula (see e.g. [3]), to show that the exponential function is operator differentiable with differential

$$d(\exp)_T(S) = \int_0^1 e^{sT} Se^{(1-s)T} ds.$$

Inserting this in the formula for f gives the desired result by Proposition 2.10. Q.E.D.

Corollary 3.3. *If $f \in C_F^1(\mathbb{R})$ such that $\widehat{f'} \in L^1(\mathbb{R})$, then for any T in $\mathcal{B}(\mathcal{H})_{sa}$*

$$\|df_T\| \leq \|\widehat{f'}\|_1.$$

Proof. We have $\mu = \widehat{f}$, and it follows from the formula in Theorem 3.2 that

$$\|df_T\| \leq \|S\| \int |t| d|\mu|(t) = \|S\| \|\widehat{f'}\|.$$

Q.E.D.

Remark 3.4. Combining Proposition 3.1 and Theorem 3.2 it follows that $C^2(I) \subset C_{op}^1(I)$ for every compact interval I . However, the estimate in Corollary 3.3 depends crucially on the choice of the extension of the function from I to \mathbb{R} , cf. Remark 3.13. The relevance of the class $C_F^1(\mathbb{R})$ was pointed out by Widom in one of the earliest references to the class $C_{op}^1(\mathbb{R})$, [33] (whose first edition is from 1984).

Remark 3.5. The estimate $\|df_T\| \leq \|\widehat{f'}\|_1$ in Corollary 3.3 can often be improved. But we see from Proposition 2.2 that $\|df_T\| \geq \|f'(T)\|$, so that this number is a lower bound. The class \mathfrak{D}_1 of operator differentiable functions f on $]0, \infty[$, for which $\|df_T\| = \|f'(T)\|$ ($= \|f' \mid \text{sp}(T)\|_\infty$) is studied in [6], [7], and [8]. It is surprisingly large, and includes the operator monotone functions, and the holomorphic functions with positive Taylor coefficients (such as monomials and exponential functions), cf. [8]. However, the function $f(t) = t^\alpha$ does not belong to \mathfrak{D}_1 if $1 < \alpha < \sqrt{2}$. Counterexamples occur already with 2×2 matrices, cf. [6, Example 2.A].

Remark 3.6. Although every operator differentiable function is continuously operator differentiable, cf. Theorem 2.6, the variation of the differential df_T is not necessarily uniform in the variable T . However, an estimate of the form

$$\|df_T - df_S\| \leq \gamma \|T - S\|$$

is obtainable if $f \in C_F^2(\mathbb{R})$ (with $\gamma = \int |t|^2 d|\mu|(t)$). The straightforward computation is left to the reader. The second differential is treated in [13], and studied in some detail for operator differentiable

functions in [15] and [4], where convexity considerations makes it essential. See also [32] and [8].

Theorem 3.7. *Let δ be the generator of a strongly continuous, one-parameter family α_t , $t \in \mathbb{R}$, of $*$ -automorphisms of a C^* -algebra \mathfrak{A} . Then $f(T) \in \mathfrak{D}(\delta)$ for every T in $\mathfrak{D}(\delta)_{\text{sa}}$ and every operator differentiable function f on an interval containing $\text{sp}(T)$, and*

$$\delta(f(T)) = df_T(f(T)).$$

In particular,

$$\|\delta(f(T))\| \leq \|df_T\| \|\delta(T)\|.$$

Proof. We know that $\alpha_t = \exp(t\delta)$, so that

$$\mathfrak{D}(\delta) = \{T \in \mathfrak{A} \mid \lim_{t \rightarrow 0} t^{-1}(\alpha_t(T) - T) \in \mathfrak{A}\}.$$

If $T \in \mathfrak{D}(\delta)_{\text{sa}}$ and f is operator differentiable we therefore have a continuous curve $s \rightarrow A(s)$ from T to $\alpha_t(T)$ given by $A(s) = \alpha_{st}(T)$. Since $T \in \mathfrak{D}(\delta)$ this curve is actually differentiable with $A'(s) = t\alpha_{st}(\delta(T))$. By Theorem 2.8 this implies that

$$f(\alpha_t(T)) - f(T) = \int_0^1 df_{\alpha_{st}(T)}(t\alpha_{st}(\delta(T))) ds.$$

Consequently $f(T) \in \mathfrak{D}(\delta)$, and

$$\begin{aligned} \delta(f(T)) &= \lim_{t \rightarrow 0} t^{-1}(f(\alpha_t(T)) - f(T)) \\ &= \lim_{t \rightarrow 0} \int_0^1 df_{\alpha_{st}(T)} \alpha_{st}(\delta(T)) ds = df_T(\delta(T)). \end{aligned}$$

Q.E.D.

Corollary 3.8. *If $f \in C_{\text{op}}^1(I)$ and $T \in \mathfrak{B}(\mathcal{H})_{\text{sa}}^I$, then for every S in $\mathfrak{B}(\mathcal{H})$ we have*

$$[f(T), S] = df_T([T, S]) = [T, df_T(S)].$$

In particular,

$$\|[f(T), S]\| \leq \|df_T\| \|[T, S]\|.$$

Proof. The first equality sign follows by taking $\delta(T) = [T, S]$ (breaking if necessary S into real and imaginary parts). The second follows from (*) in Remark 2.4. Q.E.D.

Remark 3.9. Theorem 3.7 is actually valid even if δ is only a densely defined, closed $*$ -derivation of \mathfrak{A} . The argument goes as follows: First show that $\delta(f(T)) = df_T(\delta(T))$, whenever $f \in C_F^1(\mathbb{R})$. The proof can be found in [29, 3.3.6] and uses the formula from Theorem 3.2, coupled with the fact that $e^{iT} \in \mathfrak{D}(\delta)$ whenever $T \in \mathfrak{D}(\delta)_{sa}$. After that one uses Proposition 2.11 to approximate an arbitrary function from $C_{op}^1(I)$ with functions in $C_F^1(\mathbb{R}) \upharpoonright I$. As another application of this approach we have the following:

Proposition 3.10. *If $f \in C_{op}^1(I)$ and $T \in \mathcal{B}(\mathcal{H})_{sa}^I$, then for all A, B and S in $\mathcal{B}(\mathcal{H})$, where A and B commute with T , we have*

$$df_T(ASB) = A df_T(S)B.$$

Proof. If $f \in C_F^1(\mathbb{R})$, then from the formula in Theorem 3.2 we see that df_T is an average of operators of the form

$$S \rightarrow e^{ixT} S e^{iyT}.$$

Each of these is a module operator over the commutant of T , and therefore so is df_T .

In the general case we may assume that I is compact. There is then a sequence (f_n) in $C^\infty(I)$ converging to f in d -norm. By Proposition 3.1 these functions are in $C_F^1(I)$. Since $df_{nT} \rightarrow df_T$ for every T , the desired formula follows. Q.E.D.

Proposition 3.11. *If $f \in C_{op}^1(I)$ and $g \in C_{op}^1(J)$ and R, T are commuting operators in $\mathcal{B}(\mathcal{H})_{sa}^I$ and $\mathcal{B}(\mathcal{H})_{sa}^J$, respectively, then*

$$[df_R, dg_T] = 0 \text{ in } \mathcal{B}(\mathcal{B}(\mathcal{H})).$$

Proof. Assume first that $f \in C_F^1(\mathbb{R})$ with $f(x) = \int e^{itx} d\mu(t)$. Then by Theorem 3.2 and Proposition 3.10 we have for each S in $\mathcal{B}(\mathcal{H})$ that

$$\begin{aligned} df_R(dg_T(S)) &= \int_{-\infty}^{\infty} \int_0^1 e^{itsR} dg_T(S) e^{it(1-s)R} ds \, it d\mu(t) \\ &= \int_{-\infty}^{\infty} \int_0^1 dg_T(e^{itsR} S e^{it(1-s)R}) ds \, it d\mu(t) = dg_T(df_R(S)). \end{aligned}$$

The general case follows by approximating f in d -norm with C_F^1 -functions on any compact subinterval of I containing $\text{sp}(T)$. Q.E.D.

Proposition 3.12. $C_{\text{op}}^1(I) \neq C^1(I)$.

Proof. For $0 < \alpha < 1$ (think $\alpha = \frac{1}{2}$) consider the function

$$f_\alpha(x) = |x| (\log(1 - \log|x|))^{-\alpha}$$

on the interval $I = [-e^{-1}, e^{-1}]$. Clearly $f \in C^1(I)$, the only point in dispute being $x = 0$ where $f'_\alpha(0) = 0$. However, McIntosh constructs in [21, Theorem 4] a generator δ for an automorphism group $\alpha_t = \exp(t\delta)$, together with an element T in $\mathfrak{D}(\delta)_{\text{sa}}$, such that $f_\alpha(T) \notin \mathfrak{D}(\delta)$. It follows from Theorem 3.7 that $f_\alpha \notin C_{\text{op}}^1(I)$. By translation and scaling, the statement $C^1(I) \not\subseteq C_{\text{op}}^1(I)$ holds for any interval. Q.E.D.

Remark 3.13. In view of the preceding results one may naturally wonder whether every operator differentiable function on an interval I is the restriction of a function from $C_F^1(\mathbb{R})$. Especially so since by Propositions 2.12 & 3.1 the space $C_F^1(I)$ is dense in $C_{\text{op}}^1(I)$ for the d -norm for every compact interval I .

There is at least one pitfall in the extension problem which should be mentioned: If $f \in C_{\text{op}}^1(I)$ for some compact interval I , then the trivial extension – setting f to zero outside I – may not produce an element in $C_F^1(\mathbb{R})$, even though f has (other) extensions in $C_F^1(\mathbb{R})$. For a specific example take $f(x) = (1 - x^2)_+$. On $[-1, 1]$ the function is C^∞ , hence operator differentiable (because it has a C^∞ -extension with compact support on \mathbb{R}). However, as defined, the function f does not belong to $C_F^1(\mathbb{R})$. By calculation

$$\widehat{f}(t) = 4t^{-3} \sin t - 4t^{-2} \cos t,$$

so that $\widehat{f} \in L^1(\mathbb{R})$. However, $(\text{id}) \widehat{f} \notin L^1(\mathbb{R})$ because $f' \notin C_0(\mathbb{R})$.

Perturbations 3.14. If $\{\alpha_t \mid t \in \mathbb{R}\}$ is a strongly continuous one parameter group of automorphisms of a C^* -algebra \mathfrak{A} , and δ denotes the closed, densely defined derivation of \mathfrak{A} for which $\alpha_t = \exp(t\delta)$, then, since for every f in C_{op}^1 we have

$$\alpha_t(dt_T(S)) = df_{\alpha_t(T)}(\alpha_t(S)),$$

whence $df_{\alpha_t(T)} = \alpha_t \circ df_T \circ \alpha_{-t}$, it follows that the function $t \rightarrow df_{\alpha_t(T)}$ is differentiable. The differential quotient at zero is the (possibly unbounded) operator on $\mathfrak{D}(\delta)$ given by

$$\frac{d}{dt} \Big|_{t=0} df_{\alpha_t(T)}(S) = \delta(df_T(S)) - df_T(\delta(S)).$$

If $f \in C_F^2(I)$ (or, more generally $f \in C_{op}^2$) this implies that if $T \in \mathfrak{D}(\delta)$, then a perturbation of df_T in the direction $\delta(T)$ is computable as

$$df_{(T+\varepsilon\delta(T))}(S) = df_T(S) + \varepsilon(\delta(df_T(S)) + df_T(\delta(S))) + o(\varepsilon) \|S\|.$$

In particular, if $\delta(A) = [R, A]$ for some R in \mathfrak{A} , then we find (again when $f \in C_F^2(I)$) that

$$df_{T+\varepsilon[R, T]}(S) = df_T(S) + \varepsilon([R, df_T(S)] + df_T([R, S])) + o(\varepsilon) \|S\|.$$

§ 4. Special Cases

The Löwner Kernel 4.1. For an interval I in \mathbb{R} we define for each f in $C^1(I)$ its *Löwner kernel* $f^{[1]}$ in $C(I^2)$ by

$$f^{[1]}(x, y) = \begin{cases} \frac{f(x) - f(y)}{x - y} & \text{if } x \neq y \\ f'(x) & \text{if } x = y \end{cases}$$

From this we obtain for each finite set $\{\lambda_1, \dots, \lambda_n\}$ in I an $n \times n$ matrix L , where $L_{ij} = f^{[1]}(\lambda_i, \lambda_j)$. These matrices were used by K. Löwner in [20] to show that f is operator monotone (increasing) precisely if each matrix L is positive (definite). The reason for this becomes transparent when L is seen as a differential, cf. [17, Theorem 3.2]. Thus, if $T = T^* \in \mathbf{M}_n$ and T is diagonal with eigenvalues $T_{ii} = \lambda_i$, $1 \leq i \leq n$, then for every S in \mathbf{M}_n we have

$$df_T(S)_{ij} = f^{[1]}(\lambda_i, \lambda_j) S_{ij} = L_{ij} S_{ij}.$$

In other words, $df_T(S)$ is the Schur (or Hadamard) product between S and the Löwner matrix L .

An infinite-dimensional analogue of this matrix calculation was given by Widom in [33] (and found again in [17, 3.5]): If $\mathcal{H} = L_\mu^2(I)$ for some Borel measure μ on the bounded interval I , and if $T = \text{id}$ (regarded as a multiplication operator), then for each Hilbert-Schmidt operator $S = S_k$, where k denotes the kernel in $L_{\mu \otimes \mu}^2(I^2)$ determining S , we get

$$df_T(S) = S_{f^{[1]k}}$$

for f in $C_{\text{op}}^1(I)$, i.e. $df_T(S)$ is the Hilbert-Schmidt operator with (pointwise) product kernel $f^{[1]}k$. Note that even though the above result was only established for f in $C_F^1(I)$, it does apply to any function f in $C^1(I)$, and it will describe the action of the differential (restricted to the class of Hilbert-Schmidt operators) when $f \in C_{\text{op}}^1(I)$. The reference [33] contains a succinct commentary by Peller on the ramifications of this approach to operator differentiability.

The Square Root 4.2. For elementary functions the differential can sometimes be found as a solution to an operator equation. Thus, if $n^{\text{th}}r(t) = t^{1/n}$ for n in \mathbb{N} , and similarly $n^{\text{th}}p(t) = t^n$, the fact that $n^{\text{th}}p \circ n^{\text{th}}r = \text{id}$ gives the following formula, (cf. Remark 2.5).

$$\begin{aligned} S &= d(n^{\text{th}}p)_{T^{1/n}}(d(n^{\text{th}}r)_T(S)) \\ &= \sum_{k=1}^n T^{(k-1)/n} d(n^{\text{th}}r)_T(S) T^{(n-k)/n} \end{aligned}$$

for S in $\mathcal{B}(\mathcal{H})$ and T invertible in $\mathcal{B}(\mathcal{H})_+$ (i.e. $\text{sp}(T) \subset]0, \infty[$). Thus for $S = S^*$ the element $d(n^{\text{th}}r)_T(S)$ is the unique self-adjoint solution to the operator equation

$$\sum_{k=1}^n T^{(k-1)/n} X Y^{(n-k)/n} = S.$$

In particular, with $n = 2$ (and renaming $\text{sqr}(t) = t^{1/2}$) we see that $d(\text{sqr})_T(S)$ is the unique self-adjoint solution to the operator equation

$$T^{1/2}X + XT^{1/2} = S.$$

This equation was studied in [23], and by a result due to VanDaele it follows that

$$d(\text{sqr})_T(S) = \frac{1}{2} \int_{-\infty}^{\infty} T^{it/2} T^{-1/4} S T^{-1/4} T^{-it/2} (\cosh(\pi t))^{-1} dt.$$

The Logarithm 4.3. The fact that $\exp \circ \log = \text{id}$ on $]0, \infty[$ gives the formula

$$S = d(\exp)_{\log T}(d(\log))_T(S)$$

for every invertible T in $\mathcal{B}(\mathcal{H})_+$. Inserting the expression for the differential of \exp , cf. the proof of Theorem 3.2, we see that for $S = S^*$ the differential $d(\log)_T(S)$ is the solution to the integral equation

$$(*) \quad \int_0^1 T^s X T^{1-s} ds = S.$$

In order to find an explicit formula for the differential of \log we note that since the function is operator monotone on $]0, \infty[$ it must have an integral representation of extremal functions; and a few experiments show that, in fact,

$$\log(t) = \int_0^\infty \frac{t-1}{(s+t)(s+1)} ds.$$

It follows that

$$d(\log)_T(S) = \int_0^\infty (s1 + T)^{-1} S (s1 + T)^{-1} ds.$$

The integral equation $(*)$ is also studied in [18, § 2], where other formulas for the solution, akin to VanDaele's in 4.2, are presented.

Operator Monotone Functions 4.4. Both the logarithm (in the translated form $f(x) = \log(1+x)$) and all the root functions are operator monotone on the (closed) positive half-axis. The estimates for commutators obtained from Corollary 3.6, using the operator norm of the differential, will therefore take the form

$$\|[S, f(T)]\| \leq \|f'(T)\| \|[S, T]\|,$$

cf. Remark 3.5. However, the root functions are operator monotone on \mathbb{R}_+ , but have a vertical tangent at 0, so this formula is of little help near zero. In [22, Lemma 2.1] and again in [24, Lemma 6.2] a much improved estimate, also discovered by Davidson, is given. A simple proof, presented in [25], shows that

$$\|[S, f(T)]\| \leq \frac{5}{4} f(\|[S, T]\|)$$

for every $T \geq 0$ and S with $\|S\| \leq 1$. Although the argument is only given for root functions, it holds for every operator monotone function f on \mathbb{R}_+ , because each such is also subadditive.

The constant $\frac{5}{4}$ in the estimate above is not optimal, cf. [24, Remark 6.5], the current record for the square root function being $2/\sqrt{\pi} \sim 1.128$. If S is unitary, the constant may be taken equal to 1, and this is probably also true for all other operators.

References

- [1] Ando, T., Topics on Operator Inequalities, *Lecture Notes* (mimeographed), Hokkaido University, Sapporo, 1978.
- [2] ———, Concavity of certain maps of positive definite matrices and applications to Hadamard products, *Lin. Algebra Appl.*, **26** (1979), 203–241.
- [3] Araki, H., Expansional in Banach algebras, *Ann. Sci. École Norm. Sup.*, **6** (1973), 229–241.
- [4] Araki, H. and Hansen, F., Jensen's operator inequality for functions of several variables, *Proc. Amer. Math. Soc.*, (to appear).
- [5] Bapat, J. and Sherman, S., Monotone and convex operator functions, *Trans. Amer. Math. Soc.*, **79** (1955), 58–71.
- [6] Bhatia, R. and Sinha, K. B., Variation of real powers of positive operators, *Indiana Univ. Math. J.*, **43** (1994), 913–925.
- [7] Bhatia, R. and Holbrook, J. A., Fréchet derivatives of the power function, *Indiana Univ. Math. J.* to appear.
- [8] Bist, V. and Vasudeva, H. L., Second order perturbation bounds, *Publ. R.I.M.S., Kyoto Univ.*, **33** (1997), 59–71.
- [9] Boyadzhiev, K. N., Mean value theorems for traces, *Math. Jap.*, **38** (1993), 217–224.
- [10] Bratteli, O. and Robinson, D. W. Unbounded derivations of C^* -algebras, *Comm. Math. Phys.*, **42** (1975), 253–268.
- [11] Brown, L. G. and Kosaki, H., Jensen's inequality in semi-finite von Neumann algebras, *J. Operator Theory*, **23** (1990), 3–19.
- [12] Davis, C., Notions generalizing convexity for functions defined on spaces of matrices, *Proc. Amer. Math. Soc. Symp. (Convexity)*, **7** (1963), 187–201.
- [13] Flett, T. M., *Differential Analysis*, Cambr. Univ. Press, Cambridge, 1980.
- [14] Hansen, F., Jensen's operator inequality for functions of two variables, *Proc. Amer. Math. Soc.*, **125** (1997), 2093–2102.
- [15] ———, Operator convex functions of several variables, *Publ. RIMS, Kyoto Univ.*, **33** (1997), 443–463.
- [16] Hansen, F. and Pedersen, G. K., Jensen's inequality for operators and Löwner's theorem, *Math. Ann.*, **258** (1982), 229–241.
- [17] ———, Perturbation formulas for traces on C^* -algebras, *Publ. RIMS, Kyoto Univ.*, **31** (1995), 169–178.
- [18] Hiai, F. and Kosaki, H., Means for matrices and comparison of their norms, *Indiana Univ. Math. J.*, **48** (1999), 899–936.
- [19] Hille, E. and Phillips, R. S., *Functional Analysis and Semi-Groups*, *Amer. Math. Soc. Coll. Publ.*, **31**, Providence R.I., revised edition 1957.
- [20] Löwner, K., Über monotone Matrixfunktionen, *Math. Z.*, **38** (1934), 177–216.
- [21] McIntosh, A., Functions and derivations of C^* -algebras, *J. Funct. Anal.*, **30** (1978), 264–275.
- [22] Olsen, C. L. and Pederson, G. K., Corona C^* -algebras and their applications to lifting problems, *Math. Scand.*, **64** (1989), 63–86.
- [23] Pedersen, G. K., On the operator equation $HT + TH = 2K$, *Indiana Univ. Math. J.*, **25** (1976), 1029–1033.

- [24] Pedersen, G. K., The corona construction, "Operator Theory: Proc. 1988 GPOTS-Wabash Conf.", Eds. Conway, J. B. and Morrel, B. B., *Pitman Res. Notes in Math. Sci.*, **225**, Longman Sci. & Tech., Harlow, Essex UK, (1991), 49-92.
- [25] Pedersen, G. K., A commutator inequality, "Operator Algebras, Mathematical Physics, and Low Dimensional Topologies", Eds. Herman, R. and Tanbay, B., *Res. Notes in Math.*, **5**, A. K. Peters, Wellesly MA, (1993), 233-235.
- [26] Petz, D., Spectral scale of self-adjoint operators and trace inequalities, *J. Math. Anal. Appl.*, **109** (1985), 74-82.
- [27] ———, Jensen's inequality for positive contractions on operator algebras, *Proc. Amer. Math. Soc.*, **99** (1987), 273-277.
- [28] Powers, R. T., A remark on the domain of an unbounded derivation of a C^* -algebra, *J. Funct. Anal.*, **18** (1975), 85-95.
- [29] Reed, M. and Simon, B., *Methods of Modern Mathematical Physics II: Fourier Analysis, Self-Adjointness*, Academic Press, New York, 1975.
- [30] Sakai, S., *Operator Algebras in Dynamical Systems*, Cambr. Univ. Press, Cambridge, 1991.
- [31] Suzuki, M., *Exponential Product Formulas and Quantum Analysis*, Univ. Tokyo, Tokyo, 1996.
- [32] Vasudeva, H. L., On monotone matrix functions of two variables, *Trans. Amer. Math. Soc.*, **176** (1973), 305-318.
- [33] Widom, H., *When are differentiable functions differentiable?*, "Linear and Complex Analysis Problem Book **3**, Part 1", Eds. Havin, V. P. et al., *Lecture Notes in Math.*, **1573**, Springer-Verlag, Berlin-Heidelberg-New York, (1994), 266-271.
- [34] Wigner, E. P. and von Neumann, J., Significance of Loewner's theorem in the quantum theory of collisions, *Ann. Math.*, **59** (1954), 418-433.

