DE GRUYTER Dependence Modeling 2023; 11: 20230104

Research Article

Paul Ressel*

Functions operating on several multivariate distribution functions

https://doi.org/10.1515/demo-2023-0104
received June 28, 2023; accepted September 22, 2023

Abstract: Functions f on $[0,1]^m$ such that every composition $f\circ(g_1, \ldots, g_m)$ with d-dimensional distribution functions $g_1, \ldots, g_m$ is again a distribution function, turn out to be characterized by a very natural monotonicity condition, which for $d = 2$ means ultramodularity. For $m = 1$ (and $d = 2$), this is equivalent with increasing convexity.

Keywords: multivariate distribution function, ultramodular, Bernstein polynomials, Faà di Bruno's formula, higher order, monotonicity

MSC 2020: 26E05, 26B40, 60E05

1 Introduction

Which functions of distribution functions ("d.f.s") are again d.f.s? – this is a very general question, with an obvious answer only in dimension one. If g_1 and g_2 are both univariate d.f.s, and $f: [0,1]^2 \to \mathbb{R}_+$ is just increasing (i.e., $f(x) \le f(y)$ for $x \le y$), then $[f\circ(g_1, g_2)](t) \coloneqq f(g_1(t), g_2(t))$ is again a d.f. (disregarding right continuity), but $[f\circ(g_1 \times g_2)](x) \coloneqq f(g_1(x_1), g_2(x_2))$ need not be a bivariate d.f., as the example $f \coloneqq 1_{\{(s,t)\in[0,1]^2 \mid s\le t\}}$ and $g_1(t) = g_2(t) \coloneqq (t \vee 0) \wedge 1$ shows. In this situation, f itself has to be a two-dimensional d.f. in order to guarantee that $f\circ(g_1 \times g_2)$ is also of this type.

We see already that we are confronted with two related but different questions. The first one: given m multivariate d.f.s $g_1, \ldots, g_m$ of (possibly different) dimensions $n_1, \ldots, n_m$, and $f: [0,1]^m \to \mathbb{R}_+$, under which general conditions at f is then $f\circ(g_1 \times \cdots \times g_m)$ again a d.f.? And the second question: if $n_1 = \cdots = n_m \eqqcolon d$, what are necessary and sufficient conditions for f, such that $f\circ(g_1, \ldots, g_m)$ is another d-variate d.f.?

The first question was solved some years ago: f has to be "$\mathbf{n}$-↑" with $\mathbf{n} \coloneqq (n_1, \ldots, n_m)$, a notion to be explained shortly (see [12, Theorem 12]). The second question was posed already in [5] and will be answered in this article. It turns out that the function f then has to fulfill a very natural condition, known for $d = 2$ as being an "ultramodular aggregation function."

An important role in the proof of the main result will be played by a multivariate generalization of the famous Faà di Bruno formula. In order to apply it, we have to use C^∞ approximations, in particular multivariate Bernstein polynomials.

Notations:

$\mathbb{R}_+ = [0, \infty[$, $\mathbb{N} = \{1, 2, 3, \ldots\}$, $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$,

$|\mathbf{n}| = \sum_{i=1}^d n_i$ for $\mathbf{n} \in \mathbb{N}_0^d$, $\mathbf{r}_d \coloneqq (r, r, \ldots, r) \in \mathbb{N}_0^d$ for $r \in \mathbb{N}_0$ (mostly for $r \in \{0, 1\}$), $[d] \coloneqq \{1, 2, \ldots, d\}$,

$\mathbf{1}_\alpha(i) \coloneqq \begin{cases} 1, & i \in \alpha \\ 0, & i \in \alpha^{\complement} \end{cases}$ for $\alpha \subseteq [d]$, $x_\alpha \coloneqq (x_i)_{i\in\alpha}$,

*** Corresponding author: Paul Ressel,** MGF, Kath. Universität Eichstätt-Ingolstadt, Ostenstr. 26, Eichstätt, Germany, e-mail: paul.ressel@ku.de

$(f \times g)(x, y) := (f(x), g(y))$ for mappings f, g,
$(f, g)(x) := (f(x), g(x))$ for mappings with the same domain,
$(f \otimes g)(x, y) := f(x) \cdot g(y)$ for real-valued f, g,
$e_1, \ldots, e_d$ are the usual unit vectors in $\mathbb{R}^d$,
d.f. means distribution function.

2 Some notions of multivariate monotonicity

Let $I_1, \ldots, I_d \subseteq \mathbb{R}$ be non-degenerate intervals, $I := I_1 \times \cdots \times I_d$, and let $f : I \to \mathbb{R}$ be any function. For $s \in I$ and $h \in \mathbb{R}_+^d$ such that also $s + h \in I$, put

$$(E_h f)(s) := f(s + h)$$

and $\Delta_h := E_h - E_0$, i.e., $(\Delta_h f)(s) := f(s + h) - f(s)$. Since $\{E_h \mid h \in \mathbb{R}_+^d\}$ is commutative (where defined), so is also $\{\Delta_h \mid h \in \mathbb{R}_+^d\}$. In particular, with $e_1, \ldots, e_d$ denoting standard unit vectors in $\mathbb{R}^d$, $\Delta_{h_1 e_1}, \ldots, \Delta_{h_d e_d}$ commute. As usual, $\Delta_h^0 f := f$ (also for $h = 0$, but clearly $\Delta_0 f = 0$). For $\mathbf{n} = (n_1, \ldots, n_d) \in \mathbb{N}_0^d$ and $h = (h_1, \ldots, h_d) \in \mathbb{R}_+^d$ we put

$$\Delta_h^{\mathbf{n}} := \Delta_{h_1 e_1}^{n_1} \Delta_{h_2 e_2}^{n_2} \ldots \Delta_{h_d e_d}^{n_d},$$

so that $(\Delta_h^{\mathbf{n}} f)(s)$ is defined for $s, s + \sum_{i=1}^d n_i h_i e_i \in I$.

Definition. $f : I \to \mathbb{R}$ is $\mathbf{n}$–↑ (read "$\mathbf{n}$-*increasing*") iff $(\Delta_h^{\mathbf{p}} f)(s) \geq 0$ $\forall s \in I$, $\forall h \in \mathbb{R}_+^d$, $\forall \mathbf{p} \in \mathbb{N}_0^d$ and $0 \neq \mathbf{p} \leq \mathbf{n}$, such that $s_j + p_j h_j \in I_j$ $\forall j \in [d]$.

A specially important case is $\mathbf{n} = \mathbf{1}_d$; being $\mathbf{1}_d$ -↑ is the "crucial" property of d.f.s. More precisely, $f : I \to \mathbb{R}_+$ is the d.f. of a (non-negative) measure μ, i.e., $f(s) = \mu([-\infty, s] \cap \bar{I})$ $\forall s \in I$, if and only if f is $\mathbf{1}_d$-↑ and right-continuous; c.f. [10, Theorem 7].

Let us for a moment consider the case $d = 1$. Then, $I \subseteq \mathbb{R}$, $\mathbf{n} = n \in \mathbb{N}$, we assume $n \geq 2$, and a famous old result of Boas and Widder [1, Lemma 1] shows that a continuous function $f : I \to \mathbb{R}$ is n-↑ (i.e., $\Delta_h^j f \geq 0 \quad \forall j \in [n], \ \forall h > 0$) iff

$$(\Delta_{h_1} \Delta_{h_2} \ldots \Delta_{h_j} f)(s) \geq 0$$

$\forall j \in [n], \forall h_1, \ldots, h_j > 0$ such that $s, s + h_1 + \cdots + h_j \in I$. For $n = 2$, f is 2-↑ iff it is increasing and convex (and automatically continuous on $I \backslash \{\sup I\}$).

The following definition now seems to be natural:

Definition. Let $I_1, \ldots, I_d \subseteq \mathbb{R}$ be non-degenerate intervals, $I = I_1 \times \cdots \times I_d$, $f : I \to \mathbb{R}$, and $k \in \mathbb{N}$. Then, f is called *k-increasing* ("k-↑") iff $\forall j \in [k]$, $\forall h^{(1)}, \ldots, h^{(j)} \in \mathbb{R}_+^d$, $\forall s \in I$ such that $s + h^{(1)} + \cdots + h^{(j)} \in I$

$$(\Delta_{h^{(1)}} \ldots \Delta_{h^{(j)}} f)(s) \geq 0.$$

(We do not assume f to be continuous.)

We mentioned already that a univariate f is 2-↑ iff it is increasing and convex. But also multivariate 2-↑ functions are well known: they are called *ultramodular*, mostly ultramodular aggregation functions, the latter meaning they are also increasing, and defined as functions $f : [0,1]^d \to [0,1]$ with $f(\mathbf{0}_d) = 0$ and $f(\mathbf{1}_d) = 1$. The restriction to the standard unit cube $[0,1]^d$ is not one of course, but sometimes appropriate as we will see. In our terminology, if f is k-↑, it is by definition also j-↑ for $1 \leq j \leq k$, in particular just increasing.

In this connection, also *increasing supermodular* functions should be mentioned: in the bivariate case, they coincide with $(1,1)$-↑ functions, and in higher dimensions, they are "pairwise $(1,1)$-↑" in the obvious meaning; cf. [6].

Already in dimension two increasing convexity and being 2-↑ are uncomparable properties: on $\mathbb{R}_+^2$ the product is 2-↑, but not convex; and the Euclidean norm is convex, however not 2-↑.

Remark 1. Already in 2005, Bronevich [2] introduced k-↑ functions, calling them "k-monotone." Later on, the name "strongly k-monotone" was used [5,8], a terminology usually associated with strict inequalities, therefore not really adequate.

Some simple properties of k-↑ functions are shown first.

Lemma 1. *Let $f: [0,1]^d \to \mathbb{R}$ be 2-↑. Then,*

(i) *f is continuous iff f is continuous in $\mathbf{1}_d$.*

(ii) *f is right-continuous and on $[0,1[^d$ continuous.*

(iii) *If $f(x^0) = f(\mathbf{1}_d)$ for some $x^0 \neq \mathbf{1}_d$, hence $\alpha := \{i \le d | x_i^0 = 1\} \subsetneq [d]$, then for $\alpha \neq \emptyset$ f depends only on $x_\alpha := (x_i)_{i\in\alpha}$. In case, $\alpha = \emptyset$ f is constant.*

(iv) *For each $y \in [0,1]^d \backslash [0,1[^d$, the limit*

$$f_0(y) := \lim_{\substack{x\to y \\ x<y}} f(x)$$

exists, $f_0(y) \le f(y)$, f_0 is a 2-↑ and continuous extension of $f|[0,1[^d$. If f is k-↑, so is f_0.

Proof. (i) For $x, h \in [0,1]^d$ such that also $x \pm h \in [0,1]^d$, we have (with $\mathbf{1} := \mathbf{1}_d$)

$$0 \le f(x) - f(x-h) \le f(x+h) - f(x) \le f(\mathbf{1}) - f(\mathbf{1}-h), \qquad (*)$$

from which the claim follows, f being increasing.

(ii) For any $x \in [0,1[^d$, the univariate function $[0,1] \ni t \mapsto f(tx)$ is convex, hence continuous, also in $t = 1$ (being defined and convex in a neighborhood of 1). Since f is increasing, f is continuous in x. For $x \in [0,1]^d \backslash [0,1[^d$, $x \neq \mathbf{1}_d$, let $\alpha := \{i \le d \mid x_i = 1\}$. Then, for $y \ge x$, also $y_i = 1\ \forall i \in \alpha$, and $[0,1[^{\alpha^c} \ni z \mapsto f(\mathbf{1}_\alpha, z)$ is continuous, in particular in the point x_{α^c}, implying f to be right-continuous in $x = (\mathbf{1}_\alpha, x_{\alpha^c})$.

(iii) If $\alpha = \emptyset$, then $x^0 = 1 - h^0$ for some $h^0 \in]0,1]^d$, $f(\mathbf{1}) = f(\mathbf{1} - h)\ \forall 0 \le h \le h^0$, and f is constant by (*). For $\alpha \neq \emptyset$, we define $g : [0,1]^{\alpha^c} \to \mathbb{R}_+$ by $g(z) := f(\mathbf{1}_\alpha, z)$. Also, g is 2-↑, and

$$g((x^0)_{\alpha^c}) = f(x^0) = f(\mathbf{1}_d) = g(\mathbf{1}_{\alpha^c}),$$

hence g is constant. But then, $\forall y \in [0,1]^\alpha$

$$0 \le f(y, \mathbf{1}_{\alpha^c}) - f(y, \mathbf{0}_{\alpha^c}) \le f\big(\mathbf{1}_\alpha, \mathbf{1}_{\alpha^c}\big) - f\big(\mathbf{1}_\alpha, \mathbf{0}_{\alpha^c}\big) = g\big(\mathbf{1}_{\alpha^c}\big) - g\big(\mathbf{0}_{\alpha^c}\big) = 0,$$

showing $f(y, z)$ to be independent of z.

(iv) The existence of $f_0(y)$ is clear, f being increasing and bounded. The defining inequalities for f being k-↑ prevail for f_0, for any $k \ge 2$. Since f_0 is continuous in $\mathbf{1}$, it is everywhere continuous. □

Our first theorem will state some equivalent conditions for f to be k-↑. An essential ingredient will be positive linear (or affine) mappings: a linear function $\psi : \mathbb{R}^m \to \mathbb{R}^d$ is called *positive* iff $\psi(\mathbb{R}_+^m) \subseteq \mathbb{R}_+^d$; and an affine $\varphi : \mathbb{R}^m \to \mathbb{R}^d$ is positive iff its "linear part" $\varphi - \varphi(0)$ is.

Theorem 1. *Let $I \subseteq \mathbb{R}^d$ be a non-degenerate interval, $f : I \to \mathbb{R}$, $k, d \in \mathbb{N}$. Then, there are equivalent:*

(i) *f is k-↑,*

(ii) *f is $\mathbf{n}$-↑ $\quad \forall \mathbf{n} \in \mathbb{N}_0^d$ with $0 < |\mathbf{n}| \le k$,*

(iii) *$\forall m \in \mathbb{N}$, $\forall$ non-degenerate interval $J \subseteq \mathbb{R}^m$, $\forall$ positive affine $\varphi : \mathbb{R}^m \to \mathbb{R}^d$ such that $\varphi(J) \subseteq I$, also $f \circ \varphi$ is k-↑,*

(iv) *$\forall m, J, \varphi$ as before, and $\forall \mathbf{n} \in \mathbb{N}_0^m$ with $0 < |\mathbf{n}| \le k$ the function $f \circ \varphi$ is $\mathbf{n}$-↑,*

(v) $\forall m, J, \varphi$ *as before, and* $\forall \mathbf{n} \in \{0,1\}^m$ *with* $0 < |\mathbf{n}| \le k$ *the function* $f \circ \varphi$ *is* $\mathbf{n}$-↑.

Remark 2. For $k \ge d \ge 2$ and $f \ge 0$, the aforementioned function f is not only right-continuous (Lemma 1(ii)), but also $\mathbf{1}_d$-↑, hence a d.f. on I; however, with the extra property that $f \circ \varphi$ is a d.f., too, for any positive affine $\varphi : \mathbb{R}^m \to \mathbb{R}^d$. In other words, if $f(x) = P(X \le x)$ for some d-dimensional random vector X, and $k \ge m$, then also $y \mapsto P(X \le \varphi(y))$ is an m-dimensional d.f.

Proof. We show (i) ⇔ (ii) and (i) ⇒ (iii) ⇒ (iv) ⇒ (v) ⇒ (i).

(i) ⇒ (ii) is clear.

(ii) ⇒ (i): We use induction on k, $k = 1$ being obvious. Let now $k \ge 2$ and suppose the case $k-1$ is already known. Fix some $h \in \mathbb{R}_+^d$ and consider $g := \Delta_h f$, i.e., $g(s) = f(s+h) - f(s)$. With

$$\begin{aligned} g_1(s) &:= f(s + h_1 e_1) - f(s), \\ g_2(s) &:= f(s + h_1 e_1 + h_2 e_2) - f(s + h_1 e_1), \\ &\cdots \end{aligned}$$

we have $g = g_1 + g_2 + \dots + g_d$. Each g_i is $\mathbf{n}$-↑ for any $\mathbf{n} \in \mathbb{N}_0^d$ with $|\mathbf{n}| \le k-1$; hence, $(k-1)$-↑ by assumption, and so is then g. Since $h \in \mathbb{R}_+^d$ was arbitrary, f is k-↑.

(i) ⇒ (iii): Let $j \in [k]$, $h^{(1)}, \dots, h^{(j)} \in \mathbb{R}_+^m$, $x \in J$ such that also $x + \sum_{i=1}^j h^{(i)} \in J$. Then, with $\psi := \varphi - \varphi(0)$,

$$\begin{aligned} [\Delta_{h^{(1)}} \dots \Delta_{h^{(j)}}(f \circ \varphi)](x) &= f \circ \varphi(x + h^{(1)} + \dots + h^{(j)}) \mp \dots + (-1)^j f \circ \varphi(x) \\ &= f(\varphi(0) + \psi(x + h^{(1)} + \dots + h^{(j)})) \mp \dots + (-1)^j f(\varphi(0) + \psi(x)) \\ &= f(\varphi(0) + \psi(x) + \psi(h^{(1)}) + \dots + \psi(h^{(j)})) \mp \dots + (-1)^j f(\varphi(0) + \psi(x)) \\ &= [\Delta_{\psi(h^{(1)})} \dots \Delta_{\psi(h^{(j)})} f](\varphi(0) + \psi(x)) \\ &= [\Delta_{\psi(h^{(1)})} \dots \Delta_{\psi(h^{(j)})} f](\varphi(x)) \ge 0, \end{aligned}$$

(note that $\varphi(x) + \sum_{i=1}^j \psi(h^{(i)}) = \varphi(x + \sum_{i=1}^j h^{(i)}) \in I$).

(iii) ⇒ (iv) ⇒ (v) is clear.

(v) ⇒ (i): Let $j \in [k]$, $x \in I$, $h^{(1)}, \dots, h^{(j)} \in \mathbb{R}_+^d$ such that $x + h^{(1)} + \dots + h^{(j)} \in I$. Denote by $\psi : \mathbb{R}^j \to \mathbb{R}^d$ the linear map whose matrix with respect to the standard bases is $(h^{(1)}, \dots, h^{(j)})$ ($h^{(i)}$ as column vectors), $\varphi := x + \psi$; obviously, ψ (and φ) are positive. For $J := [\mathbf{0}_j, \mathbf{1}_j] \subseteq \mathbb{R}^j$, we have $\varphi(J) \subseteq I$, since $\psi(e_i) = h^{(i)} \quad \forall i \le j$, $\varphi(\mathbf{0}_j) = x$ and $\varphi(\mathbf{1}_j) = x + \sum_{i=1}^j h^{(i)}$. By assumption,

$$0 \le [\Delta_{\mathbf{1}_j}^{\mathbf{1}_j}(f \circ \varphi)](\mathbf{0}_j) = f(x + h^{(1)} + \dots h^{(j)}) \mp \dots + (-1)^j f(x) = (\Delta_{h^{(1)}} \dots \Delta_{h^{(j)}} f)(x). \square$$

Corollary 1. *Let $I \subseteq \mathbb{R}^d$ and $B \subseteq \mathbb{R}$ be non-degenerate intervals. If $g : I \to B$ and $f : B \to \mathbb{R}$ are both k-↑, then so is $f \circ g$.*

Proof. We show Condition (iv) in Theorem 1 to hold for $f \circ g$. We know that $g \circ \varphi$ is $\mathbf{n}$-↑ for $|\mathbf{n}| \le k$. A special case of Theorem 12 in [12] implies $f \circ (g \circ \varphi)$ to be also $\mathbf{n}$-↑, and $f \circ (g \circ \varphi) = (f \circ g) \circ \varphi$. □

Theorem 2. *Let $I \subseteq \mathbb{R}^{d_1}$ and $J \subseteq \mathbb{R}^{d_2}$ be non-degenerate intervals, $f : I \to \mathbb{R}$ and $g : J \to \mathbb{R}$, both non-negative and k-↑. Then, also $f \otimes g$ is k-↑ on $I \times J$, and in case $I = J$, the product $f \cdot g$ is k-↑, too.*

Proof. We first apply (ii) of Theorem 1. For $0 \ne (\mathbf{m}, \mathbf{n}) \in \mathbb{N}_0^{d_1} \times \mathbb{N}_0^{d_2}$, $(x, y) \in I \times J$, $h^{(1)} \in \mathbb{R}_+^{d_1}$, $h^{(2)} \in \mathbb{R}_+^{d_2}$ we have

$$[\Delta_{(h^{(1)}, h^{(2)})}^{\mathbf{m},\mathbf{n}}(f \otimes g)](x, y) = (\Delta_{h^{(1)}}^{\mathbf{m}} f)(x) \cdot (\Delta_{h^{(2)}}^{\mathbf{n}} g)(y),$$

and for $|(\mathbf{m}, \mathbf{n})| = |\mathbf{m}| + |\mathbf{n}| \le k$, both factors on the right-hand side are non-negative. Since $\mathbf{m} = 0$ or $\mathbf{n} = 0$ is allowed, only $(\mathbf{m}, \mathbf{n}) \ne 0$ being required, we need in fact $f \ge 0$ and $g \ge 0$.

For $I = J$ (with $d_1 = d_2 =: d$), let $\varphi : \mathbb{R}^d \to \mathbb{R}^{2d}$ be given by $\varphi(x) := (x, x)$, a positive linear map. Then, $\varphi(I) \subseteq I \times I$, and by Theorem 1 (iii), $(f \otimes g) \circ \varphi = f \cdot g$ is also k-↑. □

We see that any monomial $f(x) = \prod_{i=1}^d x_i^{n_i}$ ($n_i \in \mathbb{N}$) is k-↑ on $\mathbb{R}_+^d$ for each $k \in \mathbb{N}$. If $c_i \in]0, \infty[$ then $\prod_{i=1}^d x_i^{c_i}$ is k-↑ on $\mathbb{R}_+^d$ at least for $c_i \geq k-1$, $i = 1, \ldots, d$.

Examples 1. (a) For $a > 0$, the function $f(x, y) \coloneqq (xy - a)_+$ is 2-↑ on $\mathbb{R}_+^2$, since $t \mapsto (t - a)_+$ is 2-↑ on $\mathbb{R}_+$, by Corollary 1. Its restriction to $[0, \sqrt{1+a}\,]^2$ is therefore the d.f. of some random vector. In [12] on page 261, it was shown that f is not (2, 1)-↑ (resp. (1, 2)-↑), but it is of course (1, 1)-↑. The tensor product $g(x, y) \coloneqq (x - a)_+ \cdot (y - b)_+$ is (2, 2)-↑ $\forall a, b > 0$; hence, certainly 2-↑, but not 3-↑ since $x \mapsto (x - a)_+$ is not.

Similarly, $(xyz - a)_+^2$ is 3-↑ on $\mathbb{R}_+^3$, for $a > 0$, and of course, $(xy - a)_+^2$ is 3-↑ on $\mathbb{R}_+^2$. We will see later on that $xy + xz + yz - xyz$ is 2-↑ on $[0, 1]^3$, but not 3-↑.

(b) Consider $f_n(t) \coloneqq t^n/(1 + t)$ for $t \geq 0$. It was shown in [7], Lemma 2.4, that f_n is n-↑ (it is not $(n + 1)$-↑). So for any non-negative n-↑ function g on any interval in any dimension, $g^n/(1 + g)$ is n-↑, too.

If g is “only” an n-dimensional d.f., then so is also $g^n/(1 + g)$ – this follows from [11], Theorem 2, but is also a special case of Theorem 6 below.

3 Approximation by Bernstein polynomials

The proof of our main result relies heavily on these special polynomials, since they inherit the monotonicity properties of interest. To define them, we introduce for $r \in \mathbb{N}$, $i \in \{0, 1, \ldots, r\}$

$$b_{i,r}(t) \coloneqq \binom{r}{i} t^i (1 - t)^{r-i}, \quad t \in \mathbb{R},$$

and for $\mathbf{i} = (i_1, \ldots, i_d) \in \{0, 1, \ldots, r\}^d$

$$B_{\mathbf{i},r} \coloneqq b_{i_1,r} \otimes \cdots \otimes b_{i_d,r}.$$

For any $f: [0, 1]^d \to \mathbb{R}$, the associated Bernstein polynomials $f^{(1)}, f^{(2)}, \ldots$ are defined by:

$$f^{(r)} \coloneqq \sum_{\mathbf{0}_d \leq \mathbf{i} \leq \mathbf{r}_d} f\left(\frac{\mathbf{i}}{r}\right) \cdot B_{\mathbf{i},r}.$$

It is perhaps not so well known, that for each continuity point x of f, we have

$$f^{(r)}(x) \to f(x), \quad r \to \infty.$$

This is shown for ex. in [14, page 296], based on the strong law of large numbers for independent Bernoulli trials. Well known is the uniform convergence of $f^{(r)}$ to f on $[0, 1]^d$ for a continuous function f.

In the following, the “upper right boundary” of $[0, 1]^d$ will play a role. Let for $\alpha \subseteq [d]$

$$T_\alpha \coloneqq \{x \in [0, 1]^d \mid x_i < 1 \Leftrightarrow i \in \alpha\}.$$

Then, $[0, 1]^d = \bigcup_{\alpha \subseteq [d]} T_\alpha$ is a disjoint union, $T_\varnothing = \{\mathbf{1}_d\}$ and $T_{[d]} = [0, 1[^d$. The union $\bigcup_{\alpha \subsetneq [d]} T_\alpha$ is called the *upper right boundary* of $[0, 1]^d$.

For $f: [0, 1]^d \to \mathbb{R}$ with Bernstein polynomials $f^{(1)}, f^{(2)}, \ldots$, let $f_{(\alpha)} \coloneqq f|T_\alpha$, where for $\alpha \neq \varnothing$ T_α may be identified with $[0, 1[^\alpha$. Since for $y \in [0, 1[^\alpha$

$$B_{\mathbf{i},r}(y, \mathbf{1}_{\alpha^c}) = \begin{cases} B_{\mathbf{i}_\alpha,r} & \text{if } \mathbf{i}_{\alpha^c} = \mathbf{r}_{\alpha^c} \\ 0 & \text{else,} \end{cases}$$

we obtain for $\varnothing \neq \alpha \subsetneq [d]$

$$(f_{(\alpha)})^{(r)}(y) = \sum_{\mathbf{i}_\alpha \leq \mathbf{r}_\alpha} f_{(\alpha)}\left(\frac{\mathbf{i}_\alpha}{r}\right) B_{\mathbf{i}_\alpha,r}(y),$$

where $f_{(\alpha)}\left(\frac{\mathbf{i}_\alpha}{r}\right) = f\left(\frac{\mathbf{i}_\alpha}{r}, \mathbf{1}_{\alpha^c}\right) = f\left(\frac{(\mathbf{i}_\alpha, \mathbf{r}_{\alpha^c})}{r}\right)$, and then

$$\begin{aligned} f^{(r)}(y, \mathbf{1}_{a^c}) &= \sum_{\mathbf{i}\le\mathbf{r}_d} f\left(\frac{\mathbf{i}}{r}\right) B_{\mathbf{i},r}(y, \mathbf{1}_{a^c}) \\ &= \sum_{\mathbf{i}_a\le\mathbf{r}_a} f\left(\frac{(\mathbf{i}_a, \mathbf{r}_{a^c})}{r}\right) B_{\mathbf{i}_a,r}(y) \\ &= (f_{(a)})^{(r)}(y). \end{aligned}$$

In other words, the restriction of f to one of the parts T_a of the upper right boundary has as its Bernstein polynomials the restrictions of the original ones to T_a. This leads to the following.

Theorem 3. *Let $f: [0,1]^d \to \mathbb{R}$ have the property that each restriction $f|T_a$ for $\varnothing \neq a \subseteq [d]$ is continuous. Then,*

$$\lim_{r\to\infty} f^{(r)}(x) = f(x) \quad \forall x \in [0,1]^d,$$

i.e., the Bernstein polynomials converge pointwise to f everywhere.

Proof. Each $x \in [0,1]^d \backslash \{\mathbf{1}_d\}$ lies in exactly one T_a, i.e., $x = (x_a, \mathbf{1}_{a^c})$ with $x_a < \mathbf{1}_a$, where $\varnothing \neq a \subseteq [d]$, and is thus a continuity point of $f_{(a)} \coloneqq f|T_a$. As already mentioned, this implies

$$(f_{(a)})^{(r)}(x_a) \to f_{(a)}(x_a) = f(x_a, \mathbf{1}_{a^c}) = f(x),$$

and we saw also that

$$(f_{(a)})^{(r)}(x_a) = f^{(r)}(x_a, \mathbf{1}_{a^c}) = f^{(r)}(x).$$

Since $f^{(r)}(\mathbf{1}_d) = f(\mathbf{1}_d)\ \forall r$, the proof is complete. □

For a function f of d variables, we will use a short notation for its partial derivatives (if they exist). Let $\mathbf{p} \in \mathbb{N}_0^d \backslash \{0\}$, then

$$f_{\mathbf{p}} \coloneqq \frac{\partial^{|\mathbf{p}|} f}{\partial x_1^{p_1} \dots \partial x_d^{p_d}},$$

complemented by $f_{\mathbf{0}_d} \coloneqq f$.

Lemma 2. *Let $f: [0,1]^d \to \mathbb{R}$ be arbitrary, $0 \neq \mathbf{p} \in \mathbb{N}_0^d$.*
(i) *If $\Delta_h^{\mathbf{p}} f \geq 0\ \forall h \in \mathbb{R}_+^d$ then $(f^{(r)})_{\mathbf{p}} \geq 0 \forall r \in \mathbb{N}$.*
(ii) *If f is in addition C^∞, then $f_{\mathbf{p}} \geq 0$.*

Proof. (i) Applying the formula for derivatives of one-dimensional Bernstein polynomials [12, p. 273]d times, we obtain

$$(f^{(r)})_{\mathbf{p}} = c_{\mathbf{p}} \cdot \sum_{\mathbf{i}\le\mathbf{r}_d-\mathbf{p}} \left(\Delta^{\mathbf{p}}_{\frac{1}{r}\cdot\mathbf{1}_d} f\right)\left(\frac{\mathbf{i}}{r}\right) b_{i_1,r-p_1} \otimes \cdots \otimes b_{i_d,r-p_d}$$

with $c_{\mathbf{p}} \coloneqq \prod_{i=1}^d r(r-1)\cdot \ldots \cdot (r-p_i+1)$. Hence, $(f^{(r)})_{\mathbf{p}} \geq 0$.

(ii) By [14, Theorem 4]$(f^{(r)})_{\mathbf{p}} \to f_{\mathbf{p}}$, even uniformly, so $f_{\mathbf{p}} \geq 0$, too. □

Theorem 4. *Let $f: [0,1]^d \to \mathbb{R}$ be a C^∞-function, $\mathbf{n} \in \mathbb{N}^d$, $k \in \mathbb{N}$. Then,*
(i) *f is $\mathbf{n}$-↑ $\Leftrightarrow$ $f_{\mathbf{p}} \geq 0\ \forall 0 \neq \mathbf{p} \leq \mathbf{n}, \mathbf{p} \in \mathbb{N}_0^d$.*
(ii) *f is k-↑ $\Leftrightarrow$ $f_{\mathbf{p}} \geq 0\ \forall 0 < |\mathbf{p}| \leq k, \mathbf{p} \in \mathbb{N}_0^d$.*

Proof. (i) "⇒": follows from Lemma 2.

"⇐": Let for $m \in \mathbb{N}$ $\sigma_m : \mathbb{R}^m \to \mathbb{R}$ be the sum function, $\sigma_{\mathbf{n}} \coloneqq \sigma_{n_1} \times \sigma_{n_2} \times \cdots \times \sigma_{n_d}$. By [13, Theorem 5], we have

$$f \text{ is } \mathbf{n}\text{-}\uparrow \iff f \circ \sigma_{\mathbf{n}} \text{ is } \mathbf{1}_{|\mathbf{n}|}\text{-}\uparrow \text{ on } J := \prod_{i=1}^{d} \left[0, \frac{1}{n_i}\right]^{n_i}.$$

The chain rule gives

$$(f \circ \sigma_{\mathbf{n}})_{\mathbf{1}_{|\mathbf{n}|}} = f_{\mathbf{n}} \circ \sigma_{\mathbf{n}} \geq 0,$$

so that for $x, x + h \in J, h \geq 0$ by Fubini's theorem

$$(\Delta_h^{\mathbf{1}_{|\mathbf{n}|}}(f \circ \sigma_{\mathbf{n}}))(x) = \int_{[x,x+h]} (f \circ \sigma_{\mathbf{n}})_{\mathbf{1}_{|\mathbf{n}|}} \mathrm{d}\lambda^{|\mathbf{n}|} \geq 0.$$

The same reasoning can be applied to $0 \lneq \mathbf{q} \leq \mathbf{1}_{|\mathbf{n}|}$, so that indeed $f \circ \sigma_{\mathbf{n}}$ is $\mathbf{1}_{\mathbf{n}}$-↑, i.e., f is $\mathbf{n}$-↑.

(ii) This follows immediately from the first equivalence in Theorem 1. □

Examples 2.

(a) $f(x, y) := x^2y - ax^2y^2 + y^2$ on $[0, 1]^2$, $0 < a \leq \frac{1}{2}$. Since $f_{\mathbf{p}} \geq 0$ for $\mathbf{p} \in \{(1, 0), (0, 1), (1, 1), (2, 0), (0, 2)\}$, f is 2-↑; but $f_{(1,2)}(x, y) = -4ax$ shows that f is neither 3-↑ nor (2, 2)-↑.

(b) $f : \mathbb{R}_+^2 \to \mathbb{R}$ is defined by $f(0, 0) := 0$ and else

$$f(x, y) := \frac{xy(x^2 - y^2)}{x^2 + y^2} + 13 \cdot (x^2 + y^2) + 3xy,$$

(see [9, p. 321]), where it is given as an example of an ultramodular function on $\mathbb{R}_+^2$ (which does not automatically include that it is increasing). However, all partial derivatives $f_{\mathbf{p}}$ with $0 < |\mathbf{p}| \leq 2$ are ≥ 0; hence, f is 2-↑ (and not 3-↑).

(c) With the abbreviation $x^\alpha := \prod_{i \in \alpha} x_i$ for $\alpha \subseteq [d]$, $x^\emptyset := 1$, a polynomial of the form

$$f(x) = \sum_{\alpha \subseteq [d]} c_\alpha x^\alpha$$

is called *multilinear*. f is affine in each variable; therefore, $f_{\mathbf{p}} = 0$ whenever $p_i > 1$ for some i. Hence, f is k-↑ iff $f_{\mathbf{p}} \geq 0 \; \forall \mathbf{p} \leq \mathbf{1}_d$ with $0 < |\mathbf{p}| \leq k$, and $\mathbf{n}$-↑ iff f is $(\mathbf{n} \wedge \mathbf{1}_d)$-↑. The example ($d = 3$)

$$f(x) := x_1x_2 + x_1x_3 + x_2x_3 - x_1x_2x_3$$

is thus 2-↑ on $[0, 1]^3$, but not 3-↑, since $f_{(1,1,1)} = -1$. And f is $(n, n, 0)$-↑ $\forall n$.

Theorem 5. *Let* $f : [0, 1]^d \to \mathbb{R}$, $\mathbf{2}_d \leq \mathbf{n} \in \mathbb{N}_0^d$, $2 \leq k \in \mathbb{N}$. *The Bernstein polynomials of f are denoted* $f^{(1)}, f^{(2)}, \ldots$.
(i) *If f is* $\mathbf{n}$-↑, *then so is each* $f^{(r)}$, *and* $f^{(r)} \to f$ *pointwise.*
(ii) *If f is* k-↑, *then so is each* $f^{(r)}$, *and* $f^{(r)} \to f$ *pointwise.*

Proof. In both cases, f is (at least) 2-↑; therefore (by Lemma 1(ii)), the restriction $f|[0, 1[^d$ is continuous, and so are the other restrictions $f|T_\alpha$ for each non-empty $\alpha \subseteq [d]$. By Theorem 3, $f^{(r)}(x) \to f(x) \; \forall x$.
(i) Lemma 2 implies $(f^{(r)})_{\mathbf{p}} \geq 0 \; \forall r$ and $\forall 0 \neq \mathbf{p} \leq \mathbf{n}$; hence, $f^{(r)}$ is $\mathbf{n}$-↑ $\forall r$.
(ii) Similarly now, $(f^{(r)})_{\mathbf{p}} \geq 0 \; \forall r$ and $\forall 0 < |\mathbf{p}| \leq k$, showing $f^{(r)}$ to be k-↑. □

Of course, a similar result holds if $[0, 1]^d$ is replaced by any non-degenerate compact interval in $\mathbb{R}^d$.

4 Main results

The proof of Theorem 6 makes use of a far-reaching simultaneous generalization of the usual multivariate chain rule and Faà di Bruno's formula. This admirable result was shown by Constantine and Savits [3, Theorem 2.1], and we present it here, keeping (almost) their notation.

Let $d, m \in \mathbb{N}$, let $g_1, \dots, g_m$ be defined and C^∞ in a neighborhood of $x^{(0)} \in \mathbb{R}^d$ (real-valued), put $g := (g_1, \dots, g_m)$, and let f be defined and C^∞ in a neighborhood of $y^{(0)} := g(x^{(0)}) \in \mathbb{R}^m$.

For $\mu, \nu \in \mathbb{N}_0^d$, the relation $\mu \prec \nu$ holds iff one of the following three assertions is true:

(i) $|\mu| < |\nu|$,

(ii) $|\mu| = |\nu|$ and $\mu_1 < \nu_1$,

(iii) $|\mu| = |\nu|, \mu_1 = \nu_1, \dots, \mu_k = \nu_k, \mu_{k+1} < \nu_{k+1}, \exists k \in [d-1]$,

(implying $\mu \neq \nu$).

Examples:

(a) $(1,3,0,4,1) \prec (1,3,1,1,3)$, here $k = 2$,

(b) $e_d \prec e_{d-1} \prec \cdots \prec e_1$,

(c) For $d = 1$ we have $\mu \prec \nu \Leftrightarrow \mu < \nu$.

We need some abbreviations:

$$D_x^\nu := \frac{\partial^{|\nu|}}{\partial x_1^{\nu_1} \dots \partial x_d^{\nu_d}} \quad \text{for } |\nu| > 0, \quad D_x^0 f := f$$

$$x^\nu := \prod_{i=1}^d x_i^{\nu_i}, \quad \nu! := \prod_{i=1}^d \nu_i!, \quad |\nu| := \sum_{i=1}^d \nu_i$$

$$g_\mu^{(i)} := (D_x^\mu g_i)(x^{(0)}), \quad g_\mu := (g_\mu^{(1)}, \dots, g_\mu^{(m)})$$

$$f_\lambda := (D_y^\lambda f)(y^{(0)})$$

$$h := f \circ g, \quad h_\nu := (D_x^\nu h)(x^{(0)}),$$

and, for $\nu \in \mathbb{N}_0^d, \lambda \in \mathbb{N}_0^m, s \in \mathbb{N}, s \le |\nu|$

$$P_s(\nu, \lambda) := \left\{ (k_1, \dots, k_s; l_1, \dots, l_s) \mid |k_j| > 0, 0 \prec l_1 \prec \cdots \prec l_s, \sum_{j=1}^s k_j = \lambda, \sum_{j=1}^s |k_j| l_j = \nu \right\},$$

where (of course) $k_j \in \mathbb{N}_0^m$ and $l_j \in \mathbb{N}_0^d$. (For some values of s, these sets may be empty.)

The announced formula by Constantine and Savits then reads

$$h_\nu = \sum_{1 \le |\lambda| \le |\nu|} f_\lambda \cdot \sum_{s=1}^{|\nu|} \sum_{P_s(\nu,\lambda)} \nu! \cdot \prod_{j=1}^s \frac{(g_{l_j})^{k_j}}{(k_j!) \cdot (l_j!)^{|k_j|}}. \tag{**}$$

This formula reduces for $d = 1$ to the classical one of Faà di Bruno from 1855 (see [3,4]).

One more result is needed, allowing general d.f.s to be "replaced" by C^∞ ones:

Lemma 3.

(i) *Let $(\Omega, \mathcal{A}, \rho)$ be a finite measure space and $\varnothing \neq \mathcal{B} \subseteq \mathcal{A}$ a finite collection of measurable sets. Then, there is another finite measure ρ_0 on $\mathcal{A}$ with finite support such that $\rho_0|\mathcal{B} = \rho|\mathcal{B}$.*

(ii) *Let F on $\mathbb{R}^d$ be the d.f. of some finite measure and $\varnothing \neq B \subseteq \mathbb{R}^d$ a finite subset. Then, there is a C^∞ d.f. $\tilde{F}$ on $\mathbb{R}^d$ such that $\tilde{F}|B = F|B$.*

Proof. (i) The set algebra generated by $\mathcal{B}$ is still finite, and thus generated by a (unique) partition $\{A_1, \dots, A_n\}$ of Ω. Choose $x_i \in A_i$ for each $i \le n$, and put $\rho_0 := \sum_{i=1}^n \rho(A_i) \cdot \varepsilon_{x_i}$.

(ii) Let F be the d.f. of ρ, i.e., $F(x) = \rho(]-\infty, x]) \; \forall x \in \mathbb{R}^d$. Then, apply (i) to $\mathcal{B} := \{]-\infty, b] \mid b \in B\}$, and denote by F_0 the d.f. of ρ_0. Since ρ_0 has finite support, Lemma 3 of [11] is applicable, whose (short) proof provides a d.f. $\tilde{F}$ as desired. □

Theorem 6. *Let* $f : [0,1]^m \to \mathbb{R}_+$ *be* d-↑ ($d \ge 2$) *and let* $g_1, \ldots, g_m : \mathbb{R}^d \to [0,1]$ *be d.f.s of* (*subprobability*) *measures on* $\mathbb{R}^d$. *Then, also* $f \circ (g_1, \ldots, g_m)$ *is a d.f. on* $\mathbb{R}^d$.

Proof. Put $g := (g_1, \ldots, g_m) : \mathbb{R}^d \to [0,1]^m$, $h := f \circ g$. By Lemma 1, also h is right-continuous, and it remains to show that h is $\mathbf{1}_d$-↑, the crucial property of a d.f. on $\mathbb{R}^d$.

A consequence of Theorem 5 is that we may assume f to be C^∞, and we first let also $g_1, \ldots, g_m$ be C^∞ functions.

Switching to the terminology in connection with the aforementioned generalized Faà di Bruno formula, we have to show $h_\nu \ge 0$ for $\nu \le \mathbf{1}_d$. Then, $|\nu| \le d$, and for $\lambda \in \mathbb{N}_0^m$ with $|\lambda| \le |\nu|$, we have $f_\lambda \ge 0$, by Theorem 4(ii). The condition

$$\sum_{j=1}^{s} |k_j|\ l_j = \nu$$

in the set $P_s(\nu, \lambda)$, together with $|k_j| > 0$ and $l_j \ne 0\ \forall j$, reduces to $|k_j| = 1 \forall j$ and

$$\sum_{j=1}^{s} l_j = \nu,$$

so that the l_j are "disjoint" in an obvious sense, i.e., $l_j \in \{0,1\}^d \setminus \{\mathbf{0}_d\}$ and $l_i \wedge l_j = \mathbf{0}_d$ for $i \ne j$. In particular, $g_{l_j} \ge 0\ \forall j$, each g_i being a d.f. Formula (**) now shows $h_\nu \ge 0$.

Now to the general case: in order to see that $h = f \circ g$ is $\mathbf{1}_d$-↑, we have to show for given $x \in \mathbb{R}^d$ and $\xi \in \mathbb{R}_+^d$

$$(\Delta_\xi^{\mathbf{1}_d} h)(x) = h(x + \xi) \mp \cdots + (-1)^d h(x) \ge 0$$

(as well as the analogue for some variables fixed, which is shown similarly).

In Lemma 3, we choose the finite set

$$\left\{ x + \sum_{i \in a} \xi_i e_i \mid a \subseteq [d] \right\} =: B$$

and find C^∞ d.f.s $\tilde{g}_1, \ldots, \tilde{g}_m$ such that $\tilde{g}_i | B = g_i | B$ for each $i \le m$. Then,

$$0 \le (\Delta_\xi^{\mathbf{1}_d} (f \circ \tilde{g}))(x) = (\Delta_\xi^{\mathbf{1}_d} h)(x),$$

thus finishing the proof. □

Remark 3. The aforementioned theorem answers positively a question in the concluding remarks of [5]. For $d = 2$, this result was shown in [6], Theorem 3.1.

Remark 4. If for a given f the conclusion of Theorem 6 holds for all d.f.s $g_1, \ldots, g_m$, then f must be d-↑. This follows from Theorem 1(v), since each component of an affine positive function φ is of course $\mathbf{1}_d$-↑.

Examples 3.

(a) We saw before that $f(x) := x_1x_2 + x_1x_3 + x_2x_3 - x_1x_2x_3$ is 2-↑ on $[0,1]^3$. Hence, for arbitrary bivariate d.f.s g_1, g_2 and g_3 also $g_1g_2 + g_1g_3 + g_2g_3 - g_1g_2g_3$ is a d.f., while f itself is not a three-dimensional d.f..

(b) Put $f_a(t) := (t - a)_+/(1 - a)$ for $t \in [0,1]$ and $a \in [0,1[$, complemented by $f_1 := 1_{\{1\}}$. Then, $\{f_a^n \mid a \in [0,1]\}$ are the "essential" extreme points for $(n+1)$-↑ functions on $[0,1]$, and $\{f_{a_1}^{n_1} \otimes \cdots \otimes f_{a_d}^{n_d} \mid a \in [0,1]^d\}$ correspondingly for $(\mathbf{n} + \mathbf{1}_d)$-↑ functions on $[0,1]^d$, cf. [12]. In the bivariate case, $f_a \otimes f_b$ is $(2,2)$-↑, in particular 2-↑, so that $f_c \circ (f_a \otimes f_b)$ is 2-↑ on $[0,1]^2$. For any bivariate d.f.s g_1 and g_2, we see that

$$\left[\frac{(g_1 - a)_+ \cdot (g_2 - b)_+}{(1 - a) \cdot (1 - b)} - c \right]_+, \quad (a, b, c) \in [0,1[^3$$

is again a bivariate d.f..

Another important property of k-↑ functions is their "universal" compatibility and composability within their class, which is made precise in the following.

Theorem 7. *Let $m, d, k \in \mathbb{N}$, $J \subseteq \mathbb{R}^m$ and $I \subseteq \mathbb{R}^d$ be non-degenerate intervals, $g = (g_1, \ldots, g_m) : I \to J$, $f : J \to \mathbb{R}$, each g_i and f being k-↑. Then, also $f \circ g$ is k-↑.*

Proof. The case $k = 1$ being obvious, let us assume $k \geq 2$. Since any non-degenerate interval is an increasing union of compact non-degenerate subintervals, we may choose $I = [0,1]^d$ and $J = [0,1]^m$.

By Theorem 1, we have to show that $h \coloneqq f \circ g$ is $\mathbf{n}$-↑ for any $\mathbf{n} \in \mathbb{N}_0^d$ such that $0 < |\mathbf{n}| \leq k$. Since the variables i with $n_i = 0$ do not enter, we may and do assume $\mathbf{n} \in \mathbb{N}^d$, in particular $k \geq d$. Then, each g_i is $\mathbf{n}$-↑, or equivalently, by [13, Theorem 5], $g_i \circ \sigma_{\mathbf{n}}$ is $\mathbf{1}_{|\mathbf{n}|}$-↑ on $\prod_{i \leq d} \left[0, \frac{1}{n_i}\right]^{n_i}$. Theorem 6 above now implies that also

$$f \circ (g_1 \circ \sigma_{\mathbf{n}}, \ldots, g_m \circ \sigma_{\mathbf{n}}) = h \circ \sigma_{\mathbf{n}}$$

is $\mathbf{1}_{|\mathbf{n}|}$-↑, which in turn means that h is $\mathbf{n}$-↑. □

Remark 5. We mentioned earlier that k-↑ functions were considered already in [2], where our Theorem 7 is stated as Theorem 2. However, the proof given there is not a real one, in my opinion: the function g disappears more or less after a few lines, the terminology and notation are nearly "chaotic," and I consider the reasoning incomprehensible. Of course, in theory, a completely "elementary" proof might be possible, but then discrete analogues of formula (**) would have to appear, and this might get "out of control." In [5, 8], Bronevich's Theorem 2 is cited, without any comments on the proof. The special case $k = 2$ is proved in [6].

An open problem

While $\mathbf{n}$-↑ functions on $[0,1]^d$, non-negative and normalized, are a Bauer simplex, with "essentially" certain powers of $\{f_{a_1} \otimes \cdots \otimes f_{a_d} \mid a \in [0,1]^d\}$ as their extreme points (Example 3(b)), not much so far is known for k-↑ functions. Let us consider $d = k = 2$ and

$$K \coloneqq \{f : [0,1]^2 \to [0,1] \mid f \text{ is } 2\text{-↑ and } f(1,1) = 1\}.$$

K is obviously convex and compact and also stable under (pointwise) multiplication. It is easy to see that each $f_c \circ (f_a \otimes f_b)$ is an extreme point of K – but that is it, for the time being.

Acknowledgments: Thanks are due to the two reviewers for their constructive suggestions that certainly improved the presentation.

Funding information: No funding is involved.

Conflict of interest: The author states no conflict of interest.

References

[1] Boas, R. P., & Widder, D. V. (1940). Functions with positive differences. *Duke Mathematical Journal*, *7*, 496–503.

[2] Bronevich, A. G. (2005). On the closure of families of fuzzy measures under eventwise aggregations. *Fuzzy Sets and Systems*, *153*, 45–70.

[3] Constantine, G. M., & Savits, T. H. (1996). A multivariate Faa di Bruno formula with applications. *Transactions of the AMS*, *348*, 503–520.

[4] Faà di Bruno, F.. (1855). Sullo sviluppo delle funzioni. *Annali di Scienze Mathematiche e Fisiche*, *6*, 479–480.

[5] Klement, E. P., Manzi, M., & Mesiar, R. (2010). Aggregation functions with stronger types of monotonicity, (pp. 418–424). In: *LNAI 6178*. New York: Springer.

[6] Klement, E. P., Manzi, M., & Mesiar, R. (2011). Ultramodular aggregation functions. *Information Sciences*, *181*, 4101–4111.
[7] Koumandos, S., & Pedersen, H. L. (2009). Completely monotonic functions of positive order and asymptotic expansions of the logarithm of Barnes double gamma function and Euler's gamma function. *Journal of Mathematical Analysis and Applications*, *355*, 33–40.
[8] Manzi, M. (2011). New construction methods for copulas and the multivariate case. *Tesi (Padova)*, BN 2013-396T.
[9] Marinacci, M., & Montrucchio, L. (2005). Ultramodular functions. *Mathematics of Operations Research*, *30*, 311–332.
[10] Ressel, P. (2011). Monotonicity properties of multivariate distribution and survival functions – With an application to Lévy-frailty copulas. *Journal of Multivariate Analysis*, *102*, 393–404.
[11] Ressel, P. (2012). Functions operating on multivariate distribution and survival functions – With applications to classical mean values and to copulas. *Journal of Multivariate Analysis*, *105*, 55–67.
[12] Ressel, P. (2014). Higher order monotonic functions of several variables. *Positivity*, *18*, 257–285.
[13] Ressel, P. (2019). Copulas, stable tail dependence functions, and multivariate monotonicity. *Dependence Modeling*, *7*, 247–258.
[14] Veretennikov, A. Y., & Veretennikova, E. V. (2016). On partial derivatives of multivariate Bernstein polynomials. *Siberian Advances in Mathematics*, *26*, 294–305.