STUDIA MATHEMATICA 223 (3) (2014)

Submultiplicative functions and operator inequalities

by

Hermann König (Kiel) and Vitali Milman (Tel Aviv)

Dedicated to Prof. A. Pietsch on the occasion of his 80th birthday

Abstract. Let $T : C^1(\mathbb{R}) \to C(\mathbb{R})$ be an operator satisfying the "chain rule *inequality*"

$$T(f \circ g) \le (Tf) \circ g \cdot Tg, \quad f, g \in C^1(\mathbb{R}).$$

Imposing a weak continuity and a non-degeneracy condition on T, we determine the form of all maps T satisfying this inequality together with $T(-\mathrm{Id})(0) < 0$. They have the form

$$Tf = \begin{cases} (H \circ f/H) f'^p, & f' \ge 0, \\ -A(H \circ f/H)|f'|^p, & f' < 0, \end{cases}$$

with $p > 0$, $H \in C(\mathbb{R})$, $A \ge 1$. For $A = 1$, these are just the solutions of the chain rule operator *equation*. To prove this, we characterize the submultiplicative, measurable functions K on $\mathbb{R}$ which are continuous at 0 and 1 and satisfy $K(-1) < 0 < K(1)$. Any such map K has the form

$$K(\alpha) = \begin{cases} \alpha^p, & \alpha \ge 0, \\ -A|\alpha|^p, & \alpha < 0, \end{cases}$$

with $A \ge 1$ and $p > 0$. Corresponding statements hold in the supermultiplicative case with $0 < A \le 1$.

1. Introduction and results. Let $S : C^1(\mathbb{R}) \to C(\mathbb{R})$ be an operator satisfying the "chain rule"

$$S(f \circ g) = (Sf) \circ g \cdot Sg, \quad f, g \in C^1(\mathbb{R}). \tag{1.1}$$

It was shown in [AKM] that if S is non-degenerate and if the image of S contains functions with negative values, then S has the form

$$Sf = \frac{H \circ f}{H} |f'|^p \operatorname{sgn} f' \tag{1.2}$$

where $H \in C(\mathbb{R})$, $H > 0$ and $p > 0$ are suitably chosen. A priori, S is not assumed to be a continuous operator, though a posteriori this is a consequence of formula (1.2). Letting G be an antiderivative of $H^{1/p} > 0$, G is a

2010 *Mathematics Subject Classification*: Primary 39B42; Secondary 26A24, 46E15.
Key words and phrases: submultiplicative functions, chain rule, operator inequalities.

DOI: 10.4064/sm223-3-3

strictly monotone C^1-function with

$$Sf = \left|\frac{d(G\circ f)}{dG}\right|^p \operatorname{sgn}\left(\frac{d(G\circ f)}{dG}\right).$$

In this sense, the solutions S of (1.1) are pth powers of some transformed derivative. Equation (1.1) is remarkably rigid and stable as shown in [KM]: The solutions of the more general equation

$$V(f\circ g) = (S_1 f)\circ g \cdot S_2 g, \qquad f, g \in C^1(\mathbb{R}),$$

for operators $V, S_1, S_2 : C^1(\mathbb{R}) \to C(\mathbb{R})$ under some weak non-degeneracy conditions are just natural modifications of the solutions S of (1.1): there are continuous functions $c_1, c_2 \in C(\mathbb{R})$ such that

$$Vf = c_1\circ f \cdot c_2 \cdot Sf, \qquad S_1 f = c_1 \circ f \cdot Sf, \qquad S_2 f = c_2 \cdot Sf.$$

Clearly, these are solutions of the more general equation, but there are no others. So (1.1) is rigid. If the chain rule is perturbed by a function B of $(x, (f\circ g)(x), g(x))$, i.e.

$$S(f\circ g) = Sf\circ g\cdot Sg + B(\cdot, f\circ g(\cdot), g(\cdot)),$$

then under weak conditions on S, we have $B = 0$, i.e. equation (1.1) is superstable.

In this paper we study another weakening of the chain rule equation showing again a very stable behavior. We replace the equality requirement by an inequality requirement: we study operators $T : C^1(\mathbb{R}) \to C(\mathbb{R})$ satisfying the chain rule *inequality*

$$T(f\circ g) \le (Tf)\circ g \cdot Tg, \qquad f, g \in C^1(\mathbb{R}), \tag{1.3}$$

and determine its solutions under mild continuity and non-degeneracy assumptions on T, if the image of T also contains functions with negative values. In fact, solutions of the operator inequality (1.3) are bounded by suitable solutions of the operator equation (1.1). This is a similar phenomenon to the one in Gronwall's inequality [G] in its (weaker) differential form, where any solution of a differential inequality is bounded by a solution of the corresponding differential equation. Again the form of solutions T of (1.3) is very similar to those for (1.1), up to some constant $A \ge 1$ in the case of $f' < 0$. This again exhibits a very stable behavior of the chain rule in operator form.

The dependence on the derivative f' in (1.2), and also for solutions of (1.3), is of power type; to prove this, we need a suitable result on submultiplicative functions on $\mathbb{R}$ (not only on $\mathbb{R}_+$) which seems to be of some independent interest. Let $I \subset \mathbb{R}$ be an open interval. Recall that $K : I \to \mathbb{R}$ is *submultiplicative* if

$$K(\alpha\beta) \le K(\alpha)K(\beta), \qquad \alpha, \beta \in I. \tag{1.4}$$

THEOREM 1.1. *Let $K : \mathbb{R} \to \mathbb{R}$ be a measurable and submultiplicative function which is continuous at* 0 *and at* 1*. Assume further that $K(-1) < 0 < K(1)$. Then there exists $p > 0$ such that*

$$K(\alpha) = \begin{cases} \alpha^p, & \alpha \geq 0, \\ -A|\alpha|^p, & \alpha < 0. \end{cases}$$

Here $K(-1) \leq -1$.

EXAMPLES. The measurability assumption is necessary since there are lots of non-measurable submultiplicative functions: Let $f : \mathbb{R} \to \mathbb{R}$ be a non-measurable *additive* function and define

$$K(\alpha) = \begin{cases} \exp(f(\ln \alpha)), & \alpha \geq 0, \\ -A \exp(f(\ln |\alpha|)), & \alpha < 0, \end{cases}$$

with $A > 1$ and $K(0) := 0$. Then K is non-measurable and submultiplicative with $K(-1) < 0 < K(1)$. By a well-known result of Banach [B] and Sierpiński [S], measurable additive functions are linear, in which case this example would have the form given in Theorem 1.1.

Also, for any $d > 0$, $c \geq 0$, $c \neq d$,

$$K(\alpha) = \begin{cases} 1, & \alpha = 1, \\ -c, & \alpha = 0, \\ -d, & \alpha \notin \{0, 1\}, \end{cases}$$

is measurable and submultiplicative with $K(-1) < 0 < K(1)$, but continuous neither at 0 nor at 1.

There are numerous measurable submultiplicative functions from $\mathbb{R}_{\geq 0}$ to $\mathbb{R}_{\geq 0}$ different from powers α^p which are continuous at 0 and 1 (cf. Remark (b) after Theorem 1.2). However, they cannot be extended to measurable submultiplicative functions on $\mathbb{R}$ which also attain negative values. Only measurable *multiplicative* maps on $\mathbb{R}_{\geq 0}$ extend to such maps on $\mathbb{R}$.

REMARK. A corresponding result holds for *supermultiplicative* functions, i.e. with $K(\alpha\beta) \geq K(\alpha)K(\beta)$, under the same assumptions. The result is formally the same, except that now $0 < A \leq 1$. The proof is similar.

Studying the "chain rule operator inequality" (1.3), we will impose a non-degeneracy and a mild continuity condition on T.

DEFINITION 1.1. An operator $T : C^1(\mathbb{R}) \to C(\mathbb{R})$ is *non-degenerate* provided that for any open interval $I \subset \mathbb{R}$ and any $x \in I$ there exists $g \in C^1(\mathbb{R})$ with $g(x) = x$, $\mathrm{Im}(g) \subset I$ and $(Tg)(x) > 1$.

DEFINITION 1.2. An operator $T : C^1(\mathbb{R}) \to C(\mathbb{R})$ is *pointwise continuous* if for any sequence of functions $f_n \in C^1(\mathbb{R})$ and $f \in C^1(\mathbb{R})$ with $f_n \to f$ and $f_n' \to f'$ uniformly on all compact subsets of $\mathbb{R}$, we have the pointwise convergence $\lim_{n\to\infty}(Tf_n)(x) = (Tf)(x)$ for all $x \in \mathbb{R}$.

The continuity assumption on T is rather weak, and the non-degeneracy assumption just means that the image of T should contain sufficiently many functions; it is a very weak surjectivity type assumption. Our main result concerning the chain rule inequality then states:

THEOREM 1.2. *Let $T : C^1(\mathbb{R}) \to C(\mathbb{R})$ be an operator such that the chain rule inequality holds:*

$$T(f \circ g) \leq (Tf) \circ g \cdot Tg, \qquad f, g \in C^1(\mathbb{R}). \tag{1.3}$$

Assume in addition that T is non-degenerate and pointwise continuous. Suppose also that there exists $x \in \mathbb{R}$ such that

$$T(-\mathrm{Id})(x) < 0.$$

Then there exists a continuous function $H \in C(\mathbb{R})$, $H > 0$, a number $p > 0$ and some $A \geq 1$ such that T has the form

$$Tf = \begin{cases} (H \circ f/H) f'^p, & f' \geq 0, \\ -A(H \circ f/H)|f'|^p, & f' < 0. \end{cases} \tag{1.5}$$

REMARKS. (a) Equation (1.5) means that $Tf \leq Sf$ with Sf as given by (1.2), i.e. any solution of the chain rule *inequality* is bounded from above by a corresponding solution of the chain rule *equality* (1.1) for which $A = 1$. Note that $-A = T(-\mathrm{Id})(0) \leq -1$ and hence $T(-\mathrm{Id})(0) = -1$ implies that $T = S$. Further, $T(-\mathrm{Id})(x) < 0$ for all $x \in \mathbb{R}$. With (1.2) and (1.5) we have precise formulas for all solutions of (1.1) and (1.3) under the assumed conditions.

(b) The condition $T(-\mathrm{Id})(x) < 0$ guarantees that there are sufficiently many negative functions in the range of T. If it is violated, there are many solutions of (1.3) different from (1.5) which only allow for *non-negative* functions in the range of T. Examples of non-negative solutions of (1.3) can be given in the form

$$Tf(x) = F(x, f(x), |f'(x)|),$$

where $F : \mathbb{R}^2 \times \mathbb{R}_+ \to \mathbb{R}_+$ is a continuous function satisfying

$$F(x, z, \alpha\beta) \leq F(y, z, \alpha) F(x, y, \beta) \tag{1.6}$$

for all $x, y, z \in \mathbb{R}$ and $\alpha, \beta \geq 0$. One may take e.g.

$$F(y, z, \alpha) = \exp(d(y, z)) \cdot \max(\alpha^p, \alpha^q),$$

or

$$F(y, z, \alpha) = \exp(d(y, z)) + \max(\alpha^p, \alpha^q)$$

where $0 < p, q$ and d is either a metric on $\mathbb{R}$ or $d(y, z) = z - y$. Moreover, given any continuous submultiplicative function $f : \mathbb{R}_+ \to \mathbb{R}_+$ and a continuous function $F : \mathbb{R}^2 \times \mathbb{R}_+ \to \mathbb{R}_+$ satisfying (1.6), the composition $f \circ F$ will also satisfy (1.6). By [GMP], for any fixed $c \geq e$ the function f given by

$f(x) = \ln(c+x)$ is submultiplicative and monotone on $\mathbb{R}_+$. The same is true for

$$f(x) = \max(x^p, x^q) \ln(c+x)^\gamma, \qquad x \geq 0,$$

where $0 < p, q$, $c \geq e$ and $\gamma \geq 0$. This yields many examples of operators T satisfying (1.3) but not $T(-\mathrm{Id})(x) < 0$ which are not of the form given by (1.5).

(c) Note that the operator T in Theorem 1.2 is not assumed to be linear. For the solution operator S of the chain rule *equation* (1.1) in [AKM] no continuity assumption is needed. For the chain rule *inequality*, we have less information. In our proof, we need a weak continuity assumption instead.

(d) A similar result is valid for the super chain rule inequality

$$T(f \circ g) \geq (Tf) \circ g \cdot Tg, \qquad f, g \in C^1(\mathbb{R}),$$

under the same assumptions as in Theorem 1.2 except for a slightly stronger condition of non-degeneracy: In addition we require that for any open interval $I \subset \mathbb{R}$ and any $x \in I$ there exists $h \in C^1(\mathbb{R})$ with $h(x) = x$, $\mathrm{Im}(h) \subset I$ and $(Th)(x) < 0$. In the case of the sub chain rule inequality (1.3), the latter fact can be proved from the assumptions (see Lemma 3.2). The solutions of the super chain rule inequality have the same form as (1.5) except that now $0 < A \leq 1$. The proof is similar to the one of Theorem 1.2, except that Lemma 3.2 is replaced by the stronger assumption of non-degeneracy.

(e) The assumption in Theorem 1.2 on $T(-\mathrm{Id})(x) < 0$ for some x will be used only in Lemma 3.2 below.

2. Submultiplicative functions on $\mathbb{R}$. To prove Theorem 1.1, we first collect a few simple facts on submultiplicative and subadditive functions.

LEMMA 2.1. *Let $K : \mathbb{R}_{>0} \to \mathbb{R}$ be submultiplicative with $K(1) > 0$. Assume that K is continuous at 1. Then $K(1) \geq 1$ and $K|_{\mathbb{R}_{>0}} > 0$.*

Proof. Since $0 < K(1) \leq K(1)^2$, we have $K(1) \geq 1$. Since K is continuous at 1, there is $\epsilon > 0$ such that $K|_{[1/(1+\epsilon),1+\epsilon]} > 0$. For any θ in $[1/(1+\epsilon), 1+\epsilon]$, $K(\theta) > 0$ and $K(1/\theta) > 0$; hence $0 < K(\theta) \leq K(1/\theta)K(\theta^2)$ implies that $K(\theta^2) > 0$, i.e. $K|_{[1/(1+\epsilon)^2,(1+\epsilon)^2]} > 0$. Inductively we conclude that $K|_{\mathbb{R}_{>0}} > 0$ since $\mathbb{R}_{>0} = \bigcup_{n\in\mathbb{N}}[1/(1+\epsilon)^n, (1+\epsilon)^n]$. ∎

LEMMA 2.2. *Let $K : \mathbb{R} \to \mathbb{R}$ be submultiplicative with $K(-1) < 0 < K(1)$. Assume that K is continuous at 0 and 1. Then $K(0) = 0$ and there is $\epsilon > 0$ such that $0 < K(x) < 1$ for all $x \in (0, \epsilon)$, and $1 < K(x) < \infty$ for all $x \in (1/\epsilon, \infty)$. Further, $K|_{\mathbb{R}_{>0}} > 0 > K|_{\mathbb{R}_{<0}}$.*

Proof. The inequality $K(0) = K((-1)\cdot 0) \leq K(-1)K(0)$ with $K(-1) < 0$ shows that $K(0) > 0$ is impossible. Hence $K(0) \leq 0$. By Lemma 2.1, $K|_{\mathbb{R}_{>0}} > 0$. Since K is assumed to be continuous at 0, we conclude that

$K(0)=0$ and hence there is $\epsilon>0$ with $0<K|_{(0,\epsilon)}<1$. But for any $\theta>0$, $1\le K(1)\le K(\theta)K(1/\theta)$, which implies that $K|_{(1/\epsilon,\infty)}>1$. Further, for any $\theta>0$, $K(-\theta)\le K(-1)K(\theta)<0$, i.e. $K|_{\mathbb{R}_{<0}}<0$. ∎

Remark. Instead of $K(-1)<0$, it suffices to assume that $K(-t_0)<0$ for some $t_0>0$.

A function $f:\mathbb{R}\to\mathbb{R}$ is *subadditive* provided that

$$f(s+t)\le f(s)+f(t),\qquad s,t\in\mathbb{R}.$$

We recall a few well-known facts on subadditive functions in the following lemma (cf. Hille–Phillips [HP, Chapter VII]):

Lemma 2.3. *Assume $f:\mathbb{R}\to\mathbb{R}$ is measurable and subadditive. Define $p:=\sup_{t<0}f(t)/t$ and $q:=\inf_{t>0}f(t)/t$. Then f is bounded on compact intervals, $-\infty<p\le q<\infty$ and $f(0)\ge 0$. Moreover, the limits $p=\lim_{t\to-\infty}f(t)/t$ and $q=\lim_{t\to\infty}f(t)/t$ both exist. If additionally $f(0)=0$ and f is continuous at 0, then f is continuous on $\mathbb{R}$.*

As a consequence of Lemma 2.3, for any $t>0$,

$$f(t)=qt+b(t),\qquad b(t)\ge 0,\qquad \lim_{t\to\infty}\frac{b(t)}{t}=0$$

and for any $t<0$,

$$f(t)=pt+a(t),\qquad a(t)\ge 0,\qquad \lim_{t\to-\infty}\frac{a(t)}{t}=0.$$

If $f\ge 0$ near ∞ and $f\le 0$ near $-\infty$, Lemma 2.3 yields $0\le p\le q<\infty$.

Proof of Theorem 1.1. (a) Assume that $K:\mathbb{R}\to\mathbb{R}$ is measurable, submultiplicative, continuous at 0 and 1, and $K(-1)<0<K(1)$. By Lemma 2.2, $K(0)=0$, $K|_{\mathbb{R}_{>0}}>0$ and $0<K<1$ in $(0,\epsilon)$ and $1<K<\infty$ in $(1/\epsilon,\infty)$ for a suitable $0<\epsilon<1$. Let $f(t):=\ln K(\exp(t))$, $t\in\mathbb{R}$. Then f is measurable and subadditive, and by Lemma 2.3 we have

$$-\infty<p:=\sup_{t<0}\frac{f(t)}{t}=\lim_{t\to-\infty}\frac{f(t)}{t}\le q:=\inf_{t>0}\frac{f(t)}{t}=\lim_{t\to\infty}\frac{f(t)}{t}<\infty.$$

Since f is negative near $-\infty$ and positive near ∞, we have $0\le p\le q<\infty$ with

$$f(t)=\begin{cases} pt+a(t), & t<0,\\ qt+b(t), & t>0,\end{cases}$$

where $a(t)\ge 0$ for $t<0$, $b(t)\ge 0$ for $t>0$ and $\lim_{t\to-\infty}a(t)/t=\lim_{t\to\infty}b(t)/t=0$. This means that for $0<\alpha<1$,

$$K(\alpha)=\exp(f(\ln(\alpha)))=\alpha^p\exp(a(\ln(\alpha)))\ge\alpha^p,$$

and for $1<\alpha<\infty$,

$$K(\alpha)=\exp(f(\ln(\alpha)))=\alpha^q\exp(b(\ln(\alpha)))\ge\alpha^q.$$

We note that

$$\lim_{\alpha\to 0}\alpha^{\epsilon}\frac{K(\alpha)}{\alpha^p}=\lim_{\alpha\to\infty}\alpha^{-\epsilon}\frac{K(\alpha)}{\alpha^q}=0\qquad\text{for any }\epsilon>0,$$

although we do not need it.

(b) We claim that $p=q>0$. To prove this, we use the fact that $K|_{\mathbb{R}_{<0}}<0<K|_{\mathbb{R}_{>0}}$ by Lemma 2.2. Hence for all $y<0<x$, by submultiplicativity,

$$K(xy)\le K(x)K(y),\qquad |K(xy)|\ge |K(y)|K(x).$$

Since $K(1)\ge 1$ and $1\le K((-1)^2)\le K(-1)^2$, we have $K(-1)\le -1$. Therefore

$$K(-1)\le K(1)K(-1),\qquad |K(-1)|\ge |K(-1)|K(1),$$

implying that $K(1)\le 1$, i.e. $K(1)=1$. Thus $f(0)=0$. Fix $t<0$ and choose $y<-1$ and $0<x<1$ with $t=xy$. Then by submultiplicativity,

$$\begin{gathered}K(t)=K((-1)x|y|)\le K(-1)K(x)K(|y|),\\ |K(t)|\ge |K(-1)|K(x)K(|y|)\ge x^p|y|^q=|t|^q x^{p-q}.\end{gathered}$$

Assuming $p\ne q$, i.e. $p<q$, and letting x tend to 0 (and hence y to $-\infty$) would give the contradiction that $|K(t)|=\infty$. Hence $0\le p=q<\infty$. In fact, $0<p=q$ since K is continuous at 0 with $K(0)=0$ and $K(x)\ge x^p$ for $0<x<1$.

(c) Now let $g(t):=\ln|K(-\exp(t))|$ for $t\in\mathbb{R}$. Then for any $s,t\in\mathbb{R}$,

$$\begin{aligned}g(s+t)&=\ln|K(-\exp(s)\exp(t))|\\ &\ge \ln|K(-\exp(s))|+\ln K(\exp(t))\\ &=g(s)+f(t)=g(s)+pt+a(t)\end{aligned}\tag{2.1}$$

with $a(t)\ge 0$ for all $t\in\mathbb{R}$ and $\lim_{t\to\pm\infty}a(t)/t=0$. For $t>0$, the function a was called b before. Since $f(0)=0$, we have $a(0)=0$. Setting $s=0$ in (2.1) and $t=-s$ in (2.1) and then renaming s as t yields

$$g(t)\ge g(0)+pt+a(t),\qquad g(0)\ge g(t)-pt+a(-t),$$

hence

$$g(0)+pt+a(t)\le g(t)\le g(0)+pt-a(-t).$$

Since $a\ge 0$ on $\mathbb{R}$, this implies that $a=0$ on $\mathbb{R}$ and $g(t)=g(0)+pt$ for all $t\in\mathbb{R}$. Also $f(t)=pt$ for all $t\in\mathbb{R}$. Therefore for all $\beta<0<\alpha$,

$$K(\alpha)=\alpha^p\exp(a(\ln(\alpha)))=\alpha^p,\qquad |K(\beta)|=\exp(g(\ln(|\beta|)))=\exp(g(0))|\beta|^p.$$

Hence $\exp(g(0))=|K(-1)|\ge 1$, i.e. $g(0)\ge 0$. Thus $K(\beta)=K(-1)|\beta|^p$, proving Theorem 1.1. ∎

3. Localization. We will now show that the operator T in Theorem 1.2 is locally defined, more precisely $(Tf)(x)$ only depends on x, $f(x)$ and $f'(x)$.

Proposition 3.1. *Let $T : C^1(\mathbb{R}) \to C(\mathbb{R})$ be non-degenerate, pointwise continuous and satisfy the chain rule inequality*

$$T(f \circ g) \le (Tf) \circ g \cdot Tg, \qquad f, g \in C^1(\mathbb{R}). \tag{3.1}$$

Assume also that there is $x_0 \in \mathbb{R}$ such that $T(-\mathrm{Id})(x_0) < 0$. Then there is a function $F : \mathbb{R}^3 \to \mathbb{R}$ such that for all $f \in C^1(\mathbb{R})$ and all $x \in \mathbb{R}$,

$$(Tf)(x) = F(x, f(x), f'(x)).$$

We first strengthen the condition of non-degeneracy.

Lemma 3.2. *Under the assumptions of Proposition 3.1, for any open interval $I \subset \mathbb{R}$ and for any $x \in I$ there exists $g \in C^1(\mathbb{R})$ with $g(x) = x$ and $\mathrm{Im}(g) \subset I$ such that $(Tg)(x) < 0$.*

Recall that non-degeneracy required the existence of a similar function g but with $(Tg)(x) > 1$.

Proof of Lemma 3.2. (a) By (3.1), $T(\mathrm{Id})(x) \le T(\mathrm{Id})(x)^2$ for any $x \in \mathbb{R}$. Hence $T(\mathrm{Id})(x) \ge 1$ or $T(\mathrm{Id})(x) \le 0$. If there were $x_0 \in \mathbb{R}$ with $T(\mathrm{Id})(x_0) \le 0$, use the fact that T is non-degenerate to find $g \in C^1(\mathbb{R})$ with $g(x_0) = x_0$ and $(Tg)(x_0) > 1$. Then

$$1 \le (Tg)(x_0) = T(g \circ \mathrm{Id})(x_0) \le (Tg)(x_0) T(\mathrm{Id})(x_0) \le 0$$

yields a contradiction. Hence $T(\mathrm{Id})(x) \ge 1$ for all $x \in \mathbb{R}$.

(b) Similarly, $T(-\mathrm{Id})(x) < 0$ for all $x \in \mathbb{R}$: Since

$$1 \le T(\mathrm{Id})(x) = T((-\mathrm{Id})^2)(x) \le T(-\mathrm{Id})(-x) T(-\mathrm{Id})(x),$$

we have $T(-\mathrm{Id})(x) \ne 0$ for all $x \in \mathbb{R}$. By assumption, there is $x_0 \in \mathbb{R}$ with $T(-\mathrm{Id})(x_0) < 0$. Hence $T(-\mathrm{Id})(x) < 0$ since else by the continuity of the function $T(-\mathrm{Id})$, there would be $y \in \mathbb{R}$ with $T(-\mathrm{Id})(y) = 0$, which we just showed to be impossible.

(c) Take any open interval $I \subset \mathbb{R}$ and $x_1 \in I$. Then there is $\epsilon > 0$ such that $J_\epsilon := (x_1 - \epsilon, x_1 + \epsilon) \subset I$. Consider $\tilde{J} := J - \{x_1\} = (-\epsilon, \epsilon)$. By non-degeneracy of T, there is a function $f \in C^1(\mathbb{R})$ with $f(0) = 0$, $\mathrm{Im}(f) \subset \tilde{J}$ and $(Tf)(0) > 1$. Now

$$T(-f)(0) = T((-\mathrm{Id}) \circ f)(0) \le T(-\mathrm{Id})(0)(Tf)(0) < 0$$

and $\mathrm{Im}(-f) \subset \tilde{J}$ since $\mathrm{Im}(f) \subset \tilde{J}$ and $\tilde{J} = -\tilde{J}$. We transport $-f$ back to J by conjugation, using the continuity assumption on T. For $y \in \mathbb{R}$, let $S_y := \mathrm{Id} + y \in C^1(\mathbb{R})$ denote the shift by y. Since for $y_n \to y$, $S_{y_n} \to S_y$ uniformly on compact sets, we conclude that $T(S_{y_n})(x) \to T(S_y)(x)$ for all $x \in \mathbb{R}$. Therefore $T(S_y)(x)$ depends continuously on y for any fixed x. Since

$$1 \le T(\mathrm{Id})(x_1) \le T(S_{x_1})(0) T(S_{-x_1})(x_1),$$

we get $T(S_{x_1})(0) \ne 0$. Using that $T(S_0)(0) = T(\mathrm{Id})(0) \ge 1 > 0$, the continuity of $T(S_y)(0)$ in y implies that $T(S_{x_1})(0) > 0$ and $T(S_{-x_1})(x_1) > 0$. Let

$g := S_{x_1} \circ (-f) \circ S_{-x_1}$. Then $g \in C^1(\mathbb{R})$, $g(x_1) = x_1$, $\mathrm{Im}(g) \subset J \subset I$ and by submultiplicativity,

$$\begin{aligned}(Tg)(x_1) &\le T(S_{x_1})((-f) \circ S_{-x_1}(x_1))T(-f)(S_{-x_1}(x_1))T(S_{-x_1})(x_1)\\ &= T(S_{x_1})(0)T(-f)(0)T(S_{-x_1})(x_1) < 0,\end{aligned}$$

since $T(-f)(0) < 0$ and $T(S_{x_1})(0) > 0$, $T(S_{-x_1})(x_1) > 0$. Hence g satisfies the assertions of Lemma 3.2. ∎

Remark. The argument in (b) also shows that $T(-\mathrm{Id})(0) \le -1$.

We next show that T is localized on intervals.

Lemma 3.3. *Under the assumptions of Proposition* 3.1, *for any open interval $I \subset \mathbb{R}$ we have:*

(a) *Let $c \in \mathbb{R}$ and $f \in C^1(\mathbb{R})$ with $f|_I = c$. Then $Tf|_I = 0$.*
(b) *Let $f \in C^1(\mathbb{R})$ with $f|_I = \mathrm{Id}|_I$. Then $Tf|_I = 1$.*
(c) *Take $f_1, f_2 \in C^1(\mathbb{R})$ with $f_1|_I = f_2|_I$ and assume that f_2 is invertible. Then $Tf_1|_I \le Tf_2|_I$. Therefore, if f_1 is also invertible, then $Tf_1|_I = Tf_2|_I$.*

Proof. (a) For the constant function c, $c \circ g = c$ for any $g \in C^1(\mathbb{R})$, hence

$$Tc(x) \le Tc(g(x))Tg(x) \quad \text{for any } x \in I.$$

By non-degeneracy of T and Lemma 3.2, we find $g_1, g_2 \in C^1(\mathbb{R})$ with $g_j(x) = x$, $\mathrm{Im}(g_j) \subset I$ $(j = 1, 2)$ and $(Tg_1)(x) > 1$, $(Tg_2)(x) < 0$. Then $Tc(x) \le Tc(x)Tg_i(x)$. Assuming $Tc(x) > 0$, we get $Tc(x) < 0$ by applying g_2. Assuming $Tc(x) < 0$, and applying g_1, we find

$$|Tc(x)| \ge |Tc(x)|\,|Tg_1(x)| > |Tc(x)|.$$

Therefore $Tc(x) = 0$ for all $x \in I$.

Now assume that $f \in C^1(\mathbb{R})$ satisfies $f|_I = c$. Choose g_1, g_2 as before. Then $f \circ g_j = c$ $(j = 1, 2)$ and hence, by what we just showed, for $x \in I$, $0 = Tc(x) \le Tf(x)Tg_j(x)$, yielding $Tf(x) = 0$ since $Tg_1(x) > 0$ and $Tg_2(x) < 0$. Thus $Tf|_I = 0$.

(b) Assume that $f \in C^1(\mathbb{R})$ satisfies $f|_I = \mathrm{Id}|_I$. Let $x \in I$ and choose again $g_1, g_2 \in C^1(\mathbb{R})$ with $g_j(x) = x$, $\mathrm{Im}(g_j) \subset I$ $(j = 1, 2)$ with $Tg_1(x) > 1$ and $Tg_2(x) < 0$. Then $f \circ g_j = g_j$ for $j = 1, 2$ and

$$Tg_j(x) = T(f \circ g_j)(x) \le Tf(x)Tg_j(x).$$

This inequality for g_1 yields $Tf(x) \ge 1$, and for g_2 we get $|Tg_2(x)| \ge Tf(x)|Tg_2(x)|$, i.e. $Tf(x) \le 1$. Hence $Tf(x) = 1$, $Tf|_I = 1$.

(c) Assume that $f_1|_I = f_2|_I$ and that f_2 is invertible. Let $g := f_2^{-1} \circ f_1$. Then $g \in C^1(\mathbb{R})$ with $f_1 = f_2 \circ g$, and since $f_1|_I = f_2|_I$, we have $g|_I = \mathrm{Id}|_I$.

By (b), $Tg|_I = 1$. Hence for $x \in I$, $g(x) = x$ and

$$Tf_1(x) = T(f_2 \circ g)(x) \le Tf_2(x)Tg(x) = Tf_2(x).$$

Therefore $Tf_1|_I \le Tf_2|_I$. If f_1 is invertible as well, clearly $Tf_1|_I = Tf_2|_I$. ∎

Proof of Proposition 3.1. (i) Let $\mathcal{C} := \{f \in C^1(\mathbb{R}) \mid f$ is invertible and $f'(x) \neq 0$ for all $x \in \mathbb{R}\}$. Fix any $x_0 \in \mathbb{R}$. For $f \in \mathcal{C}$, consider the tangent line at x_0,

$$g(x) := f(x_0) + f'(x_0)(x - x_0), \qquad x \in \mathbb{R},$$

and set

$$h(x) := \begin{cases} f(x), & x \le x_0, \\ g(x), & x > x_0. \end{cases}$$

By definition, g and h are $C^1(\mathbb{R})$-functions belonging to $\mathcal{C}$. Let $I_- := (-\infty, x_0)$ and $I_+ := (x_0, \infty)$. Then $f|_{I_-} = h|_{I_-}$ and $h|_{I_+} = g|_{I_+}$. By Lemma 3.3, $Tf|_{I_-} = Th|_{I_-}$ and $Th|_{I_+} = Tg|_{I_+}$. Since Tf, Th and Tg are continuous functions and $\{x_0\} = \overline{I_-} \cap \overline{I_+}$, we get $Tf(x_0) = Th(x_0) = Tg(x_0)$. However, g and therefore also $Tg(x_0)$ only depend on x_0, $f(x_0)$ and $f'(x_0)$. Therefore there exists a function $F : \mathbb{R}^2 \times (\mathbb{R} \setminus \{0\}) \to \mathbb{R}$ such that for all $f \in \mathcal{C}$ and all $x_0 \in \mathbb{R}$,

$$Tf(x_0) = F(x_0, f(x_0), f'(x_0)).$$

(ii) For $f \in C^1(\mathbb{R})$, let $I := (y_0, y_1)$ be an interval where f is strictly increasing with $f'(x) > 0$ for any $x \in I$ and $f'(y_0) = 0$, $f'(y_1) = 0$ (or $y_0 = -\infty$, $f'(y_1) = 0$ or $f'(y_0) = 0$, $y_1 = \infty$, with obvious modifications in the following). For sufficiently small $\epsilon_0 > 0$, $f'(y_0 + \epsilon) > 0$ and $f'(y_1 - \epsilon) > 0$ for any $0 < \epsilon \le \epsilon_0$. Define $\tilde{f} \in C^1(\mathbb{R})$ by

$$\tilde{f}(x) = \begin{cases} f(y_0), & x \le y_0, \\ f(x), & x \in I, \\ f(y_1), & x \ge y_1. \end{cases}$$

Then $\tilde{f}'(y_0) = \tilde{f}'(y_1) = 0$ and $\tilde{f}$ is the limit of some functions $\tilde{f}_\epsilon \in \mathcal{C}$ in the sense that $\tilde{f}_\epsilon \to \tilde{f}$ and $\tilde{f}'_\epsilon \to \tilde{f}'$ uniformly on compact subsets of $\mathbb{R}$ as $\epsilon \to 0$. One may choose e.g.

$$\tilde{f}_\epsilon(x) = \begin{cases} f(y_0 + \epsilon) + f'(y_0 + \epsilon)(x - (y_0 + \epsilon)), & x \le y_0 + \epsilon, \\ f(x), & x \in (y_0 + \epsilon, y_1 - \epsilon), \\ f(y_1 - \epsilon) + f'(y_1 - \epsilon)(x - (y_1 - \epsilon)), & x \ge y_1 - \epsilon. \end{cases}$$

Note that $\tilde{f}_\epsilon \in \mathcal{C}$ for any $0 < \epsilon \le \epsilon_0$ since $\tilde{f}_\epsilon$ is invertible with $\tilde{f}'_\epsilon(x) > 0$ for all $x \in \mathbb{R}$. By part (i), for any $x \in (y_0 + \epsilon, y_1 - \epsilon)$, $T\tilde{f}_\epsilon(x) = F(x, f(x), f'(x))$, since for these x, $\tilde{f}_\epsilon(x) = f(x)$ and $\tilde{f}'_\epsilon(x) = f'(x)$. By the continuity assumption on T, we have

$$T\tilde{f}(x) = \lim_{\epsilon\to 0} T\tilde{f}_\epsilon(x) = F(x, f(x), f'(x))$$

for all $x \in (y_0, y_1)$. Since $\tilde{f}_\epsilon \in \mathcal{C}$ is invertible, by Lemma 3.3(c) we have for any $x \in I_\epsilon := (y_0 + \epsilon, y_1 - \epsilon)$ with $f|_{I_\epsilon} = \tilde{f}_\epsilon|_{I_\epsilon}$,

$$Tf(x) \le T\tilde{f}_\epsilon(x) = T\tilde{f}(x) = F(x, f(x), f'(x)). \tag{3.2}$$

(iii) Hence it suffices to show that $T\tilde{f}(x) \le Tf(x)$ for all $x \in (y_0, y_1)$ since then $Tf(x) = F(x, f(x), f'(x))$ by (3.2). We may write $\tilde{f} = f \circ g$ where

$$g(x) = \begin{cases} y_0, & x \le y_0, \\ x, & x \in (y_0, y_1), \\ y_1, & x \ge y_1. \end{cases}$$

If g were in $C^1(\mathbb{R})$, then $g|_{(y_0,y_1)} = \mathrm{Id}$ and $Tg|_{(y_0,y_1)} = 1$ so that

$$T\tilde{f}(x) = T(f \circ g)(x) \le Tf(g(x))Tg(x) = Tf(x),$$

which would prove the claim. However, $g \notin C^1(\mathbb{R})$. Therefore, we approximate g by functions $g_\epsilon \in C^1(\mathbb{R})$. Let

$$g_\epsilon(x) = \begin{cases} y_0 + \epsilon/2, & x < y_0, \\ y_0 + (\epsilon^2 + (x - y_0)^2)/(2\epsilon), & y_0 \le x \le y_0 + \epsilon, \\ x, & y_0 + \epsilon \le x \le y_1 - \epsilon, \\ y_1 - (\epsilon^2 + (y_1 - x)^2)/(2\epsilon), & y_1 - \epsilon \le x \le y_1, \\ y_1 - \epsilon/2, & x \ge y_1. \end{cases}$$

Then $g_\epsilon(y_1) = y_1 - \epsilon/2$, $g'_\epsilon(y_1) = 0$, $g_\epsilon(y_1 - \epsilon) = y_1 - \epsilon$, $g'_\epsilon(y_1 - \epsilon) = 1$ and similar equations hold for y_0 and $y_0 + \epsilon$ so that $g_\epsilon \in C^1(\mathbb{R})$. Note that $f \circ g_\epsilon \to \tilde{f}$, $(f \circ g_\epsilon)' \to \tilde{f}'$ uniformly on compact sets of $\mathbb{R}$, with $f \circ g_\epsilon$, $\tilde{f} \in C^1(\mathbb{R})$. One has $g'_\epsilon = 1$ in $(y_0 + \epsilon, y_1 - \epsilon)$ and $0 \le g'_\epsilon \le 1$ in $(y_1 - \epsilon, y_1)$ and $g'_\epsilon = 0$ in (y_1, ∞). As $g_\epsilon|_{I_\epsilon} = \mathrm{Id}|_{I_\epsilon}$ where $I_\epsilon = (y_0 + \epsilon, y_1 - \epsilon)$, we have $Tg_\epsilon|_{I_\epsilon} = 1$ by Lemma 3.3(b). By the operator inequality (3.1), for $x \in I_\epsilon$,

$$T(f \circ g_\epsilon)(x) \le Tf(g_\epsilon(x))Tg_\epsilon(x) = Tf(x).$$

By the continuity assumption on T, for $x \in (y_0, y_1)$,

$$T\tilde{f}(x) = \lim_{\epsilon\to 0} T(f \circ g_\epsilon)(x) \le Tf(x).$$

Together with (3.2) we get

$$Tf(x) = F(x, f(x), f'(x)), \qquad x \in (y_0, y_1). \tag{3.3}$$

(iv) We see that formula (3.3) holds on all intervals of monotonicity of $f \in C^1(\mathbb{R})$ with non-vanishing derivative. On intervals J where f is constant, $Tf|_J = 0$ by Lemma 3.3(a), and $F(x, y, 0) = 0$ is a result of continuity arguments like $\lim_{\epsilon\to 0} T\tilde{f}_\epsilon(x) = T\tilde{f}(x)$ for boundary points of J together with $Tf|_J = 0$. Equation (3.3) thus holds on J, meaning $0 = 0$ there. Similarly, if $x \in \mathbb{R}$ is a limit point of end points x_n of intervals of monotonicity of f where $f'(x_n) = 0$ and $Tf(x_n) = 0$, we have $f'(x) = 0$ and $Tf(x) = 0$ by continuity of f' and Tf. With $F(x, y, 0) = 0$, equation (3.3) also holds for those points x, stating that $0 = 0$ there. Hence (3.3) holds for all $f \in C^1(\mathbb{R})$ and all $x \in \mathbb{R}$. ∎

4. Proof of Theorem 1.2. We now prove Theorem 1.2, giving the form of the solutions of the chain rule inequality under the assumptions stated there. By Proposition 3.1, there is $F : \mathbb{R}^3 \to \mathbb{R}$ such that for all $f \in C^1(\mathbb{R})$ and $x \in \mathbb{R}$,

$$Tf(x) = F(x, f(x), f'(x)).$$

(a) In the case of the chain rule operator equation (1.1), this yields a functional equation for F. Functional equations have been studied intensively (cf. Aczél [A]). In the case of the chain rule operator inequality (1.3), the localization of Proposition 3.1 yields the following equivalent functional *inequality* for F,

$$F(x, z, \alpha\beta) \le F(y, z, \alpha)F(x, y, \beta) \tag{4.1}$$

for all $x, y, z, \alpha, \beta \in \mathbb{R}$. Just choose $f, g \in C^1(\mathbb{R})$ with $g(x) = y$, $f(y) = z$ and $g'(x) = \beta$, $f'(y) = \alpha$. The equations $Tc = 0$, $T(\mathrm{Id}) = 1$ mean that

$$F(x, y, 0) = 0, \qquad F(x, x, 1) = 1. \tag{4.2}$$

By (4.1) and (4.2),

$$1 = F(x, x, 1) \le F(y, x, 1)F(x, y, 1),$$

hence $F(x, y, 1) \ne 0$ for all $x, y \in \mathbb{R}$. In fact, $F(x, y, 1) > 0$ since $F(x, y, 1) = T(S_{y-x})(x)$ depends continuously on y and $F(x, x, 1) = 1 > 0$. Here S_{y-x} is the shift by $y - x$ (cf. the arguments in the proof of Lemma 3.2).

By the remark after the proof of Lemma 3.2, $T(-\mathrm{Id})(0) < 0$, which means $F(0, 0, -1) < 0$. Actually, for any $x \in \mathbb{R}$,

$$F(x, x, -1) \le F(0, x, 1)F(0, 0, -1)F(x, 0, 1) < 0.$$

Fix $x_0 \in \mathbb{R}$ and let $K(\alpha) := F(x_0, x_0, \alpha)$. By (4.1) for $x = y = z = x_0$, K is submultiplicative. Further $K(-1) < 0 < K(1)$ and $K(0) = 0$. Assume that $\alpha_n, \alpha \in \mathbb{R}$ satisfy $\lim_{n\to\infty} \alpha_n = \alpha$. Consider the functions f_n, f defined by $f_n(x) := \alpha_n(x - x_0) + x_0$ and $f(x) := \alpha(x - x_0) + x_0$. Then $f_n(x_0) = f(x_0) = x_0$ and $f_n'(x) = \alpha_n \to \alpha = f'(x)$ uniformly on $\mathbb{R}$ and $f_n \to f$ uniformly on compacta. Hence by the continuity assumption on T, we have $Tf_n(x_0) \to Tf(x_0)$, which means that

$$K(\alpha_n) = F(x_0, x_0, \alpha_n) = Tf_n(x_0) \to Tf(x_0) = F(x_0, x_0, \alpha) = K(\alpha).$$

Therefore K is continuous on $\mathbb{R}$. Theorem 1.1 implies that there is $p(x_0) > 0$ such that

$$K(\alpha) = \begin{cases} \alpha^{p(x_0)}, & \alpha > 0, \\ F(x_0, x_0, -1)|\alpha|^{p(x_0)}, & \alpha < 0. \end{cases} \tag{4.3}$$

For any $x, y, z \in \mathbb{R}$,

$$F(x, x, \alpha) \le F(z, x, 1)F(z, z, \alpha)F(x, z, 1) = d(x, z)F(z, z, \alpha)$$

where $d(x, z) := F(z, x, 1)F(x, z, 1) \ge 1$ is a number independent of α. Fixing $x \ne z$, for any $\alpha > 0$ we have $\alpha^{p(x)-p(z)} \le d(x, z)$. If $p(x) \ne p(z)$, we would get a contradiction either for $\alpha \to 0$ or for $\alpha \to \infty$. Hence the exponent $p := p(x)$ is independent of $x \in \mathbb{R}$.

(b) We now study $F(x, z, \alpha)$ for $x \ne z$. By (4.1) and (4.3) for $\alpha > 0$ and $\beta \in \mathbb{R}$,

$$F(x, z, \alpha\beta) \le F(x, z, \beta)F(x, x, \alpha) = \alpha^p F(x, z, \beta)$$

and

$$F(x, z, \beta) \le F(x, z, \alpha\beta)F\left(x, x, \frac{1}{\alpha}\right) = \frac{1}{\alpha^p}F(x, z, \alpha\beta).$$

Therefore

$$F(x, z, \alpha\beta) \le \alpha^p F(x, z, \beta) \le F(x, z, \alpha\beta),$$

i.e. $F(x, z, \alpha\beta) = \alpha^p F(x, z, \beta)$. Setting $\beta = 1$ and $\beta = -1$, we find that

$$F(x, z, \alpha) = \begin{cases} F(x, z, 1)\alpha^p, & \alpha > 0, \\ F(x, z, -1)|\alpha|^p, & \alpha < 0. \end{cases} \tag{4.4}$$

We know that $F(x, z, 1) > 0$. On the other hand,

$$F(x, z, -1) \le F(0, z, 1)F(0, 0, -1)F(x, 0, 1) < 0.$$

Let $c_\pm(x, z) := F(x, z, \pm 1)$ and set $a(x, z) := |c_-(x, z)|/c_+(x, z)$. Since

$$c_-(x, z) = F(x, z, -1) \le F(x, z, 1)F(x, x, -1) \le -F(x, z, 1) = -c_+(x, z),$$

using $F(x, x, -1) \le -1$, we have $a(x, z) \ge 1$ for all $x, z, \in \mathbb{R}$. Choose $\alpha, \beta \in \{1, -1\}$ in the functional inequality (4.1) to find

$$\begin{aligned} c_+(x, z) &\le c_+(y, z)c_+(x, y) \quad (\alpha = 1,\ \beta = 1), \\ c_-(x, z) &\le c_-(y, z)c_+(x, y) \quad (\alpha = -1,\ \beta = 1), \\ c_-(x, z) &\le c_+(y, z)c_-(x, y) \quad (\alpha = 1,\ \beta = -1). \end{aligned}$$

Using these inequalities, the definition of a, and observing that $\operatorname{sgn}(c_-) = -1$, we get

$$\begin{aligned}(4.5)\qquad c_+(x,z)\max(a(y,z),a(x,y))&\\ &\le c_+(y,z)c_+(x,y)\max(a(y,z),a(x,y))\\ &= \max(|c_-(y,z)|c_+(x,y),c_+(y,z)|c_-(x,y)|)\\ &\le |c_-(x,z)| = c_+(x,z)a(x,z).\end{aligned}$$

Since $c_+(x,z) > 0$, this implies that $\max(a(y,z),a(x,y)) \le a(x,z)$ for all $x,y,z \in \mathbb{R}$. This yields $a(x,y) \le a(x,0) \le a(0,0)$ and $a(0,0) \le a(x,0) \le a(x,y)$. Therefore a is constant, $a(x,y) = a(0,0)$ for all $x,y \in \mathbb{R}$. Let $A := a(0,0) = a(x,z)$. Then $A \ge 1$ and $c_-(x,z) = -Ac_+(x,z)$. Since we now have equalities everywhere in (4.5), we conclude that

$$c_+(x,z) = c_+(y,z)c_+(x,y), \qquad x,y \in \mathbb{R}.$$

For $y = 0$, $c_+(x,z) = c_+(0,z)c_+(x,0)$, $1 = c_+(x,x) = c_+(0,x)c_+(x,0)$. Let $H(x) := c_+(0,x)$. Then $H > 0$ and $c_+(x,z) = H(z)/H(x)$. Hence by (4.4) with $F(x,z,\pm 1) = c_\pm(x,z)$,

$$F(x,z,\alpha) = \begin{cases} (H(z)/H(x))\alpha^p, & \alpha > 0,\\ -A(H(z)/H(x))|\alpha|^p, & \alpha < 0.\end{cases}$$

Note that $H(z) = F(0,z,1) = T(S_z)(0)$ depends continuously on z as we showed in the proof of Lemma 3.2. Together with (3.3) this ends the proof of Theorem 1.2. ∎

Acknowledgments. Research of H. König was supported by Minerva's Minkowski Center of Tel Aviv University. Research of V. Milman was supported by Minerva's Minkowski Center of Tel Aviv University, by the Alexander von Humboldt Foundation, by ISF grant 826/13 and by BSF grant 2012111.

References

[A] J. Aczél, *Lectures on Functional Equations and Their Applications*, Academic Press, 1966.

[AKM] S. Artstein-Avidan, H. König and V. Milman, *The chain rule as a functional equation*, J. Funct. Anal. 259 (2010), 2999–3024.

[B] S. Banach, *Sur l'équation fonctionnelle* $f(x+y) = f(x) + f(y)$, Fund. Math. 1 (1920), 123-124.

[G] Th. Gronwall, *Note on the derivatives with respect to a parameter of the solutions of a system of differential equations*, Ann. of Math. 20 (1919), 292–296.

[GMP] J. Gustavsson, L. Maligranda and J. Peetre, *A submultiplicative function*, Indag. Math. 51 (1989), 435–442.

[HP] E. Hille and R. Phillips, *Functional Analysis and Semi-Groups*, Amer. Math. Soc., Providence, 1968.

[KM] H. König and V. Milman, *Rigidity and stability of the Leibniz and the chain rule*, Proc. Steklov Inst. Math. 280 (2013), 191–207.
[S] W. Sierpiński, *Sur l'équation fonctionnelle* $f(x+y) = f(x)+f(y)$, Fund. Math. 1 (1920), 116–122.

Hermann König
Mathematisches Seminar
Universität Kiel
24098 Kiel, Germany
E-mail: hkoenig@math.uni-kiel.de

Vitali Milman
School of Mathematical Sciences
Tel Aviv University
Ramat Aviv, Tel Aviv 69978, Israel
E-mail: milman@post.tau.ac.il

Received March 25, 2014
Revised version September 11, 2014 (7917)