Funkcje wypukłe raz jeszcze.

Na początek przypomnę podstawowe narzędzie, udowodnione w zeszłym semestrze:

Lemat o monotoniczności ilorazów różnicowych

Niech $P \subset \mathbb{R}$ będzie przedziałem. Jeżeli funkcja $f: P \to \mathbb{R}$ jest wypukła (odp. ściśle wypukła), to iloraz różnicowy $I_f : P \times P \setminus \{(x,x) : x \in P\} \to \mathbb{R}$,
$I_f(x,y) = \dfrac{f(x)-f(y)}{x-y}$, jest niemalejącą (odp. rosnącą) funkcją każdego ze swych argumentów, tj.

$$\forall_{\substack{x,y,u,v \in P \\ x \neq y,\ u \neq v}} \quad \text{jeżeli } \begin{matrix} x \overset{A}{\leq} u \\ \text{i} \\ y \overset{B}{\leq} v \end{matrix} \quad \text{to} \quad I_f(x,y) \overset{C}{\leq} I_f(u,v)$$

i jeżeli f jest ściśle wypukła i którakolwiek z nierówności A, B jest ostra, to również nierówność C jest ostra.

Łatwo możemy sprawdzić, że jest też odwrotnie:

Lemat: Jeżeli I_f jest niemalejącą / rosnącą funkcją którejś* ze swoich zmiennych, to f jest funkcją wypukłą / ściśle wypukłą.

* I_f jest funkcją symetryczną, $I_f(x,y) = I_f(y,x)$, więc jeżeli jest rosnącą / niemalejącą wzgl. pierwszej zmiennej, to wzgl. drugiej też i v.v.

<u>Dowód</u>

<u>Umowa</u>: $*$ nad nierównością oznacza, że ta nierówność jest ostra, gdy I_f jest funkcją rosnącą.

Dla dowolnych $x, y \in P$, $x < y$ oraz $\alpha \in (0,1)$ punkt $t = \alpha x + (1-\alpha) y$ spełnia $x < t < y$

Stąd i z założenia o I_f

$$I_f(x,y) \overset{*}{\leq} I_f(t,y), \text{ czyli}$$

$$\frac{f(x)-f(y)}{x-y} \overset{*}{\leq} \frac{f(t)-f(y)}{t-y} = \frac{f(\alpha x+(1-\alpha)y)-f(y)}{\alpha(x-y)}$$

skąd od razu dostajemy

$$f(\alpha x + (1-\alpha)y) \overset{*}{\leq} \alpha f(x) + (1-\alpha) f(y). \quad \square$$

Korzystając z tych dwóch lematów możemy udowodnić

<u>Twierdzenie</u>: Niech $f:(a,b) \to \mathbb{R}$, $a<b$, $a,b \in \mathbb{R} \cup \{\pm\infty\}$ będzie wypukła (odp. ściśle wypukła). Wówczas:

<u>Umowa</u>: w tym wypadku nierówności ozn. $(*)$ są ostre.

a) W każdym $z \in (a,b)$ funkcja f ma skończone pochodne jednostronne $f'_-(z)$ i $f'_+(z)$, przy czym $f'_-(z) \leq f'_+(z)$.

b) jeżeli $x, y \in (a,b)$, $x<y$, to $f'_+(x) \overset{*}{\leq} f'_-(y)$

c) jeżeli $z \in (a,b)$, to $\forall_{x \in (z,b)}\ f(x) \overset{*}{\geq} f(z) + f'_+(z)(x-z)$

i $\forall_{y \in (a,z)}\ f(y) \overset{*}{\geq} f(z) + f'_-(z)(y-z)$

d) f ma w (a,b) co najwyżej przeliczalnie wiele punktów, w których nie jest różniczkowalna.

Dowód: Iloraz różnicowy $I_f(z,x)$ jest niemalejącą (rosnącą, gdy f jest ściśle wypukła) funkcją zmiennej x, więc ma w każdym punkcie granice jednostronne (być może nieskończone). W szczególności istnieją

$$f'_+(z) = \lim_{x \to z^+} I_f(z,x) \quad \text{oraz} \quad f'_-(z) = \lim_{x \to z^-} I_f(z,x)$$

Wiemy też, że gdy ustalimy $x_1 < z < x_2$, $x_1, x_2 \in (a,b)$, to dla $x \in (z, x_2)$ mamy $I_f(z,x) \overset{*}{\leq} I_f(z,x_2)$, więc $f'_+(z) \overset{*}{\leq} I_f(z,x_2)$, co wyklucza $f'_+(z) = +\infty$.

Analogicznie, $f'_-(z) \overset{*}{\geq} I_f(z,x_1)$, więc $f'_-(z) \neq -\infty$.

Na koniec, jeżeli $y < z < x$, to $I_f(z,y) \overset{*}{\leq} I_f(z,x)$; skąd, gdy przejdziemy z x i y do granicy, $y \to z^-$, $x \to z^+$, otrzymamy

$f'_-(z) \leq f'_+(z)$, a więc a)

To w szczególności wyklucza $f'_-(z) = +\infty$ i $f'_+(z) = -\infty$, zatem obie pochodne jednostronne są skończone.

b) bardzo podobnie do a): dla $0 < t < y-x$

$$\begin{array}{ccc} I_f(x, x+t) & \overset{*}{\leq} & I_f(y-t, y) \\ \big\downarrow {\scriptstyle t \to 0^+} & & \big\downarrow {\scriptstyle t \to 0^+} \\ f'_+(x) & & f'_-(y) \end{array}$$

skąd $f'_+(x) \leq f'_-(y)$. Aby dostać ostrą nierówność gdy f jest ściśle wypukła, trzeba minimalnie subtelniej:

dla $0 < t < \frac{y-x}{3}$

$$I_f(x, x+t) \overset{*}{\leq} I_f\left(x, x+\frac{y-x}{3}\right) \overset{*}{\leq} I_f\left(x, x+\frac{2}{3}(y-x)\right) \overset{*}{\leq}$$

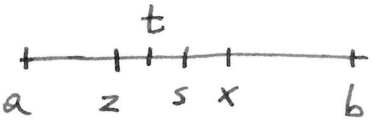
$$\overset{*}{\leq} I_f(y-t, y)$$

$$I_f(x, x+t) \xrightarrow{t \to 0^+} f'_+(x), \qquad I_f(y-t, y) \xrightarrow{t \to 0^+} f'_-(y)$$

skąd $f'_+(x) \overset{*}{\leq} f'_-(y)$.

c) udowodnimy pierwszą z nierówności, dowód drugiej jest analogiczny.

Niech $x \in (z,b)$; ustalmy $s \in (z,x)$. Wtedy, dla $t \in (z,s)$,



$$f'_+(z) \xleftarrow{t \to z^+} I_f(z,t) \overset{*}{\le} I_f(z,s) \overset{*}{\le} I_f(z,x),$$

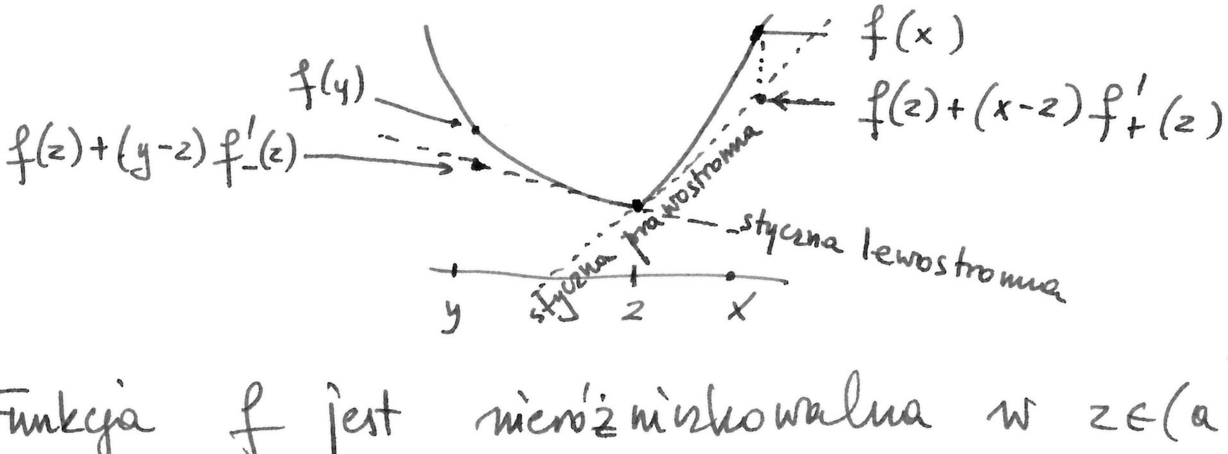
skąd dostajemy, że $f'_+(z) \le I_f(z,s) \overset{*}{\le} I_f(z,x)$,

$$f'_+(z) \overset{*}{\le} \frac{f(z)-f(x)}{z-x} = \frac{f(x)-f(z)}{x-z}$$

czyli

$$f(x) \overset{*}{\ge} \underline{f(z) + (x-z) f'_+(z)}$$

Zauważmy, że prawa strona to wzór prawostronnej stycznej do wykresu f w z.:



d) Funkcja f jest nieróżniczkowalna w $z \in (a,b)$, gdy $f'_-(z) < f'_+(z)$ (z a) wiemy, że obie pochodne jednostronne istnieją i $f'_-(z) \le f'_+(z)$).

Z b) wiemy, że jeżeli $x, y \in (a,b)$, $x<y$, to

$$f'_-(x) \overset{a)}{\le} f'_+(x) \overset{*}{\le} f'_-(y),$$

Dowód

a) Iloraz różnicowy $I_f(z,x)$ jest niemalejącą / $\overset{*}{\text{rosnącą}}$ funkcją zmiennej x, ma więc w każdym punkcie granice jednostronne, być może nieskończone.

W szczególności istnieją

$$f'_+(z)=\lim_{x\to z^+} I_f(z,x) \quad \text{oraz} \quad f'_-(z)=\lim_{x\to z^-} I_f(z,x).$$

Wiemy też, że, gdy ustalimy $x_1<z<x_2$, $x_1,x_2\in(a,b)$, to dla $x\in(z,x_2)$ $I_f(z,x)\le I_f(z,x_2)$, więc $f'_+(z)\le I_f(z,x_2)$, co wyklucza $f'_+(z)=+\infty$.

Analogicznie $f'_-(z)\ge I_f(z,x_1)$, więc $f'_-(z)\ne-\infty$.

Na koniec, jeżeli $y<z$, $x>z$, to $I_f(z,y)\overset{*}{\le} I_f(z,x)$, co, gdy przejdziemy do granicy: $y\to z^-$ i $x\to z^+$, otrzymamy

$$f'_-(z)\le f'_+(z), \text{ co wyklucza } \begin{array}{l} f'_-(z)=+\infty \\ f'_+(z)=-\infty \end{array}$$

i dowodzi a).

b) Bardzo podobnie:

$$\begin{array}{ccc} I(x,x+t) & \overset{*}{\le} & I(y-t,y) \quad \text{dla } t<y-x, \\ \downarrow t\to 0^+ & \Downarrow & \downarrow t\to 0^+ \\ f'_+(x) & \le & f'_-(y) \end{array}$$

c)

Niech $x > z$, ustalmy ~~t~~ $s \in (a,b)$ tż.

$z < s < x$. Wtedy, dla $t \in (z,s)$,

$$\begin{array}{c} I_f(z,t) \overset{*}{\leq} I_f(z,s) \overset{*}{\leq} I_f(z,x) \\ \Big\downarrow t \to z^+ \\ f'_+(z) \end{array}$$

skąd $f'_+(z) \overset{*}{\leq} I_f(z,x)$

$$f'_+(z) \overset{*}{\leq} \frac{f(z)-f(x)}{z-x} = \frac{f(x)-f(z)}{x-z}$$

czyli $f(x) \overset{*}{\geq} \underbrace{f(z) + (x-z) f'_+(z)}_{\text{styczna prawostronna do } f \text{ w } z}$

Druga nierówność, dla $x < z$, tak samo.

d) Zauważmy, że f nie jest różniczkowalna w $z \in (a,b)$, gdy $f'_-(z) < f'_+(z)$ (bo z a) wiemy, że $f'_-(z) \leq f'_+(z)$ i obie istnieją).
Mamy też z b), że jeżeli $x, y \in (a,b)$, $x < y$, to $f'_-(x) \leq f'_+(x) \overset{*}{\leq} f'_-(y)$, a więc funkcja $f'_- : (a,b) \to \mathbb{R}$ jest ~~niemalej~~ niemalejąca (* - rosnąca). Jako funkcja monotoniczna, f'_- ma w każdym punkcie granice prawostronne,

zatem funkcja $f'_-:(a,b)\to\mathbb{R}$ jest funkcją niemalejącą (rosnącą, gdy f ściśle wypukła), f'_- ma więc, jako funkcja monotoniczna, w każdym punkcie granice prawostronne, a jeżeli w pewnym punkcie z f nie jest różniczkowalna, to $f'_-(z) < f'_+(z) \leq f'_-(t)$ dla wszystkich $t>z$,

skąd $f'_-(z) < f'_+(z) \leq \lim\limits_{t\to z^+} f'_-(t)$.

To oznacza, że f'_- ma w z nieciągłość – skok.
Funkcja monotoniczna ma co najwyżej przeliczalnie wiele punktów nieciągłości. $\square$

Wnioski:

1. jeżeli f jest wypukła na (a,b) i różniczkowalna w $x_0\in(a,b)$, to wykres f leży nad styczną do wykresu w x_0; jeżeli f jest ściśle wypukła, to wykres f i styczna stykają się tylko w punkcie $(x_0, f(x_0))$.

(wynika od razu z c))

$$\forall_{x\in(a,b)} \quad f(x) \overset{*}{\geq} f(x_0) + f'(x_0)(x-x_0)$$

2. Jeżeli f jest wypukła na (a,b) i $f'(x_0)=0$, to f ma w x_0 minimum. Jeżeli f jest ściśle wypukła, to to minimum jest właściwe.

Dowód:

Z wniosku 1. wiemy, że $\forall_{x\in(a,b)}$

$$f(x) \overset{*}{\geq} f(x_0) + f'(x_0)(x-x_0) = f(x_0) \quad \square$$

3. Jeżeli f jest różniczkowalna na (a,b) i wypukła (odp. ściśle wypukła), to $f':(a,b)\to\mathbb{R}$ jest niemalejąca (odp. rosnąca).

(wynika to od razu z dowodu punktu d), gdzie wykazaliśmy, że $f'_-:(a,b)\to\mathbb{R}$ jest funkcją niemalejącą (odp. rosnącą)).

Zachodzi też twierdzenie odwrotne

Twierdzenie: Jeżeli f jest różniczkowalna na (a,b) i $f':(a,b)\to\mathbb{R}$ jest niemalejąca (odp. rosnąca), to funkcja f jest wypukła (odp. ściśle wypukła) na (a,b).

Dowód: Ustalmy $x,y\in(a,b)$ t.ż. $x<y$ i $\alpha\in(0,1)$.

Oznaczmy $t=\alpha x+(1-\alpha)y\in(x,y)$.

WARUNKI PIERWSZEGO RZĘDU NA WYPUKŁOŚĆ

Wówczas, z tw. Lagrange'a o wartości średniej,

$$I_f(x,t) = f'(\xi) \text{ dla pewnego } \xi \in (x,t),$$
$$I_f(t,y) = f'(\zeta) \text{ dla pewnego } \zeta \in (t,y).$$

umowa jak poprzednio

Stąd $$I_f(x,t) \overset{*}{\leq} I_f(t,y)$$

$$t-x=(1-\alpha)(y-x)$$
$$y-t=\alpha(y-x)$$

$$\frac{f(t)-f(x)}{t-x} \overset{*}{\leq} \frac{f(y)-f(t)}{y-t}$$

$$\frac{f(t)-f(x)}{t-x} = \frac{f(t)-f(x)}{(1-\alpha)(y-x)}, \qquad \frac{f(y)-f(t)}{y-t} = \frac{f(y)-f(t)}{\alpha(y-x)}$$

$$\alpha\big(f(t)-f(x)\big) \overset{*}{\leq} (1-\alpha)\big(f(y)-f(t)\big)$$

$$\Downarrow$$

$$f(\alpha x+(1-\alpha)y) = f(t) \overset{*}{\leq} \alpha f(x) + (1-\alpha)f(y),$$

co należało dowieść. □

<u>Wniosek</u>: Jeżeli funkcja f jest dwukrotnie różniczkowalna na (a,b), to

- f jest wypukła na (a,b) wtedy i tylko wtedy, gdy $f'' \geq 0$ na (a,b).
- f jest ściśle wypukła na (a,b) wtedy i tylko wtedy, gdy $f'' \geq 0$ na (a,b) i dla każdych $x,y \in (a,b)$, $x<y$, istnieje $\xi \in (x,y)$ t.ż. $f''(\xi) > 0$.

WARUNKI DRUGIEGO RZĘDU NA WYPUKŁOŚĆ

Dowód: Od razu z twierdzenia o pochodnych funkcji monotonicznych.

Oczywiście wszystkie te twierdzenia mają odpowiedniki dla funkcji wklęsłych i ściśle wklęsłych: trzeba w odp. miejscach pozmieniać kierunki nierówności, w miejsce funkcji rosnących/niemalejących dać malejące/nierosnące. Zostawię Państwu te dość oczywiste modyfikacje.

Newton ~ 1665
~~Abel~~ Euler przez wiele lat, 1774
Abel 1818,
Cauchy

Szereg dwumianowy Newtona

Def: Symbol Newtona $\binom{a}{n} := \frac{a(a-1)\cdot\ldots\cdot(a-n+1)}{n!}$

$$\binom{a}{0} := 1$$

można powyższym wzorem zdefiniować dla wszystkich $a \in \mathbb{R}$ i $n \in \mathbb{N} \cup \{0\}$.

Możemy od ręki sprawdzić, że $\binom{a}{n-1} + \binom{a}{n} = \binom{a+1}{n}$
oraz że $\binom{a+1}{n} = \frac{a+1}{n}\binom{a}{n-1}$.

Proste zadanka

Z1 $\forall_{a,b \in \mathbb{R}} \ \forall_{n \in \mathbb{N}\cup\{0\}} \ \sum_{k=0}^{n} \binom{a}{k}\binom{b}{n-k} = \binom{a+b}{n}$

Z2 szereg $\sum_{n=0}^{\infty} \binom{a}{n} x^n$ jest bezwzględnie zbieżny dla $x \in (-1,1)$ i dowolnego $a \in \mathbb{R}$.

Zdefiniujmy $f : \mathbb{R} \times (-1,1) \to \mathbb{R}$

$$f(a,x) = \sum_{n=0}^{\infty} \binom{a}{n} x^n.$$

(dla uproszczenia notacji będę czasem pomijał x, pisząc $f(a)$ i $f(b)$ zamiast $f(a,x)$ i $f(b,x)$ itp)

Zatem, dla ustalonego $n \in \mathbb{N}$

$$\left| \frac{f(a,x)-1}{a} - \sum_{k=1}^{\infty} \frac{(-1)^{k-1} x^k}{k} \right| =$$

$$\leq \left| \sum_{k=1}^{n} (1-a) \cdot \ldots \cdot \left(1-\frac{a}{k-1}\right) \frac{(-1)^{k-1} x^k}{k} - \sum_{k=1}^{n} \frac{(-1)^{k-1} x^k}{k} \right| +$$

$$+ \left| \sum_{k=n+1}^{\infty} (1-a) \cdot \ldots \cdot \left(1-\frac{a}{k-1}\right) \frac{(-1)^{k-1} x^k}{k} \right|$$

$$+ \left| \sum_{k=n+1}^{\infty} \frac{(-1)^{k-1} x^k}{k} \right| = A + B + C$$

Dla ustalonego n, $\lim_{a \to 0} A = 0$ (bo to skończona suma)

$B \leq \sum_{k=n+1}^{\infty} \frac{|x|^k}{\sqrt{k}}$ ← to jest szereg zbieżny (kryt. d'Alemberta)

dla $a \in (-\frac{1}{2}, \frac{1}{2})$

dla $x \in (-1,1)$, więc, przy ustalonym x, biorąc dost. duże n wiemy, że $B < \varepsilon$

Analogicznie $C \leq \sum_{k=n+1}^{\infty} \frac{|x|^k}{k} < \varepsilon$ o ile n jest dost. duże.

Ustalamy więc x, dobieramy n tak duże, by B i C były $\leq \varepsilon$. Dla tych x i n istnieje $\delta > 0$ taka, że gdy $|a| < \delta$, to również $A \leq \varepsilon$.

Stąd $A + B + C \leq 3\varepsilon$, o ile $|a| < \delta$, więc

$$\lim_{a \to 0} \frac{f(a,x)-1}{a} = \sum_{k=1}^{\infty} \frac{(-1)^{k-1} x^k}{k}.$$

Gdy $a \to 0$, k-ty wyraz tego szeregu dąży do $\frac{(-1)^{k-1} x^k}{k}$. Czy to znaczy, że

$$\lim_{a \to 0} \frac{f(a,x)-1}{a} \overset{\boxtimes}{=} \sum_{k=1}^{\infty} \frac{(-1)^{k-1} x^k}{k} \ ?$$

A priori tak być nie musi: małe odległości kolejnych wyrazów od ich granic mogą się sumować do dużej - może nawet nieskończonej - różnicy między obiema stronami równości $\boxtimes$.

Przykład, że mogą z tym być problemy:

niech $a_n(x) = \begin{cases} 1 & x \in [\frac{1}{n+1}, \frac{1}{n}) \\ 0 & \text{w p.p.} \end{cases}$

Wtedy $\forall_{x \in (0,1)} \sum_{n=1}^{\infty} a_n(x) = 1$, więc $\lim_{x \to 0} \sum_{n=1}^{\infty} a_n(x) = 1$,

choć $\forall_n \lim_{x \to 0} a_n(x) = 0$.

Musimy zatem sprawdzić starannie, czy

$$\frac{f(a,x)-1}{a} - \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} x^k \quad \text{dąży do zera.}$$

Na początek, oszacujmy dla $|a| < \frac{1}{2}$,

$$(1-a)(1-\tfrac{a}{2}) \cdot \ldots \cdot (1-\tfrac{a}{k-1}) \le e^{-a} \cdot e^{-\frac{a}{2}} \cdot \ldots \cdot e^{-\frac{a}{k-1}} \le$$

$$\le e^{\frac{1}{2}(1+\frac{1}{2}+\ldots+\frac{1}{k-1})} \le e^{\frac{1}{2}\ln k} = \sqrt{k}, \text{ skąd}$$

$$\left| (1-a) \cdot \ldots \cdot (1-\tfrac{a}{k-1}) \frac{(-1)^{k-1} x^k}{k} \right| \le \frac{|x|^k}{\sqrt{k}}$$

o ile $a \in (-\frac{1}{2}, \frac{1}{2})$.

Mamy

$$f(a)f(b) = \left(\sum_{n=0}^{\infty} \binom{a}{n} x^n\right)\left(\sum_{n=0}^{\infty} \binom{b}{n} x^n\right) =$$

$$= \sum_{n=0}^{\infty} \sum_{k=0}^{n} \binom{a}{k} x^k \binom{b}{n-k} x^{n-k} = \sum_{n=0}^{\infty} x^n \sum_{k=0}^{n} \binom{a}{k}\binom{b}{n-k} \overset{(21)}{=}$$

$$= \sum_{n=0}^{\infty} \binom{a+b}{n} x^n = f(a+b)$$

Przypomnijmy, że (twierdzenie): jeżeli $g: \mathbb{R} \to \mathbb{R}$ spełnia 1. $\forall_{a,b}\ g(a)g(b) = g(a+b)$ oraz

2. $\lim\limits_{a \to 0} \dfrac{g(a)-1}{a} = A$, to $g(a) = e^{Aa}$.

Żeby zastosować to twierdzenie do funkcji f, musimy jeszcze obliczyć (o ile w ogóle istnieje)

$$A(x) = \lim_{a \to 0} \frac{f(a,x)-1}{a}.$$

Jeżeli wyznaczymy $A(x)$, będziemy wiedzieli, że

$$f(a,x) = e^{A(x) \cdot a}.$$

$$\frac{f(a,x)-1}{a} = \frac{\sum_{k=0}^{\infty} \binom{a}{k} x^k - 1}{a} = \frac{\sum_{k=1}^{\infty} \binom{a}{k} x^k}{a} =$$

$$= \sum_{k=1}^{\infty} \frac{\cancel{a}(a-1)\cdot\ldots\cdot(a-k+1)}{\cancel{a}\cdot k!} x^k = \sum_{k=1}^{\infty} (1-a)\left(1-\frac{a}{2}\right)\cdot\ldots\cdot\left(1-\frac{a}{k-1}\right)\frac{(-1)^{k-1}x^k}{k}$$

Stąd $\quad f(a,x) = e^{A(x)\cdot a} = e^{\sum_{k=1}^{\infty} \frac{(-1)^{k-1}x^k}{k}\cdot a}$.

To nie jest szczególnie wygodne wyrażenie.

Zauważmy jednak, że dla $a=1$ $\quad \binom{a}{n} = 1$ dla $n=0$ i $n=1$, $\quad = 0$ dla $n>1$

więc $\quad f(a,x) = 1+x = e^{\ln(1+x)}$

$$f(a,x) = e^{A(x)\cdot 1}$$

Mamy więc $\quad \sum_{k=1}^{\infty} \frac{(-1)^{k-1}x^k}{k} = A(x) = \ln(1+x)$

i $\quad \sum_{k=0}^{\infty} \binom{a}{k} x^k = f(a,x) = e^{A(x)\cdot a} = e^{\ln(1+x)\cdot a} =$

$$= (1+x)^a.$$

Otrzymaliśmy w ten sposób szereg dwumianowy Newtona dla wszystkich $x\in(-1,1)$ i $a\in\mathbb{R}$ $\quad (1+x)^a = \sum_{k=0}^{\infty}\binom{a}{k}x^k$

i przy okazji rozwinięcie

$$\ln(1+x) = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}x^k}{k}.$$