Przestrzenie metryczne

Analizę, czyli badanie lokalnych właściwości zbiorów (albo tych właściwości określanie - bo jak się okaże są one do pewnego stopnia zależne od przyjętych do ich badania instrumentów) i funkcji na nich określonych, można uprawiać wtedy, gdy sens ma pojęcie wzajemnej odległości dwu elementów zbioru. Z tego powodu ważną rolę w matematyce odgrywa klasa przestrzeni zwanych przestrzeniami metrycznymi (przypomnijmy, że dla matematyka przestrzeń to taki zbiór, który jest na potrzeby chwili całym światem - trochę jak Wszechświat dla fizyka: pytanie co jest na zewnątrz Wszechświata albo matematycznej przestrzeni jest naogół, mówiąc językiem filozofii, nieprawomocne - a różne przestrzenie łączą ze sobą tylko odwzorowania, czyli abstrakcyjne przyporządkowania).

Przypomnienie: Przestrzeń metryczna jest to para: $(\mathcal{X},\, d)$, w której $\mathcal{X}$ jest zbiorem, a d funkcją, zwaną *metryką* określoną na $\mathcal{X} \times \mathcal{X}$ o wartościach w $\mathbb{R}_+$, (czyli po ludzku: maszynką z dwiema dziurami, do których wrzuca się dwa elementy zbioru $\mathcal{X}$ i dostaje w zamian rzeczywistą liczbę nieujemną, którą właśnie nazywamy odległością wzajemną tych dwu elementów) o następujących właściwościach (x, y, z są tu elementami $\mathcal{X}$):

i) $d(x,y) \geq 0$, przy czym $d(x,y) = 0$ tylko, gdy $x = y$,
ii) $d(x,y) = d(y,x)$,
iii) $d(x,y) \leq d(x,z) + d(z,y)$ ("nierówność trójkąta").

Nas najbardziej jako przestrzeń będzie interesować $\mathbb{R}^n$ (której elementami są uporządkowane n-ki liczb $(x_1, \ldots, x_n)$, które będziemy też pisać jako żywe wektory $\mathbf{x}$) i standardowa metryka zwana *euklidesową*, dana wzorem

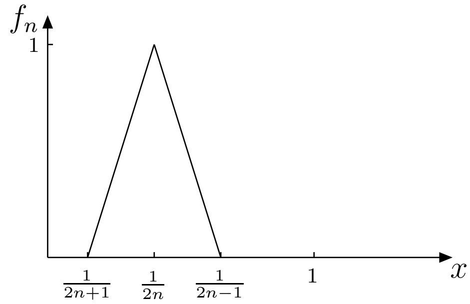
$$d(\mathbf{x},\, \mathbf{y}) = \sqrt{(x_1 - y_1)^2 + \ldots + (x_n - y_n)^2}\,,$$

ale siła matematyki polega na tym, że przestrzenie mogą być bardzo rozmaite, a ogólne reguły pozostają takie same i wszystko co z nich można wywnioskować jest prawdą niezależnie od tego, "czym" są przestrzenie, do których te reguły stosujemy. W samej przestrzeni $\mathbb{R}^n$ można też wprowadzić inne metryki, ale zanim do nich przejdziemy, rozpatrzmy (tak, aby się przestać bać matematyki - jak mówiłem kiedyś na wykładzie, trudną rzeczą jest twórcze uprawianie matematyki; zrozumieć to, co matematycy wymyślili, to już każdy może) jako przykład przestrzeń $\mathcal{X}$, którą jest[1] jednostkowy okrąg w $\mathbb{R}^2$, czyli zbiór punktów $\mathbb{R}^2$ spełniających warunek $x_1^2 + x_2^2 = 1$. Na tym zbiorze można określić jako metrykę funkcję d_1 poprzez przeniesienie nań metryki euklidesowej z $\mathbb{R}^2$ danej wzorem powyżej, albo funkcję d_2, która jest odległością od siebie dwu punktów liczoną "po łuku". Inny, bardziej "zaawansowany" przykład[2] to przestrzeń $\mathcal{C}[0,1]$ wszystkich rzeczywistych funkcji ciągłych[3] określonych na odcinku $[0,1]$ (przykład ten łatwo uogólnić na funkcje ciągłe

[1] A tu jakby można wyjść poza przestrzeń $\mathcal{X}$, w $\mathbb{R}^2$, w której $\mathcal{X}$ jest zanurzona. To dlatego, że tę prestrzeń konstruujemy przez "zanurzenie" w większej przestrzeni.

[2] Różne rzeczy dotyczące tego przykładu biorę z tetralogii Reeda i Simona, którą przy okazji bardzo wszystkim polecam, bo jej autorzy w wielu miejscach tłumaczą co i dlaczego robią; poza tym jest to dzieło zorientowane na fizykę i pozwalające zrozumieć różne matematyczne aspekty mechaniki kwantowej.

[3] Dlaczego ciągłych? Po to żeby nie trzeba było pytać w jakim sensie ta całka, która jest w definicji metryki d_2 i żeby funkcja nie robiła się nigdzie na odcinku $[0,1]$ nieskończona, co by pozbawiało sensu



Rysunek 1: Ciąg f_n rzeczywistych ciągłych funkcji na odcinku $[0, 1]$.

określone na odcinku $[a, b]$). Na tej przestrzeni można określić (m.in.) metryki (f i g to funkcje będące elementami, czyli punktami przestrzeni $\mathcal{C}[0, 1]$)

$$d_1(f, g) = \max_{x\in[0,1]} |f(x) - g(x)|, \qquad d_2(f, g) = \int_0^1 dx\, |f(x) - g(x)|\,.$$

Dopiero gdy na przestrzeni $\mathcal{X}$ jest zadana jakaś metryka, można dyskutować zagadnienie zbieżności ciągów x_n, $n = 1, 2, \dots$ elementów przestrzeni $\mathcal{X}$: mówimy, że ciąg taki zbiega do elementu $x \in \mathcal{X}$, czyli że $\lim_{n\to\infty} x_n = x$, gdy $d(x_n, x) \to 0$.

To, czy dany ciąg zbiega do czegoś w $\mathcal{X}$, czy nie, zależy jednak od przyjętej metryki d. W przykładzie pierwszym (w którym przestrzeń $\mathcal{X}$ była jednostkowym okręgiem w $\mathbb{R}^2$) zachodzą dość oczywiste nierówności (p i p' to są punkty na rozpatrywanym okręgu)

$$d_1(p, p') \leq d_2(p, p') \leq (\pi/2)\, d_1(p, p')\,,$$

co powoduje, że metryki te są równoważne: jeśli ciąg punktów p_n okręgu zbiega do punktu p w jednej tych metryk, to zbiega i w drugiej. Nie jest tak jednak w drugim przykładzie. Tu bowiem można tylko napisać nierówność

$$d_2(f, g) \leq d_1(f, g)\,,$$

i te dwie metryki nie są równoważne. Rozpatrzmy bowiem ciąg f_n rzeczywistych funkcji ciągłych na odcinku $[0, 1]$ zdefiniowanych na rysunku 1 (łatwiej spojrzeć na rysunek, niż czytać hieroglificzne wzorki). Ciąg ten w metryce d_2 jest zbieżny do funkcji $f(x) \equiv 0$. Nie jest on jednak zbieżny w metryce d_1, bo $d_1(f_n, 0) = 1$, niezależnie od n. Metryka d_1 (w której zbieżność ciągu funkcji jest zbieżnością jednostajną, podczas gdy w metryce d_2 jest ona tylko zbieżnością punktową) jest "silniejsza": każdy ciąg funkcji, który jest zbieżny w metryce d_1, jest też zbieżny w metryce d_2 (bo $d_2 \leq d_1$), ale, jak widać z powyższego przykładu, nie na odwrót.

---

definicję metryki d_1. Poza tym, gdyby dopuścić funkcje nieciągłe, to z $d_2(f, g) = 0$ nie wynikałoby, że $f = g$, tak jak tego wymaga definicja metryki: nieciągłe funkcje f i g mogłyby się bowiem różnić na zbiorach miary zero, czyli, mówiąc nieprecyzyjnie, w pojedyńczych punktach.

Pewnym zabawnym przykładem metryki, którą można wprowadzić w zasadzie na dowolnym zbiorze $\mathcal{X}$ jest tzw. metryka dyskretna:

$$d_{\text{discr}}(x,y) = \begin{cases} 0\,, & \text{gdy } \ x = y \\ 1\,, & \text{gdy } \ x \neq y \end{cases} .$$

W tej metryce zagadnienie zbieżności ciągów wygląda dość karykaturalnie (zbieżne są tylko ciągi stałe $x_n = x$), niemniej funkcja $d_{\text{discr}}(x,y)$ jest możliwą metryką, bo spełnia wszystkie podane w definicji metryki warunki.[4]

Mając metrykę na $\mathcal{X}$ można też zdefiniować ciągi Cauchyego, czyli takie, że dowolne dwa ich wyrazy, poczynając od pewnego N leżą dowolnie blisko siebie, tj. $d(x_n, x_{n'}) < \varepsilon$ dla dowolnie małego $\varepsilon > 0$, jeśli tylko $n', n > N(\varepsilon)$. Ciągi takie stanowią podstawowe narzędzie (śrubokręt i klapcążki) analizy matematycznej. Każdy ciąg zbieżny jest, jak łatwo zobaczyć, ciągiem Cauchy'ego, ale nie każdy ciąg Cauchy'ego jest zbieżny. Dlatego o takiej "dobrej" przestrzeni $\mathcal{X}$, w której każdy ciąg Cauchy'ego elementów $\mathcal{X}$ jest zbieżny (tzn., ma granicę, która jest też elementem $\mathcal{X}$) mówimy, że jest *zupełna*. Bo choć intuicyjnie wydaje nam się, że jak elementy x_n ciągu Cauchy'ego są coraz bliżej jedne drugich, to powinny do czegoś zbiegać, to wcale tak być nie musi. np. jeśli $\mathcal{X} = \mathbb{W}$ (przestrzeń liczb wymiernych) z metryką $d(x,y) = |x - y|$, a jako ciąg x_n weźmiemy

$$x_n = \left(1 + \frac{1}{n}\right)^n ,$$

to taki ciąg, którego wszystkie wyrazy są liczbami wymiernymi, nie zbiega, choć jest ciągiem Cauchy'ego, bo jak wiemy, jego granicą, jest niewymierna liczba e, ale ona nie należy do $\mathbb{W}$. I należy się tu postawić w pozycji Poldzia Kroneckera: dobry Pan Bóg dał nam zbiór liczb naturalnych; my chytrze skonstruowaliśmy z nich liczby wymierne, czyli przestrzeń $\mathbb{W}$, ale poza nią nic nie istnieje! To jest to, co wyżej było: przestrzeń $\mathbb{W}$ jest na razie całym światem - niema nic poza nim! I liczby niewymierne, czyli przestrzeń $\mathbb{R}$, trzeba dopiero stworzyć. Jak? Poprzez tzw. popołnienije Koszi przestrzeni $\mathbb{W}$, czyli przez podzielenie ciągów Cauchy'ego elementów tej przestrzeni na klasy równoważności (dwa ciągi Cauchy'ego, co wydają się zbiegać do tego samego, uznajemy za to samo), stworzeniu nowej przestrzeni $\tilde{\mathcal{X}} = \mathbb{R}$, której elementami są te klasy równoważności ciągów Cauchy'ego, które to klasy utożsamiamy z liczbami rzeczywistymi.

Oczywiście to, czy dana przestrzeń $\mathcal{X}$ jest zupełna, czy nie jest, zależy od tego, kakoj mietrikoj ana nadieliena: można np. pokazać (choć to już wyższa szkoła jazdy), że przestrzeń $\mathcal{C}[0,1]$ jest zupełna, gdy rozpatrujemy ją jako przestrzeń metryczną z metryką d_1, a niezupełna, gdy z metryką d_2.

Pojęcie wzajemnej odległości dwóch elementów zbioru, któremu sens nadaje metryka, jest potrzebne, by zdefiniować ciągłość odwzorowań. Odwzorowanie jednej przestrzeni metrycznej w drugą: $f : (\mathcal{X}, d_{\mathcal{X}}) \longrightarrow (\mathcal{Y}, d_{\mathcal{Y}})$ nazywamy ciągłym w punkcie $x \in \mathcal{X}$, gdy dla każdego ciągu $x_n \in \mathcal{X}$ (o $x_n \neq x$) zbieżnego (w metryce $d_{\mathcal{X}}$) do x ciąg $f(x_n)$

[4] Ale po co nam taka metryka? Nie wiem. Matematycy, choć genialni, są trochę jak dzieci: znajdują sobie zabawki i się nimi radośnie bawią, choć nikt poza nimi nie rozumie, z czego tak się cieszą...

jest zbieżny (w metryce $d_{\mathcal{Y}}$) do $f(x) \in \mathcal{Y}$. Powinno już być jasne, że ciągłość danego odwzorowania zależy od tego, jakie są metryki. Np. odwzorowanie $F : (\mathcal{C}[0,1], d_2) \longrightarrow (\mathcal{C}[0,1], d_1)$, które funkcji f przypisuje tę samą funkcję (czyli odwzorowanie tożsamościowe) nie jest ciągłe (rozpatrzyć przykład ciągu funkcji z rysunku 1), a odwzorowanie "odwrotne" $F^{-1} : (\mathcal{C}[0,1], d_1) \longrightarrow (\mathcal{C}[0,1], d_2)$, okazuje się, jest.

Dobrze jest też uogólnić kilka pojęć znanych z analizy uprawianej w przestrzeni $\mathbb{R}$ na dowolne przestrzenie metryczne. I tak

- *Kulą otwartą* $K(x_0, r)$ o promieniu r i środku w x_0 nazywa się zbiór wszystkich takich $x \in \mathcal{X}$, których odległość od x_0 jest mniejsza niż r. Hieroglifami: $K(x_0, r) = \{x \in \mathcal{X} |\ d(x, x_0) < r\}$.

- Zbiór $A \subset \mathcal{X}$ jest nazywany *otwartym* (w $\mathcal{X}$, przy zadanej metryce d), jeśli każdy element $x \in A$ jest środkiem jakiejś kuli otwartej $K(x, r)$ całkowicie zawartej w A (tzn. można tak dobrać r, żeby wszystkie punkty kuli należały do A). Otwartość zbioru zależy od przyjętej metryki. Np. w metryce dyskretnej każdy zbiór $A \subset \mathcal{X}$ jest otwarty, bo $K(x, 1) = \{x\} \subset A$.

- Zbiór O zwie się *otoczeniem* punktu $x_0 \in O$, jeśli istnieje kula otwarta $K(x_0, r) \subset O$. Można też ograniczyć się do otoczeń będących kulami.

- Punkt $x \in A$ jest *punktem wewnętrznym* zbioru $A \subset \mathcal{X}$ jeśli A jest otoczeniem punktu x. Czasem formułuje się to mówiąc, że istnieje kula $K(x, r) \subset A$; zbiór wszystkich punktów wewnętrznych zbioru A nazywa się jego wnętrzem (int$A$).

- Punkt $x \in A$ jest *punktem izolowanym* zbioru $A \subset \mathcal{X}$ jeśli istnieje taka kula $K(x, r)$, że $A \cap (K(x, r) - \{x\}) = \emptyset$, czyli kula o środku w x, która poza samym x innych punktów zbioru A nie zawiera.

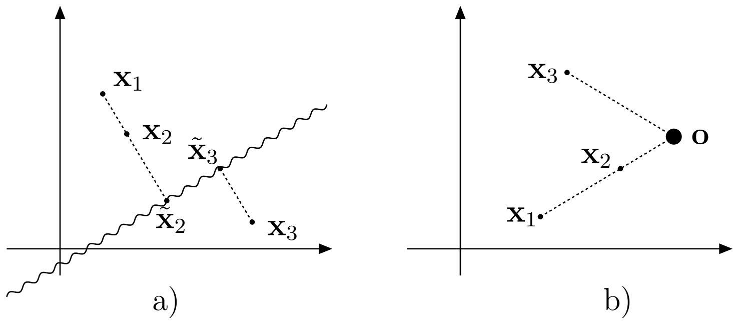
- Punkt x nazywa się *punktem skupienia* zbioru $A \subset \mathcal{X}$, gdy każda kula o środku w x (i dowolnym promieniu r) zawiera punkty należące do A. Sam punkt x może przy tym nie należeć do zbioru A. Skończony (czyli też dyskretny) zbiór punktów przestrzeni metrycznej nie posiada punktów skupienia.

- Zbiór A zwie się *zbiorem domkniętym*, gdy należą doń wszystkie jego punkty skupienia. Zbiór punktów skupienia dowolnego podzbioru A przestrzeni metrycznej jest więc zawsze domknięty. Domknięcie $\bar{A}$ zbioru $A \subset \mathcal{X}$ to sam zbiór A i wszystkie jego punkty skupienia. Domknięcie $\bar{A}$ zbioru A jest więc zbiorem domknietym.

- *Brzeg* ∂A zbioru $A \subset \mathcal{X}$ to $\partial A = \bar{A} - A$.

- Zbiór A jest ograniczony, gdy istnieje jakaś kula (o skończonym promieniu r), w której A się zawiera cały. Ograniczoność też może zależeć od przyjętej metryki: np. w metryce dyskretnej, każdy zbiór jest ograniczony.[5]

[5]Dlaczego? Zob. Zadanie 2 i pomyśl.



Rysunek 2: Dziwne metryki na $\mathbb{R}^2$: *a*) rzeczna (linia falista to rzeka) i *b*) kolejowa (gruba kropka to kolejowy wezeł centalny).

Po tych odstuplenijach powróćmy na grunt dobrze nam znanej przestrzeni $\mathbb{R}^n$. Jak już było powiedziane można w niej wprowadzić wiele metryk. Oprócz metryki dyskretnej, która jest nie jest specjalnie subtelna, możliwymi metrykami (oczywiście nie są to wszystkie możliwe!) są ($\mathbf{x} \equiv (x_1, \ldots, x_n)$)

- metryka maximum $d_\infty(\mathbf{x}, \mathbf{y}) = \max_{i=1,\ldots,n}|x_i - y_i|$,
- metryka nie wiedzieć czemu zwana metryką taxi $d_{\rm taxi}(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^n |x_i - y_i|$ (zwana jest ona także[6] metryką "miejską"),
- wspomniana już metryka euklidesowa $d_{\rm Eucl}(\mathbf{x}, \mathbf{y}) = [\sum_{i=1}^n (x_i - y_i)^2]^{1/2}$.

Dla zabawy można też np. w $\mathbb{R}^2$ wprowadzić tzw. metrykę "rzeczną" pokazaną na rysunku 2a, w której

$$d_{\rm riv}(\mathbf{x}_1, \mathbf{x}_2) = d_{\rm Eucl}(\mathbf{x}_1, \mathbf{x}_2)\,, \qquad d_{\rm riv}(\mathbf{x}_2, \mathbf{x}_3) = d_{\rm Eucl}(\mathbf{x}_1, \tilde{\mathbf{x}}_2) + d_{\rm Eucl}(\tilde{\mathbf{x}}_2, \tilde{\mathbf{x}}_3) + d_{\rm Eucl}(\tilde{\mathbf{x}}_3, \mathbf{x}_3)\,,$$

albo pokazaną na rysunku 2b metrykę "kolejową" w której

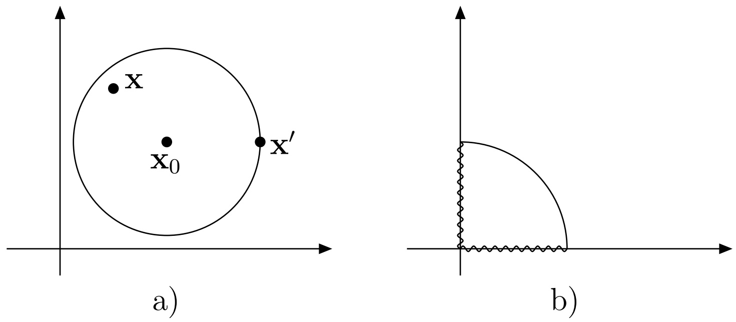
$$d_{\rm kol}(\mathbf{x}_1, \mathbf{x}_2) = d_{\rm Eucl}(\mathbf{x}_1, \mathbf{x}_2)\,, \qquad d_{\rm kol}(\mathbf{x}_2, \mathbf{x}_3) = d_{\rm Eucl}(\mathbf{x}_2, \mathbf{o}) + d_{\rm Eucl}(\mathbf{o}, \mathbf{x}_3)\,.$$

Można się zabawić sprawdzaniem, że to też są dobre metryki.

Przykłady kuli otwartej w przestrzeni metrycznej $(\mathbb{R}^2, d_{\rm Eucl})$ i zbioru A, który nie jest ani otwarty ani domknięty pokazuje rysunek 3. Zbiór A jest zdefiniowany jako $A = \{(x_1, x_2) \in \mathbb{R}^2 |\ x_1^2 + x_2^2 < 1,\ x_1 \geq 0,\ x_2 \geq 0\}$. Oba zbiory pokazane na tym rysunku są ograniczone (kula choćby dlatego, że jest zawarta sama w sobie). Inne jeszcze przykłady (już bez rysunków):

- zbiór $A = \{(x_1, x_2) \in \mathbb{R}^2 |\ a < x_1 < b,\ c < x_2 < d\}$ jest otwarty i ograniczony, bo jest np. zawarty w kuli $K((x_1^{(0)}, x_2^{(0)}), r)$, gdzie $x_1^{(0)} = \frac{1}{2}(a + b)$, $x_2^{(0)} = \frac{1}{2}(c + d)$, $r = \frac{1}{2}\sqrt{(b - a)^2 + (d - c)^2}$,

[6] No, to można już zrozumieć: niektórzy po mieście się obwożą taksówką (zamiast na rowerze...), choć akurat krakusy taksówkami nazywają prywatne samochody, a naszą warszawską taksówkę nazywają "taryfą".



Rysunek 3: Przykłady zbiorów w przestrzeni $\mathbb{R}^2$: *a*) Kula $K(\mathbf{x}_0, r)$; punkt $\mathbf{x}$ jest punktem skupienia tej kuli, punkt $\mathbf{x}'$ też, ale $\mathbf{x}$ należy do K, a $\mathbf{x}'$ nie należy; *b*) Zbiór A, który nie jest ani otwarty ani domknięty; punkty linii zaznaczonych falkami należą do A; punkty leżące na ćwierćłuku do A nie należą; i jedne i drugie są punktami skupienia zbioru A.

- zbiór $B = \{(x_1, x_2) \in \mathbb{R}^2 |\ |x_1| \leq |a|,\ |x_2| \leq |b|\}$ jest i ograniczony i domknięty,
- zbiór $C = \{(x_1, x_2) \in \mathbb{R}^2 |\ 0 < \sin x_1 < \frac{1}{2}\}$ jest otwarty i nieograniczony ($x_1 \in (0, \frac{1}{6}\pi) \cup (2\pi, \frac{13}{6}\pi) \cup \ldots \cup (\frac{5}{6}\pi, \pi) \cup (\frac{17}{6}\pi, 3\pi) \cup \ldots$)
- zbiór $D = \{(x_1, x_2) \in \mathbb{R}^2 |\ x_1, x_2 \in \mathbb{W}, 0 < x_1, x_2 < 1\}$ (liczby x_1 i x_2 są wymierne z otwartego przedziału (0,1)) nie jest domknięty bo np. jego punkt skupienia $(\frac{1}{\sqrt{2}}, \frac{1}{2})$ doń nie należy ale w każdym otoczeniu tego punktu (albo w każdej, o dowolnie małym promieniu r kuli o środku w tym punkcie) są inne punkty należące do D (dlatego jest to punkt skupienia zbioru D), np. punkty $(x_1^{(n)}, \frac{1}{2})$, gdzie $x_1^{(n)}$ to n-te dziesiętne przybliżenie liczby $\frac{1}{\sqrt{2}}$; ale nie jest też i otwarty, bo w dowolnie małym otoczeniu (dowolnie małej kuli ześrodkowanej na nim) każdego punktu (x_1, x_2) o wymiernych x_1, x_2, znajdują się także punkty o niewymiernych x_1, x_2 (pomiędzy każdymi dwiema różnymi liczbami wymiernymi znajdzie się jakaś liczba niewymierna).

Z ostatniego przykładu widać, że zbiór nieotwarty nie musi być domknięty, a niedomknięty nie musi być otwarty. Samo zaś $\mathbb{R}^2$ jako podzbiór $\mathbb{R}^2$ jest przykładem zbioru, który jest jednocześnie i otwarty i domknięty. Prawdą jest natomiast, że (i jest to stwierdzenie ogólne, dotyczące dowolnych przestrzeni metrycznych) jeśli zbiór $A \in \mathcal{X}$ jest otwarty, to zbiór $\mathcal{X} - A$ jest domknięty i na odwrót: jeśli zbiór $A \in \mathcal{X}$ jest domknięty, to zbiór $\mathcal{X} - A$ jest otwarty. Inne stwierdzenia ogólne to:

- Kula $K(x_0, r)$ zdefiniowana wyżej i trochę na wyrost nazwana "otwartą", jest rzeczywiście zbiorem otwartym niezależnie od metryki (Zadanie 7).
- Kula $\bar{K}(x_0, r) = \{x \in \mathcal{X} |\ d(x_0, x) \leq r\}$ jest zbiorem domkniętym niezależnie od metryki.
- Jeśli x_0 jest punktem skupienia zbioru $A \subset \mathcal{X}$, to każda kula $K(x_0, r)$ zawiera nieskończenie wiele punktów zbioru A.

- Jeśli zbiory O_i, $i = 1, \ldots, n < \infty$ są w $(\mathcal{X}, d)$ zbiorami otwartymi, to $O_1 \cup \ldots \cup O_n$ oraz $O_1 \cap \ldots \cap O_n$ też są w $(\mathcal{X}, d)$ zbiorami otwartymi. Jeśli chodzi o sumę zbiorów to prawdziwe jest mocniejsze stwierdzenie: mianowicie, dowolna suma (nawet nieskończonej rodziny) zbiorów otwartych jest zbiorem otwartym; ale tylko przecięcie skończonej liczby zbiorów otwartych musi być zbiorem otwartym.

- Jeśli zbiory D_i, $i = 1, \ldots, n < \infty$ są w $(\mathcal{X}, d)$ zbiorami domkniętymi, to $D_1 \cup \ldots \cup D_n$ oraz $D_1 \cap \ldots \cap D_n$ też są w $(\mathcal{X}, d)$ zbiorami domkniętymi. I tu na odwrót: to przecięcie dowolnej (nawet nieskończenie licznej) rodziny zbiorów domkniętych musi być zbiorem domkniętym, podczas gdy domknięta musi być suma tylko skończonej liczby zbiorów domkniętych.

Bardzo ważnym pojęciem ogólnym jest *zwartość* zbioru: zbiór $A \in \mathcal{X}$ jest *zbiorem zwartym*, jeśli z każdego ciągu punktów $x_n \in A$ można wyjąć podciąg zbieżny do $x \in A$. Inny sposób zdefiniowania zwartości jest jeszcze bardziej abstrakcyjny (jako, że jest on *par excellence* topologiczny). *Pokryciem otwartym* zbioru A zawartego w przestrzeni metrycznej $(\mathcal{X}, d)$ nazywa się rodzinę $\{O_s\}_{s \in S}$ zbiorów otwartych (S jest pewnym zbiorem indeksów, niekoniecznie skończonym i niekoniecznie nawet przeliczalnym) takich, że

$$A \subset \bigcup_{s \in S} O_s \, .$$

Zbiór A jest zwarty, jeśli z każdego jego pokrycia otwartego można wybrać podpokrycie skończone (tzn. zostawić tylko skończoną liczbę zbiorów O_s tak, że one nadal pokrywają A w tym sensie, że zbiór A jest zawarty w ich sumie).

Pojęcie zwartości jest jednak zbędne, gdy uprawiamy analizę w przestrzeniach $\mathbb{R}^n$, bo w takich przestrzeniach zwartość zbioru jest (na mocy twierdzenia Borela-Lebesgue'a) równoważna jego domkniętości i ograniczoności (równoważność taka nie zachodzi jednak w dowolnych przestrzeniach metrycznych, a w szczególności w przestrzeniach funkcji, takich jak $\mathcal{C}[0, 1]$ i innych, i dlatego zwartość odgrywa ważną rolę w analizie funkcjonalnej, która jest właśnie analizą uprawianą w abstrakcyjnych przestrzeniach metrycznych).

Jeśli danych jest kilka przestrzeni metrycznych $(\mathcal{X}_i, d_i)$, $i = 1, \ldots, k$, to iloczyn kartezjański $\mathcal{X}_1 \times \ldots \times \mathcal{X}_k$ można wyposażyć w naturalna metrykę iloczynową (się mówi "produktową", ale to jakoś nie po naszemu), w której odległością punktów $\mathbf{x} \equiv (x_{(1)}, \ldots, x_{(k)})$ i $\mathbf{y} \equiv (y_{(1)}, \ldots, y_{(k)})$, gdzie $x_{(i)}, y_{(i)} \in \mathcal{X}_i$, jest

$$d(\mathbf{x}, \mathbf{y}) = \left[\sum_{i=1}^{k} (d_i(x_{(i)}, y_{(i)}))^2 \right]^{1/2} .$$

Ważną klasą przestrzeni, które w sposób naturalny (matematycy używają tu swojego ulubionego słówka "kanoniczny" - to im daje poczucie legitymizacji ich działań) można przekształcić w przestrzenie metryczne są wektorowe (mówi się też liniowe) przestrzenie unormowane. Ogólnie przestrzeń liniowa unormowana, to taka para $(V, \ \| \cdot \|)$, w której

V jest przestrzenią wektorową (dla nas nad $\mathbb{R}$ lub $\mathbb{C}$), a $\|\cdot\|$ odwzorowaniem V w $\mathbb{R}_+$ o właściwościach:

i) $\|\mathbf{v}\| \geq 0$, przy czym $\|\mathbf{v}\| = 0$ tylko gdy $\mathbf{v} = \mathbf{0}$,
ii) $\|\alpha\mathbf{v}\| = |\alpha|\,\|\mathbf{v}\|$, gdzie α jest liczbą z ciała,
iii) $\|\mathbf{v}+\mathbf{u}\| \leq \|\mathbf{v}\| + \|\mathbf{u}\|$ ("nierówność trójkąta").

Jeśli w przestrzeni wektorowej jest zadany iloczyn skalarny $(\cdot|\cdot)_S$ - może będzie na wykładzie później w części algebraicznej, a na pewno o iloczynie skalarnym jest w moim skrypcie - to automatycznie jest też i norma $\|\mathbf{v}\| \equiv (\mathbf{v}|\mathbf{v})_S$. Dzięki liniowości norma zadaje też naturalną metrykę wzorem

$$d(\mathbf{v}_2, \mathbf{v}_2) \equiv \|\mathbf{v}_1 - \mathbf{v}_2\|\,.$$

Każda więc przestrzeń wektorowa z jakimś iloczynem skalarnym (a w każdej skończenie wymiarowej przestrzeni wektorowej, czyli takiej, jakimi się zajmowaliśmy w części algebraicznej, można jakiś iloczyn skalarny bardzo łatwo wprowadzić) jest w naturalny sposób przestrzenią metryczną.

Analiza w $\mathbb{R}^n$

Ciąg w $\mathbb{R}^n$ (tu będziemy mieć zawsze na myśli metrykę euklidesową) to po prostu zespół n ciągów z $\mathbb{R}$:

$$\mathbf{x}_k \equiv (x_k^1, \ldots, x_k^n)\,.$$

Ciąg (x_k, y_k) w $\mathbb{R}^2$ (żeby mniej pisać) jest zbieżny do (a, b), gdy $x_k \to a$ i jednocześnie $y_k \to b$ w sensie zwykłych ciągów z $\mathbb{R}$ (z metryką zadaną przezmoduł różnicy). Np. z trzech ciągów w $\mathbb{R}^2$

$$(\frac{1}{k},\, (-1)^k)\,, \qquad (1-\frac{1}{k},\, \frac{1}{k}\sin k)\,, \qquad (2,\, \cos\frac{1}{k})\,,$$

pierwszy jest niezbieżny (bo $(-1)^k$ nie zbiega), drugi zbieżny do $(1,0) \in \mathbb{R}^2$, a trzeci zbieżny do $(2,1) \in \mathbb{R}^2$.

Będziemy mieć do czynienia z różnymi odwzorowaniami:

- $\mathbb{R} \to \mathbb{R}$ (to było ćwiczone w pierwszym semestrze),
- $\mathbb{R}^n \to \mathbb{R}$, np. $f(x,y) = x^2 + y^2$ jako przykład $\mathbb{R}^2 \to \mathbb{R}$,
- $\mathbb{R} \to \mathbb{R}^n$ - mechanika newtonowska pojedyńczej cząstki jest naogół badaniem trajektorii cząstki $\mathbf{r}(t)$, czyli badaniem odwzorowania $\mathbb{R} \to \mathbb{R}^3$
- mogą wreszcie być odwzorowania $\mathbb{R}^m \to \mathbb{R}^n$ takie jak np. statyczne pole elektryczne punktowego ładunku Q, które jest odwzorowaniem $\mathbb{R}^3 \to \mathbb{R}^3$ danym wzorami[7]

$$E^x(x,y,z) = \frac{Q\,(x-x_0)}{[(x-x_0)^2 + (y-y_0)^2 + (z-z_0)^2]^{3/2}}\,,$$

[7] Piszemy je tu w naturalnych jednostkach Gaussa - SI to jest układ dla idiotów - système des idiots, jak sama jego nazwa wskazuje.

$$E^y(x,y,z) = \frac{Q\,(x-x_0)}{[(x-x_0)^2+(y-y_0)^2+(z-z_0)^2]^{3/2}}\,,$$
$$E^z(x,y,z) = \frac{Q\,(x-x_0)}{[(x-x_0)^2+(y-y_0)^2+(z-z_0)^2]^{3/2}}\,,$$

Jak widać odwzorowanie $\mathbb{R}^m \to \mathbb{R}^n$ to jest po prostu n odwzorowań $\mathbb{R}^m \to \mathbb{R}$. Innym przykładem odwzorowania $\mathbb{R}^m \to \mathbb{R}^n$ jest znane nam z algebry odwzorowanie liniowe m wymiarowej przestrzeni wektorowej w n-wymiarową przestrzeń wektorową, które przy ustalonych bazach obu tych przestrzeni ma postać (a_{ij} są tu stałymi)

$$\begin{pmatrix} y^1(x^1,\dots,x^m) \\ y^2(x^1,\dots,x^m) \\ \dots \\ y^n(x^1,\dots,x^m) \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ \cdot \\ x^m \end{pmatrix}.$$

Ostatni przykład jest bardzo ważny: okaże się bowiem, że pochodna (ale jak jest ona zdefiniowana, to dopiero trzeba będzie sobie powiedzieć) każdego odwzorowania $\mathbb{R}^m \to \mathbb{R}^n$ jest właśnie (w danym punkcie $\mathbb{R}^m$) takim odwzorowaniem liniowym.

Trzeba teraz zająć się zagadnieniem ciągłości odwzorowania. Jedna definicja ciągłości już była sformułowana ale ją tu jeszcze raz podamy w nieco innej formie, najpierw definiując granicę funkcji w punkcie: f_0 jest *granicą* funkcji $f(\mathbf{x})$ w punkcie $\mathbf{x}_0$, gdy[8] dla każdego $\varepsilon > 0$ można dobrać $\delta > 0$, że $|f(\mathbf{x}) - f_0| < \varepsilon$, gdy $|\mathbf{x}-\mathbf{x}_0| < \delta$ (tzw. definicja Cauchy'ego granicy funkcji). Inna definicja, tzw. Heinego (albo ciągowa, równoważna tej Cauchy'ego) jest taka: f_0 jest granicą funkcji $f(\mathbf{x})$ w punkcie $\mathbf{x}_0$, gdy f_0 jest granicą każdego ciągu punktów $f(\mathbf{x}_k)$ otrzymanego jako obraz dowolnego ciągu $\mathbf{x}_k$, byle takiego, że (i to jest tu bardzo istotny warunek!) $\mathbf{x}_k \neq \mathbf{x}_0$, zbieżnego do $\mathbf{x}_0$. Odwzorowanie f jest w punkcie $\mathbf{x}_0$ ciągłe, gdy jego granica w tym punkcie istnieje i jest równa wartości funkcji w tym punkcie (bo zawsze można sobie arbitralnie zdefiniować, że danym punkcie funkcja ma jakąś wartość). Definicja Heinego jest bardzo wygodna, gdy chcemy pokazać, że funkcja f *nie* może być ciągła w danym punkcie (bo nie ma tam granicy): wystarczy wymyślić jakiekolwiek dwa ciągi $\mathbf{x}_k$ i $\mathbf{x}'_k$ oba zbieżne do $\mathbf{x}_0$ i pokazać że $\lim f(\mathbf{x}_k) \neq \lim f(\mathbf{x}'_k)$. Trudniej jest wykazać z jej pomocą, że funkcja może być w danym punkcie ciągła, bo trzeba pokazać że granica $f(\mathbf{x}_k)$ jest ta sama dla *wszystkich możliwych* ciągów zbieżnych do $\mathbf{x}_0$. Niemniej czasem się to w miarę prosto daje zrobić.

Rozpatrzmy np. funkcję $\mathbb{R}^2 \to \mathbb{R}$ (głównie takich odwzorowaniach będziemy badać granice i ciągłość, bo jak już zauważyliśmy, odwzorowanie $\mathbb{R}^n \to \mathbb{R}^m$ to jest m odwzorowań $\mathbb{R}^n \to \mathbb{R}$ i każde z nich z osobna można badać w $\mathbf{x}_0$ pod tym kątem).

$$f(x,y) = \begin{cases} \frac{1}{x}\sin(xy), & x \neq 0 \\ 0, & x = 0 \end{cases},$$

[8]Ponieważ będziemy mieć zawsze na myśli metrykę euklidesową, zamiast pisać $d(\mathbf{x},\mathbf{y})$ będzemy odtąd pisać $|\mathbf{x}-\mathbf{y}|$.

(wzdłuż całej osi y funkcja jest zadana odrębnym wzorem) i zbadajmy jej ciągłość w punktach postaci $(0, y)$. Rozpatrzymy osobno punkt $(0,0)$ i osobno puntkty postaci $(0, y)$ z $y \neq 0$. Weźmy najpierw jakieś ciągi $x_n \to 0$ i $y_n \to 0$ (ale pamiętamy, że jednocześnie, tzn. dla tego samego n nie może być $x_n = 0$ i $y_n = 0$) i niech $c_n = f(x_n, y_n)$. Jeśli $y_n \equiv 0$, a $x_n \neq 0$ (zbiegamy do punktu $(0,0)$ wzdłuż osi x), to stale $c_n = 0$, więc jeśliby granica funkcji w $(0,0)$ miała istnieć, to musiałaby być równa zeru. Ogólniej, jeśli $x_n y_n \to 0$, co zachodzi zawsze gdy $(x_n, y_n) \to (0,0)$, to $\sin(x_n y_n) \to x_n y_n$ i $c_n \to y_n$. Więc ponieważ rozpatrujemy takie ciągi, że $y_n \to 0$, granicą funkcji w $(0,0)$ jest zero i arbitralnie przypisana funkcji (w podanym wzorze ją definiującym) wartość 0 w punkcie $(0,0)$ "pasuje": funkcja tak zdefiniowana jest w $(0,0)$ ciągła. Jednak w punktach postaci $(0, y)$ z $y \neq 0$ jest gorzej: Jeśli rozpatrujemy np. ciągi $(x_n, y_n) \to (0,1)$, to $\sin(x_n y_n) \to x_n y_n$ i $c_n \to y_n \to y = 1$. Tymczasem podany wzór przypisuje funkcji w punkcie $(0,1)$ (i wszystkim punktom na osi y) wartść zero! Zatem funkcja tak zdefiniowana jest ciągła w punkcie $(0,0)$ ale nie w punktach postaci $(0, y)$ z $y \neq 0$ (oczywiście poza osią y jest ona bezdyskusyjnie ciągła). Aby była ona ciągła (ale nie byłaby to ona, tylko już inna funkcja) trzebaby ją określić wzorem

$$f(x,y) = \begin{cases} \frac{1}{x}\sin(xy), & x \neq 0 \\ y, & x = 0 \end{cases}.$$

Drugi (typowy) przykład, to funkcja $\mathbb{R}^2 \to \mathbb{R}$ poza punktem $(0,0)$ zadana wzorem

$$f(x,y) = \frac{x\,y}{x^2 + y^2}\,.$$

Zobaczmy, czy ma ona jakąś granicę w $(0,0)$ (gdyby miała, można by ją odpowiedno zdefiniować, tak by była tam ciagła). W tym celu wystarczy rozpatrzyć klasę ciągów (x_n, y_n), w których $x_n = a/n$, $y_n = b/n$ (różne wartości a i b dają różne ciągi). Wtedy

$$f(x_n, y_n) = \frac{x_n\,y_n}{x_n^2 + y_n^2} = \frac{a\,b}{a^2 + b^2}\,,$$

niezależnie od n, a to oznacza, że

$$\lim_{n\to\infty} f(x_n, y_n) = \frac{a\,b}{a^2 + b^2}\,,$$

czyli że granica zależy od ciągu. Gdy zbiegamy do punktu $(0,0)$ wzdłuż osi x (ciągi o $b = 0$), lub wzdłuż osi y (ciągi o $a = 0$), to dostajemy w granicy zero. Ale jeśli zbiegamy wzdłuż diagonali ($a = b$) lub antydiagonali ($a = -b$), to granicą jest $1/2$ lub $-1/2$. Zatem badana funkcja nie ma granicy w punkcie $(0,0)$. Do tej samej konkluzji można dojść rozpatrując ciągi postaci $(r_n \cos\varphi_n,\ r_n \sin\varphi_n)$, co odpowiada zapisaniu tej samej funkcji w zmiennych biegunowych (fizyk powinien mieć w głowie, że zawsze można wprowadzić na płaszczyźnie, czy dowolnej przestrzeni $\mathbb{R}^n$ jakieś inne, być może dogodniejsze dla analizy danego konkretnego problemu, zmienne). Aby taki ciąg $(x_n,\ y_n)$ zbiegał do punktu $(0,0)$

wystarczy, by $r_n \to 0$; ciąg φ_n może się przy tym zachowywać dowolnie (np. jeśli φ_n rośnie monotonicznie, odpowiada to zbieganiu do punktu $(0,0)$ po jakiejś spirali). Mamy wtedy

$$f(r_n \cos\varphi_n,\ r_n \sin\varphi_n) = \frac{1}{2}\sin 2\varphi_n\,,$$

i od razu widać, że granica nie istnieje, bo $\sin 2\varphi_n$ może zbiegać do dowolnej wartości (z przedziału $[-1,1]$) albo w ogóle do niczego nie zbiegać.

Trzeci przykład to funkcja $\mathbb{R}^2 \to \mathbb{R}$ poza punktem $(0,0)$ zadana wzorem

$$f(x,y) = \frac{x^2\,y}{x^2+y^2}\,.$$

Znów zobaczmy, czy ma ona jakąś granicę w $(0,0)$. Weźmy najpierw tę samą klasę ciągów, którą wykorzystaliśmy w przykładzie drugim. Tym razem

$$f(x_n,y_n) = \frac{x_n^2\,y_n}{x_n^2+y_n^2} = \frac{a^2\,b}{n\,(a^2+b^2)} \to 0\,,$$

i granica ta nie zależy od a i b. Jest więc lepiej, ale to jeszcze niczego nie dowodzi, bo to tylko jedna klasa ciągów, która nie wyczerpuje wszystkich możliwości (do punktu $(0,0)$ można zbiegać np. po paraboli, albo po jakiejś spirali, albo jeszcze jakoś inaczej). Spróbujmy więc rozumować bardziej ogólnie: załóżmy, że $(x_n,y_n) \to (0,0)$ czyli, że $x_n \to 0$, $y_n \to 0$, co oznacza też, że i $x_n^2+y_n^2 \to 0$. Niech $c_n \equiv \max(|x_n|,|y_n|)$. Wtedy $|x_n^2 y_n| \le c_n^3$, a z kolei $x_n^2+y_n^2 \ge c_n^2$. Można więc napisać oszacowanie:

$$0 \le |f(x_n,y_n)| \le \frac{c_n^3}{c_n^2} = c_n \to 0\,,$$

które pokazuje, że na każdym ciągu zbieżnym do punktu $(0,0)$ ciąg $f_n = f(x_n,y_n)$ zbiega do zera. Granicą funkcji w tym punkcie jest więc zero i taką trzeba nadać wartość funkcji w tym punkcie, jeśli chcemy, by była ona w nim ciągła.

W tym przykładzie sztuczka z przejściem do ciągów zapisanych w zmiennych biegunowych (w $\mathbb{R}^3$ można użyć zmiennych cylindrycznych, albo sferycznych, albo jakichś jeszcze innych) daje wynik natychmiast bo

$$f(r_n \cos\varphi_n,\ r_n \sin\varphi_n) = r_n \sin\varphi_n \cos^2\varphi_n \to 0\,,$$

a dowolny ciąg (x_n,y_n) punktów $\mathbb{R}^2$ można zapisać jako ciąg $(r_n\cos\varphi_n,\ r_n\sin\varphi_n)$ i jedynym warunkiem jego zbiegania do $(0,0)$ jest $r_n \to 0$.

Rozpatrzmy jeszcze funkcję $\mathbb{R}^2 \to \mathbb{R}$ poza punktem $(0,0)$ zadaną wzorem

$$f(x,y) = \frac{x^2\,y}{x^4+y^2}\,,$$

i znów zobaczmy, czy ma ona jakąś granicę w $(0,0)$. Nie ma: jeśli weźmiemy np. $x_n \to 0$ i $y_n = ax_n \to 0$ (co obejmuje zbieganie do punktu $(0,0)$ wzdłuż każdej prostej z wyjątkiem osi y) to

$$f(x_n, y_n) = f(x_n, ax_n) = \frac{ax_n^3}{x_n^2(x_n^2 + a^2)} = \frac{ax_n}{(x_n^2 + a^2)} \to 0\,.$$

Jeśli jednak weźmiemy $x_n \to 0$ i $y_n = x_n^2 \to 0$, co odpowiada zbieganiu do punktu $(0,0)$ po paraboli (punkty takiego ciagu leżą w płaszczyźnie $\mathbb{R}^2$ na paraboli $y = x^2$) to wtedy

$$f(x_n, y_n) = f(x_n, x_n^2) = \frac{x_n^4}{x_n^4 + x_n^4} \to \frac{1}{2}\,.$$

Na tym ciągu wartość funkcji ma inną granicę! Oznacza to, (ponieważ kryterium Heinego nie jest spełnione), że badana funkcja nie ma granicy w punkcie $(0,0)$. Zauważmy też, iż przykład ten jest bardzo podobny do przykładu drugiego: gdyby podstawić $z = x^2$ to funkcja $\tilde{f}(z,y) = f(\sqrt{z},\, y)$ by była (przynajmniej na $\mathbb{R} \times \mathbb{R}_+$) tą samą funkcją, co rozpatrywana w tamtym przykładzie.

Weźmy jeszcze funkcję zdefiniowaną wzorami

$$f(x,y) = \frac{x\,y}{x-y} \qquad \text{gdy } \ x \neq y\,,$$

i $f(x,x) = 0$ i zapytajmy, czy funkcja ta może być ciągła w punkcie $(0,0)$ oraz w innych punktach postaci (x_0, x_0). Gdy weźmiemy ciąg (x_n, y_n) zbieżny do punktu $(0,0)$, np. z $x_n = 1/n$ i $y_n = -1/n$, to

$$f(x_n, y_n) = -1/2n \to 0\,.$$

Ale jeśli wziąć

$$x_n = \frac{1}{n} + \frac{1}{n^3}\,, \qquad y_n = \frac{1}{n}\,,$$

to

$$f(x_n, y_n) = n^3 \left(\frac{1}{n} + \frac{1}{n^3}\right) \frac{1}{n} = n + \frac{1}{n} \to \infty\,.$$

Zatem niema szansy, by funkcja ta mogła być ciagła w $(0,0)$, bo nie ma ona w tym pukcie granicy. W pozostałych punktach typu (x_0, x_0) jest podobnie: wystarczy wziąć np. $x_n = x_0 + 1/n$, $y_n = x_0 - 1/n$, by zobaczyć, że

$$f(x_n, y_n) = \frac{n}{2}\left(x_0^2 - \frac{1}{n^2}\right) = \frac{n}{2}\,x_0^2 - \frac{1}{2n} \to \infty\,.$$

Inny jeszcze (czasem spotykany) sposób spojrzenia na granice funkcji np. w $(0,0)$ to tzw. granice iterowane: porównujemy wynik granic branych w odwrotnej kolejności:

$$\lim_{x\to 0}\left(\lim_{y\to 0} f(x,y)\right), \qquad \text{i} \qquad \lim_{y\to 0}\left(\lim_{x\to 0} f(x,y)\right).$$

Zastosujmy go tu do zbadania granicy w $(0,0)$ funkcji

$$f(x,y) = \frac{x-y+x^2+y^2}{x+y},$$

Mamy wtedy

$$\lim_{x\to 0}\left(\lim_{y\to 0} f(x,y)\right) = \lim_{x\to 0}\left(\frac{x+x^2}{x}\right) = \lim_{x\to 0}(1+x) = 1,$$
$$\lim_{y\to 0}\left(\lim_{x\to 0} f(x,y)\right) = \lim_{y\to 0}\left(\frac{-y+y^2}{y}\right) = \lim_{y\to 0}(-1+y) = -1.$$

Ponieważ otrzymuje się w ten sposób dwa różne wyniki, wystarcza to do stwierdzenia, ża badana funkcja w $(0,0)$ granicy mieć nie może. Sposób ten jest jest jednak mało ogólny, bo odpowiada tylko (w przypadku $\lim_{x\to 0}(\lim_{y\to 0} f)$) zejściu na oś x (w dowolnym punkcie $x \neq 0$) a następnie zbiegnięciu wzdłuż osi x do zera. Jest więc to to samo, co badanie $f(x,y)$ na ciągach $(x_n, 0)$ z $x_n \to 0$. I nawet gdyby się okazało, że obie granice iterowane są takie same, to nie dowodziłoby to jeszcze, że granica funkcji w badanym punkcie istnieje.

I jeszcze ostatni przykład: zbadajmy istnienie granicy funkcji

$$f(x,y) = \frac{1}{x+y}\left(y + x\sin\frac{1}{x}\right),$$

w punktach $(x_0, -x_0)$. Znów wyobraźmy sobie ciąg $(x_n, y_n) = (x_0 + a_n, \ -x_0 + b_n)$, taki, że $a_n \to 0$ i $b_n \to 0$ (ale $(a_n, b_n) \neq (0,0)$). Wtedy

$$\begin{aligned} f(x_n, y_n) &= \frac{1}{a_n + b_n}\left\{-x_0 + b_n + (x_0 + a_n)\sin\frac{1}{x_0 + a_n}\right\} \\ &= \frac{1}{a_n + b_n}\left\{-x_0 + b_n + (x_0 + a_n)\sin\left[\frac{1}{x_0} - \frac{a_n}{x_0^2} + \dots\right]\right\}. \end{aligned}$$

W drugim kroku rozwinęliśmy argument sinusa wokół $x = x_0$ (ten krok nie przejdzie, gdy $x_0 = 0$, więc ten przypadek trzeba będzie zbadać osobno). Widać teraz, że naogół skończona granica nie istnieje (dostaje się ∞) bo licznik pozostaje niezerowy, podczas gdy mianownik dąży do zera. Ale w punktach postaci $1/x_0 = \pi/2 + 2\pi k$ sytuacja wygląda inaczej: wtedy

$$\sin\left[\frac{1}{x_0} - \frac{a_n}{x_0^2} + \dots\right] = \cos\left[-\frac{a_n}{x_0^2} + \dots\right] \to 1,$$

i licznik, gdy $n \to \infty$ zachowuje się jak $-x_0 + b_n + (x_0 + a_n)[1 + \mathcal{O}(a_n^2)] \to a_n + b_n$ i całe wyrażenie dąży do 1. Zatem w punktach $x_0 = (\pi/2 + 2\pi k)^{-1}$. które zagęszczają się koło $x = 0$ (pozostając jednak zawsze punktami izolowanymi) funkcja ma granice równe 1. Pozostaje zbadać punkt $(0, 0)$. Napiszmy $x_n = r_n \cos\varphi_n$, $y_n = r_n \sin\varphi_n$, gdzie $r_n \to 0$, ale φ_n jest zupełnie dowolnym ciągiem (który nie musi zbiegać do niczego, gdy $n \to \infty$). Wtedy

$$f(x_n, y_n) = \frac{1}{\cos\varphi_n + \sin\varphi_n} \left\{ \sin\varphi_n + \cos\varphi_n \sin\left(\frac{1}{r_n \cos\varphi_n} \right) \right\},$$

i ponieważ argument ostatniego sinusa dąży do nieskończoności (bo $r_n \to 0$, a $|\cos\varphi_n| \leq 1$), sam sinus szaleje i całe wyrażenie do niczego konkretnego nie zbiega: np. jeśli $\varphi_n = 2\pi n$ (może być taki ciąg) to $f(x_n, y_n) = \sin(1/r_n)$, co nie zbiega do niczego.

Na koniec zauważmy, że jakkolwiek nie jest w ogólności prawdą, że przy odwzorowaniu ciągłym obrazem zbioru otwartego jest zbiór otwarty (kontrprzykład to np. funkcja $f(x) = \sin x$ określona na otwartym zbiorze $(-3\pi/4,\, 3\pi/4)$ - obrazem jest zbiór $[-1, 1]$, który jest domknięty), to prawdą jest, że przy ciągłym odwzorowaniu przeciwobrazem zbioru otwartego jest zbiór otwarty.

Różne pochodne funkcji wielu zmiennych na $\mathbb{R}^n$

Niech punktami $\mathbb{R}^n$ będą $(x_1, \ldots, x_n)$. *Pochodną cząstkową* funkcji $f: \mathbb{R}^n \to \mathbb{R}$ w punkcie $(x_1^{(0)}, \ldots, x_n^{(0)})$ nazywa się granicę (jeśli granica ta istnieje) ilorazu różnicowego

$$\left.\frac{\partial f}{\partial x_k}\right|_{(x_1^{(0)},\ldots,x_n^{(0)})} = \lim_{h\to 0} \frac{f(x_1^{(0)}, \ldots, x_k^{(0)}+h, \ldots, x_n^{(0)}) - f(x_1^{(0)}, \ldots, x_k^{(0)}, \ldots, x_n^{(0)})}{h}\,.$$

Z powyższego wzoru widać, że pochodną cząstkową w punkcie $(x_1^{(0)}, \ldots, x_n^{(0)})$ można obliczyć, tylko jeśli funkcja jest w tym punkcie określona (choćby nawet w taki sposób, że nie jest tam ciągła w sensie badanym przez nas poprzednio). Z praktycznego punktu widzenia obliczanie pochodnej cząstkowej $\partial f/\partial x_k$, czasem oznaczanej też f_{x_k}, sprowadza się do potraktowania wszystkich zmiennych oprócz x_k jak stałych i obliczenia pochodnej po x_k zgodnie ze zwykłymi regułami obliczania pochodnych znanymi z operowania funkcjami jednej zmiennej. Pochodna taka zdaje sprawę z tego, jak funkcja zmienia się wzdłuż jednej konkretnej prostej (wzdłuż osi x_k) przechodzącej przez punkt $(x_1^{(0)}, \ldots, x_n^{(0)})$ i prosta ta jest związana z wyborem zmiennych $x_1, \ldots, x_n$. Ponieważ dla fizyka oczywiste jest, że zawsze można np. obrócić osie układu współrzędnych pokrywającego przestrzeń $\mathbb{R}^n$, jest też jasne, że musi istnieć naturalne uogólnienie pochodnych cząstkowych. Są nimi *pochodne kierunkowe*, które od razu tu zdefiniujemy (no bo skoro to naturalne uogólnienie...), które zdają sprawę ze zmienności funkcji wzdłuż dowolnego kierunku, niekoniecznie równoległego do którejś z osi układu $x_1, \ldots, x_n$. Aby pochodną kierunkową zdefiniować trzeba wprowadzić w $\mathbb{R}^n$ jednostkowy wektor

$$\mathbf{n} = (n_1, n_2, \ldots, n_n)\,, \qquad n_1^2 + n_2^2 + \ldots + n_n^2 = 1\,.$$

Pochodna kierunkowa, oznaczana $\nabla_{\mathbf{n}} f$, jest wtedy dana wzorem

$$(\nabla_{\mathbf{n}} f)_{(x_1^{(0)},\ldots,x_n^{(0)})} = \lim_{h\to 0} \frac{f(x_1^{(0)} + h\,n_1, \ldots, x_n^{(0)} + h\,n_n) - f(x_1^{(0)}, \ldots, x_k^{(0)}, \ldots, x_n^{(0)})}{h}\,.$$

Zdaje ona sprawę ze zmienności funkcji wzdłuż prostej mającej kierunek wektora $\mathbf{n}$ i przechodzącej przez punkt $(x_1^{(0)}, \ldots, x_n^{(0)})$. Widać, że gdy tylko jedna składowa $n_k = 1$, a pozostałe zero, jest to to samo, co zdefiniowana wyżej pochodna cząstkowa $(\partial f/\partial x_k)_{(x_1^{(0)},\ldots,x_n^{(0)})}$.

Ważne jest by rozumieć, że pochodna cząstkowa nie jest tym samym, co prawdziwa pochodna (nastajaszczaja proizwodnaja) funkcji w danym punkcie. Tę prawdziwą pochodną zdefiniujemy dalej. W przypadku funkcji $\mathbb{R} \to \mathbb{R}$, którymi zajmowaliśmy się w poprzednim semestrze,[9] istnienie pochodnej w danym punkcie x_0 automatycznie oznaczało, że funkcja jest w tym punkcie ciągła. W przypadku pochodnych cząstkowych (lub kierunkowych) funkcji $\mathbb{R}^n \to \mathbb{R}$ ($n > 1$) już tak nie jest: Weźmy np. funkcję zdefiniowaną wzorem

$$f(x,y) = \begin{cases} xy/(x^2+y^2) & \text{gdy } (x,y) \neq (0,0) \\ 0 & \text{gdy } (x,y) = (0,0) \end{cases}.$$

[9] Taka figura retoryczna - ja się nie zajmowałem, bo materiał pierwszego semestru nie jest dla mnie tak zabawny, jak drugiego.

Funkcja ta nie ma granicy w punkcie $(0,0)$, co łatwo sprawdzić (to już przećwiczyliśmy), ale żeby można było pytać o jej pochodne cząstkowe w tym punkcie ma ona w nim przypisaną wartość równą 0. Pochodne cząstkowe tej funkcji w punkcie $(0,0)$ istnieją: obliczamy bowiem z definicji

$$\left.\frac{\partial f}{\partial x}\right|_{(0,0)} = \lim_{h\to 0}\frac{f(0+h,0)-f(0,0)}{h} = \lim_{h\to 0}\frac{0-0}{h} = 0\,.$$

Dokładnie tak samo (bo funkcja jest symetryczna wzgledem zamiany $x \leftrightarrow y$) oblicza się pochodną $(\partial f/\partial y)_{(0,0)}$. Czyli obie pochodne cząstkowe tej funkcji w punkcie $(0,0)$ istnieją, choć funkcja nie jest w tym punkcie, jako funkcja dwu zmiennych, ciągła. Żeby zrozumieć dlaczego tak jest (to jest to, czego matematycy na ogół nie dopowiadają, zostawiając publikę z rozdziawionymi gębami) wystarczy zobaczyć, że obie granice tej funkcji brane "po osiach" istnieją i są właśnie równe zeru: gdy sprawdzamy granicę funkcji "po osi x", to badamy zachowanie $f(x_n,y_n)$ na ciągach postaci $(x_n,0)$ gdzie $x_n\to 0$ (ale zawsze $x_n\neq 0$), a ponieważ $f(x_n,0)\equiv 0$ granica ta jest równa zeru, czyli wzdłuż osi x (i tak samo wzdłuż osi y) funkcja jest ciągła (jest ciągła jako funkcja jednej zmiennej). To dlatego jej pochodne cząstkowe $(\partial f/\partial x)_{(0,0)}$ i $(\partial f/\partial y)_{(0,0)}$ istnieją. Ale łatwo też zobaczyć, że skoro przypisaliśmy funkcji wartość 0 w punkcie $(0,0)$, nie jest ona ciągła, gdy zbiegamy do $(0,0)$ z innego kierunku niż po którejś z dwu osi. Np. gdybyśmy chcieli obliczyć pochodną kierunkową w $(0,0)$ "po diagonali" (wzdłuż kierunku $x=y$), tzn. pochodną kierunkową odpowiadającą wzięciu $n_1=n_2=1/\sqrt{2}$, to ta pochodna nie będzie istnieć:

$$\begin{aligned}\nabla_{\mathbf{n}} f|_{(0,0)} &= \lim_{h\to 0}\frac{f(0+n_1h,\,0+n_2h)-f(0,0)}{h}\\ &= \lim_{h\to 0}\left\{\frac{1}{h}\left[\frac{(h/\sqrt{2})(h/\sqrt{2})}{\frac{1}{2}h^2+\frac{1}{2}h^2}-0\right]\right\} = \lim_{h\to 0}\frac{1}{h}\left[\frac{1}{2}-0\right]=\infty\,.\end{aligned}$$

Pochodna w tym kierunku istniałaby, gdybyśmy nadali funkcji w punkcie $(0,0)$ wartość $1/2$ (taką, jaka wynika z granicy branej w tym kierunku). Ale wtedy oczywiście w $(0,0)$ nie istniałyby pochodne $\partial f/\partial x$ i $\partial f/\partial y$. Mimo, że teraz pewnie zaczynamy lepiej rozumieć, jak to działa, to dobrze jest wiedzieć też, że nawet istnienie pochodnych kierunkowych w danym punkcie $(x_1^{(0)},\dots,x_n^{(0)})$ we wszystkich możliwych kierunkach nie jest jeszcze wystarczające, by funkcja miała w tym punkcie prawdziwą (tj, w silnym sensie) pochodną. Wynika to z tego, że w danym punkcie funkcja może być ciągła gdy do tego punktu zbiegamy po dowolnej prostej, ale nie jest ciągła, gdy do $(x_1^{(0)},\dots,x_n^{(0)})$ zbiegamy np. po paraboli, czy jakiejś spirali. Rozpatrzmy przykład funkcji (już taka była)

$$f(x,y)=\begin{cases} x^2y/(x^4+y^2) & \text{gdy } (x,y)\neq(0,0)\\ 0 & \text{gdy } (x,y)=(0,0)\end{cases}.$$

Wiemy, że taka funkcja w punkcie $(0,0)$ jest nieciągła (choć jest ciągła wzdłuż dowolnej prostej przechodzącej przez punkt $(0,0)$ - to też już ustaliliśmy). Spróbujmy jednak obliczyć jej pochodne kierunkowe w tym punkcie. Żeby od razu obliczać pochodne w dowolnym kierunku, weźmy wektor $\mathbf{n}$ o składowych $n_1 = \cos\theta \equiv c$ i $n_2 = \sin\theta \equiv s$.

$$\begin{aligned}\nabla_{\mathbf{n}} f|_{(0,0)} &= \lim_{h\to 0} \frac{f(0+hc,\,0+hs) - f(0,0)}{h} \\ &= \lim_{h\to 0} \left\{ \frac{1}{h} \left[\frac{c^2 s h^3}{c^4 h^4 + s^2 h^2} - 0 \right] \right\} = \lim_{h\to 0} \frac{c^2 s}{c^4 h^2 + s^2} = \frac{c^2}{s}\,.\end{aligned}$$

Pochodna jak widać istnieje zawsze, pozornie z wyjątkiem sytuacji, gdy $s = 0$, czyli gdy jest to pochodna wzdłuż osi x. Ale to tylko pozór, bo pochodną kierunkową w kierunku osi x, czyli po prostu pochodną cząstkową $\partial f/\partial x$ w punkcie $(0,0)$ możemy osobno obliczyć bezpośrednio z definicji

$$\left.\frac{\partial f}{\partial x}\right|_{(0,0)} = \lim_{h\to 0} \frac{f(0+h,\,0) - f(0,0)}{h} = \lim_{h\to 0} \frac{0-0}{h} = 0\,.$$

Pochodna ta jak najbardziej istnieje, a to, że musieliśmy ją obliczyć osobno pokazuje tylko, że pochodna kierunkowa w tym punkcie jest, jako funkcja kierunku, nieciągła (ktoś powiedział, że musi być?). Można też sprawdzić, co polecam jako zadanie domowe (samo obliczenie podanych tu pochodnych też należe potraktować jak ćwiczenie!), że pochodne cząstkowe potraktowane jak funkcje na $\mathbb{R}^2$

$$f_x(x,y) = \frac{2xy^3 - 2x^5 y}{(x^4+y^2)^2}\,, \qquad f_y(x,y) = \frac{x^6 - x^2 y^2}{(x^4+y^2)^2}\,,$$

nie mają w punkcie $(0,0)$ granic, czyli nie są one, jako funkcje na $\mathbb{R}^2$, w tym punkcie ciągłe.

Jeśli jednak wszystko idzie gładko, tzn. pochodne cząstkowe w punkcie $(x_1^{(0)},\ldots,x_n^{(0)})$ istnieją i wszystkie $f_{x_i}(x_1,\ldots,x_n)$ traktowane jak funkcje na $\mathbb{R}^n$ są w tym punkcie ciągłe, to pochodna kierunkowa $\nabla_{\mathbf{n}} f|_{(x_1^{(0)},\ldots,x_n^{(0)})}$ obliczona z definicji jest równa kombinacji

$$\nabla_{\mathbf{n}} f|_{(x_1^{(0)},\ldots,x_n^{(0)})} = n_1 \left.\frac{\partial f}{\partial x_1}\right|_{(x_1^{(0)},\ldots,x_n^{(0)})} + \ldots + n_n \left.\frac{\partial f}{\partial x_n}\right|_{(x_1^{(0)},\ldots,x_n^{(0)})}\,,$$

pochodnych cząstkowych funkcji f w punkcie $(x_1^{(0)},\ldots,x_n^{(0)})$.

Pochodne cząstkowe funkcji $\mathbb{R}^n \to \mathbb{R}$ wyższych rzędów

Jeśli pochodne cząstkowe pierwszego rzędu $\partial f/\partial x_k$ funkcji $\mathbb{R}^n \to \mathbb{R}$ w jakimś otoczeniu (tzn. zbiorze otwartym) punktu $(x_1^{(0)},\ldots,x_n^{(0)})$ istnieją, to każda z nich jest w tym otoczeniu pewną nową funkcją $\mathbb{R}^n \to \mathbb{R}$ i można obliczać pochodne cząstkowe tej nowej funkcji, które będą pochodnymi cząstkowymi drugiego rzędu wyjściowej funkcji f. Potem przy tych samych warunkach można obliczać pochodne cząstkowe pochodnych cząstkowych drugiego rzędu tworząc pochodne cząstkowe rzędu trzeciego itd. Pochodne te oznaczamy

$$\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)(x,y) \equiv \frac{\partial^2 f}{\partial y\,\partial x} \equiv f_{xy}(x,y)\,, \qquad \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right)(x,y) \equiv \frac{\partial^2 f}{\partial x^2} \equiv f_{xx}(x,y)\,,$$

etc. Zachodzi pytanie, czy pochodne

$$f_{xy} \equiv \frac{\partial^2 f}{\partial y\,\partial x}\,, \qquad \text{i} \qquad f_{yx} \equiv \frac{\partial^2 f}{\partial x\,\partial y}\,,$$

obliczone w punkcie $(x_1^{(0)}, \dots, x_n^{(0)})$ są sobie równe? Odpowiedź jest taka: jeśli f_{xy} i f_{yx} jako funkcje x i y (przenosi się to oczywiście na więcej zmiennych) istnieją i są ciągłe w $(x_1^{(0)}, \dots, x_n^{(0)})$, to $f_{xy}(x^{(0)}, y^{(0)}) = f_{yx}(x^{(0)}, y^{(0)})$. Matematycy, z natury trochę prestidigitatorzy, lubią możliwą nierówność tych pochodnych demonstrować na (kanonicznym w ich języku) przykładzie funkcji $\mathbb{R}^2 \to \mathbb{R}$ zadanej wzorem

$$f(x,y) = \begin{cases} xy(x^2-y^2)/(x^2+y^2) & \text{gdy } (x,y) \neq (0,0) \\ 0 & \text{gdy } (x,y) = (0,0) \end{cases}.$$

Funkcja ta jest uczciwie ciągła w $(0,0)$. Następnie obliczają oni pochodne cząstkowe tak:

$$\left.\frac{\partial f}{\partial x}\right|_{(0,y)} = \lim_{h\to 0} \frac{f(0+h,y)-f(0,y)}{h} = \lim_{h\to 0} y\,\frac{h^2-y^2}{h^2+y^2} = -y\,,$$
$$\left.\frac{\partial f}{\partial y}\right|_{(x,0)} = \lim_{h\to 0} \frac{f(x,0+h)-f(x,0)}{h} = \lim_{h\to 0} x\,\frac{x^2-h^2}{x^2+h^2} = x\,,$$

Po czym radośnie pokazują, że w punkcie $(0,0)$

$$\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = -1\,, \qquad \text{a} \qquad \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = 1\,.$$

Po takim numerze dalej nie do końca jednak wiadomo, jakie fiku-miku doprowadziło do tego szokującego rezultatu. Otóż, aby to zrozumieć, najlepiej obliczyć obie drugie mieszane pochodne w dowolnym punkcie. Po dłuższych przekształceniach (zalecam wprawienie się na tym przykładzie!) otrzymuje się

$$\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{x^6+9x^4y^2-9x^2y^4-y^6}{(x^2+y^2)^3} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right).$$

"Na wzorkach" zatem funkcje $f_{xy}(x,y)$ i $f_{yx}(x,y)$ to jest zawsze ta sama funkcja! (Czyli jak komuś wyjdą te dwie pochodne różne to nie może mówić, że tak może być!) No i teraz już wszystko jest jasne: druga mieszana pochodna jest funkcją jednorodną stopnia zero typu wielomian k-tego (tu szóstego) stopnia przez inny wielomian tego samego stopnia. Po zrobieniu paru przykładów (i Zadania 9) powinno być jasne, że takie funkcje są w punkcie $(0,0)$ zawsze nieciągłe (wyjściowa funkcja f była!) bo mają, gdy się do $(0,0)$ zbiega z różnych kierunków, różne tam granice. I ta pochodna, co wyżej dała -1, odpowiada braniu wartości $f_{xy}(x,y)$ przy zbieganiu do $(0,0)$ po osi y (bo najpierw obliczamy pochodną po x ale otrzymaną funkcję rozpatrujemy już jako funkcję y tylko, kładąc $x=0$, czyli na osi y), a ta druga, równa $+1$, odpowiada $f_{xy}(x,y)$ przy zbieganiu do $(0,0)$ po osi x. I cała tajemnica pryska!

Różniczka df funkcji f i pochodna prawdziwa

I tak dochodzimy wreszcie do prawdziwej pochodnej funkcji $f:\ \mathbb{R}^n \to \mathbb{R}$. Niech punkty $(x_1^{(0)}, \dots, x_n^{(0)})$ i $(x_1^{(0)} + dx_1 \dots, x_n^{(0)} + dx_n)$ należą do dziedziny funkcji f. Niech

$$\Delta f \equiv f(x_1^{(0)} + dx_1 \dots, x_n^{(0)} + dx_n) - f(x_1^{(0)}, \dots, x_n^{(0)}) ,$$

będzie przyrostem wartości funkcji f, gdy przesuwamy się od punktu $(x_1^{(0)}, \dots, x_n^{(0)})$ do punktu $(x_1^{(0)} + dx_1 \dots, x_n^{(0)} + dx_n)$. Fizyk pyta, jak ten przyrost Δf można przybliżyć (niezależnie od kierunku, w którym się przesunęliśmy w bok, byle infinitezymalnie). Odpowiedź matematyka jest taka, że jeśli pierwsze pochodne cząstkowe funkcji f są w otoczeniu $(x_1^{(0)}, \dots, x_n^{(0)})$ ciągłymi funkcjami, to wtedy

$$\Delta f \approx df(x_1^{(0)}, \dots, x_n^{(0)}) \equiv f_{x_1}(x_1^{(0)}, \dots, x_n^{(0)})\, dx_1 + \dots + f_{x_n}(x_1^{(0)}, \dots, x_n^{(0)})\, dx_n ,$$

tzn. to df funkcji, zdefiniowane wyrażeniem po prawej stronie, dobrze przybliża wartość Δf w tym sensie, że różnica $\Delta f - df$ zbiega do zera szybciej niż $\sqrt{(dx_1)^2 + \dots + (dx_n)^2}$. Technicznie rzecz ujmując

$$\frac{\Delta f - df}{\sqrt{(dx_1)^2 + \dots + (dx_n)^2}} \to 0 , \qquad \text{gdy} \quad \sqrt{(dx_1)^2 + \dots + (dx_n)^2} \to 0 ,$$

niezależnie od tego jak każde dx_i z osobna zbiega do zera. df, czyli właśnie *różniczka*[10] funkcji f w punkcie $(x_1^{(0)}, \dots, x_n^{(0)})$, to jest *wartość pochodnej* prawdziwej funkcji f w tymże punkcie na wektorze przyrostu $(dx_1, \dots, dx_n)$. Różniczka jest zawsze (wbijmy więc sobie do głowy to hasło!) *główną liniową częścią przyrostu*, tu przyrostu Δf, *wartości funkcji f*. Czymże zatem jest ta pochodna prawdziwa? Jest to, najogólniej rzecz ujmując, *kowektor*, czyli odwzorowanie liniowe, które odwzorowuje wektor przyrostu w przestrzeń wartości odwzorowania; w przypadku funkcji $f : \mathbb{R}^n \to \mathbb{R}$ kowektor odwzorowujący $(dx_1, \dots, dx_n)$ w $\mathbb{R}$. (I w tym miejscu widzimy, że analiza łączy się nam z algebrą i lepiej jest wrócić do odpowiednich stron mojego skryptu z algebry i to i owo sobie przyswoić). Kowektor ten, jak każde odwzorowanie liniowe przestrzeni wektorowej,[11] jest reprezentowany macierzą

[10] Czyli takie małe, co sie rusza - tak kiedyś, gdy byłem studentem, odpowiedział na pytanie "a co to jest różniczka?" zadane przez wykładowcę fizyki mój kolega; istotnie, sposób operowania przez fizyków takimi wielkościami jak dx, df, musiał prowadzić do takiej właśnie odpowiedzi...

[11] I tu robimy taki mały myk, że nagle to $\mathbb{R}^n$, a właściwie to różnice punktów tego $\mathbb{R}^n$, zaczynamy traktować jak przestrzeń wektorową. Naprawdę jest to trochę inaczej, ale bardziej logiczne przedstawienie tego wymaga przejścia do obrazka, w którym funkcja f jest w istocie zdefiniowana na pewnej *rozmaitości* różniczkowalnej M, czyli takim tworze, który jakoś można "obmacać", a obmacanie to polega na wprowadzeniu układu współrzędnych $(x_1, \dots, x_n)$ na rozmaitości całej lub przynajmniej jakimś jej kawałku; w każdym punkcie p takiej rozmaitości identyfikowanym we wprowadzonym układzie wartościami $(x_1^{(0)}, \dots, x_n^{(0)})$ współrzędnych jest "przyczepiona" pewna przestrzeń wektorowa oznaczana T_pM (w każdym punkcie jest ona trochę inna, choć wszystkie są jakoś tam do siebie podobne, czyli izomorficzne) i to właśnie w tej przestrzeni stycznej istnieją te wektorki infinitezymalnych przemieszczeń z jednego punktu rozmaitości do drugiego; z każdym układem współrzędnych (bo współrzędne można na tej samej rozma-

(tu w kanonicznej zero-jedynkowej bazie przestrzeni $\mathbb{R}^n$ traktowanej jak wektorowa) daną wzorem

$$\left(\left.\frac{\partial f}{\partial x_1}\right|_{(x_1^{(0)},\dots,\,x_n^{(0)})},\dots,\left.\frac{\partial f}{\partial x_n}\right|_{(x_1^{(0)},\dots,\,x_n^{(0)})}\right),$$

działającą na składowe $(dx_1,\dots,dx_n)$ (w tej samej bazie) wektora przesunięcia.

Zadanie
Zbadać istnienie w punkcie $(0,0)$ pochodnej odwzorowań $\mathbb{R}^2\to\mathbb{R}$

$$a)\quad g(x,y)=\sqrt{x^4+y^4}\,,\qquad b)\quad f(x,y)=\sqrt{x^2+y^2}\,.$$

W przypadku *b*) sprawdzić także z definicji istnienie pochodnej w punkcie $(x_0,y_0)\neq(0,0)$.
Rozwiązanie: a) Obie pochodne cząstkowe funkcji $g(x,y)$ istnieją

$$g_x(x,y)=\frac{2x^3}{\sqrt{x^4+y^4}}\,,\qquad g_y(x,y)=\frac{2y^3}{\sqrt{x^4+y^4}}\,,$$

i mają w punkcie $(0,0)$ granice równe 0, co łatwo sprawdzić już poznanymi sposobami. Zatem $df=0$ w tym punkcie. Prawdziwa pochodna funkcji $g(x,y)$ w punkcie $(0,0)$ istnieje (bo pochodne cząstkowe istnieją i są tam ciągłe, jeśli nadać im wartości 0) i jej wartość na przyrostach, czyli różniczka dg, jest równa 0. Mimo to warunek "dobrego" przybliżania w tym punkcie przyrostu Δg funkcji przez jej różniczkę jest spełniony. No bo:

$$\lim\frac{\Delta g-dg}{\sqrt{(dx)^2+(dy)^2}}=\lim\frac{\sqrt{(dx)^4+(dy)^4}-0\cdot dx-0\cdot dy}{\sqrt{(dx)^2+(dy)^2}}=\lim\sqrt{\frac{(dx)^4+(dy)^4}{(dx)^2+(dy)^2}}=0\,,$$

gdzie "lim" oznacza $(dx,dy)\to0$. Wyrażenie to zbiega do zera, co najłatwiej zobaczyć pisząc $dx=dr\cos\varphi$, $dy=dr\sin\varphi$ z $dr\to0$ (kąt φ może się zmieniać dowolnie w miarę, jak dr dąży do zera).

---

itości wprowadzać różne) jest w sposób naturalny stowarzyszona pewna baza $\mathbf{i}_k$, $k=1,\dots,n$ przestrzeni stycznej, wektorek $\boldsymbol{\delta}$ przesunięcia z punktu p o współrzędnych $(x_1^{(0)},\dots,x_n^{(0)})$ do sąsiedniego p' o współrzędnych $(x_1^{(0)}+\delta x_1,\dots,x_n^{(0)}+\delta x_n)$ ma postać $\boldsymbol{\delta}=\mathbf{i}_1\delta x_1+\ldots+\mathbf{i}_n\delta x_n$, a żywy kowektor (jedno-forma) $\hat{df}$, będący pochodną prawdziwą funkcji $f:\ M\to\mathbb{R}$, jest dany przez

$$\hat{df}=\left(\frac{\partial f}{\partial x_1}\right)_{(x_1^{(0)},\dots,x_n^{(0)})}dx_1+\ldots+\left(\frac{\partial f}{\partial x_n}\right)_{(x_1^{(0)},\dots,x_n^{(0)})}dx_n\,,$$

jako kowektor zapisany w bazie dx_k jedno-form dualnej do bazy tworzonej przez wektory $\mathbf{i}_k$ (dualnej, tj. takiej, że $dx_k(\mathbf{i}_j)=\delta_{kj}$). I tak robi się z tego geometria różniczkowa, która jest bardzo naturalna i opanowanie jej prostych podstaw bardzo ułatwia życie fizyka. No ale w $\mathbb{R}^n$ wszystko się trywializuje i nic z tego piękna nie widać... Trochę tu odleciałem w kosmos, więc jak kogoś to przeraża, to niech tego przypisu nie czyta.

W przypadku funkcji $f(x, y)$, pochodne cząstkowe

$$f_x(x, y) = \frac{x}{\sqrt{x^2 + y^2}}\,, \qquad f_y(x, y) = \frac{x}{\sqrt{x^2 + y^2}}\,,$$

nie są ciągłe w punkcie $(0, 0)$, co też powinno już być oczywiste. Jeśli obliczamy je w punkcie $(0, 0)$ bezpośrednio z definicji, np.

$$f_x(0, 0) = \lim_{h\to 0} \frac{f(h, 0) - f(0, 0)}{h} = \lim_{h\to 0} \frac{|h|}{h}\,,$$

to też nie możemy im nadać wartości, bo granica $h \to 0^+$ jest równa 1, a granica $h \to 0^-$ jest równa -1. Nie można więc nawet napisać prawdziwej różniczki df w tym punkcie. Gdyby się umówić, że gdy dx i dy są dodatnie, to bierzemy $f_x(0, 0) = 1 = f_y(0, 0)$, to i tak taka "różniczka" nie przybliży dobrze przyrostu Δf: wyrażenie

$$\lim \frac{\Delta f - df}{\sqrt{(dx)^2 + (dy)^2}} = \lim \frac{\sqrt{(dx)^2 + (dy)^2} - f_x dx - f_y dy}{\sqrt{(dx)^2 + (dy)^2}} = \lim \left(1 - \frac{dx + dy}{\sqrt{(dx)^2 + (dy)^2}}\right),$$

nie dąży w sposób bezwarunkowy do zera, co znów widać pisząc $dx = dr \cos\varphi$, $dy = dr \sin\varphi$ z $dr \to 0$: ponieważ kąt φ może się zmieniać dowolnie w miarę, jak dr dąży do zera, drugi człon w nawiasie jest równy $\cos\varphi + \sin\varphi$, co nie musi być jedynką.

W punkcie $(x_0, y_0) \neq (0, 0)$ obie pochodne cząstkowe f_x i f_y są ciągłe, więc twierdzenie zapewnia istnienie w takim punkcie także prawdziwej pochodnej i tym samym i różniczki df. Warto jednak zobaczyć, że różniczka ta "dobrze" przybliża przyrost funkcji. Zbadajmy więc wyrażenie $(\Delta f - df)/\sqrt{(dx)^2 + (dy)^2}$. Jest ono równe

$$\frac{1}{\sqrt{(dx)^2 + (dy)^2}} \left[\sqrt{(x_0 + dx)^2 + (y_0 + dy)^2} - \sqrt{x_0^2 + y_0^2} - \frac{x_0 dx + y_0 dy}{\sqrt{x_0^2 + y_0^2}}\right].$$

Korzystając ze starej sztuczki $\sqrt{a} - \sqrt{b} = (a - b)/(\sqrt{a} + \sqrt{b}\,)$ przepisujemy to wyrażenie w postaci

$$\frac{1}{\sqrt{(dx)^2 + (dy)^2}} \left[\frac{(dx)^2 + (dy)^2 + 2(x_0 dx + y_0 dy)}{\sqrt{(x_0 + dx)^2 + (y_0 + dy)^2} + \sqrt{x_0^2 + y_0^2}} - \frac{x_0 dx + y_0 dy}{\sqrt{x_0^2 + y_0^2}}\right].$$

Człon $(dx)^2 + (dy)^2$ na pierwszej kresce ułamkowej w nawiasie kwadratowym nawet po podzieleniu przez $\sqrt{(dx)^2 + (dy)^2}$ znika, gdy $\sqrt{(dx)^2 + (dy)^2} \to 0$, więc możemy go już pominąć; poza tym z mianowników w tymże nawiasie wyciągamy $\sqrt{x_0^2 + y_0^2}$ i, jako że jest to pewna stała, nie będziemy już tego pisać. W mianowiku pierwszego ułamka w nawiasie kwadratowym mamy wtedy

$$1 + \sqrt{1 + \varepsilon} = 2\left(1 + \frac{1}{4}\varepsilon + \ldots\right) = 2\left(1 - \frac{1}{4}\varepsilon + \ldots\right)^{-1},$$

$$\varepsilon = \frac{2(x_0 dx + y_0 dy) + (dx)^2 + (dy)^2}{x_0^2 + y_0^2}.$$

Wielokropek oznacza wyższe potęgi ε. Pozostaje zobaczyć, że

$$\frac{x_0 dx + y_0 dy}{\sqrt{(dx)^2 + (dy)^2}} \left[-\frac{1}{4}\varepsilon + \ldots \right],$$

zbiega do zera, gdy $\sqrt{(dx)^2 + (dy)^2} \to 0$. Znów jednak człon $(dx)^2 + (dy)^2$ w ε (i w wyższych potęgach ε) znika w tej granicy, nawet po podzieleniu przez $\sqrt{(dx)^2 + (dy)^2}$, a

$$\frac{(x_0 dx + y_0 dy)^2}{\sqrt{(dx)^2 + (dy)^2}},$$

też dąży do zera (to samo będzie się działo i w wyższych potęgach ε), gdy $\sqrt{(dx)^2 + (dy)^2} \to 0$, co znów najłatwiej zobaczyć pisząc $dx = dr \cos\varphi$, $dy = dr \sin\varphi$ z $dr \to 0$. Zatem w punktach $(x_0, y_0) \neq (0, 0)$ różniczka df należycie przybliża przyrost Δf funkcji.

Na koniec zajmijmy się jeszcze pochodnymi odwzorowań F przestrzeni $\mathbb{R}^n$ w przestrzeń $\mathbb{R}^m$. Każde takie odwzorowanie, to po prostu m funkcji $\mathbb{R}^n \to \mathbb{R}$, które sobie oznaczymy $F_1(x_1, \ldots, x_n)$, ..., $F_m(x_1, \ldots, x_n)$. Innymi słowy, obrazem punktu $(x_1, \ldots, x_n) \in \mathbb{R}^n$ jest m-wymiarowy wektor $(F_1, \ldots, F_m) \in \mathbb{R}^m$. Przyrost ΔF takiej funkcji, który też jest m-wymiarowym wektorem, przy przejściu od punktu $(x_1^{(0)}, \ldots, x_n^{(0)})$ do $(x_1^{(0)} + dx_1, \ldots, x_n^{(0)} + dx_n)$ jest równy

$$\Delta F = \begin{pmatrix} F_1(x_1^{(0)} + dx_1, \ldots, x_n^{(0)} + dx_n) - F_1(x_1^{(0)}, \ldots, x_n^{(0)}) \\ F_2(x_1^{(0)} + dx_1, \ldots, x_n^{(0)} + dx_n) - F_2(x_1^{(0)}, \ldots, x_n^{(0)}) \\ \ldots\ldots\ldots \\ F_m(x_1^{(0)} + dx_1, \ldots, x_n^{(0)} + dx_n) - F_m(x_1^{(0)}, \ldots, x_n^{(0)}) \end{pmatrix},$$

i jest, gdy wszystkie pochodne cząstkowe wszystkich m funkcji, są w punkcie $(x_1^{(0)}, \ldots, x_n^{(0)})$ ciągłe, dobrze przybliżany przez różniczkę

$$dF = \begin{pmatrix} \partial_{x_1} F_1 \, dx_1 + \ldots + \partial_{x_n} F_1 \, dx_n \\ \partial_{x_1} F_2 \, dx_1 + \ldots + \partial_{x_n} F_2 \, dx_n \\ \ldots\ldots\ldots \\ \partial_{x_1} F_m \, dx_1 + \ldots + \partial_{x_n} F_m \, dx_n \end{pmatrix} = \begin{pmatrix} \partial_{x_1} F_1 & \partial_{x_2} F_1 & \ldots & \partial_{x_n} F_1 \\ \partial_{x_1} F_2 & \partial_{x_2} F_2 & \ldots & \partial_{x_n} F_2 \\ \ldots & \ldots & \ldots & \ldots \\ \partial_{x_1} F_m & \partial_{x_2} F_m & \ldots & \partial_{x_n} F_m \end{pmatrix} \begin{pmatrix} dx_1 \\ dx_2 \\ \ldots \\ dx_n \end{pmatrix}.$$

gdzie wprowadziliśmy (też często używaną) notację

$$\partial_{x_i} F \equiv \frac{\partial F}{\partial x_i},$$

i gdzie wszystkie pochodne cząstkowe są obliczone w punkcie $(x_1^{(0)}, \ldots, x_n^{(0)})$. Widać tu jak na dłoni, że pochodna funkcji F w tym punkcie jest odwzorowaniem liniowym odwzorowującym n-wymiarowy wektor przyrostów w $\mathbb{R}^m$. W kanonicznej bazie zero-jedynkowej jest ona dana powyższą macierzą.

Składanie odwzorowań

Rozpatrzymy teraz problem pochodnej funkcji złożonej. Ogólna sytuacja którą mamy na myśli mówiąc o składaniu odwzorowań wygląda następująco: dane są trzy przestrzenie metryczne, którymi dla nas są zawsze $\mathbb{R}^n$-y i dwa odwzorowania, które oznaczymy sobie H i G:

$$\mathbb{R}^n \xrightarrow{H} \mathbb{R}^k \xrightarrow{G} \mathbb{R}^m .$$

Niech punktami tych przestrzeni będą

$$\begin{array}{ll} \mathbb{R}^n & (x_1, x_2, \ldots, x_n)\,, \\ \mathbb{R}^k & (y_1, y_2, \ldots, y_k)\,, \\ \mathbb{R}^m & (z_1, z_2, \ldots, z_m)\,. \end{array}$$

Przy kolejnych odwzorowaniach

$$\begin{array}{ll} y_1 = H_1(x_1, x_2, \ldots, x_n)\,, & z_1 = G_1(y_1, y_2, \ldots, y_k)\,, \\ y_2 = H_2(x_1, x_2, \ldots, x_n)\,, & z_2 = G_2(y_1, y_2, \ldots, y_k)\,, \\ \ldots\ldots\ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots\ldots\ldots \\ y_k = H_k(x_1, x_2, \ldots, x_n)\,, & z_m = G_m(y_1, y_2, \ldots, y_k)\,. \end{array}$$

Zgodnie z tym, co już wiemy o pochodnych odwzorowań, pochodna H' (zakładamy, że istnieje; czasem oznacza się ją DH, ale to się myli z różniczką, która - zgodnie z tym, czego nauczałem wyżej - jest wartością pochodnej obliczonej na wektorze przesunięcia) odwzorowania H jest macierzą $k \times n$ (k wierszy, n kolumn), tak by działając na n-wymiarowy wektor dawała wektor k-wymiarowy. Z kolei pochodna odwzorowania G (też zakładamy, że istnieje) jest macierzą $m \times k$ (m wierszy k kolumn), która działając na k-wymiarowy wektor daje wektor m-wymiarowy.

Ich złożenie $F = H \circ G$ jest odwzorowaniem z $\mathbb{R}^n$ bezpośrednio w $\mathbb{R}^m$ i wyraża się wzorami:

$$\begin{array}{c} z_1 = F_1(x_1, x_2, \ldots, x_n) = G_1(H_1(x_1, x_2, \ldots, x_n), \ldots, H_k(x_1, x_2, \ldots, x_n))\ , \\ z_2 = F_2(x_1, x_2, \ldots, x_n) = G_2(H_1(x_1, x_2, \ldots, x_n), \ldots, H_k(x_1, x_2, \ldots, x_n))\ , \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ z_m = F_m(x_1, x_2, \ldots, x_n) = G_m(H_1(x_1, x_2, \ldots, x_n), \ldots, H_k(x_1, x_2, \ldots, x_n))\ . \end{array}$$

Obliczmy teraz korzystając z powyższych wzorów pochodną po x_1 funkcji F_1. Tak jak przy obliczaniu zwykłych pochodnych stosuje się tu metoda "pochodna funkcji zewnętrznej po jej argumencie razy pochodna funkcji wewnętrznej". Zatem

$$\frac{\partial F_1}{\partial x_1} \equiv \partial_{x_1} F_1 = \partial_{y_1} G_1\, \partial_{x_1} H_1 + \ldots + \partial_{y_k} G_1\, \partial_{x_1} H_k\,.$$

Funkcja F_1 od x_1 zależy poprzez funkcję G_1, która zależy od k funkcji $H_1, \ldots, H_k$ i dopiero każda tych zależy od x_1. Uogólnijmy ten wzorek na pochodną po x_j funkcji F_i:

$$\partial_{x_j} F_i = \sum_{l=1}^{k} \partial_{y_l} G_i \, \partial_{x_j} H_l \, .$$

Jeśli teraz dobrze się przyjrzymy temu wzorkowi, to dostrzeżemy w nim wzór na mnożenie macierzy: pochodna F', zgodnie z tym, co ma ona robić powinna być macierzą $m \times n$ (m wierszy, n-kolumn; ma ona działać na n-wymiarowy wektor i dawać wektor m-wymiarowy), a powyższy wzór daje ją jako iloczyn macierzy: wzór

$$[F']^i{}_j = \sum_{l=1}^{k} [G']^i{}_l \, [H']^l{}_j \, ,$$

w którym

$$[G']^i{}_l \equiv \partial_{y_l} G_i \equiv \frac{\partial G_i}{\partial y_l} \, , \qquad [H']^l{}_j \equiv \partial_{x_j} H_l \equiv \frac{\partial H_l}{\partial x_j} \, ,$$

jest właśnie (opanowanym na algebrze) wzorem na element $[F']^i{}_j$ (tj. element stojący w i-tym wierszu i j-tej kolumnie macierzy F'. Pozostaje tylko ustalić, w jakich punktach pochodne te (czyli macierze) mają być obliczone. Jeśli pochodna F' ma być obliczona w punkcie $(x_1, \ldots, x_n)$, to w tym punkcie musi też być wzięta pochodna H'. Natomiast pochodna G', którą obliczamy mając dane funkcje $G_1(y_1, \ldots, y_k)$, ..., $G_m(y_1, \ldots, y_k)$ i różniczkując je po y_1, ..., y_k, musi być na końcu wyrażona przez $y_1 = H_1(x_1, \ldots, x_n)$, ..., $y_k = H_k(x_1, \ldots, x_n)$. Tak wiec

$$(G \circ H)'_{(x_1,\ldots,x_n)} = G'_{(H_1(x_1,\ldots,x_n),\ldots,H_k(x_1,\ldots,x_n))} \cdot H'_{(x_1,\ldots,x_n)} \, ,$$

gdzie kropka oznacza mnożenie macierzy. Ot i wszystko.

Sformułowane wyżej reguły obliczania pochodnych odwzorowań będących złożeniami innych odwzorowań sprowadzają się do znanej z teorii funkcji jednej zmiennej reguły "pochodna po jakiejś wybranej zmiennej fukcji będącej złożeniem równa się pochodnej po argumencie funkcji zewnętrznej razy pochodna po wybranej zmiennej funkcji wewnętrznej". Tyle, że teraz trzeba posumować po argumentach funkcji zewnętrznej.

Przykład
Dane są dwa odwzorowania: $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ oraz $S : \mathbb{R} \longrightarrow \mathbb{R}^2$ zadane wzorami

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} \sin x \\ \exp(x+y) \\ 1 \end{bmatrix} , \qquad S(t) = \begin{bmatrix} t \\ \ln(1+t^4) \end{bmatrix} .$$

Ich złożeniem (złożyć je można tylko w jeden sposób, mam nadzieję, że to oczywiste) $F = T \circ S$ jest odwzorowanie $F : \mathbb{R} \longrightarrow \mathbb{R}^3$, którego wzór dostajemy wstawiając do wzoru na T, zmienne x i y wyrażone przez t, tak jak dyktuje odwzorowanie S:

$$F(t) = \begin{bmatrix} \sin t \\ \exp\{t + \ln(1+t^4)\} \\ 1 \end{bmatrix} = \begin{bmatrix} \sin t \\ (1+t^4)\, e^t \\ 1 \end{bmatrix}.$$

Pochodna F' odwzorowania F jest macierzą o trzech wierszach i jednej kolumnie, czyli wektorem (bo całe F jest, jak w mechanice pojedyńczej cząstki trajektorią tej cząstki w przestrzeni, a pochodna F to "po fizycznemu" po prostu prędkość tej cząstki; wartość tej pochodnej na "wektorze" przesunięcia, czyli δt, daje zmianę położenia po czasie δt):

$$F'(t) = \begin{bmatrix} \cos t \\ (1 + 4t^3 + t^4)\, e^t \\ 0 \end{bmatrix}.$$

Sprawdźmy, jak to się ma do iloczynu macierzy będących pochodnymi odwzorowań T i S. Pochodna T' odwzorowania T jest macierzą 3×2 (trzy wiersze, dwie kolumny), a pochodna S' odwzorowania S jest macierzą 2×1 (czyli dwuwymiarowym wektorem):

$$T'(x,y) = \begin{pmatrix} \cos x & 0 \\ e^{x+y} & e^{x+y} \\ 0 & 0 \end{pmatrix}, \qquad S'(t) = \begin{pmatrix} 1 \\ 4t^3/(1+t^4) \end{pmatrix}.$$

Iloczyn (macierzowy) tych pochodnych, to

$$T' \cdot S' = \begin{pmatrix} \cos x & 0 \\ e^{x+y} & e^{x+y} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 4t^3/(1+t^4) \end{pmatrix} = \begin{pmatrix} \cos x \\ [1 + 4t^3/(1+t^4)] \exp(x+y) \\ 0 \end{pmatrix}.$$

Teraz jednak przypominamy sobie, że trzeba wyrazić x oi y przez t, tak jak dyktuje odwzorowanie S: $x = t$, $y = \ln(1+t^4)$. Po zrobieniu tego znajdujemy, że $T' \cdot S'$ jest tym samym co F'.

Zadanie
Pokazać, że funkcję $f(t,x)$ spełniającą dwuwymiarowe równanie falowe

$$\frac{1}{c^2} \frac{\partial^2 f}{\partial t^2} - \frac{\partial^2 f}{\partial x^2} = 0$$

można przedstawić w postaci sumy $f(t,x) = h_L(x - ct) + h_R(x + ct)$ dwóch funkcji z których jedna, h_L, reprezentuje falę przemieszczającą się z prędkością c bez zmiany kształtu w kierunku dodatnim osi x, a druga, h_R, falę przemieszającą się bez zmiany swojego kształtu i zupełnie niezależnie od tamtej w kierunku ujemnym osi x.
Rozwiązanie: Wyobrażamy sobie, że[12]

$$f(t,x) = \tilde{f}(v(t,x),\, u(t,x)),$$

[12]Matematyk pisze $\tilde{f}$, bo jako maszynka z dwiema dziurami funkcja ta jest inną maszynką niż funkcja f; fizyk używa tej samej litery f na oznaczenie obu funkcji, bo dla niego jest to ta sama wielkość fizyczna.

gdzie , $v = x + ct$ a $u = x - ct$. W związku z tym piszemy zodnie z regułą łańcuszkową:

$$\frac{\partial f}{\partial x} = \frac{\partial \tilde{f}}{\partial v}\frac{\partial v}{\partial x} + \frac{\partial \tilde{f}}{\partial u}\frac{\partial u}{\partial x} = \frac{\partial \tilde{f}}{\partial v} + \frac{\partial \tilde{f}}{\partial u}\,,$$
$$\frac{\partial f}{\partial t} = \frac{\partial \tilde{f}}{\partial v}\frac{\partial v}{\partial t} + \frac{\partial \tilde{f}}{\partial u}\frac{\partial u}{\partial t} = c\,\frac{\partial \tilde{f}}{\partial v} - c\,\frac{\partial \tilde{f}}{\partial u}\,.$$

Drugie pochodne obliczamy podobnie (tu jest to proste, bo pochodne $(\partial u/\partial x)$, $(\partial u/\partial t)$ etc. są stałymi):

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial \tilde{f}}{\partial v} + \frac{\partial \tilde{f}}{\partial u}\right) = \frac{\partial^2 \tilde{f}}{\partial v^2}\frac{\partial v}{\partial x} + \frac{\partial^2 \tilde{f}}{\partial v \partial u}\frac{\partial u}{\partial x} + \frac{\partial^2 \tilde{f}}{\partial u \partial v}\frac{\partial v}{\partial x} + \frac{\partial^2 \tilde{f}}{\partial u^2}\frac{\partial u}{\partial x}$$
$$= \frac{\partial^2 \tilde{f}}{\partial v^2} + 2\,\frac{\partial^2 \tilde{f}}{\partial v \partial u} + \frac{\partial^2 \tilde{f}}{\partial u^2}\,.$$

Analogicznie

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial t}\left(\frac{\partial \tilde{f}}{\partial v} - \frac{\partial \tilde{f}}{\partial u}\right)c = \left(\frac{\partial^2 \tilde{f}}{\partial v^2}\frac{\partial v}{\partial t} + \frac{\partial^2 \tilde{f}}{\partial v \partial u}\frac{\partial u}{\partial t} - \frac{\partial^2 \tilde{f}}{\partial u \partial v}\frac{\partial v}{\partial t} - \frac{\partial^2 \tilde{f}}{\partial u^2}\frac{\partial u}{\partial t}\right)c$$
$$= c^2\,\frac{\partial^2 \tilde{f}}{\partial v^2} - 2\,c^2\,\frac{\partial^2 \tilde{f}}{\partial v \partial u} + c^2\,\frac{\partial^2 \tilde{f}}{\partial u^2}\,.$$

Zatem w nowych zmiennych równanie falowe przybiera postać

$$\frac{1}{c^2}\frac{\partial^2 f}{\partial t^2} - \frac{\partial^2 f}{\partial x^2} = 4\,\frac{\partial^2 \tilde{f}}{\partial v \partial u} = 4\,\frac{\partial}{\partial v}\left(\frac{\partial \tilde{f}}{\partial u}\right) = 4\,\frac{\partial}{\partial u}\left(\frac{\partial \tilde{f}}{\partial v}\right) = 0\,,$$

z której natychmiast wynika, że $(\partial \tilde{f}/\partial u)$ nie zależy od zmiennej v, a $(\partial \tilde{f}/\partial v)$ nie zależy od zmiennej u. Zatem ze scałkowania $(\partial \tilde{f}/\partial u)$ po u (lub ze scałkowania $(\partial \tilde{f}/\partial v)$ po v) otrzymujemy $\tilde{f} = h_L(u) + h_R(v)$, co jest właśnie tym, co trzeba było wykazać.

Zadanie
Przepisać równanie różniczkowe

$$\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2 - f^2 = 0\,,$$

w zmiennych u i v zadanych związkami

$$x = \frac{u}{u^2 + v^2} \equiv \frac{u}{R^2}\,, \qquad y = \frac{v}{u^2 + v^2} \equiv \frac{v}{R^2}\,.$$

Rozwiązanie: Chodzi jak zawsze o wyobrażenie sobie, że $f(x, y) = \tilde{f}(u(x, y),\, v(x, y))$. Tak jak w porzednim zadaniu (uwaga: tu zmienną u traktuję jak piewszą, a v jak drugą

- odwrotnie, niż w poprzednim)

$$\frac{\partial f}{\partial x} = \frac{\partial \tilde{f}}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial \tilde{f}}{\partial v}\frac{\partial v}{\partial x}\,,$$
$$\frac{\partial f}{\partial y} = \frac{\partial \tilde{f}}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial \tilde{f}}{\partial v}\frac{\partial v}{\partial y}\,.$$

Tu mamy jednak pewną techniczną trudność, bo aby np. zróżniczkować u po x, musimy "odkręcić" podane wzory (wyrażają one x i y przez u i v, a potrzebujemy na odwrót) a następnie, kiedy już znajdziemy pochodne $(\partial u/\partial x)$ etc., to trzeba będzie je zurück wyrazić przez zmienne u i v (żeby równanie na $\tilde{f}$ było wyrażone w tych tylko zmiennych). Wszystko to się oczywiście daje zrobić, szczególnie w tym zadaniu, bo wzory wiążące x i y z u i v są tu proste. Niemniej przydatna może być następująca sztuczka. Niech $F : (u,v) \longrightarrow (x,y)$ (czyli F to jest to odwzorowanie $\mathbb{R}^2$ w $\mathbb{R}^2$, które wyraża x i y przez u i v), a $G : (x,y) \longrightarrow (u,v)$ (czyli G to jest F^{-1}, czyli odwzorowanie odwrotne do F wyrażające u i v przez x i y). Ponieważ $G \circ F =$id - złożenie odwzorowania i odwzorowanie odwrotnego jest odwzorowaniem identycznościowym, więc pochodna $(G \circ F)'$ jest, no właśnie: czym jest? Oczywiście macierzą jednostkową I, tu 2×2 (nie zerem, jak by ktoś mógł mniemać! W przypadku jednowymiarowym $u = G(x)$, a $x = F(u)$, więc $(F \circ G)(x) = x$ i pochodna jest równa 1; teraz uogólnijmy to sobie...). Czyli, ponieważ $(G \circ F)' = G' \cdot F' = I$, więc $G' = (F')^{-1}$ - macierz G' jest macierzą odwrotną do macierzy F' (po to m.in. się uczyliśmy odwracać macierze). Ponieważ

$$F' = \begin{pmatrix} \partial x/\partial u & \partial x/\partial v \\ \partial y/\partial u & \partial y/\partial v \end{pmatrix} = \begin{pmatrix} (v^2-u^2)/R^4 & -2uv/R^4 \\ -2uv/R^4 & (u^2-v^2)/R^4 \end{pmatrix},$$

więc (pamiętamy: odwracanie macierzy 2×2 sprowadza się do zamienienia miejscami jej elementow diagonalnych, zmienienia znaków jej elementów pozadiagonalnych i podzielenia całości przez wyznacznik odwracanej macierzy)

$$G' = \begin{pmatrix} \partial u/\partial x & \partial u/\partial y \\ \partial v/\partial x & \partial v/\partial y \end{pmatrix} = (F')^{-1} = \frac{1}{\mathrm{det}F'}\begin{pmatrix} \partial y/\partial v & -(\partial x/\partial v) \\ -(\partial y/\partial u) & \partial x/\partial u \end{pmatrix}.$$

Zauważmy, że w ten sposób dostajemy od razu pochodne $\partial u/\partial x$, etc. wyrażone przez zmienne u i v, dokładnie tak, jak tego potrzebujemy!. No to teraz tylko powypisywać wzory w naszym przypadku:

$$\mathrm{det}F' = \frac{1}{R^8}\left(-(u^2-v^2)^2 - 4u^2v^2\right) = -\frac{(u^2+v^2)^2}{R^8} = -\frac{1}{R^4}\,,$$

i

$$G' = \begin{pmatrix} \partial u/\partial x & \partial u/\partial y \\ \partial v/\partial x & \partial v/\partial y \end{pmatrix} = -R^4\begin{pmatrix} (u^2-v)^2/R^4 & 2uv/R^4 \\ 2uv/R^4 & (v^2-u^2)/R^4 \end{pmatrix}.$$

Zatem, jawnie już,

$$\frac{\partial u}{\partial x} = v^2 - u^2\,, \qquad \frac{\partial u}{\partial y} = -2uv\,, \qquad \frac{\partial v}{\partial x} = -2uv\,, \qquad \frac{\partial v}{\partial y} = u^2 - v^2\,.$$

No i teraz możemy przepisać równanie różniczkowe w zmiennych u i v:

$$
\begin{aligned}
\frac{\partial f}{\partial x} &= (v^2 - u^2)\,\frac{\partial \tilde{f}}{\partial u} - 2uv\,\frac{\partial \tilde{f}}{\partial v},\\
\frac{\partial f}{\partial y} &= -2uv\,\frac{\partial \tilde{f}}{\partial u} + (u^2 - v^2)\,\frac{\partial \tilde{f}}{\partial v}\,.
\end{aligned}
$$

Po podniesieniu do kwadratu i zebraniu do kupy (wyrazy $(\partial\tilde{f}/\partial u)(\partial\tilde{f}/\partial v)$ sie skasują) dostajemy równanie

$$
R^4(u,v)\left[\left(\frac{\partial \tilde{f}}{\partial u}\right)^2 + \left(\frac{\partial \tilde{f}}{\partial v}\right)^2\right] - f^2 = 0\,.
$$

Nic prostszego nie wyszło, ale nie o to tu chodziło.

Oczywiście wszystko można zrobić bezpośrednio też.

$$
x^2 + y^2 = \frac{u^2 + v^2}{R^4} = \frac{1}{R^2}\,,
$$

więc

$$
u = R^2 x = \frac{x}{x^2 + y^2} \equiv \frac{x}{\kappa^2}\,, \qquad v = R^2 y = \frac{y}{x^2 + y^2} \equiv \frac{y}{\kappa^2}\,,
$$

i

$$
\frac{\partial u}{\partial x} = \frac{y^2 - x^2}{\kappa^4}\,, \qquad \frac{\partial u}{\partial y} = -\frac{2xy}{\kappa^4}\,, \qquad \frac{\partial v}{\partial x} = -\frac{2xy}{\kappa^4}\,, \qquad \frac{\partial v}{\partial y} = \frac{x^2 - y^2}{\kappa^4}\,.
$$

Teraz trzeba by te wzory zurück wyrazić przez u i v, ale na nawet nie musimy tego jawnie robić, bo jak podniesiemy do kwadratu pochodne

$$
\begin{aligned}
\frac{\partial f}{\partial x} &= \frac{y^2 - x^2}{\kappa^4}\,\frac{\partial \tilde{f}}{\partial u} - \frac{2xy}{\kappa^4}\,\frac{\partial \tilde{f}}{\partial v},\\
\frac{\partial f}{\partial y} &= -\frac{2xy}{\kappa^4}\,\frac{\partial \tilde{f}}{\partial u} + \frac{x^2 - y^2}{\kappa^4}\,\frac{\partial \tilde{f}}{\partial v}\,,
\end{aligned}
$$

i dodamy do siebie te kwadraty, to dostaniemy

$$
\frac{1}{\kappa^4}\left[\left(\frac{\partial \tilde{f}}{\partial u}\right)^2 + \left(\frac{\partial \tilde{f}}{\partial v}\right)^2\right] - f^2 = 0\,,
$$

ale $\kappa^{-4} = R^4$, więc rzeczywiście wyszło to samo.

Warto poznać jeszcze jeden sposób znajdywania pochodnych $(\partial u/\partial x)$ etc. wyrażonych od razu przez zmienne u i v. Jest to w mojej terminologii sposób[13] "termodynamiczny". Zamiast jawnie "odkręcać" wzory na x i y, zapisujemy je tak

$$x = \frac{u(x,y)}{R^2(u(x,y),\, v(x,y))}\,, \qquad y = \frac{v(x,y)}{R^2(u(x,y),\, v(x,y))}\,,$$

tj. wyobrażamy sobie (ale tylko wyobrażamy, nie musimy wypisywać jawnych wzorów!), że u i v są funkcjami x i y, więc wypisane wzorki są jakby tożsamościami. Możemy je teraz zróżniczkować stronami po x (traktując x i y jak zmienne niezależne, a u i v jak ich funkcje, tak jak to wynika z zapisu), co da

$$\begin{aligned} 1 &= \frac{u_x}{R^2} - \frac{u}{R^4}\,(2uu_x + 2vv_x) = \frac{v^2-u^2}{R^4}\,u_x - \frac{2uv}{R^4}\,v_x,\\ 0 &= \frac{v_x}{R^2} - \frac{v}{R^4}\,(2uu_x + 2vv_x) = \frac{u^2-v^2}{R^4}\,v_x - \frac{2uv}{R^4}\,u_x, \end{aligned}$$

Jak widać są to dwa równania liniowe na dwie niewiadome u_x i v_x A równania liniowe też już nauczyliśmy się rozwiązywać (widzą Państwo, jak potrzebna jest algebra!). Tu akurat jest to bardzo proste, bo z drugiego mamy $v_x = 2uvu_x/(u^2 - v^2)$ i jak to wstawimy do pierwszego, to dostaniemy

$$1 = \left[\frac{v^2-u^2}{R^4} - \frac{4u^2v^2}{R^4(u^2-v^2)}\right] u_x\,,$$

skąd już łatwo znajdujemy, że $u_x = v^2 - u^2$, a potem, że $v_x = -2uv$. Analogicznie, różniczkując stronami po y wypisane wyżej tożsamości dostajemy dwa liniowe równania na u_y i v_y i rozwiązawszy je znajdujemy $v_y = u^2 - v^2$ i $u_y = -2uv$, tak jak poprzednio. Oczywiście, cała ta metoda jest tym samym, co metoda polegająca na odwracaniu macierzy F', tylko jest ona nieco inaczej sformułowana. W istocie, rozwiązywanie układów równań liniowych na pochodne u_x, v_y, etc. jest równoważne operacji odwracania macierzy (pamiętamy z algebry, że najprostszym - w moim odczuciu - sposobem znajdywania macierzy odwrotnej do danej jest potraktowanie tejże jak macierzy zmiany bazy i rozwiązanie odpowiedniego układu równań liniowych właśnie).

Zadanie

Gradient funkcji $f : \mathbb{R}^3 \longrightarrow \mathbb{R}$ rozumiany[14] "po fizycznemu" jest to wektor o składowych

$$\boldsymbol{\nabla} f = \mathbf{e}_x\, f_x + \mathbf{e}_y\, f_y + \mathbf{e}_z\, f_z\,,$$

[13] Naprawdę sposób ten opiera się na umiejętności operowania funkcjami zadanymi w sposób uwikłany (a na tym stoi połowa matematycznej - bo nie fizycznej! - części termodynamiki; przekona się o tym ten, kto będzie miał szczęście przyjść na mój wykład z termodynamiki i fizyki statystycznej). Tego będziemy się uczyć dopiero za dwa tygodnie pewnie, ale fizyk się takimi drobnostkami nie przejmuje, tylko działa!

[14] Naprawdę, to gradient nie jest wektorem, tylko kowektorem i wszystko, co tu zrobimy idzie łatwiej, gdy się to przyjmie do wiadomości. W XXI wieku dobrze by już było unowocześnić to nauczanie i wprowadzić trochę elementarnej wiedzy z zakresu geometrii różniczkowej i form różniczkowych, bo czyni to wszystko bardziej zrozumiałym, a nie jest trudniejsze od zwykłego operowania wektorami. Au contraire, wiele rzeczy się wręcz upraszcza.

gdzie $\mathbf{e}_x$, $\mathbf{e}_y$ i $\mathbf{e}_z$ są jednostkowymi wektorami (wersorami - ale trzeba pamiętać, że choć brzmi to z angielska, takiego słowa w angielskim się nie używa, podobnie jak słowa "termostat", które też nam z angielska brzmi) tworzącymi układ ortonormalny. Wyrazić gradient we współrzędnych sferycznych zdefiniowanych związkami[15]

$$\begin{aligned} x &= r\sin\theta\sin\varphi\,, \\ y &= r\sin\theta\cos\varphi\,, \\ z &= r\cos\theta\,, \end{aligned}$$

rozpisując go na wersory $\mathbf{e}_r$, $\mathbf{e}_\theta$ i $\mathbf{e}_\varphi$.

Rozwiązanie: Jak zawsze trzeba sobie wyobrazić, że

$$f(x,y,z) = \tilde{f}(r(x,y,z),\,\theta(x,y,z),\,\varphi(x,y,z))\,.$$

Wyrażamy wobec tego najpierw składowe wektora-gradientu w bazie $\mathbf{e}_x$, $\mathbf{e}_y$ i $\mathbf{e}_z$, a potem zmieniamy bazę na $\mathbf{e}_r$, $\mathbf{e}_\theta$ i $\mathbf{e}_\varphi$. Najpierw więc pierwszy punkt programu. Zaczynamy od f_x:

$$f_x = r_x\,\tilde{f}_r + \theta_x\,\tilde{f}_\theta + \varphi_x\,\tilde{f}_\varphi\,,$$

i musimy wobec tego wyznaczyć pochodne r_x, θ_x i φ_x. Można to zrobić "odkręcając" wzory definiujące układ sferyczny czyli wypisując wzorki

$$\begin{aligned} r &= \sqrt{x^2+y^2+z^2}\,, \\ \theta &= \arccos\frac{z}{\sqrt{x^2+y^2+z^2}} = \mathrm{arctg}\frac{\sqrt{x^2+y^2}}{z}\,, \\ \varphi &= \mathrm{arctg}\frac{y}{x}\,, \end{aligned}$$

różniczkując po x (a potem, przy wypisywaniu f_y po y, a przy wypisywaniu f_z po z) i na koniec wyrażając otrzymane pochodne na powrót przez r, θ i φ. Tu jednak zastosujemy mój ulubiony trick "termodynamiczny", czyli podane wzory definiujące układ sferyczny zróżniczkujemy stronami po x (a potem po y i po z) traktując r, θ i φ jak funkcje niezależnych zmiennych x, y i z. Tak czyniąc otrzymujemy układ trzech równań liniowych na r_x, θ_x i φ_x (używamy tu oznaczeń $s_\theta \equiv \sin\theta$, $c_\varphi \equiv \cos\varphi$, etc.)

$$\begin{aligned} 1 &= r_x\,s_\theta\,c_\varphi + \theta_x\,r\,c_\theta\,c_\varphi - \varphi_x\,r\,s_\theta\,s_\varphi\,, \\ 0 &= r_x\,s_\theta\,s_\varphi + \theta_x\,r\,c_\theta\,s_\varphi + \varphi_x\,r\,s_\theta\,c_\varphi\,, \\ 0 &= r_x\,c_\theta - \theta_x\,r\,s_\theta\,. \end{aligned}$$

Z ostatniego $\theta_x = (r_x/r)(c_\theta/s_\theta)$ i to do dwu pierwszych. To da

$$\begin{aligned} 1 &= r_x\left(s_\theta + \frac{c_\theta^2}{s_\theta}\right)c_\varphi - \varphi_x\,r\,s_\theta\,s_\varphi\,, \\ 0 &= r_x\left(s_\theta + \frac{c_\theta^2}{s_\theta}\right)s_\varphi + \varphi_x\,r\,s_\theta\,c_\varphi\,, \end{aligned}$$

[15]Trzeba sobie te wzorki zapamiętać i nie zamieniać sinusów z kosinusami (co czasem studenci robią), bo wprowadza to zamęt, a jest wiele ciekawszych rzeczy, na których trzeba się skupiać.

i teraz dodajemy pierwsze razy c_φ do drugiego razy s_φ i dostajemy $r_x = s_\theta c_\varphi$ (więc mamy też i θ_x), a z kolei pierwsze razy s_φ minus drugie razy c_φ da φ_x. Zatem

$$r_x = s_\theta\, c_\varphi\,, \qquad \theta_x = \frac{c_\theta\, c_\varphi}{r}\,, \qquad \varphi_x = -\frac{s_\varphi}{r\, s_\theta}\,.$$

Po zróżniczkowaniu stronami po y dostajemy układ

$$\begin{aligned} 0 &= r_y\, s_\theta\, c_\varphi + \theta_y\, r\, c_\theta\, c_\varphi - \varphi_y\, r\, s_\theta\, s_\varphi\,, \\ 1 &= r_y\, s_\theta\, s_\varphi + \theta_y\, r\, c_\theta\, s_\varphi + \varphi_y\, r\, s_\theta\, c_\varphi\,, \\ 0 &= r_y\, c_\theta - \theta_y\, r\, s_\theta\,. \end{aligned}$$

który można rozwikłać tak samo, jak poprzedni: $\theta_y = (r_y/r)(c_\theta/s_\theta)$ i

$$\begin{aligned} 0 &= r_y \left(s_\theta + \frac{c_\theta^2}{s_\theta} \right) c_\varphi - \varphi_y\, r\, s_\theta\, s_\varphi\,, \\ 1 &= r_y \left(s_\theta + \frac{c_\theta^2}{s_\theta} \right) s_\varphi + \varphi_y\, r\, s_\theta\, c_\varphi\,. \end{aligned}$$

Stąd

$$r_x = s_\theta\, s_\varphi\,, \qquad \theta_x = \frac{c_\theta\, s_\varphi}{r}\,, \qquad \varphi_x = \frac{c_\varphi}{r\, s_\theta}\,.$$

Wreszcie, zróżniczkowanie stronami po z da

$$\begin{aligned} 0 &= r_z\, s_\theta\, c_\varphi + \theta_z\, r\, c_\theta\, c_\varphi - \varphi_z\, r\, s_\theta\, s_\varphi\,, \\ 0 &= r_z\, s_\theta\, s_\varphi + \theta_z\, r\, c_\theta\, s_\varphi + \varphi_z\, r\, s_\theta\, c_\varphi\,, \\ 1 &= r_z\, c_\theta - \theta_z\, r\, s_\theta\,. \end{aligned}$$

Teraz mnożymy pierwsze przez c_φ, dodajemy do drugiego pomnożonego przez s_φ i otrzymujemy układ

$$\begin{pmatrix} s_\theta & r\, c_\theta \\ c_\theta & -r\, s_\theta \end{pmatrix} \begin{pmatrix} r_z \\ \theta_z \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

z którego łatwo dostajemy $r_z = c_\theta$ i $\theta_z = -s_\theta/r$. Podstawiajac te pochodne do $0 = r_z s_\theta c_\varphi + \theta_z r c_\theta c_\varphi - \varphi_z r s_\theta s_\varphi$ znajdujemy, że $\varphi_z = 0$. Mamy więc

$$\begin{aligned} \boldsymbol{\nabla} f = \mathbf{e}_x & \left(s_\theta\, c_\varphi\, f_r + \frac{c_\theta\, c_\varphi}{r}\, f_\theta - \frac{s_\varphi}{r\, s_\theta}\, f_\varphi \right) \\ & + \mathbf{e}_y \left(s_\theta\, s_\varphi\, f_r + \frac{c_\theta\, s_\varphi}{r}\, f_\theta + \frac{c_\varphi}{r\, s_\theta}\, f_\varphi \right) + \mathbf{e}_z \left(c_\theta\, f_r - \frac{s_\theta}{r}\, f_\theta \right)\,. \end{aligned}$$

Można teraz obliczyć kwadrat gradientu jako sumę kwadratów współczynników przy $\mathbf{e}_x$, $\mathbf{e}_y$ i $\mathbf{e}_z$. Wyjdzie

$$(\boldsymbol{\nabla} f)^2 = (f_r)^2 + \left(\frac{f_\theta}{r} \right)^2 + \left(\frac{f_\varphi}{r\, s_\theta} \right)^2\,,$$

co podpowiada, że

$$\nabla f = \mathbf{e}_r\, f_r + \mathbf{e}_\theta\, \frac{f_\theta}{r} + \mathbf{e}_\varphi\, \frac{f_\varphi}{r\, s_\theta}\,,$$

i tak w istocie jest, ale aby to ściśle pokazać, trzeba wektory $\mathbf{e}_x$, $\mathbf{e}_y$ i $\mathbf{e}_z$ wyrazić przez wektory $\mathbf{e}_r$, $\mathbf{e}_\theta$ i $\mathbf{e}_\varphi$ i wstawić do wzoru na ∇f. Wzory w odwrotną stronę dość łatwo napisać:[16] wektor jednostkowy $\mathbf{e}_r$ biegnie wzdłuż promienia, więc jego rzut na oś z idzie z kosinusem θ, a rzut na płaszczyznę xy idzie z sinusem; rzut ten należy jeszcze zrzutować na oś x i y. Z kolei wektor $\mathbf{e}_\varphi$ jest zawsze równoległy do płaszczyzny xy i jest taki sam, jak w układzie biegunowym. Wreszcie wektor $\mathbf{e}_\theta$ ma rzut na oś z równy $-s_\theta$, a jego rzut równy c_θ na płaszczyznę xy trzeba jeszcze dodatkowo zrzutować na osie x i y. Zatem

$$\begin{aligned}
\mathbf{e}_r &= \mathbf{e}_x\, s_\theta\, c_\varphi + \mathbf{e}_y\, s_\theta\, s_\varphi + \mathbf{e}_z\, c_\theta\,,\\
\mathbf{e}_\theta &= \mathbf{e}_x\, c_\theta\, c_\varphi + \mathbf{e}_y\, c_\theta\, s_\varphi - \mathbf{e}_z\, s_\theta\,,\\
\mathbf{e}_\varphi &= -\mathbf{e}_x\, s_\varphi + \mathbf{e}_y\, c_\varphi\,.
\end{aligned}$$

Żeby te wzory odwikłać najlepiej przepisać je tak jak na algebrze

$$(\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_\varphi) = (\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z) \begin{pmatrix} s_\theta\, c_\varphi & c_\theta\, c_\varphi & -s_\varphi \\ s_\theta\, s_\varphi & c_\theta\, s_\varphi & c_\varphi \\ c_\theta & -s_\theta & 0 \end{pmatrix}$$

Stojąca tu macierz zmiany bazy łączy dwa układy wektorów ortonormalnych, jest więc macierzą ortogonalną o wyznaczniku 1 (to, że 1, a nie -1 wynika z tego, że dwa te układy są zgodnie zorientowane, cokolwiek by to miało znaczyć...) i macierz do niej odwrotna jest po prostu dana przez jej transpozycję. Zatem

$$(\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z) = (\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_\varphi) \begin{pmatrix} s_\theta\, c_\varphi & s_\theta\, s_\varphi & c_\theta \\ c_\theta\, c_\varphi & c_\theta\, s_\varphi & -s_\theta \\ -s_\varphi & c_\varphi & 0 \end{pmatrix}.$$

albo

$$\begin{aligned}
\mathbf{e}_x &= \mathbf{e}_r\, s_\theta\, c_\varphi + \mathbf{e}_\theta\, c_\theta\, c_\varphi - \mathbf{e}_\varphi\ s_\varphi,\\
\mathbf{e}_y &= \mathbf{e}_r\, s_\theta\, s_\varphi + \mathbf{e}_\theta\, c_\theta\, s_\varphi + \mathbf{e}_\varphi\, c_\varphi\,,\\
\mathbf{e}_z &= \mathbf{e}_r\, c_\theta - \mathbf{e}_\theta\, s_\theta\,.
\end{aligned}$$

Oczywiście można było układ równań rozwiązać konwencjonalnie: dodając pierwsze pomnożone przez s_θ do drugiego pomnożonego przez c_θ wygaussowuje się $\mathbf{e}_z$ a potem to już jest układ dwu równań na dwie niewiadome ($\mathbf{e}_x$ i $\mathbf{e}_y$). Z kolei $\mathbf{e}_z$ dostaje się odejmując od pierwszego pomnożonego przez c_θ drugie pomnożone przez s_θ. No i jak się te wzory na $\mathbf{e}_x$, $\mathbf{e}_y$ i $\mathbf{e}_z$ wstawi do tego długiego wzoru na gradient, to wyjdzie to, co już zostało napisane wcześniej.

[16] O właśnie: elementarna wiedza z geometrii różniczkowej pozwala wzory na jednostkowe wektory związane z dowolnym krzywoliniowym układem współrzędnych wypisywać "mechanicznie", bez wyobrażania sobie, jak to naprawdę w przestrzeni wygląda...

Rozwinięcie w szereg Taylora funkcji wielu zmiennych

Jak wszyscy dobrze wiedzą (lub wiedzieć powinni), w przypadku funkcji jednej zmiennej, czyli odwzorowania z $\mathbb{R}$ w $\mathbb{R}$, mającej $n+1$ ciągłych pochodnych w całym przedziale domkniętym $[x_0,\, x_0+h]$, słuszny jest *wzór Taylora* (nie szereg, tylko wzór!)

$$\phi(x_0+h) = \phi(x_0) + \phi'(x_0)\, h + \frac{1}{2!}\, \phi''(x_0)\, h^2 + \ldots + \frac{1}{n!}\, \phi^{(n)}(x_0)\, h^n + R_{n+1}\,.$$

Wyraz R_{n+1}, zwany resztą Taylora (a może doczesnymi resztkami Taylora), jest dany wzorem (tzw. postać Lagrange'a resztek Taylora - są też inne jej postacie)

$$R_{n+1} = \frac{h^{n+1}}{(n+1)!}\, \phi^{(n+1)}(x_0+\theta h)\,, \qquad \text{gdzie} \quad 0 < \theta < 1\,.$$

Chodzi oczywiście o to, że w przedziale $(x_0,\, x_0+h)$ jest gdzieś (ale gdzie dokładnie, to tego właśnie wzór nie mówi) taki punkt,[17] że obliczona w nim $(n+1)$-sza pochodna funkcji $\phi(x)$ dopełnia wzór Taylora do uczciwej równości. Sztandarowym przykładem, jak działa wzór Taylora jest zastosowanie go z $x_0 = 0$ do funkcji

$$\phi(x) = \exp\left(-\frac{1}{x^2}\right).$$

Funkcja ta ma w $x_0 = 0$ pochodne dowolnego rzędu, ale wszystkie one, jak jeden mąż (choć to dziewczyny!), są równe zeru: $\phi^{(n)}(0) = 0$, $n = 0, 1, \ldots$ Wobec tego wszystkie wyrazy we wzorze Taylora są równe zeru oprócz - właśnie! - tej reszty, która daje dokładnie wartość $\phi(h)$.

Dowodzi się (np. u Lejka - to naprawdę bardzo przyjazna człowiekowi książeczka!), że równoważnie resztę Taylora można też zapisać w postaci

$$R_{n+1} = \frac{h^{n+1}}{n!}\, (1-\tilde{\theta})^n \phi^{(n+1)}(x_0+\tilde{\theta} h)\,, \qquad \text{gdzie} \quad 0 < \tilde{\theta} < 1\,.$$

Czasem bywa to użyteczne.

Dopiero, gdy w przedziale domkniętym $[x_0,\, x_0+h]$ istnieją pochodne dowolnego rzędu funkcji $\phi(x)$ i gdy można pokazać, iż przy $n \to \infty$ reszta R_{n+1} dąży do zera,[18] wzór Taylora staje się nieskończonym szeregiem potęgowym. Oczywiście promień zbieżności takiego szeregu jest wtedy równy conajmniej h (no bo jak pokazaliśmy, że przy danym h reszta zbiega do zera, to tak musi być; odwracając kota do góry ogonem, jeśli napiszemy nieskończony potęgowy szereg Taylora i zobaczymy, że ma on skończony promień zbieżności r, to to znaczy, że przy $|h| > r$, lub $|h| \geq r$ - zależnie od zachowania szeregu na krańcach przedziału zbieżności - reszta Taylora do zera nie zbiega).

[17] Utarła jakaś taka tradycja matematyczna, żeby wielkość parametryzującą niepewność, gdzie obliczać tę pochodną w reszcie Taylora oznaczać θ. Nie wiem skąd to poszło, ale nie będziemy tej tradycji tu łamać.

[18] Jest jasne, że w przypadku funkcji $\phi(x) = \exp(-1/x^2)$ tego właśnie nie można pokazać!

Wzór Taylora obowiązujący w przypadku funkcj wielu zmiennych, czyli odwzorowań[19] $f : \mathbb{R}^n \longrightarrow \mathbb{R}$, można otrzymać z przypomnianego wyżej wzoru Taylora dla funkcji jednej zmiennej. W tym celu stosujemy ten wzór do funkcji

$$\phi(t) = \phi(0+t) = f(\mathbf{x} + t\mathbf{h})\,,$$

potraktowanej jak funkcja jednej zmiennej t. Przyrost $\mathbf{h} = (h_1, \dots, h_n)$ jest tu ustalonym wektorem. Zakładając, że funkcja $\phi(t)$ ma w przedziale $[0,1]$ pochodne do rzędu $(n+1)$-szego włącznie,[20] mamy

$$\begin{aligned}\phi(t) = f(\mathbf{x}) + t\sum_{j=1}^{n} \left.\frac{\partial f}{\partial x_j}\right|_{\mathbf{x}} h_j + \frac{t^2}{2!}\sum_{j_1=1}^{n}\sum_{j_2=1}^{n} \left.\frac{\partial^2 f}{\partial x_{j_1}\partial x_{j_2}}\right|_{\mathbf{x}} h_{j_1}h_{j_2} + \dots \\ + \frac{t^r}{r!}\sum_{j_1=1}^{n}\dots\sum_{j_r=1}^{n} \left.\frac{\partial^r f}{\partial x_{j_1}\dots\partial x_{j_r}}\right|_{\mathbf{x}} h_{j_1}\dots h_{j_r} + R_{r+1}\,.\end{aligned}$$

Wykorzystana tu została regułka obliczania pochodnych złożenia funkcji: np.

$$\begin{aligned}\left.\frac{d\phi(t)}{dt}\right|_{t=0} &= \left[\frac{d}{dt} f(x_1 + th_1, \dots, x_n + th_n)\right]_{t=0} \\ &= \left.\frac{\partial f(x_1,\dots,x_n)}{\partial x_1}\right|_{\mathbf{x}} h_1 + \dots + \left.\frac{\partial f(x_1,\dots,x_n)}{\partial x_n}\right|_{\mathbf{x}} h_n\,,\end{aligned}$$

etc. Można teraz położyć $t = 1$ i w ten sposób otrzymuje się wzór Taylora dla funkcji wielu zmiennych.

W występujących w k-tym wyrazie $(1 \le k \le r)$ sumach pochodnych k-tego rzędu wiele wyrazów się powtarza, bo mieszane pochodne są sobie równe. Np. w wyrazie o $k = 2$ mamy

$$\left.\frac{\partial^2 f}{\partial x_1 \partial x_2}\right|_{\mathbf{x}} h_1 h_2 = \left.\frac{\partial^2 f}{\partial x_2 \partial x_1}\right|_{\mathbf{x}} h_2 h_1\,.$$

Zawsze można taki k-ty wyraz zapisać tak, by już nie było powtarzania się takich samych pochodnych:

$$\sum_{j_1=1}^{n}\dots\sum_{j_k=1}^{n} \left.\frac{\partial^k f}{\partial x_{j_1}\dots\partial x_{j_k}}\right|_{\mathbf{x}} h_{j_1}\dots h_{j_k} = \sum_{p_1=0}^{k}\dots\sum_{p_n=0}^{k} C^{(k)}_{p_1\dots p_n} \left.\frac{\partial^k f}{\partial x_1^{p_1}\dots x_n^{p_n}}\right|_{\mathbf{x}} h_1^{p_1}\dots h_n^{p_n}\,.$$

$C^{(k)}_{p_1\dots p_n}$ jest tu pewnym czynnikiem kombinatorycznym, który jest równy zeru, jeśli $p_1 + \dots + p_n \neq k$. Jeśli wydaje się to zawiłe, to najlepiej wziąć przypadek funkcji dwóch

[19] Odwzorowania $F : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ to po prostu m osobnych funkcji $F_i : \mathbb{R}^n \longrightarrow \mathbb{R}$, $i = 1, \dots, m$; wzór Taylora stosujemy do każdej z nich osobna.

[20] To czy ma, zależy oczywiście też od wektora przyrostu $\mathbf{h}$. Zawsze można ten wektor "skrócić" tak, żeby funkcja $\phi(t)$ te pochodne do rzędu $(n+1)$-szego w przedziale $[0,1]$ miała. No chyba, że funkcja $f(\mathbf{x})$ jest jakaś wredna, ale takimi się nie zajmujemy.

zmiennych ($n = 2$):

$$\sum_{j_1=1}^{2} \dots \sum_{j_k=1}^{2} \left.\frac{\partial^k f}{\partial x_{j_1} \dots \partial x_{j_k}}\right|_{\mathbf{x}} h_{j_1} \dots h_{j_r} = \sum_{p=0}^{k} \binom{k}{p} \left.\frac{\partial^k f}{\partial x_1^p \, \partial x_2^{k-p}}\right|_{\mathbf{x}} h_1^p \, h_2^{k-p} \, .$$

Czynnik $C^{(k)}_{p_1 \dots p_n}$ jest tu po prostu symbolem Newtona. k-ty wyraz wzoru Taylora można też napisać nieco symbolicznie w postaci

$$\sum_{j_1=1}^{n} \dots \sum_{j_k=1}^{n} \left.\frac{\partial^k f}{\partial x_{j_1} \dots \partial x_{j_k}}\right|_{\mathbf{x}} h_{j_1} \dots h_{j_r} = \left(h_1 \frac{\partial}{\partial x_1} + \dots + h_n \frac{\partial}{\partial x_n} \right)^k f(x_1, \dots, x_n) \equiv d^k f\big|_{\mathbf{x}} \, ,$$

rozumiejąc, że należy tu k-krotnie zadziałać na funkcję f operatorem różniczkowym

$$\left(h_1 \frac{\partial}{\partial x_1} + \dots + h_n \frac{\partial}{\partial x_n} \right) .$$

Symbol $d^k f|_{\mathbf{x}}$ można rozumieć jako różniczkę k-tego rzędu funkcji f.

Resztę R_{r+1} otrzymuje się przy tej konstrukcji z reszty

$$R_{r+1} = \frac{t^{r+1}}{(r+1)!} \, \phi^{(r+1)}(\theta t) \qquad \text{gdzie} \quad 0 < \theta < 1 \, ,$$

wzoru Taylora dla funkcji $\phi(t)$, kładąc $t = 1$:

$$R_{r+1} = \frac{1}{(r+1)!} \left(h_1 \frac{\partial}{\partial x_1} + \dots + h_n \frac{\partial}{\partial x_n} \right)^{r+1} f \Bigg|_{\mathbf{x}+\theta\mathbf{h}} .$$

Napiszmy jeszcze w pełnej krasie wzór Taylora dla funkcji $f(x, y)$ dwóch zmiennych:

$$\begin{aligned} f(x + h_x, \, x + h_y) = f(x, y) &+ \left.\frac{\partial f}{\partial x}\right|_{x,y} h_x + \left.\frac{\partial f}{\partial y}\right|_{x,y} h_y \\ &+ \frac{1}{2} \left(\left.\frac{\partial^2 f}{\partial x^2}\right|_{x,y} h_x^2 + 2 \left.\frac{\partial^2 f}{\partial x \, \partial y}\right|_{x,y} h_x h_y + \left.\frac{\partial^2 f}{\partial y^2}\right|_{x,y} h_y^2 \right) \\ &+ \dots + \frac{1}{r!} \sum_{p=0}^{r} \binom{r}{p} \left.\frac{\partial^r f}{\partial x^p \partial y^{r-p}}\right|_{x,y} h_x^p h_y^{r-p} \\ &+ \frac{1}{(r+1)!} \sum_{p=0}^{r+1} \binom{r+1}{p} \left.\frac{\partial^{r=1} f}{\partial x^p \partial y^{r+1-p}}\right|_{x+\theta h_x, y+\theta h_y} h_x^p h_y^{r+1-p} \, . \end{aligned}$$

Tak jak w przypadku funkcji jednej zmiennej, jeśli można pokazać, że R_{n+1} dąży do zera, gdy $n \to \infty$, otrzymuje się nieskończony szereg Taylora. Ponieważ szereg Taylora jest jednoznaczny, naogół zamiast tępo obliczać pochodne cząstkowe, wygodniej jest przy rozwijaniu funkcji wielu zmiennych w taki szereg skorzystać ze znanych rozwinięć kilku funkcji elementarnych. Poniżej zobaczymy to na przykładach. Rozwinięcie w szereg Taylora daje też uzasadnienie metody szukania ekstremów funkcji wielu zmiennych, czym zajmiemy się w przyszłym tygodniu.

Przykłady

Rozwinąć w szereg Taylora wokół punktu $(0,0)$ do trzeciego rzędu włącznie funkcję $f(x,y) = e^x \sin y$. Korzystamy ze znanych[21] rozwinięć funkcji exponens i funkcji sinus:

$$
\begin{aligned}
f(x,y) &= \left(1 + x + \frac{1}{2}\,x^2 + \frac{1}{6}\,x^3 + \ldots\right)\left(y - \frac{1}{6}\,y^3 + \ldots\right) \\
&= y + xy + \frac{1}{2}\,x^2 y - \frac{1}{6}\,y^3 + \ldots
\end{aligned}
$$

Sprawdzamy, że to samo wychodzi z pochodnych: $f(0,0) = 0$,

$$
\begin{aligned}
f_x(x,y) &= e^x \sin y\,, & f_x(0,0) &= 0\,, \\
f_y(x,y) &= e^x \cos y\,, & f_y(0,0) &= 1\,, \\
f_{xx}(x,y) &= e^x \sin y\,, & f_{xx}(0,0) &= 0\,, \\
f_{yy}(x,y) &= -e^x \sin y\,, & f_{yy}(0,0) &= 0\,, \\
f_{xy}(x,y) &= e^x \cos y\,, & f_{xy}(0,0) &= 1\,, \\
f_{xxx}(x,y) &= e^x \sin y\,, & f_{xxx}(0,0) &= 0\,, \\
f_{xxy}(x,y) &= e^x \cos y\,, & f_{xxy}(0,0) &= 1\,, \\
f_{xyy}(x,y) &= -e^x \sin y\,, & f_{xyy}(0,0) &= 0\,, \\
f_{yyy}(x,y) &= -e^x \cos y\,, & f_{yyy}(0,0) &= -1\,.
\end{aligned}
$$

I teraz składamy to w całość (pierwsze liczby w środkowych wyrazach w nawiasach są czynnikami kombinatorycznymi biorącymi się z symbolu newtona):

$$
\begin{aligned}
f(x,y) &= 0 + 0\cdot x + 1\cdot y + \frac{1}{2!}\left(0\cdot x^2 + 2\cdot 1\cdot xy + 0\cdot y^2\right) \\
&\quad + \frac{1}{3!}\left(0\cdot x^3 + 3\cdot 1\cdot x^2 y + 3\cdot 0\cdot xy^2 + (-1)\cdot y^3\right),
\end{aligned}
$$

co jest tym samym rozwinięciem, co uzyskane wyżej.

To samo z funkcją $f(x,y) = \ln(1+x+y^2)$: rozwijamu do trzeciego rzędu wokół punktu $(0,0)$. Najpierw sprytem: pamiętamy, że[22] $\ln(1+\varepsilon) = \varepsilon - \frac{1}{2}\,\varepsilon^2 + \frac{1}{3}\,\varepsilon^3 + \ldots$, więc, biorąc $\varepsilon = x + y^2$,

$$
\begin{aligned}
\ln(1+x+y^2) &= x + y^2 - \frac{1}{2}\,(x+y^2)^2 + \frac{1}{3}\,(x+y^2)^3 + \ldots \\
&= x + y^2 - \frac{1}{2}\,x^2 - xy^2 - \frac{1}{2}\,y^4 + \frac{1}{3}\,x^3 + x^2y^2 + xy^4 + \frac{1}{3}\,y^6 + \ldots
\end{aligned}
$$

[21] To są rzeczy, które trzeba pamiętać aż po grób.

[22] Znów: to trzeba przez sen recytować!

Ograniczając rozwinięcie do wyrazów trzeciego rzędu mamy więc

$$\ln(1+x+y^2) = x - \frac{1}{2}\,x^2 + y^2 + \frac{1}{3}\,x^3 - xy^2 + \ldots$$

Teraz to samo zw wzoru z pochodnymi: $f(0,0) = 0$ i (dobre ćwiczenie w obliczaniu pochodnych!)

$$\begin{aligned}
f_x(x,y) &= \frac{1}{1+x+y^2}\,, & f_x(0,0) &= 1\,,\\
f_y(x,y) &= \frac{2y}{1+x+y^2}\,, & f_y(0,0) &= 0\,,\\
f_{xx}(x,y) &= -\frac{1}{(1+x+y^2)^2}\,, & f_{xx}(0,0) &= -1\,,\\
f_{yy}(x,y) &= \frac{2+2x-2y^2}{(1+x+y^2)^2}\,, & f_{yy}(0,0) &= 2\,,\\
f_{xy}(x,y) &= -\frac{2y}{(1+x+y^2)^2}\,, & f_{xy}(0,0) &= 0\,,\\
f_{xxx}(x,y) &= \frac{2}{(1+x+y^2)^3}\,, & f_{xxx}(0,0) &= 2\,,\\
f_{xxy}(x,y) &= \frac{4y}{(1+x+y^2)^3}\,, & f_{xxy}(0,0) &= 0\,,\\
f_{xyy}(x,y) &= -\frac{2+2x-6y^2}{(1+x+y^2)^3}\,, & f_{xyy}(0,0) &= -2\,,\\
f_{yyy}(x,y) &= -\frac{12y+12xy-4y^3}{(1+x+y^2)^3}\,, & f_{yyy}(0,0) &= 0\,.
\end{aligned}$$

No i znów składamy to w całość

$$\begin{aligned}
f(x,y) = 0 + 1\cdot x + 0\cdot y + \frac{1}{2!}\big((-1)\cdot x^2 + 2\cdot 0\cdot xy + 2\cdot y^2\big)\\
+\frac{1}{3!}\big(2\cdot x^3 + 3\cdot 0\cdot x^2y + 3\cdot(-2)\cdot xy^2 + 0\cdot y^3\big)\,,
\end{aligned}$$

I znów jest to to samo, co poprzednio.

Rozwińmy jeszcze funkcję $f(x,y) = -x^2 + 2xy + 3y^2 - 6x - 2y + 4$ w szereg Taylora wokół punktu, dla odmiany, $(-2,1)$: $f(-2,1) = -4 - 4 + 3 + 12 - 2 + 4 = 9$ i

$$\begin{aligned}
f_x(x,y) &= -2x + 2y - 6\,, & f_x(0,0) &= 0\,,\\
f_y(x,y) &= 3x + 6y - 2\,, & f_y(0,0) &= 0\,,\\
f_{xx}(x,y) &= -2\,,\\
f_{yy}(x,y) &= 6\,,\\
f_{xy}(x,y) &= 2\,.
\end{aligned}$$

Wszystkie dalsze pochodne są po prostu równe zeru - wzór Taylora ma w tym przypadku skończoną liczbę wyrazów. Składamy w całość: $h_x \equiv x+2$, $h_y \equiv y-1$:

$$f(-2+h_x,\,1+h_y) = 9 + 0\cdot h_x + 0\cdot h_y + \frac{1}{2}\left(-2\cdot h_x^2 + 2\cdot 2\cdot h_x h_y + 6\cdot y^2\right),$$

czyli, wstawiajac $h_x = x+2$ i $h_2 = y-1$,

$$f(x,y) = 9 - (x+2)^2 + 2\,(x+2)(y-1) + 3\,(y-1)^2\,.$$

Jest to oczywiście ten sam wielomian, co funkcja $f(x,y)$, tylko inaczej pogrupowany. Każdy wielomian wielu zmiennych skończonego stopnia n można rozwinąć w szereg Taylora wokół dowolnego punktu i zawsze wyjdzie ten sam wielomian, tylko inaczej zorganizowany.

Napiszmy jeszcze wzór Taylora rzędu trzeciego uwzględniający resztę stosując go do funkcji $f(x,y) = \sin^2(x+y)$ i punktu (π,π). Oczywiście $f(\pi,\pi) = 0$. Obliczamy pochodne:

$$\begin{aligned}
&f_x(x,y) = f_y(x,y) = \sin(2x+2y)\,,\\
&f_{xx}(x,y) = f_{yy}(x,y) = f_{xy}(x,y) = 2\cos(2x+2y)\,,\\
&f_{xxx}(x,y) = f_{xxy}(x,y) = f_{xyy}(x,y) = f_{yyy}(x,y) = -4\sin(2x+2y)\,.
\end{aligned}$$

Zatem $f_x(\pi,\pi) = f_x(\pi,\pi) = 0$ $f_{xx}(\pi,\pi) = f_{yy}(\pi,\pi) = f_{xy}(\pi,\pi) = 2$ i $f_{xxx}(\pi,\pi) = f_{xxy}(\pi,\pi) = f_{xyy}(\pi,\pi) = f_{yyy}(\pi,\pi) = 0$. Kładziemy $h_x = x-\pi$, $h_y = y-\pi$:

$$f(\pi+h_x,\,\pi+h_y) = 0 + 0\cdot h_x + 0\cdot h_y + \frac{1}{2!}\left(2\cdot h_x^2 + 2\cdot 2\cdot h_x h_y + 2\cdot h_y^2\right) + 0 + R_4\,.$$

Zero przed R_4 to są wyrazy rzędu trzeciego, które są wszystkie zerami. Reszta R_4 zgodnie z przytoczonymi wcześniej wzorami jest dana przez

$$R_4 = \frac{1}{4!}\sum_{p=0}^{4}\binom{4}{p}\left.\frac{\partial^4 f}{\partial x^p\,\partial y^{4-p}}\right|_{(\pi+\theta h_x,\pi+\theta h_y)} h_x^p h_y^{4-p}\,,$$

gdzie $0 < \theta < 1$. Oczywiście cały wic w tym, że (bez szczegółowej analizy konkretnej funkcji) nie wiadomo ile ta θ ma być, ale tu możemy prosto oszacować $|R_4|$ od góry, czyli oszacować błąd popełniany przez urwanie wzoru Taylora na wyrazach trzeciego rzędu (i nieuwzględnianiu reszty). Daje się to zrobić, bo wszystkie pochodne czwartego rzędu tej badanej tu funkcji są takie same

$$\frac{\partial^4 f}{\partial x^p\,\partial y^{4-p}} = -8\cos(2x+2y)\,, \qquad p = 0,1,2,3,4\,.$$

Obliczamy je w jakimś punkcie $(\pi+\theta h_x, \pi+\theta h_y)$, ale że $|\cos(2x+2y)| \le 1$, to możemy napisać

$$\begin{aligned}
|R_4| &= \left|-\frac{1}{4!}\sum_{p=0}^{4}\binom{4}{p} h_x^p h_y^{4-p}\, 8\cos\left(2(\pi+\theta h_x) + 2(\pi+\theta h_y)\right)\right|\\
&\le \frac{8}{4!}\left|\sum_{p=0}^{4}\binom{4}{p} h_x^p h_y^{4-p}\right| = \frac{1}{3}\,|h_x+h_y|^4\,.
\end{aligned}$$

Ekstrema funkcji wielu zmiennych

Ze wzoru Taylora wynikają warunki istnienia lokalnych ekstremów funkcji wielu zmiennych. Niech bowiem $f(x_1, \dots, x_n)$ będzie funkcją ciągłą i (przynajmniej) dwakroć różniczkowalną w pewnym (otwartym) otoczeniu punktu $\mathbf{x}^* \equiv (x_1^*, \dots, x_n^*) \in \mathbb{R}^n$ o ciągłych w otoczeniu tego punktu drugich[23] pochodnych. Zastosowany do funkcji f w otoczeniu punktu $\mathbf{x}^*$ wzór Taylora drugiego rzędu (który, przypomnijmy, jest ścisłą równością) ma postać

$$
\begin{aligned}
f(x_1^* + h_1, \dots, x_n^* + h_n) &= f(x_1^*, \dots, x_n^*) + (f_{x_1}, \dots, f_{x_n})|_{\mathbf{x}^*} \begin{pmatrix} h_1 \\ \cdot \\ \cdot \\ h_n \end{pmatrix} \\
&+ \frac{1}{2}\,(h_1, \dots, h_n) \begin{pmatrix} f_{x_1x_1} & \dots & f_{x_1x_n} \\ \cdot & \dots & \cdot \\ \cdot & \dots & \cdot \\ f_{x_nx_1} & \dots & f_{x_nx_n} \end{pmatrix}_{\mathbf{x}^*+\theta\mathbf{h}} \begin{pmatrix} h_1 \\ \cdot \\ \cdot \\ h_n \end{pmatrix}.
\end{aligned}
$$

Pochodne drugiego rzędu są tu obliczone gdzieś na linii pomiędzy punktem $\mathbf{x}^*$ i punktem $\mathbf{x}^* + \mathbf{h}$ ($0 \le \theta \le 1$, jak to we wzorze Taylora). Jeśli wszystkie pochodne cząstkowe pierwszego rzędu obliczone w punkcie $\mathbf{x}^*$ są równe zeru (znika cały człon liniowy w wektorze $\mathbf{h}$ w pierwszej linii powyższego wzoru Taylora), a forma kwadratowa drugich pochodnych funkcji f *obliczonych w punkcie* $\mathbf{x}^*$ (a nie w punkcie $\mathbf{x}^* + \theta\mathbf{h}$!)

$$
Q_{\mathbf{x}^*}(\mathbf{h}) = Q_{ij}(\mathbf{x}^*) h_i h_j \equiv \left.\frac{\partial^2 f}{\partial x_i\, \partial x_j}\right|_{\mathbf{x}^*} h_i h_j \equiv (h_1, \dots, h_n) \begin{pmatrix} f_{x_1x_1} & \dots & f_{x_1x_n} \\ \cdot & \dots & \cdot \\ \cdot & \dots & \cdot \\ f_{x_nx_1} & \dots & f_{x_nx_n} \end{pmatrix}_{\mathbf{x}^*} \begin{pmatrix} h_1 \\ \cdot \\ \cdot \\ h_n \end{pmatrix},
$$

jest dodatnio (ujemnie) określona, tzn. wyrażenie to jest zawsze dodatnie (ujemne) dla każdego wektora przemieszczenia $\mathbf{h}$, to funkcja ma w punkcie $\mathbf{x}^*$ lokalne minimum (maksimum). Istotnie: ponieważ założyliśmy, że drugie pochodne funkcji f są ciągłe w otoczeniu punktu $\mathbf{x}^*$, to zawsze można wybrać tak małe otoczenie (czyli także dostatecznie "krótki" wektor $\mathbf{h}$), żeby w całym tym otoczeniu, więc także i w nieznanym a priori punkcie $\mathbf{x}^* + \theta\mathbf{h}$, forma kwadratowa drugich pochodnych była też dodatnio (ujemnie) określona (przez analogię: jeśli $f(0) > 0$, a funkcja f jednej zmiennej jest ciągła w otoczeniu punktu 0, to zawsze można dobrać taki $|\varepsilon|$, że $f(\varepsilon) > 0$ również; kluczowa jest tu ciągłość). Zatem przy dostatecznie krótkich wektorach $\mathbf{h}$ dodatnio (ujemnie) określona jest też forma kwadratowa (w której drugie pochodne są obliczone w punkcie $\mathbf{x}^* + \theta\mathbf{h}$) w wypisanym wyżej wzorze Taylora drugiego rzędu i tym samym jest jasne że dla dowolnego punktu $\mathbf{x}^* + \mathbf{h}$ należącego do tego otoczenia $f(\mathbf{x}^* + \mathbf{h}) > f(\mathbf{x}^*)$, czyli w $\mathbf{x}^*$ funkcja ma lokalne minimum ($f(\mathbf{x}^* + \mathbf{h}) < f(\mathbf{x}^*)$ i funkcja ma w $\mathbf{x}^*$ lokalne maksimum).

[23] Że pierwsze są ciągłe wynika z istnienia drugich pochodnych.

Punkty $\mathbf{x}^*$, w których znikają wszystkie pierwsze pochodne cząstkowe funkcji f nazywa się jej *punktami krytycznymi* lub punktami *stacjonarnymi* (ta druga nazwa, to dlatego, że przy odchyleniu się od takiego punktu wartość funkcji z dokładnością do pierwszego rzędu nie zmienia się, czyli pozostaje stacjonarna). Jeśli spełnione są założenia o ciągłości drugich pochodnych w otoczeniu takiego punktu, a forma kwadratowa drugich pochodnych ma w punkcie $\mathbf{x}^*$ sygnaturę mieszaną (zob. skrypt do algebry), tj. ileś plusów i ileś minusów (i, być może jakieś zera), to punkt taki nazywa się punktem *siodłowym* funkcji f, co wynika z tego, że przy pewnych wyborach kierunku wektora $\mathbf{h}$ zachodzi ostra nierówność $f(\mathbf{x}^* + \mathbf{h}) > f(\mathbf{x}^*)$, a przy innych wyborach kierunku zachodzi ostra nierówność $f(\mathbf{x}^* + \mathbf{h}) < f(\mathbf{x}^*)$: w szczególnym przypadku funkcji $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ ma ona, na trójwymiarowym wykresie, w otoczeniu punktu $\mathbf{x}^*$ kształt mniej lub bardziej przypominający siodło.

Może też się zdarzyć, że w punkcie krytycznym $\mathbf{x}^*$ forma kwadratowa drugich pochodnych ma sygnaturę z zerami. Jeśli jest to sygnatura typu kilka plusów, kilka minusów i zero (lub zera), to w punkcie takim funkcja nie ma ekstremum, bo i tak (znów przy założonej ciągłości drugich pochodnych!) są kierunki (te "plusowe"), w których $f(\mathbf{x}^* + \mathbf{h}) > f(\mathbf{x}^*)$ i są inne kierunki (te "minusowe"), w których $f(\mathbf{x}^* + \mathbf{h}) < f(\mathbf{x}^*)$. Mniej jasna jest sytuacja, gdy sygnaturą formy kwadratowej drugich pochodnych w punkcie krytycznym jest kilka plusów i zera (kilka minusów i zera) lub też same zera (forma Q zerowa, tzn. dająca zero na każdym wektorze). W tym przypadku nie można się odwołać do ciągłości (zero w sygnaturze formy w punkcie $\mathbf{x}^*$ naogół przestaje być zerem już w sąsiednim punkcie) i bez dokładniejszej analizy nie wiadomo, jaka jest sygatura formy w (nieznanym a priori) punkcie $\mathbf{x}^* + \theta\mathbf{h}$ (w którym bierze się formę kwadratową drugich pochodnych we wzorze Taylora drugiego rzędu). Czasem można jednak orzec, czy w takim punkcie krytycznym jest ekstremum. Np. jeśli w punkcie krytycznym $\mathbf{x}^*$ forma kwadratowa drugich pochodnych jest całkowicie zerowa (znikają w tym punkcie wszystkie drugie pochodne cząstkowe), a można wypisać wzór Taylora trzeciego rzędu bo trzecie pochodne funkcji f istnieją i są one ciągłe w jakimś otwartym otoczeniu punktu $\mathbf{x}^*$ i nie są wszystkie tożsamościowo (w tym otoczeniu) zerowe, to funkcja f nie ma w punkcie $\mathbf{x}^*$ ekstremum, bo forma trójliniowa

$$\left.\frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k}\right|_{\mathbf{x}^* + \theta\mathbf{h}} h_i h_j h_k \,,$$

nie znika (znów dzięki założonej ciągłości trzecich pochodnych) na przynajmniej niektórych wektorach, a jej wartość zmienia znak przy zmianie zwrotu wektora $\mathbf{h}$, tak iż jeśli $f(\mathbf{x}^* + \mathbf{h}) < f(\mathbf{x}^*)$, to $f(\mathbf{x}^* - \mathbf{h}) > f(\mathbf{x}^*)$. W takim przypadku (zerowej formy kwadratowej drugich pochodnych) dopiero znikanie w $\mathbf{x}^*$ także wszystkich trzecich pochodnych cząstkowych stwarza szansę na istnienie w takim punkcie minimum lub maksimum. Czy funkcja ma rzeczywiście w takim punkcie ekstremum, to zależy od charakteru formy tetraliniowej[24]

$$T(\mathbf{h}) = T_{ijkl} h_i h_j k_k h_l \,, \qquad T_{ijkl} \equiv \left. \frac{\partial^4 f}{\partial x_i \, \partial x_j \, \partial x_k \, \partial x_l} \right|_{\mathbf{x}^*} ,$$

ale charakteru takich form nie nauczyliśmy się badać, więc takimi przypadkami się nie będziemy tu zajmować. Jeszcze bardziej skomplikowane są sytuacje, gdy sygnaturą formy kwadratowej drugich pochodnych w punkcie krytycznym $\mathbf{x}^*$ są same plusy i zera lub same minusy i zera. Ekstremum może wtedy w takim punkcie istnieć, ale żeby to stwierdzić, trzeba zbadać zachowanie formy trójliniowej trzecich i tetraliniowej czwartych pochodnych na takich wektorach $\mathbf{h}$, na których forma kwadratowej drugich pochodnych w punkcie krytycznym $\mathbf{x}^*$ się zeruje (znów forma trójliniowa trzecich pochodnych musi na takich wektorach znikać, aby była szansa na istnienie ekstremum).

W całej procedurze szukania ekstremów funkcji wielu zmiennych najtrudniejszym krokiem jest znalezienie punktów krytycznych, gdyż trzeba znaleźć (wszystkie) rozwiązania układu n równań, które naogół nie są liniowe. Wymaga to zwykle trochę "sprytu boiskowego". Pozostałe kroki są już bezproblemowe. Charakter form kwadratowych nauczyliśmy się już badać w części algebraicznej. Można to robić albo diagonalizując ją metodą Lagrange'a, albo (najczęściej) stosując kryterium "minorowe".

Wreszcie, trzeba mieć świadomość, że funkcja f może mieć ekstrema, których istnienia nie da się zbadać stosując wzór Taylora, bo np. mogą one się znajdować w punktach, w których np. funkcja nie jest różniczkowalna, albo nie ma pochodnych wyższych rzędów. Jakieś przykłady tego typu będą niżej.

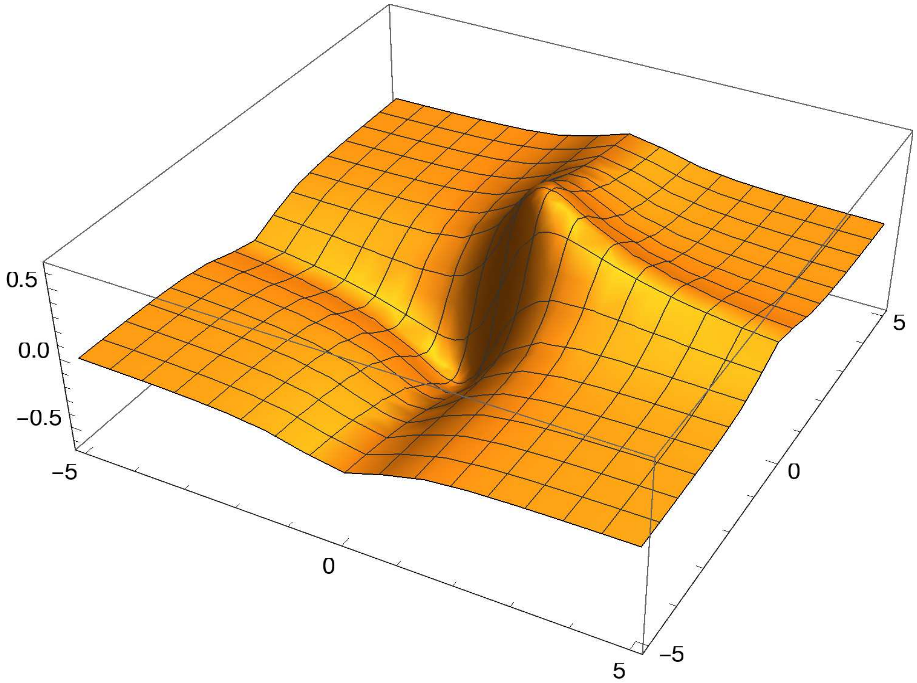
Zadanie EX.Z1
Znaleźć punkty krytyczne funkcji określonej na $\mathbb{R}^2$ o wartościach rzeczywistych

$$f(x,y) = \frac{x+y}{(1+x^2)(1+y^2)} \,.$$

i zbadać, czy są one minimami, maksimami, czy punktami siodłowymi.
Rozwiązanie: Znajdujemy najpierw pierwsze pochodne cząstkowe. Ponieważ funkcja ma symetrię $x \leftrightarrow y$, wystarczy znaleźć jedną, np. f_x, a f_y można będzie otrzymać zamieniając w f_x miejscami x z y-kiem. W ten sposób dostajemy

$$f_x(x,y) = \frac{1-x^2-2xy}{(1+x^2)^2(1+y^2)} \,,$$
$$f_y(x,y) = \frac{1-y^2-2xy}{(1+x^2)(1+y^2)^2} \,.$$

[24] T jak tetradrachma (jak w wierszu Kawafisa "Orofernes": *Ten, który na tetradrachmie ma twarz jak gdyby rozjaśnioną uśmiechem...*) - taka moneta z greckiej, a właściwie hellenistycznej, starożytności, albo jak tetrarchia - system rządów wymyślony przez Dioklecjana, czy tetralogia - już była wspomniana w tym skrypcie tetralogia Reeda i Simona.



Rysunek 4: Kształt funkcji z Zadania EXZ1. Oś x biegnie od lewej ukosem nieco w dół, a oś y ukosem w prawo do góry.

W punktach krytycznych, czyli tam, gdzie funkcja może mieć ekstrema, $f_x = 0$ i $f_y = 0$. Ponieważ mianowniki f_x i f_y są zawsze dodatnie, rozwiązujemy układ dwóch równań:

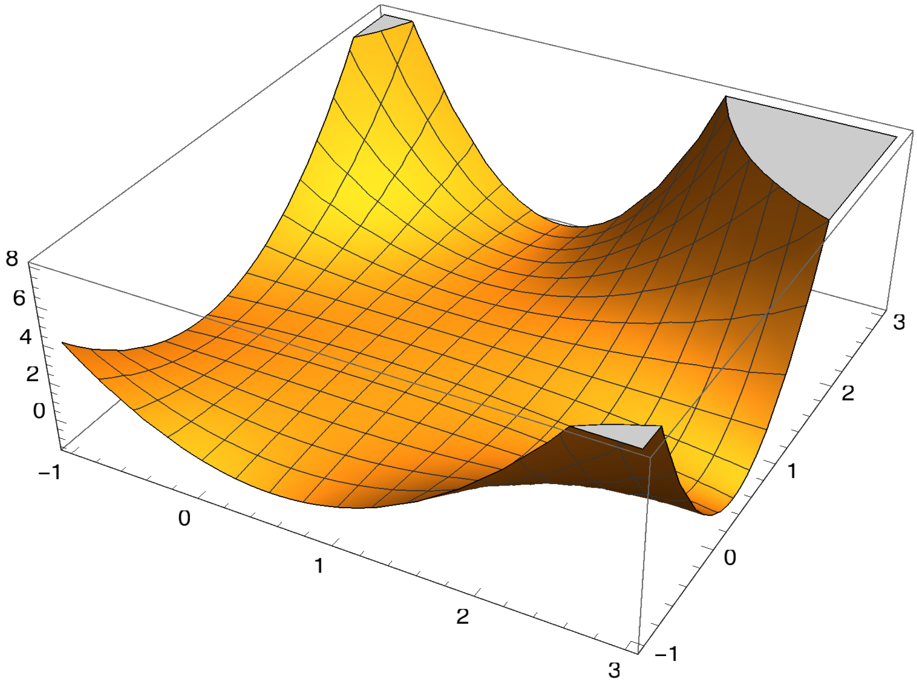
$$1 - x^2 - 2xy = 0\,,$$
$$1 - y^2 - 2xy = 0\,.$$

Odejmujemy jedno od drugiego stronami i znajdujemy, że w punktach krytycznych $x^2 = y^2$, czyli albo $x = y$, albo $x = -y$. Jeśli $x = -y$, to wyżej wypisane równania $1 + x^2 = 0$ nie mają rozwiązań. Zatem punkty krytyczne są tylko w $x = y = \pm 1/\sqrt{3}$.

Trzeba teraz obliczyć drugie pochodne cząstkowe. Ponieważ funkcja jest trochę skomplikowana, jest to dobre ćwiczenie w różniczkowaniu (znów f_{xx} i f_{yy} różnią się zamianą $x \leftrightarrow y$). Znajdujemy:

$$f_{xx} = \left.\frac{2(-3x - y + x^3 + 3x^2y)}{(1+x^2)^3(1+y^2)}\right|_{x=y} = \frac{8x(x^2-1)}{(1+x^2)^4}\,,$$
$$f_{yy} = \left.\frac{2(-x - 3y + 3xy^2 + y^3)}{(1+x^2)(1+y^2)^3}\right|_{x=y} = \frac{8x(x^2-1)}{(1+x^2)^4}\,,$$
$$f_{xy} = \left.\frac{2(-x - y + x^2y + xy^2)}{(1+x^2)^2(1+y^2)^2}\right|_{x=y} = \frac{4x(x^2-1)}{(1+x^2)^4}\,.$$

Obliczamy teraz w każdym z dwu punktów krytycznych, $x_1 = y_1 = 1/\sqrt{3}$ oraz $x_2 = y_2 = -1/\sqrt{3}$, macierz drugich pochodnych. Widać, że macierze te różnią się znakami



Rysunek 5: Kształt funkcji z Zadania EXZ2 w pobliżu punktu $(a, b) = (1/2, 1/2)$. Maksimum w tym punkcie jest słabo widoczne i pewnie bez analizy pochodnych trudno by było być pewnym, że ono tam rzeczywiście jest.

tylko (tzn. wszystkie elementy macierzy drugich pochodnych w punkcie $x_1 = y_1 = 1/\sqrt{3}$ różnią się tylko o znak od odpowiadających im elementów macierzy drugich pochodnych w punkcie $x_2 = y_2 = -1/\sqrt{3}$). Jak łatwo obliczyć, w $x_1 = y_1 = 1/\sqrt{3}$

$$Q(x_1, y_1) = -\frac{4}{\sqrt{3}}\,\frac{2}{3}\left(\frac{3}{4}\right)^4 \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

Macierz ta jest ujemnie określona, zatem w punkcie $x_1 = y_1 = 1/\sqrt{3}$ funkcja ma maksimum lokalne. Zatem w $x_2 = y_2 = -1/\sqrt{3}$ macierz $Q(x_2, y_2)$ jest dodatnio określona i tam funkcja ma lokalne minimum. Że tak jest rzeczywiście pokazuje wykres 4.

Zadanie EX.Z2

Znaleźć punkty krytyczne funkcji określonej na $\mathbb{R}^2$ o wartościach rzeczywistych

$$f(x, y) = (x^2 - 2ax)(y^2 - 2by) = xy(x - 2a)(y - 2b)\,,$$

i zbadać, czy są one minimami, maksimami, czy punktami siodłowymi.

Rozwiązanie: Tu możnaby myśleć, że skoro badana funkcja jest iloczynem funkcji $g(x) = x(x-2a)$, która ma minimum w punkcie $x = a$ (bo to zwykła funkcja kwadratowa) i funkcji $h(y)$, która ma minimum w punkcie $y = b$, to $f(x, y)$ też musi mieć minimum w punkcie (a, b). Nie jest to jednak takie proste. Pochodne cząstkowe

$$f_x(x, y) = 2y\,(x - a)(y - 2b)\,,$$

$$f_y(x,y) = 2x\,(x-2a)(y-b)\,,$$

zerują się bowiem w kilku punktach, z których (a,b), jest tylko jednym z możliwych:

$$\begin{array}{ccccc} i) & ii) & iii) & iv) & v) \\ (a,b)\,, & (0,0)\,, & (2a,2b)\,, & (2a,0)\,, & (0,2b)\,. \end{array}$$

Drugie pochodne

$$\begin{aligned} f_{xx}(x,y) &= 2y\,(y-2b)\,, \\ f_{yy}(x,y) &= 2x\,(x-2a)\,, \\ f_{xy}(x,y) &= 4(x-a)(y-b)\,, \end{aligned}$$

Dają w tych punktach odpowiednio formy kwadratowe

$$Q_{i)} = \begin{pmatrix} -2b^2 & 0 \\ 0 & -2a^2 \end{pmatrix}, \quad Q_{ii),iii)} = \begin{pmatrix} 0 & 4ab \\ 4ab & 0 \end{pmatrix}, \quad Q_{iv),v)} = \begin{pmatrix} 0 & -4ab \\ -4ab & 0 \end{pmatrix}.$$

Pierwsza forma kwadratowa jest zawsze (z wyjatkiem sytuacji, gdy $a = b = 0$) ujemnie określona z czego wynika, że odwrotnie niż możnaby myśleć, w punkcie (a,b) funkcja ma lokalne maksimum. W pozostałych czterech punktach formy kwadratowe drugich pochodnych są nieokreślone. Aby się o tym przekonać (jeśli nie ufamy kryterium "minorowemu") wystarczy zbadać je na dwóch prostych wektorach przesunięć $\mathbf{h} = (h,h)$ i $\mathbf{h} = (h,-h)$; weźmy np. $Q_{ii),iii)}$:

$$(h,h)\begin{pmatrix} 0 & 4ab \\ 4ab & 0 \end{pmatrix}\begin{pmatrix} h \\ h \end{pmatrix} = 8ab\,h^2\,, \qquad (h,-h)\begin{pmatrix} 0 & 4ab \\ 4ab & 0 \end{pmatrix}\begin{pmatrix} h \\ -h \end{pmatrix} = -8ab\,h^2\,.$$

Teraz kilka mniej typowych zadań.

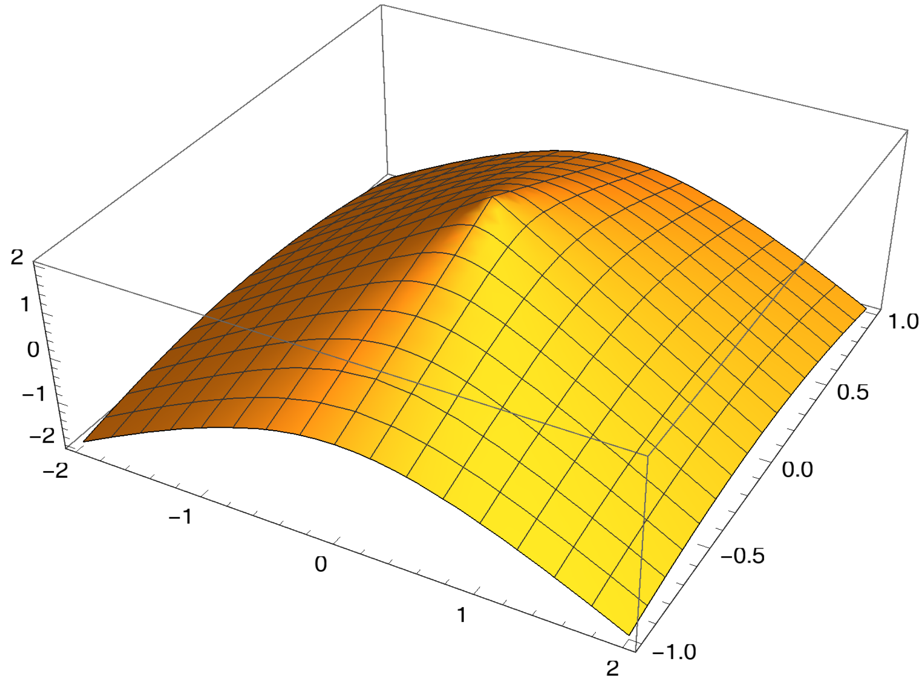
Zadanie EX.Z3
Znaleźć ekstrema funkcji określonej na $\mathbb{R}^2$ o wartościach rzeczywistych

$$f(x,y) = 2 - \sqrt{3x^2+4y^2}\,.$$

Rozwiązanie: Pochodne cząstkowe tej funkcji

$$f_x(x,y) = -\frac{3x}{\sqrt{3x^2+4y^2}}\,, \qquad f_y(x,y) = -\frac{4y}{\sqrt{3x^2+4y^2}}\,,$$

są różne od zera wszędzie, poza punktem $(0,0)$. W tym punkcie są one nieciągłe: funkcja ma w punkcie $(0,0)$ pochodne kierunkowe w każdym kierunku, ale nie jest w tym punkcie różniczkowalna w sensie mocnym. Wobec tego, nie można tu odwołać się do reguły, że ekstrema funkcji znajdują się w punktach, w których znikają wszystkie pochodne cząstkowe. Niemniej jest oczywiste, że $f(x,y) < f(0,0) = 2$ we wszystkich punktach $(x,y) \neq (0,0)$. Zatem funkcja ma w punkcie $(0,0)$ ma maksimum i jest to nawet maksimum globalne.



Rysunek 6: Kształt funkcji z Zadania EXZ3 w pobliżu punktu $(0, 0)$. Widać. że w tym punkcie funkcja ma "dziubek", więc nie jest tam różniczkowalna.

Kształt funkcji jest pokazany na rysunku 6. Widać, że nieróżniczkowalność funkcji w mocnym sensie w punkcie $(0, 0)$ wynika z tego, że ma ona tam "dziubek".

Zadanie EX.Z4
Znaleźć ekstrema funkcji określonej na $\mathbb{R}^2$ o wartościach rzeczywistych wzorem

$$f(x, y) = x^8 - y^4 \,.$$

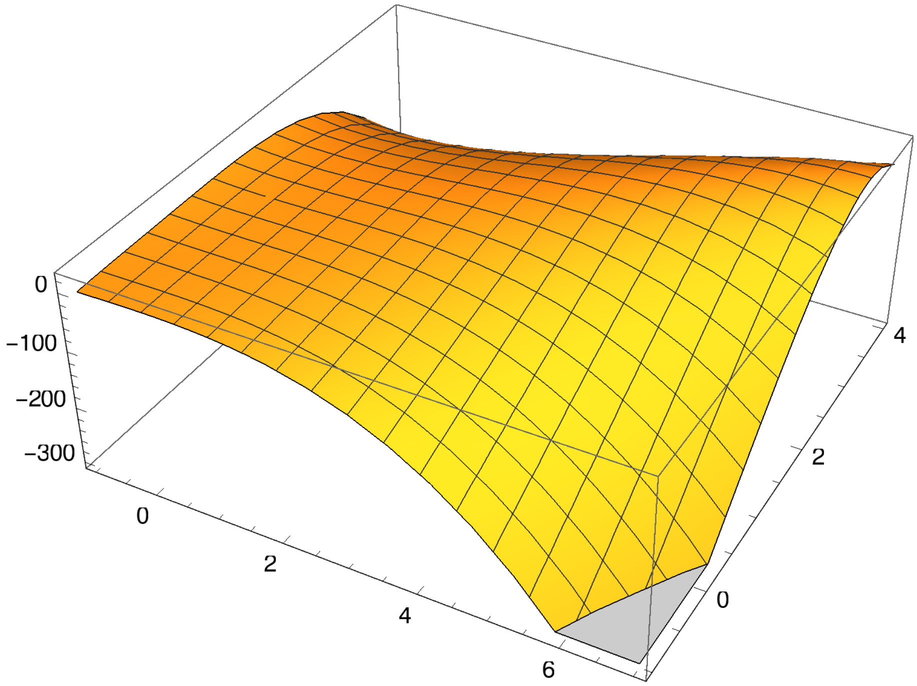
Rozwiązanie: Pochodne cząstkowe tej funkcji

$$f_x(x, y) = 8x^7 \,, \qquad f_y(x, y) = -4y^3 \,,$$

istnieją na całym $\mathbb{R}^2$ i są ciągłe, podobnie jak drugie pochodne cząstkowe

$$f_{xx}(x, y) = 56x^6 \,, \qquad f_{yy}(x, y) = -12y^2 \,, \qquad f_{xy}(x, y) = 0 \,.$$

Obie pierwsze pochodne cząstkowe zerują się tylko w punkcie $(0, 0)$ - jest więc to jedyny punkt krytyczny badanej funkcji - ale forma kwadratowa drugich pochodnych jest w tym punkcie całkowicie zerowa. Zatem nie można na jej podstawie określić charakteru tego punktu krytycznego. Można jednak to zrobić posługując się zdrowym rozsądkiem. Wystarczy zauważyć, że $f(0, 0) = 0$, oraz, że we wszystkich punktach postaci $(\varepsilon, 0)$, które mogą, gdy $\varepsilon \to 0$ leżeć w dowolnie małym otwartym otoczeniu punktu krytycznego, wartość funkcji jest większa od zera, a we wszystkich punktach postaci $(0, \varepsilon)$, też mogących



Rysunek 7: Kształt funkcji z Zadania EXZ5.

leżeć w w dowolnie małym otwartym otoczeniu punktu krytycznego, wartość funkcji jest mniejsza od zera. Zatem punkt krytyczny jest tylko punktem siodłowym.

Zadanie EX.Z5
Znaleźć punkty krytyczne funkcji określonej na $\mathbb{R}^2$ o wartościach rzeczywistych wzorem

$$f(x, y) = 3x^2y - x^3 - y^4 \,,$$

i zbadać ich charakter.
Rozwiązanie: Pochodne cząstkowe tej funkcji

$$f_x(x, y) = 6xy - 3x^2 \,, \qquad f_y(x, y) = 3x^2 - 4y^3 \,,$$

zerują się jednocześnie w punkcie $(0, 0)$ i w punkcie $(6, 3)$. Widać to w zasadzie od ręki: jeśli $f_x = 0$ bo $x = 0$, to wtedy zerowanie się f_y wymaga by i $y = 0$; jeśli zaś $(x, y) \neq (0, 0)$, to $f_x = 0$ daje $x = 2y$, co wstawione do $f_y = 0$ daje $y = 3$. Są więc dwa punkty krytyczne. Drugie pochodne

$$f_{xx}(x, y) = 6y - 6x \,, \qquad f_{yy}(x, y) = -12y^2 \,, \qquad f_{xy}(x, y) = 6x \,,$$

Dają one w punkcie $(0, 0)$ całkowicie zerową macierz formy kwadratowej. Ponieważ jednak funkcja jest wielomianem, czyli wokół każdego punktu może być dokładnie reprezentowana skończonym wielomianem, można spojrzeć na pochodne cząstkowe trzeciego rzędu.

Ponieważ kilka z nich, np. f_{xxx}, f_{xxy}, jest w punkcie $(0,0)$ niezerowych, funkcja nie może mieć w tym punkcie ekstremum.[25]

W punkcie $(6,3)$ macierz formy kwadratowej drugich pochodnych ma postać

$$\begin{pmatrix} -18 & 36 \\ 36 & -108 \end{pmatrix} = -18 \begin{pmatrix} 1 & -2 \\ -2 & 6 \end{pmatrix}.$$

Macierz ta jest ujemnie określona ($-$macierz jest dodatnio określona), zatem w tym punkcie funkcja ma maksimum lokalne. Kształt tej funkcji pokazuje rysunek 7.

Zadanie EX.Z6

Znaleźć ekstrema funkcji określonej na $\mathbb{R}^2$ o wartościach rzeczywistych wzorem

$$f(x,y) = \sin x \sin y \sin(x+y)\,.$$

Rozwiązanie: Pierwsze pochodne

$$\begin{aligned} f_x(x,y) &= \cos x \sin y \sin(x+y) + \sin x \sin y \cos(x+y) = \sin y \sin(2x+y)\,, \\ f_y(x,y) &= \sin x \cos y \sin(x+y) + \sin x \sin y \cos(x+y) = \sin x \sin(x+2y)\,, \end{aligned}$$

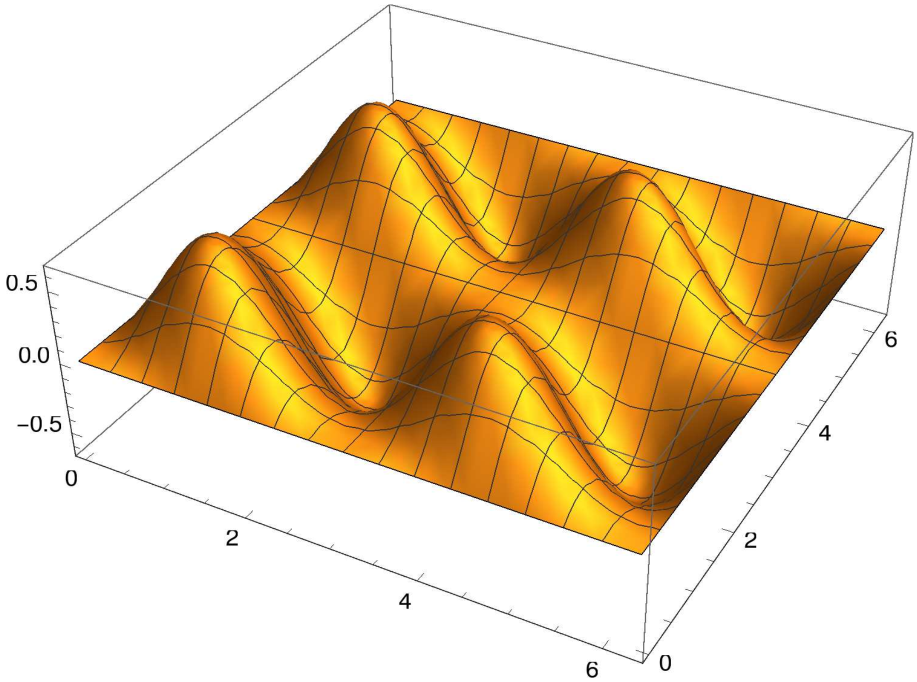
zerują się albo tam, gdzie $\sin x = \sin y = 0$, czyli w punktach postaci $(x,y) = (k\pi, l\pi)$, gdzie k i l są dowolnymi liczbami całkowitymi, lub tam, gdzie $2x+y = p\pi$ i $x+2y = r\pi$ (p i r też są dowolnymi liczbami całkowitymi). Odejmując stronami te dwa równanka znajdujemy, że $y = x - \pi(p-r)$, czyli x może się różnić od y o całkowitą wielokrotność π i wstawiając ten związek do któregokolwiek z tych dwu równanek dostajemy, że $x = \frac{\pi}{3}(2p-r)$, a $y = \frac{\pi}{3}(2r-p)$. W zasadzie trzeba jeszcze rozpatrzyć rozwiązania takie, w których $x = k\pi$ i $2x+y = n\pi$ (lub na odwrót), ale to jest to samo, co $x = k\pi$ i $y = n\pi - 2k\pi \equiv l\pi$, wiec one nie dają nic nowego. Zatem część z tych punktów, np. punkt o $p = r = 0$, pokrywa się z uprzednio już znalezionymi, ale są też i nowe. Ponieważ badana funkcja jest, jak łatwo się zorientować, biperiodyczna, tzn. $f(x+l\pi,\, y+k\pi) = f(x,y)$, przy dowolnych całkowitych k i l, wystarczy zbadać tylko charakter punktów krytycznych leżących w "komórce fundamentalnej" tj. w obszarze $[0,\pi) \times [0,\pi)$. Wynika z tego natychmiast, że rozpatrywać trzeba tylko punkty o $p = r$ (bo inaczej albo x albo y jest ujemne). W sumie więc w na płaszczyźnie $\mathbb{R}^2$ punkty krytyczne tworzą całą sieć, ale w $[0,\pi) \times [0,\pi)$ są tylko trzy

$$(0,0), \qquad (\frac{\pi}{3}, \frac{\pi}{3}), \qquad (\frac{2\pi}{3}, \frac{2\pi}{3}).$$

Drugie pochodne po prostych przekształceniach można zapisać w prostej formie

$$\begin{aligned} f_{xx}(x,y) &= 2 \sin y \cos(2x+y)\,, \\ f_{yy}(x,y) &= 2 \sin x \cos(x+2y)\,, \\ f_{xy}(x,y) &= \sin(2x+2y)\,. \end{aligned}$$

[25] W przypadku zerowania się w punkcie krytycznym formy kwadratowej drugich pochodnych warunkiem koniecznym, choć nie dostatecznym!, istnienia w tym punkcie ekstremum jest - w przypadku funkcji, które w otoczeniu tego punktu mogą być reprezentowane wzorem Taylora czwartego rzędu - znikanie wszystkich trzecich pochodnych cząstkowych.



Rysunek 8: Funkcja z Zadania EX.Z6 w kształcie "korobki dla jajic" w dźwięcznym języku rosyjskim.

Widać więc, że we wszystkich punktach typu $(k\pi,\, l\pi)$, w szczególności w punkcie $(0,0)$, macierz formy kwadratowej drugich pochodnych jest zerowa i trzeba charakter takich punktów zbadać jakoś inaczej. Jest jednak oczywiste, jeśli $x = y = \varepsilon$ i $|\varepsilon| \ll 1$, to $f(\varepsilon,\varepsilon) \approx 2\varepsilon^3$ i jeśli $\varepsilon > 0$ to funkcja jest dodatnia, a jeśli $\varepsilon < 0$, to ujemna. To wystarcza, by stwierdzić, że w dowolnie małym otwartym otoczeniu punktu $(0,0)$ (a zatem i wszystkich punktów $(k\pi,\, l\pi)$ z uwagi na biperiodyczność) leżą punkty, w których wartość funkcji jest większa i mniejsza od $f(0,0)$. Zatem funkcja nie ma w tym punkcie ekstremum.

Pozostaje zbadać charakter punktów krytycznych typu

$$(\frac{\pi}{3}, \frac{\pi}{3}), \qquad (\frac{2\pi}{3}, \frac{2\pi}{3}).$$

Macierze formy kwadratowej drugich pochodnych mają w tych punktach postacie

$$-\frac{\sqrt{3}}{2}\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \qquad \frac{\sqrt{3}}{2}\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

Jest więc jasne, że "komórce fundamentalnej" $[0,\pi) \times [0,\pi)$ funkcja ma maksimum w punkcie $(\frac{\pi}{3}, \frac{\pi}{3})$ i minimum w punkcie $(\frac{2\pi}{3}, \frac{2\pi}{3})$ oraz, że w punkcie $(0,0)$ jest siodło. Struktura ta powtarza się biperiodycznie na całej płaszczyźnie $\mathbb{R}^2$, jak to widać z rysunku 8.

Zadanie EX.Z7
Znaleźć punkty krytyczne funkcji o wartościach rzeczywistych określonej wzorem

$$f(x,y,z) = x^2 + \frac{y}{x} + \frac{z^2}{y} + \frac{1}{z}\,,$$

na $\mathbb{R}^3$ z wyłączeniem płaszczyzn $x=0$, $y=0$ i $z=0$ i zbadać ich charakter.
Rozwiązanie: Trzy równania wyznaczające punkty krytyczne

$$\begin{aligned} f_x(x,y,z) &= 2x - \frac{y}{x^2} = 0\,, \\ f_y(x,y,z) &= \frac{1}{x} - \frac{z^2}{y^2} = 0\,, \\ f_z(x,y,z) &= \frac{2z}{y} - \frac{1}{z^2} = 0\,, \end{aligned}$$

sprowadzają się, po pomnożeniu odpowiednio przez x^2, xy^2 i yz^2 (co wolno zrobić, bo $x \neq 0$, $y \neq 0$ i $z \neq 0$) do $2x^3 = y$, $y^2 = xz^2$ i $2z^3 = y$. Z połączenia pierwszego z trzecim dostaje się $x^3 = z^3$, czyli $x = z$ i teraz pierwsze, $2x^3 = y$, w połączeniu z drugim zamienionym na $2y^2 = 2x^3$ daje $y = 0$, co jest wykluczone, lub $y = \frac{1}{2}$. Zatem $x^3 = \frac{1}{4}$, czyli $x = z = 1/4^{1/3}$. Jest więc tylko jeden punkt krytyczny.

Drugie pochodne

$$\begin{aligned} f_{xx} &= 2 + \frac{2y}{x^3}\,, & f_{yy} &= \frac{2z^2}{y^3}\,, & f_{zz} &= \frac{2}{y} + \frac{2}{z^3}\,, \\ f_{xy} &= -\frac{1}{x^2}\,, & f_{xz} &= 0\,, & f_{yz} &= -\frac{2z}{y^2}\,, \end{aligned}$$

dają w punkcie krytycznym macierz

$$\begin{pmatrix} 6 & -2^{4/3} & 0 \\ -2^{4/3} & 2^{8/3} & -2^{7/3} \\ 0 & -2^{7/3} & 12 \end{pmatrix}.$$

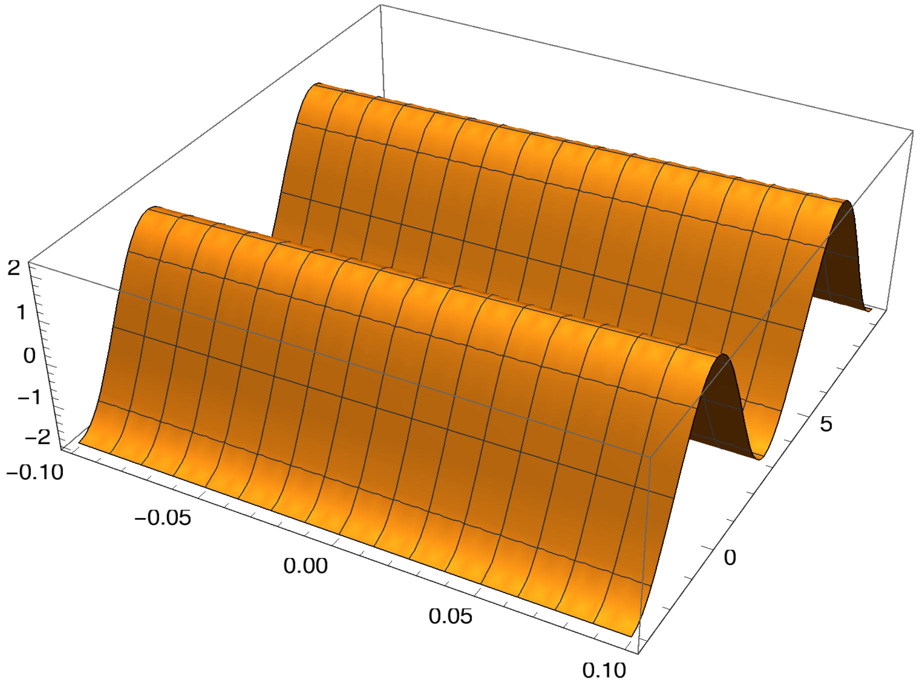
Wszystkie trzy minory tej macierzy M_{11}, M_{22} i M_{33} są dodatnie, więc funkcja ma w punkcie krytycznym minimum lokalne.

Zadanie
Pokazać, że funkcja

$$f(x,y) = (1 + e^x)\cos y + x\,e^x\,,$$

ma nieskończenie wiele lokalnych minimów, ale żadnego maksimum.
Rozwiązanie: Warunki wyznaczające punkty krytyczne

$$\begin{aligned} f_x(x,y) &= e^x\,(1 + x + \cos y) = 0\,, \\ f_y(x,y) &= -\,(1 + e^x)\sin y = 0\,, \end{aligned}$$



Rysunek 9: Funkcja mająca same minima i żadnych maksimów (oś x pokazana w pobliżu $x = 0$).

mają jako swoje rozwiązania punkty $y = n\pi$ ($f_y = 0$) i, wobec tego (bo $\cos(n\pi) = (-1)^n$, co wszyscy powinni wiedzieć), $x = -1 - (-1)^n$. Punktów krytycznych jest więc (przeliczalnie) nieskończenie wiele. Sprawdzamy drugie pochodne (druga równość zachodzi w punktach krytycznych):

$$\begin{aligned} f_{xx}(x,y) &= e^x\,(1 + x + \cos y) + e^x = e^x\,, \\ f_{yy} &= -\,(1 + e^x)\cos y = (-)^{n+1}\,(1 + e^x)\,, \\ f_{xy} &= -e^x \sin y = 0\,. \end{aligned}$$

Zatem pierwszy minor M_{11} macierzy formy drugich pochodnych jest w każdym punkcie krytycznym dodatni, ale drugi minor M_{22} jest dodatni tylko w punktach o nieparzystych n. W tych punktach są więc minima lokalne, a w punktach o parzystych n punkty siodłowe. Zatem rzeczywiście funkcja ma nieskończenie wiele minimów i żadnych maksimów. Jak to jest możliwe? Ano tak, jak pokazuje rysunek 9 - po prostu garby funkcji są nieco pochylone w kierunku x-owym i dlatego są tylko siodłami. W minimach natomiast w kierunku x-owym są dołki, bo są to funkcje $f(x,\pi) = -1 - e^x(1-x) \approx -1 - (1 + x + \ldots)(1-x) \approx -2 + x^2$.

Jeszcze inny typ zagadnień ilustruje

Zadanie (Ekstrema na zbiorze zwartym)

Znaleźć największą i najmniejszą wartość funkcji $f(x,y) = x^2 - 2y^2$ na zbiorze punktów (x,y) ograniczonych warunkiem $x^2 + y^2 \leq 36$.

Rozwiązanie: Punktem krytycznym funkcji jest punkt $(0, 0)$. W punkcie tym forma kwadratowa drugich pochodnych ma postać

$$\begin{pmatrix} 2 & 0 \\ 0 & -4 \end{pmatrix},$$

funkcja nie ma więc tam ekstremum. Jakieś maksimum i jakieś minimum funkcja jednak mieć musi, bo jak pisał wielki poeta (który Młodzieży kochana?!):

"Zbiorze szlachetny, cny zwarty zbiorze,
na Tobie człek funkcję określić sobie może,
co jeśli ciągła, zachowa Twą zwartość,
i przyjmie po drodze swą największą wartość!"

Skoro jednak wewnątrz zbioru (tzn. na zbiorze otwartym, $x^2+y^2 < 36$) ekstremów funkcja nie ma, muszą one być na brzegu zbioru. Jeśli wstawimy $y^2 = 36-x^2$, otrzymamy funkcję $f(x) = 3x^2 - 72$, która ma minimum równe -72 w $x = 0$, czyli minimalną swą wartość funkcja $f(x, y)$ przyjmuje w punktach $(0, \pm 6)$. Powinna ona jednak mieć gdzieś i wartość największą. Oczywiście podstawiając $y^2 = 36 - x^2$ zgubiliśmy[26] punkty $(\pm 6, 0)$. W nich to właśnie funkcja przyjmuje swe wartości największe, równe 36.

[26] Bo jak się zaraz dowiemy zajmując się tzw. funkcjami zadanymi w sposób uwikłany, w tych punktach warunek $F(x, y) = x^2 - y^2 - 36 = 0$ nie wyznacza funkcji $y = f(x)$, za to wyznacza funkcję $x = x(y)$.

Odwzorowania zadane w sposób uwikłany

Sytuacja którą się tu zajmiemy jest następująca: w przestrzeni $\mathbb{R}^n$ jest wyróżniony przez jakieś m warunków ($m < n$) mających postać równości (zawsze można je tak zapisać)

$$F_1(x_1, \ldots, x_n) = 0\,, \qquad \ldots, \qquad F_m(x_1, \ldots, x_n) = 0\,,$$

pewien zbiór E. Warunki te, przyjmujemy, są wszystkie niezależne, tzn. spełnienie tylko kilku z nich nie powoduje, że któryś z pozostałych jest już automatycznie spełniony.[27] Ogólniej jeszcze, dane jest odwzorowanie F zbioru $D \subset \mathbb{R}^n$ w $\mathbb{R}^m$ ($m < n$), czyli właśnie m funkcji F_i, o których zakładamy to co wyżej. Zakładamy też, że odwzorowanie F jest na zbiorze D ciągłe. Zbiór $E \subset D$ jest zadany jako przeciwobraz punktu $\mathbf{0} \in \mathbb{R}^m$, tzn. $F|_E \equiv \mathbf{0} \in \mathbb{R}^m$ (po ludzku: odwzorowanie F ograniczone do zbioru E jest tożsamościowo zerowe, albo: F odzwzorowuje cały zbiór E w jeden punkt przestrzeni $\mathbb{R}^m$, mianowicie w jej zero). Interesuje nas wtedy pytanie, kiedy zbiór E definiuje (w sposób uwikłany - stąd właśnie nazwa "odwzorowania uwikłane") uczciwe odwzorowanie $f : \mathbb{R}^{n-m} \to \mathbb{R}^m$. A jeśli definiuje, to jak obliczać pochodne cząstkowe tego odwzorowania? Pochodne cząstkowe mogą być potrzebne, bo jak mamy zadane w ten sposób np. odwzorowanie $f : \mathbb{R}^3 \to \mathbb{R}^1$ (tzn. sytuację odpowiadającą $n = 4$, $m = 1$ w naszych ogólnych rozważaniach), to możemy chcieć znać jego ekstrema, a do tego, jak już wiemy z poprzednego rozdzialiku, trzeba umieć badać pochodne cząstkowe tego odwzorowania. Żeby dać jakiś przykład: weźmy znane nam z algebry równanie liniowe

$$\begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \cdot & \ldots & \cdot \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \cdot \\ \cdot \\ x_n \end{pmatrix} - \begin{pmatrix} b_1 \\ \cdot \\ b_m \end{pmatrix} = \begin{pmatrix} 0 \\ \cdot \\ 0 \end{pmatrix},$$

w którym $m < n$, zapisane w trochę nietypowy sposób. W tej postaci lewa strona jest pewnym odwzorowaniem $F : \mathbb{R}^n \to \mathbb{R}^m$, a punkty $(x_1, \ldots, x_n) \in \mathbb{R}^n$ spełniające to równanie (z wektorem zerowym po prawej) tworzą właśnie w $\mathbb{R}^n$ zbiór E. Załóżmy, że rząd macierzy po lewej stronie jest równy dokładnie m. Jest to właśnie żądanie (liniowej) niezależności układu m równań. W takiej sytuacji jest (na podstawie tego, co wiemy z algebry - algebra to potęga!) jasne, że $F^{-1}(\mathbf{0})$, czyli właśnie zbiór E, zadaje odwzorowanie $f : \mathbb{R}^{n-m} \longrightarrow \mathbb{R}^m$. Pamiętamy przecież: możemy w macierzy wybrać niezerowy minor rzędu m - załóżmy (jak zwykle dla prostoty zapisu), że tworzy go pierwsze m kolumn

[27] W szczególności oznacza to, że nie są one liniowo zależne, tzn. żądanie, by

$$\lambda_1 F_1(x_1, \ldots, x_n) + \ldots + \lambda_m F_m(x_1, \ldots, x_n) \equiv 0\,,$$

(znak $\equiv$ oznacza tu że kombinacja po lewej stronie ma być funkcją tożsamościowo, a więc na całym $\mathbb{R}^n$, a nie tylko na E, równą zeru) pociąga za sobą równość $\lambda_1 = \ldots = \lambda_m = 0$; jednak sformułowane żądanie jest ogólniejsze, bo wyklucza też sytuacje takie, że np. $F_2(x_1, \ldots, x_n) = (F_1(x_1, \ldots, x_n))^2$, czy $F_4(x_1, \ldots, x_n) = [2F_1(x_1, \ldots, x_n) - 3F_2^2(x_1, \ldots, x_n)]^{1/3}$, czy coś w tym guście. Warunki te można pominąć, wtedy należy to trochę bardziej zawile sformułować - nie będziemy się tu w to wdawać.

macierzy; $n-m$ zmiennych $x_{m+1}, \ldots, x_n$ wraz z wektorem $\mathbf{b}$ przenosimy na prawą stronę równania i np. kramersiętami rozwiązujemy układ liniowych równań na niewiadome x_1, ..., x_m. Rozwiązanie istnieje dla dowolnych wartości zmiennych $x_{m+1}, \ldots, x_n$ (i dowolnego wektora $\mathbf{b}$ - dlaczego? sprawdźcie Państwo swoje rozumienie algebry!). W rezultacie otrzymujemy jako rozwiązanie związki

$$
\begin{aligned}
x_1 &= f_1(\mathbf{b}, x_{m+1}, \ldots, x_n)\,, \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
x_m &= f_m(\mathbf{b}, x_{m+1}, \ldots, x_n)\,,
\end{aligned}
$$

czyli właśnie jawną funkcję $f : \mathbb{R}^{n-m} \longrightarrow \mathbb{R}^m$ (wektor $\mathbf{b}$ jest tu nieistotny dla nas; inny wektor $\mathbf{b}$, to po prostu inne odwzorowanie F i inne zatem f). Przykład ten, jakkolwiek szczególny - bo daje się tu odwzorowanie f napisać jawnie - pokazuje też, że nie zawsze da się np. dostać funkcję

$$
\begin{aligned}
x_{n-m+1} &= \tilde{f}_1(\mathbf{b}, x_1, \ldots, x_{n-m})\,, \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
x_n &= \tilde{f}_m(\mathbf{b}, x_1, \ldots, x_{n-m})\,,
\end{aligned}
$$

bo może sie zdarzyć, że minor stopnia m utworzony z m ostatnich kolumn macierzy problemu jest akurat zerowy.

Widać też, że można by tu było założyć, iż rząd macierzy problemu jest równy $r < m$. Przy założeniu, że macierz rozszerzona (z dostawioną kolumną wektora $\mathbf{b}$) też jest rzędu r było by wtedy po prostu $r < m$ niezależnych równań na n niewiadomych i, zakładając, że r pierwszych kolumn i r pierwszych wierszy macierzy tworzy niezerowy minor stopnia r, (zgodnie z procedurą opisaną w skrypcie do algebry) odrzucilibyśmy $m - r$ ostatnich równań otrzymując, po wykorzystaniu kramersiąt, jawne odwzorowanie $f : \mathbb{R}^{n-r} \longrightarrow \mathbb{R}^r$

$$
\begin{aligned}
x_1 &= f_1(\mathbf{b}, x_{r+1}, \ldots, x_n)\,, \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
x_r &= f_r(\mathbf{b}, x_{r+1}, \ldots, x_n)\,.
\end{aligned}
$$

W przykładzie powyższym zawsze się daje otrzymać jawnie m (albo r) jakichś funkcji $n - m$ (albo $n - r$) zmiennych. Wobec tego pochodne można już potem sobie obliczać normalnie. W ogólnym przypadku jednak, kiedy odwzorowanie $F : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ $(m < n)$ nie jest liniowe, naogół odpowiednich związków nie daje się jawnie rozwikłać, ale daje się powiedzieć, że w otoczeniu jakiegoś punktu $\mathbf{x} \in D$ odwzorowanie F zadaje jakąś uczciwą funkcję f $n-m$ (lub $n-r$) którychś x-ów w m (lub r; ale sytuacje, w których nie wszystkie równania są niezależne wykluczyliśmy z założenia, więc już dalej nie będziemy mącić tym ogólniejszym przypadkiem) pozostałych x-ów, czyli właśnie funkcję $f : \mathbb{R}^{n-m} \longrightarrow \mathbb{R}^m$, i daje się obliczać pochodne takiej funkcji i to dowolnego rzędu bez operowania jawnymi wzorami na funkcję f. Chodzi więc o to, by umieć powiedzieć, w otoczeniu których punktów należących do "poziomicy zerowej" $F^{-1}(\mathbf{0}) \subset D$, czyli do przeciwobrazu zera przestrzeni $\mathbb{R}^m$, definiowanej przez dane odwzorowanie $F : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ odwzorowanie to po pierwsze definiuje funkcję $f : \mathbb{R}^{n-m} \longrightarrow \mathbb{R}^m$ i po drugie, by umieć obliczać pochodne takich funkcji. Po co to fizykowi? A no np. dlatego, że cała termodynamika fenomenologiczna - przepiękna i fascynująca teoria, którą mam okazję dręczyć studentów na wykładzie na czwartym roku - wykorzystuje odwzorowania uwikłane. Np. zadana jest entropia S jako funkcja $S = f(U, V, n)$, a my chcemy mieć raczej funkcję $U = g(S, V, n)$ i ją różniczkować. Wtedy pytaniem jest, czy wszędzie $F(S, U, V, n) \equiv f(U, V, n) - S$ - funkcja z $\mathbb{R}^4$ w $\mathbb{R}^1$ - zadaje, poprzez warunek $F(S, U, V, n) = 0$, funkcję $U = g(S, V, n)$, czyli właśnie funkcję z $\mathbb{R}^3$ w $\mathbb{R}$ i jak obliczać pochodne U, nie umiejąc jawnie napisać na U wzoru. W termodynamice naogół milcząco zakłada się, że wszędzie tak się daje i to w dowolną stronę, np. że daje się też napisać $V = h(S, U, n)$. Ale matematyk musi wszystko mieć udowodnione... Jeszcze przykładzik z termodynamiki. Wszyscy znają równanie stanu gazu doskonałego $p\,V = nRT$. Poniewa jest ono proste,[28] można z niego jawnie wyznaczyć p jako funkcję V, T i n, albo wyznaczyć V jako funkcję p, T i n i sobie te wielkości różniczkować np. żeby obliczyć mierzalne współczynniki izotermicznej ściśliwości i termicznej rozszerzalności objętościowej

$$k_T \equiv -\frac{1}{V}\left(\frac{\partial V}{\partial p}\right)_{T,n}, \qquad \alpha_p \equiv \frac{1}{V}\left(\frac{\partial V}{\partial T}\right)_{p,n}.$$

Ale są bardziej skomplikowane równania stanu. Np. równanie VdW (Van der Waalsa)

$$\left(p + \frac{an^2}{V^2}\right)(V - nb) = nRT\,,$$

I przy takim równaniu obliczenie powyższych współczynnkików, to już nie jest takie oczywiste...

W każdym razie ważne jest żeby sobie uświadomić: a) to jest użyteczne, b) to *nie jest trudne!* Zdrowy fizyczny rozum wystarcza, żeby sobie z tym dawać radę. Żeby nie komplikować zbytnio, będziemy się niżej zajmować prawie wyłącznie przypadkiem, $m = 1$, tj. kiedy odwzorowanie F zbioru D (którym może być całe $\mathbb{R}^n$ lub jakiś duży otwarty podzbiór $\mathbb{R}^n$) jest dane jedną funkcją $F(x_1, \dots, x_n)$. Na początek będziemy się zajmować przypadkiem $n = 2$, czyli będziemy badać, kiedy $F(x, y) = 0$ zadaje funkcję $y = y(x)$, albo funkcję $x = x(y)$. Zbiór $E \subset D \subset \mathbb{R}^2$ można wtedy hieroglifami zapisać tak

$$E = \{(x, y) \in \mathbb{R}^2 :\, F(x, y) = 0\}\,.$$

Np. jeśli $F(x, y) = x^2 + y^2 + 1$, to E jest zbiorem pustym i sprawa jest bezprzedmiotowa; jeśli zaś $F(x, y) = x^2 + y^2/4 - 1$, to zbiór E jest elipsą. Pytamy w tym przypadku, kiedy przez punkt $(x_0, y_0) \in E$ przechodzi krzywa, dająca się zapisać jako ciągła funkcja $y = y(x)$, której wszystkie punkty należą do zbioru E. Czy jest jedna taka krzywa, czy

[28]Na drugim roku jest jakiś wykład niby z termodynamiki i po tym wykładzie studenci myślą, że wszystko jest gazem doskonałym i wszyscy chodzą jak (doskonale) zagazowani... A termodynamika to jest teoria obejmująca wszystkie układy fizyczne, nie tylko gazy!

może kilka? To samo pytanie można postawić zastępując $y = y(x)$ funkcją $x = x(y)$. Na tym drugim przykładzie widać wszystko jasno: możemy napisać

$$y = \sqrt{4 - 4x^2}\,, \qquad \text{lub} \qquad y = -\sqrt{4 - 4x^2}\,,$$

i wiemy, że przez dany punkt (x_0, y_0) leżący na rzeczonej elipsie przechodzi albo jedna z tych krzywych albo druga. Wyjątki stanowią punkty $(1, 0)$ i $(-1, 0)$ bo tam wykres elipsy na płaszczyźnie xy jest pionowy i nie jest wykresem żadnej funkcji $y = y(x)$ (inaczej: dwie funkcje zdefiniowane powyższymi wzorami się w tych punktach łączą). Za to przez te dwa punkty spokojnie przechodzą wykresy funkcji $x = x(y)$: funkcji $x = \sqrt{1 - y^2/4}$ przez pierwszy i funkcji $x = -\sqrt{1 - y^2/4}$ przez drugi, które to funkcje są uczciwymi ciągłymi funkcjami wszędzie oprócz punktów $(0, -2)$ i $(0, 2)$ elipsy.

Ogólnie sprawę załatwia takie twierdzenie (dowód był na wykładzie, ale skupmy się i przeczytajmy z uwagą): jeśli odwzorowanie $F(x, y)$ ma w otoczeniu punktu $(x_0, y_0) \in E$, tj. takiego, że $F(x_0, y_0) = 0$, ciągłe pochodne cząstkowe F_x i F_y i ta druga pochodna *nie znika* w tym punkcie, tj. $F_y(x_0, y_0) \neq 0$, to w D istnieje (choćby nie wiem jak malutkie) takie otoczenie otwarte punktu (x_0, y_0), w którym odwzorowanie F wyznacza jedną i tylko jedną funkcję $y = y(x)$, która jest ciągła i jej pochodna y' jest w tym otoczeniu dana (na pierwszy rzut oka dziwnym, ale jak zaraz zobaczymy zupełnie oczywistym) wzorem

$$y'(x) = -\left.\frac{F_x(x, y)}{F_y(x, y)}\right|_{y=y(x)}\,.$$

Dziwny dopisek $_{y=y(x)}$ oznacza, że prawą stronę należy obliczyć nie w dowolnym punkcie (x, y) tego otoczenia, tylko w punkcie należącym do zbioru E (czyli do przecięcia E z tym otoczeniem). Z tego wzoru widać, że warunek $F_y(x_0, y_0) \neq 0$ jest tu jakoś kluczowy (z ciągłości funkcji F i jej pochodnych wynika, że jak $F_y(x_0, y_0) \neq 0$ to $F_y \neq 0$ także w pewnym otwartym otoczeniu punktu (x_0, y_0) więc powyższy wzór jest dobry w tym otoczeniu). Oczywiście wszystko można odwrócić i jeżeli $F_x(x_0, y_0) \neq 0$, to w pewnym otoczeniu punktu (x_0, y_0) istnieje analogiczna jednoznaczna funkcja $x = x(y)$ i $x'(y) = -F_y(x, y)/F_x(x, y)|_{x=x(y)}$.

Przykłady

1) Niech $F(x, y) = -10 + \exp(2x - 3y)$. To jest przykład, w którym można rozwiązać warunek $F(x, y) = 0$ i dostać jawnie funkcję $y = y(x)$:

$$y(x) = \frac{2}{3}\,x - \frac{1}{3}\ln 10\,,$$

więc $y'(x) = \frac{2}{3}$. Ale zobaczmy, jak działa podany wzór na pochodną funkcji zadanej w sposób uwikłany. Pochodne cząstkowe funkcji $F(x, y)$

$$F_x(x, y) = 2\,e^{2x-3y}\,, \qquad F_y(x, y) = -3\,e^{2x-3y}\,,$$

nie zerują się nigdzie (na całym $\mathbb{R}^2$) więc zgodnie z twierdzeniem przez każdy punkt zbioru $E = F^{-1}(0)$ przechodzi dokładnie jedna krzywa $y = y(x)$ (i oczywiście dokładnie jedna krzywa $x = x(y)$). Pochodna funkcji $y = y(x)$ zadanej równością $F(x, y) = 0$, którą daje podany wcześniej wzór jest równa

$$y'(x) = -\frac{F_x(x,y)}{F_y(x,y)} = \frac{2}{3}\,,$$

i jest oczywiście tą samą pochodną, co obliczona bezpośrednio.
2) Bardziej skomplikowany przykład. Niech

$$F(x,y) = x\,2^y - x^2y^2 + (1-x^2)\sin y\,.$$

Pytamy, czy przez punkt $(0, 0)$ przechodzi jakaś krzywa będąca wykresem funkcji $y = y(x)$ zadanej warunkiem $F(x, y) = 0$? A jak tak, to jaka jest pochodna tej funkcji? Najpierw sprawdzamy, czy punkt $(0, 0)$ w ogóle należy do zbioru E (czyli, czy pytanie ma sens, albo, w języku filozofów, czy jest prawomocne). Należy, bo $F(0, 0) = 0$. To teraz trzeba sprawdzić pochodną cząstkową F_y:

$$F_y(x,y) = x\,2^y \ln 2 - 2x^2y + (1-x^2)\cos y\,,$$

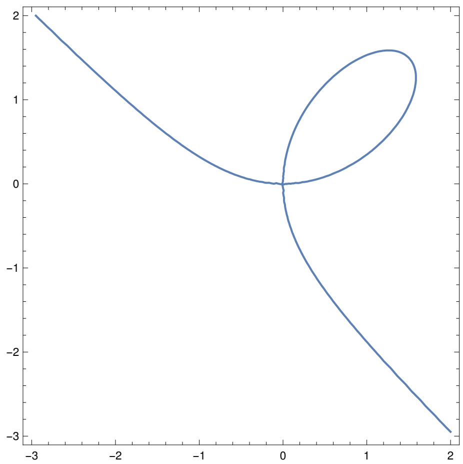
i $F_y(0, 0) = 1$. Ponieważ w punkcie $(0, 0)$ pochodna F_y nie znika i punkt ten należy do $F^{-1}(0)$, przechodzi przezeń dokładnie jedna krzywa będąca w jakimś otoczeniu tego punktu wykresem uczciwej funkcji $y = y(x)$, choć jawnie nie jesteśmy w stanie napisać wzoru na tę funkcję. Poza tym, $y(0) = 0$. No i, ponieważ $F_x(x, y) = 2^y - 2xy^2 - 2x \sin y$, więc w tym otoczeniu punktu $(0, 0)$

$$y'(x) = -\frac{2^y - 2xy^2 - 2x\sin y}{x\,2^y \ln 2 - 2x^2y + (1-x^2)\cos y}\,.$$

Po lewej stronie y' jest napisane jako funkcja x tylko, a po prawej występuje i x i y. Należałoby oczywiście po prawej stronie wstawić $y = y(x)$, ale nie potrafimy tego praktycznie zrobić, bo jak już mówiliśmy nie jesteśmy w stanie rozwikłać wzory $F(x, y) = 0$ względem y. No coż, takie jest życie z funkcjami uwikłanymi... Ale jedno zawsze można: można warunek $F(x, y) = 0$ wykorzystać żeby jakoś sobie prawą stronę inaczej zapisać. Np. możemy zawsze napisać $x\,2^y = x^2y^2 - (1 - x^2)\sin y$ i przekształcić powyższy wzór na pochodną w

$$y'(x) = -\frac{2^y - 2xy^2 - 2x\sin y}{(x^2y^2 - (1-x^2)\sin y)\ln 2 - 2x^2y + (1-x^2)\cos y}\,.$$

Jest to równie dobra i całkowicie równoważna postać tej pochodnej w tym sensie, że jeśli prawe strony tych dwu wzorów obliczymy w jakimś punkcie (x_0, y_0) należącym do E, czyli takim, że $F(x_0, y_0) = 0$, to oba wzory dadzą tę samą liczbę (pochodną), czyli to samo nachylenie krzywej $y = y(x)$ w punkcie x_0, w którym $y(x_0) = y_0$. Oczywiście, ponieważ w $(0, 0)$ nie znika też pochodna $F_x(x, y)$, można krzywą wyznaczaną przez warunek $F(x, y) =$



Rysunek 10: Liść Kartezjusza (feuille de Descartes), czyli krzywa będąca zbiorem E punktów spełniających równość $x^3 + y^3 - 3xy = 0$.

0 traktować w otoczeniu tego punktu jak funkcję $x = x(y)$ i wtedy pochodna tej funkcji jest równa

$$x'(y) = -\frac{F_y(x,y)}{F_x(x,y)} = -\frac{x\,2^y \ln 2 - 2x^2y + (1-x^2)\cos y}{2^y - 2xy^2 - 2x\sin y}$$
$$= -\frac{(x^2y^2 - (1-x^2)\sin y)\ln 2 - 2x^2y + (1-x^2)\cos y}{2^y - 2xy^2 - 2x\sin y}\,.$$

3) Przykład sztandarowy, bardzo pouczający, wzięty z Lejka. Niech

$$F(x,y) = x^3 + y^3 - 3xy\,.$$

Zbiór punktów $(x,y) \in \mathbb{R}^2$ spełniających równość $F(x,y) = 0$ jest tzw. liściem Kartezjusza. Jest jasne, że gdy $|x| \gg 1$, warunek $F(x,y) = 0$ ma jedno rozwiązanie rzeczywiste $y \approx -x$ (pozostałe dwa rozwiązania - bo to w końcu jest równanie trzeciego stopnia na y - są przy takich x-ach zespolone). Mniej oczywiste, ale można też do tego dojść rozumem, jest to, że tak jest zawsze, gdy $x < 0$. Zato gdy $x > 0$, ale x nie jest za duże, istnieją trzy rzeczywiste rozwiązania: wygląda to tak jak na rysunku 10: krzywa idąca od góry od ujemnych x-ów przechodzi przez $(0,0)$, robi w pierwszej ćwiartce płaszczyzny xy zawijasa, ponownie wraca do punktu $(0,0)$ i dalej już biegnie, gdy x rośnie od zera, w dół czwartej ćwiartki. Jest więc przynajmniej od razu jasne, że punkt $(0,0)$ jest jakoś trefny, bo wychodzą z niego cztery linie będące obrazem zbioru E, a nie dwie, i te jakby dwie krzywe przechodzące przez punkt $(0,0)$ są w każdym, nawet najmniejszym *otwartym* otoczeniu tego punktu. To zobaczmy, jak to wygląda z punktu widzenia twierdzenia o funkcji uwikłanej? Pochodna cząstkowa po y funkcji $F(x,y)$

$$F_y(x,y) = 3\,(y^2 - x)\,,$$

zeruje się na całej paraboli $x = y^2$, ale z punktu widzenia istnienia funkcji uwikłanej nas interesują tylko te zera pochodnej F_y, które są punktami zbioru E, czyli są rozwiązaniami układu dwóch równań

$$F_y(x,y) = 3\,(y^2 - x) = 0\,,$$
$$F(x,y) = x^3 + y^3 - 3xy = 0\,.$$

Podstawiając z pierwszego $x = y^2$ do drugiego dostajemy równanie $y^3(y^3 - 2) = 0$. Zatem istotnymi z punkt widzenia twierdzenia o funkcji uwikłanej punktami są $(0,0)$ i $(2^{2/3},\,2^{1/3})$, będące rozwiązaniami tego układu dwóch równań. W należącym do E punkcie $(2^{2/3},\,2^{1/3})$ zeruje się F_y więc przez ten punkt nie przechodzi uczciwa funkcja $y = y(x)$: istotnie: to jest na tym zawijasie, który wykres zbioru E robi w pierwszej ćwiartce, punkt położony najbardziej na prawo - krzywa reprezentująca zbiór E "staje tam dęba", czyli w samym punkcie $(2^{2/3},\,2^{1/3})$ biegnie dokładnie pionowo. To nie może być więc w tym punkcie funkcja $y = y(x)$! Ale oczywiście taka krzywa może jak najbardziej być wykresem uczciwej funkcji $x = x(y)$. I rzeczywiście: w punkcie $(2^{2/3},\,2^{1/3})$ nie znika pochodna F_x, $F_x(2^{2/3},\,2^{1/3}) \neq 0$ i twierdzenie o istnieniu funkcji uwikłanej nie zabrania przechodzenia przez ten punkt funkcji $x = x(y)$ - przeciwnie, mówi ono, że przez ten punkt i w jego otoczeniu biegnie jedna i tylko jedna funkcja $x = x(y)$. Analogiczny jest punkt $(2^{1/3},\,2^{2/3})$, również należący do E (jest to punkt położony najwyżej na krzywej z rysunku 10 w pierwszej ćwiartce), w którym zeruje się na odmianę F_x, a $F_y \neq 0$: w tym punkcie (i jego otoczeniu) istnieje funkcja $y = y(x)$, ale nie istnieje funkcja $x = x(y)$. Natomiast charakter punktu $(0,0)$ jest inny. Łatwo zobaczyć, że w tym punkcie zeruje się również F_x. Skoro więc w punkcie $(0,0)$ należącym do zbioru E zerują się obie pochodne, F_x i F_y, to z twierdzenia wynika, że nie może przez ten punkt przechodzić ani funkcja $y = y(x)$, ani funkcja $x = x(y)$. Tak właśnie odzwierciedla się zauważona wyżej "trefność" punktu $(0,0)$. bardziej matematycznie rzecz ujmując, jest to punkt osobliwy (ale dlaczego "osobliwy" ma być lepiej niż "trefny"?).

Zajmijmy się teraz zrozumieniem, skąd się bierze ten dziwny wzór na pochodną y' funkcji $y = y(x)$ zadanej w sposób uwikłany. Chodzi oczywiście o to, że jak się ten wzór rozumie, to nie trzeba nic pamiętać - można go zawsze natychmiast sobie odtworzyć (i życie staje się lżejsze). Można ten wzór zrozumieć na przynajmniej dwa sposoby. Po pierwsze, jeśli funkcja $y = y(x)$ jest zadana przez równość $F(x,y) = 0$ (w otoczeniu jakiegoś przyzwoitego punktu) to znaczy, że $F(x,\,y(x)) \equiv 0$ - wyrażenie po lewej jest, jako funkcja x-a, tożamościowo zerowe. Liczymy pochodną tej tożsamościowo zerowej funkcji x-a (czyli pochodna też jest zero tożsamościowo) stosując regułki, które już znamy:

$$\begin{aligned} 0 = \frac{d}{dx}\,F(x,\,y(x)) &= \left.\frac{\partial F(x,y)}{\partial x}\right|_{y=y(x)} + \left.\frac{\partial F(x,y)}{\partial y}\right|_{y=y(x)} \frac{dy}{dx} \\ &\equiv F_x(x,\,y(x)) + F_y(x,\,y(x))\,\frac{dy}{dx}\,. \end{aligned}$$

Wyliczamy stąd $y' = dy/dx$ i mamy ten "dziwny" wzór.

Drugi sposób widzenia skąd się bierze wzór na pochodną $y'(x)$ jest bardziej zabawny, ale wielce kształcący. Będzie on nam bardzo pomocny przy rozumieniu równań różniczkowych, a poza tym różne tożsamości termodynamiczne (że znów wskoczę na ulubionego konika), które studentów straszą po nocach, dają się w ten sposób zrozumieć i całkowicie oswoić. Mamy sobie funkcję dwóch zmiennych $F(x,y)$ i jesteśmy na płaszczyźnie xy w jakimś punkcie $(x_0,y_0)\in E$ czyli takim, że $F(x_0,y_0)=0$. Piszemy teraz zwykłą dwuwymiarową różniczkę funkcji F w tym punkcie

$$dF = F_x(x_0,y_0)\,dx + F_y(x_0,y_0)\,dy\,.$$

Różniczka dF mówi nam, o ile mniej więcej - pamiętamy, to jest główna liniowa część prawdziwej zmiany ΔF - wartość funkcji F zmieni się, gdy z (x_0,y_0) (gdzie $F=0$) przesuniemy się do jakiegoś innego punktu $(x_0+dx,\ y_0+dy)$. Na ogół jakoś się zmieni. Ale możemy zapytać, jak kroczek dy w kierunku y musi być skorelowany z kroczkiem dx w kierunku x, by wartość funkcji F pozostała ta sama, co w (x_0,y_0) (czyli równa zeru)? Tzn. jak muszą te składowe wektora przesunięcia być ze sobą skorelowane, by przesunąć się z (x_0,y_0) do sąsiedniego punktu zbioru E? No, po prostu tak, żeby $dF=0$, czyli

$$\left.\frac{dy}{dx}\right|_{x_0,y_0} = -\frac{F_x(x_0,y_0)}{F_y(x_0,y_0)}\,.$$

A ten stosunek dy do dx to jest właśnie to, co zwiemy pochodną y' funkcji $y=y(x)$. I tak musi być w każdym punkcie zbioru E, w otoczeniu którego istnieje i przez który przechodzi funkcja $y=y(x)$. Widać też, co się dzieje w punktach trefnych: jeśli $F_y=0$, to, musi być $dx=0$ czyli trzeba z (x_0,y_0) się przesunąć pionowo, żeby wartość F się nie zmieniła i dlatego nie istnieje tam funkcja $y=y(x)$, zato istnieje funkcja $x=x(y)$ (i ma w tym punkcie zerową pochodną). Jeśli zaś i $F_x=0$ i $F_y=0$, to niema żadnej korelacji dx-a z dy-kiem i można się przesunąć w dowolnym kierunku więc w sumie nie wiadomo, gdzie i dlatego taki punkt jest trefny. Zobaczymy też, że taka sama jest przyczyna zbiegania się w niektórych punktach krzywych całkowych równań różniczkowych pierwszego rzędu (jak się na te równania patrzy w sposób fizyczny, czyli tak jak tu, a nie poprzez te matematyczne hieroglify).

Zauważmy jeszcze, że ostatni wzór w termodynamice zapisalibyśmy jako

$$\left(\frac{\partial y}{\partial x}\right)_F = -\frac{F_x}{F_y} \equiv -\frac{(\partial F/\partial x)_y}{(\partial F/\partial y)_x}\,.$$

Te subskrypciki (fuj, jaki anglicyzm!) $_F$, $_x$, $_y$ w termodynamice się dopisuje, żeby pamiętać, która zmienna jest trzymana stała. Powstaje z tego taki "szokujący" wzorek

$$\left(\frac{\partial x}{\partial y}\right)_F \left(\frac{\partial F}{\partial x}\right)_y \left(\frac{\partial y}{\partial F}\right)_x = -1\,.$$

W moich wykładach z termodynamiki - zgodnie z moją manierą nadawania wszystkiemu oswajających nazw - zwie się to "the shocking relation" - relacją szokującą. No ale teraz już niema w niej nic szokującego.

Idąc za ciosem wyprowadźmy jeszcze wzorek na drugą pochodną funkcji $y=y(x)$ w punkcie, w którym ta funkcja istnieje. $y=y(x)$ ma w otoczeniu takiego punktu i drugą pochodną, jeśli drugie pochodne cząstkowe F_{xx}, F_{xy} i F_{yy} w otoczeniu tego punktu istnieją i są ciągłe. Co to jest druga pochodna? No, jest to pochodna pierwszej pochodnej. Zatem

$$\frac{d}{dx}\,y'(x) = \frac{d}{dx}\left[-\frac{F_x(x,\,y(x))}{F_y(x,\,y(x))}\right] = -\frac{1}{F_y}\left(\frac{d}{dx}\,F_x(x,\,y(x))\right) + \frac{F_x}{F_y^2}\left(\frac{d}{dx}\,F_y(x,\,y(x))\right),$$

i dalej łańcuszkowo:

$$\frac{d}{dx}\,y'(x) = -\frac{1}{F_y}\left(F_{xx} + F_{xy}\,\frac{dy}{dx}\right) + \frac{F_x}{F_y^2}\left(F_{yx} + F_{yy}\,\frac{dy}{dx}\right),$$

teraz przypominamy sobie, że $dy/dx = y' = -F_x/F_y$, wstawiamy i po małym uporządkowaniu mamy

$$\frac{d}{dx}\,y'(x) \equiv y''(x) = -\left.\frac{F_{xx}\,F_y^2 - 2F_{xy}\,F_x\,F_y + F_{yy}\,F_x^2}{F_y^3}\right|_{(x,y=y(x))}.$$

Prawda, że śliczny wzorek? Nawet da się go zapamietać, ale po co? Przecież można go zawsze w minutkę wyprowadzić...

Możemy ten wzorek sprawdzić. Weźmy pierwszy przerabiany przykład, czyli odwzorowanie $F(x,y) = -10 + \exp(2x - 3y)$. Warunek $F(x,y) = 0$ definiuje funkcję:

$$y(x) = \frac{2}{3}\,x - \frac{1}{3}\ln 10\,,$$

więc $y''(x) \equiv 0$. A co da wyprowadzony wyżej wzór? Pamiętamy, że

$$F_x(x,y) = 2\,e^{2x-3y}\,, \qquad F_y(x,y) = -3\,e^{2x-3y}\,,$$

i wobec tego

$$F_{xx}(x,y) = 4\,e^{2x-3y}\,, \qquad F_{yy}(x,y) = 9\,e^{2x-3y}\,, \qquad F_{xy}(x,y) = -6\,e^{2x-3y}\,.$$

Czyli:

$$F_{xx}\,F_y^2 - 2F_{xy}\,F_x\,F_y + F_{yy}\,F_x^2 = \left[4\cdot(-3)^2 - 2\cdot(-6)\cdot(2)\cdot(-3) + 9\cdot(2)^2\right]\left(e^{2x-3y}\right)^3 = 0\,.$$

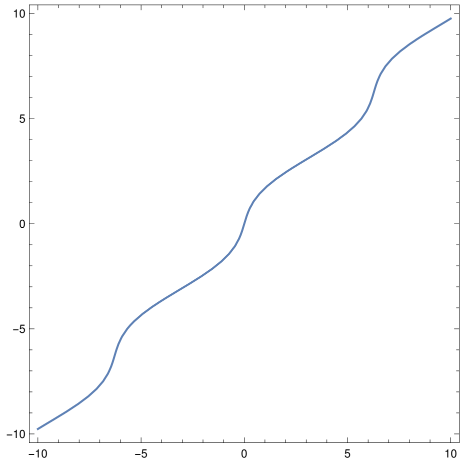
Zgadza się.

Inny przykład. Znajdźmy pierwszą i drugą pochodną funkcji $y = y(x)$ zadanej w sposób uwikłany warunkiem $F(x,y) = x\,e^y - y + 1 = 0$ i wyraźmy te pochodne przez zmienną y tylko. (Po co? A tak, dla wprawy). Obliczamy:

$$F_x(x,y) = e^y\,, \qquad F_y(x,y) = x\,e^y - 1\,,$$

więc

$$y'(x) = -\frac{F_x}{F_y} = \frac{e^y}{1 - x\,e^y} = \frac{e^y}{2 - y}\,.$$



Rysunek 11: Rysunek zbioru $E \subset \mathbb{R}^2$ zadanego równością $y - \varepsilon \sin y - x = 0$, gdy $\varepsilon = 0.7$.

W ostatnim kroku w mianowniku wykorzystaliśmy (już wiemy, że wolno) to, że na krzywej $x\,e^y - y + 1 = 0$, czyli $x\,e^y = -1 + y$. Teraz drugie pochodne:

$$F_{xx}(x,y) = 0\,, \qquad F_{yy} = x\,e^y\,, \qquad F_{xy}(x,y) = e^y\,.$$

To do tego ślicznego wzoru:

$$y''(x) = -\frac{1}{F_y^3}\left\{x\,e^y\,(e^y)^2 - 2\,e^y\,e^y\,(x\,e^y - 1)\right\} = -\frac{e^{2y}}{F_y^3}\,(2 - x\,e^y) = -\frac{e^{2y}}{F_y^3}\,(3 - y)\,,$$

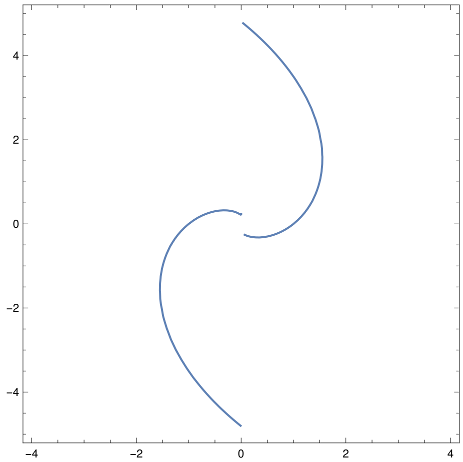
no i jeszcze w mianowniku $F_y = x\,e^y - 1 = y - 2$. Ostatecznie

$$y''(x) = \frac{3-y}{(2-y)^3}\,e^{2y}\,.$$

Zadanie Uwik.1
Jaka może być wartość parametru ε, żeby wzórek $y - \varepsilon \sin y = x$ definiował *globalnie*, czyli na całym $\mathbb{R}$ określoną funkcję $y = y(x)$?
Rozwiązanie: Niewątpliwie jeśli $\varepsilon = 0$, jest to przyzwoita funkcja $y(x) = x$. Ale jak $|\varepsilon|$ rośnie, to na tę prostą $y = x$ nakładają się falki (zob. rysunek 11) i przy zbyt dużym $|\varepsilon|$ te falki powodują, że jednemu x-owi już nie odpowiada jeden tylko y-ek. Pytanie, jak duży może być jeszcze $|\varepsilon|$, żeby tak nie było? Napiszmy $F(x,y) = y - \varepsilon \sin y - x$ i zastosujmy to, co już umiemy, czyli twierdzenie o funkcji uwikłanej. Warunkiem, by $F(x,y) = 0$ wyznaczało wszędzie (dla wszystkich $x \in \mathbb{R}$) przyzwoitą funkcję $y = y(x)$ jest, by nigdzie na zbiorze E nie zerowała się pochodna F_y. Ale

$$F_y(x,y) = 1 - \varepsilon \cos y\,,$$

i widać, że nie może się ona zerować, jeśli $|\varepsilon| < 1$ (bo cosinus nie bywa większy niż 1). Ot i wszystko.



Rysunek 12: Rysunek zbioru $E \subset \mathbb{R}^2$ zadanego równością $\ln\sqrt{x^2+y^2} - \mathrm{arctg}(y/x) = 0$.

Zadanie Uwik.2
W których punktach płaszczyzny xy odwzorowanie

$$F(x,y) = \ln\sqrt{x^2+y^2} - \mathrm{arctg}\frac{y}{x},$$

nie zadaje, poprzez warunek $F(x,y)=0$ funkcji $y=y(x)$?
Rozwiązanie: Znów obliczamy pochodne F:

$$F_x(x,y) = \frac{x}{x^2+y^2} - \left(-\frac{y}{x^2+y^2}\right) = \frac{x+y}{x^2+y^2},$$
$$F_y(x,y) = \frac{y}{x^2+y^2} - \left(\frac{x}{x^2+y^2}\right) = \frac{-x+y}{x^2+y^2}.$$

Widać, że F_y zeruje się na linii $y=x$, ale oczywiście interesują nas tylko te punkty, w których linia ta przecina się ze zbiorem E wyznaczanym przez warunek $F(x,y)=0$. Szukamy zatem rozwiązań równania $F(x,x)=0$, czyli

$$\ln\sqrt{2x^2} = \frac{\pi}{4}.$$

Zatem punkty w których $F(x,y)=0$ nie definiuje funkcji $y=y(x)$, to

$$x = y = \pm\frac{1}{\sqrt{2}}\,e^{\pi/4}.$$

A jak w ogóle wygląda zbiór E? Jeśli przejdziemy do zmiennych biegunowych, to naiwnie można by sądzić, że $F(x,y)=0$ daje po prostu $\ln r = \varphi$, czyli spiralę $r(\varphi) = \exp(\varphi)$. Ale to nie jest do końca tak, bo jak się tak napisze, to się wydaje, że zmienna φ może się zmieniać od $-\infty$ do $+\infty$. Tymczasem, żeby funkcja $F(x,y)$ była dobrze określoną funkcją trzeba się zdecydować na jakąś gałąź funkcji arctg. Konwencjonalnie przyjmuje się, że $\operatorname{arctg}(0) = 0$ (inne gałęzie tej funkcji odpowiadają przyjęciu $\operatorname{arctg}(0) = k\pi$ z jakimś całkowitym $k \neq 0$) i tak sobie przyjmijmy. Wtedy arctg zmienia się od $-\pi/2$ do $+\pi/2$. Zatem we wzorze $r = \exp(\varphi)$ w zmiennych biegunowych $-\pi/2 \leq \varphi \leq \pi/2$, ale to dotyczy wtedy tylko półpłaszczyzny $x > 0$; drugiej półpłaszczyzny w taki sposób nie obejmiemy, a $F(x, y) = 0$ ma i tam rozwiązanie. Teraz zobaczmy, jaki się dostaje y z warunku $F(x, y) = 0$, gdy $x \to 0^+$. Są dwie możliwości: jest rozwiązanie z $y > 0$ - wtedy $\operatorname{arctg}(y/x)$ przy skończonym y dąży do $\pi/2$ i widać, że rozwiązaniem warunku $\ln\sqrt{y^2} = \operatorname{arctg}(y/x) \to \pi/2$ jest $y = \exp(\pi/2)$. Jest też i drugie rozwiązanie z $y < 0$: wtedy $\operatorname{arctg}(y/x)$ przy skończonym y dąży do $-\pi/2$ i widać, że rozwiązaniem warunku $\ln\sqrt{y^2} = \operatorname{arctg}(y/x) \to -\pi/2$ jest $y = -\exp(-\pi/2)$. Jest też jasne, że dwa te punkty: $(0, -e^{-\pi/2})$ i $(0, e^{\pi/2})$ muszą być końcami krzywej $r = \exp(\varphi)$ i że pierwszy odpowiada $r = e^{-\pi/2}$, $\varphi = -\pi/2$, a drugi $r = e^{\pi/2}$, $\varphi = \pi/2$. Analogicznie, gdy $x \to 0^-$, jest rozwiązanie ujemne $y = -\exp(\pi/2)$ i dodatnie $y = \exp(-\pi/2)$. Warunek $F(x, y) = 0$ wyznacza więc na płaszczyźnie xy dwie rozłączne krzywe: jedną położoną w obszarze $x \geq 0$ i drugą w obszarze $x \leq 0$. Całość wygląda tak, jak na rysunku 12. Znalezione wyżej punkty, w których $F(x, y) = 0$ nie definiuje funkcji $y = y(x)$ są to: punkt najbardziej położony na prawo na prawej gałęzi rysunku 12 i punkt położony najbardziej na lewo na lewej gałęzi tego rysunku. W obu tych punktach krzywe "stają dęba". Z kolei $F(x, y) = 0$ nie definiuje funkcji $x = x(y)$ w punktach

$$x = -y = \pm\frac{1}{\sqrt{2}}\, e^{-\pi/4}\,.$$

położonych na przecięciu krzywych z rysunku 12 z prostą $y = -x$. W punktach tych zeruje się pochodna F_x, a krzywe z rysunku 12 mają zerowe nachylenie w stosunku do osi x.

Zadanie Uwik.3

Znaleźć punkty krytyczne funkcji (liczba pojedyńcza lub mnoga - zaraz zostanie wyjaśnione dlaczego) $y = y(x)$ zadanej w sposób uwikłany warunkiem

$$F(x, y) = 8x^2 - 8xy + y^4 - 4y = 0\,,$$

i zbadać, czy są one jej ekstremami.

Rozwiązanie: W zasadzie wszystko sprowadza się do tego, co w szkole: trzeba znaleźć te x-y, w których zeruje się pochodna $y'(x)$, a potem zobaczyć, jaki jest znak drugiej pochodnej $y''(x)$ w takich punktach (lub, co tu mniej wygodne, sprawdzić czy pochodna $y'(x)$ zmienia znak i z jakiego na jaki przy przejściu przez taki punkt). Jedyna różnica z tym, co było w szkole (a może już tego w szkole niema? może to za trudne dla ministrów?) jest taka, że teraz mamy pewien szczególny sposób znajdywania pochodnych $y'(x)$ i $y''(x)$. No i jest jeszcze jedna sprawa: jeśli się znajdzie ze dwa punkty krytyczne funkcji (punkty, w których $y' = 0$), to naogół nie jest jasne, czy są to dwa punkty krytyczne tej samej funkcji $y = y(x)$, czy dwu różnych funkcji $y = y_1(x)$ i $y = y_2(x)$ zadawanych tym samym warunkiem. Czasem jest to łatwo ustalić. A czasem nie. Życie z funkcjami uwikłanymi ma swoje uroki...

Ponieważ $y'(x) = -F_x/F_y$, a funkcja $y = y(x)$ jest funkcją tylko tam, gdzie $F_y \neq 0$, przeto szukanie zer pochodnej y' sprowadza się do rozwiązania układu równań

$$\begin{aligned} F_x(x,y) &= 16x - 8y = 0\,, \\ F(x,y) &= 8x^2 - 8xy + y^4 - 4y = 0\,. \end{aligned}$$

Z pierwszego $y = 2x$, to do drugiego i dostajemy równanie

$$2x^4 - x^2 - x = 0\,.$$

Nie zrażamy się tym, że jest ono czwartego stopnia, tylko czujnie rzucamy na nie okiem i widzimy, że jednym z pierwiastków jest $x = 1$, a drugim oczywiście $x = 0$. Więc (nie zaczyna się zdania od "więc"!) piszemy

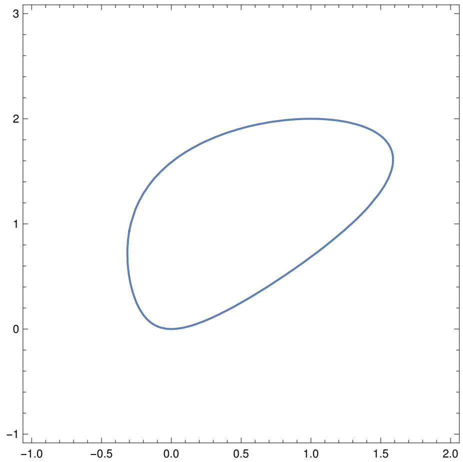
$$2x^4 - x^2 - x = x\,(x-1)\,(2x^2 + ax + 1) = 0\,,$$

i łatwo ustalamy, że $a = 2$. Ponieważ Δ dwumianu kwadratowego w nawiasie jest ujemna, innych rzeczywistych rozwiązań niż $x = 0$ i $x = 1$ niema. Zatem punkty, w których zeruje się (na zbiorze E) pochodna F_x to $(0,0)$ i $(1,2)$, czyli $y'(0) = 0$ w miejscu, gdzie $y(0) = 0$ i $y'(1) = 0$ w miejscu, gdzie $y(1) = 2$. No i właśnie: mamy dwa punkty krytyczne, ale bez dalszego wnikania w sprawę nie wiemy, czy to punkty krytyczne tej samej funkcji, czy dwóch różnych. Zato od razu mamy wartości funkcji w tych punktach. No to teraz drugie pochodne. W ogólności są one dany tym ślicznym (dla niektórych może strasznym?) wzorem $y'' = -(F_{xx}F_y^2 - 2F_{xy}F_xF_y + F_{yy}F_x^2)/F_y^3$. Ale tu przyjemna niespodzianka: w punktach krytycznych funkcji $y = y(x)$, w których chcemy znać pochodną y'', zeruje się F_x! Więc ten śliczno-straszny wzór sprowadza się do $y''|_{y'=0} = -F_{xx}/F_y$! A nas interesuje i tak tylko znak tego wyrażenia. Wszystko jest więc już banalnie proste: $F_{xx}(x,y) = 16$, $F_y(x,y) = 4y^3 - 8x - 4$ i w punkcie $(0,0)$ $-F_{xx}/F_y = 4$ - funkcja $y = y(x)$ ma tu minimum lokalne równe 0; z kolei w punkcie $(1,2)$ $-F_{xx}/F_y = -4/5$ i w tym punkcie krytycznym funkcja $y = y(x)$ ma lokalne maksimum równe 2. Jeśli sobie narysujemy zbiór E wyznaczony warunkiem $F(x,y) = 0$, to wyglada on tak, jak na rysunku 13. Widać z rysunku, że są to ekstrema dwóch różnych funkcji (albo jak kto woli, dwóch różnych gałęzi tej samej funkcji), bo to, że jest to ciągła jedna zamknięta krzywa nic nie znaczy: są jak widać dwa punkty (x_1, y_1) i (x_2, y_2), gdzie ta krzywa staje dęba - w tych punktach $F(x,y) = 0$ nie wyznacza funcji $y = y(x)$ i trzeba to właśnie interpretować tak, że w tych punktach x_1 i x_2, które nie należą do dziedziny żadnej z funkcji $y = y_1(x)$ i $y = y_2(x)$ te dwie różne funkcje (lub dwie gałęzie jednej funkcji) tylko się zbiegają.

Teraz uogólnimy te metody na więcej zmiennych, tj. na przypadek funkcji $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ zadanych w sposób uwikłany jako zerowa poziomica odwzorowania $F : \mathbb{R}^{n+1} \longrightarrow \mathbb{R}$. Niech

$$F(x,y,z) = e^z - xyz - 1\,.$$

Pytamy, kiedy warunek $F(x,y,z) = 0$ wyznacza np. funkcję $z = z(x,y)$? (Można też pytać o fukcję $y = y(x,z)$ lub $x = x(y,z)$). Odpowiedź jest oczywiście taka, że tam, gdzie $F_z \neq 0$, tzn. w tych punktach zbioru $E = F^{-1}(0) \subset \mathbb{R}^3$, w których $F_z(x,y,z) \neq 0$.



Rysunek 13: Rysunek zbioru $E \subset \mathbb{R}^2$ zadanego równością $8x^2 - 8xy + y^4 - 4y = 0$.

Obliczmy pochodne cząstkowe $\partial z/\partial x$ i $\partial z/\partial y$ funkcji $z = z(x, y)$ zdefiniowanej warunkiem $F(x, y, z) = 0$ w punkcie $(x_0, y_0, z_0) = (2, 1, 0)$. Punkt ten, jak łatwo zobaczyć, należy do zbioru E i w punkcie tym pochodna $F_z(x, y, z) = e^z - xy$ jest równa $1 - 2 = -1 \neq 0$, więc postawiony problem ma sens: w tym punkcie i w jego jakimś otwartym otoczeniu funkcja $z = z(x, y)$ istnieje. W celu obliczenia pochodnych $\partial z/\partial x$ i $\partial z/\partial y$ posłużymy się zdrowym fizycznym rozsądkiem. Napiszemy w tym punkcie różniczkę funkcji F:

$$dF = F_x(x_0, y_0, z_0)\, dx + F_y(x_0, y_0, z_0)\, dy + F_z(x_0, y_0, z_0)\, dz\,,$$

i zapytajmy znów, jak trzeba skorelować składowe dx, dy i dz wektora przemieszczenia, by nie "spaść" z poziomicy zerowej funkcji F, tj. by przemieścić się do sąsiedniego punktu zbioru E. Tzn. jak skorelować dx, dy i dz, by dostać $dF = 0$. Oczywiście to jest teraz jeden warunek na trzy składowe. Jeśli jednak "przytrzymamy" y, tzn., położymy $dy = 0$, to aby dostać $dF = 0$, musimy mieć

$$\frac{dz}{dx} = -\frac{F_x(x_0, y_0, z_0)}{F_z(x_0, y_0, z_0)}\,.$$

Ale ten stosunek po lewej jest dokładnie tym, czym jest pochodna cząstkowa funkcji $z(x, y)$ po x: stosunkiem zmiany dz wielkości z, gdy maciupko, o dx (myślmy o granicy $dx \to 0$), zmienimy x trzymając y ustalone. Analogiczne rozumowanie dotyczy stosunku dz/dy, gdy nie zmienia się x (czyli gdy $dx = 0$). W przypadku rozpatrywanej funkcji $F_x = -yz$, $F_y = -xz$ ale, że rozpatrujemy punkt $(2, 1, 0)$, to $F_x(x_0, y_0, z_0) = 0$, $F_y(x_0, y_0, z_0) = 0$, czyli

$$z_x(x_0, y_0) = 0\,, \qquad z_y(x_0, y_0) = 0\,.$$

Przypadkiem, jak się wydaje, trafiliśmy na punkt krytyczny funkcji $z = z(x, y)$. Zaraz się nim zajmiemy.

To samo rozumowanie musi oczywiście być słuszne w każdym punkcie (x, y, z) zbioru E, w którym nie znika F_z i dzięki temu funkcja $z = z(x, y)$ istnieje. Zatem mamy ogólne wzory

$$\left(\frac{\partial z}{\partial x}\right)_y (x, y) = -\left.\frac{F_x(x, y, z)}{F_z(x, y, z)}\right|_{z=z(x,y)}, \qquad \left(\frac{\partial z}{\partial y}\right)_x (x, y) = -\left.\frac{F_y(x, y, z)}{F_z(x, y, z)}\right|_{z=z(x,y)}.$$

(Dla szpanu - używa się jeszcze tego ślicznego słowa? - napisaliśmy te pochodne zgodnie z manierą termodynamiczną zaznaczając która zmienna jest trzymana ustalona). Oczywiście te same wzorki możemy wyprowadzić zauważając, że $z = z(x, y)$ jest taką funkcją, że $F(x,\ y,\ z(x, y)) \equiv 0$ tzn. jest to funkcja x i y tożsamościowo równa zeru. Możemy obliczać pochodne cząstkowe tej tożsamościowo zerowej funkcji stosując już opanowane (mam nadzieję!) reguły słuszne dla pochodnych funkcji złożonych:

$$0 = \frac{\partial}{\partial x} F(x,\ y,\ z(x, y)) = F_x(x, y, z)|_{z=z(x,y)} + F_z(x, y, z)|_{z=z(x,y)}\ \frac{\partial z}{\partial x},$$

i stąd wywikłujemy tę samą pochodną $\partial z/\partial x$, co dana wzorem wypisanym wyżej.

No to "idąc za ciosem" (powtarzam się, ale to też element przemyślanej taktyki - ma na celu uświadomienie Państwu, że wszystko idzie tak samo jak poprzednio) wyprowadźmy wzory na drugie pochodne cząstkowe funkcji $z = z(x, y)$.

$$\begin{aligned}\frac{\partial^2 z}{\partial x^2} &= \frac{\partial}{\partial x}\left[-\frac{F_x(x,\ y,\ z(x, y))}{F_z(x,\ y,\ z(x, y))}\right] = -\frac{1}{F_z}\frac{\partial F_x}{\partial x} + \frac{F_x}{F_z^2}\frac{\partial F_z}{\partial x} \\ &= -\frac{1}{F_z}\left(F_{xx} + F_{xz}\frac{\partial z}{\partial x}\right) + \frac{F_x}{F_z^2}\left(F_{zx} + F_{zz}\frac{\partial z}{\partial x}\right).\end{aligned}$$

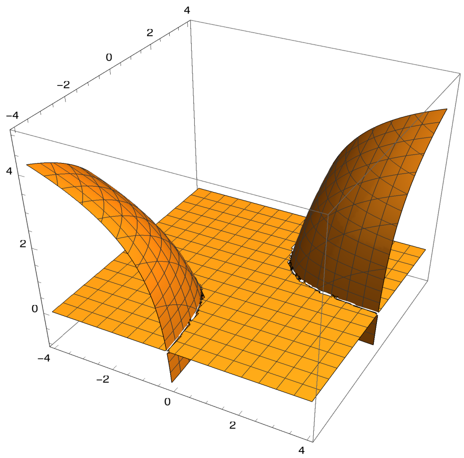
Teraz wstawiamy tu $\partial z/\partial x = -F_x/F_z$ i po małym fiku-miku mamy

$$\begin{aligned}\frac{\partial^2 z}{\partial x^2} &= -\left.\frac{F_{xx}\,F_z^2 - 2F_{xz}\,F_x\,F_z + F_{zz}\,F_x^2}{F_z^3}\right|_{z=z(x,y)}, \\ \frac{\partial^2 z}{\partial y^2} &= -\left.\frac{F_{yy}\,F_z^2 - 2F_{yz}\,F_y\,F_z + F_{zz}\,F_y^2}{F_z^3}\right|_{z=z(x,y)}, \\ \frac{\partial^2 z}{\partial x\,\partial y} &= -\left.\frac{F_{xy}\,F_z^2 - F_{xz}\,F_y\,F_z - F_{yz}\,F_x\,F_z + F_{zz}\,F_x\,F_y}{F_z^3}\right|_{z=z(x,y)}.\end{aligned}$$

Drugi wzór został otrzymany przez zamianę w pierwszym $_x$ na $_y$. Trzeci zostawiam Państwu do samodzielnego wyprowadzenia.

Jak już mamy i wzory na drugie pochodne funkcji $z = z(x, y)$ zadanej w sposób uwikłany, to możemy szaleć: w szczególności możemy badać punkty krytyczne takich funkcji i sprawdzać, czy są w nich ich ekstrema. W tym przypadku znów znajdujemy przyjemne uproszczenie: w punktach krytycznych $F_x = 0$ i $F_y = 0$ i te straszne ogólne wzory wypisane wyżej redukują się do

$$\left.\frac{\partial^2 z}{\partial x^2}\right|_{z_x=z_y=0} = -\left.\frac{F_{xx}}{F_z}\right|_{z=z(x,y)},$$



Rysunek 14: Rysunek zbioru $E \subset \mathbb{R}^3$ zadanego równością $e^z - xyz - 1 = 0$.

$$\left.\frac{\partial^2 z}{\partial y^2}\right|_{z_x=z_y=0} = -\left.\frac{F_{yy}}{F_z}\right|_{z=z(x,y)},$$
$$\left.\frac{\partial^2 z}{\partial x^2}\right|_{z_x=z_y=0} = -\left.\frac{F_{xy}}{F_z}\right|_{z=z(x,y)}.$$

Prawda, że to bardzo przyjemne?

No to teraz możemy wrócić do funkcji $z = z(x, y)$ zadanej w sposób uwikłany warunkiem $F(x, y, z) = e^z - xyz - 1 = 0$ i zbadać charakter przypadkiem znalezionego jej punktu krytycznego $(2, 1, 0)$. Znajdujemy jednak, że $F_{xx}(x, y, z) = 0$, $F_{yy}(x, y, z) = 0$ i $F_{xy}(x, y, z) = z$, więc w badanym punkcie krytycznym forma kwadratowa drugich pochodnych funkcji $z = z(x, y)$ jest całkowicie zerowa. Zatem nic nie można powiedzieć. Ale jeśli przypatrzymy się uważniej warunkowi $F(x, y, z) = e^z - xyz - 1 = 0$, to zobaczymy, że jest on spełniony przez wszystkie punkty leżące w $\mathbb{R}^3$ na płaszczyźnie $z = 0$. Punkt, który badaliśmy jest właśnie jednym z nich i w sposób oczywisty funkcja $z = z(x, y)$ jest po prostu funkcją tożsamościowo równą zeru. Ale zawód nas spotkał!!! Ale nie narzekajmy. Przy okazji czegoś się nauczyliśmy! Wyprowadziliśmy ogólne wzory na pochodne i drugie pochodne takich funkcji, a to już coś! Poza tym, warunek $F(x, y, z) = e^z - xyz - 1 = 0$ nie jest taki trywialny: jeśli zapuścimy Mathematicę każąc jej wymalować powierzchnie $z = z(x, y)$ będące rozwiązaniem tego warunku, to zobaczymy rysunek 14. Widać, że warunek $e^z - xyz - 1 = 0$ definiuje też i inne funkcje $z = z(x, y)$. Ich powierzchnie przecinają się z powierzchnią funkcji $z(x, y) \equiv 0$. Tak może być. Oczywiście w takich punktach przecinania się powierzchni zerują się wszystkie pochodne F_x, F_y i F_z i warunek $F(x, y, z) = 0$ nie wyznacza w otoczeniu takich punktów żadnej funkcji: ani $z = z(x, y)$, anie $x = x(y, z)$, ani $y = y(x, z)$. Tak samo jak miało to miejsce w punkcie samoprzecinania się liścia Kartezjusza.

Zadanie Uwik.4 ("cieciurzynka" - tzn. wzięte z zadań dr G. Cieciury)
Zbadać punkty krytyczne funkcji $z = z(x, y)$ w obszarze $x, y \neq 0$ zadanej warunkiem

$$F(x,y,z) = (x+z)(y+z)\left(1+\frac{z}{xy}\right) - 8 = 0\,.$$

Rozwiązanie: Punkty krytyczne wyznaczają równości

$$\begin{aligned}
F_x(x,y,z) &= (y+z)\left[\left(1+\frac{z}{xy}\right) - \frac{z}{x^2y}\,(x+z)\right] = (y+z)\left(1-\frac{z^2}{x^2y}\right) = 0\,,\\
F_y(x,y,z) &= (x+z)\left[\left(1+\frac{z}{xy}\right) - \frac{z}{y^2x}\,(y+z)\right] = (x+z)\left(1-\frac{z^2}{y^2x}\right) = 0\,,\\
F(x,y,z) &= (x+z)(y+z)\left(1+\frac{z}{xy}\right) - 8 = 0\,.
\end{aligned}$$

Jak zwykle w takich zadaniach, nawet z funkcjami zadanymi jawnie!, a co dopiero, gdy są zadane w sposób uwikłany, trudnym etapem jest rozwiązanie tych równań. Reszta to już rutynowe czynności. Jeśli w pierwszym lub drugim równaniu wybralibyśmy rozwiązanie $y = -z$ i/lub $x = -z$, to nie będzie spełnione trzecie. Zatem $y \neq -z$ i $x \neq -z$. Musi zatem być $z^2 = x^2y$ i $z^2 = y^2x$, co razem oznacza, że[29] $x = y$, czyli $z^2 = x^3 = y^3$. Żeby nie operować pierwiastkami przejdziemy do niezależnej zmiennej t zdefiniowanej przez $z = t^3$ i $x = y = t^2$. Podstawiamy to do ostatniego z równań sprowadzając je tym samym do postaci

$$\left(t^3+t^2\right)^2\left(1+\frac{1}{t}\right) = 8\,.$$

czyli po prostu do $(t^2 + t)^3 = 8$, albo,

$$t^2 + t - 2 = (t-1)(t+2) = 0\,.$$

Zatem punkty krytyczne, to

$$(1,1,1) \qquad \text{oraz} \qquad (4,4,-8)\,.$$

W punktach tych pochodna

$$F_z(x,y,z) = (y+z)\left(1+\frac{z}{xy}\right) + (x+z)\left(1+\frac{z}{xy}\right) + (x+z)(y+z)\,\frac{1}{xy}\,,$$

w której można położyć $z = t^3$, $x = y = t^2$, co sprowadza ją do

$$F_z(t^2,t^2,t^3) = 2(t^2+t^3)\left(1+\frac{1}{t}\right) + \frac{1}{t^4}\left(t^2+t^3\right)^2 = (2t+1)(1+t)^2\,,$$

[29] To, że takie punkty należy sprawdzić, wynika z symetrii funkcji $F(x,y,z) = F(y,x,z)$; jednak przy symetrii punkty krytyczne mogłyby też występować parami: (a,b,c) i (b,a,c); teraz już wiemy, że takich tu niema.

jest równa 12 (w punkcie $(1,1,1)$, $t=1$) i -3 (w punkcie $(4,4,-8)$, $t=-2$). Nie jest wiec równa zeru (trzeba to zawsze sprawdzić! nie zapominać o tym!!!), czyli w tych punktach rzeczywiście warunek $F(x,y,z)=0$ definiuje fukcje $z=z(x,y)$ (a czy to jest ta sama funkcja w obu punktach, czy dwie różne, tego już nie wiemy...). Reszta, jako się rzekło, to już rutyna.

$$
\begin{aligned}
F_{xx}(x,y,z) &= 2(y+z)\,\frac{z^2}{x^3y}\,,\\
F_{yy}(x,y,z) &= 2(x+z)\,\frac{z^2}{y^3x}\,,\\
F_{xy}(x,y,z) &= \left(1-\frac{z^2}{x^2y}\right)+(y+z)\,\frac{z^2}{x^2y^2}=1+\frac{z^3}{x^2y^2}\,.
\end{aligned}
$$

Więc, gdy $x=y=t^2$, $z=t^3$, $F_{xx}=F_{yy}=2(1+t)$, $F_{xy}=1+t$. Pamiętając, że w punktach krytycznych $z_{xx}=-F_{xx}/F_z$, $z_{yy}=-F_{yy}/F_z$, $z_{xy}=-F_{xy}/F_z$, dostajemy w tych punktach następujące macierze form kwadratowych

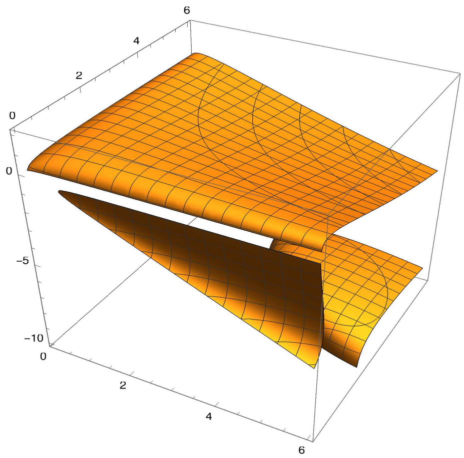
$$
-\frac{1}{12}\begin{pmatrix}4 & 2\\ 2 & 4\end{pmatrix}, \qquad \frac{1}{3}\begin{pmatrix}-2 & -1\\ -1 & -2\end{pmatrix}.
$$

Pierwsza z nich, odpowiadająca punktowi $(1,1,1)$, jest ujemnie określona i tam funkcja $z=z(x,y)$ ma lokalne maksimum. Druga, odpowiadająca punktowi $(4,4,-8)$, jest też ujemnie określona i tam funkcja $z=z(x,y)$ ma też lokalne maksimum. Żeby nie żyć w niepewności, czy to są ekstrema tej samej funkcji, czy dwóch różnych, zaprzęgamy znów Mathematicę i otrzymujemy rysunek 15. Warunek $F(x,y,z)=0$ definiuje w pokazanym obszarze aż cztery różne funkcje $z=z_i(x,y)$, $i=1,2,3,4$ (czwarta nie jest tu widoczna, ale można ją zobaczyć obracając wykres w Mathematice, lub po prostu wnosić o jej istnieniu odwołując sie do symetrii $F(x,y,z)$ w zmiennych x i y), z czego tylko dwie (ten "dach" na rysunku 15 i ten placek na dole) mają (to ustaliliśmy analitycznie) punkty krytyczne, które są ich maksimami.

Na koniec, żebyśmy mieli poczucie, że umiemy sobie radzić w każdej uwikłanej (byle nie w sprzeczności) sytuacji rozpatrzmy jeszcze przypadek dwóch warunków na trzy zmienne, czyli odwzorowanie $F:\ \mathbb{R}^3\longrightarrow\mathbb{R}^2$, które przez $F^{-1}(\mathbf{0})=E\subset\mathbb{R}^3$ zadaje funkcję $f:\mathbb{R}\longrightarrow\mathbb{R}^2$, czyli po prostu krzywą w $\mathbb{R}^2$ sparametryzowaną jedną zmienną. Niech

$$
\begin{aligned}
F_1(x,y,z) &= xz+\ln y+y\ln z\,,\\
F_2(x,y,z) &= x-y+z\,.
\end{aligned}
$$

Znajdźmy w punkcie $(0,1,1)$ pochodne dy/dx i dz/dx funkcji $x\to(y(x),\,z(x))$ zadanej warunkami $F_1(x,y,z)=0$, $F_2(x,y,z)=0$. Każdy z tych warunków z osobna definiuje (najpewniej) jakąś powierzchnię w $\mathbb{R}^3$ (no, czasem kilka rozłącznych, albo przecinających się - to już wiemy z poprzednich przykładów), a punkty spełniające oba te warunki leżą na przecięciu takich powierzchni, które to przecięcie najbardziej typowo jest właśnie krzywą,



Rysunek 15: Rysunek zbioru $E \subset \mathbb{R}^3$ zadanego równością $(x+z)(y+z)(1+z/xy)-8=0$.

którą można sparametryzować np. x-em i o to tu chodzi. Podany punkt $(0,1,1)$ należy jak widać do zbioru E bo spełnia oba warunki: $F_1(0,1,1)=0$, $F_2(0,1,1)=0$. Aby obliczyć pochodne dy/dx i dz/dx w punkcie $(0,1,1)$ piszemy w tym punkcie różniczki obu tych funkcji (numer funkcji F teraz będziemy pisać u góry żeby na dole mieć miejsce na symbol pochodnej cząstkowej)

$$\begin{aligned} dF^1 &= F^1_x\,dx + F^1_y\,dy + F^1_z\,dz\,, \\ dF^2 &= F^2_x\,dx + F^2_y\,dy + F^2_z\,dz\,, \end{aligned}$$

obliczając pochodne cząstkowe F^1_x, etc. w punkcie $(0,1,1)$. Znów pytamy jak skorelować składowe dx, dy, dz wektora przemieszczenia, żeby nadal $F^1=0$ i $F^2=0$, czyli, żeby $dF^1=0$ i $dF^2=0$. Jawnie

$$\begin{aligned} dF^1 &= z\,dx + \left(\frac{1}{y} + \ln z\right) dy + \left(\frac{y}{z} + x\right) dz\,, \\ dF^2 &= dx - dy + dz\,. \end{aligned}$$

W punkcie $(0,1,1)$ warunki $dF^1=0$ i $dF^2=0$ dają związki

$$\begin{aligned} dx + dy + dz &= 0\,, \\ dx - dy + dz &= 0\,. \end{aligned}$$

Ich rozwiązaniem jest oczywiście $dz=-dx$, $dy=0\cdot dx$. Zatem w punkcie $(0,1,1)$, w którym $y(0)=1$, $z(0)=1$

$$\frac{dy}{dx}=0\,, \qquad \frac{dz}{dx}=-1\,.$$

W dowolnym punkcie, aby dostać te pochodne, rozwiązywalibyśmy wypisany wyżej ogólny układ równań, który można przekształcić do postaci

$$\begin{pmatrix} F^1_y/F^1_x & F^1_z/F^1_x \\ F^2_y/F^2_x & F^2_z/F^2_x \end{pmatrix} \begin{pmatrix} dy \\ dz \end{pmatrix} = \begin{pmatrix} -dx \\ -dx \end{pmatrix},$$

skąd

$$\begin{pmatrix} dy \\ dz \end{pmatrix} = \frac{F_x^1 F_x^2}{F_y^1 F_z^2 - F_z^1 F_y^2} \begin{pmatrix} F_z^2/F_x^2 & -F_z^1/F_x^1 \\ -F_y^2/F_x^2 & F_y^1/F_x^1 \end{pmatrix} \begin{pmatrix} -dx \\ -dx \end{pmatrix},$$

Dostalibyśmy więc np.

$$\frac{dy}{dx} = \frac{F_x^1 F_x^2}{F_y^1 F_z^2 - F_z^1 F_y^2} \left(\frac{F_z^1}{F_x^1} - \frac{F_z^2}{F_x^2} \right).$$

Przy okazji widać, w jakich punktach przestrzeni $\mathbb{R}^3$ warunki $F^1(x,y,z) = 0$ i $F^2(x,y,z) = 0$ wyznaczają funkcję $\mathbb{R} \longrightarrow \mathbb{R}^2$: w takich, w których nie znika $F_y^1 F_z^2 - F_z^1 F_y^2$. Te same pochodne dy/dx i dz/dx można też dostać i naszą drugą metodą, tj. pisząc dwie tożsamościowo równe zeru funkcje $F^1(x,\, y(x),\, z(x)) \equiv 0$, $F^2(x,\, y(x),\, z(x)) \equiv 0$, różniczkując te tożsamości po x, co da

$$\begin{aligned} 0 &= F_x^1 + F_y^1\, y' + F_z^1\, z' \,, \\ 0 &= F_x^2 + F_y^2\, y' + F_z^2\, z' \,, \end{aligned}$$

i rozwiązując ten układ liniowych równań względem $y' = dy/dx$ i $z' = dz/dx$. Ale jest to ten sam układ, co poprzednio napisany. Różniczkując wypisane wyżej obie równości jeszcze raz stronami po x i wstawiając potem już wyliczone y' i z', dostaniemy zaś układ dwu równań na drugie pochodne y'' i z''. Naprawdę już wszystko możemy!

Ogólnie, jeśli poziomica zerowa $F^{-1}(\mathbf{0}) \subset \mathbb{R}^n$, $\mathbf{0} \in \mathbb{R}^m$ odwzorowania $F : \mathbb{R}^n \longrightarrow \mathbb{R}^m$, gdzie $m < n$, ma lokalnie, w otoczeniu jakiegoś punktu $(x_1^{(0)}, \ldots, x_n^{(0)})$ spełniającego równości

$$F^1(x_1^{(0)}, \ldots, x_n^{(0)}) = 0\,, \quad \ldots, \quad F^m(x_1^{(0)}, \ldots, x_n^{(0)}) = 0\,,$$

wyznaczać "przyzwoitą" funkcję $f : \mathbb{R}^{n-m} \longrightarrow \mathbb{R}^m$, np.

$$\begin{aligned} x_1 &= x_1(x_{m+1}, \ldots, x_n)\,, \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ x_m &= x_m(x_{m+1}, \ldots, x_n)\,, \end{aligned}$$

to w punkcie $(x_1^{(0)}, \ldots, x_n^{(0)})$ to wyznacznik macierzy

$$\begin{pmatrix} \partial F^1/\partial x_1 & \ldots & \partial F^1/\partial x_m \\ \cdot & \ldots & \cdot \\ \partial F^m/\partial x_1 & \ldots & \partial F^m/\partial x_m \end{pmatrix} \equiv \frac{\partial(F^1, \ldots, F^m)}{\partial(x_1, \ldots, x_m)}\,,$$

powinien nie być zerowy. Widać, że jest to uogólnienie tego, co mieliśmy, gdy warunek $F(x,y) = 0$ miał wyznaczać funkcję $y = y(x)$ lub gdy warunek $F(x,y,z) = 0$ miał wyznaczać $z = z(x,y)$. W przypadku "rozwikływania" odwzorowania $F : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ o $m < n$

można to skonfrontować z przykładem odwzorowania liniowego, od którego rozpoczęliśmy ten rozdział: nieznikanie powyższego wyznacznika jest tam właśnie warunkiem, by odpowiedni minor stopnia m macierzy A problemu nie znikał. Przy okazji wspomniany był tam problem niezależności warunków $F^1(x_1, \ldots, x_n) = 0$, ..., $F^m(x_1, \ldots, x_n) = 0$. Tym z kolei rządzi macierz pochodnej F' funkcji F:

$$\begin{pmatrix} \partial F^1/\partial x_1 & \ldots & \partial F^1/\partial x_n \\ \cdot & \ldots & \cdot \\ \partial F^m/\partial x_1 & \ldots & \partial F^m/\partial x_n \end{pmatrix}.$$

Warunki $F^1(x_1, \ldots, x_n) = 0$, ..., $F^m(x_1, \ldots, x_n) = 0$ są w otoczeniu punktu $(x_1^{(0)}, \ldots, x_n^{(0)})$ niezależne, gdy rząd tej macierzy jest maksymalny, tj. równy m (znów można to skonfrontować z przykładem odwzorowania liniowego i uwagami poczynionymi na jego temat).

Ekstrema warunkowe czyli związane.
Zagadnieniem wiążącym się naturalnie z odwzorowaniami uwikłanymi jest szukanie lokalnych ekstremów funkcji $F(x_1,\dots,x_n)$ określonej na jakimś podzbiorze (ograniczonym lub nieograniczonym) $E\subset\mathbb{R}^n$, który to podzbiór jest zadany jako poziomica zerowa $G^{-1}(\mathbf{0})\subset\mathbb{R}^n$ pewnego odwzorowania $G:\mathbb{R}^n\to\mathbb{R}^m$ $(m<n)$. Mówiąc bardziej przyziemnym językiem, szukamy ekstremum funkcji $F(x_1,\dots,x_n)$ przy ubocznych warunkach $G^1(x_1,\dots,x_n)=0\ ,\ \dots,\ G^m(x_1,\dots,x_n)=0$ zwanych także *więzami*. Gdyby w otoczeniu jakiegoś punktu $(x_1^*,\dots,x_n^*)\in E$ dało się *jawnie* wywikłać m którychś zmiennych x w funkcji pozostałych $n-m$ zmiennych x, czego warunkiem koniecznym jest by macierz G' pochodnej odwzorowania G w tym punkcie była rzędu m (co będziemy tu zawsze zakładać), np. gdyby dało się otrzymać funkcje $x_1=x_1(x_{m+1},\dots,x_n)$, $\dots$, $x_m=x_m(x_{m+1},\dots,x_n)$, czego warunkiem koniecznym jest (patrz koniec poprzedniego rozdziału), by nie znikał wyznacznik macierzy $m\times m$ utworzonej z pierwszych m kolumn macierzy G' pochodnej odwzorowania G w tym punkcie, to można by było po prostu badać funkcję $n-m$ zmiennych

$$f(x_{m+1},\dots,x_n)=F(x_1(x_{m+1},\dots,x_n),\dots,x_m(x_{m+1},\dots,x_n),\,x_{m+1},\dots,x_n)\,,$$

stosując znane już nam standardowe metody. Gdyby nie dało się wywikłać $x_1,\dots,x_m$, to, przy przyjętym założeniu, że rząd macierzy G' jest równy m, dałoby się wywikłać jawnie m którychś innych zmiennych, np. $x_{n-m+1},\dots,x_n$ w funkcji $x_1,\dots,x_{n-m}$. Naogół się jednak nie daje tego zrobić jawnie. Co więcej, widzieliśmy na przykładach, że czasem w otoczeniu jednego punktu zbioru E da się wywikłać m jednych x-ów, a w otoczeniu innego punktu zbioru E daje się wywikłać tylko m innych zmiennych, a my byśmy chcieli problem szukania lokalnych ekstremów funkcji f na zbiorze E rozwiązać jakoś globalnie, w tym sensie, żeby nie musieć dzielić zbioru E na kawałki i w jednym wywikływać jawnie, lub tylko niejawnie, jednych zmiennych, a w innym innych. Do tego służy metoda wykorzystująca *mnożniki Lagrange'a.* (Mnożniki te znajdują zastosowanie w wielu problemach, np. w mechanice, gdy rozpatruje się siły reakcji, a także - jakże by inaczej! - w mojej ulubionej termodynamice. Kto do mnie trafi na wykład ten zobaczy).

Najpierw jak wyznaczamy na zbiorze E punkty krytyczne? Załóżmy, że tak jak wyżej wywikłaliśmy m pierwszych zmiennych x w funkcji pozostałych $n-m$. Chodzi jednak o to, że tylko zakładamy; nie będziemy musieli tego robić jawnie i, co więcej, końcowe wzory nie będą wyróżniać żadnej z n zmiennych x, więc będą słuszne globalnie na całym zbiorze E, pod warunkiem, że w każdym jego punkcie, któreś m zmiennych x się daje wywikłać (a to, jak już było wyżej, jest żądaniem, by w każdym punkcie E macierz G' była rzędu m, a nie mniejszego). Punkty krytyczne są więc wyznaczone przez zerowanie się pochodnych funkcji f po niezależych x-ach, czyli przez układ $n-m$ równań

$$\begin{gathered}\frac{\partial}{\partial x_{m+1}}\,F(x_1(x_{m+1},\dots,x_n),\dots,x_m(x_{m+1},\dots,x_n),\,x_{m+1},\dots,x_n)=0\,,\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ \frac{\partial}{\partial x_n}\,F(x_1(x_{m+1},\dots,x_n),\dots,x_m(x_{m+1},\dots,x_n),\,x_{m+1},\dots,x_n)=0\,.\end{gathered}$$

Obliczając np. pochodną po x_{m+1} musimy wiec różniczkować F po x_1 i x_1 po x_{m+1}, F po x_2 i x_2 po x_{m+1} itd. To mam nadzieję jest jasne. Ale, i teraz nam się przydaje wiedza o odwzorowaniach uwikłanych, skoro $x_1 = x_1(x_{m+1}, \ldots, x_n)$, ..., $x_m = x_m(x_{m+1}, \ldots, x_n)$ biorą się z rozwikłania warunków $G^1(x_1, \ldots, x_n) = 0$, ..., $G^m(x_1, \ldots, x_n) = 0$, to

$$
\begin{gathered}
G^1(x_1(x_{m+1}, \ldots, x_n), \ldots, x_m(x_{m+1}, \ldots, x_n), x_{m+1}, \ldots, x_n) \equiv 0\,, \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
G^m(x_1(x_{m+1}, \ldots, x_n), \ldots, x_m(x_{m+1}, \ldots, x_n), x_{m+1}, \ldots, x_n) \equiv 0\,,
\end{gathered}
$$

są tożsamościami względem niezależnych zmiennych $x_{m+1}, \ldots, x_n$ i różniczkując je po tych zmiennych dostajemy zera. Możemy powyższe m tożsamości zróżniczkować po x_{m+1} i to da m liniowych równań na m pochodnych $\partial x_1/\partial x_{m+1}$, ..., $\partial x_m/\partial x_{m+1}$. Jeśli wypisane wyżej tożsamości kropniemy po x_{m+2}, to dostaniemy m liniowych równań na m pochodnych $\partial x_1/\partial x_{m+2}$, ..., $\partial x_m/\partial x_{m+2}$, itd. Razem różniczkując te tożsamości po $x_{m+1}, \ldots, x_n$ dostaniemy $n-m$ liniowych układów po m równań na m pochodnych każdy i w ten sposób możemy wyliczyć wszystkie potrzebne nam pochodne od $\partial x_1/\partial x_{m+1}$, po $\partial x_m/\partial x_n$. Łatwo dostrzec w tym po prostu zastosowanie tego, co było na końcu poprzedniego rozdziału.

Jednak robi się to wszystko zawiłe, nie tyle z powodu komplikacji samej materii, co z powodu nadmiaru literek i wskaźników, a wiadomo: nec Hercules contra plures, czyli i Herkules dupa,[30] gdy wskaźników kupa - nawet Einstein by teorii względności nie wymyślił, gdyby nie wpadł na pomysł konwencji sumacyjnej (już jej używaliśmy na algebrze to i tu użyjemy). Pora więc dokonać skoku technologicznego i zapisać to wszystko w zwarty sposób. Mamy więc m warunków $G^a = 0$, $a = 1, \ldots, m$, a zmienne $(x_1, \ldots, x_n)$ zapiszemy jako $(y_1, \ldots, y_m, x_{m+1}, \ldots, x_n)$, albo jeszcze krócej, jako[31] (y^a, x^i), $a = 1, \ldots, m$ oraz $i = m+1, \ldots, n$. Będziemy też pisać[32] $F'_i = \partial F/\partial x^i$ oraz $F'_b = \partial F/\partial y^b$. Wszyscy dorośli tak piszą, to my też. No to teraz warunki wyznaczające ekstrema funkcji $f(x^i)$ zapiszą się w jedej linijce (przypominam, że po wskaźniku występującym dwa razy, raz na górze, raz na dole, sumujemy; jeśli ten sam wskaźnik - ta sama literka robiąca za wskaźnik - pojawi się więcej niż dwa razy, to znaczy, że coś schrzaniliśmy po drodze - tak naucza wujek Albercik)

$$
f'_i \equiv F'_i + F'_a \frac{\partial y^a}{\partial x^i} = 0\,, \qquad i = m+1, \ldots, n\,,
$$

a owe $n-m$ układów po m równań liniowych na m niewiadomych pochodnych każdy też się zapiszą w jednej linii:

$$
[G']^b{}_i + [G']^b{}_a \frac{\partial y^a}{\partial x^i} = 0\,, \qquad i = m+1, \ldots, n\,.
$$

[30] Przepraszam za wulgaryzm, ale inaczej się nie rymuje; zresztą, jak mówił T. Konwicki - pisarz zdaje się dziś zupełnie zapoznany, a dla mojego pokolenia "kultowy" - "Czasem brzydkie słowo jest jak łyk świeżego powietrza." Sytuacja ogólna jest taka, że aż chce się mu przytaknąć.

[31] Umieszczamy tu wskaźnik i u góry, bo x^i i y^a mają sens składowych (w bazie kanonicznej) wektora z $\mathbb{R}^n$ i w ten sposób będziemy mieli "algebraiczny" porządek w interesie.

[32] Poprzednio pochodne cząstkowe funkcji, np. $f = f(x, y, z)$, pisaliśmy jako f_x, f_y; teraz jednak będzie wygodniej pochodną opatrywać też primem, czyli pisać f'_x itd., bo pochodne funkcji $f(x^1, x^2, x^3)$ będziemy teraz pisać jako $f'_j \equiv \partial f/\partial x^j$; trzeba trochę się wykazywać inteligencją i rozumieć co jest pochodną, a co ma wskaźnik bo jest wektorem "od urodzenia"...

Wskaźnik b się zmienia od 1 do m i numeruje równania w i-tym układzie; po a jest sumowanie od 1 do m. Wskaźnik $i = m+1, \ldots, n$ numeruje te $n-m$ układów równań. G' to jest pochodna[33] odwzorowania G; jest ona macierzą $m \times n$ (tj. ma m wierszy, n kolumn - to już wiemy!). Ponieważ założyliśmy, że da się wyznaczyć y-ki (czyli $x_1, \ldots, x_m$) - jeszcze raz to sobie przypomnijmy) jako funkcje x-ów (czyli $x_{m+1}, \ldots, x_n$), to znaczy, że wyznacznik macierzy (kwadratowej, $m \times m$) $[G']^b{}_a$ jest niezerowy i można ją odwrócić.[34] Zatem

$$\frac{\partial y^a}{\partial x^i} = -\left[(G')^{-1}\right]^a{}_b \, [G']^b{}_i \,, \qquad i = m+1, \ldots, n \,.$$

No to teraz to wstawiamy do warunków $f_i' = 0$ na punkty krytyczne i dostajemy

$$F_i' - F_a' \left[(G')^{-1}\right]^a{}_b \, [G']^b{}_i = 0 \,, \qquad i = m+1, \ldots, n \,.$$

Te z kolei równania możemy zapisać w chytry sposób wprowadzając właśnie te zapowiedziane mnożniki Lagrange'a:

$$F_i' - \lambda_b \, [G']^b{}_i = 0 \,, \qquad i = m+1, \ldots, n \,,$$

gdzie $\lambda_b \equiv F_a' \left[(G')^{-1}\right]^a{}_b$ są właśnie mnożnikami Lagrange'a numerowanymi wskaźnikiem $b = 1, \ldots, m$. Pamiętamy przy tym - potośmy wałkowali te funkcje uwikłane! - że rozwiązania tych wszystkich warunków nas interesują nie na całym $\mathbb{R}^n$, tylko na $G^{-1}(\mathbf{0}) \subset \mathbb{R}^n$, czyli gdy spełnione są warunki $G^a = 0$, $a = 1, \ldots, m$. No to teraz już możemy wszystko zapisać elegancko: Aby znaleźć punkty krytyczne funkcji f odwzorowującej $E = G^{-1}(\mathbf{0}) \subset \mathbb{R}^n$ w liczby rzeczywiste, tworzymy pomocniczą funkcję

$$\tilde{F}(x^1, \ldots, x^n) \equiv F(x^1, \ldots, x^n) - \lambda_b(x^1, \ldots, x^n) \, G^b(x^1, \ldots, x^n) \,,$$

i rozwiązujemy układ równań

$$\begin{aligned} \frac{\partial \tilde{F}(x^1, \ldots, x^n)}{\partial x^i} &= 0 \,, \qquad i = 1, \ldots, n \,, \\ G^a(x^1, \ldots, x^n) &= 0 \,, \qquad a = 1, \ldots, m \,, \end{aligned}$$

który można równoważnie zapisać jako

$$\begin{aligned} F_i' - \lambda_b \, [G']^b{}_i &= 0 \,, \qquad i = 1, \ldots, n \,, \\ G^a(x^1, \ldots, x^n) &= 0 \,, \qquad a = 1, \ldots, m \,. \end{aligned}$$

Mnożniki λ_b są w zasadzie funkcjami x-ów, ale gdy je różniczkujemy, to ich pochodne są mnożone przez funkcje G^b, które mają być i tak zero; więc z praktycznego punktu

[33] Logiczniej by ją było pisać jako $(G^{a\prime})_b$, bo to jest n pochodnych m różnych funkcji, ale wtedy by trudniej zobaczyć w tym zapisie macierz.

[34] Założyliśmy, że macierz G' ma rząd maksymalny, czyli m, więc któreś m x-ów da się wyrazić przez pozostałe; jak nie te to inne.

widzenia, można tu traktować mnożniki Lagrange'a jak stałe. Zatem jest to $n+m$ równań na $n+m$ niewiadomych: n niewiadomych x^i oraz m niewiadomych λ_b. Rozwiązaniem jest n wartości x^i_* i m wartości λ^*_b. Ten nowy układ równań jest tym samym co układ wypisany wcześniej, bo te definicje mnożników Lagrange'a, $\lambda_b \equiv F_a\left[(G')^{-1}\right]^a{}_b$, można zapisać (zurück odwracając kwadratową macierz pochodnych) jako

$$F'_a - \lambda_b \left[G'\right]^b{}_a = 0\,, \qquad a = 1, \ldots, m\,,$$

i te m równań w połączeniu z $n-m$ równaniami $F'_i - \lambda_b\left[G'\right]^b{}_i = 0$, $i = m+1, \ldots, m$ daje razem takie same równania wypisane wyżej tylko z numerującym je wskaźnikiem i biegnącym już od 1 do n. Widać teraz, że otrzymany układ równań nie wyróżnia żadnych m zmiennych x spośród pozostałych i tym samym nie musimy deklarować, które x-y da się wyrazić przez pozostałe. Uff. Zaszalałem. Taki mam bezkompromisowy charakter.[35] Jeśli kogoś te straszne wzory przerażają, to niech o nich natychmiast zapomni i przyswoi z notatek wykładowcy prosty przypadek jednej funkcji i jednego warunku. No i wystarczy, by pamiętał ogólny praktyczny przepis.

Wyprowadziliśmy więc praktyczny przepis na znajdywanie na zbiorze E punktów krytycznych funkcji. Teraz jeszcze musimy podać kryterium, kiedy w takich punktach są rzeczywiście ekstrema lokalne funkcji na zbiorze E. Gdy nie było żadnych warunków, tzn., gdy pytaliśmy o ekstrema funkcji zadanej na $\mathbb{R}^n$ (lub na jakimś otwartym podzbiorze $\mathbb{R}^n$), to badaliśmy określoność (sygnaturę) formy kwadratowej drugich pochodnych funkcji w znalezionym punkcie krytycznym. Opierało się to na wzorze Taylora i fakcie, że pierwsze pochodne funkcji w punkcie krytycznym znikały. Teraz nie możemy się bezpośrednio do tej metody odwołać, bo pierwsze pochodne funkcji w punkcie krytycznym naogół nie znikają. Znikają owszem, pierwsze, pochodne tej pomocniczej funkcji $\tilde{F}$, ale to nie to samo - nas interesują ekstrema $f \equiv F|_E$ (funkcji F obciętej do zbioru E), a nie funkcji $\tilde{F}$ na $\mathbb{R}^n$! Są tu różne podejścia. Jedno najprostsze jest takie: jeśli zbiór $E = G^{-1}(\mathbf{0})$ jest zwarty (w $\mathbb{R}^n$ oznacza to, że E jest domknięty i ograniczony), a interesują nas tylko ekstrema globalne (czyli największa i najmniejsza wartość funkcji na E), to po prostu można sprawdzić wartość funkcji F w każdym z tych punktów, bo twierdzenie Weierstrassa (*Zbiorze szlachetny,...* itd., to właśnie to!) mówi że funkcja ciągła na zbiorze zwartym osiąga swe kresy, i wybrać wartość największą i najmniejszą.

Drugi sposób propagowany przez Wykładowcę[36] (i zapewne uzasadniony na wykładzie) ale w formie przez niego zaprezentowanej *stosowalny tylko do przypadku funkcji na* $\mathbb{R}^2$ *i jednego warunku ubocznego* polega na potraktowaniu funkcji $\tilde{F}$ jak funkcji x-ów (wszystkich n x-ów) i dodatkowo funkcji m mnożników Lagrange'a (pisałem, że z praktycznego punktu widzenia one są jak stałe, a teraz będziemy je uważać za niezależne zmienne):

$$\tilde{F}(x^1, \ldots, x^n, \lambda_1, \ldots, \lambda_m) \equiv F(x^1, \ldots, x^n) - \lambda_b\, G^b(x^1, \ldots, x^n)\,,$$

[35] Ale też uważam, że czasem zajęcie się ogólnym przypadkiem nie jest jakoś specjalnie trudniejsze, a rzuca trochę więcej światła na ogólną metodę i dlatego warto trochę zainwestować.

[36] Pisałem to, gdy wykładał prof. J. Wojtkiewicz. Może to teraz nieaktualne.

i zbadaniu, gdy $n = 2$, $m = 1$, *tylko jednej* liczby, mianowicie wyznacznika macierzy drugich pochodnych tej funkcji, jako funkcji $n + m = 3$ zmiennych (macierz formy kwadratowej drugich pochodnych jest więc w tym podejściu macierzą 3×3), wziętych w punktach krytycznych, tj. $x^1_*, \dots, x^2_*$, λ^*. Jeśli ten wyznacznik jest dodatni, to w punkcie krytycznym funkcja ma maksimum lokalne (a nie minimum, jak by można myśleć!), a jeśli jest ujemny, to lokalne minimum.

Jeszcze inny sposób, najczęściej wykorzystywany, bo nie jest ograniczony do przypadku $n = 2$, $m = 1$, polega na badaniu macierzy formy kwadratowej drugich pochodnych funkcji $\tilde{F}$ (nie F!), ale teraz traktowanej, jak funkcja tylko n zmiennych $(x^1, \dots, x^n)$ (mnożniki Lagrange'a traktujemy tu jak zupełnie stałe). W tym podejściu macierz tej formy obliczona w punkcie $(x^1_*, \dots, x^n_*)$ - mnożnikom Lagrange'a nadajemy w tym podejściu wartości $\lambda^*_1, \dots, \lambda^*_m$ - jest macierzą $n \times n$. Nie badamy jednak jej określoności na wszystkich możliwych wektorach przyrostów $(h^1, \dots, h^n)$, a tylko na wektorach $(h^1, \dots, h^n)$ stycznych do warunku więzów, tzn., tylko na wektorach spełniających m warunków

$$\begin{array}{c} \left.\dfrac{\partial G^1}{\partial x^1}\right|_{\mathbf{x}^*} h^1 + \ldots + \left.\dfrac{\partial G^1}{\partial x^n}\right|_{\mathbf{x}^*} h^n = 0\,, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \left.\dfrac{\partial G^m}{\partial x^1}\right|_{\mathbf{x}^*} h^1 + \ldots + \left.\dfrac{\partial G^m}{\partial x^n}\right|_{\mathbf{x}^*} h^n = 0\,. \end{array}$$

W naszej zwartej notacji warunki te można napisać po prostu jako

$$[G'_{\mathbf{x}^*}]^a{}_i h^i = 0\,, \quad a = 1, \dots, m\,.$$

Wektory $\mathbf{h} \equiv (h^1, \dots, h^n)$ spełniające tych m liniowych warunków (albo jeden macierzowy warunek) mają tylko $n - m$ niezależnych składowych, bo m ich składowych można wyrazić przez m pozostałych (znów trzeba w macierzy G' wymiaru $m \times n$, o której założyliśmy, że jest rzędu m, wybrać podmacierz $m \times m$ i to wybór tej podmacierzy wyznacza, które składowe wektorów $\mathbf{h}$ się da wyrazić przez pozostałe). Z praktycznego punktu widzenia badanie formy kwadratowej Q drugich pochodnych na wektorach $\mathbf{h}$ spełniających powyższe warunki, można przeprowadzić następująco: do wyrażenia

$$Q(\mathbf{h}) = Q_{11}\, h^1 h^1 + 2Q_{12}\, h^1 h^2 + \ldots + Q_{nn}\, h^n h^n\,,$$

będącego jawnym rozpisaniem działaniem formy Q na wektor $\mathbf{h}$, podstawiamy wyznaczone z tych dodatkowych warunków składowe $h^1 = h^1(h^{m+1}, \dots, h^{m+1})$, ..., $h^m = h^m(h^{m+1}, \dots, h^{m+1})$, ... (albo inne, jeśli akurat wyznaczyliśmy/dało się wyznaczyć m innych). Ponieważ związki te są liniowe, całość nadal będzie miała postać formy kwadratowej $\tilde{Q}$ działającej już na "krótszy" wektor przyrostów $\tilde{\mathbf{h}}$ mający tylko $n - m$ składowych. Można to nazwać *redukowaniem formy* do formy kwadratowej zadanej wyłącznie na wektorach stycznych do powierzchni więzów. Formę $\tilde{Q}$ możemy znów zwinąć do postaci macierzowej i zastosować do jej badania metodę minorową. To, że drugie pochodne badamy na wektorach przemieszczeń stycznych do powierzchni więzów jest zrozumiałe: tylko takie przemieszczenia nie wyprowadzają nas poza zbiór E, a to na zbiorze E tylko interesują nas ekstrema funkcji F. Oczywiście, jeśli się zdarzy (a zdarza się to dość często), że cała forma kwadratowa Q drugich pochodnych funkcji $\tilde{F}$ jest w danym punkcie krytycznym $(x^1_*, \dots, x^n_*)$, $\lambda^*_1, \dots, \lambda^*_m$ określona (ma sygnaturę z samymi plusami, albo z samymi minusami), to na wszystkich wektorach przemieszczeń, więc także i na wektorach stycznych do powierzchni więzów, daje wartości tylko jednego znaku i nie trzeba już badać jej specjalnie na wektorach stycznych do więzów.

Nasuwa się naturalne pytanie, dlaczego ta metoda działa? Dlaczego nie wystarczy np. badanie na wektorach stycznych do więzów formy kwadratowej drugich pochodnych (wziętych w punkcie krytycznym) samej funkcji F (a nie $\tilde{F}$). To akurat jest jasne: bo pierwsze pochodne funkcji F nie muszą znikać w punkcie krytycznym, więc nie działa orgument, że o kierunku zmiany wartości funkcji przey przesunięciu się z $\mathbf{x}_*$ o $\mathbf{h}$ decyduje $Q(\mathbf{h})$ nawet jeśli formę Q będziemy badać tylko na wektorach $\mathbf{h}$ stycznych do więzów. Ale dlaczego akurat przepis z funkcją $\tilde{F}$ działa? Też się zastanawiałem i wymyśliłem kiedyś taki tego dowód. Wyobraźmy sobie, że szukamy na $\mathbb{R}^n$ (a nie na E) ekstremów funkcji

$$\tilde{F}(\mathbf{x}) = F(\mathbf{x}) - \lambda_a(\mathbf{x})\, G^a(\mathbf{x})\,,$$

w której m funkcji $\lambda_a(\mathbf{x})$ jest zupełnie dowolnymi funkcjami na $\mathbb{R}^n$, ograniczonymi tylko warunkiem by $\lambda_a(\mathbf{x}^*) = \lambda^*_a$, gdzie $\mathbf{x}^*$ i λ^*_a to jest badany punkt krytyczny na E i odpowiadające mu wartości mnożników Lagrange'a w naszym wyjściowym problemie. Funkcja $\tilde{F}(\mathbf{x})$ jest tak skonstruowana, że po pierwsze na zbiorze E przyjmuje dokładnie te same wartości, co funkcja F rozpatrywana w wyjściowym problemie, a po drugie, na $\mathbb{R}^n$ ma te same punkty krytyczne co F na zbiorze E (być może ma także i inne, poza E, ale to nie jest tu ważne). No bo warunki na punkty krytyczne

$$\tilde{F}'_i \equiv F'_i - (\lambda'_a)_i\, G^a - \lambda_a\, G'_i = 0\,,$$

są spełniane przez punkty $\mathbf{x}^* \in E$ dzięki temu, że $G^a = 0$ na E i że o funkcjach $\lambda_a(\mathbf{x})$ założyliśmy, iż $\lambda_a(\mathbf{x}^*) = \lambda^*_a$. Wypiszmy teraz (w zwartej notacji) drugie pochodne tej funkcji:

$$\tilde{F}''_{ij} = F''_{ij} - (\lambda'_a)_i\, (G')^a{}_j - (\lambda'_a)_j\, (G')^a{}_i - \lambda_a\, (G'')^a{}_{ij} - (\lambda''_a)_{ij}\, G^a\,.$$

Teraz wszystko jest standardowo: pierwsze pochodne funkcji $\tilde{F}$ są w punkcie $\mathbf{x}^*$ równe zeru, więc możemy, szukać ekstremów $\tilde{F}$ odwołując się do wzoru Taylora i stosując go do wektorów przesunięć $\mathbf{h}$ stycznych do więzów (bo one nie wyprowadzają poza zbiór E, a na E funkcja $\tilde{F}$ przyjmuje takie same wartości jak funkcja F) badamy w $\mathbf{x}^*$ powyższą formę kwadratową drugich pochodnych. W punktach krytycznych na E ostatni człon znika (bo $G^a = 0$ na E). W szczególności, z punktu widzenia ekstremów na zbiorze E tylko (przypomnijmy, że na E funkcja $\tilde{F}$ ma te same wartości, co F) interesuje nas, jak ta forma działa na wektory przemieszczeń styczne do więzów, tzn., interesuje nas dodatnia lub ujemna określoność wyrażenia $h^i \tilde{F}''_{ij}\, h^j$ na wektorach stycznych. Ale na takich wektorach $(G')^a{}_i h^i = 0$ więc dwa wyrazy z pochodnymi funkcji λ_a wypadają i wszystko sprowadza się do tego, co zostało podane w praktycznym przepisie.

Tak jak w przypadku szukania ekstremów zwykłych, najtrudnejszym elementem w całej tej zabawie jest rozwiązanie układu równań wyznaczających punkty krytyczne na E. Wymaga to często pewnej dozy sprytu i "orientacji w terenie". Gdy te punkty już są znalezione, reszta jest sprawą rutynowych czynności. Teraz możemy już przejść do praktycznych przykładów.

Zadanie Wex.1
Znaleźć ekstrema funkcji $f(x, y) = xy$ na zbiorze $E \subset \mathbb{R}^2$ zadanym warunkiem $G(x, y) = x^2 + y^2 - 2 = 0$.
Rozwiązanie: Tworzymy funkcję pomocniczą

$$\tilde{f}(x, y) = xy - \lambda(x^2 + y^2 - 2)\,,$$

Obliczamy pochodne cząstkowe tej funkcji: $\tilde{f}_x = y - 2\lambda x$, $\tilde{f}_y = x - 2\lambda y$ i przyrównujemy je do zera wraz z warunkiem $G = 0$:

$$\begin{aligned} &y - 2\lambda x = 0\,, \\ &x - 2\lambda y = 0\,, \\ &x^2 + y^2 - 2 = 0\,. \end{aligned}$$

Szukamy rozwiązań tego układu równań mnożąc pierwsze przez y drugie przez x i odejmując pierwsze od drugiego. Daje to $x^2 = y^2$, czyli $y = \pm x$. Ostatnie równanie w obu przypadkach daje $2x^2 = 2$, czyli $x = \pm 1$. Są więc aż cztery punkty i każdemu z nich, poprzez pierwsze lub drugie równanie, odpowiada pewna wartość mnożnika λ:

$$\begin{array}{cccc} (1,1)\,, & (1,-1)\,, & (-1,1)\,, & (-1,-1)\,. \\ \lambda = \dfrac{1}{2}\,, & \lambda = -\dfrac{1}{2}\,, & \lambda = -\dfrac{1}{2}\,, & \lambda = \dfrac{1}{2}\,. \end{array}$$

Zbiór $E = G^{-1}(0)$ jest w tym przypadku zbiorem zwartym (i bez brzegu, bo to okrąg) i możemy od razu zobaczyć, że w pierwszym i ostatnim punkcie funkcja f przyjmuje wartość 1, a dwóch środkowych punktach przyjmuje wartość -1. W punkcie pierwszym i ostatnim funkcja f ma więc maksima (globalne na E), a w dwóch środkowych ma minima (też globalne na E). Zobaczmy jednak, jak działają te dwa różne przepisy sprawdzania, czy w swoim punkcie krytycznym funkcja f ma na E ekstremum (lokalne). Najpierw ostatnia (powszechniej stosowana) metoda. Obliczamy macierz drugich pochodnych funkcji $\tilde{f}$ w zmiennych x i y, traktując mnożnik λ jak stałą: $\tilde{f}_{xx} = \tilde{f}_{yy} = -2\lambda$, $\tilde{f}_{xy} = 1$. Macierz formy kwadratowej w pierwszym i w ostatnim punkcie ma więc postać

$$\begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Ogólnie rzecz biorąc jest to macierz o sygnaturze $(-, 0)$, ale wiemy, że to co jest istotne, to to, jakie wartości daje ona na wektorach stycznych do więzów. W punktach $(1, 1)$ i $(-1, -1)$ pochodna $G' = (2x, 2y)$ to macierz (kowektor)

$$G'_{(1,1)} = (2, 2)\,, \qquad G'_{(-1,-1)} = (-2, -2)\,.$$

Składowe (h_1, h_2) wektorów przemieszczeń stycznych do więzów w każdym z tych punktów spełniają taki sam warunek $2h_1 + 2h_2 = 0$. Określoność formy drugich pochodnych funkcji $\tilde{f}$ należy zatem w punktach $(1,1)$ i $(-1,-1)$ badać tylko na wektorach przemieszczeń postaci $(h, -h)$:

$$(h, -h) \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} h \\ -h \end{pmatrix} = -4h^2 \,.$$

Jej wartość jest więc na takich wektorach zawsze ujemna, co oznacza, że w punktach $(1,1)$ oraz $(-1,-1)$ funkcja ma na zbiorze E lokalne maksimum. Analogicznie w środkowych dwóch punktach krytycznych, tj. w $(1,-1)$ i $(-1,1)$, pochodna $G' = (2x, 2y)$ to macierz (kowektor)

$$G'_{(1,-1)} = (2, -2)\,, \qquad G'_{(-1,1)} = (-2, 2)\,,$$

i macierz formy kwadratowej drugich pochodnych funkcji $\tilde{F}$ w tych punktach trzeba badać na wektorach (stycznych do więzów), które są postaci (h, h):

$$(h, h) \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} h \\ h \end{pmatrix} = 4h^2 \,.$$

Wartość tej formy na wektorach przemieszczeń stycznych do więzów jest zawsze dodatnia i wobec tego na E funkcja f ma w tych punktach minima.

Sprawdźmy jeszcze jak działa metoda zalecana na wykładzie. Teraz zajmujemy się funkcją $\tilde{F}(x, y, \lambda) = xy - \lambda(x^2 + y^2 - 2)$ i znajdujemy jej drugie pochodne w trzech zmiennych:

$$\begin{aligned} &\tilde{F}_{xx} = -2\lambda\,, \qquad &&\tilde{F}_{yy} = -2\lambda\,, \qquad &&\tilde{F}_{\lambda\lambda} = 0\,, \\ &\tilde{F}_{xy} = 1\,, &&\tilde{F}_{x\lambda} = -2x\,, &&\tilde{F}_{y\lambda} = -2y\,. \end{aligned}$$

W punkcie $(1,1)$ i w punkcie $(-1,-1)$, gdy $\lambda = 1/2$, macierz 3×3 formy kwadratowej drugich pochodnych mają postacie

$$\begin{pmatrix} -1 & 1 & -2 \\ 1 & -1 & -2 \\ -2 & -2 & 0 \end{pmatrix}, \qquad \begin{pmatrix} -1 & 1 & 2 \\ 1 & -1 & 2 \\ 2 & 2 & 0 \end{pmatrix},$$

i w obu przypadkach wyznacznik jest równy 4, czyli dodatni (maksimum). Z kolei w punktach $(1,-1)$ i $(-1,1)$, gdy $\lambda = -1/2$, macierze te mają postacie

$$\begin{pmatrix} 1 & 1 & -2 \\ 1 & 1 & 2 \\ -2 & 2 & 0 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 1 & 2 \\ 1 & 2 & -1 \\ 2 & -2 & 0 \end{pmatrix},$$

i ich wyznaczniki są równe -4, czyli są ujemne (minima). Widać, że i ta metoda działa!

W przykładzie tym warunek $G(x,y)=0$ daje się oczwiście odwikłać, ale nie globalnie. Wyznacza on dwie funkcje

$$g_+(x)=+\sqrt{2-x^2}\,,\qquad \text{oraz}\qquad g_-(x)=-\sqrt{2-x^2}\,,$$

które są dobrymi funkcjami jedna w otoczeniu punktów $(1,1)$ i $(-1,1)$, a druga w otoczeniu punktów $(1,-1)$ i $(-1,-1)$. Punktami, w których $G(x,y)=0$ nie wyznacza żadnej dobrej funkcji $y=y(x)$, są punkty $(\pm\sqrt{2},\,0)$ (warunek $G(x,y)=0$ wyznacza tam dobre funkcje $x=x(y)$), ale tu na szczęście dwie funkcje $g_+(x)$ i $g_-(x)$ są dobrze zdefiniowane na dwóch kawałkach E obejmujących razem wszystkie punkty krytyczne. Na górnej połowie E_+ zbioru E (górnej, tj. tej, gdzie $y>0$) funkcja $f_+=F|_{E_+}$ może być zapisana jako funkcja

$$f_+(x)=x\,g_+(x)=x\sqrt{2-x^2}\,.$$

Jej pochodna

$$f'_+(x)=\sqrt{2-x^2}-\frac{x^2}{\sqrt{2-x^2}}=\frac{2-2x^2}{\sqrt{2-x^2}}\,,$$

zeruje się w punktach $x=1$ i $x=-1$ (którym odpowiada $y=1$), a jej druga pochodna

$$f''_+(x)=\frac{-4x}{\sqrt{2-x^2}}+\frac{2x\,(1-x^2)}{(2-x^2)^{3/2}}\,,$$

jest w tych punktach odpowiednio ujemna (maksimum lokalne) i dodatnia (minimum lokalne). Analogicznie na E_- mamy funkcję $f_-(x)=x\,g_-(x)$, która różni się od $f_+(x)$ znakiem i wobec tego ma w punktach $x=1$ i $x=-1$ (którym odpowiada $y=-1$) odpowiednio minimum i maksimum.

Zadanie Wex.2

Jaka jest maksymalna możliwa objętość prostopadłościanu wpisanego w elipsę o półosiach równych a, b i c?

Rozwiązanie: Jeśli rozsądnie umieścimy środek elipsy w punkcie $(0,0,0)$ tak, by jej osie główne były osiami x, y i z, to funkcją, której ekstremum trzeba znaleźć jest $V=8xyz$, ale technicznie wygodniej będzie szukać maksimum funkcji $(V/8)^2=x^2y^2z^2$, bo w końcu kwadrat jednej ósmej objętości jest maksymalny wtedy, gdy sama objętość jest też maksymalna. Warunkiem ubocznym jest tu równanie elipsy

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1\,.$$

Można zrobić jeszcze jeden myk polegający na przeskalowaniu zmiennych: zamiast operować x, y i z, będziemy operować zmiennymi $\tilde{x}=x/a$, $\tilde{y}=y/b$, $\tilde{z}=z/c$ (tych tyld nie będziemy pisać; odpowiada to mierzeniu odległości wzdłuż osi x w jednostkach a, wzdłuż osi y w jednostkach b i wzdłuż osi z w jednostkach c - kto nam może tego zabronić?!). Tworzymy więc pomocniczą funkcję

$$\tilde{F}(x, y, x) = x^2y^2z^2 - \lambda\,(x^2 + y^2 + z^2 - 1)\,,$$

(bo $(V/8)^2 = a^2b^2c^2\tilde{x}^2\tilde{y}^2\tilde{z}^2$, a czynnik $a^2b^2c^2$ nie wpływa na położenie maksimum) i postępujemy według regulaminu: przyrównujemy do zera jej pochodne i warunek uboczny

$$\begin{aligned}
&\tilde{F}_x(x, y, z) = 2x\left(y^2z^2 - \lambda\right) = 0\,,\\
&\tilde{F}_y(x, y, z) = 2y\left(x^2z^2 - \lambda\right) = 0\,,\\
&\tilde{F}_z(x, y, z) = 2z\left(x^2y^2 - \lambda\right) = 0\,,\\
&G(x, y, z) = x^2 + y^2 + z^2 - 1 = 0\,.
\end{aligned}$$

Oczywiście rozwiązania z $x = 0$ lub $y = 0$ lub $z = 0$ odpowiadają zerowej objętości, czyli minimum (absolutnemu) funkcji F i są nieinteresujące. Można więc trzy pierwszr równania zastąpić równaniami $y^2z^2 = \lambda$, $x^2z^2 = \lambda$ i $x^2y^2 = \lambda$. Jeśli te z kolei poodejmować parami jedno od drugiego, to dostaniemy $x^2 = y^2 = z^2$. Znów, ujemnych x, y i z nie rozpatrujemy - funkcje F i $\tilde{F}$ są funkcjami tylko x^2, y^2 i z^2, więc i tak, nawet gdyby nie myśleć o x, y i z jak o długościach boków prostopadłościanu, to punkty krytyczne o ujemnych warościach tych współrzędnych miałyby taki sam chakter, jak te o dodatnich tylko - więc

$$x = y = z = \frac{1}{\sqrt{3}}\,, \qquad \lambda = \frac{1}{9}\,.$$

Drugie pochodne funkcji $\tilde{F}$ są następujące:

$$\begin{aligned}
&\tilde{F}_{xx}(x, y, z) = 2(y^2z^2 - \lambda), \quad &&\tilde{F}_{yy}(x, y, z) = 2(x^2z^2 - \lambda), \quad &&\tilde{F}_{zz}(x, y, z) = 2(x^2y^2 - \lambda),\\
&\tilde{F}_{xy}(x, y, z) = 4xyz^2\,, &&\tilde{F}_{xz}(x, y, z) = 4xzy^2\,, &&\tilde{F}_{yz}(x, y, z) = 4yzx^2\,,
\end{aligned}$$

i macierz Q formy kwadratowej drugich pochodnych w punkcie krytycznym ma postać

$$Q = \frac{4}{9}\begin{pmatrix} 0 & 1 & 1\\ 1 & 0 & 1\\ 1 & 1 & 0 \end{pmatrix}\,,$$

Forma nie jest więc bezwzględnie dodatnio określona, ale na wektorach stycznych musi taka być, bo zbiór E jest zbiorem zwartym, jakieś maksimum funkcja F na nim mieć musi, a jedynym kandydatem jest znaleziony punkt krytyczny. Ale zobaczmy na tym przykładzie, jak działa w praktyce przedstawione w Regulaminie redukowanie formy do formy określonej na wektorach stycznych tylko. Na dowolnym wektorze przemieszczenia $\mathbf{h} = (h_x, h_y, h_z)$ (pominiemy dla przerzystości te $4/9$)

$$Q(\mathbf{h}) = 2h_xh_y + 2h_xh_z + 2h_yh_z\,.$$

Pochodna $G' = (2x, 2y, 2z)$ warunku $G(x, y, z) = x^2 + y^2 + z^2 - 1$ działając w punkcie krytycznym na wektor $\mathbf{h}$ przemieszczenia zgodny z więzami ma dawać zero. Zatem wektory takie muszą spełniać warunek

$$\frac{2}{\sqrt{3}}\,(1,\,1,\,1)\begin{pmatrix} h_x \\ h_y \\ h_z \end{pmatrix} = 0\,.$$

czyli np. $h_z = -h_x - h_y$. Wstawiamy to do $Q(\mathbf{h})$ i porządkujemy:

$$\begin{aligned} Q &= 2h_xh_y + 2(h_x + h_y)(-h_x - h_y) \\ &= -2h_x^2 - 2h_y^2 - 2h_xh_y = (h_x, h_y)\begin{pmatrix} -2 & -1 \\ -1 & -2 \end{pmatrix}\begin{pmatrix} h_x \\ h_y \end{pmatrix}. \end{aligned}$$

I teraz metoda "minorowa" mówi, że $-Q$ jest formą dodatnio określoną, więc Q jest formą ujemnie określoną, a to znaczy, że w punkcie $x = y = z = 1/\sqrt{3}$ funkcja F ma na E maksimum, tak jak to już ustaliliśmy fizycznym rozumowaniem. Jeśli teraz odskalujemy spowrotem zmienne, to wynik jest taki, że maksymalną objętość ma wpisany w elipsę o półosiach a, b i c prostopadłościan o $x = a/\sqrt{3}$, $y = b/\sqrt{3}$, $z = c/\sqrt{3}$. Maksymalna objętość prostopadłościanu wpisanego w taką elipsę jest równa $8abc/3\sqrt{3}$.

Można też na tym przykładzie sprawdzić metodę z wykładu ustalania charakteru punktu krytycznego. Jeśli $\tilde{F}$ potraktujemy jak funkcję x, y, z i λ, to dojdą drugie pochodne $\tilde{F}_{x\lambda} = -2x$, $\tilde{F}_{y\lambda} = -2y$, $\tilde{F}_{z\lambda} = -2z$, $\tilde{F}_{\lambda\lambda} = 0$. W punkcie krytycznym macierz drugich pochodnych wygląda wtedy tak

$$Q = \frac{4}{9}\begin{pmatrix} 0 & 1 & 1 & -\alpha \\ 1 & 0 & 1 & -\alpha \\ 1 & 1 & 0 & -\alpha \\ -\alpha & -\alpha & -\alpha & 0 \end{pmatrix}, \qquad \alpha = \frac{3}{2}\sqrt{3}\,.$$

Wyznacznik tej macierzy (no trzeba trochę poLaplasować) jest równy $-3\alpha^2(4/9)^4$. Zatem widać na tym przykładzie, że kryterium podane na wykładzie nie działa w przypadku funkcji na $\mathbb{R}^n$ o $n > 2$ i/lub większej liczby warunków ubocznych.

Również i to zadanie można rozwiązać zwykłą metodą szukania ekstremum (maksimum) funkcji dwóch zmiennych. Można bowiem z równania elipsy wywikłać z^2 i szukać ekstremów funkcji

$$V^2 = F(x, y) = 64x^2y^2z^2 = 64c^2x^2y^2\left(1 - \frac{x^2}{a^2} - \frac{y^2}{b^2}\right).$$

Znów lepiej jest zdefiniować nowe zmienne $u = x^2/a^2$, $v = y^2/b^2$, i szukać w obszarze $0 \le u \le 1$, $0 \le v \le 1$ ekstremum funkcji

$$f(u, v) = uv(1 - u - v) = uv - uv^2 - u^2v\,.$$

Szukamy jej puktów krytycznych rozwiązując równania

$$f_u = v - 2uv - v^2 = 0\,,$$
$$f_v = u - 2uv - u^2 = 0\,.$$

Odejmując jedno od drugiego dostajemy $(u - v) - (u^2 - v^2) = (u - v)[1 - u - v] = 0$. Zatem albo $u = v$ i wtedy $u - 3u^2 = 0$ i $u = 0$ lub $u = 1/3$, albo $v = 1 - u$ i to podstawione np. do $f_v = 0$ da $u = 1$ (i $v = 0$) lub $u = 0$ (i $v = 1$). $u = 0$ lub $v = 0$ oznacza zerową objętość, czyli minimum funkcji, i te możliwości odrzucamy. Zostaje więc $u = v = 1/3$. Drugie pochodne $f_{uu} = -2v$, $f_{vv} = -2u$ i $f_{uv} = 1 - 2u - 2v$ dają w badanym punkcie krytycznym następującą macierz Q formy kwadratowej drugich pochodnych

$$Q = \begin{pmatrix} -2/3 & -1/3 \\ -1/3 & -2/3 \end{pmatrix},$$

która jest ujemnie określona (sygnatura $(-,-)$), co oczywiście oznacza, że w punkcie krytycznym funkcja ma maksimum (tu jest to maksimum absolutne). Oczywiście dostajemy w wyjściowych zmiennych te same $x = 1/\sqrt{3}$, $y = b/\sqrt{3}$ i $z = c/\sqrt{3}$.

Zadanie Wex.3 (część zadania 82 ze skryptu do algebry)
Znaleźć odległość w $\mathbb{E}_3$ między prostymi l_1 i l_2 zadanymi następująco:

$$l_1: \qquad \{(x_1, x_2, x_3) \in \mathbb{E}_3 : \;\; x_1 + x_2 = 1,\; x_1 + 2x_2 + x_3 = 2\}\,.$$

$$l_2: \qquad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix} + t \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}.$$

Rozwiązanie: W skrypcie do algebry to zadanie jest też rozwiązane czysto geometrycznie, ale tu rozwiążemy je analitycznie. Wprowadzamy w tym celu funkcję

$$f(x_1, x_2, x_3, t) = (x_1 - t)^2 + (x_2 - t - 2)^2 + (x_3 - 2t)^2\,,$$

będącą kwadratem odległości punktu X o współrzędnych (x_1, x_2, x_3) od punktu na prostej l_2 scharakteryzowanego parametrem t. Minimalizujemy zatem funkcję czterech zmiennych. Warunkiem dodatkowym jest to, że punkt X musi leżeć na prostej l_1, co oznacza, że współrzędne (x_1, x_2, x_3) muszą spełniać warunki (uprościliśmy tu drugi z warunków zadających prostą l_2 odejmując odeń pierwszy)

$$g_1(x_1, x_2, x_3, t) = x_1 + x_2 - 1 = 0\,,$$
$$g_2(x_1, x_2, x_3, t) = x_2 + x_3 - 1 = 0\,,$$

(chociaż warunki g_1 i g_2 dotyczą tylko współrzędnych punktu X, to mimo to, należy je formalnie traktować jak funkcje wszystkich zmiennych, ze względu na które minimalizujemy funkcję f). Znów tworzymy funkcję pomocniczą zależną od dwu mnożników Lagrange'a λ_1 i λ_2:

$$F(x_1, x_2, x_3, t) = f(x_1, x_2, x_3, t) + 2\lambda_1 g_1(x_1, x_2, x_3, t) + 2\lambda_2 g_2(x_1, x_2, x_3, t)\,,$$

(żeby się ładniej liczby komponowały przyjęliśmy za mnożniki Lagrange'a $2\lambda_1$ i $2\lambda_2$) i przyrównujemy do zera jej pochodne cząstkowe po x_1, x_2, x_3 i t:

$$\begin{aligned}
F'_{x_1} &= 2(x_1 - t) + 2\lambda_1 = 0\,,\\
F'_{x_2} &= 2(x_2 - t - 2) + 2\lambda_1 + 2\lambda_2 = 0\,,\\
F'_{x_3} &= 2(x_3 - 2t) + 2\lambda_2 = 0\,,\\
F'_{t} &= -2(x_1 - t) - 2(x_2 - t - 2) - 4(x_3 - 2t) = 0\,.
\end{aligned}$$

W połączeniu z warunkami ubocznymi daje to układ sześciu równań

$$\begin{aligned}
&x_1 - t + \lambda_1 = 0\,,\\
&x_2 - t + \lambda_1 + \lambda_2 = 2\,,\\
&x_3 - 2t + \lambda_2 = 0\,,\\
&x_1 + x_2 + 2x_3 - 6t = 2\,,\\
&x_1 + x_2 = 1\,,\\
&x_2 + x_3 = 1\,.
\end{aligned}$$

Aby je systematycznie rozwiązać wyznaczamy z pierwszych trzech $x_1 = t - \lambda_1$, $x_2 = t - \lambda_1 - \lambda_2 + 2$, $x_3 = 2t - \lambda_2$ i wstawiamy do pozostałych trzech. Pierwsze z nich daje wtedy

$$2\lambda_1 + 3\lambda_2 = 0\,,$$

a pozostałe

$$\begin{aligned}
2t - 2\lambda_1 - \lambda_2 &= -1\,,\\
3t - \lambda_1 - 2\lambda_2 &= -1\,.
\end{aligned}$$

Po wyeliminowaniu λ_1 otrzymujemy dwa równania

$$\begin{aligned}
2t + 2\lambda_2 &= -1\,,\\
3t - \frac{1}{2}\lambda_2 &= -1\,,
\end{aligned}$$

których rozwiązaniem są $t = -\frac{5}{14}$, $\lambda_2 = -\frac{2}{14}$; dalej już łatwo: $\lambda_1 = \frac{3}{14}$ oraz

$$x_1 = x_3 = -\frac{8}{14}\,, \qquad x_2 = \frac{22}{14}\,.$$

Wartość minimalizowanej funkcji $f(x_1, x_2, x_3, t)$ w tym punkcie wynosi $\frac{1}{14}$ (tak jak nam to wyszło w zadaniu 81). Macierz Q drugich pochodnych w zmiennych x_1, x_2, x_3 i t funkcji $F(x_1, x_2, x_3, t, \lambda_1, \lambda_2)$ ma postać

$$Q = \begin{pmatrix} 2 & 0 & 0 & -2 \\ 0 & 2 & 0 & -2 \\ 0 & 0 & 2 & -4 \\ -2 & -2 & -4 & 12 \end{pmatrix}.$$

Nie jest ona dodatnio określona bo największy jej minor jest równy zeru (aby to zobaczyć wystarczy do ostatniego wiersza dodać dwa pierwsze wiersze i podwojony trzeci, co da macierz górnotrójkątną mającą na diagonali trzy dodatne wyrazy i jeden zerowy). Ale macierz tę trzeba badać na wektorach stycznych. Ponieważ gradientami więzów $g_1(x_1, x_2, x_3, x_4) = 0$ i $g_2(x_1, x_2, x_3, x_4) = 0$ są tu kowektory

$$g_1' = (1,\, 1,\, 0,\, 0)\,, \qquad g_2' = (0,\, 1,\, 1,\, 0)\,,$$

najogólniejszy wektor styczny do więzów jest postaci

$$\mathbf{h} = \begin{bmatrix} h \\ -h \\ h \\ k \end{bmatrix},$$

na tym wektorze i forma drugich pochodnych daje

$$Q(\mathbf{h}) = 6h^2 - 6hk + 12k^2 = 6\,(h, k) \begin{pmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & 2 \end{pmatrix} \begin{pmatrix} h \\ k \end{pmatrix}.$$

Widać więc (metodą minorową np.), że jest to wielkość zawsze dodatnia. Znaleziony punkt krytyczny jest wiec minimum funkcji f.

Możemy też wykorzystać to, co zrozumieliśmy przy okazji dyskusji warunków dostatecznych istnienia ekstremów warunkowych (konstrukcja wektorów przemieszczenia stycznych do powierzchni więzów) i rozwiązać takie kształcące zadanie.

Zadanie Wex.4
Znaleźć wektory styczne do zanurzonej w $\mathbb{R}^3$ powierzchni (będącej elipsoidą) zadanej równaniem

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1\,.$$

Rozwiązanie: Traktujemy to równanie jak warunek więzów w poprzednich zadaniach

$$G(x, y, z) = \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} - 1 = 0\,,$$

i teraz wiemy, że w dowolnym punkcie tej powierzchni tj. punkcie $P \equiv (x, y, z)$ takim, że $G(x, y, z) = 0$, wektory styczne $\mathbf{t}_P$ mają tę właściwość, że zeruje się na nich kowektor (forma liniowa), którym jest (tłumaczyliśmy to sobie jakiś czas temu) pochodna G' obliczona w punkcie P, czyli

$$G'_P \equiv G'(x, y, z) = (2x/a^2,\ 2y/b^2,\ 2z/c^2)\,.$$

Jeśli wektor $\mathbf{t}_P$ ma (w kanonicznej bazie zero-jedynkowej) składowe t_P^x, t_P^y i t_P^z, to muszą one spełniać warunek

$$2x\, t_P^x/a^2 + 2y\, t_P^y/b^2 + 2z\, t_P^z/c^2 = 0\,.$$

Jest to jeden warunek na trzy składowe, więc w każdym punkcie $P = (x, y, z)$ są (znów algebra!) dwa liniowo niezależne takie wektory rozpinające razem w $\mathbb{R}^3$ dwuwymiarową podprzestrzeń oznaczaną T_PM (M oznacza tu "manifold", po naszemu 'rozmaitość", jaką jest elipsoida, a T jest od "tangent"). Wektory te można oczywiście wybrać na wiele sposobów. Jednak jeśli wprowadzimy jakąś parametryzację powierzchni (choćby tylko lokalną), czyli mówiąc językiem geometrii różniczkowej, układ współrzędnych na rozmaitości M w otoczeniu jej punktu P, to każdy taki układ współrzędnych (parametryzacja) w naturalny sposób wyróżnia w punkcie P pewną bazę przestrzeni T_PM. Zobaczmy to na przykładzie. Wprowadźmy dwa parametry $0 \leq \theta \leq \pi$ i $0 \leq \varphi < 2\pi$ wzorami

$$\begin{aligned} x &= a \sin\theta \cos\varphi \equiv a\, s_\theta\, c_\varphi\,, \\ y &= b \sin\theta \sin\varphi \equiv b\, s_\theta\, s_\varphi\,, \\ z &= c \cos\theta \qquad\quad \equiv c\, c_\theta\,. \end{aligned}$$

Jest jasne, że dla dowolnych wartości θ i φ (z podanego ich zakresu) otrzymuje się x, y i z spełniające warunek $G(x, y, z) = 0$. To właśnie oznacza sparametryzować rozmaitość M (zanurzoną w $\mathbb{R}^n$; można to bardziej uabstrakcyjnić, ale są od tego inni specjaliści) zadaną jakimś warunkiem lub warunkami. Łatwo też sprawdzić, że dwa wektory (zapisane tu w kolumienkach, zawierających ich składowe w kanonicznej zero-jedynkowej bazie $\mathbf{e}_x$, $\mathbf{e}_y$, $\mathbf{e}_z$ przestrzeni $\mathbb{R}^3$, w której wszystko tu się rozgrywa - uabstrakcyjnienie polega m.in. na tym, że można się obywać bez takiej przestrzeni $\mathbb{R}^n$, ale tego już nie musimy tu zgłębiać)

$$\mathbf{t}_\theta = \begin{pmatrix} a\, c_\theta\, c_\varphi \\ b\, c_\theta\, s_\varphi \\ -c\, s_\theta \end{pmatrix}, \qquad \mathbf{t}_\varphi = \begin{pmatrix} -a\, s_\theta\, s_\varphi \\ b\, s_\theta\, c_\varphi \\ 0 \end{pmatrix},$$

spełniają automatycznie warunki $G'_P \cdot \mathbf{t}_\theta = 0$, $G'_P \cdot \mathbf{t}_\varphi = 0$, czyli są wektorami stycznymi do powierzchni w punkcie P (identyfikowanym teraz wartościami parametrów θ i φ).

A jak te wektory zostały otrzymane? A no po prostu pierwszy przez różniczkowanie związków definiujących układ współrzędnych po θ, a drugi przez ich różniczkowanie po φ. Dlatego też fachowi geometrzy różniczkowi oznaczają te wektory tak:

$$\mathbf{t}_\theta = \frac{\partial}{\partial\theta}\,, \qquad \mathbf{t}_\varphi = \frac{\partial}{\partial\varphi}\,.$$

Nie będziemy tu ich naśladować, żeby nas te "gołe" pochodne nie straszyły po nocach (mogą nas gonić i chcieć zróżniczkować!). Ale patent na konstruowanie wektorów stycznych do rozmaitości, gdy jest ona sparametryzowana jakimiś współrzędnymi, jest uniwersalny. Fizyk może go zresztą łatwo przyswoić bez konsultowania się z fachowymi geometrami. Wystarczy wyobrazić sobie, że zmieniamy np. parametr θ trzymając parametr φ ustalony. Wytyczamy w ten sposób w przestrzeni $\mathbb{R}^3$ pewną krzywą leżącą oczywiście na powierzchni elipsoidy (na rozmaitości). No i teraz możemy sobie wyobrazić, że θ to jest taki parametr jak czas t w mechanice, który też parametryzuje trajektorię. A każdy wie, że pochodna położenia po czasie to jest wektor prędkości, który jest zawsze styczny do toru. I tu jest tak samo!

Dodajmy jeszcze, że taki układ współrzędnych, czyli odwzorowanie $\mathbb{R}^k$ w $\mathbb{R}^n$, gdzie $k < n$, definiuje w $\mathbb{R}^n$ powierzchnię (rozmaitość), która jest *regularna* (tzn. ma wymiar k) w danym punkcie P, gdy wszystkie k wektorów stycznych uzyskanych z powyższego przepisu tworzy w punkcie P układ wektorów liniowo niezależych, bo jeśli nie są one liniowo niezależne, to znaczy, że rozpinana przez nie (pod)przestrzeń styczna ma wymiar mniejszy niż k, a to znaczy, że w tym punkcie powierzchnia (rozmaitość) ma jakąś osobliwość, czyli robi jakieś siupy...

ODE, czyli równania różniczkowe zwyczajne.

Równania różniczkowe to jest obszerny dział matematyki, który kiedyś, gdy fizyka sprowadzała się do mechaniki klasycznej, elektrodynamiki, elastomechaniki, hydrodynamiki itd., był podstawowym działem dla każdego fizyka (no, przede wszystkim teoretyka). Dziś trzeba umieć i algebrę[37] i geometrię różniczkową i coś wiedzieć o analizie funkcjonalnej i o topologii i analizie zespolonej, teorii prawdpodobieństwa, teorii grup... Ale równania różniczkowe pozostają wciąż jakąś wspólną dla wszystkich podstawą.

Równania różniczkowe mogą być różne. Przede wszystkim można je podzielić na ODE i PDE. ODE, czyli równania różniczkowe zwyczajne, są równaniami wyznaczającymi nieznaną funkcję jednej zmiennej na podstawie zadanego związku tej funkcji z jej pochodną lub pochodnymi i podanego warunku lub warunków (zwanych naogół warunkami początkowymi). Nazywają się zwyczajnymi, bo w związkach tych wystepują pochodne zwyczajne. ODE można jeszcze dalej podzielić na równania wyznaczające funkcje odwzorowujące $\mathbb{R}$ w $\mathbb{R}$ i takie, które wyznaczają funkcje odwzorowujące $\mathbb{R}$ w $\mathbb{R}^n$ o $n > 1$. Drugi ich podział, to na równania różniczkowe pierwszego rzędu (są one związkami między funkcją i jej pierwszą pochodną) i wyższych rzędów (tu wiążą one funkcję z jej wyższymi pochodnymi aż do pewnego skończonego rzędu r). Jak się przekonamy, równanie różniczkowe rzędu r wiążące funkcję odwzorowującą $\mathbb{R}$ w $\mathbb{R}$, można zawsze przekształcić w równanie różniczkowe pierwszego rzędu na funkcję odwzorowującą $\mathbb{R}$ w $\mathbb{R}^r$ (ogólniej: ODE rzędu r na funkcję odwzorowującą $\mathbb{R}$ w $\mathbb{R}^n$ można zapisać jako ODE pierwszego rzędu na funkcję odwzorowującą $\mathbb{R}$ w $\mathbb{R}^{n\cdot r}$). Pewną szczególną klasę równań stanowią równania zwyczajne pierwszego rzędu o zmiennych rozdzielonych, a drugą (klasy te nie są rozłączne) równania liniowe (mogą być i wyższego rzędu), tzn. takie, w których związek funkcji i jej pochodnych jest liniowy w samej funkcji; rozszerzeniem tej klasy są równania liniowe z niejednorodnością, tzn. mające postać równości pewnego operatora (różniczkowego) działającego na funkcję i pewnej zadanej (ustalonej z góry) funkcji. Te dwie klasy równań są szczególnie łatwe i nimi się będziemy tu sporo zajmować.

Osobny dział stanowią PDE - czyli równania różniczkowe wiążące odwzorowania z $\mathbb{R}^n$ w $\mathbb{R}$ (można też rozpatrywać takie dotyczące odwzorowania z $\mathbb{R}^n$ w $\mathbb{R}^m$ - takimi są np. znane każemu fizykowi równania Maxwella) z ich pochodnymi cząstkowymi pierwszego rzędu tylko - wtedy mamy do czynienia z PDE pierwszego rzędu - albo i wyższego rzędu. Okazuje się, że PDE pierwszego rzędu daje się sprowadzić do równań różniczkowych zwyczajnych. Zwie się to metodą charakterystyk. Działa ona trochę prościej, gdy rozpatrywane PDE jest liniowe, lub liniowe z niejednorodnością, ale stosuje się ona także do nieliniowych PDE pierwszego rzędu. Kiedyś tę metodę sobie przyswoiłem (w wersji stosowalnej do liniowych PDE pierwszego rzędu jest ona potrzebna przy analizie tzw. grupy renormalizacji w kwantowej teorii pola i fizyce statystycznej) więc jakieś dwa proste przykłady jej zastosowania będą na deser (tylko nie mówić wykładowcy...). Dalej

[37] Z jaką rezerwą podchodzili fizycy do algebry, kiedy nagle okazało się, że mechanikę kwantową można - to zrobił Heisenberg - sformułować w języku takich dziwnych obiektów, jak macierze! Dopiero wtedy w Getyndze (a gdzieżby indziej to się mogło wydarzyć, jak nie tam, gdzie był patronat Hilberta?) usiedli Max Born z Pascualem Jordanem i zaczęli się uczyć od tamtejszych znakomitych matematyków tego, co Państwo dziś już mają w małym palcu...

rozciąga się po horyzont, i pewnie nawet dalej, świat skomplikowanych równań różniczkowych cząstkowych wyższych rzędów; te drugiego rzędu się jeszcze jakoś klasyfikuje na równania typu parabolicznego, hiperbolicznego itd., a co z wyższego rzędu, to już tylko spece wiedzą. To tak, żebyśmy mieli jakąś świadomość, w co się zagłębiamy. "Und somit fangen wir an", jak mówi Thomas Mann kończąc wstęp do *Der Zauberberg*.

Zaczniemy oczywiście od równań różniczkowych zwyczajnych pierwszego rzędu wiążących nieznaną funkcję $y = y(x)$ z jej pierwszą pochodną $y'(x)$. Związek taki może być jawny

$$y'(x) = f(x, y)\,,$$

zakłada się wtedy, że $f(x, y)$ jest odwzorowaniem ciągłym pewnego (otwartego) zbioru $D \subset \mathbb{R}^2$ w $\mathbb{R}$, lub mieć postać uwikłaną

$$F(x, y, y') = 0\,,$$

o której zakłada się, że $F(x, y, z)$ jest odwzorowaniem ciągłym pewnego (otwartego) zbioru $E \subset \mathbb{R}^3$ w $\mathbb{R}$. Jak już wiemy, przy założeniu, że $F_z \neq 0$, równość $F = 0$ wyznacza (lokalnie) funkcję $z = z(x, y)$, czyli daje się wtedy postać drugą sprowadzić do pierwszej, więc zajmiemy się głównie tą pierwszą.

Całką równania takiego jak te wyżej nazywa się każdą różniczkowalną (na pewnym podzbiorze $\mathbb{R}$) funkcję $y = y(x)$ taką, że jej wykres (takie szkolne pojęcie) leży w zbiorze $D \subset \mathbb{R}^2$ i

$$y'(x) - f(x, y(x)) \equiv 0\,, \qquad \text{lub} \quad F(x, y(x), y'(x)) \equiv 0\,.$$

Wykres $y = y(x)$ jest *krzywą całkową równania różniczkowego*, ale, jak to sobie zaraz powiemy, mogą być też krzywe całkowe równania, które nie są wykresem funkcji $y = y(x)$. Powstaje zaraz oczywiste dla matematyka pytanie, czy równanie różniczkowe ma rozwiązanie przechodzące przez zadany punkt (x_0, y_0) płaszczyzny $\mathbb{R}^2$, tj. takie, że $y(x_0) = x_0$, a jak ma, to czy jednoznaczne i jak zbiór takich rozwiązań przechodzących przez różne punkty skatalogować. Fizyk formułuje ten pierwszy problem w postaci pytania o rozwiązanie spełniające zadany *warunek początkowy* $y(x_0) = x_0$ (mało się zastanawiając nad jednoznacznością).

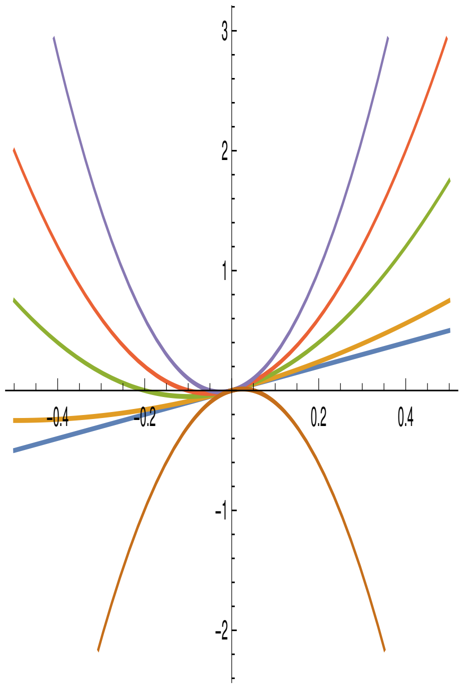
Zadanie Ode.1 Zanim przytoczymy stosowne twierdzenia, rozpatrzmy taki przykład (z nieocenionego Lejka). Niech równanie ma postać (tę drugą)

$$x\,y' + x - 2y = 0\,.$$

Można sprawdzić przez bezpośrednie podstawienie do wypisanego wyżej równania, że rozwiązaniem jego jest funkcja[38]

$$y(x) = C\,x^2 + x\,,$$

[38]Za niedługo stanie sie jasne, skąd się ona bierze. Rozpatrywane równanie jest bowiem przykładem równania liniowego z niejednorodnością.



Rysunek 16: Będące uczciwymi funkcjami krzywe całkowe równania $y' = 2(x/y) - 1$.

w której C jest zupełnie dowolną stałą (rzeczywistą). Jest więc to cała rodzina całek wypisanego równania różniczkowego. Jeśli się spojrzy na rysunek 16, to widać, że gdy $C > 0$, jest to rodzina parabol idących "do góry", a gdy $C < 0$, rodzina parabol idących "w dół", ale każda z tych parabol przechodzi przez punkt $(0, 0) \in \mathbb{R}^2$; gdy $C = 0$, jest to prosta rozdzielająca jakby te dwie rodziny parabol. Jest więc mniej więcej oczywiste, że przez każdy punkt płaszczyzny $\mathbb{R}^2$ z wyjątkiem punktów $(0, y)$ przechodzi tylko jedna z tych parabol; przez punkt $(0, 0)$ przechodzi ich nieskończenie wiele (każda z całek tego równania), a przez punkty $(0, y)$ o $y \neq 0$ nie przechodzi żadna krzywa całkowa. Widać więc, że punkty $(0, y)$ są jakoś nienormalne: czasem mogą dziać się jakieś takie siupy. Spróbujemy jednak takie zachowanie rozwiązań zrozumieć po "fizycznemu".

Stosowne twierdzenie wyjęte z Lejka i odnoszące się do pierwszego, jawnego, sformułowania problemu, brzmi tak: Jeśli odwzorowanie $f(x, y)$ i jego pochodna $f_y(x, y)$ są ciągłe w punkcie $(x, y) \in \mathbb{R}^2$, to przez punkt ten (i jakieś jego otoczenie) przechodzi dokładnie jedna krzywa całkowa równania $y' = f(x, y)$. (Jak to wyrazić w przypadku równania danego w sprosób uwikłany jako $F(x, y, y') = 0$, powinno być jasne, gdy się wie, kiedy warunek $F(x, y, z) = 0$ wyznacza regularną funkcję $z = z(x, y)$).

W świetle tego Lejkowego twierdzenia jest jasne, że punkty $(0, y)$ są, w przypadku rozpatrywanego wyżej przykładu, trefne: funkcja (uzyskana po przedstawieniu równania w pierwszej postaci)

$$f(x, y) = \frac{2y}{x} - 1\,,$$

jako odwzorowanie $f:\mathbb{R}^2 \longrightarrow \mathbb{R}$, nie jest (to już wiemy! - po to uczyliśmy się analizować ciągłość takich odwzorowań) ciągła w żadnym z punktów postaci $(0,y)$. Przytoczone twierdzenie nie rzuca jednak światła na pytanie, dlaczego punkt $(0,0)$ jest inny niż punkty $(0,y)$ o $y \neq 0$. Sprawa wyjaśnia się trochę (ale nie do końca), jeśli popatrzymy na funkcję $f(x,y)$ jak na mającą zadawać na płaszczyźnie $\mathbb{R}^2$ *pole kierunków*, albo pole nachyleń krzywej $y=y(x)$ - wiemy bowiem, że pochodna $y'(x)$ jest nachyleniem takiej krzywej w punkcie x, czyli tangensem kąta nachylenia w stosunku do osi x prostej stycznej do wykresu tej krzywej w punkcie (x,y). Krzywa całkowa równania różniczkowego ma w danym punkcie (x,y) mieć nachylenie takie, jak zadaje funkcja $f(x,y)$. No a jaką wartość ma w punktach $(0,y)\in\mathbb{R}^2$ funkcja $f(x,y)=2y/x-1$, czyli jakie nachylenie ona w tych punktach zadaje? No, równe $\pm\infty$, gdy $y\neq 0$, czyli takiej funkcji $y=y(x)$ przechodzącej przez punkt $(0,y)$ nie może być. A w punkcie $(0,0)$? Dowolne! Bo wiemy, że gdy do punktu $(0,0)$ zbiegamy na płaszczyźnie $\mathbb{R}^2$ z różnych kierunków po prostych, to dostajemy skończoną wartość funkcji $f(x,y)=2x/y-1$, tyle że zależną od kierunku zbiegania. Można więc zrozumieć, dlaczego przez punkty $(0,y)$ o $y\neq 0$ nie przechodzi żadna krzywa całkowa postaci $y=y(x)$ badanego równania, ale jeszcze nie jest jasne, dlaczego wszystkie takie krzywe przechodzą przez punkt $(0,0)$. Na razie rozumiemy tylko, że ma to jakoś związek z tym, że funkcja $f(x,y)$ daje się w punkcie $(0,0)$ "uciąglić" w kierunkach "po prostych". Jednak samo to jeszcze nie wystarcza, bo, jak się okazuje, działa to tylko tylko w jedną stronę: jeśli przez dany punkt płaszczyzny $\mathbb{R}^2$ przechodzą jakieś krzywe całkowe, to z tego wynika tylko, że w tym punkcie funkcja daje się w jakichś kierunkach uciąglić (tu w dowolnym kierunku); w szczególności zdarza się (są takie przykłady w zadaniach do samodzielnej zabawy), że przez dany punkt przechodzi nieskończenie wiele rozwiązań $y=y(x)$ równania różnieczkowego $y'=f(x,y)$ i każde z nich ma w tym punkcie inne nachylenie (w tym przykładzie tu wszystkie rozwiązania przechodzące przez $(0,0)$ mają w tym punkcie to samo nachylenie) - to oznacza, że funkcję $f(x,y)$ można uciąglić w tym punkcie w dowolnym kierunku (tzn. nie może być ona uczciwie ciągła w tym punkcie jako funkcja $f(x,y)$ na $\mathbb{R}^2$). Jednak badany tu przykład pokazuje, iż z faktu, że funkcja daje się uciąglić w danym punkcie w dowolnym kierunku nie wynika, że będzie przez ten punkt przechodzić nieskończenie wiele rozwiązań o różnych w tym punkcie nachyleniach ani że w ogóle jakieś będą przezeń przechodzić (przykłady są w Zadaniu 38c, d i f). Warunek ciągłości $f_y(x,y)$ też jest istotny. Kluczowe tu jest także to, żeby, gdy do punktu $(0,0)$ zbiegamy po prostej $y=ax$, tzn. mającej w $(0,0)$ nachylenie o tangensie równym a, funkcja $f(x,y)$ stawała się w tym kierunku ciągła jeśli nadamy jej w $(0,0)$ wartość dokładnie a. Te dwie rzeczy: nachylenie kierunku zbiegania i konieczna do ciągłości wartość funkcji w tym punkcie nie muszą być ty samym! I tu właśnie mamy tego ilustrację: gdy zbiegamy do $(0,0)$ po prostej $y=x$ o nachyleniu 1, to zgadza się ono z wartością uciąglającą funkcję $f(x,y)$ w tym kierunku; przy zbieganiu zaś po prostej $y=ax$ z $a\neq 1$ wartość uciąglająca fukcję $f(x,y)$ w takim kierunku jest równa $2a-1\neq a$. I dlatego wszystkie całki badanego równania przechodzące przez punkt $(0,0)$ mają w $x=0$ nachylenie równe 1 (widać więc, że mamy jakieś wytłumaczenie tego, dlaczego przez punkt $(0,0)$ przechodzą krzywe o nachyleniu 1, ale nie tego, dlaczego jest ich nieskończenie wiele...).

Ponadto fakt, że w punktach $(0, y)$ o $y \neq 0$ nachylenie krzywej całkowej musiałoby być nieskończone, pozwala spojrzeć na równanie różniczkowe nieco inaczej i rozszerzyć zbiór jego rozwiązań. Jeśli bowiem przepiszemy równanie w postaci

$$dy = f(x, y)\, dx\,,$$

albo postaci

$$p(x, y)\, dx + q(x, y)\, dy = 0\,,$$

w której $p(x, y)$ i $q(x, y)$ są jakimiś dwiema funkcjami ograniczonymi tylko warunkiem, by $-p(x, y)/q(x, y) = f(x, y)$ (tam, gdzie $f(x, y)$ jest dobrze określona i ciągła), i popatrzymy na nie jak na związek korelujący "kroczki" dy i dx, o jakie wolno nam przesunąć się z punktu (x, y) w bok (już ten punkt widzenia w tym skrypcie wykorzystywaliśmy - nie na darmo ja mam nawyki wyrobione na termodynamice i kwantowej teorii pola!), to tym samym dopuścimy do gry rozwiązania - dalej będziemy je zwać krzywymi całkowymi[39] równania różniczkowego - które nie są (przynajmniej nie wszędzie) funkcjami $y = y(x)$, tylko jakimiś krzywymi na płaszczyźnie xy. Można wtedy jako rozwiązania rozpatrywać krzywe zapisane w postaci parametrycznej $x = x(\tau)$, $y = y(\tau)$; spełniają one wtedy równanie

$$p(x(\tau),\, y(\tau))\, \frac{dx}{d\tau} + q(x(\tau),\, y(\tau))\, \frac{dy}{d\tau} = 0\,.$$

Jeśli badane tu równanie napiszemy w postaci

$$x\, dy = (2y - x)\, dx\,,$$

to zobaczymy, że krzywą całkową jest też prosta $x = 0$. Tzn. możemy rozpocząć spacer z dowolnego punktu na osi y, ale powyższy związek mówi, że kroczek wykonywany z takiego punktu musi być o $dx = 0$ - nie wolno iść ani o włos w prawo lub lewo - tylko wzdłuż osi y! Zatem i przez punkty położone na osi y, punkty $(0, y)$ o $y \neq 0$ przechodzi krzywa całkowa, tylko nie jest ona funkcją w szkolnym znaczeniu. Oczywiście z punktu $(0, 0)$ nadal udaje się ruszyć w dowolnym kierunku, bo w tym punkcie korelacja "kroczków" ma postać $0\, dy = 0\, dx$, czyli nie narzuca żadnego warunku. Mamy w zasadzie całkowitą wolność wyboru, w którym kierunku ruszymy.[40]

Zadanie Ode.2 Inny przykład. Rozpatrzmy równanie różniczkowe:

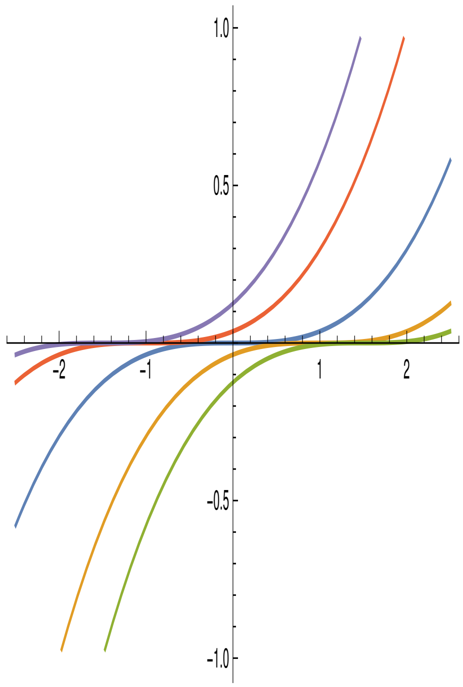
$$y' = y^{2/3}(x)\,.$$

Możemy je łatwo scałkować, tzn. znaleźć (jakieś) rozwiązania, dzieląc stronami przez $y^{2/3}$ i całkując stronami po dx:

$$\int dx\, \frac{y'(x)}{y^{2/3}(x)} \equiv \int dx\, \frac{dy}{dx}\, y^{-2/3}(x) = \int dx\,.$$

[39] W samej nazwie "krzywa" całkowa, a nie "funkcja" jest już przecież ta możliwość ukryta!

[40] Ale po ruszeniu w dowolnym kierunku daje się tu iść dalej tylko wtedy, gdy wybrany kierunek nie jest niezgodny z kierunkiem zadawanym przez $f(x, y)$ w punkcie ciut obok punktu $(0, 0)$ tak jest tu tylko, gdy ruszymy z $(0, 0)$ w kierunku o nachyleniu równym 1.



Rysunek 17: Rodzina krzywych całkowych równania $y' = y^{2/3}$.

Po lewej stronie dokonujemy zamiany zmiennej całkowania[41] z x na $y = y(x)$; czynnik $dx\, y'$ jest akurat tym, co potrzebne, czyli dy, i mamy po tym po lewej stronie całkę

$$\int dy(x)\, y^{-2/3}(x) = 3\, y^{1/3} + C_1\,.$$

Całka po prawej daje $x + C_2$ i scalając stałe $C \equiv C_1 - C_2$ mamy rodzinę rozwiązań (całek) równania różniczkowego:

$$y(x) = \left(\frac{x - C}{3}\right)^3.$$

Jak widać z rysunku 17 każda z tych krzywych całkowych przecina oś x-ów w punkcie $x = C$. Łatwo jednak zauważyć, że $y(x) \equiv 0$ jest też rozwiązaniem badanego równania[42] i to rozwiązaniem, które nie odpowiada żadnej stałej C w znalezionym powyżej rozwiązaniu, choć jest uczciwą funkcją $y = y(x)$. Znów winę za to, że przez każdy punkt na osi x przechodzą dwie (a nie jedna) krzywa całkowa, można zrzucić na niespełnienie założeń Lejkowego twierdzenia: pochodna $f_y(x, y) = (2/3)y^{-1/3}$ nie jest bowiem ciągła (ani nawet nie może mieć granicy) w punktach typu $(x, 0)$. Można to też widzieć tak, że gdy patrzymy

[41] Oczywiście zwykle bezrefleksyjnie równanie takie całkuje się “przenosząc dx pochodzące z dy/dx na prawą stronę”. Tu jednak chciałem pokazać, że można na to spojrzeć bardziej ortodoksyjnie.

[42] Znów łatwo wpaść na to pisząc równanie w postaci $dy = y^{2/3}\, dx$ - widać wtedy, że gdy jesteśmy na osi x, gdzie $y = 0$, korelacja kroczków jest taka, że $dy = 0$, a dx nie jest niczym ograniczone.

na punkty $(x, |\varepsilon|)$ i $(x, -|\varepsilon|)$, czyli tuż nad osią x i tuż pod nią, to w obu tych punktach nachylenie krzywej dąży do zera, gdy $\varepsilon \to 0$ i ma ten sam znak: krzywe otrzymane z całkowania "zlewają" się na moment z krzywą $y \equiv 0$; w tych punktach to właśnie krzywe $y = ((x - C)/3)^3$ są w jakiś sposób patologiczne, a regularnym rozwiązaniem jest $y \equiv 0$.

Rodzinę $y = y(x, C)$ rozwiązań równania różniczkowego (zwyczajnego pierwszego rzędu) parametryzowaną stałą C nazywa się *całką ogólną* tego równania. Przykład powyżej pokazał, że całka ogólna nie musi obejmować wszystkich rozwiązań (nawet wszystkich, które są uczciwymi funkcjami $y = y(x)$, a nie tylko krzywymi całkowymi).

Pewną wyróżnioną podklasę zwyczajnych równań różniczkowych pierwszego rzędu stanowią *równania o zmiennych rozdzielonych*, tj. postaci[43]

$$\frac{dy}{dx} = -\frac{p(x)}{q(y)}, \qquad \text{lub} \quad q(y)\,dy + p(x)\,dx = 0\,,$$

Całki ogólne takich równań są od razu dane "w kwadraturach" - ta nieco archaicznie brzmiąca nazwa (spotykana np. w nieśmiertelnym podręczniku do mechaniki Rubinowicza i Królikowskiego) oznacza po prostu, że rozwiązania takich równań są "od ręki" dane w postaci konkretnych całek (a czy całki te się da analitycznie wykonać, to już inna historia). Naogół jednak otrzymuje się w ten sposób związek $G(x, y) = 0$, czyli funkcję $y = y(x)$, albo $x = x(y)$ w postaci uwikłanej; ale jako sposób zadania krzywych całkowych jest to zupełnie wystarczające. Drugi nasz przykład był właśnie równaniem różniczkowym takiego rodzaju (a pierwszy - nie).

Zadanie Ode.3
Znaleźć wszystkie krzywe całkowe równania różniczkowego

$$y' = (2x - 1)\,y\,.$$

Rozwiązanie: Jest to właśnie równanie o zmiennych rozdzielonych. Możemy je przepisać w postaci dogodnej do scałkowania

$$\int \frac{dy}{y} = \int dx\,(2x - 1)\,.$$

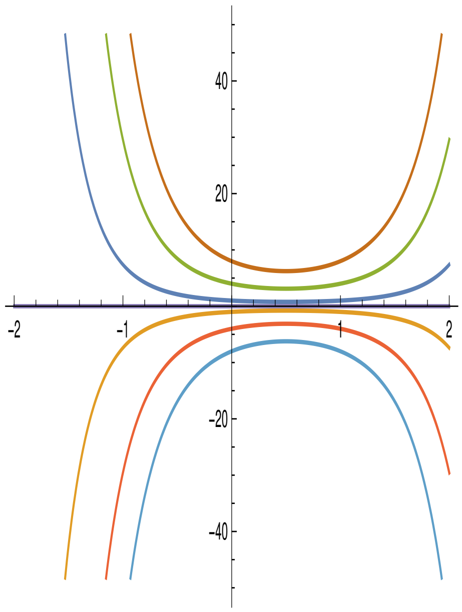
Wykonujemy całki (każdą jako całkę nieoznaczoną, ale dwie stałe całkowania możemy złączyć w jedną) i dostajemy

$$\ln|y| = x^2 - x + C\,,$$

czyli

$$y = \pm\exp\big(C + x^2 - x\big) \equiv \tilde{C}\,\exp(x^2 - x)\,.$$

[43] Minus jest dla porządku w interesie, żeby nazwy funkcji p i q były takie jak poprzednio; jest jasne, że można by to było zapisać też jako $dy/dx = Q(y)P(x)$ lub $Q(y)/\tilde{P}(x)$, czy jakoś podobnie.



Rysunek 18: Rodzina krzywych całkowych równania $y' = (2x-1)\,y$.

Zastąpiliśmy tu multiplikatywny czynnik $\pm e^C$ przez stałą $\tilde{C}$, dopuszczając oba jej znaki; oczywiście stała $\tilde{C} = 0$ odpowiada $C = -\infty$. Trochę więc to podejrzane (bo czy uczciwa stała całkowania może być nieskończona?), ale możemy zauważyć, że w ten sposób w całkę ogólną włączyliśmy krzywą całkową $y \equiv 0$, która też jest rozwiązaniem wyjściowego równania, ale nie była objęta całką ogólną ze stałą w eksponencie. Rodzina rozwiązań tego równania jest pokazana na rysunku 18. Możemy też zobaczyć, że przez każdy punkt płaszczyzny xy przechodzi teraz dokładnie jedna funkcja $y = y(x)$ będąca rozwiązaniem równania. Istotnie: jeśli chcemy, mieć funkcję przechodzącą przez punkt (x_0, y_0) dobieramy odpowiednio stałą $\tilde{C}$:

$$\tilde{C} = y_0 \exp(x_0 - x_0^2)\,.$$

Musi tak być, bo w równaniu zapisanym w postaci $dy/dx = f(x,y)$, funkcja $f(x,y)$ jest super przyzwoitą funkcją: ciągłą i nawet różniczkowalną ma całym $\mathbb{R}^2$ i to nieskończenie wiele razy.

Zadanie Ode.4 (przykład z Krysickiego-Włodarskiego)
Znaleźć wszystkie rozwiązania (krzywe całkowe) równania

$$y' = \frac{2\,x\,y^2}{1+x^2}\,.$$

Rozwiązanie: Równanie jest równaniem o zmiennych rozdzielonych i całkujemy je standardowo

$$\int \frac{dy}{y^2} = \int dx\,\frac{2x}{1+x^2}\,, \qquad \text{czyli} \qquad -\frac{1}{y} = C + \ln(1+x^2)\,.$$

Stąd

$$y = -\frac{1}{C + \ln(1 + x^2)} .$$

Znów $C = \infty$ odpowiada formalnie rozwiązaniu $y \equiv 0$. Tu $x \equiv 0$ nie jest rozwiązaniem równania $(1 + x^2)\, dy = 2xy^2 dy$, ale i tak przez każdy punkt (x_0, y_0) przechodzi dokładnie jedna krzywa całkowa równania i to zawsze będąca uczciwą funkcją $y = y(x)$, bo funkcja $f(x, y) = 2xy^2/(1 + x^2)$ jest na całym $\mathbb{R}^2$ regularna i jej pochodna $f_y(x, y)$ też taka jest. Aby rozwiązanie przechodziło przez punkt (x_0, y_0), stała C musi być wyznaczona z warunku

$$y_0 = -\frac{1}{C + \ln(1 + x_0^2)} ,$$

który zawsze ma rozwiązanie $C = -1/y_0 - \ln(1 + x_0^2)$ (trzeba oczywiście przyjąć, że $y_0 = 0$ odpowiada $C = \infty$, ale to jest właśnie to, co dopuściliśmy już). Zatem rozwiązanie przechodzące przez punkt (x_0, y_0) ma postać

$$y = \frac{1}{1/y_0 + \ln(1 + x_0^2) - \ln(1 + x^2)} .$$

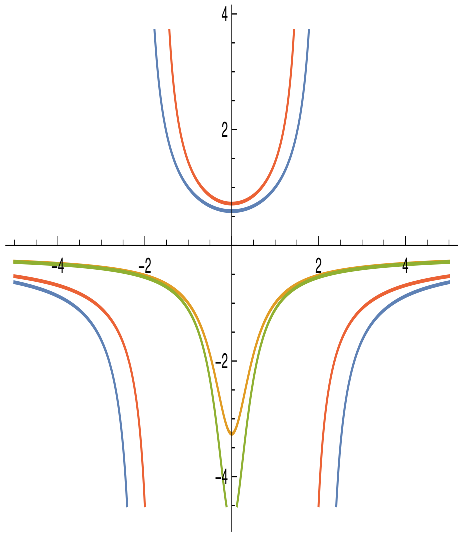
Charakter tych funkcji zależy jednak od punktu (x_0, y_0). Trzeba rozpatrzyć różne przypadki. Czynnik $\ln(1 + x^2)$ w mianowniku jest zawsze nieujemny, więc jeśli $1/y_0 + \ln(1 + x_0^2) > 0$, co zachodzi zawsze, gdy $y_0 > 0$, tj. gdy punkt, przez który rozwiązanie ma przechodzić leży w górnej półpłaszczyźnie ($y_0 = 0$ już mamy załatwione - odpowiada mu zawsze, niezależnie od x_0, rozwiązanie $y \equiv 0$), mianownik rozwiązania zawsze zeruje się w dwóch punktach, co oznacza, że rozwiązanie przechodzące przez dowolny punkt (x_0, y_0) o $y_0 > 0$ ma dwie pionowe asymptoty, tj. nigdy nie sięga w x-ach dalej w lewo niż pewne x_- i w prawo niż pewne x_+, gdzie

$$x_\mp = \mp\sqrt{e^{1/y_0}(1 + x_0^2) - 1} .$$

Sytuacja, gdy $1/y_0 + \ln(1 + x_0^2) > 0$ może też zachodzić, gdy $y_0 = -|y_0| < 0$, ale to zależy teraz od wartości $\ln(1 + x_0^2)$. Ogólnie, gdy $y_0 < 0$ są możliwe trzy przypadki (myślmy o ustalonym x_0 i analizujmy różne ujemne y_0): $-1/|y_0| + \ln(1 + x_0^2) > 0$, $-1/|y_0| + \ln(1 + x_0^2) < 0$ i taki przypadek krytyczny, gdy ten caly czynnik jest równy zeru. Krzywa odpowiadająca temu przypadkowi, dana wzorem

$$y = -\frac{1}{\ln(1 + x^2)} ,$$

leży (zależnie od tego, czy $x_0 > 0$, czy $x_0 < 0$ - przypadek $x_0 = 0$ nie może dać $-1/|y_0| + \ln(1 + x_0^2) = 0$) w prawej lub lewej dolnej ćwiartce, i biegnie od $x = \infty$ (od $x = -\infty$), gdzie dąży do zera od dołu, do $x = 0^\pm$, gdzie dąży do $-\infty$, czyli ma w $x = 0$ asymptotę pionową. Krzywa ta rozdziela (pamiętamy: x_0 sobie ustaliliśmy)



Rysunek 19: Rodzina krzywych całkowych równania $y' = 2xy^2/(1 + x^2)$. Trzeba tu uwzględnić to, że Mathematica maluje przy zadanym $x_0 = 1$ i y_0 także kawałki krzywych za asymptotami, co jest mylące. W szczególności pozorne jest widoczne na rysunku utożsamienie asymptot pionowych krzywych leżących w dolnej i górnej półpłaszczyźnie (położenie asymptot, czy w ogóle ich istnienie, zależy od wyboru punktu (x_0, y_0)). Dlatego trzeba się wczytać w przedstawioną w tekście dyskusję.

krzywe należące do dwóch pozostałych przypadków. Krzywe odpowadające pierwszemu z nich, zachodzącemu, gdy $-1/|y_0| + \ln(1 + x_0^2) > 0$, biegną jeśli $x_0 > 0$ ($x_0 < 0$) od $x = \infty$ ($x = -\infty$) przez (x_0, y_0) do pewnego x_+ (x_-) - danego wzorem wypisanym wyżej, w którym mają asymptotę pionową (nie dochodzą więc nigdy do $x = 0$. Z kolei krzywe odpowiadające przypadkowi, gdy $-1/|y_0| + \ln(1 + x_0^2) < 0$, biegną od $x = -\infty$ przez (x_0, y_0) do $x = +\infty$ mając minimum w $x = 0$. Można się w tym zorientować wnikliwie wpatrując się w rysunek 19.

Często na pozór beznadziejne równanie różniczkowe udaje się rozwiązać przy pomocy jakiegoś sprytnego chwytu. Czasem jest to przedefiniowanie szukanej funkcji, czasem zamiana zmiennych, a czasem jedno i drugie razem. Jeden prosty chwyt daje się zawsze zastosować, gdy funkcja $f(x, y)$ w równaniu $y' = f(x, y)$ jest funkcją jednorodną stopnia zerowego, tj. taką, że $f(\lambda x, \lambda y) = f(x, y)$. Funkcja taka musi być bowiem funkcją ilorazu y/x, tj. $f(x, y) = \tilde{f}(y/x)$. Podstawienie $y = x\,u(x)$ sprowadza wtedy równanie do postaci

$$u + x\,u' = \tilde{f}(u)\,, \qquad \text{czyli} \quad x\,u' = -u + \tilde{f}(u)\,,$$

t.j. do równania o zmiennych rozdzielonych. Poza tym, przy szukaniu sprytnych podstawień niema jednak (chyba - może matematycy mają jakieś tajne sposoby) żadnych ogólnych reguł postępowania i wszystko opiera się na sprycie boiskowym i orientacji w terenie (a w tym matematycy górują nad resztą ludzkości).

Zadanie Ode.5
Rozwiązać równanie różniczkowe

$$2\,y' + y^2 + \frac{1}{x^2} = 0\,.$$

Rozwiązanie: W podanej postaci nie jest to ani równanie o zmiennych rozdzielonych, ani też równanie liniowe (tymi zajmiemy się za niżej), bo występuje w nim y^2. Ale zamiast szukać funkcji $y = y(x)$ poszukajmy funkcji $u = u(x)$ wiążącej się z $y = y(x)$ przez

$$y = \frac{u(x)}{x}\,.$$

Podstawiamy to do wyjściowego równania i mamy

$$\frac{2u'}{x} - \frac{2u}{x^2} + \frac{u^2}{x^2} + \frac{1}{x^2} = 0\,.$$

Po pomnożeniu stronami przez x^2 przybiera to postać

$$2x\,u' - 2u + u^2 + 1 \equiv 2x\,u' + (u-1)^2 = 0\,,$$

a to już jest równanie o rozdzielonych zmiennych. Możemy je więc scałkować:

$$-2\int \frac{du}{(u-1)^2} = \int \frac{dx}{x}\,,$$

co daje

$$\frac{2}{u-1} = C + \ln|x| \equiv \ln|\tilde{C}\,x|\,,$$

czyli $u(x) = 1 + 2/\ln|\tilde{C}\,x|$. Możemy teraz napisać szukaną funkcję $y = y(x)$:

$$y(x) = \frac{1}{x} + \frac{2}{x\ln|\tilde{C}\,x|}\,.$$

Sprawdźmy (samokontroli nigdy dość!):

$$2y' = -\frac{2}{x^2} - \frac{4}{x^2\ln^2|\tilde{C}\,x|}\Big(1 + \ln|\tilde{C}\,x|\Big)\,,$$
$$y^2 = \frac{1}{x^2} + \frac{4}{x^2\ln|\tilde{C}\,x|} + \frac{4}{x^2\ln^2|\tilde{C}\,x|}\,.$$

Jak to dodamy i dodamy jeszcze $1/x^2$ to rzeczywiście wyjdzie zero. Oczywiście rozwiązaniem wyjściowego równania jest też $y(x) = 1/x$. Odpowiada ono dopuszczeniu w znalezionej całce ogólnej $|\tilde{C}| = \infty$.

Przy przejściu do funkcji u trzeba było wykluczyć punkty o $x = 0$ (w równaniu $y' = f(x, y) = -(y^2 + 1/x^2)/2$ funkcja $f(x, y)$ jest nieciągła w punktach $(0, y)$ więc nie mogą przez nie przechodzić funkcje $y = y(x)$, ale może jakaś funkcja $x = x(y)$ może?). Zobaczmy, co z nimi. Przepiszmy więc wyjściowe równanie w postaci

$$2x^2 dy = -(x^2 y^2 + 1)\, dx\,.$$

Teraz widać, że jest ono także spełniane przez prostą $x \equiv 0$.

Przechodzimy teraz do drugiego typu równań różniczkowych, które daje się rozwiązywać systematycznie. Są to równania (na razie wciąż pierwszego rzędu) liniowe z niejednorodnością, które ogólnie dają się zapisać w postaci[44]

$$y' + g(x)\, y = f(x)\,.$$

$f(x)$ i $g(x)$ są tu zadanymi z góry funkcjami. Niejednorodnością jest funkcja $f(x)$. Zauważmy, że gdy $f(x) \equiv 0$, jest to równanie o zmiennych rozdzielonych. Szukana funkcja $y = y(x)$ występuje tu w pierwszej potędze i na tym polega liniowość tego równania. Dzięki temu lewą jego stronę można zapisać w ogólnej postaci

$$D(x)\, y(x)\,,$$

gdzie $D(x)$ jest *operatorem różniczkowym.* Powinno nam się to zacząć kojarzyć z algebrą: fukcję y można traktować jak element przestrzeni wektorowej (i coż, że nieskończenie wymiarowej?) funkcji, a $D(x)$ można traktować jak odwzorowanie liniowe tej przestrzeni w nią samą. Ten sposób widzenia jest tu przydatny, bo zaraz zobaczymy, że rozwiązania takich równań mają taką samą strukturę, jak rozwiązania liniowych równań typu $F{\cdot}\mathbf{x} = \mathbf{b}$, którymi zajmowaliśmy się w ramach algebry.

Równania takie rozwiązuje się metodą *uzmiennienia stałej całkowania.* Zademonstrujemy ją tu na prostym przykładzie.[45]

Zadanie Ode.6
Znaleźć rozwiązanie (całkę ogólną) równania

$$y' = \frac{y}{x} + 3x\,.$$

Rozwiązanie: Najpierw rozwiązujemy równanie jednorodne, tj.

$$y' + g(x)\, y = 0\,.$$

[44] Zawsze można równanie tak napisać, żeby najwyższa pochodna szukanej funkcji - tu jest to pierwsze pochodna - nie była przez nic mnożona.

[45] Równanie różniczkowe użyte jako przykład na samym początku tego rozdziału też zostało rozwiązane metodą uzmiennienia stałej. Po przetrawieniu poniższego przykładu mogą Państwo wrócić do początku i samemu sobie tamto równanie rozwiązać.

czyli w danym przypadku równanie

$$y' = \frac{y}{x}\,.$$

Ponieważ, jak już zauważyliśmy, jest to równanie o zmiennych rozdzielonych, i to takie jak w zadaniu 38c, od razu wypisujemy jego rozwiązanie (całkę ogólną)

$$y_{\rm hom} = C\,x\,.$$

Dodaliśmy dopisek "hom" od homogeneous, czyli jednorodne, żeby pamiętać czego to jest rozwiązanie. Teraz następuje drugi krok od którego bierze się nazwa metody: zastępujemy stałą C nieznaną funkcją $h(x)$ i funkcję $y = x\,h(x)$ podstawiamy jako Ansatz[46] do wyjściowego równania:

$$\frac{d}{dx}\,(x\,h(x)) = h(x) + x\,h'(x) = \frac{1}{x}\,(x\,h(x)) + 3x\,.$$

Dzieje się wtedy to, co zawsze się dzieje w takich przypadkach (na tym polega cały wic tej metody!): wyrazy z $h(x)$ po obu stronach równania się nawzajem redukują i zostajemy z

$$x\,h'(x) = 3x\,, \qquad \text{czyli} \quad h'(x) = 3\,.$$

Dostajemy więc równanie na funkcję h, które ma zawsze postać równania o zmiennych rozdzielonych i to takiego, w którym pochodna h' nieznanej funkcji jest po prostu równa pewnej funkcji x. Równanie takie całkuje się zazwyczaj łatwo, a w każdym razie daje się rozwiązanie napisać jako kwadraturę. Tu jest to banalne: $h(x) = 3x + C'$ (zaraz zobaczymy, że pisanie stałej dowolnej nie jest tu konieczne: można wziąć jakiekolwiek rozwiązanie). I teraz następuje ostatni krok: dodajemy do siebie dwie części: $y_{\rm hom}$ i ten Ansatz, który nazywamy $y_{\rm inhom} = x\,h(x)$ (od inhomogeneous):

$$y = y_{\rm hom} + y_{\rm inhom} = C\,x + x\,(3x + C') = 3x^2 + \tilde{C}x\,.$$

Można sprawdzić podstawiając do wyjściowego równania, że jest to rzeczywiście rozwiązanie. Widać też, że rzeczywiście w $y_{\rm inhom}$ można było pominąć stałą C' bo i tak połączyła się ona ze stałą C w $y_{\rm hom}$ (zawsze się tak dzieje). Jest to kompletna całka ogólna wyjściowego równania. Stałą $\tilde{C}$ można dobrać tak, by spełnić warunek początkowy[47] $y(x_0) = y_0$.

[46] Tak to się nazywa w ogólnosłowiańskim języku niemieckim.

[47] Oczywiście stosują się tu wszystkie przewałkowane już reguły dotyczące tego, kiedy przez dany punkt (x_0, y_0) przechodzi zawsze dokładnie jedno rozwiązanie; czasem nie przechodzi żadne i wtedy nie można narzucić takiego warunku, czasem przechodzi więcej niż jedno i wtedy takie warunki początkowe są nie do przyjęcia, a czasem akurat tego rozwiązania, co przechodzi przez (x_0, y_0), całka ogólna otrzymana metodą uzmiennienia stałej może nie obejmować. W szczególności w rozpatrywanym przykładzie warunki początkowe z $x = 0$ muszą być trefne bo $f(x, y) = y/x + 3x$ jest nieciągła na osi y. Przepisując jednak (całe) wyjściowe równanie w postaci

$$x\,dy = (y + 3x^2)\,dx\,,$$

widzimy, że jego rozwiązaniem jest jeszcze prosta $x = 0$ i to jest właśnie krzywa całkowa przechodząca przez punkty $(0, y_0)$.

Uzyskana metodą uzmiennienia stałej całka ogólna ma dokładnie taką samą strukturę, jak najogólniejsze rozwiązanie algebraicznego problemu $F \cdot \mathbf{x} = \mathbf{b}$: jest ono sumą najogólniejszego (tzn. tu zależnego od stałej dowolnej) rozwiązania równania jednorodnego i *jakiegokolwiek* (dlatego stałą C' można było od początku pominąć) rozwiązania pełnego równania niejednorodnego. Przypomnijmy więc, że najogólniejsze rozwiązanie algebraicznego równania $F \cdot \mathbf{x} = \mathbf{b}$ miało postać $\mathbf{x} = C_1\mathbf{x}_1 + \ldots + C_r\mathbf{x}_r + \mathbf{x}_{\rm inhom} \equiv \mathbf{x}_{\rm hom} + \mathbf{x}_{\rm inhom}$, gdzie $\mathbf{x}_i$, $i = 1, \ldots r$ były wektorami takimi, że $F \cdot \mathbf{x}_i = \mathbf{0}$, a $\mathbf{x}_{\rm inhom}$ był jakimkowiek wektorem spełniajacym równanie $F \cdot \mathbf{x}_{\rm inhom} = \mathbf{b}$. Choć powinno to po algebrze być oczywiste, ale może warto jeszcze raz zobaczyć, dlaczego rozwiązanie równania liniowego z niejednorodnością ma taką strukturę. Wyobraźmy sobie, że mamy dwie różne funkcje (każda z nich może odpowiadać innym warunkom początkowym np.) $y = y_1(x)$ i $y = y_2(x)$, spełniające to samo liniowe równanie z niejednorodnością:

$$
\begin{aligned}
y_1' + g(x)\, y_1 &= f(x)\,, \\
y_2' + g(x)\, y_2 &= f(x)\,.
\end{aligned}
$$

Odejmijmy te równania jedno od drugiego stronami:

$$
\frac{d}{dx}\,(y_1' - y_2') + g(x)\,(y_1' - y_2') = 0\,.
$$

Różnica $y_1 - y_2$ spełnia więc równanie jednorodne. Mówiąc inaczej: każde dwa rozwiązania równania niejednorodnego różnią się od siebie o jakieś rozwiązanie równania jednorodnego. Więc jak mamy jakiekolwiek rozwiązanie równania niejednorodnego (a to właśnie znajdujemy uzmienniając stałą, ale można by było je zamiast tego znaleźć na śmietniku, czy ściągnąć od kolegi/koleżanki), to wystarczy doń dodać naogólniejsze (w sensie całki ogólnej) rozwiązanie równania jednorodnego, by dostać w ten sposób najogólniejsze rozwiązanie równania niejednorodnego, czyli *skonstruować jego całkę ogólną*. Można też napisać formalny wzór na całkę ogólną równania niejednorodnego $y' + g(x)y = f(x)$. Ma on postać

$$
\begin{aligned}
y &= e^{-G(x)} \left(C + \int^x dx'\, f(x')\, e^{G(x')} \right), \\
G(x) &= \int dx\, g(x)\,.
\end{aligned}
$$

Zamiast go pamiętać (ale warto go sobie wyprowadzić jako sprawdzian, czy się wszystko zrozumiało!), lepiej jest w każdym konkretnym przypadku postępować według podanego wyżej schematu.

Powiedzmy też od razu tutaj, żeby potem już nie powracać do tego, że ten sam schemat pozostaje słuszny w przypadku liniowych równań różniczkowych rzędu r z niejednorodnością, tj. postaci

$$
\frac{d^r y}{dx^r} + g_{r-1}(x)\,\frac{d^{r-1}y}{dx^{r-1}} + \ldots + g_1(x)\,\frac{dy}{dy} + g_0(x)\,y = f(x)\,.
$$

Tak samo jak w przypadku równań liniowych pierwszego rzędu z niejednorodnością, również i tu, każde dwa rozwiązania powyższego równania niejednorodnego różnią się o jakieś rozwiązanie równania jednorodnego z funkcją f zastąpioną przez zero. I dlatego tak samo wystarczy znaleźć (kupić na czarnym rynku albo pchlim targu) jakiekolwiek jedno (choćby już mocno używane) rozwiązanie równania niejednorodnego i dodać do niego najogólniejsze rozwiązanie równania jednorodnego (całka ogólna tego rówanania w tym przypadku zależy od r stałych dowolnych: $C_1, \dots, C_r$), by skonstruować kompletną całkę ogólną liniowego równania niejednorodnego rzędu r z niejednorodnością, która wobec tego ma postać

$$y = y_{\rm inhom}(x) + y_{\rm hom}(x, C_1, \dots, C_r)\,.$$

Zobaczymy to dalej na przykładzie równania drugiego rzędu, ale takiego prostego, w którym funkcje $g_1(x)$ i $g_0(x)$ są stałymi. W innych przypadkach, gdy funkcje $g_1(x)$ i $g_0(x)$ są nietrywialnymi funkcjami metoda pozostaje w mocy, ale naogół trudno jest znaleźć całkę ogólną równania jednorodnego. (Gdy wszystkie funkcje $g_{r-1}, \dots, g_0$ są stałymi są proste sposoby - też je poznamy właśnie na przykładzie takiego równania o $r = 2$ - by ogólną całkę równania jednorodnego skonstruować). Ostatnia uwaga jest taka, że jeżeli niejednorodność równania linowego ma postać sumy dwóch funkcji, $f(x) = f_1(x) + f_2(x)$, to rozwiązanie $y_{\rm inhom}(x)$ równania niejednorodnego też ma postać sumy $y^{(1)}_{\rm inhom}(x) + y^{(2)}_{\rm inhom}(x)$, gdzie $y^{(1)}_{\rm inhom}(x)$ jest rozwiązaniem równania liniowego z niejednorodnością $f_1(x)$, a $y^{(2)}_{\rm inhom}(x)$ jest rozwiązaniem tegoż równania z niejednorodnością $f_2(x)$.

Zadanie Ode.7
Znaleźć rozwiązanie równania

$$y' \cos x + 2y \sin x = 2 \sin x\,,$$

przechodzące przez punkt (x_0, y_0).
Rozwiązanie: Jest to właśnie liniowe równanie pierwszego rzędu z niejednorodnością. Postępujemy więc standardowo. Najpierw znajdujemy całkę ogólną równania jednorodnego

$$y' + 2y\,{\rm tg}x = 0\,.$$

Rozdzielamy zmienne i całkujemy

$$\int \frac{dy}{y} = -2 \int dx\, \frac{\sin x}{\cos x}\,, \qquad \text{czyli} \quad \ln|y| = 2\ln|\cos x| + C = \ln|\tilde{C}\cos^2 x|\,.$$

Stąd (zdejmując moduł upychamy znak $\pm$ w stałą $\tilde{C}$)

$$y_{\rm hom} = \tilde{C}\cos^2 x\,.$$

Teraz szukamy $y_{\rm inhom}$ metodą uzmiennienia stałej, tj. podstawiamy do wyjściowego równania Ansatz $y = h(x)\cos^2 x$. Daje to

$$\big(h' \cos^2 x - 2h\cos x \sin x\big)\cos x + 2h(x)\cos^2 x \sin x = 2\sin x\,.$$

Wyrazy z gołym h się redukują i dostajemy na h równanie

$$h' = \frac{2\sin x}{\cos^3 x}\,,$$

które daje $h = 1/\cos^2 x$. Zatem $y_{\rm inhom} = h(x)\cos^2 x = 1$ po prostu (co można było zauważyć!) i pełne rozwiązanie wyjściowego równania (jego całka ogólna) ma postać

$$y = y_{\rm inhom} + y_{\rm hom} = 1 + \tilde{C}\cos^2 x\,.$$

Możemy teraz dobrać $\tilde{C}$ tak, by rozwiązanie spełniało warunek początkowy, tj. by krzywa przechodziła przez punkt (x_0, y_0):

$$y = 1 + \frac{y_0 - 1}{\cos^2 x_0}\cos^2 x\,.$$

Widać jednak, że jest to możliwe tylko wtedy, gdy $x_0 \neq (\pi/2) + k\pi$. Patrząc na całkę ogólną $y = 1 + \tilde{C}\cos^2 x$ widzimy też, że wszystkie obejmowane przez nią funkcje (czyli nieskończenie wiele rozwiązań) przechodzą przez punkty $(\frac{\pi}{2} + k\pi, 1)$ i wszystkie one mają w tych punktach zerowe nachylenie. Można to zrozumieć pisząc wyjściowe, niejednorodne równanie w kanonicznej postaci[48] $y' = 2(1-y){\rm tg}\,x$. Widać z niej, że funkcja występująca po prawej stronie jest, jako funkcja na $\mathbb{R}^2$, nieciągła w punktach $x = \frac{\pi}{2} + k\pi$ i $y \neq 1$, ale w punktach $(\frac{\pi}{2} + k\pi, 1)$ można ją zawsze dookreślić tak, by była ciągła wzdłuż jednego, dowolnie wybranego kierunku; niemniej tylko gdy uciąglamy ją wzdłuż osi x, to wartość uciąglająca - czyli zero[49] - pasuje do nachylenia "kierunku uciąglania". I dlatego wszystkie rozwiązania $y = y(x)$ przechodzące przez punkty $(\frac{\pi}{2}+k\pi, 1)$ muszą mieć zerowe nachylenie. Z kolei przez punkty $(\frac{\pi}{2} + k\pi, y)$ o $y \neq 1$ nie przechodzi żadna funkcja $y = y(x)$, ale jeśli wyjściowe równanie przepisać w formie $\cos x\,dy = 2(1-y)\sin x\,dx$, to stanie się jasne, że jego rozwiązaniami są też proste $x = \frac{\pi}{2} + k\pi$, które są przyzwoitymi funkcjami $x = x(y)$.

Zajmiemy się teraz równaniami różniczkowymi rzędu $r > 1$ mającymi wyznaczać jedną funkcję (jednej zmiennej) i układami r równań różniczkowych pierwszego rzędu na r funkcji. Oczywiście tylko pewnymi szczególnie prostymi klasami problemów należących do tych dwóch typów. Rozpatrzymy też może jeden układ równań drugiego rzędu na kilka funkcji (jako rozszerzenie rozważań na przypadek układu p równań r-tego rzędu na p szukanych funkcji).

Ogólnie równanie różniczkowe rzędu r na jedną funkcję $y = y(x)$ ma jedną z dwu postaci

$$y^{(r)} = f(x, y, y', \ldots, y^{(r-1)})\,, \qquad \text{lub} \quad F(x, y, y', \ldots, y^{(r)}) = 0\,.$$

---

[48] Oczywiście widać, że jest to równanie o zmiennych rozdzielonych i można je było rozwiązać także bez używania metody uzmienniania stałej. Oczywiście wynik będzie ten sam - zalecam sprawdzenie!

[49] W pobliżu punktu $x = x_0$, gdzie $x_0 = \frac{\pi}{2} + k\pi$, funkcja tangens zachowuje się jak $1/(x - x_0)$, wiec gdy badamy $f(x, y)$ na ciągach postaci $x = x_0 + a/n$, $y = 1 + b/n$, czyli zbiegamy do punktu $(x_0, 1)$ wzdłuż prostych o tangensie kąta nachylenia równym b/a, wartość uciąglająca $f(x, y)$ w $(x_0, 1)$ jest równa $-2b/a$.

Całką takiego równania nazywamy każdą funkcję $y = y(x)$, spełniającą te równania tożsamościowo na jakimś otwartym podzbiorze $\mathbb{R}$. Całką ogólną zaś - rodzinę całek sparametryzowaną r dowolnymi stałymi $C_1, \dots, C_r$, czyli funkcję $y = y(x, C_1, \dots, C_r)$. Zależność od r stałych dowolnych pozwala (naogół) narzucić na całkę równania r-tego rzędu r lokalnych warunków początkowych

$$y(x_0) = y_0, \quad y'(x_0) = y'_0, \quad \dots \quad y^{(r-1)}(x_0) = y_0^{(r-1)} .$$

Przez "lokalne warunki" rozumiemy tu to, że wszystkie one są zadane w jednym i tym samym punkcie x_0. Przy odpowiednich warunkach spełnianych przez fukcję $f(x, y_1, \dots, y_{r-1})$ w punkcie $(x_0, y'_0, \dots, y_0^{(r-1)})$ (lub funkcję $F(x, y_1, \dots, y_r)$ w punkcie $(x_0, y'_0, \dots, y_0^{(r)})$ przy sformułowaniu uwikłanym) i w jego jakimś otwartym otoczeniu istnienie spełniającej takie warunki całki równania jest zagwarantowane przez odpowiednie twierdzenia.[50] Najprostszym przypadkiem takich równań są równania liniowe z niejednorodnością, o których już mówiliśmy wcześniej. Tu zajmiemy się takimi równaniami o stałych współczynnikach, tzn. postaci

$$\frac{d^r y}{dx^r} + a_{r-1} \frac{d^{r-1} y}{dx^{r-1}} + \dots + a_1 \frac{dy}{dx} + a_0 \, y = f(x) ,$$

i drugiego rzędu, tj. o $r = 2$, bo w fizyce najczęściej występują równania drugiego rzędu.

Równanie różniczkowe rzedu r na jedną funkcję łatwo przekształcić w r równań pierwszego rzędu na r funkcji, podstawiając po prostu $y = y_1$, $y' = y_2$, $\dots$, $y^{(r-1)} = y_r$. Otrzymujemy wtedy układ równań pierwszego rzędu

$$\begin{aligned} y'_1 &= y_2 , \\ y'_2 &= y_3 , \\ &\dots\dots\dots\dots\dots\dots \\ y'_r &= f(x, y_1, y_2, \dots, y_r) , \end{aligned}$$

który jest szczególnym przypadkiem ogólnego układu r równań pierwszego rzędu

$$\begin{aligned} y'_1 &= f_1(x, y_1, y_2, \dots, y_r) , \\ y'_2 &= f_2(x, y_1, y_2, \dots, y_r) , \\ &\dots\dots\dots\dots\dots\dots\dots \\ y'_r &= f_r(x, y_1, y_2, \dots, y_r) , \end{aligned}$$

na r funkcji. Układy takie, nawet ten szczególny, wypisany wyżej, są dość skomplikowane i dlatego zajmiemy się tylko przypadkiem, w którym prawe strony od nieznanych funkcji

[50] Nie zawsze takie lokalne warunki są naturalne: np. w zagadnieniach wariacyjnych (typu: znaleźć kształt zjeżdżalni po której masa m najszybciej zjedzie z wysokości h w polu grawitacyjnym - to jest klasyczy problem brachistochrony) które prowadzą do równań drugiego rzędy naturalne są warunki $y(x_1) = y_1$, $y(x_2) = y_2$, czyli nielokalne. Przy takich warunkach jednak twierdzenia matematyczne o istnieniu rozwiązania są często bezradne.

$y_1, \ldots, y_r$ zależą liniowo. Można je wtedy, gdy chodzi o jednorodne równania zapisać macierzowo

$$\frac{d}{dx}\begin{pmatrix} y_1 \\ y_2 \\ \cdot \\ y_r \end{pmatrix} = \begin{pmatrix} f_{11}(x) & f_{12}(x) & \ldots & f_{1r}(x) \\ f_{21}(x) & f_{22}(x) & \ldots & f_{2r}(x) \\ \cdot & \cdot & \ldots & \cdot \\ f_{r1}(x) & f_{r2}(x) & \ldots & f_{rr}(x) \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ \cdot \\ y_r \end{pmatrix}.$$

Albo krótko w postaci $d\mathbf{y}/dx = F(x)\cdot\mathbf{y}$. Formalne rozwiązanie takiego układu ma postać

$$\mathbf{y}(x) = P_x \exp\left\{\int_{x_0}^{x} dx'\, F(x)\right\}\cdot\mathbf{y}_0\,,$$

w której $\mathbf{y}_0 = \mathbf{y}(x_0)$ jest wektorem warunków początkowych, a P_x oznacza operację uporządkowania iloczynu macierzy $F(x_1)\cdot\ldots\cdot F(x_n)$ występującego w n-tym wyrazie

$$\frac{1}{n!}\int_{x_0}^{x} dx'_1 \ldots \int_{x_0}^{x} dx'_n\, F(x_1)\cdot\ldots\cdot F(x_n)\,,$$

rozwinięcia funkcji exponens[51] w kolejności (idąc od prawej strony) od najmniejszej wartości zmiennej x'_i do największej. W ogólności macierzy działającej na wektor $\mathbf{y}_0$ nie daje się jednak jawnie wypisać i dlatego rozwiązanie to nazwaliśmy formalnym. Problem upraszacza się, gdy macierz F w równaniu $d\mathbf{y}/dx = F\cdot\mathbf{y}$ nie zależy od zmiennej x. Rozwiązanie jest wtedy dane prostym wzorem

$$\mathbf{y} = \exp\{(x - x_0)\,F\}\cdot\mathbf{y}_0\,.$$

Zauważmy też, że uogólniając macierz F działającą na wektory pewnej r-wymiarowej przestrzeni wektorowej (tu nad ciałem $\mathbb{R}$, ale mogła by być i nad $\mathbb{C}$) do operatora liniowego $\hat{H}$ działającego w pewnej nieskończonej przestrzeni wektorowej nad $\mathbb{C}$ ("działającego", tzn. odwzorowującego tę przestrzeń w nią samą), z iloczynem skalarnym $(\cdot|\cdot)_S$ i zupełnej (tj. takiej w której wszystkie Cauchy-ciągi mają swoje granice) i zmieniając nazwę zmiennej niezależnej z x na t, otrzymujemy typowy problem ewolucji czasowej wektora stanu w mechnice kwantowej. Operator $\hat{H}$ jest w takim przypadku operatorem energii, zwanym Hamiltonianem, a rozwiązywane równanie - równaniem Schrödingera

$$i\hbar\,\frac{d}{dt}\,|\psi(t)\rangle = \hat{H}|\psi(t)\rangle\,,$$

Takimi równaniami zajmiemy się teraz (po to na algebrze uczyliśmy się znajdywać eksponensy macierzy). Naturalnym uogólnieniem są układy liniowe równań z niejednorodnością postaci

$$\frac{d\mathbf{y}}{dx} = F\cdot\mathbf{y} + \mathbf{b}(x)\,.$$

[51] Macierz $F(x_1)$, …, $F(x_n)$ są, gdy x-y w nich są różne, różnymi macierzami i nie są przemienne: $F(x_1)\cdot F(x_2)$ nie jest tą samą macierzą, co $F(x_2)\cdot F(x_1)$; operacja P_x definiuje więc, w jakim porządku mają one być mnożone w różnych podobszarach całego obszaru $[x_0, x]\times\ldots\times[x_0, x]\subset\mathbb{R}^n$ całkowania po $dx_1\ldots dx_n$.

w których w ogólności macierz F mogłaby zależeć od zmiennej x, ale tu będziemy rozpatrywać tylko takie, w których macierz F jest stała (niezależna od x). Równania takie można rozwiązywać omówioną już wyżej metodą uzmienienia stałej: uzmiennioną stałą jest w tym przypadku wektor $\mathbf{h}(x)$ zastępujący wektor warunku początkowego $\mathbf{y}_0$ w rozwiązaniu (całce ogólnej) równania jednorodnego (tzn. spełniającego równanie z wektorem $\mathbf{b} \equiv \mathbf{0}$). Do takiego równania z macierzą F i wektorze $\mathbf{b}(x)$ szczególnych postaci

$$F = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \cdot & \cdot & \cdot & \dots & \cdot \\ 0 & 0 & 0 & \dots & 1 \\ -a_0 & -a_1 & -a_2 & \dots & -a_{r-1} \end{pmatrix}, \qquad \mathbf{b}(x) = \begin{pmatrix} 0 \\ 0 \\ \cdot \\ 0 \\ f(x) \end{pmatrix},$$

sprowadza się wypisane wyżej równanie liniowe rzędu r o stałych współczynnikach:

$$\frac{d^r y}{dx^r} + a_{r-1}\,\frac{d^{r-1}y}{dx^{r-1}} + \ldots + a_1\,\frac{dy}{dx} + a_0\, y = f(x)\,.$$

Zadanie Ode.8

Znaleźć ogólne rozwiązanie jednorodnego liniowego układu trzech równań różniczkowych pierwszego rzędu[52]

$$\frac{d}{dt}\begin{pmatrix} y_1(t) \\ y_2(t) \\ y_3(t) \end{pmatrix} = \begin{pmatrix} 2 & -1 & 1 \\ 0 & 1 & 0 \\ 1 & -1 & 2 \end{pmatrix}\begin{pmatrix} y_1(t) \\ y_2(t) \\ y_3(t) \end{pmatrix}.$$

Podać także rozwiązanie spełniające warunki początkowe: $y_1(t=t_0) = a$, $y_2(t=t_0) = b$, $y_3(t=t_0) = c$.

Rozwiązanie: Oznaczmy przez F występującą w tym równaniu różniczkowym macierz. Rozwiązaniem takiego równania, tj. równania postaci $\dot{\mathbf{y}}(t) = F\cdot\mathbf{y}(t)$, jest zawsze

$$\mathbf{y}(t) = e^{(t-t_0)F}\cdot\mathbf{y}(t_0)\,,$$

tj. macierz $e^{(t-t_0)F}$ działająca na wektor warunków początkowych. Nie musimy zatem znajdywać całej macierzy $\exp(tF)$; wystarczy znaleźć jej działanie na podany wektor warunków początkowych. Szukamy zatem najpierw wartości własnych macierzy F:

$$W_F(\lambda) = \det\begin{pmatrix} 2-\lambda & -1 & 1 \\ 0 & 1-\lambda & 0 \\ 1 & -1 & 2-\lambda \end{pmatrix} = -(\lambda-1)(\lambda-2)^2 + (\lambda-1) = -(\lambda-1)^2(\lambda-3)\,.$$

Odpowiadającymi im wektorami własnymi są:

$$\lambda = 3: \quad \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \qquad \lambda = 1: \quad \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \qquad \text{oraz} \quad \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}.$$

[52] Jako że w problemach fizycznych zmienną niezależną w tego typu równaniach jest czas, zmienimy tu x na t.

Mimo że druga z wartości własnych ma krotność dwa, to istnieją dwa odpowiadające jej liniowo niezależne wektory własne. Rozwiązanie będzie miało zatem charakter czysto eksponencjalny. Rozkładamy dowolny wektor warunków początkowych zadanych w $t = t_0$ na znalezione wektory własne:

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \frac{a-b+c}{2} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + \frac{a+b-c}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \frac{-a+b+c}{2} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}.$$

Stąd

$$\begin{pmatrix} y_1(t) \\ y_2(t) \\ y_3(t) \end{pmatrix} = e^{(t-t_0)F} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \frac{a-b+c}{2} e^{3(t-t_0)} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$
$$+ \frac{a+b-c}{2} e^{(t-t_0)} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \frac{-a+b+c}{2} e^{(t-t_0)} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix},$$

bo działanie $e^{(t-t_0)F}$ na każdy z wektorów własnych sprowadza się do pomnożenia go przez czynnik $e^{(t-t_0)\lambda}$, gdzie λ jest odpowiadającą temu wektorowi wartością własną.

Oczywiście skoro znamy działanie $e^{(t-t_0)F}$ na dowolny wektor (a, b, c), możemy bez większych trudności znaleźć i samą macierz $e^{(t-t_0)F}$. W tym celu przepisujemy (kładąc $t_0 = 0$ dla uproszczenia)

$$\begin{aligned} e^{tF} \begin{pmatrix} a \\ b \\ c \end{pmatrix} &= \frac{1}{2} \begin{pmatrix} (a-b+c)\, e^{3t} + (a+b-c)\, e^{t} \\ (a+b-c)\, e^{t} + (-a+b+c)\, e^{t} \\ (a-b+c)\, e^{3t} + (-a+b+c)\, e^{t} \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} e^{3t} + e^{t} & -e^{3t} + e^{t} & e^{3t} - e^{t} \\ 0 & 2\, e^{t} & 0 \\ e^{3t} - e^{t} & -e^{3t} + e^{t} & e^{3t} + e^{t} \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix}. \end{aligned}$$

Stojąca w ostatniej linii macierz jest właśnie macierzą e^{tF}.

Zauważmy też, że $y_2(t) = e^{t-t_0} b \equiv e^{t-t_0} y_2(0)$, bo równanie różniczkowe wyznaczające $y_2(t)$ było w istocie niezależne od $y_1(t)$ i $y_3(t)$. Można więc było najpierw rozwiązać niezależne równanie $\dot{y}_2 = y_2$, i otrzymane jego rozwiązanie wstawić jako jawną już funkcję do układu dwóch równań na y_1 i y_3, który w ten sposób przybrałby postać

$$\frac{d}{dt} \begin{pmatrix} y_1(t) \\ y_3(t) \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} y_1(t) \\ y_3(t) \end{pmatrix} - \begin{pmatrix} b\, e^{t-t_0} \\ b\, e^{t-t_0} \end{pmatrix},$$

czyli układu równań liniowych z niejednorodnością, który ogólniej, oznaczając $\tilde{F}$ występującą w nim macierz, zapiszmy w postaci

$$\frac{d}{dt} \mathbf{y} = \tilde{F} \cdot \mathbf{y} + \mathbf{b}\,.$$

Pozwala nam to zademonstrować na tym przykładzie metodę uzmienniania stałej w wersji wektorowej. Macierz $\tilde{F}$, ma tu wartości własne 3 i 1, którym odpowiadają oczywiste wektory własne. Rozwiązujemy najpierw liniowe równanie jednorodne z jakimś dowolnym warunkiem początkowym y_1^0, y_3^0:

$$\begin{pmatrix} y_1(t) \\ y_3(t) \end{pmatrix}_{\text{hom}} = e^{(t-t_0)\tilde{F}} \begin{pmatrix} y_1^0 \\ y_3^0 \end{pmatrix} = \frac{y_1^0 + y_3^0}{2}\, e^{3(t-t_0)} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \frac{y_1^0 - y_3^0}{2}\, e^{(t-t_0)} \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Teraz jednak już potrzebna będzie jawna postać macierzy $\exp\{(t-t_0)\tilde{F}\}$, więc ją szybko znajdujemy:

$$e^{(t-t_0)\tilde{F}} = \frac{1}{2} \begin{pmatrix} e^{3(t-t_0)} + e^{t-t_0} & e^{3(t-t_0)} - e^{t-t_0} \\ e^{3(t-t_0)} - e^{t-t_0} & e^{3(t-t_0)} + e^{t-t_0} \end{pmatrix}.$$

Następnie, do równania niejednorodnego podstawiamy Ansatz

$$\mathbf{y}_{\text{inhom}} = e^{(t-t_0)\tilde{F}} \cdot \mathbf{h}(t)\,.$$

Ponieważ

$$\frac{d}{dt}\left(e^{(t-t_0)\tilde{F}} \cdot \mathbf{h}(t)\right) = \tilde{F} \cdot e^{(t-t_0)\tilde{F}} \cdot \mathbf{h}(t) + e^{(t-t_0)\tilde{F}} \cdot \frac{d}{dt}\mathbf{h}(t)\,,$$

człony $F \cdot e^{(t-t_0)\tilde{F}} \cdot \mathbf{h}(t)$ się redukują i otrzymujemy

$$e^{(t-t_0)\tilde{F}} \cdot \frac{d}{dt}\mathbf{h}(t) = \mathbf{b}(t)\,, \qquad \text{czyli} \quad \mathbf{h}(t) = \int dt\, e^{-(t-t_0)\tilde{F}} \cdot \mathbf{b}(t)\,.$$

Jawnie, w rozpatrywanym tu przypadku,

$$\begin{aligned} \mathbf{h}(t) &= -\frac{1}{2}\int dt \begin{pmatrix} e^{-3(t-t_0)} + e^{-(t-t_0)} & e^{-3(t-t_0)} - e^{-(t-t_0)} \\ e^{-3(t-t_0)} - e^{-(t-t_0)} & e^{-3(t-t_0)} + e^{-(t-t_0)} \end{pmatrix} \begin{pmatrix} b\, e^{t-t_0} \\ b\, e^{t-t_0} \end{pmatrix} \\ &= -\int dt \begin{pmatrix} b\, e^{-2(t-t_0)} \\ b\, e^{-2(t-t_0)} \end{pmatrix} = \frac{1}{2} \begin{pmatrix} b\, e^{-2(t-t_0)} \\ b\, e^{-2(t-t_0)} \end{pmatrix}. \end{aligned}$$

(Całka z wektora to po prostu całki z jego "pięterek"). Zatem

$$\begin{aligned} \mathbf{y}_{\text{inhom}} = e^{(t-t_0)\tilde{F}} \cdot \mathbf{h} = \frac{1}{4} &\begin{pmatrix} e^{3(t-t_0)} + e^{t-t_0} & e^{3(t-t_0)} - e^{t-t_0} \\ e^{3(t-t_0)} - e^{t-t_0} & e^{-3(t-t_0)} + e^{t-t_0} \end{pmatrix} \begin{pmatrix} b\, e^{-2(t-t_0)} \\ b\, e^{-2(t-t_0)} \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} b\, e^{t-t_0} \\ b\, e^{t-t_0} \end{pmatrix}. \end{aligned}$$

Składamy teraz pełne rozwiązanie: $\mathbf{y} = \mathbf{y}_{\text{hom}}(t, y_1^0, y_3^0) + \mathbf{y}_{\text{inhom}}(t)$:

$$\mathbf{y}(t) = \frac{1}{2} \begin{pmatrix} e^{3(t-t_0)} + e^{t-t_0} & e^{3(t-t_0)} - e^{t-t_0} \\ e^{3(t-t_0)} - e^{t-t_0} & e^{-3(t-t_0)} + e^{t-t_0} \end{pmatrix} \begin{pmatrix} y_1^0 \\ y_3^0 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} b\, e^{t-t_0} \\ b\, e^{t-t_0} \end{pmatrix}.$$

Teraz dopiero, mając całość, możemy dobrać stałe y_1^0 i y_3^0 tak, by spełnić warunki $y_1(t_0) = a$ i $y_3(t_0) = c$. Po położeniu $t = t_0$ macierz działająca na y_1^0 i y_3^0 staje się macierzą jednostkową (musi tak być) i odczytujemy, że $y_1^0 = a - \frac{1}{2}b$ i $y_3^0 = c - \frac{1}{2}b$. No i teraz możemy odczytać, że

$$
\begin{aligned}
y_1(t) &= \frac{1}{2}\left(e^{3(t-t_0)} + e^{t-t_0}\right)a - \frac{1}{2}\left(e^{3(t-t_0)} - e^{t-t_0}\right)b + \frac{1}{2}\left(e^{3(t-t_0)} - e^{t-t_0}\right)c\,,\\
y_3(t) &= \frac{1}{2}\left(e^{3(t-t_0)} - e^{t-t_0}\right)a - \frac{1}{2}\left(e^{3(t-t_0)} - e^{t-t_0}\right)b + \frac{1}{2}\left(e^{3(t-t_0)} + e^{t-t_0}\right)c\,.
\end{aligned}
$$

Jest to oczywiście to samo, co jako $y_1(t)$ i $y_3(t)$ otrzymaliśmy pierwszym sposobem.

Zadanie Ode.9

Znaleźć rozwiązanie układu równań różniczkowych

$$
\begin{aligned}
\dot{y}_1 &= y_1 - 3\,y_2 + 4\,y_3\\
\dot{y}_2 &= 4\,y_1 - 7\,y_2 + 8\,y_3\\
\dot{y}_3 &= 6\,y_1 - 7\,y_2 + 7\,y_3\,.
\end{aligned}
$$

Rozwiązanie: Podany układ równań można przepisać w postaci macierzowej

$$\dot{\mathbf{y}}(t) = F\cdot\mathbf{y}(t)\,,$$

z macierzą

$$F = \begin{pmatrix} 1 & -3 & 4\\ 4 & -7 & 8\\ 6 & -7 & 7 \end{pmatrix}.$$

Rozwiązanie ma oczywiście postać

$$\mathbf{y}(t) = e^{(t-t_0)F}\cdot\mathbf{y}(t_0)\,,$$

Trzeba zatem znaleźć macierz e^{tF} lub jej działanie na dowolny wektor (a, b, c) warunków początkowych $(y_1(t_0), y_2(t_0), y_3(t_0))$ zadanych w $t = t_0$. Macierz F była już przedmiotem Zadania 78 w notatkach do algebry. Macierz e^{tF} została tam znaleziona dwoma sposobami. Możemy więc napisać "od ręki" (dla uproszczenia kładziemy $t_0 = 0$):

$$
\begin{pmatrix} y_1(t)\\ y_2(t)\\ y_3(t) \end{pmatrix} = \begin{pmatrix} e^{3t} - 2t\,e^{-t} & -e^{3t} + e^{-t} + t\,e^{-t} & e^{3t} - e^{-t}\\ 2e^{3t} - 2e^{-t} - 4t\,e^{-t} & -2e^{3t} + 3e^{-t} + 2t\,e^{-t} & 2e^{3t} - 2e^{-t}\\ 2e^{3t} - 2e^{-t} - 2t\,e^{-t} & -2e^{3t} + 2e^{-t} + t\,e^{-t} & 2e^{3t} - e^{-t} \end{pmatrix} \begin{pmatrix} a\\ b\\ c \end{pmatrix}.
$$

Przypomnijmy jeszcze, że czynniki liniowe w t (w $t - t_0$ jeśli $t_0 \neq 0$) pochodzą z tego, że macierz F ma jedną dwukrotną wartość własną, której to wartości własnej odpowiada tylko jeden wektor własny (a nie dwa) i zachodzi konieczność rozkładania wektora warunków początkowych na dwa wektory własne i jeden wektor pierwiastkowy (lub stosowania sztuczki z różniczkowaniem w twierdzeniu C-H).

Zadanie Ode.10
Rozwiązać liniowe równanie różniczkowe z niejednorodnością

$$\frac{d}{dt}\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} -1 & -1 & 1 \\ 1 & 0 & 0 \\ -1 & -1 & 1 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} + \begin{pmatrix} t \\ 2 \\ 1+t \end{pmatrix}.$$

Rozwiązanie: Jest to równanie pierwszego rzędu liniowe (względem szukanego wektora $\mathbf{y}(t)$) z niejednorodnością. Ma ono postać

$$\dot{\mathbf{y}}(t) = F\cdot\mathbf{y}(t) + \mathbf{f}(t)\,.$$

Zgodnie z ogólną metodą rozwiązywania takich równań (już przećwiczoną w zadaniu Ode.8), szukamy najpierw rozwiązania równania jednorodnego $\dot{\mathbf{y}}(t) = F\cdot\mathbf{y}(t)$. Znajdujemy w tym celu pierwiastki wielomianu charakterystycznego

$$W_F(\lambda) = -\lambda^3 - \lambda$$

macierzy F stojącej po prawej stronie. Są nimi $\lambda_1 = 0$ oraz $\lambda_2 = i$, $\lambda_3 = -i$. Odpowiadającymi im wektorami własnymi są

$$\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \qquad \begin{pmatrix} 1 \\ -i \\ 1 \end{pmatrix}, \qquad \begin{pmatrix} 1 \\ i \\ 1 \end{pmatrix}.$$

Ogólne rozwiązanie równania jednorodnego ma postać $\mathbf{y}(t) = \exp(tF)\cdot\mathbf{y}_{\mathrm{const}}$, gdzie $\mathbf{y}_{\mathrm{const}} = (a,b,c)$ jest jakimś dowolnym wektorem (którego na razie nie należy utożsamiać z wektorem warunków początkowych $\mathbf{y}(t=0) = \mathbf{y}_0$; to, czy takie utożsamienie będzie można zrobić, zależy od wyboru szczególnego rozwiązania równania niejednorodnego). Ponieważ macierz F jest diagonalizowalna, aby znaleźć e^{tF} rozkładamy dowolny wektor $\mathbf{y}_{\mathrm{const}}$ na wektory własne F:

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = (c-a)\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} + \frac{1}{2}[a + i(a+b-c)]\begin{pmatrix} 1 \\ -i \\ 1 \end{pmatrix} + \frac{1}{2}[a - i(a+b-c)]\begin{pmatrix} 1 \\ i \\ 1 \end{pmatrix},$$

i działamy nań macierzą e^{tF}:

$$\begin{aligned}
e^{tF}\cdot\begin{pmatrix} a \\ b \\ c \end{pmatrix} &= (c-a)\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} + \frac{1}{2}[a + i(a+b-c)]e^{it}\begin{pmatrix} 1 \\ -i \\ 1 \end{pmatrix} + \frac{1}{2}[a - i(a+b-c)]e^{-it}\begin{pmatrix} 1 \\ i \\ 1 \end{pmatrix} \\
&= \begin{pmatrix} \frac{1}{2}a[(1+i)e^{it} + (1-i)e^{-it}] + \frac{i}{2}b(e^{it} - e^{-it}) - \frac{i}{2}c(e^{it} - e^{-it}) \\ c - a + \frac{1}{2}a[(1-i)e^{it} + (1+i)e^{-it}] + \frac{1}{2}b(e^{it} + e^{-it}) - \frac{1}{2}c(e^{it} + e^{-it}) \\ c - a + \frac{1}{2}a[(1+i)e^{it} + (1-i)e^{-it}] + \frac{i}{2}b(e^{it} - e^{-it}) - \frac{i}{2}c(e^{it} - e^{-it}) \end{pmatrix} \\
&= \begin{pmatrix} \frac{1}{2}[(1+i)e^{it} + (1-i)e^{-it}] & \frac{i}{2}(e^{it} - e^{-it}) & -\frac{i}{2}(e^{it} - e^{-it}) \\ \frac{1}{2}[(1-i)e^{it} + (1+i)e^{-it}] - 1 & \frac{1}{2}(e^{it} + e^{-it}) & -\frac{1}{2}(e^{it} + e^{-it}) + 1 \\ \frac{1}{2}[(1+i)e^{it} + (1-i)e^{-it}] - 1 & \frac{i}{2}(e^{it} - e^{-it}) & -\frac{i}{2}(e^{it} - e^{-it}) + 1 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix}
\end{aligned}$$

Stojąca tu macierz jest właśnie macierzą e^{tF}. Jest ona jawnie rzeczywista. Można ją też przepisać w postaci

$$e^{tF} = \begin{pmatrix} \cos t - \sin t & -\sin t & \sin t \\ -1+\cos t + \sin t & \cos t & 1-\cos t \\ -1+\cos t - \sin t & -\sin t & 1+\sin t \end{pmatrix}.$$

Aby znaleźć szczególne rozwiązanie równania niejednorodnego, możemy posłużyć się metodą uzmienniania stałej. "Stałą" jest w tym przypadku stały wektor $\mathbf{y}_{\mathrm{const}} = (a, b, c)$. Szukamy więc rozwiązania w postaci

$$\mathbf{y}_{\mathrm{inhom}}(t) = e^{tF} \cdot \mathbf{v}(t)\,.$$

Wstawiamy ten Ansatz do równania niejednorodnego $\dot{\mathbf{y}}(t) = F \cdot \mathbf{y}(t) + \mathbf{f}(t)$ i otrzymujemy

$$e^{tF} \cdot \dot{\mathbf{v}}(t) = \mathbf{f}(t)\,.$$

Jak zwykle zdarzyło się to, co się zawsze w takim przypadku zdarza, tj. część wyrazów skróciła się. Zatem

$$\dot{\mathbf{v}}(t) = \left(e^{tF}\right)^{-1} \cdot \mathbf{f}(t)\,,$$

ale odwrotnością macierzy e^{tF} jest oczywiście macierz[53] e^{-tF}, więc na szczęście nie musimy bawić się w odwracanie macierzy 3×3... Wektorowa funkcja $\mathbf{v}(t)$ jest zatem dana przez całkę

$$\mathbf{v}(t) = \int^t dt'\, e^{-t'F} \cdot \mathbf{f}(t')\,,$$

a całe najogólniejsze rozwiązanie wyjściowego równania niejednorodnego ma postać

$$\mathbf{y}(t) = e^{tF} \cdot \mathbf{y}_{\mathrm{const}} + e^{tF} \cdot \int^t dt'\, e^{-t'F} \cdot \mathbf{f}(t')\,.$$

Jawnie:

$$e^{-tF} \cdot \mathbf{f}(t) = \begin{pmatrix} \cos t + \sin t & \sin t & -\sin t \\ -1+\cos t - \sin t & \cos t & 1-\cos t \\ -1+\cos t + \sin t & \sin t & 1-\sin t \end{pmatrix} \begin{pmatrix} t \\ 2 \\ 1+t \end{pmatrix} = \begin{pmatrix} t\cos t + \sin t \\ 1+\cos t - t\sin t \\ 1+\sin t + t\cos t \end{pmatrix},$$

[53] Proszę sprawdzić bezpośrednio, że $e^{tF} \cdot e^{-tF}$, czyli

$$\begin{pmatrix} \cos t - \sin t & -\sin t & \sin t \\ -1+\cos t + \sin t & \cos t & 1-\cos t \\ -1+\cos t - \sin t & -\sin t & 1+\sin t \end{pmatrix} \begin{pmatrix} \cos t + \sin t & \sin t & -\sin t \\ -1+\cos t - \sin t & \cos t & 1-\cos t \\ -1+\cos t + \sin t & \sin t & 1-\sin t \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

- jest to dobry sprawdzian, czy dobrze wyznaczyliśmy macierz e^{tF}! Drugim takim sprawdzianem jest jest równość $\frac{d}{dt}\left(e^{tF}\right) = F \cdot e^{tF}$.

więc

$$\mathbf{v}(t) = \int^t dt'\, e^{-t'F} \cdot \mathbf{f}(t') = \begin{pmatrix} t\sin t \\ t + t\cos t \\ t + t\sin t \end{pmatrix},$$

i wreszcie

$$\mathbf{y}_{\text{inhom}}(t) = e^{tF} \cdot \mathbf{v}(t) = \begin{pmatrix} \cos t - \sin t & -\sin t & \sin t \\ -1 + \cos t + \sin t & \cos t & 1 - \cos t \\ -1 + \cos t - \sin t & -\sin t & 1 + \sin t \end{pmatrix} \begin{pmatrix} t\sin t \\ t + t\cos t \\ t + t\sin t \end{pmatrix} = \begin{pmatrix} 0 \\ 2t \\ t \end{pmatrix}.$$

Jeśli chcemy, by stały wektor $\mathbf{y}_{\text{const}}$ miał interpretację warunku początkowego w $t = 0$, to całkę w powyższych wzorach należy napisać jako oznaczoną:

$$\mathbf{y}(t) = e^{tF} \cdot \mathbf{y}(0) + e^{tF} \cdot \int_0^t dt'\, e^{-t'F} \cdot \mathbf{f}(t'),$$

tak, by $\mathbf{y}_{\text{inhom}}(0) = \mathbf{0}$; jak widać znalezione wyżej rozwiązanie $\mathbf{y}_{\text{inhom}}(t)$ spełnia ten warunek. Całe rozwiązanie z warunkiem $\mathbf{y}(0) = (a, b, c)$ ma zatem postać

$$\begin{pmatrix} y_1(t) \\ y_2(t) \\ y_3(t) \end{pmatrix} = \begin{pmatrix} a\cos t - (a + b - c)\sin t \\ c - a + a\sin t + (a + b - c)\cos t \\ c - a + a\cos t - (a + b - c)\sin t \end{pmatrix} + \begin{pmatrix} 0 \\ 2t \\ t \end{pmatrix}.$$

Przedstawiona tu metoda jest zupełnie ogólna i zadziałałaby także w przypadku, gdyby macierz F nie była diagonalizowalna. Skoro jednak występująca w zadaniu macierz F jest diagonalizowalna, można podać znacznie prostszy sposób znalezienia najogólniejszej postaci rozwiązania. Jak poprzednio wypisujemy najpierw ogólne rozwiązanie równania jednorodnego:

$$\mathbf{y}_{\text{hom}}(t) = A \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} + Z\, e^{it} \begin{pmatrix} 1 \\ -i \\ 1 \end{pmatrix} + Z^*\, e^{-it} \begin{pmatrix} 1 \\ i \\ 1 \end{pmatrix}.$$

Jest ono jawnie rzeczywiste, jeśli Z jest zespoloną stałą.

Aby znaleźć jakieś szczególne rozwiązanie równania niejednorodnego, rozkładamy niejednorodność na znalezione wektory własne macierzy F:

$$\begin{pmatrix} t \\ 2 \\ 1 + t \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} + \frac{1}{2}(t + i) \begin{pmatrix} 1 \\ -i \\ 1 \end{pmatrix} + \frac{1}{2}(t - i) \begin{pmatrix} 1 \\ i \\ 1 \end{pmatrix}.$$

Szczególnego rozwiązanie równania niejednorodnego szukamy uzmienniając stałe A i Z w wypisanym wyżej rozwiązaniu równania jednorodnego

$$\mathbf{y}_{\text{inhom}}(t) = C(t) \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} + D(t)\, e^{it} \begin{pmatrix} 1 \\ -i \\ 1 \end{pmatrix} + D^*(t)\, e^{-it} \begin{pmatrix} 1 \\ i \\ 1 \end{pmatrix}.$$

Po wstawieniu tego do wyjściowego równania niejednorodnego wyrazy bez pochodnych funkcji $C(t)$ i $D(t)$ redukują się i otrzymujemy (kropki oznaczają pochodne po t)

$$\dot{C}(t)\begin{pmatrix}0\\1\\1\end{pmatrix}+\dot{D}(t)\,e^{it}\begin{pmatrix}1\\-i\\1\end{pmatrix}+\dot{D}^{*}(t)\,e^{-it}\begin{pmatrix}1\\i\\1\end{pmatrix}=\begin{pmatrix}0\\1\\1\end{pmatrix}+\frac{t+i}{2}\begin{pmatrix}1\\-i\\1\end{pmatrix}+\frac{t-i}{2}\begin{pmatrix}1\\i\\1\end{pmatrix}.$$

Ponieważ wektory własne macierzy F są liniowo niezależne, daje to następujące równania

$$\dot{C}(t)=1\,,\qquad \dot{D}(t)\,e^{it}=\frac{1}{2}(t+i)\,,\qquad \dot{D}^{*}(t)\,e^{-it}=\frac{1}{2}(t-i)$$

(trzecie równanie jest sprzężeniem zespolonym drugiego). Stąd

$$C(t)=t\,,\qquad D(t)=\frac{1}{2}\int dt\,e^{-it}\,(t+i)=\frac{i}{2}\,t\,e^{-it}\,.$$

Pełne rozwiązanie równania niejednorodego ma zatem postać

$$\mathbf{y}(t)=(A+t)\begin{pmatrix}0\\1\\1\end{pmatrix}+\left(Z\,e^{it}+\frac{i}{2}\,t\right)\begin{pmatrix}1\\-i\\1\end{pmatrix}+\left(Z^{*}\,e^{-it}-\frac{i}{2}\,t\right)\begin{pmatrix}1\\i\\1\end{pmatrix}.$$

Wstawiając tu wyznaczone już wcześniej (przy okazji rozwiązywania tego układu równań metodą ogólną) współczynniki rozkładu warunku początkowego $\mathbf{y}(0)=(a,b,c)$ na wektory własne macierzy F: $A=c-a$, $Z=\frac{1}{2}[a+i(a+b-c)]$ można się przekonać, że jest to to samo rozwiązanie, co uzyskane poprzednio.

Uproszczoną metodą zademonstrowaną wyżej można się także posłużyć do rozwiązania następującego zadania

Zadanie Ode.11
Rozwiązać układ liniowych równań różniczkowych drugiego rzędu

$$\begin{aligned}\frac{d^2}{dt^2}\,y_1(t)+\omega^2\,(\ 2y_1-y_2-y_3)&=0\\ \frac{d^2}{dt^2}\,y_2(t)+\omega^2(-y_1+2y_2-y_3)&=0\\ \frac{d^2}{dt^2}\,y_3(t)+\omega^2(-y_1-y_2+2y_3)&=0\,.\end{aligned}$$

Rozwiązanie: Układ ten można zapisać w postaci macierzowej

$$\frac{d^2}{dt^2}\begin{pmatrix}y_1\\y_2\\y_3\end{pmatrix}=\omega^2\begin{pmatrix}-2&1&1\\1&-1&1\\1&1&-2\end{pmatrix}\begin{pmatrix}y_1\\y_2\\y_3\end{pmatrix}.$$

Wartościami własnymi występującej tu macierzy są $\lambda_1 = 0$ (pojedyncza) i $\lambda_2 = -3$ (podwójna), a odpowiadającymi im wektorami własnymi są

$$\lambda_1 : \quad \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \qquad \lambda_2 : \quad \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \quad \text{i} \quad \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}.$$

Mimo że jedna wartość własna ma krotność dwa, macierz jest diagonalizowalna. (Oczywiście zamiast podanych dwu wektorów własnych odpowiadających $\lambda_2 = -3$, można by wziąć jakieś inne dwie kombinacje liniowe tychże). Rozwiązania możemy więc szukać w potaci Ansatzu

$$\begin{pmatrix} y_1(t) \\ y_2(t) \\ y_3(t) \end{pmatrix} = a(t) \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + b(t) \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + c(t) \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}.$$

Po wstawieniu do równania otrzymujemy

$$\frac{d^2 a(t)}{dt^2} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \frac{d^2 b(t)}{dt^2} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + \frac{d^2 c(t)}{dt^2} \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} = -3\omega^2 \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} - 3\omega^2 \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}.$$

Skorzystaliśmy tu z faktu, że macierz działając na wektor własny mnoży go przez odpowiadającą mu wartość własną. Ponieważ wektory własne są liniowo niezależne, muszą zachodzić równości

$$\frac{d^2 a(t)}{dt^2} = 0\,, \qquad \frac{d^2 b(t)}{dt^2} = -3\omega^2 b(t)\,, \qquad \frac{d^2 c(t)}{dt^2} = -3\omega^2 c(t)\,,$$

których rozwiązania są oczywiste: $a(t) = A_1 + A_2 t$, $b(t) = B_1 \cos\omega t + B_2 \sin\omega t$ i $c(t) = C_1 \cos\omega t + C_2 \sin\omega t$.

Przy okazji warto się zastanowić, jak należałoby rozwiązać podobne liniowe równanie drugiego rzędu w przypadku, gdyby macierz nie była diagonalizowalna (tj. byłoby mniej wektorów własnych niż potrzeba). Równanie

$$\frac{d^2}{dt^2}\, \mathbf{y}(t) = F \cdot \mathbf{y}(t)\,,$$

w którym $\mathbf{y}(t)$ jest n-wymiarowym wektorem, a F macierzą $n \times n$ można oczywiście przerobić na równanie pierwszego rzędu:

$$\frac{d}{dt} \begin{pmatrix} \mathbf{y}(t) \\ \dot{\mathbf{y}}(t) \end{pmatrix} = \begin{pmatrix} 0 & I \\ F & 0 \end{pmatrix} \begin{pmatrix} \mathbf{y}(t) \\ \dot{\mathbf{y}}(t) \end{pmatrix},$$

czyli na standardowe równanie macierzowe postaci $\dot{\mathbf{v}}(t) = \tilde{F} \cdot \mathbf{v}(t)$, którego rozwiązaniem jest

$$\mathbf{v}(t) = e^{t\tilde{F}} \cdot \mathbf{v}_{\text{const}}\,, \quad \text{gdzie} \quad \tilde{F} = \begin{pmatrix} 0 & I \\ F & 0 \end{pmatrix}, \quad \text{a} \quad \mathbf{v}_{\text{const}} = \begin{pmatrix} \mathbf{y}_0 \\ \dot{\mathbf{y}}_0 \end{pmatrix}.$$

Istnieje też prosty związek między wielomianami charakterystycznymi wyjściowej macierzy F i macierzy $\tilde{F}$:

$$W_{\tilde{F}}(\tilde{\lambda}) = W_F(\tilde{\lambda}^2)\,.$$

Jeśli więc $\lambda_1, \dots, \lambda_k$ są wartościami własnymi o krotnościach $n_1, \dots, n_k$ ($n_1 + \ldots + n_k = n$) macierzy F, to wartościami własnymi $\tilde{\lambda}$ macierzy $\tilde{F}$ są $\pm\sqrt{\lambda_1}, \dots, \pm\sqrt{\lambda_k}$ o takich samych krotnościach.[54] Można też podać związek między wektorami własnymi macierzy $\tilde{F}$ i macierzy F: jeśli $\mathbf{w}_{(i)}$ jest (n-wymiarowym) wektorem własnym macierzy F odpowiadającym jej wartości własnej λ_i, to wektorami własnymi macierzy $\tilde{F}$ odpowiadającymi wartościom własnym $\pm\sqrt{\lambda_i}$ są ($2n$-wymiarowe) wektory

$$\tilde{\mathbf{w}}_{(\pm i)} \equiv \begin{pmatrix} \mathbf{w}_{(i)} \\ \pm\sqrt{\lambda_i}\mathbf{w}_{(i)} \end{pmatrix}.$$

Jasne jest więc, że jeśli macierz F jest diagonalizowalna tj. ma n wektorów własnych, to istnieje[55] $2n$ wektorów własnych macierzy $\tilde{F}$ i ogólne rozwiązanie równania $\dot{\mathbf{v}}(t) = \tilde{F}\cdot\mathbf{v}(t)$ można napisać od ręki

$$\mathbf{v}(t) = \sum_{\pm i} A_{\pm i}\, e^{t\tilde{\lambda}_{\pm i}}\, \tilde{\mathbf{w}}_{(\pm i)}\,,$$

gdzie $\tilde{\lambda}_{\pm i} \equiv \pm\sqrt{\lambda_i}$. Suma rozciąga się tu na $2n$ wektorów własnych $\tilde{\mathbf{w}}_{(\pm i)}$ macierzy $\tilde{F}$. Kto wprawny, ten od razu widzi, że "górne pięterko" rozwiązania $\mathbf{v}(t)$, tzn. n górnych składowych tego wektora, jest tym samym rozwiązaniem, które uzyskaliśmy wyżej stosując uproszczoną metodę. Podejście poprzez sprowadzenie wyjściowego równania drugiego rzędu do równania pierwszego rzędu z macierzą $\tilde{F}$ zadziała jednak także i wtedy, gdy uproszczone zawiedzie, tj. gdy macierz F okaże się niediagonalizowalna. Oczywiście musimy wtedy szukać wektorów pierwiastkowych macierzy $2n\times 2n$, tj. wektorów (liniowo niezależnych od wektorów własnych), na których zeruje się macierz

$$\begin{pmatrix} \mp\sqrt{\lambda_i}I & I \\ F & \mp\sqrt{\lambda_i}I \end{pmatrix}\begin{pmatrix} \mp\sqrt{\lambda_i}I & I \\ F & \mp\sqrt{\lambda_i}I \end{pmatrix} = \begin{pmatrix} F+\lambda_i I & \mp 2\sqrt{\lambda_i}I \\ \mp 2\sqrt{\lambda_i}F & F+\lambda_i I \end{pmatrix},$$

lub stosować sztuczki różniczkowaniem w twierdzeniu Cayleya-Hamiltona. Cała procedura może być uciążliwa, niemniej w "principie" daje się przeprowadzić.

[54] Wyjątek stanowi wartość własna równa zeru: jeśli macierz F ma taką wartość własną o krotności n_0, to macierz $\tilde{F}$ ma wartość własną $\tilde{\lambda} = 0$ o krotności $2n_0$.

[55] Znów przypadkiem szczególnym jest występowanie zerowej wartości własnej macierzy F: macierz $\tilde{F}$ ma wtedy tylko jeden wektor własny odpowiadający $\tilde{\lambda} = 0$ i trzeba znaleźć wektor pierwiastkowy. Nietrudno zobaczyć, że tymi dwoma wektorami są

$$\tilde{\mathbf{w}}_{(0)} = \begin{pmatrix} \mathbf{w}_{(0)} \\ \mathbf{0} \end{pmatrix} \quad \text{(w. własny)} \qquad \text{oraz} \qquad \tilde{\mathbf{w}}_{(p)} = \begin{pmatrix} \mathbf{w}_{(0)} \\ \mathbf{w}_{(0)} \end{pmatrix} \quad \text{(w. pierwiastkowy, bo } \tilde{F}\cdot\tilde{\mathbf{w}}_{(p)} = \tilde{\mathbf{w}}_{(0)}\text{)}.$$

Patrząc na rozwiązane właśnie zadanie widzimy dlaczego tak jest: musi wtedy w rozwiązaniu wystąpić człon potęgowy (w tym zadaniu liniowy) w t, a to właśnie, jak już wiemy, gwarantuje wektor pierwiastkowy.

Zadanie Ode.12
Rozwiązać liniowe równanie różniczkowe drugiego rzędu

$$\frac{d^2}{dt^2}\begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} = \begin{pmatrix} 2 & 2 \\ -2 & 6 \end{pmatrix}\begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix}.$$

Rozwiązanie: Występująca tu macierz F ma jedną podwójną wartość własną $\lambda = 4$ i tylko jeden wektor własny

$$\mathbf{w} = \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

i sztuczka z rozpisaniem $\mathbf{y}(t)$ na wektory własne M nie uda się. Trzeba przerobić wyjściowe równanie drugiego rzędu na równanie pierwszego rzędu

$$\frac{d}{dt}\begin{pmatrix} y_1(t) \\ y_2(t) \\ \dot{y}_1(t) \\ \dot{y}_2(t) \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 2 & 2 & 0 & 0 \\ -2 & 6 & 0 & 0 \end{pmatrix}\begin{pmatrix} y_1(t) \\ y_2(t) \\ \dot{y}_1(t) \\ \dot{y}_2(t) \end{pmatrix}.$$

Wektorami własnym figurującej w tej postaci równania macierzy $\tilde{F}$ są

$$\tilde{\lambda}_{(+)} = 2: \quad \tilde{\mathbf{w}}_{(+)} = \begin{pmatrix} 1 \\ 1 \\ 2 \\ 2 \end{pmatrix} \qquad \text{oraz} \quad \tilde{\lambda}_{(-)} = -2: \quad \tilde{\mathbf{w}}_{(-)} = \begin{pmatrix} 1 \\ 1 \\ -2 \\ -2 \end{pmatrix},$$

ale trzeba znaleźć odpowiadające tym wartościom własnym wektory pierwiastkowe. Szukamy zatem rozwiązań równań (macierze $(\tilde{F} \mp 2I)^2$ możemy obliczyć korzystając z podanego wyżej ogólnego wzoru)

$$\begin{pmatrix} 6 & 2 & -4 & 0 \\ -2 & 10 & 0 & -4 \\ -8 & -8 & 6 & 2 \\ 8 & -24 & -2 & 10 \end{pmatrix}\begin{pmatrix} a_{(+)} \\ b_{(+)} \\ c_{(+)} \\ d_{(+)} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

oraz

$$\begin{pmatrix} 6 & 2 & 4 & 0 \\ -2 & 10 & 0 & 4 \\ 8 & 8 & 6 & 2 \\ -8 & 24 & -2 & 10 \end{pmatrix}\begin{pmatrix} a_{(-)} \\ b_{(-)} \\ c_{(-)} \\ d_{(-)} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Znajdujemy[56] w ten sposób dwa wektory pierwiastkowe $\tilde{\mathbf{u}}_{(+)}$ i $\tilde{\mathbf{u}}_{(-)}$ odpowiadające wartościom własnym $\tilde{\lambda}_{(+)} = 2$ i $\tilde{\lambda}_{(+)} = -2$:

$$\tilde{\mathbf{u}}_{(+)} = \begin{pmatrix} 2 \\ 0 \\ 3 \\ -1 \end{pmatrix}, \qquad \tilde{\mathbf{u}}_{(-)} = \begin{pmatrix} 2 \\ 0 \\ -3 \\ 1 \end{pmatrix}.$$

[56] Ponieważ do wektora pierwiastkowego można zawsze dodać wielokrotność odpowiadającego tej samej wartości własnej wektora własnego, można szukać rozwiązań, w których $b_{(\pm)} = 0$.

Przedstawiamy następnie dowolny wektor (warunku początkowego) w postaci kombinacji liniowej $\alpha_{(+)}\tilde{\mathbf{w}}_{(+)} + \beta_{(+)}\tilde{\mathbf{u}}_{(+)} + \alpha_{(-)}\tilde{\mathbf{w}}_{(-)} + \beta_{(-)}\tilde{\mathbf{u}}_{(-)}$ wektorów własnych i pierwiastkowych:

$$\begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \alpha_{(+)} \begin{pmatrix} 1 \\ 1 \\ 2 \\ 2 \end{pmatrix} + \beta_{(+)} \begin{pmatrix} 2 \\ 0 \\ 3 \\ -1 \end{pmatrix} + \alpha_{(-)} \begin{pmatrix} 1 \\ 1 \\ -2 \\ -2 \end{pmatrix} + \beta_{(-)} \begin{pmatrix} 2 \\ 0 \\ -3 \\ 1 \end{pmatrix}.$$

Dość łatwo znajdujemy, że

$$\alpha_{(\pm)} = \frac{1}{2}b \pm \frac{1}{16}(c + 3d)\,,$$
$$\beta_{(\pm)} = \frac{1}{4}(a - b) \pm \frac{1}{8}(c - d)\,.$$

Rozwiązaniem równania $\dot{\mathbf{v}}(t) = \tilde{F} \cdot \mathbf{v}(t)$ jest

$$\begin{aligned} \mathbf{v}(t) &= e^{t\tilde{F}} \cdot (\alpha_{(+)}\,\tilde{\mathbf{w}}_{(+)} + \beta_{(+)}\,\tilde{\mathbf{u}}_{(+)} + \alpha_{(-)}\,\tilde{\mathbf{w}}_{(-)} + \beta_{(-)}\,\tilde{\mathbf{u}}_{(-)}) \\ &= \alpha_{(+)}\,e^{2t}\,\tilde{\mathbf{w}}_{(+)} + \beta_{(+)}\,e^{2t}\,e^{t(\tilde{F}-2I)} \cdot \tilde{\mathbf{u}}_{(+)} \\ &\quad + \alpha_{(-)}\,e^{-2t}\,\tilde{\mathbf{w}}_{(-)} + \beta_{(-)}\,e^{-2t}\,e^{t(\tilde{F}+2I)} \cdot \tilde{\mathbf{u}}_{(-)} \end{aligned}$$

Jak zwykle działanie $e^{t\tilde{F}}$ na wektory pierwiastkowe, po wyfaktoryzowaniu czynnika $e^{t\tilde{\lambda}}$ możemy znaleźć przez rozwinięcie $e^{t(\tilde{F}-\tilde{\lambda})}$ w szereg Taylora: ponieważ $(\tilde{F} - \tilde{\lambda})^2$ zeruje się na wektorze pierwiastkowym (odpowiadającym $\tilde{\lambda}$), szereg urywa się i dostajemy

$$e^{t(\tilde{F}-2I)} \cdot \tilde{\mathbf{u}}_{(+)} = \tilde{\mathbf{u}}_{(+)} + t(\tilde{F} - 2I) \cdot \tilde{\mathbf{u}}_{(+)} = \begin{pmatrix} 2-t \\ -t \\ 3-2t \\ -1-2t \end{pmatrix},$$
$$e^{t(\tilde{F}+2I)} \cdot \tilde{\mathbf{u}}_{(-)} = \tilde{\mathbf{u}}_{(-)} + t(\tilde{F} + 2I) \cdot \tilde{\mathbf{u}}_{(-)} = \begin{pmatrix} 2+t \\ t \\ -3-2t \\ 1-2t \end{pmatrix},$$

Ogólne rozwiązanie ma zatem postać

$$\begin{aligned} \begin{pmatrix} y_1(t) \\ y_2(t) \\ \dot{y}_1(t) \\ \dot{y}_2(t) \end{pmatrix} &= e^{2t} \left\{ \alpha_{(+)} \begin{pmatrix} 1 \\ 1 \\ 2 \\ 2 \end{pmatrix} + \beta_{(+)} \begin{pmatrix} 2-t \\ -t \\ 3-2t \\ -1-2t \end{pmatrix} \right\} \\ &\quad + e^{-2t} \left\{ \alpha_{(-)} \begin{pmatrix} 1 \\ 1 \\ -2 \\ -2 \end{pmatrix} + \beta_{(-)} \begin{pmatrix} 2+t \\ t \\ -3-2t \\ 1-2t \end{pmatrix} \right\}. \end{aligned}$$

Nietrudno dostrzec, że "dolne pięterka" są, tak jak być powinno, pochodnymi odpowiednich "górnych pięterek". Możemy zatem napisać rozwiązania

$$\begin{aligned} y_1(t) &= \alpha_{(+)}\,e^{2t} + \beta_{(+)}\,(2-t)\,e^{2t} + \alpha_{(-)}\,e^{-2t} + \beta_{(-)}\,(2+t)\,e^{-2t}\,, \\ y_2(t) &= \alpha_{(+)}\,e^{2t} - \beta_{(+)}\,t\,e^{2t} + \alpha_{(-)}\,e^{-2t} + \beta_{(-)}\,t\,e^{-2t}\,. \end{aligned}$$

Sprawdzenie, że spełniają one układ równań $d^2y_1/dt^2 = 2y_1 + 2y_2$, $d^2y_1/dt^2 = -2y_1 + 6y_2$ pozostawiamy jako proste ale obowiązkowe (wyrabianie niezbędnych nawyków!) ćwiczenie.

Oczywiście podany układ równań można też rozwiązać prościej: Dodając i odejmując od siebie wyjściowe równania

$$\begin{aligned} \frac{d^2}{dt^2}\,y_1 &= 2y_1 + 2y_2 \\ \frac{d^2}{dt^2}\,y_2 &= -2y_1 + 6y_2 \end{aligned}$$

i wprowadzając nowe zmienne $\xi = y_1 + y_2$, $\eta = y_1 - y_2$ otrzymujemy równoważne równania

$$\begin{aligned} \frac{d^2}{dt^2}\,\xi &= 4\xi - 4\eta \\ \frac{d^2}{dt^2}\,\eta &= 4\eta\,. \end{aligned}$$

Drugie z tych równań daje się natychmiast rozwiązać:

$$\eta(t) = 2\beta_{(+)}\,e^{2t} + 2\beta_{(-)}\,e^{-2t}\,.$$

Dowolne stałe oznaczyliśmy $2\beta_{(\pm)}$, żeby od razu uzyskać tę samą postać rozwiązania układu, co poprzednio. Pierwsze równanie jest wtedy równaniem liniowym z niejednorodnością. Jego rozwiązanie jest więc, jak zwykle, sumą ogólnego rozwiązania $\xi_{\rm hom}(t)$ równania jednorodnego $d^2\xi/dt^2 = 4\xi$ i jakiegoś szczególnego rozwiązania $\xi_{\rm inhom}(t)$ równania niejednorodnego. Ogólna metoda szukania takiego szczególnego rozwiązania jest podana w zadaniach Ode.13 i Ode.14. Tu postaramy się je znaleźć stosując prosty Ansatz oparty na tym, że w takich sytuacjach, gdy macierz problemu liniowego jest niediagonalizowalna, w rozwiązaniu występują wyrazy proporcjonalne do $t\,e^{\lambda t}$. Podstawiamy zatem

$$\xi_{\rm inhom}(t) = (A_+ + B_+ t)\,e^{2t} + (A_- + B_- t)\,e^{-2t}\,.$$

do równania $d^2\xi/dt^2 = 4\xi - 8\beta_{(+)}e^{2t} - 8\beta_{(-)}e^{-2t}$ i mamy

$$\begin{aligned} 4(A_+ + B_+ + B_+ t)\,e^{2t} + 4(A_- - B_- + B_- t)\,e^{-2t} &= 4(A_+ + B_+ t)\,e^{2t} + 4(A_- + B_- t)\,e^{-2t} \\ &\quad - 8\beta_{(+)}e^{2t} - 8\beta_{(-)}e^{-2t}\,. \end{aligned}$$

Widać, że równanie będzie spełnione, jeśli podstawimy $B_+ = -2\beta_{(+)}$ i $B_- = 2\beta_{(-)}$. Stałe $A_\pm$ są dowolne, bo to są po prostu stałe w ogólnym rozwiązaniu $\xi_{\rm hom}(t)$ równania jednorodnego. Zapiszemy je jako $A_\pm = 2(\alpha_{(\pm)} + \beta_{(\pm)})$.

Mamy więc rozwiązania

$$\begin{aligned}
\xi(t) &= 2\,(\alpha_{(+)} + \beta_{(+)} - \beta_{(+)}t)\,e^{2t} + 2\,(\alpha_{(-)} + \beta_{(-)} + \beta_{(-)}t)\,e^{-2t} \\
\eta(t) &= \qquad\qquad 2\,\beta_{(+)}\,e^{2t} \qquad\qquad + 2\,\beta_{(-)}\,e^{-2t}\,.
\end{aligned}$$

Biorąc połowę sumy i połowę różnicy dostajemy

$$\begin{aligned}
y_1(t) &= \alpha_{(+)}\,e^{2t} + \beta_{(+)}(2-t)\,e^{2t} + \alpha_{(-)}\,e^{-2t} + \beta_{(-)}(2+t)\,e^{-2t} \\
y_2(t) &= \alpha_{(+)}\,e^{2t} - \beta_{(+)}\,t\,e^{2t} + \alpha_{(-)}\,e^{-2t} + \beta_{(-)}\,t\,e^{-2t}\,,
\end{aligned}$$

czyli te same rozwiązania, co poprzedno.

Rozpatrzymy teraz ogólne równanie drugiego rzędu liniowe i liniowe z niejednorodnością o stałych współczynnikach na jedną tylko funkcję. Użytą tu metodę nie polegającą na przekształceniu takiego równania w układ dwóch rówanań pierwszego rzędu można też rozciągnąć na analogiczne równania wyższego rzędu. Jest to proste w przypadku równań jednorodnych i trochę trudniejsze w przypadku niejednorodnych.

Zadanie Ode.13
Rozwiązać liniowe równanie różniczkowe (l.r.r.) drugiego rzędu

$$\frac{d^2}{dt^2}\,y + a\,\frac{d}{dt}\,y + by = 0\,,$$

o stałych i rzeczywistych współczynnikach a i b, przyjmując jako warunki początkowe[57] $y(0) = y_0$ oraz $\dot{y}(0) = \dot{y}_0$ (równanie różniczkowe zwyczajne rzędu n wymaga zadania n warunków początkowych). Rozpatrzyć wszystkie możliwe przypadki.
Rozwiązanie: Postulujemy rozwiązanie postaci[58] $y(t) = A\exp(\lambda t)$. Daje to równanie kwadratowe na λ (także zwane równaniem charakterystycznym l.r.r.):

$$\lambda^2 + a\lambda + b = 0\,.$$

Równanie to może mieć dwa różne pierwiastki rzeczywiste, jeden rzeczywisty pierwiastek podwójny lub dwa wzajemnie sprzężone pierwiastki zespolone.

- Jeśli pierwiastkami są rzeczywiste i różne λ_1 i λ_2, to ogólnym rozwiązaniem jest

  $$y(t) = A_1\,e^{\lambda_1 t} + A_2\,e^{\lambda_2 t}\,.$$

  Z podanych warunków początkowych wyznaczamy stałe A_1 i A_2:

  $$\begin{aligned}
  A_1 + A_2 &= y(0) = y_0\,, \\
  \lambda_1 A_1 + \lambda_2 A_2 &= \dot{y}(0) = \dot{y}_0\,,
  \end{aligned}$$

[57] Ogólniejszymi warunkami byłyby $y(t_0) = y_0$ oraz $\dot{y}(t_0) = \dot{y}_0$; czasem w zastosowaniach warunkami mogą być $y(t_1) = y_1$ oraz $\dot{y}(t_2) = \dot{y}_2$.

[58] Fizyk wykształcony na *Feynmana Wykładach z Fizyki* zwykł podstawiać raczej $y(t) = A\exp(i\omega t)$, ale wynik końcowy oczywiście od tego nie zależy.

Dostajemy stąd

$$y(t) = \frac{1}{\lambda_2 - \lambda_1}\left[(\lambda_2\, y_0 - \dot{y}_0)\, e^{\lambda_1 t} - (\lambda_1\, y_0 - \dot{y}_0)\, e^{\lambda_2 t}\right]$$
$$\equiv \frac{1}{\lambda_2 - \lambda_1}\left[y_0\left(\lambda_2\, e^{\lambda_1 t} - \lambda_1\, e^{\lambda_2 t}\right) - \dot{y}_0\left(e^{\lambda_1 t} - e^{\lambda_2 t}\right)\right].$$

- Jeśli jest jeden pierwiastek podwójny $\lambda = -\frac{1}{2}a$, tj. gdy $a^2 = 4b$ ($\Delta = 0$), wówczas rozwiązaniem jest

$$y(t) = (A + B\,t)\, e^{\lambda t}.$$

Sprawdźmy to:

$$\frac{d}{dt}\, y(t) = \lambda\,(A + B\,t)\, e^{\lambda t} + B\, e^{\lambda t},$$
$$\frac{d^2}{dt^2}\, y(t) = \lambda^2\,(A + B\,t)\, e^{\lambda t} + 2\lambda B\, e^{\lambda t}.$$

Wstawiamy to do lewej strony równania i mamy

$$\lambda^2\,(A + B\,t)\, e^{\lambda t} + 2\lambda B\, e^{\lambda t} + a\left[\lambda\,(A + B\,t)\, e^{\lambda t} + B\, e^{\lambda t}\right] + b\,(A + B\,t)\, e^{\lambda t}.$$

Ponieważ A i B są dwiema dowolnymi stałymi, wyrazy je mnożące powinny zerować się niezależnie. Istotnie: współczynnik przy $Ae^{\lambda t}$ jest równy $\lambda^2 + \lambda a + b = 0$; współczynnikiem przy $Be^{\lambda t}$ jest zaś $(\lambda^2 + \lambda a + b)t + 2\lambda + a$; znika on, gdyż $\lambda = -\frac{1}{2}a$. Z warunków początkowych mamy w tym przypadku:

$$A = y_0, \qquad B + \lambda A = \dot{y}_0,$$

i stąd

$$y(t) = [y_0 + (\dot{y}_0 - \lambda\, y_0)\, t]\, e^{\lambda t}.$$

- Wreszcie, gdy są dwa pierwiastki zespolone λ i λ^* rozwiązanie ma ogólną postać

$$y(t) = (A + iB)\, e^{\lambda t} + (A - iB)\, e^{\lambda^* t}.$$

Jest ono, jak łatwo zobaczyć, rzeczywiste ($y^*(t) = y(t)$). Znów są dwie stałe dowolne, które należy wyznaczyć z warunków początkowych:

$$2A = y_0, \qquad (A + iB)\lambda + (A - iB)\lambda^* = \dot{y}_0.$$

Stąd $B = -i[2\dot{y}_0 - (\lambda + \lambda^*)y_0]/[2(\lambda - \lambda^*)]$ i rozwiązanie uwzględniające warunki początkowe ma postać:

$$y(t) = \frac{1}{2}\, y_0\left(e^{\lambda t} + e^{\lambda^* t}\right) + \frac{2\dot{y}_0 - (\lambda + \lambda^*)y_0}{2(\lambda - \lambda^*)}\left(e^{\lambda t} - e^{\lambda^* t}\right).$$

Na późniejszy użytek przepiszemy je jeszcze tak

$$y(t) = \frac{y_0}{\lambda^* - \lambda}\left(\lambda^* e^{\lambda t} - \lambda\, e^{\lambda^* t}\right) + \frac{\dot{y}_0}{\lambda^* - \lambda}\left(e^{\lambda^* t} - e^{\lambda t}\right).$$

Dobrze jest zobaczyć, że rozwiązanie w przypadku, gdy jest jeden pierwiastek rzeczywisty podwójny można otrzymać jako granicę $\lambda_2 \to \lambda_1 \equiv \lambda$ pierwszego przypadku i jako granicę $\mathrm{Im}\lambda \to 0$ przypadku trzeciego. W pierwszym przypadku obliczamy granice (wykorzystując definicję pochodnej jako ilorazu różnicowego)

$$
\begin{aligned}
\lim_{\lambda_2\to\lambda_1\equiv\lambda} \frac{e^{\lambda_2 t}-e^{\lambda_1 t}}{\lambda_2-\lambda_1} &= \frac{d}{d\lambda}\, e^{\lambda t} = t\, e^{\lambda t} \\
\lim_{\lambda_2\to\lambda_1\equiv\lambda} \frac{\lambda_2\, e^{\lambda_1 t}-\lambda_1\, e^{\lambda_2 t}}{\lambda_2-\lambda_1} &= \lim_{\lambda_2\to\lambda_1} \frac{1}{\lambda_2-\lambda_1}\left[(\lambda_2-\lambda_1)e^{\lambda_1 t}+\lambda_1 e^{\lambda_1 t}-(\lambda_1-\lambda_2)e^{\lambda_2 t}-\lambda_2 e^{\lambda_2 t}\right] \\
&= \lim_{\lambda_2\to\lambda_1}\left\{e^{\lambda_1 t}+e^{\lambda_2 t}-\frac{\lambda_2\, e^{\lambda_2 t}-\lambda_1\, e^{\lambda_1 t}}{\lambda_2-\lambda_1}\right\} = 2\, e^{\lambda t}-\frac{d}{d\lambda}\left(\lambda\, e^{\lambda t}\right) \\
&= 2\, e^{\lambda t}-e^{\lambda t}-\lambda\, t\, e^{\lambda t} = (1-\lambda\, t)\, e^{\lambda t}\,.
\end{aligned}
$$

Zatem

$$
\begin{aligned}
\lim_{\lambda_2\to\lambda_1} y(t) &= \lim_{\lambda_2\to\lambda_1} \frac{1}{\lambda_2-\lambda_1}\left[y_0\left(\lambda_2\, e^{\lambda_1 t}-\lambda_1\, e^{\lambda_2 t}\right)+\dot{y}_0\left(e^{\lambda_2 t}-e^{\lambda_1 t}\right)\right] \\
&= y_0\,(1-\lambda\, t)\, e^{\lambda t}+\dot{y}_0\, t\, e^{\lambda t}\,,
\end{aligned}
$$

co jest tym samym wynikiem, który uzyskaliśmy w punkcie drugim. Podobnie w trzecim przypadku, gdy $\lambda = \xi + i\eta$ (i ξ utożsamimy z pojedynczym rzeczywistym pierwiastkiem równania charakterystycznego) granica $\mathrm{Im}\lambda \equiv \eta \to 0$ daje

$$
\begin{aligned}
\lim_{\eta\to 0} y(t) &= \lim_{\eta\to 0}\left\{\frac{1}{2}\, y_0\, e^{\xi t}\left(e^{i\eta t}+e^{-i\eta t}\right)+\frac{2\dot{y}_0-2\xi\, y_0}{4i\eta}\, e^{\xi t}\left(e^{i\eta t}-e^{-i\eta t}\right)\right\} \\
&= y_0\, e^{\xi t}+(\dot{y}_0-\xi\, y_0)\, e^{\xi t}\lim_{\eta\to 0}\frac{e^{i\eta t}-e^{-i\eta t}}{2i\eta} \\
&= y_0\, e^{\xi t}+(\dot{y}_0-\xi\, y_0)\, e^{\xi t}\lim_{\eta\to 0}\frac{\sin(\eta t)}{\eta} = \left[y_0+(\dot{y}_0-\xi\, y_0)\, t\right] e^{\xi t}\,,
\end{aligned}
$$

czyli ten sam rezultat, co poprzednio. Pokazuje to, że przy ustalonych warunkach początkowych rozwiązanie jest ciągłą funkcją parametrów a i b równania.

Zadanie Ode.14
Rozwiązać to samo, co w Zadaniu Ode.13 równanie różniczkowe

$$
\frac{d^2}{dt^2}\, y + a\, \frac{d}{dt}\, y + by = 0\,,
$$

$y(0) = y_0$, $\dot{y}(0) = \dot{y}_0$, sprowadzając je do równania pierwszego rzędu.
Rozwiązanie: Wprowadzamy oznaczenia: $y_1 \equiv y$ oraz $y_2 \equiv \dot{y} = \dot{y}_1$. Mamy wtedy

$$
\begin{aligned}
\frac{d}{dt}\, y_1 &= y_2\,, \\
\frac{d}{dt}\, y_2 &= \frac{d^2}{dt^2}\, y = -a\frac{d}{dt}\, y - by \equiv -ay_2 - by_1\,,
\end{aligned}
$$

lub to samo macierzowo

$$\frac{d}{dt}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -b & -a \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \equiv A \cdot \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.$$

Rozwiązaniem równania z podanymi warunkami początkowymi $y_1(0) = y_1^0 \equiv y_0$, $y_2(0) = y_2^0 \equiv \dot{y}_0$ jest wtedy[59]

$$\begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} = e^{tA} \cdot \begin{pmatrix} y_0 \\ \dot{y}_0 \end{pmatrix}.$$

Aby znaleźć eksponens macierzy A szukamy jej wektorów własnych, rozwiązując jak zwykle jej równanie charakterystyczne

$$\det\begin{pmatrix} -\lambda & 1 \\ -b & -a-\lambda \end{pmatrix} = \lambda^2 + \lambda a + b = 0\,,$$

Wyszło oczywiście to samo równanie charakterystyczne, co w poprzedniej metodzie. Załóżmy najpierw, że ma ono dwa różne pierwiastki λ_1 i λ_2. Równanie wyznaczające wektory własne

$$\begin{pmatrix} -\lambda_{1,2} & 1 \\ -b & -a-\lambda_{1,2} \end{pmatrix}\begin{pmatrix} \alpha_{1,2} \\ \beta_{1,2} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

daje

$$\mathbf{w}_1 = \begin{pmatrix} 1 \\ \lambda_1 \end{pmatrix}, \qquad \mathbf{w}_2 = \begin{pmatrix} 1 \\ \lambda_2 \end{pmatrix},$$

(ponieważ dwa te równania muszą być liniowo zależne rozwiązujemy górne kładąc po prostu $\alpha_{1,2} = 1$). Rozkładamy następnie warunek początkowy na wektory własne $\mathbf{w}_1$ i $\mathbf{w}_2$

$$\begin{pmatrix} y_0 \\ \dot{y}_0 \end{pmatrix} = \zeta_1\begin{pmatrix} 1 \\ \lambda_1 \end{pmatrix} + \zeta_2\begin{pmatrix} 1 \\ \lambda_2 \end{pmatrix} \equiv \begin{pmatrix} 1 & 1 \\ \lambda_1 & \lambda_2 \end{pmatrix}\begin{pmatrix} \zeta_1 \\ \zeta_2 \end{pmatrix},$$

i wyznaczywszy współczynniki rozkładu ζ_1 i ζ_2 działamy na wektor warunków początkowych macierzą $\exp(tA)$:

$$\begin{aligned}\begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} &= e^{tA} \cdot \left\{\frac{\lambda_2\, y_0 - \dot{y}_0}{\lambda_2 - \lambda_1}\begin{pmatrix} 1 \\ \lambda_1 \end{pmatrix} - \frac{\lambda_1\, y_0 - \dot{y}_0}{\lambda_2 - \lambda_1}\begin{pmatrix} 1 \\ \lambda_2 \end{pmatrix}\right\} \\ &= \frac{\lambda_2\, y_0 - \dot{y}_0}{\lambda_2 - \lambda_1}\; e^{tA} \cdot \begin{pmatrix} 1 \\ \lambda_1 \end{pmatrix} - \frac{\lambda_1\, y_0 - \dot{y}_0}{\lambda_2 - \lambda_1}\; e^{tA} \cdot \begin{pmatrix} 1 \\ \lambda_2 \end{pmatrix}.\end{aligned}$$

[59]Ogólniej, gdy warunki zadane są w t_0: $y_1(t_0) = y_1^0$, $y_2(t_0) = y_2^0$, rozwiązanie ma postać

$$\begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} = e^{(t-t_0)A} \cdot \begin{pmatrix} y_1^0 \\ y_2^0 \end{pmatrix}.$$

Ponieważ $\exp(tA)$ działa tu na wektory własne macierzy A otrzymujemy

$$\begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} = \frac{\lambda_2\, y_0 - \dot{y}_0}{\lambda_2 - \lambda_1}\; e^{t\lambda_1} \begin{pmatrix} 1 \\ \lambda_1 \end{pmatrix} - \frac{\lambda_1\, y_0 - \dot{y}_0}{\lambda_2 - \lambda_1}\; e^{t\lambda_2} \begin{pmatrix} 1 \\ \lambda_2 \end{pmatrix}.$$

Górna składowa wektora stojącego po prawej stronie daje

$$y_1(t) \equiv y(t) = \frac{1}{\lambda_2 - \lambda_1}\left[(\lambda_2\, y_0 - \dot{y}_0)\, e^{t\lambda_1} - (\lambda_1\, y_0 - \dot{y}_0)\, e^{t\lambda_2}\right],$$

co jest rozwiązaniem znalezionym w pierwszym punkcie poprzedniego zadania. Dolna zaś składowa wektora stojącego po prawej stronie daje po prostu pochodną tego rozwiązania:

$$y_2(t) \equiv \dot{y}(t) = \frac{1}{\lambda_2 - \lambda_1}\left[\lambda_1\,(\lambda_2\, y_0 - \dot{y}_0)\, e^{t\lambda_1} - \lambda_2\,(\lambda_1\, y_0 - \dot{y}_0)\, e^{t\lambda_2}\right].$$

Zauważmy jeszcze, że rozwiązanie to pozostaje słuszne nawet jeśli równanie charakterystyczne ma dwa pierwiastki zespolone: podstawiając tu $\lambda_1 = \lambda$ i $\lambda_2 = \lambda^*$ i przegrupowując wyrazy otrzymujemy od razu rozwiązanie dla tego przypadku wypisane w poprzednim zadaniu. Jest tak dlatego, że - jak już kiedyś zauważyliśmy - macierz $\exp(tA)$ wychodzi rzeczywista nawet wtedy, gdy "po drodze" do jej znalezienia trzeba rozszerzyć przestrzeń wektorową nad ciałem $\mathbb{R}$ do przestrzeni nad ciałem $\mathbb{C}$.

Rozpatrzmy jeszcze przypadek, gdy jest jeden podwójny pierwiastek $\lambda = -\frac{1}{2}a$ równania charakterystycznego, tj. gdy $b = \frac{1}{4}a^2$. Równanie wyznaczające wektor własny ma wtedy postać

$$\begin{pmatrix} -\lambda & 1 \\ -b & -a-\lambda \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \equiv \begin{pmatrix} \frac{1}{2}a & 1 \\ -\frac{1}{4}a^2 & -\frac{1}{2}a \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Jest wtedy tylko jeden wektor własny, jako który można wziąć wektor

$$\begin{pmatrix} 2 \\ -a \end{pmatrix}.$$

Rozkład przestrzeni wektorowej na podprzestrzenie pierwiastkowe jest w tym przypadku trywialny, gdyż cała prestrzeń wektorowa jest po prostu jedną podprzestrzenią pierwiastkową; w związku z tym

$$(A - \lambda I)^2 = \begin{pmatrix} \frac{1}{2}a & 1 \\ -\frac{1}{4}a^2 & -\frac{1}{2}a \end{pmatrix} \begin{pmatrix} \frac{1}{2}a & 1 \\ -\frac{1}{4}a^2 & -\frac{1}{2}a \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

i jako drugi wektor rozpinający tę podprzestrzeń pierwiastkową (czyli całą przestrzeń) możemy wybrać dowolny wektor liniowo niezależny od wektora własnego; np. może być to wektor

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Rozkładamy teraz wektor warunków początkowych na te dwa wektory i znajdujemy rozwiązanie

$$
\begin{aligned}
\begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} &= e^{tA} \cdot \left\{ \frac{1}{2}\, y_0 \begin{pmatrix} 2 \\ -a \end{pmatrix} + \left[\frac{1}{2}\, a y_0 + \dot{y}_0 \right] \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\} \\
&= \frac{1}{2}\, y_0\, e^{\lambda t} \begin{pmatrix} 2 \\ -a \end{pmatrix} + \left(\frac{1}{2}\, a y_0 + \dot{y}_0 \right) e^{\lambda t} \left[I + t(A - \lambda I) \right] \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} .
\end{aligned}
$$

W pierwszym wyrazie macierz $\exp(tA)$ działa na wektor własny macierzy A, co sprowadza się do pomnożenia tego wektora przez $e^{\lambda t}$. W drugim wyrazie zastosowaliśmy stary chwyt polegający na napisaniu

$$
e^{tA} = e^{t\lambda I + t(A-\lambda I)} = e^{t\lambda I} e^{t(A-\lambda I)} = e^{\lambda t} e^{t(A-\lambda I)} = e^{\lambda t} \left\{ I + t(A - \lambda I) + \frac{t^2}{2}(A - \lambda I)^2 + \dots \right\}
$$

i zauważeniu, że $(A-\lambda I)^2$ i wszystkie wyższe potęgi są tu po prostu macierzami zerowymi. Zatem

$$
\begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} = \frac{1}{2}\, y_0\, e^{\lambda t} \begin{pmatrix} 2 \\ -a \end{pmatrix} + \left(\frac{1}{2}\, a y_0 + \dot{y}_0 \right) e^{\lambda t} \left[\begin{pmatrix} 0 \\ 1 \end{pmatrix} + t \begin{pmatrix} 1 \\ -a/2 \end{pmatrix} \right]
$$

i górna składowa daje uzyskane już w poprzednim zadaniu rozwiązanie, a dolna jego pochodną.

Zadanie Ode.15
Znaleźć ogólne rozwiązanie liniowego równania różniczkowego drugiego rzędu o stałych współczynnikach z niejednorodnością

$$
\frac{d^2}{dt^2}\, y + 2\gamma\, \frac{d}{dt}\, y + \omega_0^2 y = f(t)\,,
$$

w przypadku, gdy $\omega_0^2 > \gamma^2$ (zgodnie z tradycją fizyczną stałe a i b z porzednich zadań oznaczyliśmy tu odpowiednio 2γ i ω_0^2).
Rozwiązanie: Rozwiązujemy najpierw równanie jednorodne

$$
\frac{d^2}{dt^2}\, y + 2\gamma\, \frac{d}{dt}\, y + \omega_0^2 y = 0\,,
$$

podstawiając doń (tym razem dajmy upust naszym fizykalnym nawykom) $y(t) = A\exp(i\Omega t)$. Dostajemy na Ω równanie kwadratowe

$$
-\Omega^2 + 2i\gamma\,\Omega + \omega_0^2 = 0\,,
$$

które, gdy $\omega_0^2 > \gamma^2$, ma dwa różne pierwiastki

$$
\Omega = i\gamma \pm \sqrt{\omega_0^2 - \gamma^2} \equiv i\gamma \pm \omega\,.
$$

Ogólnym rozwiązaniem jest więc

$$y_{\rm hom}(t) = \left(Z\, e^{i\omega t} + Z^*\, e^{-i\omega t}\right) e^{-\gamma t}\,.$$

w którym $Z \in \mathbb{C}$ jest zespoloną stałą całkowania (czyli są to dwie rzeczywiste stałe dowolne) i które wygodnie będzie przepisać w postaci

$$y_{\rm hom}(t) = C_1\, y_1(t) + C_2\, y_2(t)\,,$$

czyli w postaci kombinacji liniowej z rzeczywistymi współczynnikami C_1 i C_2 dwóch liniowo niezależnych rzeczywistych rozwiązań

$$y_1(t) = e^{-\gamma t}\sin\omega t\,, \qquad y_2(t) = e^{-\gamma t}\cos\omega t\,,$$

równania jednorodnego

Szukamy następnie jak zwykle jakiegokolwiek (tzw. rozwiązania szczególnego) rozwiązania równania niejednorodnego. Możemy próbować znaleźć je metodą uzmiennienia stałych całkowania w rozwiązaniu równania jednorodnego. Piszemy więc

$$y_{\rm inhom}(t) = A_1(t)\, y_1(t) + A_2(t)\, y_2(t)\,.$$

Rozwiązanie równania jednorodnego zależy od dwu stałych, więc uzmienniliśmy tu je obie. Po wstawieniu tego, jak to się nazywa, Ansatzu do równania otrzymujemy

$$\begin{aligned} &A_1''\, y_1 + 2A_1'\, y_1' + A_1\, y_1'' + 2\gamma\,(A_1'\, y_1 + A_1\, y_1') + \omega_0^2 A_1\, y_1 \\ +&A_2''\, y_2 + 2A_2'\, y_2' + A_2\, y_2'' + 2\gamma\,(A_2'\, y_2 + A_2\, y_2') + \omega_0^2 A_2\, y_2 = f(t)\,. \end{aligned}$$

Jak zwykle część wyrazów wypada, bo $y_1(t)$ i $y_2(t)$ spełniają równanie jednorodne i mamy

$$\begin{aligned} &A_1''\, y_1 + 2A_1'\, y_1' + 2\gamma A_1'\, y_1 \\ +&A_2''\, y_2 + 2A_2'\, y_2' + 2\gamma A_2'\, y_2 = f(t)\,. \end{aligned}$$

Nie wygląda to jednak wesoło, bo mamy jedno równanie (i to drugiego rzędu) na dwie nieznane funkcje $A_1(t)$ i $A_2(t)$. Można by wprawdzie powiedzieć, że skoro szukamy jakiegokolwiek rozwiązania równania niejednorodnego, to możemy np. przyjąć, że $A_2 \equiv 0$, ale to dałoby nam na $A_1(t)$ równanie niejednorodne drugiego rzędu (i to gorsze niż to, którego rozwiązania właśnie szukamy, bo o współczynnikach będących funkcjami t). Sztuczka polega na tym, by na szukane funkcje $A_1(t)$ i $A_2(t)$ narzucić dodatkowy warunek

$$A_1'\, y_1 + A_2'\, y_2 = 0\,,$$

którego konsekwencją jest też związek

$$A_1''\, y_1 + A_1'\, y_1' + A_2''\, y_2 + A_2'\, y_2' = 0\,.$$

Równanie niejednorodne przybiera wtedy postać

$$A_1'\, y_1' + A_2'\, y_2' = f(t)\,.$$

Poza tym, narzucony związek pozwala wtedy wyrazić np. A_2' przez A_1': $A_2' = -(y_1/y_2)A_2'$ lub $A_1' = -(y_2/y_1)A_2'$. Otrzymujemy wtedy na $A_1(t)$ i $A_2(t)$ równania

$$A_1' = -\frac{y_2}{y_1y_2' - y_1'y_2}\, f(t)\,, \qquad A_2' = \frac{y_1}{y_1y_2' - y_1'y_2}\, f(t)\,.$$

Mianownik $W(y_1, y_2) = y_1y_2' - y_1'y_2$ nazywa się wrońskianem.[60] Jest on w tym przypadku równy

$$W(y_1, y_2) = y_1y_2' - y_1'y_2 = -\omega\, e^{-2\gamma t}\,.$$

Zatem

$$A_1' = \frac{1}{\omega}\, f(t)\, e^{\gamma t} \cos\omega t\,, \qquad A_2' = -\frac{1}{\omega}\, f(t)\, e^{\gamma t} \sin\omega t\,.$$

Najogólniejsze rozwiązanie równania niejednorodnego ma zatem postać

$$\begin{aligned} y(t) = C_1\, e^{-\gamma t} \sin\omega t + C_2\, e^{-\gamma t} \cos\omega t + \frac{1}{\omega}\, e^{-\gamma t} \sin\omega t \int^t dt'\, f(t')\, e^{\gamma t'} \cos\omega t' \\ -\frac{1}{\omega}\, e^{-\gamma t} \cos\omega t \int^t dt'\, f(t')\, e^{\gamma t'} \sin\omega t'\,. \end{aligned}$$

Dolne granice całek mogą być dowolne - wpływają one tylko na redefinicje dowolnych stałych C_1 i C_2 i razem z nimi są wyznaczane przez warunki początkowe.

Zadanie Ode.16
Otrzymać bez sztuczek z Wrońskianami rozwiązanie niejednorodnego liniowego równania różniczkowego drugiego rzędu o stałych espółczynnikach przerabiając je na liniowy układ dwóch równań pierwszego rzędu z niejednorodnością.
Rozwiązanie: Jest ono bardzo podobne do tego, co zostało zrobione w Zadaniu Ode.10. Równanie

$$\frac{d^2}{dt^2}\, y + a\, \frac{d}{dt}\, y + b\, y = f(t)\,,$$

przepisujemy w formie

$$\frac{d}{dt}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -b & -a \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \begin{pmatrix} 0 \\ f(t) \end{pmatrix},$$

wprowadzając oznaczenia $y_1 = y$, $y_2 = y_1'$. Równanie jednorodne zostało już rozwiązane w Zadaniu Ode.14. Tu założymy, że stojąca tu macierz F ma dwie różne wartości własne λ_1

[60] Józef Maria Hoene-Wroński (1776-1853) - polski fizyk, matematyk i filozof. Jeden z przedstawicieli polskiego mesjanizmu. Pamiętamy: Mickiewicz, Towiański i te sprawy - zob. Görgy Spiro "Mesjasze". Choć, jak twierdzi Miłosz (w "Ziemi Ulro"), Hoene-Wroński "wierszoklety i jego mistycznej bandy" nie znosił...

i λ_2 i, co za tym idzie, dwa wektory własne. Potrzebną macierz $\exp(tF)$ możemy znanym sposobem odczytać działając nią na wypisany już w zadaniu Ode.14 rozkład

$$\begin{pmatrix} a \\ b \end{pmatrix} = \frac{a\lambda_1 - b}{\lambda_2 - \lambda_1} \begin{pmatrix} 1 \\ \lambda_1 \end{pmatrix} - \frac{a\lambda_2 - b}{\lambda_2 - \lambda_1} \begin{pmatrix} 1 \\ \lambda_2 \end{pmatrix},$$

dowolnego wektora na wektory własne. Otrzymujemy w ten sposób

$$e^{tF} = \frac{1}{\lambda_2 - \lambda_1} \begin{pmatrix} \lambda_2\, e^{t\lambda_1} - \lambda_1\, e^{t\lambda_2} & e^{t\lambda_2} - e^{t\lambda_1} \\ \lambda_1\lambda_2(e^{t\lambda_1} - e^{t\lambda_2}) & \lambda_2\, e^{t\lambda_2} - \lambda_1\, e^{t\lambda_1} \end{pmatrix}.$$

Jak zawsze pełne rozwiązanie jest sumą ogólnego rozwiązania $\mathbf{y}_{\mathrm{hom}}$ równania jednorodnego i jakiegoś rozwiązania równania niejednorodnego, które konstruujemy podstawiając do równania niejednorodnego Ansatz $\mathbf{y}_{\mathrm{inhom}} = \exp(tF) \cdot \mathbf{h}(t)$. Daje to na $\mathbf{h}$ równanie:

$$\begin{pmatrix} h_1 \\ h_2 \end{pmatrix} = \int dt\, e^{-tF} \cdot \begin{pmatrix} 0 \\ f(t) \end{pmatrix}.$$

Po jawnym zadziałaniu macierzą $\exp(-tF)$ da to

$$h_1(t) = \frac{1}{\lambda_2 - \lambda_1} \int dt \left(e^{-t\lambda_2} - e^{-t\lambda_1}\right) f(t)\,,$$
$$h_2(t) = \frac{1}{\lambda_2 - \lambda_1} \int dt \left(\lambda_2 e^{-t\lambda_2} - \lambda_1 e^{-t\lambda_1}\right) f(t)\,.$$

Otrzymujemy więc wektor $\mathbf{y}_{\mathrm{inhom}}(t)$ w postaci

$$\frac{1}{(\lambda_2 - \lambda_1)^2} \begin{pmatrix} \lambda_2\, e^{t\lambda_1} - \lambda_1\, e^{t\lambda_2} & e^{t\lambda_2} - e^{t\lambda_1} \\ \lambda_1\lambda_2(e^{t\lambda_1} - e^{t\lambda_2}) & \lambda_2\, e^{t\lambda_2} - \lambda_1\, e^{t\lambda_1} \end{pmatrix} \int^t dt' \begin{pmatrix} (e^{-t'\lambda_2} - e^{-t'\lambda_1}) f(t') \\ (\lambda_2 e^{-t'\lambda_2} - \lambda_1 e^{-t'\lambda_1}) f(t') \end{pmatrix}.$$

Wygląda trochę zawile, ale trzeba cierpliwie wypisać jawnie pierwsze pięterko tego wektora (tylko pierwsze jest naprawdę potrzebne, bo to składowa y_1 wektora $\mathbf{y} = \mathbf{y}_{\mathrm{hom}} + \mathbf{y}_{\mathrm{inhom}}$ jest szukaną funkcją $y(t)$; składowa y_2 powinna dać pochodną szukanej funkcji $y(t)$, czyli powinien zachodzić związek $y_2 = y_1'$, który może być użyty do sprawdzenia poprawności rachunków)

$$y_1^{\mathrm{inhom}} = \frac{1}{(\lambda_2 - \lambda_1)^2} \left\{ \left(\lambda_2\, e^{t\lambda_1} - \lambda_1\, e^{t\lambda_2}\right) \int^t dt' \left(e^{-t'\lambda_2} - e^{-t'\lambda_1}\right) f(t') \right.$$
$$\left. + \left(e^{t\lambda_2} - e^{t\lambda_1}\right) \int^t dt' \left(\lambda_2\, e^{-t'\lambda_2} - \lambda_1\, e^{-t'\lambda_1}\right) f(t') \right\}.$$

Z ośmiu wyrazów cztery się parami redukują i zostaje

$$y_1^{\mathrm{inhom}} = \frac{1}{(\lambda_2 - \lambda_1)^2} \left\{ -\lambda_2\, e^{t\lambda_1} \int^t dt'\, e^{-t'\lambda_1} f(t') - \lambda_1\, e^{t\lambda_2} \int^t dt'\, e^{-t'\lambda_2} f(t') \right.$$
$$\left. + \lambda_2\, e^{t\lambda_2} \int^t dt'\, e^{-t'\lambda_2} f(t') + \lambda_1\, e^{t\lambda_1} \int^t dt'\, e^{-t'\lambda_1} f(t') \right\},$$

a po zebraniu razem takich samych całek

$$y_{\rm inhom} = \frac{1}{\lambda_2 - \lambda_1} \left\{ -e^{t\lambda_1} \int^t dt'\, e^{-t'\lambda_1}\, f(t') + e^{t\lambda_2} \int^t dt'\, e^{-t'\lambda_2}\, f(t') \right\}.$$

To zaś jest tym samym, co już było znalezione jako $y_{\rm inhom}$ w zadaniu Ode.15, tylko trochę ogólniej zapisane. Istotnie, w tamtym Zadaniu założyliśmy, że równanie charakterystyczne ma dwa wzajemnie sprzężone pierwiastki zespolone. W notacji tego zadania odpowiadałoby to położeniu $\lambda_1 = -\gamma + i\omega$, i $\lambda_2 = -\gamma - i\omega$ (gdzie $\omega = \sqrt{\omega_0^2 - \gamma^2}$). W takim przypadku otrzymane tu rozwiązanie ma postać

$$y_{\rm inhom} = \frac{i}{2\omega} \left\{ -e^{-\gamma t + i\omega t} \int^t dt'\, e^{\gamma t' - i\omega t'}\, f(t') + e^{-\gamma t - i\omega t} \int^t dt'\, e^{\gamma t' + i\omega t'}\, f(t') \right\}.$$

Po napisaniu $e^{\pm i\omega t} = \cos\omega t \pm i \sin\omega t$ połowa wyrazów się zredukuje i zostanie dokładnie to, co jako $y_{\rm inhom}$ zostało otrzymane w Zadaniu Ode.15.

Zadanie Ode.17
Znaleźć jakieś rozwiązanie $y_{\rm inhom}(t)$ równania

$$y'' + a\, y' + b\, y = f \cos(\Omega t + \delta)\,,$$

o rzeczywistych współczynnikach a i b.
Rozwiązanie: W przypadku niejednorodności, która ma postać funkcji sinus lub kosinus, albo ogólnie kombinacji eksponensów (ewentualnie mnożonych przez wielomian) zamiast wykorzystywać ogólne wzory wyprowadzone w poprzednich zadaniach, prościej jest zastąpić powyższe równanie równaniem

$$z'' + a\, z' + b\, z = \hat{f}\, e^{i\Omega t}\,,$$

w którym z może przyjmować wartości zespolone, a $\hat{f} = f\, e^{i\delta}$. Powinno być jasne, że część rzeczywista ${\rm Re}(z_{\rm inhom})$ rozwiązania tego równania będzie spełniać wyjściowe równanie (z $f \cos(\Omega t + \delta)$ jako niejednorodnością), a część urojona ${\rm Im}(z_{\rm inhom})$ rozwiązania będzie spełniać wyjściowe równanie z $f \sin(\Omega t + \delta)$ jako niejednorodnością.

Aby znaleźć rozwiązanie $z_{\rm inhom}$ wystarczy do równania podstawić jako Ansatz $z_{\rm inhom}(t) = \mathbb{A}\, e^{i\Omega t}$. Czynniki $e^{i\Omega t}$ wtedy wypadają i zostaje równanie algebraiczne na $\mathbb{A}$, którego rozwiązaniem jest

$$\mathbb{A} = \frac{\hat{f}}{-\Omega^2 + ia\,\Omega + b} = \frac{\hat{f}}{(\Omega^2 - b)^2 + a^2\Omega^2}\,(b - \Omega^2 - ia\,\Omega)\,.$$

Zatem

$$\begin{aligned} z_{\rm inhom}(t) = \frac{f}{(\Omega^2 - b)^2 + a^2\Omega^2} & \big\{ \big[(b - \Omega^2) \cos(\Omega t + \delta) + a\,\Omega \sin(\Omega t + \delta)\big] \\ & + i \big[(b - \Omega^2) \sin(\Omega t + \delta) - a\,\Omega \cos(\Omega t + \delta)\big] \big\}\,. \end{aligned}$$

I teraz można sobie wziąć część rzeczywistą albo urojoną - co tam komu potrzebne...

Zadania do samodzielnej zabawy

Zadanie 1.
Sprawdzić, że metryka dyskretna jest rzeczywiście metryką (tj. że spełnia ona konieczne warunki).

Zadanie 2.
Zbadać jak wygląda w metryce dyskretnej kula otwarta $K(x_0, r)$ i kula-bar zdefiniowana ogólnie (na razie bez związku z domknięciem) wzorem $\bar{K}(x_0, r) = \{x \in \mathcal{X} | \ d(x_0, x) \leq r\}$, w zależności od wartości jej promienia $r \in [0, \infty)$.

Zadanie 3.
Niech $\mathcal{X} = \mathbb{R}^1$ z metryką Euklidesową, która jest tu tym samym, co metryka taxi $d(x, y) = |x - y|$. Niech $A = \{x = 1/n, \ n \in \mathbb{N}\}$ (tu przyjmujemy, że zero nie jest liczbą naturalną). Znaleźć wszystkie punkty izolowane zbioru A, wszystkie punkty skupienia (i w przypadku każdego orzec, czy należy on do A, czy nie) i punkty wewnętrzne. Orzec także czy zbiór A jest otwarty oraz czy jest on domknięty.

Zadanie 4.
Zbadać jak w $(\mathbb{R}^2, d_\infty)$ i w $(\mathbb{R}^2, d_{\text{taxi}})$ wyglądają kule otwarte $K(x_0, r)$.

Zadanie 5.
Jaka jest odległość między należącymi do $\mathcal{C}[0, \pi]$ funkcjami $f = \sin x$ i $g = \cos x$ w metryce d_1 i w metryce d_2?

Zadanie 6.
Niech $(\mathcal{X}, d)$ będzie przestrzenią metryczną. Wykazać, że jeśli $x' \in K(x, r)$, gdzie $x \in \mathcal{X}$, a $r > 0$, to $r' \equiv r - d(x, x') > 0$ i $K(x', r') \subset K(x, r)$.

Zadanie 7.
Niech $(\mathcal{X}, d)$ będzie przestrzenią metryczną. Udowodnić (od razu strach w oczach! - to najpierw trzy zdrowaśki odmówić, a potem jeszcze raz spokojnie przeczytać...) korzystając z poprzedniego zadania że każda kula $K(x, r)$ w $\mathcal{X}$ (tak jak ją zdefiniowaliśmy na stronie 4 tych notatek) jest zbiorem otwartym (też w sensie zdefiniowanym tuż pod definicja kuli).

Zadanie 8.
Uzasadnić, że jeśli O_s, $s \in S$ (S jest pewnym zbiorem indeksów, niekoniecznie nawet przeliczalnym), jest rodziną zbiorów otwartych, to

$$O = \bigcup_{s \in S} O_s \,,$$

jest też zbiorem otwartym. Pokazać też, że przecięcie skończonej liczby zbiorów otwartuch jest zbiorem otwartym i podać jakiś prosty przykład przecięcia nieskończonej liczby zbiorów otwartych, które nie jest zbiorem otwartym

Zadanie 9.
Zbadać granice w punkcie $(0,0)$ funkcji $\mathbb{R}^2 \to \mathbb{R}$ zadanych wzorami

$$\begin{aligned}
&a) \quad f(x,y) = \frac{x^2-y^2}{x^2+y^2}\,,\\
&b) \quad f(x,y) = \frac{x^3+y^3}{x^2+y^2}\,,\\
&c) \quad f(x,y) = \frac{x^2y^2}{x^2+y^2}\,,\\
&d) \quad f(x,y) = \frac{2x^2+3y^2+x^3+y^3}{x^2+y^2}\,,\\
&e) \quad f(x,y) = (x^2+y^2)\ln(x^2+y^2)\,,\\
&f) \quad f(x,y) = \frac{1}{x^2+y^2}\,,\\
&g) \quad f(x,y) = \frac{x^2y^2}{x^2y^2+(x-y)^2}\,,\\
&h) \quad f(x,y) = x\sin\frac{1}{y}\,,
\end{aligned}$$

(w ostatnim przykładzie chodzi o zbadanie granicy funkcji w punktach $(x_0,0)$).

Zadanie 10.
Niech $(V_1, \|\cdot\|_1)$ i $(V_2, \|\cdot\|_2)$ będą dwiema przestrzeniami wektorowymi i niech $F: V_1 \to V_2$ będzie odwzorowaniem liniowym. Pokazać, że jeśli odwzorowanie F jest ciągłe (w metrykach d_1 i d_2 zadawanych przez normy $\|\cdot\|_1$ i $\|\cdot\|_2$) w jednym punkcie-wektorze przestrzeni V_1, to jest ciągłe wszędzie.
Wskazówka: Zastosować rozumowanie *ad absurdum*.

Zadanie 11.
Obliczyć[61] wszystkie pochodne cząstkowe pierwszego rzędu i drugiego (sprawdzając przy tym, że mieszane pochodne drugiego rzędu są takie same, tzn. że np. $\partial^2 f/\partial x\partial y = \partial^2 f/\partial y\partial x$) poniższych funkcji $\mathbb{R}^n \to \mathbb{R}$ ($n = 2$ lub 3). Jak ktoś chce się bardziej ćwiczyć, może i pochodne trzeciego rzędu poobliczać.

$$\begin{aligned}
&a) \quad f(x,y) = x^2 - y + 3y^2 + x^3y^3 - x\sin y\,,\\
&b) \quad f(x,y) = \frac{xy}{y-1}\,,\\
&c) \quad f(x,y) = \operatorname{arctg}\left(\frac{y}{x}\right),\\
&d) \quad f(x,y,z) = (3x^2y+z)^n\,,
\end{aligned}$$

[61]Zadanie jest idiotyczne, ale spełniać ma ono tę samą rolę, co codzienne rypanie na pianinie znanego kawałka "kurki trzy" przez początkujących pianistów (zaawansowani też to rypią na rozgrzewkę) - ręka musi się ułożyć i rozruszać.

$$
\begin{array}{ll}
e) & f(x,y) = \ln(x^2+y^2)\,, \\
f) & f(x,y,z) = \sqrt{x^2+y^2+z^2}\,, \\
g) & f(x,y,z) = \dfrac{1}{\sqrt{x^2+y^2+z^2}}\,, \quad \text{oblicz} \quad f_{xx}+f_{yy}+f_{zz} \\
h) & f(u,v) = \ln(u+\ln v)\,, \\
i) & f(x,y,z) = \ln\left(\dfrac{a-\sqrt{x^2+y^2+z^2}}{a+\sqrt{x^2+y^2+z^2}}\right), \\
j) & f(x,y) = x^y\,, \\
k) & f(x,y) = (\ln x)^{\sin y}\,, \\
l) & f(x,y,z) = (x\,\mathrm{tg} z)^{\ln y}\,, \\
m) & f(x,y,z) = x^{y^z}\,, \\
n) & f(x,y) = \displaystyle\int_a^{xy} dt\, g(t)\,.
\end{array}
$$

Zadanie 12.
Obliczyć bezpośrednio z definicji (tj. jako granicę odpowiedniego ilorazu różnicowego)
a) pochodne cząstkowe f_x i f_y w punkcie $(3,4)$ funkcji

$$f(x,y) = x+y-\sqrt{x^2+y^2}\,,$$

b) pochodne cząstkowe f_x i f_y w punkcie $(1,2)$ funkcji

$$f(x,y) = \ln\left(x+\frac{y}{2x}\right),$$

c) pochodną kierunkową w punkcie $(2,1)$ w kierunku wektora $\mathbf{n} = (1,\,3)/\sqrt{10}$ funkcji

$$f(x,y) = x^2+2y^2,$$

i sprawdzić, że jest ona równa odpowiedniej kombinacji pochodnych cząstkowych funkcji f obliczonych w tym samym punkcie,
d) pochodną kierunkową w punkcie $(1,1)$ w kierunku wektora $\mathbf{n} = (1,\,-3)/\sqrt{10}$ funkcji

$$f(x,y) = \sin\left(\frac{\pi}{2}\,(x+y)\right),$$

i sprawdzić, że jest ona równa odpowiedniej kombinacji pochodnych cząstkowych funkcji f obliczonych w tym samym punkcie.

Zadanie 13.
Pokazać, że funkcje $f:\mathbb{R}^2 \to \mathbb{R}$

$$a) \qquad f(x,y) = \begin{cases} x^2y/(x^2+y^2) & \text{gdy } (x,y) \neq (0,0) \\ 0 & \text{gdy } (x,y) = (0,0) \end{cases},$$

$$b) \qquad f(x,y) = \begin{cases} x^3y/(x^4+y^2) & \text{gdy } (x,y) \neq (0,0) \\ 0 & \text{gdy } (x,y) = (0,0) \end{cases},$$

choć są w punkcie $(0,0)$ ciągłe (pokazać to) i mają w tym punkcie pochodne kierunkowe w każdym kierunku (też pokazać), nie są tam nieróżniczkowalne, tzn. df nie przybliża należycie przyrostu Δf tych funkcji.

Zadanie 14.
a) Znaleźć pochodną fukcji $f : \mathbb{R}^2 \to \mathbb{R}$ zadanej wzorem

$$f(x,y) = \frac{x}{1+y^2}\,,$$

w punkcie $(x,y) = (2,1)$ i obliczyć jej wartość na wektorze przesunięcia $(dx,dy) = (\delta\lambda,\, \frac{1}{2}\delta\lambda)$, oraz na wektorze przesunięcia $(dx,dy) = (\delta\lambda,\, \delta\lambda)$,
b) W jakim kierunku w punkcie $(1,1)$ funkcja $f : \mathbb{R}^2 \to \mathbb{R}$ zadana wzorem

$$f(x,y) = 2x^2 + 3y^2\,,$$

rośnie najszybciej?
c) Obliczyć z definicji w dowolnym punkcie (x,y) dziedziny odwzorowania $F : \mathbb{R}^2_+ \longrightarrow \mathbb{R}^3$ danego wzorem

$$F\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x\sin(x+2y) \\ y^2\cos(2x-y) \\ y^{2x} \end{bmatrix},$$

pochodną kierunkową w kierunku wektora $\mathbf{n} = (1,\,1)/\sqrt{2}$. Sprawdzić, że pochodna kierunkowa obliczona w ten sposób jest taka sama, jak otrzymana z kombinacji liniowej pochodnych cząstkowych.
d) Znaleźć pochodną fukcji $F : \mathbb{R}^2 \to \mathbb{R}^2$ zadanej wzorem

$$F\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x^2 - y^2 \\ \exp(x-y) \end{bmatrix},$$

w punkcie $(x,y) = (1,-1)$ i obliczyć jej wartość na wektorze przesunięcia $(dx,dy) = (-\delta\lambda,\, 2\delta\lambda)$.
e) Znaleźć pochodną fukcji $F : \mathbb{R}^3 \to \mathbb{R}^2$ zadanej wzorem

$$F\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} \sqrt{x^2+z^2} \\ \ln(1+y^2+z^2) \end{bmatrix},$$

w punkcie $(x,y) = (1,3,0)$ i obliczyć jej wartość na wektorze przesunięcia $(dx,dy,dz) = (0,\,\delta\lambda,\,\delta\lambda)$.

Zadanie 15. (Coś bardziej praktycznego)
Podać równanie hiperpłaszczyzny w $\mathbb{R}^4$ (dokładniej to w $A\mathbb{R}^4$ - czterowymiarowej przestrzeni afinicznej - zob. mój skrypt do algebry) stycznej w punkcie $(1,2,3)$ do "wykresu" funkcji $f : \mathbb{R}^3 \longrightarrow \mathbb{R}$ danej wzorem $f(x,y,z) = x^2 + y^3 + z^2$.

Wskazówka: Najpierw pomyśleć jak znajdujemy równanie prostej w $\mathbb{R}^2$ stycznej w jakimś konkretnym x_0 do wykresu funkcji $y = f(x)$, a potem spróbować to uogólnić na równanie płaszczyzny w $\mathbb{R}^3$ stycznej w jakimś konkretnym (x_0, y_0) do powierzchni funkcji $z = f(x, y)$ (kiedy jeszcze możemy sobie to wyobrazić), a na koniec przenieść metodę na problem w czterech wymiarach.

Zadanie 16.
Pokazać, że funkcja $f(x, y)$ postaci

$$f(x, y) = \phi\left(\frac{y}{x}\right) - x^2 - y^2\,,$$

spełnia równanie różniczkowe

$$x\,\frac{\partial f}{\partial x} + y\,\frac{\partial f}{\partial y} = g(x, y)\,,$$

w którym $g(x, y)$ jest pewną funkcją x i y, której postać należy podać.

Zadanie 17.
Przepisać dwuwymiarowy laplasjan

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}\,,$$

w zmiennych u i v zadanych związkami

$$x = \frac{u}{u^2 + v^2} \equiv \frac{u}{R^2}\,, \qquad y = \frac{v}{u^2 + v^2} \equiv \frac{v}{R^2}\,.$$

Zadanie 18.
Stosując bezpośrednio regułkę różniczkowania łańcuszkowego, sprawdzić bezpośrednim rachunkiem, że wyrażone przez nowe zmienne $u = u(x, y)$ i $v = v(x, y)$ pochodne mieszane

$$\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right), \qquad \text{oraz} \qquad \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right),$$

są dane tymi samymi wyrażeniami.[62] Obliczyć jawnie tę pochodna w przypadku, gdy dane są wzory $x = x(u, v)$ i $y = y(u, v)$ i są one takie same jak w poprzednim zadaniu.

Zadanie 19.
Pokazać, że funkcja dwóch zmiennych $f(x, y)$ dana wzorem

$$f(x, y) = y\,\phi(x^2 - y^2),$$

[62] Potraktować to zadanie jak wprawkę w różniczkowaniu łańcuszkowym. Zazwyczaj studenci mają z tym jakieś klopoty.

spełnia równanie różniczkowe

$$\frac{1}{x}\frac{\partial f}{\partial x}+\frac{1}{y}\frac{\partial f}{\partial y}-g(x,y)\,f=0\,,$$

w którym $g(x,y)$ jest pewną konkretną funkcją, której postać należy podać.

Zadanie 20.
Pokazać, że funkcja trzech zmiennych $H(x,y,z)$ dana wzorem

$$H(x,y,z)=x^2\,F\Big(\frac{z}{x},\,\frac{y}{x}\Big)\,,$$

spełnia równanie różniczkowe

$$x\,\frac{\partial H}{\partial x}+y\,\frac{\partial H}{\partial y}+z\,\frac{\partial H}{\partial z}-g(x,y,z)\,H=0\,,$$

w którym $g(x,y,z)$ jest pewną konkretną funkcją, której postać należy podać.

Zadanie 21.
Rozwinąć w szereg Taylora do wyrazów trzeciego rzędu włącznie wokół punktu (x_0,y_0) funkcje

$$\begin{array}{lll}
a) & f(x,y)=\sin(x+y)\,, & (x_0,y_0)=(0,0)\,,\\
b) & f(x,y)=e^{x^2}\cos y\,, & (x_0,y_0)=(0,0)\,,\\
c) & f(x,y)=\ln(1+x+2y)\,, & (x_0,y_0)=(0,0)\,,\\
d) & f(x,y)=x+y+\sqrt{x^2+y^2}\,, & (x_0,y_0)=(1,0)\,,\\
e) & f(x,y)=x^y\,, & (x_0,y_0)=(1,0)\,.
\end{array}$$

Otrzymać rozwinięcie raz wykorzystując znane rozwinięcia funkcji jednej zmiennej i drugi raz obliczając pochodne cząstkowe.

Zadanie 22.
Znaleźć punkty krytyczne następujących funkcji dwu zmiennych (dziedzina niektórych z nich jest w oczywisty sposób trochę mniejsza niż $\mathbb{R}^2$).

$$\begin{array}{ll}
a) & f(x,y)=x^2+y^2-\lambda xy\,,\\
b) & f(x,y)=3x^3+3x^2y-y^3-15x\,,\\
c) & f(x,y)=x^2+xy+y^2+\dfrac{a^3}{x}+\dfrac{a^3}{y}\,,\quad a>0\,,\\
d) & f(x,y)=(x+y)^4+(x-y)^6\,,\\
e) & f(x,y)=x-2y-3\,\mathrm{arctg}\dfrac{x}{y}+\ln\sqrt{x^2+y^2}\,,\\
f) & f(x,y)=\sin(x+y)-\sin x-\sin y\,,\\
g) & f(x,y)=x^4-y^4-4xy^2-2x^2\,,
\end{array}$$

i zbadać ich charakter. Jeśli funkcja zależy od rzeczywistego parametru λ, to zbadać punkty krytyczne w zależności od jego wartości.

Zadanie 23.
Pokazać, że funkcja

$$f(x,y) = 3x^4 - 4x^2y + y^2\,,$$

nie ma w punkcie $(0,0)$ minimum jako funkcja dwóch zmiennych, ale ma w nim minumum, jeśli jest ograniczona do dowolnej prostej przechodzącej przez ten punkt.

Zadanie 24.
Znaleźć największą i najmniejszą wartość funkcji

$$f(x,y) = x^2y - 8x - 4y\,,$$

w trójkącie (z brzegami) o wierzchołkach w punktach $(0,0)$, $(4,0)$ i $(0,4)$.

Zadanie 25. (z nieśmiertelnego Krysickiego-Włodarskiego)
Sprawdzić, czy wypisane niżej odwzorowania $F : \mathbb{R}^2 \longrightarrow \mathbb{R}$ wyznaczają, przez warunek $F(x,y) = 0$, funkcję $y = y(x)$ w otoczeniu podanych punktów:

$$\begin{aligned}
&a) \quad F(x,y) = -1 - xy + \mathrm{tg}(x+y)\,, \qquad (0,\frac{\pi}{4})\,,\\
&b) \quad F(x,y) = -1 + y - 3xy + 2\sin x + \sin y\,, \qquad (\frac{\pi}{6},0)\,,\\
&c) \quad F(x,y) = \frac{x}{y} - \frac{4}{x} - y + 1\,, \qquad (2,1)\,,\\
&d) \quad F(x,y) = \frac{x+2y}{x-1} + 3xy - 2\,, \qquad (2,0)\,,
\end{aligned}$$

W każdym z tych przyypadków, jeśli funkcja $y = y(x)$ istnieje, obliczyć jej pochodne y' i y'' w podanym punkcie.

Zadanie 26.
W jakich punktach płaszczyzny xy warunek $F(x,y) = 0$ *nie* wyznacza funkcji $y = y(x)$ i/lub funkcji $x = x(y)$? W punktach, w których wyznacza funkcję $y = y(x)$ obliczyć jej pochodne y' i y''

$$\begin{aligned}
&a) \quad F(x,y) = x^{2/3} + y^{2/3} - a^{2/3}\,,\\
&b) \quad F(x,y) = x^2 - y^2 + 3xy - 1\,.
\end{aligned}$$

Spróbować narysować sobie zbiór $E = F^{-1}(0) \subset \mathbb{R}^2$.

Zadanie 27.
Znaleźć punkty krytyczne (i zbadać ich charakter, tj. powiedzieć, czy w punktach tych jest minimum, maksimum, czy punkt przegięcia) funkcji $y = y(x)$ zadanych w sposób

uwikłany warunkami

$$
\begin{aligned}
&a) \quad F(x,y) = x^2 - 2x - 2y + y^2 + 1 = 0\,,\\
&b) \quad F(x,y) = y^3 + 2xy + x^2 = 0\,,
\end{aligned}
$$

Jeśli można, napisać te funkcje jawnie i sprawdzć otrzymane wnioski "szkolnym" sposobem.

Zadanie 28.
Czy odwzorowanie $\mathbb{R}^3$ w $\mathbb{R}$ dane wzorem

$$
F(x,y,z) = z^3 - xyz - 2\,,
$$

zadaje w otoczeniu punktu $(1,3,2)$ funkcję $z = z(x,y)$? Jeśli zadaje, to obliczyć wartość w tym punkcie wszystkich pierwszych i drugich pochodnych cząstkowych tej funkcji.

Zadanie 29. ("cieciurzynka")
Pokazać, że jeśli $\phi : \mathbb{R} \longrightarrow \mathbb{R}$ jest funkcją różniczkowalną w otoczeniu zera i jej pochodna jest tam ciągła, to warunek $F(x,y,z) = 0$, w którym odwzorowanie $F : \mathbb{R}^3 \longrightarrow \mathbb{R}$ jest dane wzorem

$$
F(x,y,z) = \phi\big(x\,e^{z} - y\,e^{-z}\big) - z\,,
$$

definiuje w otoczeniu punktu $(0,0) \in \mathbb{R}^2$ pewną funkcję $z : \mathbb{R}^2 \longrightarrow \mathbb{R}$, spełniającą równanie różniczkowe cząstkowe

$$
\frac{\partial z}{\partial x} + g(z)\,\frac{\partial z}{\partial y} = 0\,,
$$

w którym $g(z)$ jest pewną funkcją z, którą należy podać.

Zadanie 30. (znów "cieciurzynki")
Znaleźć punkty krytyczne funkcji dwóch zmiennych $z = z(x,y)$ zadanej w sposób uwikłany warunkiem $F(x,y,z) = 0$

$$
\begin{aligned}
&a) \quad F(x,y,z) = 6z^3 - 7(x^3 - 3x)z + (2x + y)^2 - 20\,,\\
&b) \quad F(x,y,z) = z^3 + z + \frac{14xz}{1 + x^2} + (2x - y)^2 + 9\,,\\
&c) \quad F(x,y,z) = 3z^3 - 7z\cos(x + y) + \frac{20x}{1 + x^2}\,,
\end{aligned}
$$

i zbadać ich charakter (minimum lokalne, maksimum lokalne funkcji, jej punkt siodłowy?).

Zadanie 31
Niech $F^1(x,y,z,t) = y^2 + t^2 - 2xz$ i $F^2(x,y,z,t) = x^3 + y^3 + t^3 - z^3$ zadają razem w otoczeniu punktu $(1,-1,1,1) \in \mathbb{R}^4$ odwzorowanie $F : \mathbb{R}^4 \longrightarrow \mathbb{R}^2$. Pokazać, że $F^{-1}(\mathbf{0}) \subset \mathbb{R}^4$ zadaje w sposób uwikłany $x = x(y,t)$ i $z = z(y,t)$, tj. funkcję $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ i obliczyć pochodną

(tj. macierz pochodnej - pamiętamy, że pochodna funkcji $\mathbb{R}^n \longrightarrow \mathbb{R}^m$ jest macierzą $m \times n$ - ma m wierszy i n kolumn) tej funkcji w tym punkcie.

Zadanie 32
Znaleźć ekstrema (warunkowe) funkcji przy podanych warunkach

$$\begin{array}{lll} a) & F(x,y) = x^2 + y^2\,, & G(x,y) = x^3 + y^3 - 16 = 0\,, \\ b) & F(x,y) = 2x^2y^2\,, & G(x,y) = x^4 + y^4 - 1 = 0\,, \end{array}$$

Zadanie 33
Znaleźć na paraboli o równaniu $y^2 = 6x$ punkt ekstremalnie oddalony (w sensie zwykłej metryki Euklidesowej) od punktu $P = (3, 12)$.

Zadanie 34
Podać jakie musi mieć wymiary prostopadłościenne akwarium o objętości $32\,\mathrm{m}^3$, by zużyte na nie szkło miało jak najmniejsze pole powierzchni.

Zadanie 35
Znaleźć ekstrema funkcji $F(x,y,z) = xy^2z^3$ na zbiorze E zadanym warunkiem $G(x,y,z) = x + 2y + 3x - 1 = 0$. Rozwiązać ten problem zarówno metodą szukania ekstremów warunkowych, jak i zwyczajnie, eliminując jedną ze zmiennych.

Zadanie 36
Znaleźć przynajmniej jeden punkt krytyczny funkcji $F(x,y,z) = xyz$ na zbiorze $E \subset \mathbb{R}^3$ zadanym dwoma warunkami: $G^1(x,y,z) = x+y+z-5 = 0$ i $G^2(x,y,z) = xy+xz+yz-8 = 0$ i zbadać jego charakter.

Zadanie 37
Pewna powierzchnia zanurzona w $\mathbb{R}^3$ jest zadana parametrycznie wzorami

$$\begin{aligned} x &= a \sin \omega t \cos \xi\,, \\ y &= a \sin \omega t \sin \xi\,, \\ z &= z_0 + a \cos \omega t + a \ln[\mathrm{tg}(\omega t)]\,. \end{aligned}$$

Zmienne $0 \le t < \infty$ i $-\infty < \xi < \infty$ są więc układem współrzędnych na tej powierzchni. Podać wektory styczne do tej powierzchni i zbadać, w jakich punktach nie jest ona regularna.

Zadanie 38
Znaleźć wszystkie krzywe całkowe zwyczajnych równań różniczkowych pierwszego rzędu o zmiennych rozdzielonych (jeśli podany jest warunek, to podać także konkretne rozwiązanie spełniające ten warunek):

$$\begin{array}{lll} a) & y' = (1 + y^2)\,x\,, & y(7) = 1\,, \\ b) & y' = x\,y\,, & \end{array}$$

$$
\begin{aligned}
&c) \quad y' = y/x\,,\\
&d) \quad y' = -y/x\,,\\
&e) \quad y' = y/x^3\,, \qquad y(1) = 1\,,\\
&f) \quad y' = y/x^2\,,\\
&g) \quad y'\,\sin x = y\cos x\,,
\end{aligned}
$$

Przedyskutować charakter rozwiązania w zależności od warunku początkowego (x_0, y_0). Jeśli takowe występują, skomentować przypadki punktów, przez które przechodzi więcej, lub mniej niż dokładnie jedno rozwiązanie.

Zadanie 39

Rozwiązać (jeśli się da) równania różniczkowe dokonując w nich zaproponowanych podstawień (lub jakimś innym sposobem) następujące równania różniczkowe:

$$
\begin{aligned}
&a) \quad y' = \cos(x+y)\,, \qquad y(x) = -x + z(x)\,,\\
&b) \quad y' = -\frac{y}{x} + \frac{\cos(xy)}{x^2}\,, \qquad y(x) = \frac{u(x)}{x}\,,\\
&c) \quad x^2 y'' + x\,y' + y = 0\,, \qquad x = e^t\,,\\
&d) \quad (1-x^2)\,y'' - x\,y' + \omega^2 y = 0\,, \qquad x = \cos t\,,\\
&e) \quad x^4\,y'' + 2x^3\,y' + x^2 y = 0\,, \qquad x = 1/t\,,\\
&f) \quad y'' + (y')^2 = 2\,e^{-y}\,, \qquad y(0) = \ln 2\,,\ y'(0) = 1\,,\\
&g) \quad y'' = 2\,y^3\,, \qquad y(0) = 1\,,\ y'(0) = 1\,,\\
&h) \quad y' + \sin y + x\cos y + x = 0\,, \qquad u = \mathrm{tg}(y/2)\,,\\
&i) \quad y' = y/x + \sqrt{(y/x)^2 - 1}\,, \qquad y(1) = 2\,.
\end{aligned}
$$

O punkcie $f)$ powiem tylko, że trzeba napisać równanie na przedefiniowaną odpowiednio funkcję (ale tej samej zmiennej x). Proszę spróbować samemu zgadnąć! Co do punktu $g)$, to każdy fizyk, który wykorzystywał zasadę zachowania energii w problemach mechanicznych, powinien od razu skojarzyć, co trzeba zrobić. W punkcie $i)$ zauważyć, że prawa strona jest funkcją jednorodną stopnia zerowego. Przedyskutować tu zależność rozwiązania of warunku początkowego $y(x_0) = y_0$. Znaleźć także rozwiązanie nieobejmowane przez całkę ogólną.

Zadanie 40

Rozwiązać liniowe równania różniczkowe z niejednorodnością

$$
\begin{aligned}
&a) \quad x\,y' - y = 2x^3\,,\\
&b) \quad y' + \frac{xy}{1+x^2} = -\frac{1}{2x(1+x^2)}\,,\\
&c) \quad x\,y'' + y' = 4x\,, \qquad y(-1) = 0\,,\ y'(-1) = 0\,,\\
&d) \quad y'\sin x + y\cos x = \sin 2x\,, \qquad y(0) = 0\,,\\
&e) \quad y' = \frac{4y}{x} + x\sqrt{y}\,, \qquad y(1) = 1\,,
\end{aligned}
$$

$$f)\quad y' = \frac{y}{x}\ln\left|\frac{y}{x}\right|, \qquad y(1) = \ln 2\,.$$

W punktach $e)$ i $f)$ równania liniowe z niejednorodnością otrzymuje się po sprytnym podstawieniu.

Zadanie 41

Wyznaczyć taką krzywą $y = y(x)$, żeby odległość od punktu $(0,0)$ stycznej do tej krzywej w każdym jej punkcie $(x,\, y(x))$ była równa odciętej[63] tego punktu, czyli x.

Zadanie 42

Uzasadnić, że równanie Bernoulliego[64]

$$y' = y\,a(x) + y^n\,b(x)\,,$$

w którym $a(x)$ i $b(x)$ są zadanymi funkcjami (załóżmy, że regularnymi w jakimś obszarze $\mathbb{R}$), a n jest liczbą całkowitą, można rozwiązać poznanymi już metodami. Korzystając z tego rozwiązać równanie

$$3x\,y^2y' = 2y^3 + x^3\,.$$

Zadanie 43

Rozwiązać układy równań różniczkowych pierwszego rzędu z podanymi warunkami początkowymi

$$a)\quad \frac{d}{dt}\begin{pmatrix} y_1\\ y_2\\ y_3\end{pmatrix} = \begin{pmatrix} 2 & -1 & 2\\ 5 & -3 & -3\\ -1 & 0 & -2\end{pmatrix}\begin{pmatrix} y_1\\ y_2\\ y_2\end{pmatrix}, \qquad \begin{pmatrix} y_1(0)\\ y_2(0)\\ y_2(0)\end{pmatrix} = \begin{pmatrix} 1\\ 0\\ 1\end{pmatrix},$$

$$b)\quad \frac{d}{dt}\begin{pmatrix} y_1\\ y_2\\ y_3\end{pmatrix} = \begin{pmatrix} -2 & 1 & 0\\ 1 & -2 & 1\\ 0 & 1 & -2\end{pmatrix}\begin{pmatrix} y_1\\ y_2\\ y_2\end{pmatrix},$$

$$c)\quad \frac{d}{dt}\begin{pmatrix} y_1\\ y_2\\ y_3\end{pmatrix} = \begin{pmatrix} 1 & 0 & 0\\ 1 & -1 & -1\\ -2 & 4 & 3\end{pmatrix}\begin{pmatrix} y_1\\ y_2\\ y_2\end{pmatrix}, \qquad \begin{pmatrix} y_1(0)\\ y_2(0)\\ y_2(0)\end{pmatrix} = \begin{pmatrix} 1\\ 1\\ 1\end{pmatrix}.$$

Ostatnie równanie rozwiązać na dwa sposoby tak jak zadanie Ode.8 w tekście.

Zadanie 44

Rozwiązać układy liniowych równań różniczkowych pierwszego rzędu z niejednorodnością z podanymi warunkami początkowymi

$$a)\quad \frac{d}{dx}\begin{pmatrix} y_1\\ y_2\end{pmatrix} = \begin{pmatrix} 0 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix} y_1\\ y_2\end{pmatrix} + \begin{pmatrix} \sin x\\ 2\cos x\end{pmatrix}, \qquad \begin{pmatrix} y_1(0)\\ y_2(0)\end{pmatrix} = \begin{pmatrix} 2\\ 0\end{pmatrix},$$

$$b)\quad \frac{d}{dt}\begin{pmatrix} y_1\\ y_2\\ y_3\end{pmatrix} = \begin{pmatrix} 3 & -1 & 0\\ 0 & 3 & -1\\ -2 & 5 & -1\end{pmatrix}\begin{pmatrix} y_1\\ y_2\\ y_3\end{pmatrix} + \begin{pmatrix} e^t\\ 0\\ 0\end{pmatrix}, \qquad \begin{pmatrix} y_1(0)\\ y_2(0)\\ y_3(0)\end{pmatrix} = \begin{pmatrix} 1\\ 1\\ 3\end{pmatrix}.$$

[63] Były kiedyś takie śmieszne nazwy ”rzędna” i “odcięta”. Może nawet nadal są?

[64] Ale którego?! Tylu ich było!

Układ a) rozwiązać na dwa sposoby: raz jako równanie macierzowe, a drugi raz jako dwa niezależne równania otrzymane przez wzięcie odpowiednich kombinacji liniowych równań tworzących powyższy układ.

Zadanie 45
Znaleźć całkę ogólną równania różniczkowego

$$5\,y'' - 6\,y' + 5\,y = \sin\frac{4}{5}x\,.$$

Zadanie 46
Znaleźć całkę ogólną równania różniczkowego

$$y'' - 3\,y' + 2\,y = \sin\bigl(e^{-x}\bigr)\,.$$

Zadanie 47
Znaleźć całkę ogólną liniowego jednorodnego równania różniczkowego

$$y^{(7)} - 3\,y^{(6)} + 5\,y^{(5)} - 7\,y^{(4)} + 7\,y^{(3)} - 5\,y'' + 3\,y' + y = 0\,.$$

Zadanie 48
Znaleźć całkę ogólną równania różniczkowego

$$y''' + y'' + y' + y = x\,e^{-x} + \sin x\,.$$

Waskazówka: Spróbować zgadnąć rozwiązanie równania niejednorodnego. Jeśli proste podstawienie nie zadziała, to rozpatrzyć najpierw równanie $y' + y = x\,e^{-x}$.

Zadanie 49
Znaleźć najogólniejsze rozwiązanie liniowego równania różniczkowego drugiego rzędu

$$y'' + a\,y' + \frac{1}{4}\,a^2\,y = f(t)\,,$$

z dowolną niejednorodnością $f(t)$.

Odpowiedzi i podpowiedzi

Zadanie 1: Niewątpliwie $d_{\rm discr}(x,y)=d_{\rm discr}(y,x)$ oraz $x=y$, gdy $d_{\rm discr}(x,y)=0$. Tylko nierówność trójkąta wymaga sprawdzenia "case by case":
i) $x=y=z$, wtedy $0 =\leq 0+0$,
ii) $x=y\neq z$, wtedy $0\leq 1+1$,
iii) $x=z\neq y$, wtedy $1\leq 1+0$,
iv) $x\neq y$, $x\neq z$, $z\neq y$, wtedy $1\leq 1+1$. I to już chyba wszystko, bo przypadek $y=z\neq x$ jest taki sam jak *iii)*.

Zadanie 2: $K(x_0,r)$: jeśli $r=0$, to $K(x_0,r)=\emptyset$ (bo zgodnie z jednym z warunków spełnianych przez metrykę jako taką, żadna odległość, od zera mniejsza być nie może), $K(x_0,r)=\{x_0\}$ jeśli $0<r\leq 1$ (bo odległość $d_{\rm discr}(x_0,x)$ jest albo równa zeru, albo 1) i wreszcie $K(x_0,r)=\mathcal{X}$, gdy $r>1$ (bo przy tej metryce każdy punkt jest od x_0 odległy o nie więcej niż 1). $\bar{K}(x_0,r)$: jeśli $0\leq r<1$, to $\bar{K}(x_0,r)=\{x_0\}$ oraz $\bar{K}(x_0,r)=\mathcal{X}$, gdy $r\geq 1$.

Zadanie 3: Wszystkie punkty zbioru A są izolowane, bo zawsze można wybrać kulę $K(\frac{1}{n},r)$, o $r<\frac{1}{n(n+1)}$, do której oprócz $x=\frac{1}{n}$ żadne inne punkty zbioru A nie należą. Jedynym punktem skupienia zbioru A jest punkt $x=0\in\mathbb{R}$ (tylko dowolna kula o środku w $x=0$ zawsze zawiera punkty typu $1/n$), ale on do zbioru A nie należy. Zbiór A nie posiada punktów wewnętrznych: A nie jest bowiem otoczeniem żadnego ze swoich elementów bo dowolna kula o środku w $1/n$ zawiera w sobie punkty, które nie są tej postaci (nie należą więc do A). Zbiór A nie jest otwarty, bo ani jeden (a musiałyby wszystkie być) nie jest środkiem kuli całkowicie zwartej w A (czyli składającej się tylko z punktów typu $1/n$). Nie jest też domknięty, bo jego punkt skupienia 0 doń nie należy.

Zadanie 4: Najprościej rozpatrzyć kulę o środku w punkcie $(0,0)$ i promieniu $r=1$ (każda inna wygląda jak taka tylko przesunięta i przeskalowana). W metryce d_∞ $K((0,0),1)=\{(x_1,x_2)\in\mathbb{R}^2|\max(|x_1|,|x_2|)<1\}$. Jest to wnętrze kwadratu o środku w $(0,0)$ i bokach długości 2 równoległych do odpowiednich osi. W metryce $d_{\rm taxi}$ $K((0,0),1)=\{(x_1,x_2)\in\mathbb{R}^2|\,|x_1|+|x_2|<1\}$. Jest to też wnętrze kwadratu który jest zawarty pomiędzy prostymi $y=\pm x\pm 1$

Zadanie 5: W metryce d_1

$$d_1(f,g)={\rm max}_{x\in[0,\pi]}|\sin x-\cos x|=\frac{1}{\sqrt{2}}-\left(-\frac{1}{\sqrt{2}}\right)=\sqrt{2}\,.$$

W metryce d_2:

$$\begin{aligned}d_2(f,g)&=\int_0^\pi dx\,|\sin x-\cos x|=\int_0^{\pi/4}dx\,(\cos x-\sin x)+\int_{\pi/4}^\pi dx\,(\sin x-\cos x)\\&=\left(\frac{1}{\sqrt{2}}+\frac{1}{\sqrt{2}}-1\right)-\left(-1-\frac{1}{\sqrt{2}}-\frac{1}{\sqrt{2}}\right)=2\sqrt{2}\,.\end{aligned}$$

Zadanie 6: Jak zwykle z początku natłok hieroglifów przyprawia o zawrót głowy, ale jak wszystko rozebrać na części (a najlepiej zrobić rysunek na kartce, czyli w $\mathbb{R}^2$), to się okazuje, że to oczywiste. Z definicji kuli wynika bowiem, że $d(x,x') < r$, więc rzeczywiście $r' \equiv r - d(x,x') > 0$. Weźmy teraz jakiś dowolny punkt $y \in K(x',r')$ i zastosujmy nierówność trójkąta:

$$d(y,x) \leq d(y,x') + d(x',x)\,.$$

Ale $d(y,x') < r'$, bo $y \in K(x',r')$, a z kolei $d(x',x) < r - r'$ z tego, co wyżej. Zatem $d(y,x) < r$, co oznacza, że każdy punkt $K(x',r')$ jest od x odległy o mniej niż r, czyli należy też do $K(x,r)$, więc istotnie $K(x',r') \subset K(x,r)$.

Zadanie 7: Zbiór, w tym przypadku $K(x,r)$, jest otwarty, gdy każdy jego punkt y jest środkiem jakiejś kuli $K(y,r_y)$ całkowicie zawartej w tym zbiorze (można tak dobrać r_y), czyli tu w $K(x,r)$. To weźmy jako $r_y = r - d(x,y)$ i wtedy z tego, co pokazane zostało w Zadaniu poprzednim wynika już, iż $K(y,r_y) \subset K(x,r)$. Zatem istotnie, ponieważ y był dowolnym punktem $K(x,r)$, kula $K(x,r)$ jest zbiorem otwartym (niezależnie od tego, jaka jest metryka d).

Zadanie 8: Jeśli

$$x \in O = \bigcup_{s\in S} O_s\,,$$

to $x \in O_s$, dla przynakmniej jednej wartości s (x należy do przynajmniej jednego ze zbiorów wchodzącegi w skład sumy). Zatem ten O_s jest otoczeniem x. Ponieważ jednak $O_s \subset O$, więc i O jest otoczeniem x. Tak więc O jest otoczeniem każdego swojego punktu, więc jest otwarty. Aby dowieść otwartości przecięcia skończonej liczby zbiorów otwartych, wystarczy rozpatrzyć przecięcie dwóch takich zbiorów. Jeśli $x \in O = O_1 \cap O_2$, to $x \in O_1$ i $x \in O_2$. Zatem i O_1 i O_2 są otoczeniami x i stąd także $O_1 \cap O_2$ jest otoczeniem x. Zatem zbiór O jest otoczeniem każdego swojego punktu, co oznacz, że jest otwarty. Jeśli jednak w $\mathbb{R}$ weźmiemy jako rodzinę zbiorów otwartych odcinki otwarte $O_n = (-1/n,\, 1/n)$, gdzie $n \in \mathbb{N}$ (znów 0 nie uważamy za liczbę naturalną), to

$$\bigcap_{n\in\mathbb{N}} O_n = \{0\}\,,$$

tj. przecięcie jest zbiórem złożonym z punktu zero tylko, który to zbiór nie jest otwarty.

Zadanie 9: a) Granica nie istnieje. Zbadać $f(x_n,y_n)$ z $x_n = a/n$, $y_n = b/n$. b) Granica jest równa 0; wziąć $x_n = r_n \cos\varphi_n$, $y_n = \sin\varphi_n$ z $r_n \to 0$. c) Jak w poprzednim. d) Granica nie istnieje: wziąć $x_n = a/n$, $y_n = b/n$, co da

$$f(x_n y_n) = \frac{2a^2 + 3b^2 + (a^2+b^2)/n}{a^2+b^2}\,,$$

i już widać, że granica zależy od a i b, czyli od wyboru ciągu. e) Granica istnieje i jest równa 0. f) Nie istnieje. g) Nie istnieje. Wziąć raz ciąg $(1/n,\, 1/n)$, a drugi raz

$(1/n,\ -1/n)$ i raz dostanie się granicę 1, a drugi raz 0. h) W punktach $(x_0, 0)$ z $x_0 \neq 0$ granica oczywiście nie istnieje. W punkcie $(0,0)$ granica istnieje i jest równa 0.

Zadanie 10: Treść zadania jest dłuższa niż to, co trzeba zrobić (tylko trzeba pomyśleć i się nie bać). Jeśli odwzorowanie F jest ciągłe w jakimś punkcie-vektorze $\mathbf{v}_0$, to znaczy, że gdy weźmiemy dowolny ciąg $\mathbf{v}_n$ wektorów zbieżny w normie $\|\cdot\|_1$ do $\mathbf{v}_0$ (tj. taki, że $\|\mathbf{v}_n - \mathbf{v}_0\|_1 \to 0$) to ciąg $F(\mathbf{v}_n)$ jest w sensie normy w V_2 zbieżny do $F(\mathbf{v}_0)$ (czyli $\|F(\mathbf{v}_n) - F(\mathbf{v}_0)\|_2 \to 0$). Załóżmy teraz, że odwzorowanie F nie jest ciągłe w jakimś pukcie-wektorze $\mathbf{v}_0'$, co zgodnie z kryterium Heinego ciągłości oznacza, że istnieje przynajmniej jeden ciąg wektorów $\mathbf{v}_n'$ zbieżny do $\mathbf{v}_0'$ w normie $\|\cdot\|_1$ ale $F(\mathbf{v}_n')$ nie zbiega (w normie $\|\cdot\|_2$) do $F(\mathbf{v}_0')$. No ale biorąc wektor $\mathbf{w} = \mathbf{v}_0 - \mathbf{v}_0'$ możemy stworzyć ciąg $\mathbf{w} + \mathbf{v}_n'$ zbiegający do $\mathbf{v}_0$ i wobec tego, z założonej ciągłości odwzorowania F w punkcie wektorze $\mathbf{v}_0$ musimy mieć zbieganie ciągu $F(\mathbf{w} + \mathbf{v}_n')$ do wektora $F(\mathbf{v}_0)$, który, na mocy liniowości F, jest równy $F(\mathbf{v}_0') + F(\mathbf{w})$. Z drugiej jednak strony wzięliśmy (co jest możliwe, jeśli F nie jest ciągłe w $\mathbf{v}_0'$) taki ciag $\mathbf{v}_n'$, że $F(\mathbf{v}_n')$ nie zbiega do $F(\mathbf{v}_0')$, czyli $F(\mathbf{v}_n' + \mathbf{w}) = F(\mathbf{v}_n') + F(\mathbf{w})$ nie może zbiegać do $F(\mathbf{v}_0') + F(\mathbf{w})$. Sprzeczność. Koniec. Zauważmy jeszcze, że to rozumowanie jest bardzo ogólne i pozostaje słuszne zawsze, niezależnie czym są przestrzenie wektorowe V_1 i V_2: mogą one być nawet nieskończenie-wymiarowe a nawet, najbardziej "przepastne", czyli nieośrodkowe (a to takie się nam pojawiają w kwantowej teorii pola.

Zadanie 11:

$$
\begin{aligned}
a)\quad & f_x = 2x + 3x^2y^3 - \sin y\,, \quad f_y = -1 + 6y + 3x^3y^2 - x\cos y\,,\\
& f_{xx} = 2 + 6xy^3\,, \quad f_{yy} = 6 + 6x^3y + x\cos y\,, \quad f_{yx} = f_{xy} = 9x^2y^2 - \cos y\,,\\
b)\quad & f_x = \frac{1}{y-1}\,, \quad f_y = -\frac{x}{(y-1)^2}\,,\\
& f_{xx} = 0\,, \quad f_{yy} = \frac{2x}{(y-1)^3}\,, \quad f_{yx} = f_{xy} = -\frac{1}{(y-1)^2}\,,\\
c)\quad & f_x = -\frac{y}{x^2+y^2}\,, \quad f_y = \frac{x}{x^2+y^2}\,,\\
& f_{xx} = \frac{2xy}{(x^2+y^2)^2}\,, \quad f_{yy} = -\frac{2xy}{(x^2+y^2)^2}\,, \quad f_{yx} = f_{xy} = \frac{y^2-x^2}{(x^2+y^2)^2}\,,\\
& f_{yyx} = f_{yxy} = f_{xyy} = \frac{6x^2y - 2y^3}{(x^2+y^2)^3}\,, \quad f_{yxx} = f_{xyx} = f_{xxy} = \frac{2x^3 - 6xy^2}{(x^2+y^2)^3}\,,\\
d)\quad & f_x = n(3x^2y+z)^{n-1}6xy\,, \quad f_y = n(3x^2y+z)^{n-1}3x^2\,, \quad f_z = n(3x^2y+z)^{n-1}\,,\\
& f_{xx} = n(n-1)(3x^2y+z)^{n-2}36x^2y^2 + n(3x^2y+z)^{n-1}6y\,,\\
& f_{yy} = n(n-1)(3x^2y+z)^{n-2}9x^4\,, \quad f_{zz} = n(n-1)(3x^2y+z)^{n-2}\,,\\
& f_{yx} = f_{xy} = n(n-1)(3x^2y+z)^{n-2}18x^3y + n(3x^2y+z)^{n-1}6x\,,\\
& f_{zx} = f_{xz} = n(n-1)(3x^2y+z)^{n-2}6xy\,, \quad f_{yz} = f_{zy} = n(n-1)(3x^2y+z)^{n-2}3x^2\,,\\
e)\quad & f_x = \frac{2x}{x^2+y^2}\,, \quad f_y = \frac{2y}{x^2+y^2}\,,
\end{aligned}
$$

$$f_{xx} = \frac{2y^2 - 2x^2}{(x^2+y^2)^2}\,, \quad f_{yy} = \frac{2x^2 - 2y^2}{(x^2+y^2)^2}\,, \quad f_{yx} = f_{xy} = -\frac{4xy}{(x^2+y^2)^2}\,,$$

f) $f_x = \dfrac{x}{r}\,, \quad f_y = \dfrac{y}{r}\,, \quad f_z = \dfrac{z}{r}\,, \quad r \equiv \sqrt{x^2+y^2+z^2}\,,$

$$f_{xx} = \frac{y^2+z^2}{r^3}\,, \quad f_{yy} = \frac{x^2+z^2}{r^3}\,, \quad f_{zz} = \frac{x^2+y^2}{r^3}\,,$$

$$f_{yx} = f_{xy} = -\frac{xy}{r^3}\,, \quad f_{zx} = f_{xz} = -\frac{xz}{r^3}\,, \quad f_{yz} = f_{zy} = -\frac{yz}{r^3}\,,$$

g) $f_x = -\dfrac{x}{r^3}\,, \quad f_y = -\dfrac{y}{r^3}\,, \quad f_z = -\dfrac{z}{r^3}\,, \quad r \equiv \sqrt{x^2+y^2+z^2}\,,$

$$f_{xx} = \frac{2x^2 - y^2 - z^2}{r^5}\,, \quad f_{yy} = \frac{2y^2 - x^2 - z^2}{r^5}\,, \quad f_{zz} = \frac{2z^2 - x^2 - y^2}{r^5}\,,$$

$$f_{yx} = f_{xy} = \frac{3xy}{r^5}\,, \quad f_{zx} = f_{xz} = \frac{3xz}{r^5}\,, \quad f_{yz} = f_{zy} = \frac{3yz}{r^5}\,,$$

h) $f_u = \dfrac{1}{u + \ln v}\,, \quad f_v = \dfrac{1}{v\,(u + \ln v)}\,,$

$$f_{uu} = -\frac{1}{(u+\ln v)^2}\,, \quad f_{vv} = -\frac{1 + u + \ln v}{v^2(u+\ln v)^2}\,, \quad f_{uv} = f_{vu} = -\frac{1}{v\,(u+\ln v)^2}\,,$$

i) $f_x = -\dfrac{2ax}{r\,(a^2 - r^2)}\,, \quad f_y = -\dfrac{2ay}{r\,(a^2 - r^2)}\,, \quad f_z = -\dfrac{2az}{r\,(a^2 - r^2)}\,,$

$$f_{xx} = -\frac{2a\,[2x^4 + x^2(y^2+z^2) + (a^2 - y^2 - z^2)(y^2+z^2)]}{r^3(a^2-r^2)^2}\,,$$

$$f_{xy} = f_{yx} = -\frac{2axy(a^2 - 3r^2)}{r^3(a^2-r^2)^2}\,, \qquad \text{reszta pochodnych z symetrii,}$$

j) $f_x = y\,x^{y-1}\,, \quad f_y = x^y \ln x\,,$

$$f_{xx} = (y^2 - y)x^{y-2}\,, \quad f_{yy} = x^y \ln^2 x\,, \quad f_{xy} = f_{yx} = (1 + y\ln x)x^{y-1}\,,$$

k) $f_x = \dfrac{\sin y}{x}\,(\ln x)^{-1+\sin y}\,, \quad f_y = (\ln x)^{\sin y}\cos y\,\ln(\ln x)\,,$

$$f_{xx} = \frac{\sin y}{x^2}\,(-1 + \sin y - \ln x)\,(\ln x)^{-2+\sin y}\,,$$

$$f_{yy} = \left(-\sin y + \cos^2 y \ln\ln x\right)(\ln x)^{\sin y}\ln\ln x\,,$$

$$f_{xy} = f_{yx=}\frac{\cos y}{x}\,(1 + \sin y \ln\ln x)\,(\ln x)^{-1+\sin y}\,,$$

l) $f_x = (x\,\mathrm{tg}z)^{-1+\ln y}(\mathrm{tg}z)\ln y\,, \;\; f_y = \dfrac{\ln(x\,\mathrm{tg}z)}{y}\,(x\,\mathrm{tg}z)^{\ln y}\,, \;\; f_z = \dfrac{x\ln y}{\cos^2 z}\,(x\,\mathrm{tg}z)^{-1+\ln y}\,,$

$$f_{xx} = \mathrm{tg}^2 z(-1+\ln y)\ln y\,(x\,\mathrm{tg}z)^{-2+\ln y}\,, \;\; f_{yy} = \frac{-1 + \ln(x\,\mathrm{tg}z)}{y^2}\,\ln(x\,\mathrm{tg}z)\,(x\,\mathrm{tg}z)^{\ln y}\,,$$

$$f_{zz} = -\frac{4}{\sin^2 2z}\,(\cos 2z - \ln y)\ln y\,(x\,\mathrm{tg}z)^{\ln y}\,,$$

$$f_{xy} = f_{yx} = \frac{1+\ln y}{xy}\,\ln(x\,\mathrm{tg}z)\,(x\,\mathrm{tg}z)^{\ln y}\,, \quad f_{xz} = f_{zx} = \frac{\ln y}{\cos^2 z}\,(x\,\mathrm{tg}z)^{-1+\ln y}\,,$$

$$f_{zy} = f_{yz} = \frac{4(1+\ln y)}{y \sin^2 2z} \ln(x \,\mathrm{tg} z)\,(x\,\mathrm{tg} z)^{\ln y}\,,$$

$m)$ $f_x = y^z x^{-1+y^z}\,,\quad f_y = x^{y^z} z\, y^{-1+z} \ln x\,,\quad f_z = x^{y^z} y^z (\ln z)(\ln y)\,,$

$$f_{xx} = y^z(-1+y^z)\, x^{-2+y^z}\,,\quad f_{yy} = x^{y^z} y^{-2+z}(-1+z+y^z z \ln x)\, z \ln x\,,$$
$$f_{zz} = x^{y^z} y^z (1+y^z \ln x)(\ln x)(\ln y)^2\,,\quad f_{xy} = f_{yx} = x^{-1+y^z} y^{-1+z} z\,(1+y^z \ln x)\,,$$
$$f_{xz} = f_{zx} = x^{-1+y^z} y^z (1+y^z \ln x)(\ln y)\,,$$
$$f_{yz} = f_{zy} = x^{y^z} y^{-1+z}(1+z\,(1+y^z \ln y) \ln y) \ln x\,,$$

$n)$ $f_x = y\, g(xy)\,,\quad f_y = x\, g(xy)\,,$

$$f_{xx} = y^2 g'(xy)\,,\quad f_{yy} = x^2 g'(xy)\,,\quad f_{xy} = f_{yx} = g(xy) + xy\, g'(xy)\,,$$

W przykładzie $g)$ pozornie $f_{xx}+f_{yy}+f_{zz} = 0$ w całej przestrzeni $\mathbb{R}^3$. Równość ta jednak nie zachodzi w punkcie $(0,0,0)$, gdyż w przeciwnym razie prawo Gaussa znane nam (zapewne) ze szkoły nie mogło by być prawdziwe: funkcja f jest tu bowiem proporcjonalna do potencjału elektrostatycznego ładunku punktowego umieszczonego w $(0,0,0)$ i wobec tego pochodne f_x, f_y, f_z są składowymi wytwarzanego przez taki ładunek pola elektrycznego $\mathbf{E}$, a suma $f_{xx}+f_{yy}+f_{zz}$ jest dywergencją $\boldsymbol{\nabla}\cdot\mathbf{E}$ tegoż pola. Prawo Gaussa jednak mówi, że $\boldsymbol{\nabla}\cdot\mathbf{E} \propto \rho$, gdzie ρ jest gęstością ładunku elektrycznego i całka po objętości V z dywergencji pola elektrycznego jest proporcjonalna do ładunku zawartego w V (całkowa wersja wypisanego wyżej prawa Gaussa, które jest jednym z równań Maxwella)

$$\int_V d^3\mathbf{r}\, \boldsymbol{\nabla}\cdot\mathbf{E} = \int d\mathbf{s}\cdot\mathbf{E} \propto Q\,.$$

Jeśli więc $Q \neq 0$, nie może znikać i $\boldsymbol{\nabla}\cdot\mathbf{E} \propto \rho$. W istocie, funkcja f nie jest określona w punkcie $(0,0,0)$ i wszystkie wykonane różniczkowania są słuszne tylko poza tym punktem; nadanie sensu różniczkowaniu także i w punkcie $(0,0,0)$ wymaga przejścia do tzw. dystybucji, czyli uogólnionych funkcji - to się zwykle wykłada w ramach Analizy III. Zachodzi wtedy wzór

$$f_{xx} + f_{yy} + f_{zz} = -4\pi\delta^{(3)}(\mathbf{r})\,,$$

gdzie ta dystrybucja $\delta^{(3)}(\mathbf{r})$ zwana deltą Diraca to jest taka niby funkcja, która jest równa zeru wszędzie oprócz punktu $(0,0,0)$, gdzie jest ona równa nieskończoności i to tak zmyślnie, że całka z niej po dowolnym trójwymiarowym obszarze obejmującym punkt $(0,0,0)$ jest równa 1 (czego to ludzie nie wymyśla!). Podobnie jest z funkcją z punktu $e)$: pozornie $f_{xx}+f_{yy} = 0$, ale naprawdę $f_{xx}+f_{yy} = 2\pi\delta^{(2)}(\mathbf{r})$.

Zadanie 12:

$a)$
$$f_x(3,4) = \lim_{h\to 0} \frac{(3+h)+4-\sqrt{(3+h)^2+16}-7-\sqrt{25}}{h}$$
$$= 1 - \lim_{h\to 0} \frac{\sqrt{25+6h+h^2}-\sqrt{25}}{h} = 1 - \lim_{h\to 0} \frac{6h+h^2}{h\,(\sqrt{25+6h+h^2}+5)} = \frac{2}{5}\,.$$

Obliczenie $f_y(3,4)$ jest analogiczne i daje $f_y(3,4) = 1/5$.

$$b)\quad f_x(1,2) = \lim_{h\to 0} \frac{1}{h}\left[\ln\left(1+h+\frac{2}{2(1+h)}\right) - \ln 2\right]$$
$$= \lim_{h\to 0} \frac{1}{h}\ln\left[1+\frac{h}{2}+\frac{1}{2}\left(\frac{1}{1+h}-1\right)\right] = \lim_{h\to 0}\frac{1}{h}\ln\left[1+\frac{h}{2}-\frac{h}{2(1+h)}\right]$$
$$= \lim_{h\to 0}\frac{1}{h}\left(\frac{h}{2}-\frac{h}{2(1+h)}+\mathcal{O}(h^2)\right) = \lim_{h\to 0}\left(\frac{1}{2}-\frac{1}{2(1+h)}+\mathcal{O}(h^2)\right) = 0\,,$$
$$f_y(1,2) = \lim_{h\to 0}\frac{1}{h}\left[\ln\left(1+\frac{2+h}{2}\right)-\ln 2\right] = \lim_{h\to 0}\frac{1}{h}\ln\left(1+\frac{h}{4}\right) = \frac{1}{4}\,.$$

$$c)\quad (\nabla_{\mathbf{n}} f)(2,1) = \lim_{h\to 0}\frac{1}{h}\left[\left(2+\frac{h}{\sqrt{10}}\right)^2 + 2\left(1+\frac{3h}{\sqrt{10}}\right)^2 - 2^2 - 2\cdot 1^2\right]$$
$$= \lim_{h\to 0}\frac{1}{h}\left(\frac{4h}{\sqrt{10}}+\frac{h^2}{10}+\frac{12h}{\sqrt{10}}+\frac{18h^2}{10}\right) = \frac{16}{\sqrt{10}}\,.$$

Pochodne cząstkowe $\partial f/\partial x = 2x$, $\partial f/\partial y = 4y$ sa w punkcie $(2,1)$ równe odpowiednio 4 i 4 i kombinacja $4\,n_1 + 4\,n_2$ jest równa obliczonej wyżej z definicji pochodnej kierunkowej.

$$d)\quad (\nabla_{\mathbf{n}} f)(1,1) = \lim_{h\to 0}\frac{1}{h}\left\{\sin\left[\frac{\pi}{2}\left(1+\frac{h}{\sqrt{10}}+1-\frac{3h}{\sqrt{10}}\right)\right] - \sin\pi\right\}$$
$$= \lim_{h\to 0}\frac{1}{h}\sin\left[\frac{\pi}{2}\left(2-\frac{2h}{\sqrt{10}}\right)\right] = \lim_{h\to 0}\frac{1}{h}\sin\left(\frac{\pi h}{\sqrt{10}}\right) = \frac{\pi h}{\sqrt{10}}\,,$$

co zgadza się z odpowiednią kombinacją liniową pochodnych cząstkowych obliczonych w punkcie $(1,1)$:

$$\left.\frac{\partial f}{\partial x}\right|_{(1,1)} = \frac{\pi}{2}\cos\left(\frac{\pi}{2}\,(x+y)\right)\Big|_{(1,1)} = -\frac{\pi}{2} = \left.\frac{\partial f}{\partial y}\right|_{(1,1)}\,,$$

i

$$n_1\left.\frac{\partial f}{\partial x}\right|_{(1,1)} + n_2\left.\frac{\partial f}{\partial y}\right|_{(1,1)} = \frac{\pi}{2}\left(\frac{1}{\sqrt{10}}-\frac{3}{\sqrt{10}}\right) = \frac{\pi}{\sqrt{10}}\,.$$

Zadanie 13: a) Ciągłość tej fukcji w punkcie $(0,0)$ jest oczywista. Abyo bliczyć z definicji jej pochodne kierunkowe w punkcie $(0,0)$, bierzemy jednostkowy wektor $\mathbf{n}$ o składowych $n_1 = \cos\theta \equiv c$ i $n_2 = \sin\theta \equiv s$ i obliczamy granicę:

$$\nabla_{\mathbf{n}} f|_{(0,0)} = \lim_{h\to 0}\frac{f(hc,hs)-f(0,0)}{h} = c^2 s\,.$$

Stąd od razu widać, że $f_x(0,0) = 0$ i $f_y(0,0) = 0$ bo pierwsza pochodna cząstkowa odpowiada kierunkowej z $c = 1$, $s = 0$, a druga z $c = 0$, $s = 1$. Zatem $df = 0$ niezależnie od przyrostu (dx, dy), podczas gdy

$$\Delta f = f(0 + dx,\, 0 + dy) - f(0,0) = \frac{(dx)^2 dy}{[(dx)^2 + (dy)^2]}\,,$$

czyli

$$\frac{\Delta f - df}{\sqrt{(dx)^2 + (dy)^2}} = \frac{(dx)^2 dy}{[(dx)^2 + (dy)^2]^{3/2}}\,,$$

nie dąży do zera, gdy $\sqrt{(dx)^2 + (dy)^2} \to 0$. A wszystko to dlatego, że pochodne $f_x(x,y)$ i $f_y(x,y)$

$$f_x = \frac{2xy^3}{(x^2 + y^2)^2}\,, \qquad f_y = \frac{x^4 - x^2y^2}{(x^2 + y^2)^2}\,,$$

traktowane jak fukcje na $\mathbb{R}^2$ nie są ciągłe w punkcie $(0,0)$, co już każdy powinien umieć sprawdzić. Oznacza to że ta "ŕożniczka" df prawdziwą różniczką nie była. Widać tu też, że pochodna kierunkowa w kierunku jednostkowego wektora $\mathbf{n}$ nie jest w punkcie $(0,0)$ równa $n_1 f_x(0,0) + n_2 f_y(0,0)$ bo ta kombinacja jest zawsze równa zero.

b) Tu ciągłość funkcji w punkcie $(0,0)$ sprawdzamy rozpatrując np. ciągi $x_n \to 0$ i $y_n \to 0$ i pisząc

$$f(x_n, y_n) = x_n\, \frac{(x_n^2/y_n)}{1 + (x_n^2/y_n)^2} = x_n\, \frac{(y_n/x_n^2)}{1 + (y_n/x_n^2)^2}\,.$$

Niezależnie od tego, jak zachowują się ciągi $c_n \equiv x_n^2/y_n$, lub $d_n \equiv y_n/x_n^2$, oba wyrażenia dążą do zera, gdy $x_n \to 0$. Zatem funkcja f jest w punkcie $(0,0)$ ciągła. Obie pochodne cząstkowe f jako funkcje na $\mathbb{R}^2$

$$f_x(x,y) = \frac{3x^2y^3 - x^6y}{(x^4 + y^2)^2}\,, \qquad f_y(x,y) = \frac{x^7 - x^3y^2}{(x^4 + y^2)^2}\,,$$

są w punkcie $(0,0)$ ciągłe jeśli badać ich ciągłość wzdłuż prostych, tj. na ciągach $x_n = a/n$, $y_n = b/n$ (albo ogólniej $x_n = af(n)$, $y_n = bf(n)$, $f(n) \to 0$):

$$f_x(x_n, y_n) = \frac{3a^2b^3 - a^6b/n^2}{n\,(b^2 + a^4/n^2)^2} \to 0\,, \qquad f_y(x_n, y_n) = \frac{-a^3b^2 + a^7/n^2}{n\,(b^2 + a^4/n^2)^2} \to 0\,,$$

z wyjątkiem f_y badanej wzdłuż osi x, kiedy to $b = 0$ - pochodna ta w tym kierunku dąży do ∞. Zatem obie pochodne cząstkowe $(\partial f/\partial x)_{(0,0)}$ i $(\partial f/\partial y)_{(0,0)}$ istnieją, bo ta druga odpowiada granicy f_y ale branej wzdłuż osi y (a nie x, gdzie ujawnia się jej nieciągłość). Tak więc ściśle rzecz biorąc pochodne cząstkowe nie są ciągłe ale pochodne kierunkowe istnieją w każdym kierunku, co jak zwykle można sprawdzić biorąc $n_1 = c$, $n_2 = s$ $(c^2 + s^2 = 1)$:

$$\nabla_{\mathbf{n}} f|_{(0,0)} = \lim_{h\to 0} \frac{1}{h}\left[\frac{c^3 s h^4}{c^4 h^4 + s^2 h^2} - 0\right] = \lim_{h\to 0} \frac{c^3 s h}{c^2 h^4 + s^2} = 0\,,$$

z wyjątkiem przypadku $s = 0$, czyli pochodnej kierunkowej w kierunku osi x, czyli po prostu pochodnej cząstkowej $(\partial f/\partial x)_{(0,0)}$, kiedy to powyższe wyrażenie dąży do nieskończoności; znów jednak pochodna $(\partial f/\partial x)_{(0,0)}$ jak najbardziej istnieje

$$(\partial f/\partial x)_{(0,0)} = \lim_{h\to 0} \frac{0-0}{h} = 0\,.$$

tylko pochodne kierunkowe jako funkcje kierunku nie są ciągłe. Czyli prawdziwa pochodna nie istnieje (bo pochodne cząastkowe $f_x(x, y)$ i $f_y(x, y)$ nie są w punkcie $(0, 0)$ ciągłe), nie istnieje zatem i prawdziwa różniczka. Różniczka-Ersatz, czyli $(\partial f/\partial x)_{(0,0)}dx+(\partial f/\partial y)_{(0,0)}dy = 0\cdot dx+0\cdot dy$ nie przybliża należycie przyrostu Δf funkcji, bo wyrażenie

$$\frac{\Delta f - 0\cdot dx - 0\cdot dy}{\sqrt{(dx)^2 + (dy)^2}} = \frac{(dx)^3 dy}{[(dx)^4 + (dy)^2]\sqrt{(dx)^2 + (dy)^2}}\,,$$

nie dąży bezwarunkowo do zera: np. gdy $dy = (dx)^2 \to 0$, wyrażenie to ma postać

$$\frac{(dx)^5}{2(dx)^5\sqrt{1 + (dx)^2}}\,,$$

i ewidentnie nie dąży do zera.

Zadanie 14:
a) Funkcja $f(x, y) = x/(1 + y^2)$ jest przyzwoicie ciągła i ciągłe są w każdym punkcie $\mathbb{R}^2$ także jej pochodne cząstkowe

$$f_x(x, y) = \frac{1}{1 + y^2}\,, \qquad f_y = -\frac{2xy}{(1 + y^2)^2}\,.$$

Jej pochodna f' w dowolnym punkcie (x_0, y_0) zatem istnieje i jest dana przez pochodne cząstkowe:

$$f'(x_0, y_0) = (f_x(x_0, y_0),\ f_y(x_0, y_0))\ .$$

Jest ona oczywiście kowektorem (jedno-formą liniową), czyli odwzorowaniem z $\mathbb{R}^2$ (ale właściwie to nie z tego $\mathbb{R}^2$, które jest dziedziną samej funkcji i tu gra rolę rozmaitości, tylko tej przestrzeni wektorowej, w której żyją wektory przemieszczeń, czyli przestrzeni stycznej w danym punkcji (x_0, y_0) do rozmaitości; no ale na naszym poziomie wszystko się zlewa w jedno...) w $\mathbb{R}$, czyli w przestrzeń, z której wartości przyjmuje funkcja. W punkcie $(2,1)$ pochodna f jest kowektorem o składowych (w kanonicznej zero-jedynkowej bazie $\mathbb{R}^2$)

$$f'(2,2) = \left(\frac{1}{2},\,-1\right)\,,$$

i po zadziałaniu na wektor przyrostu $(\delta\lambda,\,\frac{1}{2}\delta\lambda)$ daje zero:

$$f'(2,2)\cdot\begin{pmatrix}\delta\lambda\\ \frac{1}{2}\delta\lambda\end{pmatrix} = \left(\frac{1}{2},\,-1\right)\cdot\begin{pmatrix}\delta\lambda\\ \frac{1}{2}\delta\lambda\end{pmatrix} = 0\,.$$

Oznacza to, że w tym kierunku, przy przejściu od punktu $(2,1)$ do punktu $(2+\delta\lambda,\,1+\frac{1}{2}\delta\lambda)$, funkcja w pierwszym przybliżeniu nie zmienia się (główna liniowa część zmiany wartości funkcji w tym kierunku jest równa zeru). Zadziaławszy zaś na wektor $(\delta\lambda,\,\delta\lambda)$ pochodna ta daje $-\delta\lambda/2$, co jest główną liniową częścią zmiany wartości przy przejściu od punktu $(2,1)$ do punktu $(2+\delta\lambda,\,1+\delta\lambda)$.
b) Znów funkcja jest najprzyzwoitsza z możliwych i jej pochodna w punkcie $(1,1)$ jest kowektorem

$$f'(1,1) = (4,\,6)\,.$$

Aby się dowiedzieć, w którym kierunku, ruszając z tego punktu, należy się przemieścić, by mieć najbardziej "pod górkę" (do szkoły oczywiście), obliczamy wartość tej pochodnej na przyroście $(dx,dy) = (\delta\lambda\cos\theta,\,\delta\lambda\sin\theta)$ przyjmując, że $\delta\lambda\geq 0$ (znak $\delta\lambda$ jest "załatwiany" przez kąt θ):

$$f'(1,1)\cdot\begin{pmatrix}\delta\lambda\cos\theta\\ \delta\lambda\sin\theta\end{pmatrix} = 2\,\delta\lambda\,(2\cos\theta + 3\sin\theta)\,.$$

Prawa strona daje główną liniową część zmiany wartości funkcji[65] przy przesunięciu się z puntu $(1,1)$ o odległość $\sqrt{2}\delta\lambda$ w kierunku zadanym przez kąt θ. Trzeba zatem znaleźć maksimum prawej strony jako funkcji $\theta\in[0,2\pi)$. To już umiemy:

$$\frac{d}{d\theta}\,2\,\delta\lambda\,(2\cos\theta + 3\sin\theta) = 0\,,\qquad \text{gdy}\quad \mathrm{tg}\,\theta = \frac{3}{2}\,.$$

W zakresie $\theta\in[0,2\pi)$ są oczywiście dwa rozwiązania tego warunku różniące się o π: w jednym druga pochodna jest ujemna (maksimum) i w tym to kierunku funkcja rośnie najszybciej, a w dugim druga pochodna jest dodatnia (minimum) i w kierunku odpowiadającym tej wartości θ funkcja najszybciej maleje. Oczywiście kierunki najszybszego wzrostu funkcji i najszybszego jej malenia są przeciwne (odpowiednie kąty θ różnią się o π), bo f' wyznaczające (gdy funkcja jest przyzwoita, czyli gdy ma pochodną prawdziwą) zmianę funkcji w dowolnym kierunku jest odwzorowaniem liniowym, a zmiana $\theta\to\theta\pm\pi$

[65] Przypomnijmy, że im mniejsza wartość $\delta\lambda$, tym lepiej ta główna liniowa część zmiany przybliża rzeczywistą zmianę Δf wartości funkcji.

odpowiada pomnożeniu wektora przesunięcia przez -1. Oczywiście jest też kierunek, a nawet dwa, różniące się znów o π, w którym funkcja nie rośnie wcale: są to takie kierunki, że wektorek przesunięcia (dx, dy) jest równoległy (w danym punkcie, tu w $(1,1)$) do gradientu[66] funkcji, czyli do $(f_x(1,1), f_y(1,1))$.
c) Odwzorowanie $\mathbb{R}^2_+ \longrightarrow \mathbb{R}^3$ to po prostu trzy osobne odwzorowania $\mathbb{R}^2_+ \longrightarrow \mathbb{R}$ umieszczone na trzech "pięterkach". Na każdym pięterku z osobna robimy to samo, co w zadaniu 12. Tak więc piszemy

$$F\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} f_1(x,y) \\ f_2(x,y) \\ f_3(x,y) \end{bmatrix} = \begin{bmatrix} x\sin(x+2y) \\ y^2\cos(2x-y) \\ y^{2x} \end{bmatrix},$$

i obliczamy z definicji najpierw pochodną kierunkową funkcji $f_1(x,y)$:

$$\begin{aligned}
\nabla_{\mathbf{n}} f_1 &= \lim_{h\to 0}\frac{1}{h}\left[\left(x+\frac{h}{\sqrt{2}}\right)\sin\left(x+\frac{h}{\sqrt{2}}+2y+\frac{2h}{\sqrt{2}}\right) - x\sin(x+2y)\right] \\
&= \lim_{h\to 0}\frac{1}{h}\left\{\left(x+\frac{h}{\sqrt{2}}\right)\left[\sin(x+2y)\cos\frac{3h}{\sqrt{2}} + \cos(x+2y)\sin\frac{3h}{\sqrt{2}}\right] - x\sin(x+2y)\right\} \\
&= \lim_{h\to 0}\frac{1}{h}\left[x\cos(x+2y)\sin\frac{3h}{\sqrt{2}} + \frac{h}{\sqrt{2}}\sin(x+2y) + \mathcal{O}(h^2)\right] \\
&= \frac{1}{\sqrt{2}}\sin(x+2y) + \frac{3}{\sqrt{2}}x\cos(x+2y)\,.
\end{aligned}$$

Widać, że jest to to samo, co

$$n_1\frac{\partial f_1}{\partial x} + n_2\frac{\partial f_1}{\partial y} = \frac{1}{\sqrt{2}}\left[\sin(x+2y) + x\cos(x+2y)\right] + \frac{1}{\sqrt{2}}2x\cos(x+2y)\,.$$

Analogicznie postępujemy z drugim pięterkiem, czyli funkcją $f_2(x,y)$. Pochodna kierunkowa trzeciego pięterka jest może trochę trudniejsza

$$\begin{aligned}
\nabla_{\mathbf{n}} f_3 &= \lim_{h\to 0}\frac{1}{h}\left[\left(y+\frac{h}{\sqrt{2}}\right)^{2(x+h/\sqrt{2})} - y^{2x}\right] \\
&= \lim_{h\to 0}\frac{1}{h}\left[\exp\left\{2\left[x+\frac{h}{\sqrt{2}}\right]\left[\ln y + \ln\left(1+\frac{h}{\sqrt{2}\,y}\right)\right]\right\} - y^{2x}\right] \\
&= \lim_{h\to 0}\frac{1}{h}\left[e^{2x\ln y}e^{\sqrt{2}\,h\ln y}\exp\left\{2\left(x+\frac{h}{\sqrt{2}}\right)\left(\frac{h}{\sqrt{2}\,y}+\dots\right)\right\} - y^{2x}\right] \\
&= y^{2x}\lim_{h\to 0}\frac{1}{h}\left[(1+\sqrt{2}\,h\ln y+\dots)\left\{1+2\left(x+\frac{h}{\sqrt{2}}\right)\left(\frac{h}{\sqrt{2}\,y}+\dots\right)\right\} - 1\right] \\
&= y^{2x}\left(\sqrt{2}\,\ln y + \sqrt{2}\,\frac{x}{y}\right).
\end{aligned}$$

[66] Tak zwykły fizyk zwykle nazywa pochodną prawdziwą funkcji $f : \mathbb{R}^n \longrightarrow \mathbb{R}$. Znów tu mącimy bo po pierwsze, co to znaczy, że dwa wektory są równoległe? A po drugie, jak kowektor może być równoległy do wektora? No właśnie to są skutki tego, że w $\mathbb{R}^n$ wszystko się zlewa w jedno...

Znów jest to to samo, co

$$n_1 \frac{\partial f_3}{\partial x} + n_2 \frac{\partial f_3}{\partial y} = \frac{1}{\sqrt{2}}\, 2\, e^{2x} \ln y + \frac{1}{\sqrt{2}}\, 2\, e^{2x}\, \frac{x}{y}\,.$$

Ostatecznie więc pochodna kierunkowa w kierunku $\mathbf{n}$ obliczona w dowolny punkcie (x, y) jest wektorem

$$\nabla_{\mathbf{n}} F = \begin{bmatrix} \nabla_{\mathbf{n}} f_1 \\ \nabla_{\mathbf{n}} f_2 \\ \nabla_{\mathbf{n}} f_3 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} \sin(2x+y) + \frac{3}{\sqrt{2}}\, x \cos(2x+y) \\ \sqrt{2}\, y \cos(2x-y) - \frac{1}{\sqrt{2}}\, y^2 \sin(2x-y) \\ y^{2x} \left(\sqrt{2}\, \ln y + \sqrt{2}\, \frac{x}{y} \right) \end{bmatrix},$$

który, aby uzyskać główną liniową część zmiany (przyrostu) wektorowej wartości funkcji przy przesunieciu się na płaszczyźnie $\mathbb{R}^2_+$ z punktu (x, y) do punktu $(x + n_1\delta\lambda,\, y + n_2\delta\lambda)$, należy pomnożyć przez $\delta\lambda$.

d) W tym przypadku pochodna w konkretnym punkcie jest odwzorowaniem liniowym reprezentowanym macierzą

$$F'(1,-1) = \begin{pmatrix} f^1_x & f^1_y \\ f^2_x & f^2_y \end{pmatrix}_{(1,-1)} = \begin{pmatrix} 2x & -2y \\ e^{x-y} & -e^{x-y} \end{pmatrix}_{(1,-1)} = \begin{pmatrix} 2 & 2 \\ e^2 & -e^2 \end{pmatrix}.$$

W działaniu na wektor przesunięcia $(dx, dy) = (-\delta\lambda,\, 2\delta\lambda)$ daje ona różniczkę funkcji,

$$F'(1,-1) \cdot \begin{pmatrix} -\delta\lambda \\ 2\,\delta\lambda \end{pmatrix} = \begin{pmatrix} 2\,\delta\lambda \\ -3\delta\lambda\, e^2 \end{pmatrix},$$

czyli główną liniową część zmiany wektora-wartości odwzorowania F przy przesunięciu się z punktu $(1, -1)$ do punktu $(1 + \delta\lambda,\, -1 - 2\delta\lambda)$.

e) Tu też pochodna w konkretnym punkcie jest odwzorowaniem liniowym reprezentowanym macierzą

$$\begin{aligned} F'(1,3,0) &= \begin{pmatrix} f^1_x & f^1_y & f^1_z \\ f^2_x & f^2_y & f^2_z \end{pmatrix}_{(1,3,0)} \\ &= \begin{pmatrix} x/\sqrt{x^2+z^2} & 0 & z/\sqrt{x^2+z^2} \\ 0 & 2y/(1+y^2+z^2) & 2z/(1+y^2+z^2) \end{pmatrix}_{(1,3,0)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3/5 & 0 \end{pmatrix}. \end{aligned}$$

W działaniu na wektor przesunięcia $(dx, dy, dz) = (0,\, \delta\lambda,\, \delta\lambda)$ daje ona różniczkę funkcji,

$$F'(1,3,0) \cdot \begin{pmatrix} \delta\lambda \\ \delta\lambda \end{pmatrix} = \begin{pmatrix} 0 \\ 3\delta\lambda/5 \end{pmatrix},$$

czyli główną liniową część zmiany wektora-wartości odwzorowania F przy przesunięciu się z punktu $(1, 3, 0)$ do punktu $(1,\, 3 + \delta\lambda,\, \delta\lambda)$.

Zadanie 15: W $\mathbb{R}^2$ punkt przez który musi przechodzić prosta ma współrzędne $(x_0,\, f(x_0))$. Prostą przechodzącą przez taki punkt najłatwiej zadać parametrycznie, tj. podając dwie składowe wektora $\mathbf{t}$, do którego prosta ta ma być równoległa. Wiadomo, ze szkoły, że nachylenie prostej stycznej jest dane przez pochodną. Zatem wektor taki ma składowe $[1,\ (df/dx)_{x_0}]$ i równanie prostej można napisać w formie

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x_0 \\ f(x_0) \end{pmatrix} + \begin{bmatrix} 1 \\ (df/dx)_{x_0} \end{bmatrix} \xi\,,$$

gdzie $\xi \in (-\infty,\ \infty)$ jest rzeczywistym parametrem. (Jak ktoś chce, to może z równania $x = x_0 + \xi$ wyrazić $\xi = x - x_0$ i wstawić do drugiego równania uzyskując równanie prostej w formie szkolnej, tj. $ax + by + c = 0$.)

Uogólnienie na $\mathbb{R}^3$ jest proste. Jeśli ustalimy $y = y_0$ to mamy, zmieniając x, funkcję $y = f(x, y_0)$ i płaszczyzna styczna w (x_0, y_0) do powierzchni będącej wykresem $y = f(x, y)$ musi być rówoległa do wektora $\mathbf{t}_1 = [1,\ 0,\ (\partial f/\partial x)_{(x_0,y_0)}]$, który w kierunku x-owym odgrywa tę samą rolę, co poprzedno wektor $[1,\ (df/dx)_{x_0}]$. Analogicznie jest w kierunku y-owym mamy wektor $\mathbf{t}_2 = [0,\ 1,\ (\partial f/\partial y)_{(x_0,y_0)}]$. Zatem parametrycznie zadajemy płaszczyzne styczna w (x_0, y_0) do powierzchni będącej wykresem $y = f(x, y)$ wzorami

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \\ f(x_0, y_0) \end{pmatrix} + \begin{bmatrix} 1 \\ 0 \\ (\partial f/\partial x)_{(x_0,y_0)} \end{bmatrix} \xi_1 + \begin{bmatrix} 0 \\ 1 \\ (\partial f/\partial y)_{(x_0,y_0)} \end{bmatrix} \xi_2\,,$$

gdzie $\xi_1, \xi_2 \in (-\infty,\ \infty)$ są dwoma rzeczywistymi parametrami. (Znów można z równań $x = x_0 + \xi_1$ i $y = y_0 + \xi_2$ wyrazić $\xi_1 = x - x_0$, $\xi_2 = y - y_0$ i wstawić do trzeciego równania uzyskując równanie płaszczyzny w formie szkolnej, tj. $ax + by + cz + d = 0$.)

Nietrudno teraz już zadać parametrycznie w $\mathbb{R}^4$ hiperpłaszczyznę styczną do "wykresu" funkcji $t = f(x, y, z)$ sa teraz trzy wektory $\mathbf{t}_1$, $\mathbf{t}_2$ i $\mathbf{t}_2$:

$$\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \\ z_0 \\ f(x_0, y_0, z_0) \end{pmatrix} \\ + \begin{bmatrix} 1 \\ 0 \\ 0 \\ (\partial f/\partial x)_{(x_0,y_0,z_0)} \end{bmatrix} \xi_1 + \begin{bmatrix} 0 \\ 1 \\ 0 \\ (\partial f/\partial y)_{(x_0,y_0,z_0)} \end{bmatrix} \xi_2 + \begin{bmatrix} 0 \\ 0 \\ 1 \\ (\partial f/\partial z)_{(x_0,y_0,z_0)} \end{bmatrix} \xi_3\,,$$

gdzie $\xi_1, \xi_2, \xi_3 \in (-\infty,\ \infty)$ są trzema rzeczywistymi parametrami. Zatem w przypadku podanej funkcji hiperpłaszczyzna jest zadana parametrycznie wzorami

$$\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 18 \end{pmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \\ 2 \end{bmatrix} \xi_1 + \begin{bmatrix} 0 \\ 1 \\ 0 \\ 12 \end{bmatrix} \xi_2 + \begin{bmatrix} 0 \\ 0 \\ 1 \\ 6 \end{bmatrix} \xi_3\,,$$

albo jednym liniowym równaniem $2x + 12y + 6z - t - 26 = 0$.

Zadanie 16: Trzeba obliczyć pochodne $(\partial f/\partial x)$ i $(\partial f/\partial y)$, wstawić do lewej strony i zobaczyć co tam wyjdzie.

$$\begin{aligned}\frac{\partial f}{\partial x} &= \left(\frac{d\phi(t)}{dt}\right)_{t=y/x}\frac{\partial}{\partial x}\left(\frac{y}{x}\right) - 2x = \phi'\left(\frac{y}{x}\right)\left(-\frac{y}{x^2}\right) - 2x\,,\\ \frac{\partial f}{\partial y} &= \left(\frac{d\phi(t)}{dt}\right)_{t=y/x}\frac{\partial}{\partial y}\left(\frac{y}{x}\right) - 2y = \phi'\left(\frac{y}{x}\right)\left(\frac{1}{x}\right) - 2y\,.\end{aligned}$$

Zatem lewa strona równania jest równa

$$x\left[-\frac{y}{x^2}\,\phi'\left(\frac{y}{x}\right) - 2x\right] + y\left[\frac{1}{x}\,\phi'\left(\frac{y}{x}\right) - 2y\right] = -2(x^2+y^2)\,.$$

Zatem w równaniu, które ma spełniać funkcja $f(x,y)$ funkcja $g(x,y)$ po prawe stronie musi być równa właśnie $-2(x^2+y^2)$.

Zadanie 17: Znów wyobrażamy sobie, że $f(x,y) = \tilde{f}(u(x,y),\ v(x,y))$. Pierwsze pochodne $\partial f/\partial x$ i $\partial f/\partial y$ zostały znalezione w zadaniu w tekście. Korzystając z tamtych wzorów piszemy

$$\begin{aligned}\frac{\partial^2 f}{\partial x^2} &= \frac{\partial}{\partial x}\left(\frac{\partial\tilde{f}}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial\tilde{f}}{\partial v}\frac{\partial v}{\partial x}\right)\\ &= \frac{\partial u}{\partial x}\frac{\partial}{\partial x}\left(\frac{\partial\tilde{f}}{\partial u}\right) + \frac{\partial v}{\partial x}\frac{\partial}{\partial x}\left(\frac{\partial\tilde{f}}{\partial v}\right) + \frac{\partial^2 u}{\partial x^2}\frac{\partial\tilde{f}}{\partial u} + \frac{\partial^2 v}{\partial x^2}\frac{\partial\tilde{f}}{\partial v}\\ &= \frac{\partial u}{\partial x}\left[\frac{\partial}{\partial u}\left(\frac{\partial\tilde{f}}{\partial u}\right)\frac{\partial u}{\partial x} + \frac{\partial}{\partial v}\left(\frac{\partial\tilde{f}}{\partial u}\right)\frac{\partial v}{\partial x}\right]\\ &\quad+\frac{\partial v}{\partial x}\left[\frac{\partial}{\partial u}\left(\frac{\partial\tilde{f}}{\partial v}\right)\frac{\partial u}{\partial x} + \frac{\partial}{\partial v}\left(\frac{\partial\tilde{f}}{\partial v}\right)\frac{\partial v}{\partial x}\right] + \frac{\partial^2 u}{\partial x^2}\frac{\partial\tilde{f}}{\partial u} + \frac{\partial^2 v}{\partial x^2}\frac{\partial\tilde{f}}{\partial v}\,.\end{aligned}$$

Po uporządkowaniu:

$$\frac{\partial^2 f}{\partial x^2} = \left(\frac{\partial u}{\partial x}\right)^2\frac{\partial^2\tilde{f}}{\partial u^2} + \left(\frac{\partial v}{\partial x}\right)^2\frac{\partial^2\tilde{f}}{\partial v^2} + \left(\frac{\partial u}{\partial x}\right)\left(\frac{\partial v}{\partial x}\right)\frac{\partial^2\tilde{f}}{\partial u\partial v} + \frac{\partial^2 u}{\partial x^2}\frac{\partial\tilde{f}}{\partial u} + \frac{\partial^2 v}{\partial x^2}\frac{\partial\tilde{f}}{\partial v}\,.$$

W analogiczny sposób dostaniemy

$$\frac{\partial^2 f}{\partial y^2} = \left(\frac{\partial u}{\partial y}\right)^2\frac{\partial^2\tilde{f}}{\partial u^2} + \left(\frac{\partial v}{\partial y}\right)^2\frac{\partial^2\tilde{f}}{\partial v^2} + \left(\frac{\partial u}{\partial y}\right)\left(\frac{\partial v}{\partial y}\right)\frac{\partial^2\tilde{f}}{\partial u\partial v} + \frac{\partial^2 u}{\partial y^2}\frac{\partial\tilde{f}}{\partial u} + \frac{\partial^2 v}{\partial y^2}\frac{\partial\tilde{f}}{\partial v}\,.$$

Trzeba teraz prawe strony wyrazić przez u i v. Pierwsze pochodne $u_x = (\partial u\partial x)$, etc. zostały przez u i v wyrażone (na trzy sposoby) w zadaniu w tekście. Trzeba jeszcze tylko wyrazić drugie pochodne u_{xx}, i v_{xx}. Oczywiście można to zrobić wypisując jawnie wzory na $u = u(x,y)$, $v = v(x,y)$, dwukrotnie je różniczkując po x i potem wyrażając zurück przez zmienne u i v. Ale bardziej pouczającee jest zastosować metodę "termodynamiczną". W zadaniu w tekście otrzymaliśmy dwa związki

$$
\begin{aligned}
1 &= \frac{v^2-u^2}{R^4}\,u_x - \frac{2uv}{R^4}\,v_x,\\
0 &= \frac{u^2-v^2}{R^4}\,v_x - \frac{2uv}{R^4}\,u_x,
\end{aligned}
$$

w których $R = R(u(x,y), v(x,y))$. Drugie z nich można od razu pomnożyć stronami przez R^4, co uprości robotę. Możemy je zróżniczkować jeszcze raz stronami po x co da

$$
\begin{aligned}
0 &= (v^2-u^2)\,u_{xx} - 2uvv_{xx} - 2u(u_x^2+v_x^2)\\
&\quad - \frac{2}{R^2}\left[(v^2-u^2)\,u_x - 2uvv_x\right](2uu_x + 2vv_x)\,,\\
0 &= (u^2-v^2)\,v_{xx} - 2uvu_{xx} - 2v(u_x^2+v_x^2)\,.
\end{aligned}
$$

Ponieważ u_x i v_x już mamy wyrażone przez u i v, dwa te równania są liniowymi równaniami na u_{xx} i v_{xx}. Wstawiając do nich $u_x = v^2 - u^2$ i $v_x = -2uv$ (korzystamy z wyników zadania w tekscie) otrzymujemy równania, które możemy zapisać tak

$$
\begin{pmatrix} u^2-v^2 & 2uv \\ -2uv & u^2-v^2 \end{pmatrix}
\begin{pmatrix} u_{xx} \\ v_{xx} \end{pmatrix}
=
\begin{pmatrix} 2uR^4 \\ 2vR^4 \end{pmatrix}.
$$

Wyznacznik macierzy po lewej stronie jest równy R^4 i mamy

$$
\begin{pmatrix} u_{xx} \\ v_{xx} \end{pmatrix}
= \frac{1}{R^4}
\begin{pmatrix} u^2-v^2 & -2uv \\ 2uv & u^2-v^2 \end{pmatrix}
\begin{pmatrix} 2uR^4 \\ 2vR^4 \end{pmatrix}
=
\begin{pmatrix} 2u(u^2-3v^2) \\ 2v(3u^2-v^2) \end{pmatrix}.
$$

Oczywiście ponieważ w tym przypadku jest łatwo odkręcić wzory i dostać

$$
u = \frac{x}{x^2+y^2} \equiv \frac{x}{\kappa^2}\,, \qquad v = \frac{y}{x^2+y^2} \equiv \frac{y}{\kappa^2}\,,
$$

łatwo jest znaleźć u_{xx} i v_{xx} "na piechotę. $u_x = (y^2-x^2)/\kappa^4$, $v_x = -2xy/\kappa^4$ i (pamiętamy, że $R^2 = 1/\kappa^2$)

$$
\begin{aligned}
u_{xx} &= -\frac{2x}{\kappa^4} - \frac{2(y^2-x^2)\,2x}{\kappa^6} = \frac{2x(x^2-3y^2)}{\kappa^6} = 2u(u^2-3v^2)\,,\\
v_{xx} &= -\frac{2y}{\kappa^4} + \frac{2xy\,2x}{\kappa^6} = \frac{2y(3x^2-y^2)}{\kappa^6} = 2v(3u^2-v^2)\,,
\end{aligned}
$$

tak jak poprzednio.

W zupełnie analogiczny sposób można otrzymać równania wyznaczające u_{yy} i v_{yy} różniczkując po y związki

$$
\begin{aligned}
0 &= \frac{v^2-u^2}{R^4}\,u_y - \frac{2uv}{R^4}\,v_y,\\
1 &= \frac{u^2-v^2}{R^4}\,v_y - \frac{2uv}{R^4}\,u_y,
\end{aligned}
$$

uzyskane z różniczkowania po y wzorów wiążących x i y z u i v. Dostaniemy po takich samych przekształceniach, jak te wyżej,

$$u_{yy} = 2u(3v^2 - u^2)\,, \qquad v_{yy} = 2v(v^2 - 3u^2)\,.$$

Zbierając wszystko razem otrzymujemy

$$f_{xx} = (v^2 - u^2)^2 \tilde{f}_{uu} + 4u^2v^2 \tilde{f}_{vv} - 2uv(v^2 - u^2)\tilde{f}_{vu} + 2u(u^2 - 3v^2)\tilde{f}_u + 2v(3u^2 - v^2)\tilde{f}_v\,,$$
$$f_{yy} = 4u^2v^2 \tilde{f}_{uu} + (u^2 - v^2)^2 \tilde{f}_{vv} - 2uv(u^2 - v^2)\tilde{f}_{vu} + 2u(3v^2 - u^2)\tilde{f}_u + 2v(v^2 - 3u^2)\tilde{f}_v\,.$$

Jak te dwa wzory dodamy stronami, to się okaże, że

$$f_{xx} + f_{yy} = R^4(u,v)(\tilde{f}_{uu} + \tilde{f}_{vv})\,.$$

Przeszliśmy tu dość długą drogę aby pokazać wszystko w detalach w sytuacji, gdy nie jest tak łatwo odwracać związki łączące dwa zespoły zmiennych. Tu, ponieważ nie jest to trudne, można było jednak zrobić wszystko trochę prościej. Np. wyrażając f_{xx} przez u i v można było do wzoru

$$\frac{\partial^2 f}{\partial x^2} \equiv f_{xx} = \frac{\partial}{\partial x}\left(\frac{\partial \tilde{f}}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial \tilde{f}}{\partial v}\frac{\partial v}{\partial x}\right) \equiv \frac{\partial}{\partial x}\left(u_x\,\tilde{f}_u + v_x\,\tilde{f}_v\right),$$

od razu podstawić $\partial u/\partial x \equiv u_x = (y^2 - x^2)/\kappa^4$, $\partial v/\partial x \equiv v_x = -2xy/\kappa^4$. Moglibyśmy wtedy u_x i v_v różniczkować po x jawnie i dopiero na końcu wyrazić rezultat tych różniczkowań przez zmienne u i v:

$$f_{xx} = -\frac{4x}{\kappa^6}\left[(y^2 - x^2)\,\tilde{f}_u - 2xy\,\tilde{f}_v\right] + \frac{1}{\kappa^4}\left[-2x\,\tilde{f}_u - 2y\,\tilde{f}_v + (y^2 - x^2)\,\frac{\partial \tilde{f}_u}{\partial x} - 2xy\,\frac{\partial \tilde{f}_v}{\partial x}\right].$$

Po małym uporządkowaniu daje to

$$\begin{aligned} f_{xx} = &\frac{1}{\kappa^6}\left[2x(x^2 - 3y^2)\,\tilde{f}_u + 2y(3x^2 - y^2)\,\tilde{f}_v\right] \\ &+\frac{1}{\kappa^4}\left[(y^2 - x^2)\left(u_x\,\tilde{f}_{uu} + v_x\,\tilde{f}_{uv}\right) - 2xy\left(u_x\,\tilde{f}_{vu} + v_x\,\tilde{f}_{vv}\right)\right], \end{aligned}$$

co po podstawieniu znów $u_x = (y^2 - x^2)/\kappa^4$, $v_x = -2xy\kappa^4$ i wyrażeniu x i y przez zmienne u i v (co sprowadza się do zastąpienia x przez u, a y przez v i położeniu $\kappa = 1$ - dlaczego? już to przerabialiśmy!) da na f_{xx} ten sam wzór, co wypisany już wyżej.

Zadanie 18: Piszemy

$$\begin{aligned} \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) &= \frac{\partial}{\partial y}\left(\frac{\partial \tilde{f}}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial \tilde{f}}{\partial v}\frac{\partial v}{\partial x}\right) \\ &= u_x\left(u_y\,\tilde{f}_{uu} + v_y\,\tilde{f}_{uv}\right) + u_{xy}\,\tilde{f}_u \\ &\quad +v_x\left(u_y\,\tilde{f}_{vu} + v_y\,\tilde{f}_{vv}\right) + v_{xy}\,\tilde{f}_v\,. \end{aligned}$$

Druga zaś pochodna, to

$$
\begin{aligned}
\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) &= \frac{\partial}{\partial x}\left(\frac{\partial \tilde{f}}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial \tilde{f}}{\partial v}\frac{\partial v}{\partial y}\right) \\
&= u_y\left(u_x\,\tilde{f}_{uu} + v_x\,\tilde{f}_{uv}\right) + u_{yx}\,\tilde{f}_u \\
&\quad + v_y\left(u_x\,\tilde{f}_{vu} + v_x\,\tilde{f}_{vv}\right) + v_{yx}\,\tilde{f}_v\,.
\end{aligned}
$$

Widać, że jest to to samo.

Aby jawnie podać pochodną mieszaną f_{xy} wyrażoną przez u i v musimy znaleźć pochodne $u_{xy} = u_{yx}$ i $v_{xy} = v_{yx}$ wyrażone przez te właśnie zmienne (wszystkie inne potrzebne elementy już mamy). Znów, tu jest to łatwe, bo możemy łatwo jawnie napisać wzory $u = x/\kappa^2(x,y)$ i $v = y/\kappa^2(x,y)$. Ale znów zrobimy to okrężną drogą, czyli "sposobem termodynamicznym", aby zobaczyć jak sobie radzić, gdy nie da się łatwo napisać wzorów $u = u(x,y)$ i $v = v(x,y)$ (a znane są wzory dające $x = x(u,v)$ i $y = y(u,v)$. Mamy z poprzedniego zadania dwa układy równań

$$
\begin{aligned}
1 &= \frac{v^2-u^2}{R^4}\,u_x - \frac{2uv}{R^4}\,v_x, \\
0 &= \frac{u^2-v^2}{R^4}\,v_x - \frac{2uv}{R^4}\,u_x,
\end{aligned}
$$

otrzymany ze zróżniczkowania tożsamości $x = x(u(x,y),\, v(x,y))$ i $y = y(u(x,y),\, v(x,y))$ po x, oraz drugi układ

$$
\begin{aligned}
0 &= \frac{v^2-u^2}{R^4}\,u_y - \frac{2uv}{R^4}\,v_y, \\
1 &= \frac{u^2-v^2}{R^4}\,v_y - \frac{2uv}{R^4}\,u_y,
\end{aligned}
$$

otrzymany przez zróżniczkowanie tożsamości $x = x(u(x,y), v(x,y))$ i $y = y(u(x,y), v(x,y))$ po y. Każdy z tych dwu układów może nam posłużyć do wyznaczenia u_{xy} i v_{xy}. Powinny one dać to samo. Posłużymy się tu pierwszym z nich (zalecając studentom samodzielne wyprowadzenie wzorów na u_{xy} i v_{xy} z drugiego układu równań). Różniczkujemy więc pierwsze dwa równania stronami po y co daje

$$
\begin{aligned}
0 &= (v^2-u^2)u_{xy} - 2uvv_{xy} + 2v\,(u_xv_y - u_yv_x) - 2u(u_xu_y + v_xv_y) \\
&\quad - \frac{2}{R^2}\left[(v^2-u^2)u_x - 2uvv_x\right](2uu_y + 2vv_y)\,, \\
0 &= (u^2-v^2)v_{xy} - 2uvu_{xy} + 2u\,(u_yv_x - u_xv_y) - 2v(u_xu_y + v_xv_y)\,.
\end{aligned}
$$

Podstawiamy tu teraz $u_x = (y^2-x^2)/\kappa^4$, $v_x = -2xy/\kappa^4$ oraz $u_y = -2xy/\kappa^4$, $v_y = (x^2-y^2)/\kappa^4$ albo lepiej $u_x = v^2-u^2$, $v_x = -2uv$, $u_y = -2uv$, $v_y = u^2-v^2$ (zobacz zadanie w tekście), co od razu uwidacznia, że $u_xu_y + v_xv_y = 0$, a $u_xv_y - u_yv_x = -R^4$.

Dostajemy stąd liniowy układ dwóch równań na u_{xy} i v_{xy}, kt'ory znów zapiszemy w postaci macierzowej

$$\begin{pmatrix} u^2 - v^2 & 2uv \\ -2uv & u^2 - v^2 \end{pmatrix} \begin{pmatrix} u_{xy} \\ v_{xy} \end{pmatrix} = \begin{pmatrix} 2vR^4 \\ -2uR^4 \end{pmatrix}.$$

Zatem

$$\begin{pmatrix} u_{xy} \\ v_{xy} \end{pmatrix} = \frac{1}{R^4} \begin{pmatrix} u^2 - v^2 & -2uv \\ 2uv & u^2 - v^2 \end{pmatrix} \begin{pmatrix} 2vR^4 \\ -2uR^4 \end{pmatrix} = \begin{pmatrix} 2v\,(3u^2 - v^2) \\ 2u(3v^2 - u^2) \end{pmatrix}.$$

To samo można oczywiście było dostać różniczkując bezpośrednio $u_x = (y^2 - x^2)/\kappa^4$ oraz $v_x = -2xy/\kappa^4$ po y i wyrażając rezultat tych różniczkowań przez zmienne u i v. Zatem

$$\begin{aligned} f_{xy} &= 2uv(u^2 - v^2)\left(\tilde{f}_{uu} - \tilde{f}_{vv}\right) - (u^4 - 6u^2v^2 + v^4)\tilde{f}_{uv} \\ &\quad +2v(3u^2 - v^2)\tilde{f}_u + 2u(3v^2 - u^2)\tilde{f}_v\,. \end{aligned}$$

Zadanie 19: Niech $z = x^2 - y^2$. Obliczamy pochodne funkcji $f(x,y) = y\,\phi(x^2 - y^2)$:

$$\begin{aligned} \frac{\partial f}{\partial x} &= \frac{\partial}{\partial x}\,(y\,\phi(z(x,y)) = y\,\left.\frac{d\phi(z)}{dz}\right|_{z=x^2-y^2}\frac{\partial z}{\partial x} = 2xy\,\phi'(x^2 - y^2)\,, \\ \frac{\partial f}{\partial y} &= \frac{\partial}{\partial y}\,(y\,\phi(z(x,y)) = \phi(x^2 - y^2) + y\,\left.\frac{d\phi(z)}{dz}\right|_{z=x^2-y^2}\frac{\partial z}{\partial y} \\ &= \phi(x^2 - y^2) - 2y^2\,\phi'(x^2 - y^2)\,. \end{aligned}$$

Wykorzystując te pochodne w podanej kombinacji, znajdujemy

$$\begin{aligned} \frac{1}{x}\frac{\partial f}{\partial x} + \frac{1}{y}\frac{\partial f}{\partial y} &= 2y\,\phi'(x^2 - y^2) + \frac{1}{y}\,\phi(x^2 - y^2) - 2y\,\phi'(x^2 - y^2) \\ &= \frac{1}{y}\,\phi(x^2 - y^2) = \frac{1}{y^2}\,f(x,y)\,. \end{aligned}$$

Zatem funkcja $g(x,y) = 1/y^2$.

Zadanie 20: Wprowadźmy oznaczenia $u = z/x$ i $v = y/x$. Obliczamy następnie pochodne:

$$\begin{aligned} \frac{\partial H}{\partial x} &= 2x\,F\Big(\frac{z}{x}, \frac{y}{x}\Big) + x^2\left[\frac{\partial F(u,v)}{\partial u}\right]_{u=z/x,v=y/x}\frac{\partial u}{\partial x} + x^2\left[\frac{\partial F(u,v)}{\partial v}\right]_{u=z/x,v=y/x}\frac{\partial v}{\partial x} \\ &= 2x\,F\Big(\frac{z}{x}, \frac{y}{x}\Big) + x^2\,[F_x(u,v)]_{u=z/x,v=y/x}\left(-\frac{z}{x^2}\right) + x^2\,[F_v(u,v)]_{u=z/x,v=y/x}\left(-\frac{y}{x^2}\right) \\ &= 2x\,F\Big(\frac{z}{x}, \frac{y}{x}\Big) - z\,[F_x(u,v)]_{u=z/x,v=y/x} - y\,[F_v(u,v)]_{u=z/x,v=y/x}\,. \end{aligned}$$

I analogicznie

$$\frac{\partial H}{\partial y} = x^2 \left[F_v(u,v)\right]_{u=z/x, v=y/x} \frac{1}{x} = x \left[F_v(u,v)\right]_{u=z/x, v=y/x} \,,$$
$$\frac{\partial H}{\partial z} = x^2 \left[F_u(u,v)\right]_{u=z/x, v=y/x} \frac{1}{x} = x \left[F_u(u,v)\right]_{u=z/x, v=y/x} \,.$$

Gdy wiec wstawimy te pochodne do kombinacji

$$x\,\frac{\partial H}{\partial x} + y\,\frac{\partial H}{\partial y} + z\,\frac{\partial H}{\partial z} \,,$$

wszystkie wyrazy z F_u i F_v ulegną redukcji o kombinacja ta okaże się równa $2x^2F(z/x, y/x)$. Zatem w równaniu spełnianym przez H funkcja $g(x,y,z)$ musi być stała i równa 2.

Zadanie 21:

$$a) \quad \sin(x+y) = x + y - \frac{1}{6}\,x^3 - \frac{1}{2}\,x^2y - \frac{1}{2}\,xy^2 - \frac{1}{6}\,y^3 + \ldots,$$
$$b) \quad e^{x^2}\cos y = 1 + x^2 - \frac{1}{2}\,y^2 + \frac{1}{2}\,x^4 - \frac{1}{2}\,x^2y^2 + \frac{1}{24}\,y^4 + \ldots,$$
$$c) \quad \ln(1+x+2y) = x + 2y - \frac{1}{2}\,x^2 - 2xy - 2y^2 + \frac{1}{3}\,x^3 + 2x^2y + 4xy^2 + \frac{8}{3}\,y^3 + \ldots,$$
$$d) \quad x + y + \sqrt{x^2+y^2} = 2 + 2(x-1) + y + \frac{1}{2}\,y^2 - \frac{1}{2}\,(x-1)y^2 + \ldots,$$
$$d) \quad x^y = 1 + (x-1)\,y - \frac{1}{2}\,(x-1)^2y + \frac{1}{3}\,(x-1)^2y + \frac{1}{2}\,(x-1)^2y^2 + \ldots$$

Zadanie 22:

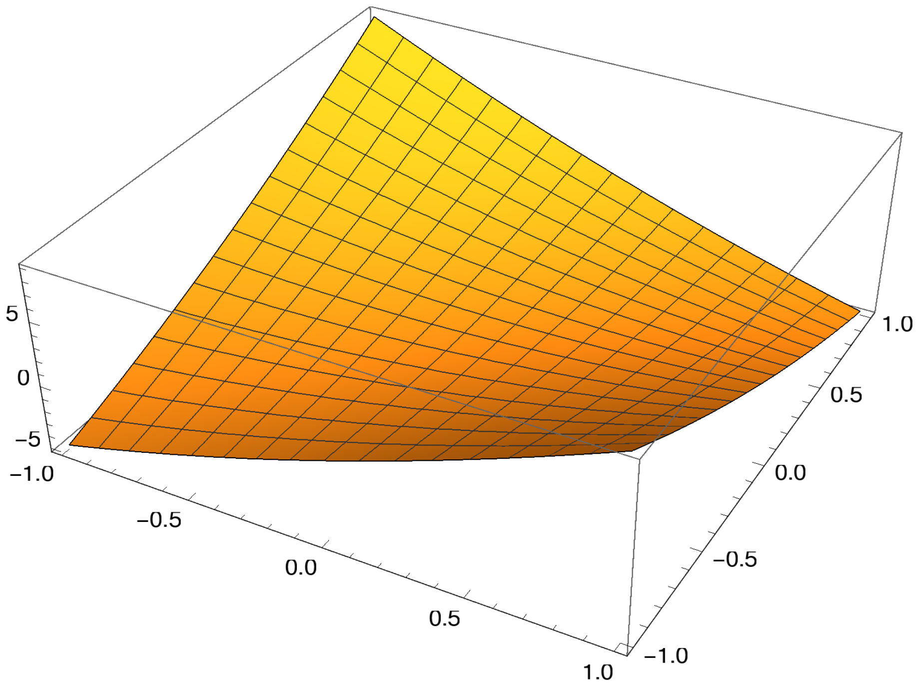
$a)$ Równania $f_x = 2x - \lambda y =$, $f_y = -\lambda x + 2y = 0$ mają tylko jedno rozwiązanie $x = 0 = 0$. Punktem krytycznym jest więc tylko punkt $(0,0)$. Drugie pochodne są stałe (takie same w każdym punkcie). Forma kwadratowa drugich pochodnych jest więc wszędzie taka sama, więc i w punkcie krytycznym też ma ona postać

$$\begin{pmatrix} 2 & -\lambda \\ -\lambda & 2 \end{pmatrix}.$$

Jej minory: $M_{11} = 2 > 0$, $M_{22} = 4 - \lambda^2$. Zatem jeśli $|\lambda| < 2$, funkcja $= x^2 + y^2 - \lambda xy$ ma w punkcie $(0,0)$ minimum lokalne. Jest też jasne, że ekstremum, jeśli istnieje, musi być w punkcie $(0,0)$, bo badana funkcja jest funkcją jednorodną stopnia drugiego, tzn. taką, że

$$f(\xi x,\, \xi y) = \xi^2\, f(x,y)\,.$$

Widać stąd, że gdy $|\xi| > 1$, wartość funkcji w punkcie $(\xi x,\, \xi y)$ jest większa niż w (x,y), a gdy $|\xi| < 1$ mniejsza - funkcja nie może więc mieś w (x,y) lokalnego ekstremum; wyjątkiem jest punkt $(0,0)$ bo wtedy $(\xi \cdot 0,\, \xi \cdot 0)$ to jest ten sam punkt i argument nie działa.



Rysunek 20: Kształt funkcji z Zadania 22a, gdy $\lambda = 7$. W punkcie $(0,0)$ można się domyślić punktu siodłowego.

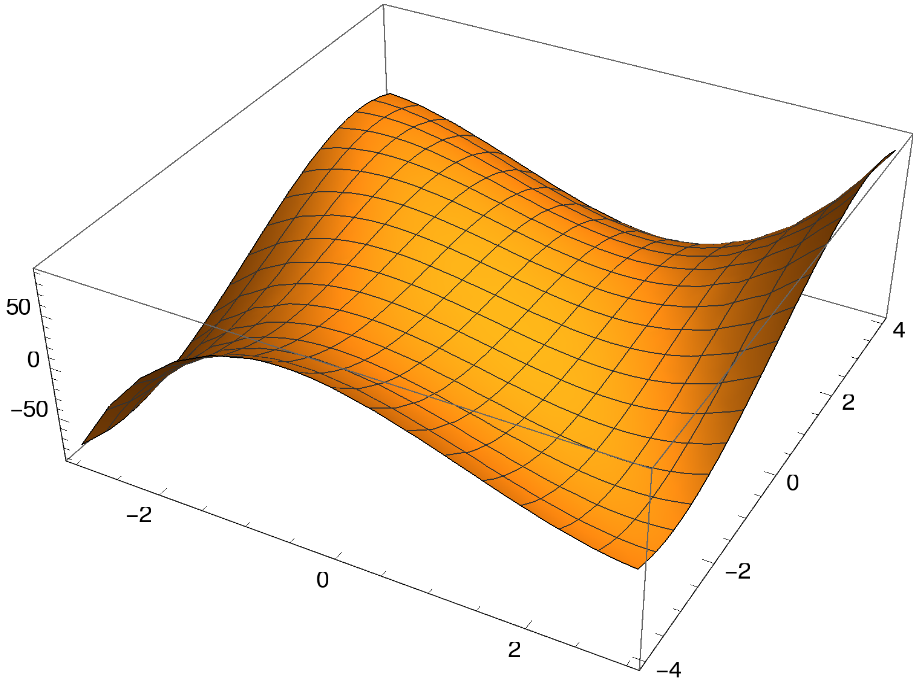
Jeśli $|\lambda| > 2$, drugi minor jest ujemny i forma kwadratowa drugich pochodnych, jest nieokreślona - w punkcie $(0,0)$ jest punkt siodłowy - co trochę widać (jak już się wie) z rysunku 20. Szczególny przypadek zachodzi, gdy $|\lambda| = 2$: forma kwadratowa drugich pochodnych ma wtedy sygnaturę $(+,0)$, czyli istnieje kierunek płaski na wektorach przesunieć skierowanych wzdłuż tego kierunku forma daje zero, a na wszystkich innych (skierowanych choćby tylko nieznacznie w bok od tego kierunku) daje wartość dodatnia. Zwykle tak nie jest, bo wyrazy trzeciego rzędu w $\mathbf{h}$ "zaginają" jakoś taki płaski kierunek, ale tu, ponieważ cała funkcja f jest tu wielomianem tylko drugiego stopnia, jest tościśle płaski kierunek, bo wtedy po prostu $f(x,y) = (x \pm y)^2$ ($\pm$ zależnie od znaku parametru λ) i jest jasne, że funkcja jest całkowicie stała wzdłuż linii $x = y$ lub $x = -y$.

b) Drugie z równanń

$$f_x(x,y) = 9x^2 + 6xy - 15 = 0\,,$$
$$f_y(x,y) = 3x^2 - 3y^2 = 0\,,$$

wyznaczających punkty krytyczne ma jako rozwiązanie $x = y$ bądź $x = -y$. Jeśli $x = y$, to pierwsze równanie sprowadza się do $15x^2 - 15 = 0$. Istnieją więc dwa punkty krytyczne o $x = y$: $(1,1)$ i $(-1,-1)$. Z kolei jeśli $x = -y$, to pierwsze równanie sprowadza się do $3x^2 - 15 = 0$ i są dwa punkty krytyczne o $x = -y$: $(\sqrt{5},\, -\sqrt{5}\,)$ i $(-\sqrt{5},\, \sqrt{5}\,)$. Drugie pochodne cząstkowe są równe

$$f_{xx} = 18x + 6y\,, \qquad f_{yy} = -6y\,, \qquad f_{xy} = 6\,,$$



Rysunek 21: Kształt funkcji z Zadania 22b. Oś x-ów jest od -3 do 3; oś y-ków od -4 do 4.

i macierze form kwadratowych w kolejnych punktach krytycznych są równe

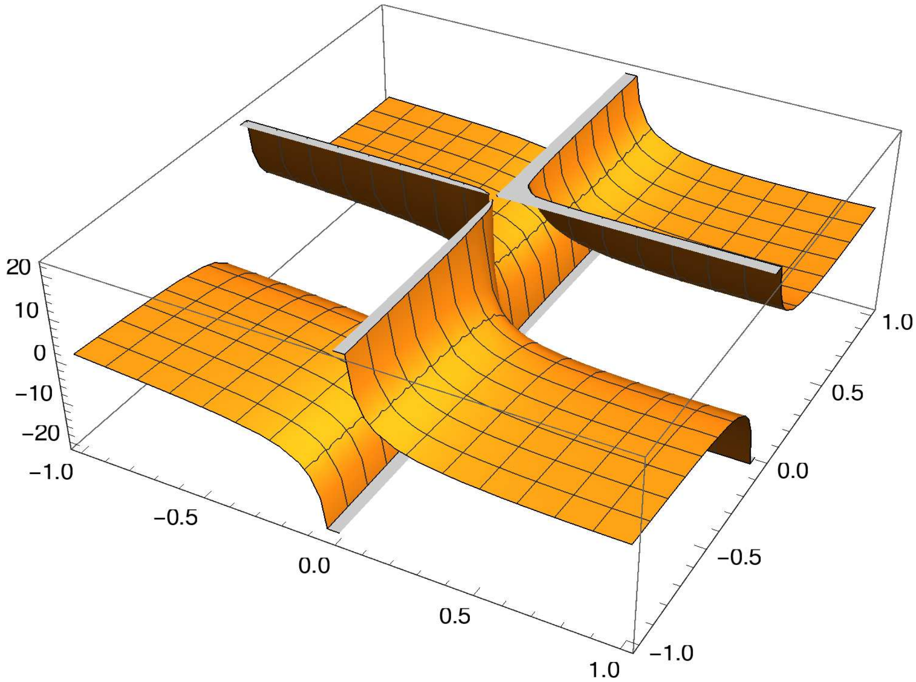
$$\begin{pmatrix} 24 & 6 \\ 6 & -6 \end{pmatrix}, \quad \begin{pmatrix} -24 & -6 \\ -6 & 6 \end{pmatrix}, \quad 2\sqrt{5}\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}, \quad 2\sqrt{5}\begin{pmatrix} -2 & -1 \\ -1 & -1 \end{pmatrix}.$$

Pierwsze dwie formy są nieokreślone, więc punkty $(1,1)$ i $(-1,-1)$ są punktami siodłowymi badanej funkcji. W punkcie $(\sqrt{5},\, -\sqrt{5}\,)$ forma drugich pochodnych jest dodatnio oreślona - w tym punkcie jest więc minimum, i ujemnie określona (bo forma $-Q$ jest tam dodatnio określona) w punkcie $(-\sqrt{5},\, \sqrt{5}\,)$, w którym jest zatem maksimum. Można to chyba dostrzeć z rysunku 21.

c) Tu znalezienie punktów krytycznych wymaga trochę sprytu. Wyznaczające punkty krytyczne równania

$$f_x(x,y) = 2x + y - \frac{a^3}{x^2} = 0\,,$$
$$f_y(x,y) = 2y + x - \frac{a^3}{y^2} = 0\,,$$

są bowiem trochę skomplikowane. Ponieważ funkcja f jest symetryczna w swoich dwóch argumentach, $f(x,y) = f(y,x)$ jest jasne, że rozwiązania tych równań muszą albo być postaci (x^*, x^*), albo występować parami: (x^*, y^*) i (y^*, x^*) (tzn. jeśli rozwiązaniem jest np. (a,b) to rozwiyazaniem musi też być punkt (b,a)). Znajdźmy więc najpierw



Rysunek 22: Kształt funkcji z Zadania 22c, gdy $a = 1$.

rozwiązania typu (x^*, x^*). Oba powyższe równania stają się wtedy tym samym równaniem $3x - a^3/x^2 = 0$ i jedynym takim punktem krytycznym jest punkt $(a/3^{1/3}, a/3^{1/3})$. Drugie pochodne

$$f_{xx} = 2 + 2a^3/x^3\,, \qquad f_{yy} = 2 + 2a^3/y^3\,, \qquad f_{xy} = 1\,,$$

tworzą w tym punkcie formę

$$\begin{pmatrix} 8 & 1 \\ 1 & 8 \end{pmatrix},$$

która jest oczywiście dodatnio określona. W punkcie $(a/3^{1/3}, a/3^{1/3})$ jest zatem minimum lokalne.

Aby pokazać, że innych punktów krytycznych badana funkcja nie ma, uprościmy sobie wpierw wzory dokonując przeskalowania zmiennych $x/a \to x$, $x = y/a \to y$, tak iż funkcja przybierze postać

$$x^2 + xy + y^2 + \frac{a^3}{x} + \frac{a^3}{y} \to a^2 \left(x^2 + xy + y^2 + \frac{1}{x} + \frac{1}{y} \right).$$

Ponieważ multiplikatywna stała a^2 nie ma wpływu na położenie punktów krytycznych, można ją pominąć. Równania wyznaczające punkty krytyczne, po pomnożeniu ich stronami przez odpowiednio x^2 i y^2 mają teraz postać

$$2x^3 + x^2 y = 1\,,$$
$$2y^3 + y^2 x = 1\,.$$

Odejmując jedno od drugiego dostajemy

$$2(x^3 - y^3) + xy(x - y) = (x - y)\left[2(x^2 + xy + y^2) + xy\right] = 0\,.$$

Ponieważ rozwiązanie o $x = y$ już znamy, możemy założyć, że $x \neq y$ i rozwiązywać układ równań

$$\begin{aligned} &2x^3 + x^2 y = 1\,, \\ &2x^2 + 3xy + 2y^2 = 0\,. \end{aligned}$$

Z pierwszego wyznaczamy y: $y = (1 - 2x^3)/x^2$ i wstawiamy do drugiego, co daje

$$2x^2 + 3x\,\frac{1 - 2x^3}{x^2} + 2\frac{(1 - 2x^3)^2}{x^4} = \frac{1}{x^4}\left(2x^6 + 3x^3(1 - 2x^3) + 2(1 - 2x^3)^2\right) = 0\,.$$

Podstawiamy teraz $t = x^3$ i przyrównujemy do zera zawartość nawiasu (czynnik $1/x^4$ nie może być równy zeru):

$$2t^2 + 3t - 6t^2 + 2 - 8t + 8t^2 \equiv 4t^2 - 5t + 2 = 0\,.$$

Ponieważ $\Delta = 25 - 32$ jest ujemna, równanie to nie ma rzeczywistych rozwiąń. Oznacza to, że układ równań $f_x = 0$, $f_y = 0$ nie ma innych rozwiązań niż $x = y = 1/2^{1/3}$ (czyli, w nieprzeskalowanych zmiennych, $a/2^{1/3}$). Potwierdza to rzut oka na rysunek 22. Funkcja $f(x, y)$ nie jest określona wzdłuż osi x i osi y - przy zbieganiu do punktów leżących na tych osiach funkcja dąży do $+\infty$ bądź $-\infty$. Wyjątkiem jest punkt $(0, 0)$ - przy zbieganiu do tego punktu wzdłuż linii $x = y$ granicą funkcji jest zero.

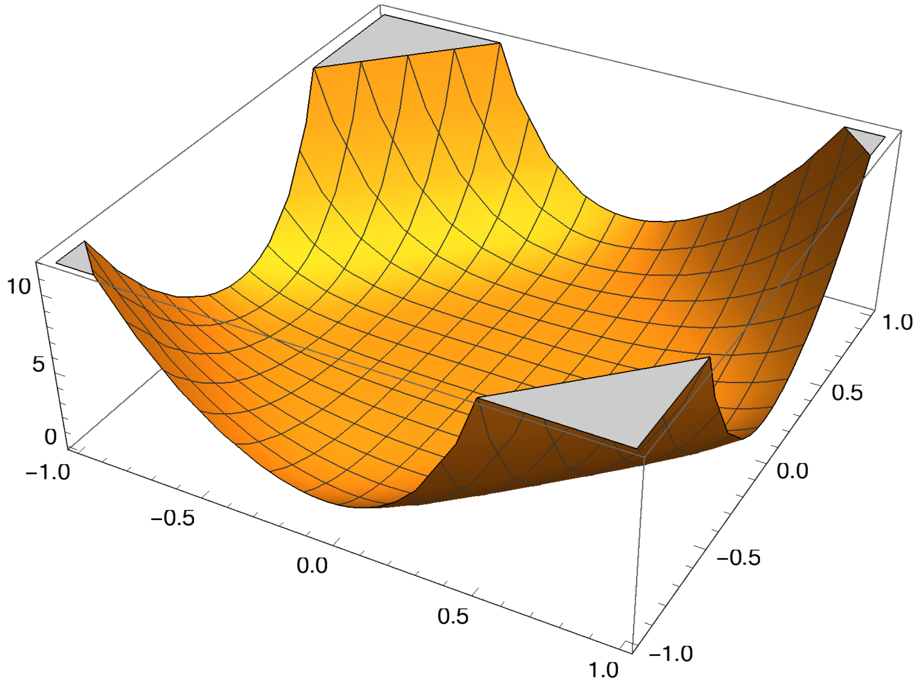
d) Pierwsze pochodne cząstkowe funkcji

$$f_x(x, y) = 4(x + y)^3 + 6(x - y)^5\,, \qquad f_y(x, y) = 4(x + y)^3 - 6(x - y)^5\,,$$

istnieją, są ciągłe na całym $\mathbb{R}^2$ i znikają tylko w punkcie $(0, 0)$ (o czym się można przekonać biorąc sumę $8(x + y)^3 = 0$ i różnicę $12(x - y)^5 = 0$ równań $f_x = 0$ i $f_y = 0$). W tymże punkcie jednak macierz formy kwadratowej drugich pochodnych jest całkowicie zerowa i nie można na jej podstawie określić charakteru punktu krytycznego $(0, 0)$. Niemniej, widać, że $f(0, 0) = 0$, a w dowolnym punkcie $(x, y) \neq (0, 0)$ wartość funkcji jest większa od zera, więc jest tak i w dowolnie małym otoczeniu otwartym punktu krytycznego. W punkcie tym znajduje się zatem minimum badanej funkcji i jest to jej minimum globalne, co widać z rysunku 23.

e) Przyrównanie do zera pierwszych pochodnych cząstkowych tej funkcji

$$\begin{aligned} f_x(x, y) &= 1 - \frac{3y}{x^2 + y^2} + \frac{x}{x^2 + y^2} = \frac{x^2 + x + y^2 - 3y}{x^2 + y^2}\,, \\ f_y(x, y) &= -2 + \frac{3x}{x^2 + y^2} + \frac{y}{x^2 + y^2} = \frac{-2x^2 - 2y^2 + 3x + y}{x^2 + y^2}\,, \end{aligned}$$



Rysunek 23: Kształt funkcji z Zadania 22d.

daje, bo mianowniki nie są istotne, układ równań

$$x^2 + x + y^2 - 3y = 0\,,$$
$$2x^2 + 2y^2 - 3x - y = 0\,.$$

Po pomnożeniu pierwszego przez 2 i odjęciu od drugiego znajduje się, że $x = y$ i wstawienie tego do pierwszego da $x^2 - x = 0$. Pochodne zerują się więc w punktach $(0, 0)$ i $(1, 1)$. Pierwszy z tych punktów leży jednak pozza dziedziną funkcji (bo $\ln 0$ jest wielkością źle określoną, więc jedym przwdziwym punktem krytycznym jest punkt $(1, 1)$.

Drugie pochodne cząstkowe (znów dobre ćwiczenie w ich liczeniu)

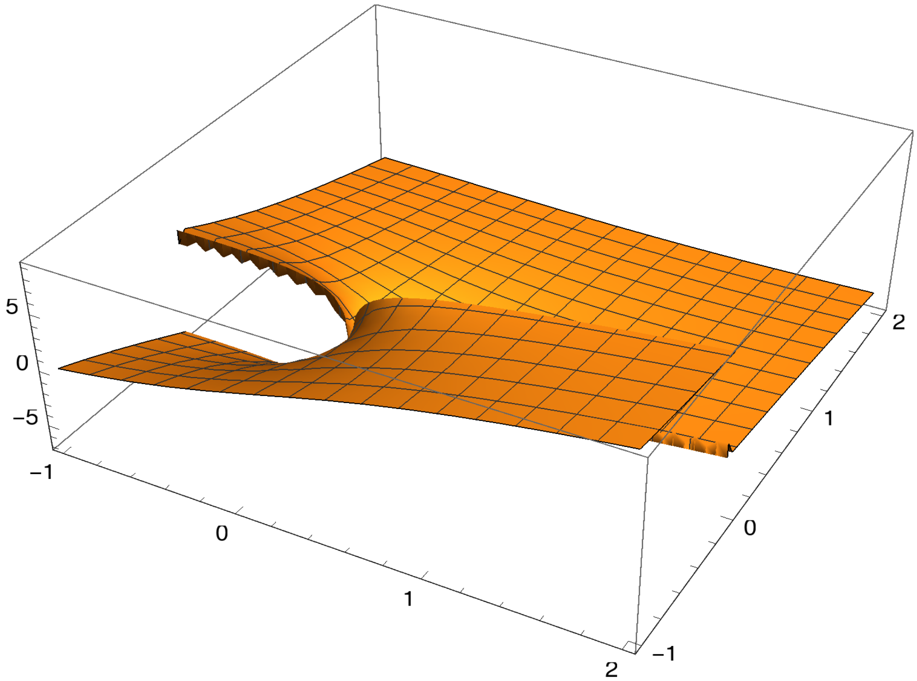
$$f_{xx} = \frac{-x^2 + y^2 + 6xy}{(x^2 + y^2)^2}\,, \qquad f_{yy} = \frac{x^2 - y^2 - 6xy}{(x^2 + y^2)^2}\,, \qquad f_{xy} = \frac{-3x^2 + 3y^2 - 6xy}{(x^2 + y^2)^2}\,,$$

dają w punkcie $(1, 1)$ macierz

$$\frac{1}{4}\begin{pmatrix} 6 & -1 \\ -1 & -6 \end{pmatrix},$$

która nie jest ani dodatnio ani ujemnie określona. W punkcie tym zatem funkcja ma punkt siodłowy. Funkcja poza tym jest, jak widać z rysunku 24, nieciągła na osi y, bo nieciągły jest arctg, który dąży do $\pm\pi/2$, gdy $y \to 0^{\pm}$.

f) Funkcja jest biperiodyczna, tzn. $f(x+k\pi, y+l\pi) = f(x, y)$ przy dowolnych całkowitych k i l, oraz symetryczna $f(x, y) = f(y, x)$. Wystarczy więc znaleźć jej punkty krytyczne



Rysunek 24: Kształt funkcji z Zadania 22e.

leżące nad diagonalą w kwadracie $[0, 2\pi) \times [0, 2\pi)$. Przyrównanie do zera pierwszych pochodnych cząstkowych tej funkcji daje równania

$$f_x(x, y) = \cos(x + y) - \cos x = 0\,, \qquad f_y(x, y) = \cos(x + y) - \cos y = 0\,,$$

równoważne równaniom $\cos x = \cos y = \cos(x + y)$. W obszarze $[0, 2\pi) \times [0, 2\pi)$ pierwsze oznacza albo $x = y$ albo $y = 2\pi - x$ (wystarczy popatrzyć na wykres cosinusa). Druga możliwość oznacza, że $\cos x = 1$ czyli $x = 0$. Pierwsza zaś to $\cos x = \cos 2x$ czyli $\cos x = 2\cos^2 x - 1$. Daje to na $t = \cos x$ równanie kwadratowe

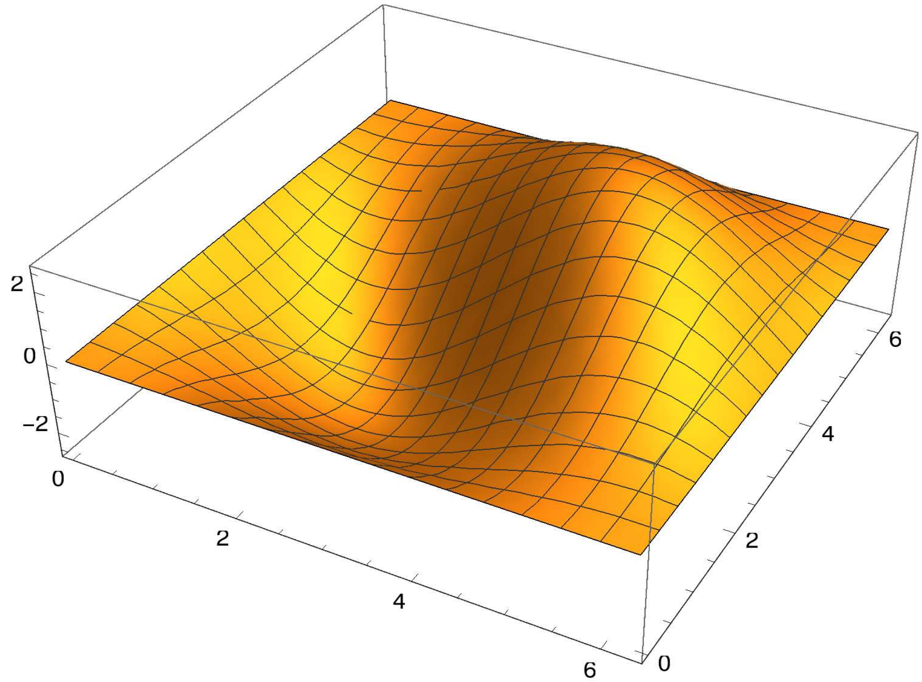
$$2t^2 - t - 1 = 2(t - 1)(t + \frac{1}{2}) = 0\,.$$

Rozwiązanie $\cos x = 1$ czyli $x = y = 0$ jest równoważne $x = 0$ i $y = 2\pi$, bo funkcja jest biperiodyczna. Drugie zaś $\cos x = -\frac{1}{2}$ daje dwie możliwości $x = y = \frac{2}{3}\pi$ i $x = y = \frac{4}{3}\pi$. Zatem w obszarze $[0, 2\pi) \times [0, 2\pi)$ funkcja ma trzy punkty krytyczne: $(0, 0)$, $(\frac{2}{3}\pi,\, \frac{2}{3}\pi)$ i $(\frac{4}{3}\pi,\, \frac{4}{3}\pi)$. Drugie pochodne

$$f_{xx}(x, y) = \sin x - \sin(x + y)\,, \quad f_{yy}(x, y) = \sin y - \sin(x + y)\,, \quad f_{xy}(x, y) = -\sin(x + y)\,,$$

dają w tych trzech punktach następujące formy kwadratowe

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \qquad \frac{\sqrt{3}}{2}\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \qquad -\frac{\sqrt{3}}{2}\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix},$$



Rysunek 25: Kształt funkcji z Zadania 22f w obszarze $[0, 2\pi) \times [0, 2\pi)$.

W punkcie $(\frac{2}{3}\pi,\, \frac{2}{3}\pi)$ funkcja ma zatem minimum lokalne (forma jest dodatnio określona), a w punkcie $(\frac{4}{3}\pi,\, \frac{4}{3}\pi)$ lokalne maksimum. Gorzej jest z punktem $(0, 0)$, bo forma jest zerowa. Można jednak zobaczyć, że jest to punkt siodłowy:

$$f(0 + h_x,\, 0 + h_y) \approx -\frac{1}{2}\, h_x h_y (h_x + h_y)\,,$$

jak wynika z rozwinięć funkcji sinus. Jeśli np. $h_x = h_y = \varepsilon$, to $f(0 + h_x,\, 0 + h_y) < f(0, 0) = 0$; jeśli zaś np. $h_x = \varepsilon > 0$, a $h_y = -\frac{1}{2}\varepsilon$, to $f(0 + h_x,\, 0 + h_y) = \frac{1}{8}\varepsilon^3 > f(0, 0) = 0$. Wszystkie te ustalenia potwierdza wykres funkcji pokazany na rysunku 25.

g) Pierwsze pochodne cząstkowe funkcji $f(x, y) = x^4 - y^4 - 4xy^2 - 2x^2$

$$f_x(x, y) = 4(x^3 - x - y^2)\,, \qquad f_y(x, y) = -4y(y^2 + 2x)\,,$$

zerują się w trzech punktach

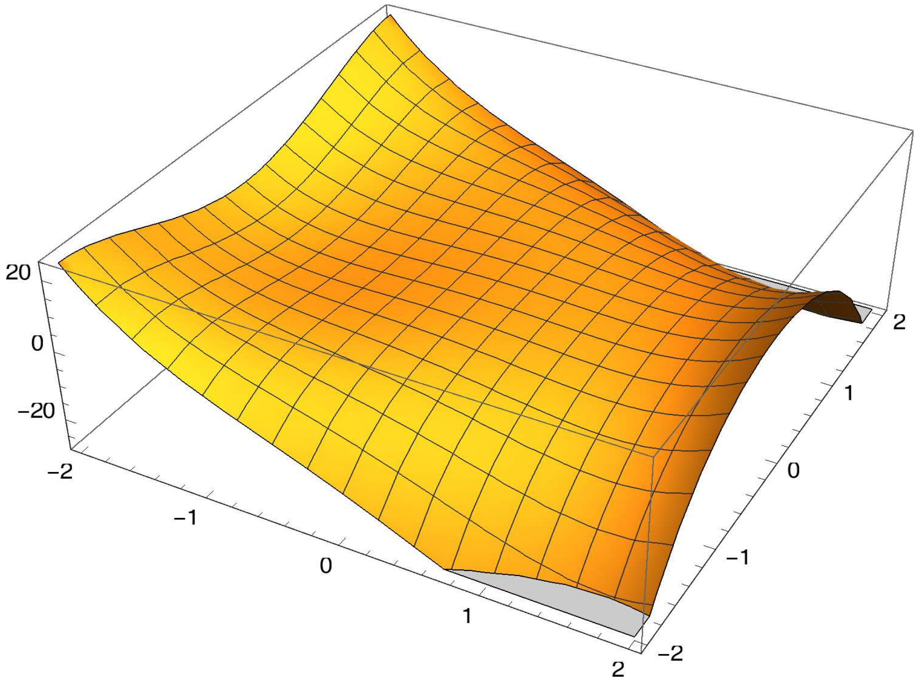
$$(0, 0)\,, \qquad (1, 0)\,, \qquad (-1, 0)\,.$$

Drugie pochodne cząstkowe

$$f_{xx}(x, y) = 12x^2 - 4\,, \qquad f_{yy}(x, y) = -12^2 - 8x\,, \qquad f_{xy} = -8y\,,$$

dają w tych trzech punktach krytycznych następujące macierze form kwadratowych

$$\begin{pmatrix} -4 & 0 \\ 0 & 0 \end{pmatrix}, \qquad \begin{pmatrix} 8 & 0 \\ 0 & -8 \end{pmatrix}, \qquad \begin{pmatrix} 8 & 0 \\ 0 & 8 \end{pmatrix}.$$



Rysunek 26: Kształt funkcji z Zadania 22g.

Druga macierz ma sygnaturę mieszaną, więc punkt $(1,0)$ jest punktem siodłowym. Trzecia macierz jest dodatio określona i wobec tego w punkcie $(-1,0)$ jest lokalne minimum. Na podstawie pierwszej macierzy, która ma sygnaturę $(-,0)$ nie można wykluczyć, że w punktcie $(0,0)$ funkcja ma lokalne maksimum (na wszystkich wektorach przemieszczeń macierz ta daje wartości ujemne lub zero, ale nigdy dodatnie). Trzeba tu pogłowkować. Np. $f(\varepsilon,0)=\varepsilon^4-2\varepsilon^2$ i przy dostatecznie małych wartościach $|\varepsilon|$, mamy $0=f(0,0)>f(\varepsilon,0)$. Ale

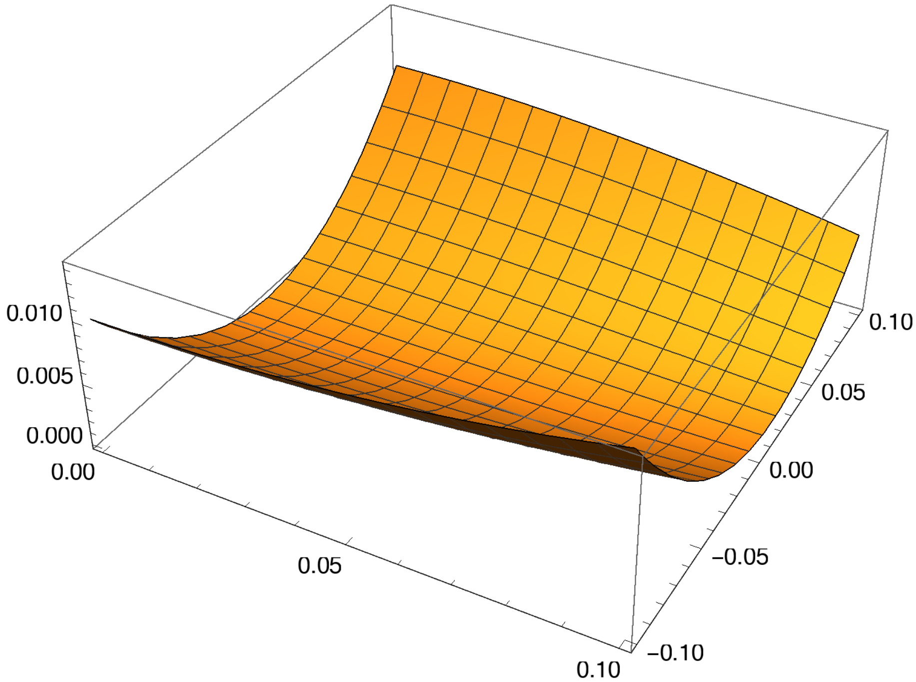
$$f(-\varepsilon^2,\varepsilon)=\varepsilon^8+\varepsilon^4>0\,.$$

Zatem w dowolnie małym otwartym otoczeniu punktu $(0,0)$ zawsze są punkty, w których wartość f jest większa i takie, w których jest mniejsza niż w punkcie $(0,0)$. Zatem w punkcie tym funkcja nie ma extremum. Nie jest to łatwo zobaczyć patrząc na kształt tej funkcji pokazany na rysunku 26!

Zadanie 23: Pierwsze pochodne $f_x(x,y)=12x^3-8xy$ i $f_y(x,y)=-4x^2+2y$ rzeczywiście znikają w punkcie $(0,0)$ - jest więc to punkt krytyczny - ale drugie pochodne $f_{xx}(x,y)=36x^2-8y$, $f_{yy}=2$ i $f_{xy}(x,y)=-8x$ dają w tym punkcie macierz

$$\begin{pmatrix}0 & 0\\ 0 & 2\end{pmatrix},$$

która nie jest dodatnio ani ujemnie określona. Nie jest jednak z tej postaci jasne, czy jest to tylko wypłaszczenie funkcji, które może zostać "zagięte" do góry przez dalsze wyrazy



Rysunek 27: Kształt funkcji z Zadania 22 w pobliżu punktu $(0,0)$. Ponieważ $f(-x,y) = f(x,y)$, pokazana jest tylko dodatnia część osi x.

rozwinięcia, czy rzeczywiście jest to punkt siodłowy. Zaraz się tym zajmiemy. Najpierw jednak sprawdźmy, że rzeczywiście w punkcie $(0,0)$ jest minimum na każdej prostej przezeń przechodzącej. Podstawmy zatem do wzoru $y = ax$, czyli zbadajmy jako funkcję x funkcję

$$g(x) = f(x,\, ax) = 3x^4 - 4ax^3 + a^2x^2\,.$$

Oczywiście $g'(0) = 12x^3 - 12ax^2 + 2a^2a = 0$ i $g'' = 36x^2 - 24ax + 2a^2 = 2a^2 > 0$. Wzdłuż prostych $y = ax$ rzeczywiście funkcja $g(x)$ ma w $x = 0$ minimum. trzeba jeszcze sprawdzić prostą, której wzór $y = ax$ nie obejmuje, tj. prostą $x = 0$. Wtedy $h(y) = f(0,y) = y^2$ i jest jasne, że ma ona minimum w zerze.

Pozostaje jeszcze wyjaśnić, co się w punkcie $(0,0)$ naprawdę dzieje. W tym celu najlepiej rozpatrzyć funkcję $\tilde{f}(z,y) = f(\sqrt{x},\, y)$, tj. podstawić $x^2 = z$. Funkcja $\tilde{f}(z,y)$ zdaje sprawę z zachowywania się $f(x,y)$ tylko w obszarze $x \geq 0$ ale to nam wystarczy. Funkcja $\tilde{f}(z,y)$ ma oczywiście punkt krytyczny w $(0,0)$ ale macierz formy kwadratowej jej drugich pochodnych w tym punkcie ma postać

$$\begin{pmatrix} 6 & -4 \\ -4 & 2 \end{pmatrix}.$$

I teraz jest jasne, że może ona na wektorach przesunięć dawać wartości zarówno dodatnie, jak i ujemne (np. gdy $(h_z, h_y) = (\varepsilon,\, 2\varepsilon)$ daje ona $-2\varepsilon^2$, a gdy $(h_z, h_y) = (\varepsilon, 0)$ daje $6\varepsilon^2$).

Oznacza to, że w dowolnie małym otwartym otoczeniu punktu $(x, y) = (0, 0)$ są punkty, w których wartość funkcji jest ujemna, np. punkty typu $(\sqrt{\varepsilon},\ 2\varepsilon)$ ale jeśli z ustalonego takiego puktu do $(0, 0)$ wytyczymy prostą, to pomiędzy nimi tuż przy $(0, 0)$ są już tylko punkty, w których wartość f jest dodatnia. Punkty, w których wartość f jest ujemna biegną bowiem do punktu $(0, 0)$ po paraboli. Konia z rzędem temu, kto to zobaczy z wykresu funkcji pokazanego na rysunku 27!

Zadanie 24: Najpierw sprawdzamy, czy funkcja ma jakieś ekstrema wewnątrz trójkąta. Pierwsze pochodne cząstkowe

$$f_x(x, y) = 2xy - 8\,, \qquad f_y(x, y) = x^2 - 4\,,$$

zerują się w punktach $(-2, -2)$ i $(2, 2)$. Pierwszy punkt leży poza trójkątem, a drugi leży dokładnie na jednym jego z boków. Zatem wewnątrz trójkąta funkcja nie ma ekstremów. Może je tylko mieć na brzegu. Na brzegu $y = 0$, funkcja $f(x, 0) = -8x$ jest malejąca i najmniejszą wartość -32 przyjmuje w $x = 4$, czyli w punkcie $(4, 0)$. Na brzegu $x = 0$, funkcja $f(0, y) = -4y$ też jest malejąca i najmniejszą wartość -16 przyjmuje w $x = 4$ czyli w punkcie $(0, 4)$. Trzeci bok, ten na którym wypadł punkt krytyczny, jest wyznaczony równaniem $y = 4 - x$. Na tym boku funkcja jest dana wzorem

$$h(x) = f(x, 4 - x) = -x^3 + 4x^2 - 4x - 16\,,$$

(oczywiście $h(0) = f(0, 4) = -16$, a $h(4) = f(4, 0) = -32$). Jej pochodna

$$h'(x) = -3x^2 + 8x - 4 = -3(x - 2)(x - \frac{2}{3})\,,$$

zeruje się, gdy $x = 2$ i tam druga pochodna, $h''(x) = -6x + 8$ jest ujemna, oraz gdy $x = 2/3$ i tam $h''(2/3) = 4$ jest dodatnia. Zatem funkcja $h(x)$ ma lokalne minimum, równe $h(2/3) \approx -17$ w $x = 2/3$ i lokalne maksimum, równe $h(2) = -16$ w punkcie $x = 2$. Zatem najmniejszą wartością funkcji $f(x, y)$ jest -32 osiąganą w $(4, 0)$ a największą, 0 osiąganą w punkcie $(0, 0)$ - tego punku wcześniej nie sprawdziliśmy, ale nie należy o nim zapominać!

Zadanie 25:
a) Sprawdzamy: $F(0, \frac{\pi}{4}) = 0$, czyli $(0, \frac{\pi}{4}) \in F^{-1}(0)$.

$$F_x(x, y) = \frac{1}{\cos^2(x + y)} - y\,, \qquad F_y(x, y) = \frac{1}{\cos^2(x + y)} - x\,,$$

i $F_y(0, \frac{\pi}{4}) = 2 \neq 0$, więc $F = 0$ definiuje w otoczeniu tego punktu taką funkcję $y = y(x)$, że $y(0) = \frac{\pi}{4}$. Poza tym, $F_x(0, \frac{\pi}{4}) = 2 - \frac{\pi}{4}$, czyli $y'(0) = \frac{\pi}{8} - 1$.

$$F_{xx}(x, y) = F_{yy}(x, y) = \frac{2\,\mathrm{tg}(x + y)}{\cos^2(x + y)}\,, \qquad F_{xy}(x, y) = \frac{2\,\mathrm{tg}(x + y)}{\cos^2(x + y)} - 1\,,$$

i $F_{xx}(0, \frac{\pi}{4}) = F_{yy}(0, \frac{\pi}{4}) = 4$, $F_{xy}(0, \frac{\pi}{4}) = 3$. Zatem

$$y''(0) = -\frac{1}{2^3}\left[4 \cdot (2)^2 - 2 \cdot 3 \cdot 2 \cdot (2 - \pi/4) + 4 \cdot (2 - \pi/4)^2\right] = -1 + \frac{\pi}{8} - \frac{\pi^2}{32}\,.$$

b) $F(\frac{\pi}{6},0) = -1+1 = 0$, czyli $(\frac{\pi}{6},0) \in F^{-1}(0)$.

$$F_x(x,y) = -3y + 2\cos x\,, \qquad F_y(x,y) = 1 - 3x + \cos y\,,$$

Ponieważ $F_y(\frac{\pi}{6},0) = 1 - \frac{\pi}{2} + 1 \neq 0$, więc $F = 0$ definiuje w otoczeniu tego punktu taką funkcję $y = y(x)$, że $y(\frac{\pi}{6}) = 0$. Poza tym, $F_x(\frac{\pi}{6},0) = \sqrt{3}$, czyli $y'(\frac{\pi}{6}) = -\sqrt{3}/(2-\frac{\pi}{2})$.

$$F_{xx}(x,y) = -2\sin x\,, \qquad F_{yy}(x,y) = -\sin y\,, \qquad F_{xy}(x,y) = -3\,.$$

Stad $F_{xx}(\frac{\pi}{6},0) = -1$, $F_{yy}(\frac{\pi}{6},0) = 0$. Zatem

$$y''(\frac{\pi}{6}) = -[-1\cdot(2-\pi/2)^2 - 2\cdot(-3)\cdot\sqrt{3}\cdot(2-\pi/2)]/(2-\pi/2)^3\,.$$

c) $F(\frac{\pi}{6},0) = 2-2-1+1+1 = 0$, czyli $(2,1) \in F^{-1}(0)$.

$$F_x(x,y) = \frac{1}{y} + \frac{4}{x^2}\,, \qquad F_y(x,y) = -\frac{x}{y^2} - 1\,.$$

$F_y(2,1) = -3 \neq 0$, więc $F = 0$ definiuje w otoczeniu tego punktu taką funkcję $y = y(x)$, że $y(2) = 1$. $F_x(2,1) = 2$, więc $y'(2) = 2/3$.

$$F_{xx}(x,y) = -\frac{8}{x^3}\,, \qquad F_{yy}(x,y) = \frac{2x}{y^3}\,, \qquad F_{xy}(x,y) = -\frac{1}{y^2}\,.$$

$F_{xx}(2,1) = -1$, $F_{yy}(2,1) = 4$, $F_{xy}(2,1) = -1$.

$$y''(2) = \frac{1}{3}\left[-1\cdot(-3)^2 - 2\cdot(-1)\cdot 2\cdot(-3) + 4\cdot 2^2\right] = -\frac{5}{3}\,.$$

d) $F(2.0) = 0$, czyli $(2,0) \in F^{-1}(0)$.

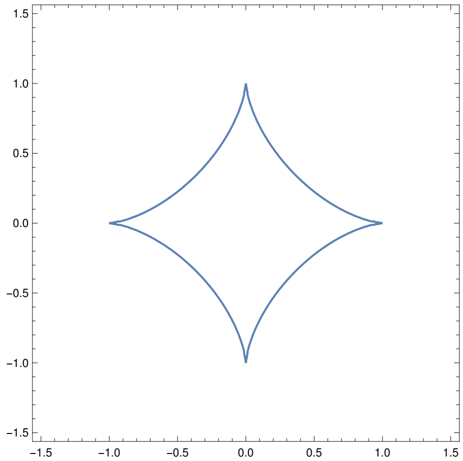
$$F_x(x,y) = -\frac{1+2y}{(x-1)^2} + 3y\,, \qquad F_y(x,y) = \frac{2}{x-1} + 3x\,.$$

$F_y(2,0) = 2+6 = 8 \neq 0$, więc $F = 0$ definiuje w otoczeniu tego punktu taką funkcję $y = y(x)$, że $y(2) = 0$. $F_x(2,0) = -1$, więc $y'(2) = 1/8$.

$$F_{xx}(x,y) = \frac{4y+2}{(x-1)^3}\,, \qquad F_{yy}(x,y) = 0 \qquad F_{xy}(x,y) = -\frac{2}{(x-1)^2} + 3\,.$$

$F_{xx}(2,0) = 1/4$, $F_{xy}(2,0) = 5/2$.

$$y''(2) = -\frac{1}{8^3}\left[\frac{1}{4}\cdot 8^2 - 2\cdot\frac{5}{2}\cdot 8\cdot(-1)\right].$$



Rysunek 28: Zbiór E z Zadania 26a.

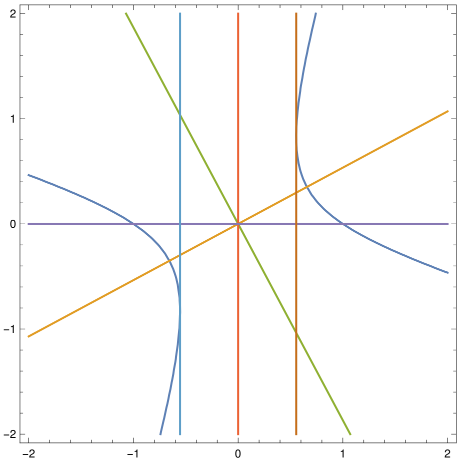
Zadanie 26:
a) Pochodne F: $F_x(x,y) = (2/3)x^{-1/3}$, $F_y(x,y) = (2/3)y^{-1/3}$. Punktami, w których $F = 0$ nie wyznacza funkcji $y = y(x)$ są rozwiązania układu

$$x^{2/3} + y^{2/3} = a^{2/3}\,, \qquad y^{-1/3} = 0\,.$$

Takich niema. Ale nie należy zapominać, że aby w ogóle w otoczeniu jakiegoś punktu można było pytać o fukcję $y = y(x)$, czy $x = x(y)$, muszą w tym punkcie istnieć i być ciągłe (w sensie wałkowanej przez nas ciągłości funkcji wielu zmiennych) obie pochodne $F_x(x,y)$ i $F_y(x,y)$. Tu w punktach $(0, \pm a)$, które należą do zbioru E, nie istnieje F_x, a w unktach $(\pm a, 0)$, które też należą do zbioru E, nie istnieje F_y. Zatem w tych punktach $F = 0$ nie wyznacza ani $y = y(x)$, ani $x = x(y)$. Jak wygląda zbiór E? Jest to jakby kwadrat z wierzchołkami położonymi symetrycznie na osiach x i y, któremu boki się wklęsły do wewnątrz. Dlaczego nie na zewnątrz? Jak sobie względem y (przyjmując, że $y > 0$ np.) rozwiążemy uogólniony warunek $|x|^p + |y|^p = 1$ (zawsze można przeskalować zmienne, żeby się tego a pozbyć), który niewątpliwie przy $p = 1$ wyznacza właśnie kwadrat: $y = (1 - x^p)^{1/p}$ i na obliczymy $y' = -x^{p-1}(1 - x^p)^{-1+1/p}$, to widzimy, że w $x = 0$ pochodna ta się zeruje, jeśli $p > 1$, czyli funkcja $y = y(x)$ dochodzi do wierzchołka płasko (kwadrat jest wypuczony na zewnątrz), jest równa -1, gdy $p = 1$ (prawdziwy kwadrat właśnie) i jest równa $-\infty$, gdy $p < 1$, co właśnie oznacza, że kwadrat się wklęsł do wewnątrz. Gdy $p = 2/3$ wygląda to jak na rysunku 28. Widać teraz, że rzeczywiście wszystkie cztery wierzchołki są punktami, w których nie może istnieć $y = y(x)$, ani $x = x(y)$, bo krzywa będąca zbiorem E ma tam dziubki. Teraz pochodne y' i y'' w punktach zbioru E, które nie są wierzchołkami

$$y' = -\left(\frac{y}{x}\right)^{1/3}, \qquad y'' = -\frac{y^3}{3}\left[(x^2y)^{-2/3} + (y^2x)^{-2/3}\right].$$

b) To jest obrócona hiperbola. $F_x = 2x + 3y$, $F_y = 3x - 2y$. Punkty, w których $F = 0$ nie



Rysunek 29: Zbiór E z Zadania 26b. Pokazane są oba układy współrzędnych: pierwotny i obrócony o kąt θ; dwie dodatkowe linie pionowe stykają się z ramionami hiperboli w punktach, w których warunek $x^2 - y^2 + 3xy = 1$ nie zadaje funkcji $y = y(x)$: w tych punktach krzywa (z punktu widzenia układu xy) "staje dęba".

wyznacza $y = y(x)$ to rozwiązania układu

$$\begin{aligned} F_y &= 3x - 2y = 0\,, \\ F &= x^2 - y^2 + 3xy - 1 = 0 \end{aligned}$$

Daje to $x = \pm 2/\sqrt{13}$, $x = y = \pm 2/\sqrt{13}$. Żeby zobaczyć dlaczego, wprowadźmy nowe osie x' i y' wzorami

$$\begin{aligned} x &= x' \cos\theta - y' \sin\theta \equiv x' c_\theta - y' s_\theta\,, \\ y &= x' \sin\theta + y' \cos\theta \equiv x' s_\theta + y' c_\theta\,, \end{aligned}$$

i podstawmy do $F(x, y) = F(x' c_\theta - y' s_\theta,\, x' s_\theta + y' c_\theta) = 0$. Da to

$$x'^2(c_\theta^2 - s_\theta^2 + 3 s_\theta c_\theta) - y'^2(c_\theta^2 - s_\theta^2 + 3 s_\theta c_\theta) + x'y'(3c_\theta^2 - s_\theta^2) - 4 s_\theta c_\theta) = 1\,.$$

Można teraz dobrać kąt θ tak, by wyzerować człon z $x'y'$:

$$\mathrm{tg}\, 2\theta = \frac{3}{2}\,,$$

Równanie $F(x, y) = 0$ przybierze wtedy postać

$$x'^2 \left(\frac{13}{3}\, s_\theta c_\theta\right) - y'^2 \left(\frac{13}{3}\, s_\theta c_\theta\right) = 1\,.$$

To już jest "szkolne" równanie hiperboli. Jak duży jest kąt θ? Można i do tego dojść (bez kalkulatora): warunek

$$\frac{2 s_\theta c_\theta}{c_\theta^2 - s_\theta^2} = \frac{3}{2}\,,$$

jest równoważny równości

$$t_\theta^2 + \frac{4}{3}\,t_\theta - 1 = 0\,,$$

której rozwiązaniem (mniejszym) jest

$$t_\theta \equiv \operatorname{tg}\theta = \frac{-2+\sqrt{13}}{3}\,.$$

Ponieważ $\sqrt{13} < 4$, $\operatorname{tg}\theta < 1$, czyli obrót jest o mniej niż $\pi/4$. I rzeczywiście, rysunek 29 krzywej $x^2 - y^2 + 3xy = 1$ pokazuje, że jest to"szkolna" hiperbola w ździebko pochylonym układzie.

Zadanie 27:
a) $F_x(x,y) = 2x - 2$, $F_y(x,y) = 2y - 2$. $F_x = 0$ w punkcie $x = 1$, któremu odpowiada y będący rozwiązaniem warunku $F(1,y) = 0$, czyli $-2y + y^2 = 0$. Rozwiązaniami warunków $F_x = 0$ i $F = 0$ są więc dwa punkty $(1,0)$ i $(1,2)$. W obu tych punktach pochodna F_y nie jest równa zeru, więc w ich otoczeniach warunek $F = 0$ definiuje - teraz to jest oczywiste, bo x jest ten sam - dwie różne funkcje: $y = y_1(x)$ w otoczeniu punktu $(1,0)$ i $y = y_2(x)$ w otoczeniu punktu $(1,2)$. Oczywiście $y_1(1) = 0$, a $y_2(1) = 2$. Dla obu tych funkcji $x = 1$ jest punktem krytycznym, bo $y_1'(1) = 0$ i $y_2'(1) = 0$. Ponieważ $F_{xx}(x,y) = 2$, a $F_y(x,y) = 2y - 2$, drugie pochodne tych funkcji są w $x = 1$ równe

$$y_1''(1) = -\frac{F_{xx}(1,0)}{F_y(1,0)} = 1\,, \qquad y_2''(1) = -\frac{F_{xx}(1,2)}{F_y(1,2)} = -1\,.$$

Funkcja $y = y_1(x)$ ma więc w $x = 1$ (lokalne) minimum, a funkcja $y = y_1(x)$ ma w tym punkcie (lokalne) maksimum.

Warunek $F(x,y) = y^2 - 2y + x^2 - 2x + 1$ można w tym przypadku zresztą jawnie rozwiązać (bo to zwykłe równanie kwadraticznoje), dostając

$$y = 1 \mp \sqrt{1 - (x^2 - 2x + 1)}\,.$$

Widać, że $-$ daje funkcję $y = y_1(x)$, bo dla $x = 1$ $y = 1 - 1 = 0$, a $+$ daje funkcję $y = y_2(x)$. "Szkolne" pochodne tych funkcji są równe

$$y_1'(x) = -\frac{2-2x}{2\sqrt{1-(x^2-2x+1)}}\,, \qquad y_2'(x) = \frac{2-2x}{2\sqrt{1-(x^2-2x+1)}}\,.$$

Obie zerują się w $x = 1$ i, ponieważ różnią się tylko znakiem, gdy pochodna y_1' zmienia się z ujemnej na dodatnią, przy przejściu x-a (z lewa na prawo) przez punkt $x = 1$ (co zapewnia, że w $x = 1$ jest lokalne minimum funkcji y_1), to pochodna y_1' zmienia się z dodatniej na ujemną.

b) $F_x(x,y) = 2y + 2x$, $F_y(x,y) = 3y^2 + 2x$. Warunek $F_x = 0$ jest spełniony, gdy $y = -x$, a warunek $F(x,-x) = 0$ ma jako rozwiązania $x = 0$ oraz $x = -1$, czyli punktami podejrzanymi są punkty $(0,0)$ i $(-1,1)$ zbioru $F^{-1}(0) \subset \mathbb{R}^2$. Ale w punkcie $(0,0)$ zeruje się pochodna F_y, więc w tym punkcie (i jego otoczeniu) warunek $F(x,y)=0$ nie wyznacza funkcji $y=y(x)$. Zostaje zatem do zbadania tylko punkt $(-1,1)$, w którym $F_y = 1$. $F_{xx}=2$ i $y''(-1) = -F_{xx}(-1,1)/F_y(-1,1) = -2$. Funkcja $y=y(x)$, taka że $y(-1)=1$, ma więc w punkcie -1 maksimum lokalne.

Zadanie 28: Po pierwsze sprawdzamy, czy pytanie ma sens, tj. czy podany punkt $(1,3,2)$ należy do poziomicy zerowej odwzorowania $F(x,y,z) = z^3 - xyz - 2$. Należy, bo $F(1,3,2)=0$. Pochodne cząstkowe F

$$F_x(x,y,z) = -yz\,, \quad F_y(x,y,z) = -xz\,, \quad F_z(x,y,z) = 3z^2 - xy\,,$$

mają w punkcie $(1,3,2)$ wartości: $F_x=-6$, $F_y=-2$, $F_z=9$. Ponieważ F_z w tym puncie nie znika, a wszystkie te pochodne są, jako funkcje na $\mathbb{R}^3$ ciągłe w otoczeniu (dowolnym zresztą) tego punktu, więc warunek $F(x,y,z)=0$ wyznacza w otoczeniu punktu $(1,3,2)$ funkcję $z=z(x,y)$. Jej pierwsze pochodne cząstkowe są tam równe

$$z_x(1,3)|_{z=2} = -\frac{F_x(1,3,2)}{F_z(1,3,2)} = \frac{2}{3}\,, \qquad z_y(1,3)|_{z=2} = -\frac{F_y(1,3,2)}{F_z(1,3,2)} = \frac{2}{9}\,.$$

(Piszemy $z_x(1,3)|_{z=2}$, a nie $z_x(1,3)$ po prostu, bo może się zdarzyć, że jest jakieś inne jeszcze rozwiązanie warunku $F(1,3,z)=0$ z $z=z_0\neq 2$, i wtedy by były dwie różne funkcje $z=z_1(x,y)$ i $z_2(x,y)$ zdefiniowane w otoczeniu tego samego punktu $(1,3)\in\mathbb{R}^2$). Drugie pochodne cząstkowe odwzorowania F są w punkcie $(1,3,2)$ równe

$$\begin{array}{lll} F_{xx}(1,3,2)=0\,, & F_{yy}(1,3,2)=0\,, & F_{zz}(1,3,2)=12\,,\\ F_{xy}(1,3,2)=-2\,, & F_{xz}(1,3,2)=-3\,, & F_{yz}(1,3,2)=-1\,. \end{array}$$

Aby obliczyć drugie pochodne cząstkowe funkcji $z=z(x,y)$ w punkcie $(1,3,2)$ wstawiamy te liczby do wyprowadzonych wzorów

$$\begin{aligned} z_{xx} &= -\frac{F_{xx}\,F_z^2 - 2F_{xz}\,F_x\,F_z + F_{zz}\,F_x^2}{F_z^3}\,,\\ z_{yy} &= -\frac{F_{yy}\,F_z^2 - 2F_{yz}\,F_y\,F_z + F_{zz}\,F_y^2}{F_z^3}\,,\\ z_{xy} &= -\frac{F_{xy}\,F_z^2 - F_{xz}\,F_y\,F_z - F_{yz}\,F_x\,F_z + F_{zz}\,F_x\,F_y}{F_z^3}\,. \end{aligned}$$

i dostajemy $z_{xx}(1,3)|_{z=2} = -4/27$, $z_{yy}(1,3)|_{z=2} = -4/243$, $z_{xy}(1,3)|_{z=2} = 42/243$. (Może się nie pomyliłem).

Zadanie 29: Brzmi to zawile (bo zadania w zbiorku ś.p. G. Cieciury są zawsze w taki skomplikowany sposób formułowane - ja i tak je sprowadzam do ludzkiego...), ale jest banalnie proste. Jeśli funkcja $\phi(t)$ jest różniczkowalna w okolicy $t=0$, to jest tam przyzwoita, nie robi siupów, ma ciągłą pochodną $\phi'(t)$, i jedyny problem, że nie wiemy, jak wygląda. Ale to pestka bo poza tym wszystko normalnie działa. Obliczamy pochodne cząstkowe odwzorowania F korzystając ze znanych już reguł

$$
\begin{aligned}
F_x(x,y,z) &= e^z\,\phi'\big(x\,e^z - y\,e^{-z}\big)\,,\\
F_y(x,y,z) &= -\,e^z\,\phi'\big(x\,e^z - y\,e^{-z}\big)\,,\\
F_z(x,y,z) &= \big(x\,e^z + y\,e^{-z}\big)\,\phi'\big(x\,e^z - y\,e^{-z}\big) - 1\,.
\end{aligned}
$$

Ponieważ $\phi(t)$ i $\phi'(t)$ jako się rzekło są przyzwoite, więc $\phi'(0)$ istnieje i ma skończoną wartość. Zatem w punkcie $(0,0,z_0)$, gdzie z_0 jest takie, że $F(0,0,z_0)=0$, pochodna F_z jest po prostu równa -1 (bo czynnik przed ϕ' jest równy zeru), a nie zero i twierdzenie o funkcji uwikłanej gwarantuje, że w otoczeniu $(0,0)\in\mathbb{R}^2$ istnieje funkcja $z=z(x,y)$ taka, że $z(0,0)=z_0$ i mająca w tym otoczeniu pochodne

$$
z_x(x,y) = -\left.\frac{F_x(x,y,z)}{F_z(x,y,z)}\right|_{z=z(x,y)}\,, \qquad z_y(x,y) = -\left.\frac{F_(x,y,z)}{F_z(x,y,z)}\right|_{z=z(x,y)}\,.
$$

Równanie różniczkowe ma więc postać

$$
\frac{\partial z}{\partial x} + g(z)\,\frac{\partial z}{\partial y} = -\frac{1}{F_z}\left\{e^z\,\phi'\big(x\,e^z - y\,e^{-z}\big) - g(z)\,e^{-z}\,\phi'\big(x\,e^z - y\,e^{-z}\big)\right\} = 0\,.
$$

Jest jasne, że funkcja $z(x,y)$ je spełnia, jeśli $g(z)=e^{2z}$.

Zadanie 30:
a) Obie pochodne cząstkowe odwzorowania F

$$
F_x(x,y,z) = -21(x^2-1)z + 4(2x+y)\,, \qquad F_y(x,y,z) = 2(2x+y)\,,
$$

zerują się jednocześnie, gdy $y=-2x$ i $x=\pm1$. Wartość z wyznacza warunek

$$
F(\pm1,\mp2,z) = 6z^3 \pm 14z - 20 = 0\,.
$$

Gdy $x=1$

$$
6z^3 + 14z - 20 = (z-1)(6z^2+6z+20) = 0\,,
$$

a gdy $x=-2$

$$
6z^3 - 14z - 20 = (z-2)(6z^2+12z+10) = 0\,.
$$

W obu przypadkach Δ trójmianu w drugim nawiasie jest ujemna. Zatem punktami krytycznymi są

$$
(1,-2,1)\,, \qquad (-1,2,2)\,.
$$

Tzn, $(1,-2)$ i $(-2,2)$ są punktami krytycznymi funkcji $z=z(x,y)$ (jednej funkcji, a może dwóch różnych? tego się właśnie przy funkcjach zadanych w sposób uwikłany nie wie...), a w tych punktach $z(1,-2) = 1$ i $z(-1,2) = 2$. Reszta to rutyna. Pochodna $F_z(x,y,z) = 18z^2 - 7(x^3 - 3x)$ jest w tych punktach równa odpowiednio 32 i 58,

$$F_{xx}(x,y,z) = 8 - 42xz\,, \qquad F_{yy}(x,y,z) = 2\,, \qquad F_{xy}(x,y,z) = 4\,,$$

i stąd macierze form kwadratowych drugich pochodnych w tych punktach krytycznych mają postać

$$Q_{(1,-2,1)} = -\frac{1}{32}\begin{pmatrix} -34 & 4 \\ 4 & 2 \end{pmatrix}, \qquad Q_{(-1,2,2)} = -\frac{1}{58}\begin{pmatrix} 92 & 4 \\ 4 & 2 \end{pmatrix}.$$

Pierwsza forma jest nijaka i w punkcie $(1,-2)$ funkcja $z = z(x,y)$ ma punkt siodłowy; druga forma jest ujemnie określona i punkcie $(-1,2)$ funkcja $z = z(x,y)$ (ta sama, albo inna) ma lokalne maksimum.

b) Obie pochodne cząstkowe odwzorowania F

$$F_x(x,y,z) = 14z\,\frac{1-x^2}{(1+x^2)^2} + 4(2x-y)\,, \qquad F_y(x,y,z) = -2(2x-y)\,,$$

zerują się jednocześnie, gdy $y = 2x$ i $x = \pm 1$ lub $z = 0$. Jednak $z = 0$ jest sprzeczne z $F = 0$ (bo $F = 9$, gdy $z = 0$ i $y = 2x$), więc są tylko dwie możliwości. Gdy $(x,y) = (1,2)$, to

$$F(1,2,z) = z^3 + 8z + 9 = (z+1)(z^2 - z + 9) = 0\,,$$

co daje $z = -1$ (bo Δ trójmianu jest ujemna); gdy $(x,y) = (-1,-2)$, wówczas

$$F(1,2,z) = z^3 - 6z - 9 = (z+3)(z^2 - 3z + 3) = 0\,,$$

i znów jedynym pierwiastkiem jest $z = -3$. Zatem punktami krytycznymi są

$$(1,2,-1)\,, \qquad (-1,-2,-3)\,,$$

tzn., $(1,2)$ i $(-1,-2)$, a w tych punktach $z(1,2) = -1$ i $z(-1,-2) = -3$. Pochodna $F_z(x,y,z) = 3z^2 + 1 + 14z/(1+x^2)$ jest w tych punktach równa odpowiednio 11 i 21.

$$F_{xx}(x,y,z) = 2zx\,\frac{x^2-3}{(1+x^2)^3} + 8\,, \qquad F_{yy}(x,y,z) = 2\,, \qquad F_{xy}(x,y,z) = -4\,,$$

i stąd macierze form kwadratowych drugich pochodnych w tych punktach krytycznych mają postać

$$Q_{(1,2,-1)} = -\frac{1}{11}\begin{pmatrix} 15 & -4 \\ -4 & 2 \end{pmatrix}, \qquad Q_{(-1,-2,-3)} = -\frac{1}{21}\begin{pmatrix} -13 & -4 \\ -4 & 2 \end{pmatrix}.$$

Pierwsza forma jest ujemnie określona i w punkcie $(1,2)$ funkcja $z = z(x,y)$ ma lokalne maksimum; druga forma jest nijaka (mieszana sygnatura) i punkcie $(-1,-2)$ funkcja $z = z(x,y)$ (ta sama, albo inna) ma punkt siodłowy.

c) Obie pochodne cząstkowe odwzorowania F

$$F_x(x,y,z) = 7z\sin(x+y) + 20\,\frac{1-x^2}{(1+x^2)^2}\,, \qquad F_y(x,y,z) = 7z\sin(x+y)\,,$$

zerują się jednocześnie, gdy $x+y = k\pi$ ($k \in \mathbb{Z}$) i $x = \pm 1$. Są więc takie możliwości: gdy $(x,y) = (1, -1+2k\pi)$, to

$$F(1,-1+2k\pi,z) = 3z^3 - 7z + 10 = (z+2)(3z^2-6z+5) = 0\,,$$

co daje $z=-2$ (bo Δ trójmianu jest ujemna); gdy $(x,y) = (1,-1+\pi+2k\pi)$, to

$$F(1,-1+\pi+2k\pi,z) = 3z^3 + 7z + 10 = (z+1)(3z^2-3z+10) = 0\,,$$

i $z=-1$ (znów $\Delta < 0$). Z kolei, gdy $(x,y) = (-1, 1+2k\pi)$, wówczas

$$F(-1,1+2k\pi,z) = 3z^3 - 7z - 10 = (z-2)(3z^2+6z+5) = 0\,,$$

i jedynym pierwiastkiem jest $z=2$ (znów $\Delta < 0$) i wreszcie, gdy $(x,y) = (-1, 1+\pi+2k\pi)$, to

$$F(-1,1+\pi+2k\pi,z) = 3z^3 + 7z - 10 = (z-1)(3z^2+3z+10) = 0\,,$$

i $z=1$ (bo $\Delta < 0$). Istnieją więc cztery serie punktów krytycznych

$$(1,-1+2k\pi,-2)\,, \quad (1,-1+\pi+2k\pi,-1)\,, \quad (-1,1+2k\pi,2)\,, \quad (-1,1+\pi+2k\pi,1)\,.$$

ale wszystkie punkty każdej z serii mają identyczny charakter (periodyczność kosinusa!) więc wystarczy położyć $k=0$. Pochodna $F_z(x,y,z) = 9z^2 - 7\cos(x+y)$ jest w tych czterech seriach punktów równa odpowiednio 29, 16, 29 i 16.

$$\begin{aligned} F_{xx}(x,y,z) &= 7z\cos(x+y) + 40x\,\frac{x^2-3}{(1+x^2)^3}\,, \\ F_{yy}(x,y,z) &= F_{xy}(x,y,z) = 7z\cos(x+y)\,, \end{aligned}$$

i stąd macierze form kwadratowych drugich pochodnych w tych punktach krytycznych mają postać

$$Q_{(1,-1,-2)} = -\frac{1}{29}\begin{pmatrix} -24 & -14 \\ -14 & -14 \end{pmatrix}, \qquad Q_{(1,-1+\pi,-1)} = -\frac{1}{16}\begin{pmatrix} -3 & 7 \\ 7 & 7 \end{pmatrix},$$
$$Q_{(-1,1,2)} = -\frac{1}{29}\begin{pmatrix} 24 & 14 \\ 14 & 14 \end{pmatrix}, \qquad Q_{(-1,1+\pi,1)} = -\frac{1}{16}\begin{pmatrix} 3 & -7 \\ -7 & -7 \end{pmatrix}.$$

Pierwsza forma jest dodatnio określona i w punkcie $(1,-1)$ funkcja $z = z(x,y)$ ma lokalne minimum; druga forma jest nijaka (mieszana sygnatura) i punkcie $(1,-1+\pi)$ funkcja $z = z(x,y)$ (ta sama, albo inna) ma punkt siodłowy. Trzecia macierz jest ujemnie określona i tam (jakaś) funkcja ma maksimum, a czwarta jest nieokreślona i punkt $(-1, 1+\pi)$ jest punktem siodłowym (jakiejś) funkcji $z = z(x, y)$.

Zadanie 31: Najpierw sprawdzamy, czy $F(1,-1,1,1) = \mathbf{0} \in \mathbb{R}^2$, czyli, czy $F^1(1,-1,1,1) = 0$ i $F^2(1,-1,1,1) = 0$. Zgadza się. Jeśli w otoczeniu tego punktu F zadaje funkcje $x = x(y,t)$ i $z = z(y,t)$, to

$$F^1(x(y,t),\, y,\, z(y,t),\, t) \equiv 0\,, \qquad F^2(x(y,t),\, y,\, z(y,t),\, t) \equiv 0\,,$$

Symbol $\equiv$ oznacza, że są wyrażenia po lewej są tożsamościowo zerami ze względu na zmienne y i t. Aby znaleźć pochodne x_y i z_y różniczkujemy te tożamości po y:

$$\begin{aligned}
\frac{\partial}{\partial y} F^1(x(y,t),\, y,\, z(y,t),\, t) &= 2y - 2zx_y - 2xz_y = 0\,,\\
\frac{\partial}{\partial y} F^2(x(y,t),\, y,\, z(y,t),\, t) &= 3x^2x_y + 3y^2 - 3z^2z_y = 0\,.
\end{aligned}$$

Daje to liniowe równania na x_y i x_t:

$$\begin{pmatrix} z & x \\ 3x^2 & -3z^2 \end{pmatrix} \begin{pmatrix} x_y \\ z_y \end{pmatrix} = \begin{pmatrix} y \\ -3y^2 \end{pmatrix},$$

Stąd (macierze 2×2 odwracamy już w pamięci!)

$$\begin{pmatrix} x_y \\ z_y \end{pmatrix} = \frac{1}{-3z^3 - 3x^3} \begin{pmatrix} -3z^2 & -x \\ -3x^2 & z \end{pmatrix} \begin{pmatrix} y \\ -3y^2 \end{pmatrix},$$

skąd

$$x_y = \frac{z^2y - y^2x}{x^3 + z^3}\,, \qquad z_y = \frac{x^2y + y^2z}{x^3 + z^3}\,.$$

Równania wyznaczające x_t i z_t otrzymane z

$$\begin{aligned}
\frac{\partial}{\partial t} F^1(x(y,t),\, y,\, z(y,t),\, t) &= 2t - 2zx_t - 2xz_t = 0\,,\\
\frac{\partial}{\partial t} F^2(x(y,t),\, y,\, z(y,t),\, t) &= 3x^2x_t + 3t^2 - 3z^2z_t = 0\,.
\end{aligned}$$

mają taką samą strukturę (z powodu symetrii $F^1(x,y,z,t)$ i $F^2(x,y,z,t)$ względem zamiany $y \leftrightarrow t$), więc

$$x_t = \frac{z^2t - t^2x}{x^3 + z^3}\,, \qquad z_t = \frac{x^2t + t^2z}{x^3 + z^3}\,.$$

W punkcie $(1,-1,1,1)$ daje to następującą macierz pochodnych cząstkowych (która to macierz jest pochodną funkcji $\mathbb{R}^2 \longrightarrow \mathbb{R}^2$):

$$\begin{pmatrix} x_y & x_t \\ z_y & z_t \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Zadanie 32:

a) Tworzymy funkcję $\tilde{F}(x,y)=x^2+y^2-\lambda(x^3+y^3-16)$ i przyrównujemy do zera jej obie pochodne cząstkowe wraz z warunkiem:

$$\begin{aligned}\tilde{F}_x(x,y) &= 2x-3\lambda x^2=0\,,\\ \tilde{F}_y(x,y) &= 2y-3\lambda y^2=0\,,\\ G(x,y) &= x^3+y^3-16=0\,.\end{aligned}$$

Z pierwszego $x=0$ lub $\lambda x=2/3$, i podobnie z drugiego $y=0$ lub $\lambda y=2/3$. Pukt $(0,0)$ odpada, bo $G(0,0)\neq 0$, więc zostają trzy możliwości. Ogólnie, macierz formy Q drugich pochodnych ma postać

$$Q=\begin{pmatrix}2-6\lambda x & 0\\ 0 & 2-6\lambda y\end{pmatrix},$$

a pochodna warunku G to $G'=(3x^2,3y^2)$.

Gdy $x=0$ i $\lambda y=2/3$, warunek $G(0,y)=0$ daje $y=2^{4/3}$. Macierz Q drugich pochodnych ma na diagonali 2 i -2, czyli ma sygnaturę $(+,-)$. Jednak warunek $G'_{\mathbf{x}_*}\cdot\mathbf{h}=0$, z $G'=(0,3{\cdot}2^{8/3})$ w badanym punkcie mówi, że istotne są tylko wartości formy Q drugich pochodnych na wektorach postaci $(h,0)$, a na takich Q przyjmuje tylko wartości dodatnie. Zatem w punkcie $(0,\,2^{4/3})$ funkcja F ma na E minimum. Symetria $x\leftrightarrow y$ mówi, że tak samo jest w punkcie $(2^{4/3},\,0)$.

Gdy $\lambda x=\lambda y=2/3$, czyli $x=y$, warunek $G(x,x)=0$ daje $x=y=2$, zatem $\lambda=1/3$ i macierz formy Q ma na diagonali -2 i -2. Jest więc ona w tym punkcie ujemnie określona i nie trzeba dodatkowo badać jej na wektorach stycznych. W punkcie $(2,2)$ funkcja F ma na E maksimum lokalne.

b) Tworzymy funkcję $\tilde{F}(x,y)=2x^2y^2-\lambda(x^4+y^4-1)$ i przyrównujemy do zera jej obie pochodne cząstkowe wraz z warunkiem:

$$\begin{aligned}\tilde{F}_x(x,y) &= 4xy^2-4\lambda x^3=0\,,\\ \tilde{F}_y(x,y) &= 4x^2y-4\lambda y^3=0\,,\\ G(x,y) &= x^4+y^4-1=0\,.\end{aligned}$$

Z pierwszego $x=0$ lub $\lambda x^2=y^2$ i podobnie z drugiego $y=0$ lub $\lambda y^2=x^2$. Pukt $(0,0)$ odpada, bo $G(0,0)\neq 0$. Jeśli $x\neq 0$ i $y\neq 0$, to wtedy $\lambda=x^2/y^2=y^2/x^2$, czyli $x^4=y^4$ i są cztery możliwości $(2^{-1/4},\,2^{-1/4})$, $(2^{-1/4},\,-2^{-1/4})$, $(-2^{-1/4},\,2^{-1/4})$, $(-2^{-1/4},\,-2^{-1/4})$. Wszystkim tym punktom odpowiada $\lambda=1$. Rozwiązaniami są też punkty $(0,\pm1)$ i $(\pm1,0)$, którym odpowiada $\lambda=0$. Razem jest więc 8 punktów krytycznych.

Ogólnie, macierz formy Q drugich pochodnych ma postać

$$Q=\begin{pmatrix}4y^2-12\lambda x^2 & 8xy\\ 8xy & 4x^2-12\lambda y^2\end{pmatrix},$$

a pochodna warunku G to $G'=(4x^3,4y^3)$. Gdy $x=0$, $y=\pm1$ i $\lambda=0$ macierz ta ma postać

$$Q=\begin{pmatrix}4 & 0\\ 0 & 0\end{pmatrix},$$

ale warunek $G' \cdot \mathbf{h} = 0$ ogranicza wektory, na których należy badać jej określoność do postaci $(h, 0)$, a na takich wektorach ma ona zawsze wartość dodatnią, czyli w punktach takich funkcja F ma na E minima. Tak samo jest w punktach $(\pm 1, 0)$.

Z kolei w czterech punktach, którym odpowiada $\lambda = 1$, macierz drugich pochodnych przyjmuje postać

$$Q = 4\sqrt{2}\begin{pmatrix} -1 & \pm 1 \\ \pm 1 & -1 \end{pmatrix},$$

przy czym wyrazy pozadiagonalne są dodatnie, gdy $y = x = \pm 2^{-1/4}$ i ujemne, gdy $y = -x = \pm 2^{-1/4}$. Jednak warunek $G' \cdot \mathbf{h} = 0$ ogranicza wektory, na których należy badać określoność macierzy drugich pochodnych do postaci $(h, -h)$, gdy $y = x$ i do postaci (h, h), gdy $y = -x$, a na takich wektorach macierz Q jest we wszystkich tych punktach ujemnie określona:

$$(h, \mp h)\begin{pmatrix} -1 & \pm 1 \\ \pm 1 & -1 \end{pmatrix}\begin{pmatrix} h \\ \mp h \end{pmatrix} = -4h^2.$$

Zatem we wszystkich tych czterech punktach funkcja F ma na E lokalne maksimum.

Zadanie 33: Dowolny punkt o współrzędnych (x, y) jest od punktu $P = (3, 12)$ oddalony (najlepiej operować kwadratami odległości - unikniemy pierwiastków) o

$$D \equiv d^2 = (x - 3)^2 + (y - 12)^2 .$$

Jeśli punkt (x, y) ma leżeć na paraboli, to szukamy ekstremum $D(x, y)$ (jeśli D ma ekstremum, to $d > 0$ też) przy ubocznym warunku $G(x, y) = y^2 - 6x$. Dalej wszystko idzie już regulaminowo: tworzymy funkcję

$$\tilde{D}(x, y) = (x - 3)^2 + (y - 12)^2 - \lambda\left(y^2 - 6x\right),$$

i rozwiązujemy układ równań

$$\begin{aligned} &\tilde{D}_x(x, y) = 2(x - 3) + 6\lambda = 0 , \\ &\tilde{D}_y(x, y) = 2(y - 12) - 2\lambda y = 0 , \\ &y^2 - 6x = 0 , \end{aligned}$$

albo $x = 3(1 - \lambda)$, $y = 12/(1 - \lambda)$ i $y^2 = 6x$ czyli

$$\frac{144}{(1 - \lambda)^2} = 18(1 - \lambda) .$$

Zatem $(1 - \lambda)^3 = 8$, a stąd $\lambda = -1$, $x = 6$, $y = 6$. Drugie pochodne funkcji $\tilde{D}(x, y)$: $D_{xx}(x, y) = 2$, $D_{yy}(x, y) = 2(1-\lambda)$, $D_{xy}(x, y) = 0$ dają w znalezionym punkcie krytycznym dodatnio określoną formę kwadratową drugich pochodnych. Zatem punkt $P' = (6, 6)$ jest położony najbliżej rzeczonej paraboli. Można też zauważyć, że wektor biegnący od P

do P' o składowych $[6-3,\ 6-12] = [3,\ -6]$ jest prostopadły do wektora stycznego w punkcie P' do paraboli. To akurat można ustalić "po szkolnemu": parabola o równaniu $y = \sqrt{6x}$ ma jako funkcja pochodną $\sqrt{3/2x}$ i w punkcie $x = 6$ tangens nachylenia równy $1/2$ Zatem wektor styczny do paraboli w punkcie P' ma x=ową składową dwa razy większą niż y-kową, np. wektor $[2, 1]$ jest stczny w P' do paraboli i jest (znów "po szkolnemu") prostopadły do wektora $P'P$. Po cichu wprowadziliśmy tu kanoniczny iloczyn skalarny w $\mathbb{R}^2$.

Problem ten można oczywiście rozwiązać inaczej po prostu wstawiając do funkcji $D(x, y)$ funkcję $x = x(y) = \frac{1}{6}\,y^2$ wywikłaną z równania paraboli i szukając zwykłego ekstremum funkcji $D(y)$:

$$\frac{d}{dy}\,D(y) = 2\left(\frac{1}{6}\,y^2 - 3\right)\frac{1}{3}\,y + 2(y-12) = 0\,,$$

co prowadzi do rówanania $y^3 - 216 = 0$, czyli $y = 6$. Poza tym

$$\frac{d^2}{dy^2}\,D(y) = \frac{1}{3}\,y^2\,,$$

więc jest to minimum.

Zadanie 34:
Jeśli krawędzie akwarium mają długości x, y i z, to jego objętość jest równa xyz. Za to pole powierzchni zużytego szkła jest dane funkcją

$$S(x, y, z) = xy + 2xz + 2yz\,.$$

Dlaczego nie $2xy + 2xz + 2yz$? - żeby dać kotu szansę... Problem sprowadza się zatem do zminimalizowania funkcji $S(x, y, z)$ przy ubocznym warunku $V(x, y, z) = xyz - 32$. Tworzymy więc funkcję

$$\tilde{S}(x, y, z) = xy + 2xz + 2yz - \lambda(xyz - 32)\,,$$

i rozwązujemy rówanania

$$\begin{aligned}
\tilde{S}_x(x, y, z) &= \ \ y + 2z - \lambda yz = 0\,,\\
\tilde{S}_y(x, y, z) &= \ \ x + 2z - \lambda xz = 0\,,\\
\tilde{S}_x(x, y, z) &= 2x + 2y - \lambda xy = 0\,,\\
V(x, y, z) &= xyz - 32 = 0\,.
\end{aligned}$$

Odejmując np. od drugiego pierwsze dostajemy $(x - y)[1 - \lambda z] = 0$. Zatem abo $\lambda z = 1$, albo $x = y$ (symetria problemu względem zamiany $x \leftrightarrow y$ podpowiada, że to to właśnie będzie właściwym rozwiązaniem). Jeśli $\lambda z = 1$, to dwa pierwsze równania zgodnie dadzą

$$\begin{aligned}
\tilde{S}_x(x, y, z) &= \ \ y + 2z - y = 2z = 0\,,\\
\tilde{S}_y(x, y, z) &= \ \ x + 2z - x = 2z = 0\,,
\end{aligned}$$

czyli $z = 0$. To zaś oznacza, że $\lambda = \infty$ i nie da się spełnić $\tilde{S}_z(x, y, z) = 0$. Zatem zostaje tylko rozwiązanie z $x = y$. Równania $\tilde{S}_x(x, y, z) = 0$ i $\tilde{S}_y(x, y, z) = 0$ stają się tym samym i układ równań redukuje się do

$$\begin{aligned}
&x + 2z - \lambda xz = 0\,, \\
&x\,(4 - \lambda x) = 0\,, \\
&x^2 z - 32 = 0\,.
\end{aligned}$$

Ponieważ $x = 0$ jest sprzeczne z trzecim, drugie daje $\lambda x = 4$, a to wstawione do pierwszego daje $x = 2z$. To zaś użyte w ostatnim mówi, że $4z^3 = 32$, czyli $z = 2$ i stąd już $x = y = 4$ i $\lambda = 1$. Drugie pochodne

$$\begin{aligned}
&\tilde{S}_{xx}(x, y, z) = 0\,, &&\tilde{S}_{yy}(x, y, z) = 0\,, &&\tilde{S}_{zz}(x, y, z) = 0\,, \\
&\tilde{S}_{xy}(x, y, z) = 1 - \lambda z\,, &&\tilde{S}_{xz}(x, y, z) = 2 - \lambda y\,, &&\tilde{S}_{yz}(x, y, z) = 2 - \lambda x\,,
\end{aligned}$$

dają w znalezionym punkcie krytycznym nastepująca macierz Q formy kwadratowej

$$Q = \begin{pmatrix} 0 & -1 & -2 \\ -1 & 0 & -2 \\ -2 & -2 & 0 \end{pmatrix}.$$

W ogólności nie jest to macierz określona, ale trzeba ją badać na wektorach (h_x, h_y, h_z) stycznych do powierzchni więzów, tj. spełniających warunek

$$(yz,\, xz,\, xy)|_{(4,4,2)} \begin{pmatrix} h_x \\ h_y \\ h_z \end{pmatrix} = (8,\, 8,\, 16) \begin{pmatrix} h_x \\ h_y \\ h_z \end{pmatrix} = 0\,,$$

czyli takich, że np. $h_z = -\frac{1}{2}(h_x + h_y)$. Podstawiając ten związek do

$$Q(\mathbf{h}) = -2h_x h_y - 4h_x h_z - 4h_y h_z\,,$$

dostajemy

$$-2h_x h_y + 2(h_x + h_y)(h_x + h_y) = 2h_x^2 + 2h_y^2 + 2h_x h_y = (h_x, h_y) \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} h_x \\ h_y \end{pmatrix}.$$

Forma zredukowana do wektorów stycznych do powierzchni wiezów jest wiec dodatnio określona i w punkcie krytycznym $(4, 4, 2)$ funkcja $S(x, y, z) = xy + 2xz + 2yz$ ma na zbiorze $E = \{(x, y, z) \in \mathbb{R}^3 : xyz = 32\}$ minimum.

Zadanie to można rozwiązać także zwyczajnie, bo z warunku ubocznego można bez trudu wywikłać dowolną ze zmiennych. Oczywiście najlepiej wywikłać zmienną z, by utrzymać w jawnej postaci wspomnianą już symetrię problemu względem zamiany $x \leftrightarrow y$. Zatem $z = 32/xy$ i pole powierzchni, które chcemy zmininalizować wyraża się wzorem

$$s(x, y) = xy + \frac{64}{x} + \frac{64}{y}\,.$$

Przyrównane do zera pierwsze pochodne funkcji $s(x, y)$

$$s_x(x, y) = y - \frac{64}{x^2},$$
$$s_y(x, y) = x - \frac{64}{y^2},$$

dają równania $x^2y = 64$ i $xy^2 = 64$, a ponieważ funkcja $s(x, y)$ jest określona na $\mathbb{R}^2$ w wyłączeniem osi x i osi y (gdy $x = 0$ lub $y = 0$ akwarium jest tak płaskie, że nawet rybka rozduszona przez kota się nie zmieści) więc rozwiązaniem jest $x = y = 4$. W tym punkcie drugie pochodne $s_{xx} = 128/x^3$, $s_{yy} = 128/y^3$ i $s_{xy} = 1$ tworzą formę kwadratową o macierzy

$$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix},$$

(nieprzypadkowo jest ona tą samą macierzą, którą otrzymaliśmy wyżej w wyniku redukcji formy drugich pochodnych, gdy szukaliśmy minimum warunkowego). Forma ta jest dodatnio okreĺona, czyli w punkcie $x = y = 4$ funkcja $s(x, y)$ ma minimum.

Zadanie 35: Najpierw posłużymy się metodą ekstremów warunkowych. Tworzymy funkcję $\tilde{F}(x, y, z) = xy^2z^3 - \lambda(x + 2y + 3z - 1)$ i przyrównujemy do zera jej oierwsze pochodne cząstkowe wraz z warunkiem

$$\begin{aligned} \tilde{F}_x(x, y, z) &= \ \ y^2z^3 - \lambda = 0\,, \\ \tilde{F}_y(x, y, z) &= 2(xyz^3 - \lambda) = 0\,, \\ \tilde{F}_z(x, y, z) &= 3(xy^2z^2 - \lambda) = 0\,, \\ G(x, y, z) &= x + 2y + 3z - 1 = 0\,. \end{aligned}$$

Równania $y^2z^3 = xyz^3 = xy^2z^2 = \lambda$ mają kilka rozwiązań.
i) Jeśli $x = 0$, to $\lambda = 0$ i albo $y = 0$ albo $z = 0$. Daje to dwa punkty $(0, 0, \frac{1}{3})$ oraz $(0, \frac{1}{2}, 0)$.
ii) Jeśli $y = 0$, to $\lambda = 0$ i jest cała rodzina rozwiązań $(1 - 3\alpha, 0, \alpha)$ z dowolnym α.
ii) Jeśli $z = 0$, to $\lambda = 0$ i jest cała rodzina rozwiązań $(1 - 2\alpha, \alpha, 0)$ z dowolnym α.
Widać też, że punkt $(0, 0, \frac{1}{3})$ należy do rodziny rozwiązań $(1 - 3\alpha, 0, \alpha)$ i odpowiada $\alpha = 1/3$, a punkt $(0, \frac{1}{2}, 0)$ należy do rodziny rozwiązań $(1-2\alpha, \alpha, 0)$ i odpowiada $\alpha = 1/2$. Jeśli ani x ani y ani z nie znika, to $y^2z^3 = \lambda = xyz^3$ daje $x = y$ i z $y^2z^3 = \lambda = xy^2z^2 \equiv y^3z^2$ mamy $y = z$, czyli $x = y = z = 1/6$ i $\lambda = 1/6^5$, ale wartoćć λ nie jest istotna dalej. Drugie pochodne funkcji $\tilde{F}$ (które tu są tożsame z drugimi pochodnymi funkcji F)

$$\begin{aligned} &\tilde{F}_{xx}(x, y, z) = 0\,, && \tilde{F}_{yy}(x, y, z) = 2xz^3\,, && \tilde{F}_{zz}(x, y, z) = 6xy^2z\,, \\ &\tilde{F}_{xy}(x, y, z) = 2yz^3\,, && \tilde{F}_{xz}(x, y, z) = 3y^2z^3\,, && \tilde{F}_{yz}(x, y, z) = 6xyz^2\,, \end{aligned}$$

Dają w punkcie $x = y = z = 1/6$ macierz Q formy kwadratowej

$$Q = \frac{1}{6^4}\begin{pmatrix} 0 & 2 & 3 \\ 2 & 2 & 6 \\ 3 & 6 & 6 \end{pmatrix}.$$

Ma ona sygnaturę $(0, -, +)$, nie jest więc nijak określona. Wektory $\mathbf{h}$ styczne do powierzchni więzów, na których trzeba ją badać spełniają warunek

$$G'|_{x=y=z=1/6} \cdot \begin{pmatrix} h_x \\ h_y \\ h_z \end{pmatrix} = (1,\, 2,\, 3) \begin{pmatrix} h_x \\ h_y \\ h_z \end{pmatrix} = h_x + 2h_y + 3h_z = 0\,.$$

Wstawiając $h_x = -2h_y - 3h_z$ do

$$Q(\mathbf{h}) = \frac{1}{6^4}\left(2h_y^2 + 6h_z^2 + 4h_xh_y + 6h_xh_z + 12h_yh_z\right),$$

dostajemy

$$Q(\mathbf{h}) = \frac{1}{6^3}\left(-h_y^2 - 2h_z^2 - 2h_yh_z\right) = -\frac{1}{6^3}\,(h_y, h_z)\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} h_y \\ h_z \end{pmatrix}.$$

Ponieważ wyrażenie to jest ujemnie określone, w punkcie $x = y = z = 1/6$ funkcja F ma na zbiorze E maksimum lokalne. W punktach $(0, 0, 1/3)$ i $(0, 1/2, 0)$ cała macierz drugich pochodnych funkcji $\tilde{F}$ jest zerowa i nie można na jej podstawie ustalić charakteru tych punktów krytycznych. Tak samo zerowa jest macierz drugich pochodnych w każdym punkcie krytycznym z rodziny $(1 - 2\alpha,\, \alpha, 0)$. W punktach krytycznych z rodziny $(1 - 3\alpha,\, 0\,\alpha)$ niezerowy jest tylko środkowy element tej macierzy (z wyjątkiem punktów $\alpha = 0$ i $\alpha = 1/3$, gdzie i ten element znika), więc też na wektorach postaci $\mathbf{h} = (0, 0, h)$ daje ona zero i nie można na jej podstawie określić charakteru tych punktów.

Jeśli z warunku $G(x, y, z) = 0$ wyznaczymy np. $x = 1 - 2y - 3z$ to badamy ekstrema zwykłej funkcji

$$f(y, z) = y^2z^3 - 2y^3z^3 - 3y^2z^4\,.$$

Przyrównanie do zera jej pierwszych pochodnych cząstkowch

$$\begin{aligned} f_y(y, z) &= 2yz^3 - 6y^2z^3 - 6y^2z^4 = 2yz^3(1 - 3y - 3z) = 0\,, \\ f_z(y, z) &= 3y^2z^2 - 6y^3z^2 - 12y^2z^3 = 3y^2z^3(1 - 2y - 4z) = 0\,, \end{aligned}$$

daje następujące punkty $(0, \alpha)$, $(\alpha, 0)$ z dowolnym α oraz $(1/6, 1/6)$. Widać, że odpowiadają one dokładnie znalezionym w metodzie ekstremów warunkowych dwóm rodzinom punktów $(1-2\alpha, \alpha, 0)$ i $(1-3\alpha, 0, \alpha)$ i punktowi $(1/6, 1/6, 1/6)$. Drugie pochodne funkcji $f(y, z)$

$$\begin{aligned} f_{yy}(y, z) &= 2z^2 - 12yz^3 - 6z^4\,, \\ f_{zz}(y, z) &= 6y^2z - 12y^3z - 36y^2z^2\,, \\ f_{yz}(y, z) &= 6yz^2 - 18y^2z^2 - 24yz^3\,, \end{aligned}$$

dają w punkcie $(1/6, 1/6)$ macierz formy kwadratowej

$$Q = -\frac{1}{6^3}\begin{pmatrix} 6 & 6 \\ 6 & 12 \end{pmatrix},$$

Rysunek 30: Kształt funkcji z Zadania 35. Oś y w zakresie $(-2, 2)$, oś z w zakresie $(-1, 1)$. Maksimum w punkcie $(1/6, 1/6)$ jest słabo widoczne.

która jest ujemnie określona - w punkcie tym funkcja f ma lokalne maksimum. Tak jak w poprzedniej metodzie, macierz drugich pochodnych w punktach rodziny $(1-2\alpha, \alpha, 0)$ jest całkowicie zerowa, a w punktach rodziny $(1-3\alpha, 0, \alpha)$ ma niezerowy tylko element Q_{yy} z wyjątkiem $\alpha = 1/3$, kiedy jest całkowicie zerowa. Jest jasne, że funkcja $f(z, y) = y^2 z^3 (1-2y-3z)$ nie może mieć uczciwych ekstremów, bo $f(0, \alpha) = f(\alpha, 0) = 0$ i znalezione dwie rodziny punktów krytycznych tworzą jej tzw. płaskie kierunki: z każdego takiego punktu można się przymieścić do dowolnie jemu bliskiego punktu sąsiedniego, w którym wartość funkcji jest taka sama. Kształt tej funkcji jest pokazany na rysunku 30.

Zadanie 36: Postępujemy regulaminowo: tworzymy funkcję

$$\begin{aligned} \tilde{F}(x, y, z) &= F(x, y, z) - \lambda_2 G^1(x, y, z) - \lambda_2 G^2(x, y, z) \\ &= xyz - \lambda_1(x + y + z - 5) - \lambda_2(xy + xz + yz - 8)\,, \end{aligned}$$

której pierwsze pochodne cząstkowe

$$\begin{aligned} \tilde{F}_x(x, y, z) &= yz - \lambda_1 - \lambda_2(y + z)\,, \\ \tilde{F}_y(x, y, z) &= xz - \lambda_1 - \lambda_2(x + z)\,, \\ \tilde{F}_z(x, y, z) &= xy - \lambda_1 - \lambda_2(x + y)\,, \end{aligned}$$

przyrównane do zera tworzą, wraz z warunkami ubocznymi układ pięciu równań wyznaczający punkty krytyczne i odpowiadające im wartości λ_1 i λ_2 (czyli razem wyznaczający pięć niewiadomych):

$$\begin{aligned} &yz - \lambda_1 - \lambda_2(y + z) = 0\,, \\ &xz - \lambda_1 - \lambda_2(x + z) = 0\,, \\ &xy - \lambda_1 - \lambda_2(x + y) = 0\,, \\ &x + y + z - 5 = 0\,, \\ &xy + xz + yz - 8 = 0\,. \end{aligned}$$

Z uwagi na symetrię funkcji względem permutacji zmiennych $F(y,z,x) = F(x,y,z) = F(x,z,y) = \ldots$, wydaje się naturalne szukanie punktu krytycznego postaci (x,x,x). Takie rozwiązanie wypisanego wyżej układu równań jednak nie istnieje, bo z dwóch ostatnich dostajemy wtedy równania $x = 5/3$ i $x^2 = 8/3$, które są sprzeczne. Trzeba zatem szukać innych rozwiązań. Np. postaci (a,a,b), (a,b,a), (b,a,a) respektujących w mniej trywialny sposób symetrię problemu.[67] Jeśli odejmiemy pierwsze równanie od drugiego, dostaniemy $(x-y)(z-\lambda_2) = 0$. Zatem albo $x = y$, albo $z = \lambda_2$. Zjamijmy się najpierw tą pierwszą możliwością. Dwa ostatnie równania redukują się do $z = 5 - 2x$ oraz $x^2 + 2xz - 8 = 0$. Razem dają one na x równanie kwadratowe

$$x^2 + 2x\,(5-2x) - 8 = -3x^2 + 10x - 8 = (x-2)(-3x+4) = 0\,.$$

Możliwe są więc rozwiązania $(2,2,1)$ i $(4/3,\ 4/3,\ 7/3)$. Z drugiego i trzeciego równania, które stają się teraz układem dwóch liniowych równań na λ_1 i λ_2 odczytujemy, że pierwszemu odpowiada $\lambda_1 = -4$ i $\lambda_2 = 2$, a drugiemu $\lambda_1 = -16/3$ i $\lambda_2 = 4$. Drugie pochodne funkcji $\tilde{F}$

$$\begin{aligned}
&\tilde{F}_{xx}(x,y,z) = 0\,, && \tilde{F}_{yy}(x,y,z) = 0\,, && \tilde{F}_{zz}(x,y,z) = 0\,,\\
&\tilde{F}_{xy}(x,y,z) = z - \lambda_2\,, && \tilde{F}_{xz}(x,y,z) = y - \lambda_2\,, && \tilde{F}_{yz}(x,y,z) = z - \lambda_2\,,
\end{aligned}$$

dają w punkcie krytycznym $(2,2,1)$, któremu odpowiada $\lambda_2 = 2$ następującą macierz Q formy kwadratowej drugich pochodnych

$$Q_{(2,2,1)} = \begin{pmatrix} 0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Jej określoność należy jednak badać na wektorach $\mathbf{h}$ spełniających warunki

$$\begin{aligned}
(G^1)'_{(2,2,1)} \cdot \mathbf{h} = (1,\ 1,\ 1) \begin{pmatrix} h_x \\ h_y \\ h_z \end{pmatrix} = 0\,,\\
(G^2)'_{(2,2,1)} \cdot \mathbf{h} = (3,\ 3,\ 4) \begin{pmatrix} h_x \\ h_y \\ h_z \end{pmatrix} = 0\,,
\end{aligned}$$

których rozwiązaniem jest $h_z = 0$, $h_y = -h_x$. Na takich wektorach przemieszczeń forma kwadratowa drugich pochodnych

$$Q_{(2,2,1)}(\mathbf{h}) = (h,\ -h,\ 0) \begin{pmatrix} 0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} h \\ -h \\ 0 \end{pmatrix} = 2h^2,$$

[67] W języku teorii grup powiedzielibyśmy, że rozwiązanie postaci (a,a,a) - gdyby istniało - stanowiłoby trywialną reprezentację grupy symetrii; rozwiązania postaci (a,a,b), (a,b,a), (b,a,a) stanowią zaś razem jej reprezentację trójwymiarową (co to jest reprezentacja grupy - zob. skrypt do algebry).

jest zawsze dodatnia, co oznacza, że funkcja F na zbiorze E ma w punkcie $(2,2,1)$, a także, z uwagi na jej symetrię, w punktach $(2,1,2)$ i $(1,2,2)$, minimum lokalne.

Analogicznie w punkcie $(4/3,\ 4/3,\ 7/3)$, któremu odpowiada $\lambda_2 = 4$ macierz formy kwadratowej drugich pochodnych ma postać

$$Q_{(4/3,4/3,7/3)} = \frac{1}{3}\begin{pmatrix} 0 & -5 & -8 \\ -5 & 0 & -8 \\ -8 & -8 & 0 \end{pmatrix},$$

a wektory, na których należy badać jej określoność spełniają warunki

$$(G^1)'_{(2,2,1)} \cdot \mathbf{h} = (1,\ 1,\ 1)\begin{pmatrix} h_x \\ h_y \\ h_z \end{pmatrix} = 0\,,$$
$$(G^2)'_{(2,2,1)} \cdot \mathbf{h} = \frac{1}{3}(11,\ 11,\ 8)\begin{pmatrix} h_x \\ h_y \\ h_z \end{pmatrix} = 0\,,$$

których rozwiązaniem znów jest jest $h_z = 0$, $h_y = -h_x$. Na takich wektorach

$$Q_{(4/3,4/3,7/3)}(\mathbf{h}) = (10/3)h^2\,,$$

i znów w punkcie $(4/3, 4/3, 7/3)$ oraz w $7/3, 4/3, 4/3)$ i $(4/3, 7/3, 4/3)$ funkcja F ma na zbiorze E (lokalne) minimum.

Można też pokazać, że ostatnia możliwość, $z = \lambda_2$ nie prowadzi do żadych rozwiązań. Dwa pierwsze równania redukują się wtedy do jednego i tego samego i ukkład równań przechodzi w ($a \equiv x + y$, $b \equiv xy$):

$$\lambda_1 + \lambda_2^2 = 0\,, \quad \lambda_1 + a\lambda_2 - b = 0\,, \quad \lambda_2 = 5 - a\,, \quad a\lambda_2 + b = 8\,,$$

Ostatnie z pomocą drugiego można wtedy przerobić na $\lambda_1 = 2b - 8$ i teraz korzystając z tego i z trzeciego podstawiamy do pierwszego i drugiego λ_1 i λ_2, co daje dwa równania:

$$\begin{aligned} 2b - 8 + (5-a)^2 \quad &\equiv a^2 + 2b - 2a + 17 = 0\,, \\ 2b - 8 + a(5-a) - b &\equiv -a^2 + b + 5a - 8 = 0\,. \end{aligned}$$

No to jeszcze je do siebie dodać, co da $3b+3a+9 = 0$, czyli $b = -a-3$ i to do pierwszego:

$$a^2 + 2(-a-3) - 2a + 17 \equiv a^2 - 4a + 11 = 0\,,$$

i teraz już widać, że niema rozwiązań.

Zadanie 37: Składowe (w kanonicznej bazie przestrzeni $\mathbb{R}^3$) wektorów stycznych do podanej powierzchni otrzymujemy wóżniczkując definiujące powierzchnię związki:

$$\mathbf{t}_t = \begin{pmatrix} a\omega\cos\omega t\cos\xi \\ \omega\cos\omega t\sin\xi \\ -a\omega\sin\omega t + a\omega/\sin\omega t \end{pmatrix}, \qquad \mathbf{t}_\xi = \begin{pmatrix} -a\sin\omega t\sin\xi \\ \omega\cos\omega t\cos\xi \\ 0 \end{pmatrix}.$$

Wektory te mogą przestać być liniowo niezależne tylko w takich punktach, w których zeruje się ostatnia składowa wektora $\mathbf{t}_t$. To zaś zdarza się tam, gdzie $\sin^2 \omega t = 1$, czyli gdy $\omega t = \frac{\pi}{2} + k\pi$. W takich punktach $\cos \omega t = 0$ i cały wektor $\mathbf{t}_t = \mathbf{0}$ i rzeczywiście w takich punktach wymiar (pod)przestrzeni $T_P M$ staje się równy 1, a nie 2, jak w innych punktach. Punkty takie są więc punktami osobliwymi badanej powierzchni. Zauważmy też, że w punktach, w których $\omega t = l\pi$ ostatnia składowa wektora $\mathbf{t}_t$ staje się nieskończona. W tych punktach jednak $z \to \infty$, czyli w kierunku z-owym powierzchnia rozciąga się do nieskończoności, która jest osiągana w skończonym "czasie t" - nie są to więc punkty osobliwe powierzchni jako takiej, tylko raczej pewna osobliwość układu współrzędnych, którymi pokryliśmy tę powierzchnię. Takie rzeczy się zdarzają często przy badaniu różnych czasoprzestrzeni w ramach ogólnej teorii względności.

Zadanie 38:
a) Rozdzielamy zmienne i całkujemy

$$\int \frac{dy}{1+y^2} = \int dx\, x\,.$$

Całkować już umiemy, zatem dostajemy

$$\operatorname{arctg} y = \frac{1}{2}\, x^2 + C\,, \qquad \text{czyli} \quad y = \operatorname{tg}\left(\frac{1}{2}\, x^2 + C\right).$$

Widać, że stała C jest określona modulo π, bo funkcja tangens jest okresowa. Oczywiście wydaje się, że każda funkcja (scharakteryzowana konkretną wartością C z przedziału $[0, \pi)$) to jest nieskończenie wiele krzywych całkowych (bo tak Mathematica maluje funkcję tangens). Ale to nie o to chodzi: chodzi o to, że przez każdy punkt przebiega jedna i tylko jedna krzywa całkowa (bo funkcja $f(x, y)$ w $dy/dx = f(x, y)$ jest super regularna). Np. przez punkt $(7, 1)$ przebiega krzywa scharakteryzowana przez $C = \pi/4 - 49/2$, i tylko to się liczy, bo równanie należy właśnie rozumieć w sensie warunku ograniczającego ruchy, gdy się jest w jakimś punkcie: gdy jesteśmy w punkcie $(7, 1)$ to możemy iść tylko po jednej z nieskończenie wielu rozłącznych krzywych, jakie daje formalny wzór $y = \operatorname{tg}(\frac{1}{2}x^2 - 49/4 + \pi/4)$. Krzywą tę można też dostać wykonując całki oznaczone, tj. pisząc

$$\int_{y_0}^{y(x)} \frac{d\xi}{1+\xi^2} = \int_{x_0}^{x} d\eta\, \eta\,,$$

i całkując już bez konieczności dobierania później stałych.

b) Tak jak poprzednio, rozdzielamy zmienne i całkujemy

$$\int \frac{dy}{y} = \int dx\, x\,,$$

dostając

$$\ln|y| = \frac{1}{2}\, x^2 + C\,, \qquad \text{czyli} \quad y = \pm\exp\left(\frac{1}{2}\, x^2 + C\right).$$

$y \equiv 0$ też jest rozwiązaniem - włączamy je w całkę ogólną pisząc $\pm e^C = \tilde{C}$ ($y \equiv 0$ teraz odpowiada $\tilde{C} = 0$). Przez każdy punkt płaszczyzny xy przechodzi dokładnie jedna krzywa całkowa (wszystkie one są uczciwymi funkcjami $y = y(x)$) bo funkcja $f(x, y) = xy$ jest regularna na całym $\mathbb{R}^2$.

c) Znów rozdzielamy zmienne i całkujemy

$$\int \frac{dy}{y} = \int \frac{dx}{x}\,,$$

dostając

$$\ln|y| = \ln|x| + C \equiv \ln(|\tilde{C}\,x|)\,, \qquad \text{czyli} \quad \ln\left|\frac{y}{\tilde{C}x}\right| = 0\,,$$

co jest równoważne $|y/\tilde{C}x| = 1$, czyli po prostu

$$y = \tilde{C}\,x\,.$$

Widać, że przez każdy punkt płaszczyzny xy z wyjątkiem punktów $(0, y)$ przechodzi jedna funkcja $y = y(x)$. Przez punkty na osi x też, bo robiąc manewry ze stałą całkowania (przechodząc od C do $\tilde{C}$) chytrze upchnęliśmy w całkę ogólną rozwiązanie $y \equiv 0$, którego pierwotna forma tej całki nie obejmowała. Przez punkty $(0, y)$ o $y \neq 0$ nie przechodzi żadna funkcja obejmowana przez całkę ogólną, a przez punkt $(0, 0)$ przechodzi ich nieskończenie wiele. Znów oznacza to, że w punkcie $(0, 0)$ funkcja $f(x, y) = x/y$ może być dookreślona tak, by być ciągłą w dowolnym (ale tylko jednym) kierunku, a w punktach $(0, y)$ o $y \neq 0$ nie może. Ale oczywiście jak się pytamy nie o funkcje tylko o krzywe całkowe, to równania zapisane w postaci $x\,dy = y\,dx$ jest zupełnie symetryczne względem zamiany $x \leftrightarrow y$ i punktu na osi y (i $y \neq 0$) powinny mieć taki sam charakter, jak punkty na osi x (i $x \neq 0$). I rzeczywiście tak jest, bo $x \equiv 0$ jest też krzywą całkową. Rozwiązania przechodzące przez dowolny punkt (x_0, y_0) nie leżący na żadnej z osi można oczywiście dostać wykonując całki oznaczone:

$$\int_{y_0}^{y(x)} \frac{d\xi}{\xi} = \int_{x_0}^{x} \frac{d\eta}{\eta}\,,$$

co daje $\ln|y/y_0| = \ln|x/x_0|$.

d) Analogicznie jak w poprzednim przypadku rozdzielenie zmiennych prowadzi, po scałkowaniu, do wyniku

$$y = \frac{C}{x}\,, \qquad \text{albo, w symetrycznej postaci} \quad x\,y = C\,.$$

Teraz przez punkt $(0, 0)$ przechodzą tylko dwie krzywe całkowe $y = 0$ i $x = 0$ (obie odpowiadające $C = 0$). Warto ten przykład skonfrontować z poprzednim: w obu tych przypadkach funkcję $f(x, y)$ w równaniu $y' = f(x, y)$ można w punkcie $(0, 0)$ dookreślić

Rysunek 31: Krzywe całkowe równania $y' = -y/x$.

tak, by była w nim ciągła wzdłuż dowolnie wybranego kierunku, czyli może ona w tym punkcie wyznaczać dowolne nachylenie krzywej całkowej. Niemniej tam przez punkt ten przechodziło nieskończenie wiele krzywych (każda o innym nachyleniu w tym punkcie), a tu tylko dwie. Tak może być, bo możliwość dookreślenia funkcji w punkcie w dowolny sposób jest tylko warunkiem koniecznym na to, by przez taki punkt mogło przechodzić nieskończenie wiele krzywych całkowych, ale nie warunkiem dostatecznym. W istocie w przykładzie teraz rozptrywanym krzywe całkowe obejmowane przez wypisaną wyżej całkę ogólną są pokazanymi na rysunku 31 hiperbolami, które pozornie biegną ku punktowi $(0,0)$, a zawsze ostatecznie go "omijają", odchylając się w bok.[68] Z punktu $(0,0)$ można tylko "ruszyć" wzdłuż osi x lub y; próba ruszenia (np. przez numeryczne całkowanie równania) w ciut innym kierunku jest niewykonalna, bo tuż po wyjściu z punktu $(0,0)$ w jakimś innym kierunku kończy się natychmiast (tj. w punkcie infinitezymalnie bliskim punktowi $(0,0)$) natrafieniem na kierunek całkowicie inny niż ten, w którym zamierzaliśmy się udać.[69]

e) W tym przykładzie rozdzielenie zmiennych prowadzi do całek

$$\int \frac{dy}{y} = \int \frac{dx}{x^3}, \qquad \text{co daje} \quad \ln|y| = -\frac{1}{2x^2} + C\,,$$

czyli

$$y(x) = \pm \exp\left(-\frac{1}{2x^2} + C\right) \equiv \tilde{C} \exp\left(-\frac{1}{2x^2}\right).$$

[68]Wygląda to jak bieg stałych sprzężenia w pobliżu odpychającego punktu stałego w równaniach grupy renormalizacji...

[69]Jeśli jeszcze chwilę o tym pomyśleć, to zauważy się, że różnica między dwoma przypadkami polega na tym, że *wartość* funkcji y/x wynikająca z ciągłości wzdłuż wybranego kierunku jest dokładnie taka, że zadaje ona dokładnie to nachylenie, które odpowiada wybranemu kierunkowi zbiegania do $(0,0)$; w przypadku funkcji $-y/x$ - jej *wartość* zadaje nachylenie odpowiadające prostej ortogonalnej do kierunku zbiegania.

Rysunek 32: Krzywe całkowe równania $y' = y/x^3$ obejmowane przez całkę ogólną.

Znów zastąpiliśmy czynnik $\pm e^C$ przez stałą $\tilde{C}$ dopuszczając oba jej znaki i znów w ten sposób włączyliśmy w całkę ogólną funkcję $y \equiv 0$ (które teraz odpowiada $\tilde{C} = 0$), której pierwotnie całka ta nie obejmowała (jeśli nie dopuszczało się $C = -\infty$). Rodzinę funkcji obejmowanych przez całkę zupełną pokazuje rysunek 32. Przez wszystkie punkty płaszczyzny xy z wyjątkiem punktów $(0, y)$ przechodzi tylko jedna funkcja $y = y(x)$; przez punkty $(0, y)$ o $y \neq 0$ nie przechodzi żadna, a przez punkt $(0, 0)$ przechodzi ich nieskończenie wiele (wszystkie rozwiązania obejmowane prze całkę ogólną). Jest to spowodowane tym, że $f(x, y) = y/x^3$ nie jest ciągła (z żadnego kierunku) w punktach $(0, y)$ o $y \neq 0$, a w punkcie $(0, 0)$ można jej nadać wartość 0 czyniąc ją ciągłą tylko wzdłuż osi x - wszystkie krzywe całkowe obejmowane przez całkę ogólną mają w $x = 0$ nachylenie równe zero (pamiętamy: funkcja $\exp(-1/x^2)$ ma w zerze wszystkie pochodne równe zero). Oczywiście równanie zapisane w postaci $x^3 dy = y\, dx$ ma także jako rozwiązanie prostą $x \equiv 0$. Rozwiązanie spełniające warunek $y(1) = 1$ to $y(x) = \exp((1 - 1/x^2)/2)$; odpowiada ono $\tilde{C} = e^{1/2}$.

f) W tym przykładzie rozdzielenie zmiennych prowadzi do całek

$$\int \frac{dy}{y} = \int \frac{dx}{x^2}, \qquad \text{co daje} \quad \ln|y| = -\frac{1}{x} + C\,,$$

czyli

$$y(x) = \pm \exp\left(-\frac{1}{x} + C\right) \equiv \tilde{C} \exp\left(-\frac{1}{x}\right).$$

Znów tym manewrem ze stałą całkowania włączyliśmy do całki ogólnej rozwiązanie $y \equiv 0$. W tym przypadku jednak wszystkie rozwiązania obejmowane przez całką ogólną są nieciągłe w $x = 0$, ale nadal przez każdy punkt (x_0, y_0) płaszczyzny xy, z wyjątkiem punktów $(0, y)$ o $y \neq 0$ przechodzi dokładnie jedno rozwiązanie $y = y_0 \exp(-1/x + 1/x_0)$. Rozwiązania te pokazuje rysunek 33. Teraz też funkcja $f(x, y) = y/x^2$ może być "uciąglona"

Rysunek 33: Krzywe całkowe równania $y' = y/x^3$ obejmowane przez całkę ogólną.

w punkcie $(0,0)$ wzdłuż kierunku $y = 0$ przez nadanie jej tam wartości 0 (w punktach $(0,y)$ o $y \neq 0$, funkcja $f(x,y) = y/x^2$ oczywiście, nie może być "uciąglona"). Umożliwia to zbieganie się w $(0,0)$ krzywych całkowych $y = y(x)$ z obszaru $x \geq 0$; wszystkie one mają tam nachylenie równe zero - pochodna funkcji $y = \tilde{C}\exp(-1/x)$ ma w $x = 0$ wartość zero.[70] Poza tym, równanie przepisane w postaci $x^2 dy = y\, dx$ ma jako swoje rozwiązanie także prostą $x \equiv 0$.

g) Oczywiście $y \equiv 0$ jest możliwym rozwiązaniem. Aby rozdzielić zmienne zakładamy, że $y \neq 0$, że $x \neq n\pi$ (na razie szukamy takich rozwiązań). Rozdzielenie prowadzi do całek

$$\int \frac{dy}{y} = \int dx\, \frac{\cos x}{\sin x}\,, \qquad \text{czyli} \qquad \ln|y| = \ln|\sin x| + C = \ln|\tilde{C}\sin x|\,.$$

Zatem (znów manewr ze stałą całkowania pozwolił włączyć do całki ogólnej rozwiązanie $y \equiv 0$)

$$y = \tilde{C}\sin x \,= y_0\,\frac{\sin x}{\sin x_0}\,.$$

Druga postać pokazuje, że jeżeli tylko $x_0 \neq n\pi$, stałą $\tilde{C}$ można tak dobrać, by rozwiązanie przechodziło przez punkt (x_0, y_0). A co z punktami o $(x_0, y) = (n\pi, y)$? Jeśli równanie przepiszemy w formie

$$\sin x\, dy = y\cos x\, dx\,,$$

[70]Znów można próbować zrozumieć, dlaczego w punkcie $(0,0)$ nie zbiegają się krzywe całkowe z obszaru $x < 0$. Gdy ruszamy z punktu $(0,0)$ płasko w kierunku dodatnich x-ów, nachylenie zerowe jest tylko infinitezymalnie różne od nachyleń zadawanych przez funkcję y/x^2 tuż pod osią x-ów i tuż nad nią; gdy zaś ruszymy z punktu $(0,0)$ z zerowym nachyleniem w kierunku ujemnych x-ów, nachylenie to różni się od tego, jakie zadaje funkcja y/x^2 tuż pod osią x-ów i tuż nad nią: np. gdybyśmy chcieli "wystartować i wznieść" się nad oś x-ów, to taka wznosząca się krzywa całkowa musiała by mieć nad osią nachylenie ujemne, a funkcja y/x^2 nad osią x zadaje nachylenie dodatnie!

Rysunek 34: Funkcje $z(x) = 2\mathrm{arctg}(x + C)$, $C = 0 \pm 2$.

to od razu zobaczymy, że krzywymi całkowymi są też proste $x(y) = n\pi$. Ostatecznie więc przez wszystkie punkty $(x, y) \neq (n\pi, 0)$ przechodzi tylko jedna krzywa całkowa tego równania, a przez punkty $(n\pi, 0)$ przechodzi ich nieskończenie wiele. Wynika to z tego, że $f(x, y) = y\,\mathrm{ctg}x$ jako funkcja na $\mathbb{R}^2$ jest nieciągła w punktach $(n\pi, y)$, ale że $\mathrm{ctg}x \sim 1/x$, gdy $x \approx 0$ (wszystkie pozostałe punkty o $x = n\pi$ są dzięki periodyczności funkcji $\mathrm{ctg}x$ takie same - wystarczy więc rozpatrzyć $x = 0$), funkcję $f(x, y)$ można dookreślić w punktach $(n\pi, 0)$ tak by była ciągła w (jednym) dowolnie wybranym kierunku. To dlatego przez punkty te punkty może przechodzić nieskończenie wiele funkcji $y = y(x)$. Za to przez punkty $(n\pi, y)$ o $y \neq 0$ może przechodzić tylko jedna funkcja $x = x(y)$, bo w równaniu $dx/dy = g(y, x) = \mathrm{tg}x/y$ funkcja $g(y, x)$ jest w takich punktach zupełnie regularna.

Zadanie 39:
a) Po podstawieniu $y = -x + z(x)$ do $y' = \cos(x + y)$ dostajemy $z' = 1 + \cos z$ i to już jest równanie o zmiennych rozdzielonych. Całkujemy

$$\int \frac{dz}{1 + \cos z} = \int dx\,.$$

To wymaga przypomnienie sobie stosowanych do obliczania takich całek podstawień... Pamiętają Państwo? Podstawia się $t = \mathrm{tg}(z/2)$. Wtedy

$$dz = \frac{2dt}{1 + t^2}\,, \quad \sin z = \frac{2t}{1 + t^2}\,, \quad \cos z = \frac{1 - t^2}{1 + t^2}\,.$$

Całka po lewej przechodzi wtedy w

$$\int \frac{dz}{1 + \cos z} = \int dt = \mathrm{tg}(z/2)\,.$$

Zatem

$$y(x) = -x + 2\,\mathrm{arctg}(x + C)\,.$$

Sprawdźmy to.

$$y' = -1 + \frac{2}{1 + (x + C)^2}\,,$$
$$\cos(x + y) = \cos(2\,\mathrm{arctg}(x + C)) = -1 + 2\cos^2\left(\mathrm{arctg}(x + C)\right),$$

ale $\cos^2\alpha = (1+\mathrm{tg}^2\alpha)^{-1}$, więc

$$\cos(x+y) = -1 + \frac{2}{1+\mathrm{tg}^2[\mathrm{arctg}(x+C)]} = -1 + \frac{2}{1+(x+C)^2}\,.$$

Zgadza się.

No a jak to jest z przechodzeniem krzywych całkowych przez punkty płaszczyzny xy? W równaniu $y' = \cos(x+y)$ funkcja po prawej stronie jest na całym $\mathbb{R}^2$ regularna i zgodnie z twierdzeniem, przez każdy jej punkt powinna przechodzić jedna krzywa będąca uczciwą funkcją $y = y(x)$. To samo powinno być prawdą i na płaszczyźnie xz (i jedno implikuje drugie) i to łatwiej przeaanalizować. Na pozór tak nie jest, bo przecież funkcje $z = 2\,\mathrm{arctg}(x+C)$ wyglądają tak, jak na rysunku 34. No ale to dlatego, że to jest tylko funkcja $\mathrm{arctg}x$ zdefiniowana konwencjonalnie; wiemy, że można ją zdefiniować tak, by przyjmowała wartości nie z przedziału $(-\pi/2, \pi/2)$, tylko z dowolnego z przedziałów $(n\pi - \pi/2, n\pi + \pi/2)$, bo $\mathrm{tg}(\alpha + n\pi) = \mathrm{tg}\alpha$. Ale to jeszcze nie załatwia sprawy punktów (x, z) o $z = -\pi + 2k\pi$: żadna z tych funkcji $2\mathrm{arctg}(x+C)$ nie przechodzi przez takie punkty. Niemniej, jak spojrzymy jeszcze raz na równanie $z' = 1 + \cos z$, to zobaczymy, że jego rozwiązaniami są również (nieobejmowane przez całkę ogólną) funkcje $z(x) = \pi + 2k\pi$ (bo wtedy $\cos z = -1$). I teraz już wszystko jest w porządku: przez każdy punkt płaszczyzny xz (a co ta tym idzie i płaszczyzny xy) przechodzi dokładnie jedno rozwiązanie równania $z' = -1 + \cos z$ (równania $y' = \cos(x+z)$).

b) Podstawiamy $y(x) = u(x)/x$ do $y' = -y/x + \cos(xy)/x^2$:

$$\frac{u'}{x} - \frac{u}{x^2} = -\frac{u}{x^2} + \frac{\cos u}{x^2}\,.$$

Równanie upraszcza się do równania $x\,u' = \cos u$ o zmiennych rozdzielonych. Całkujemy więc

$$\int \frac{du}{\cos u} = \int \frac{dx}{x}\,,$$

Znów całka po lewej stronie wymaga podstawienia $t = \mathrm{tg}(u/2)$:

$$\int \frac{du}{\cos u} = 2\int \frac{dt}{1-t^2} = \int dt\left(\frac{1}{1+t} + \frac{1}{1-t}\right) = \ln|1+t| - \ln|1-t|\,.$$

Zatem

$$\ln\frac{|1+t|}{|1-t|} = C + \ln|x| = \ln|\tilde{C}\,x|\,,$$

i stąd, odwracając kota do góry ogonem,

$$\frac{\tilde{C}\,x+1}{\tilde{C}\,x-1} = t = \mathrm{tg}(u/2)\,.$$

Ostatecznie więc

$$u = 2\,\mathrm{arctg}\left(\frac{\tilde{C}\,x+1}{\tilde{C}\,x-1}\right).$$

Jak zwykle sprawdzamy. Lewa strona równania $x\,u' = \cos u$:

$$\begin{aligned} x\,u' &= \frac{2x}{1+(\tilde{C}\,x+1)^2/(\tilde{C}\,x-1)^2}\left[\frac{\tilde{C}\,(\tilde{C}\,x+1)}{(\tilde{C}\,x-1)^2} - \frac{\tilde{C}\,(\tilde{C}\,x-1)}{(\tilde{C}\,x-1)^2}\right] \\ &= \frac{4\tilde{C}\,x}{(\tilde{C}\,x+1)^2+(\tilde{C}\,x-1)^2}\,. \end{aligned}$$

Z kolei prawa strona to:

$$\begin{aligned} \cos u &= \cos\left[2\,\mathrm{arctg}\left(\frac{\tilde{C}\,x+1}{\tilde{C}\,x-1}\right)\right] = -1+2\cos^2\left[\mathrm{arctg}\left(\frac{\tilde{C}\,x+1}{\tilde{C}\,x-1}\right)\right] \\ &= -1+\frac{2}{1+\mathrm{tg}^2[\mathrm{arctg}((\tilde{C}\,x+1)/(\tilde{C}\,x-1))]} \\ &= \frac{2}{1+(\tilde{C}\,x+1)^2/(\tilde{C}\,x-1)^2} - 1 = \frac{4\tilde{C}\,x}{(\tilde{C}\,x+1)^2+(\tilde{C}\,x-1)^2}\,. \end{aligned}$$

Zgadza się. Rozwiązaniem wyjściowego równania jest więc

$$y(x) = \frac{2}{x}\,\mathrm{arctg}\left(\frac{\tilde{C}\,x+1}{\tilde{C}\,x-1}\right).$$

Znów możemy zapytać, czy przez każdy punkt płaszczyzny xy przechodzi jedna krzywa całkowa? W wyjściowym równaniu prawa strona nie jest regularna na osi y, tj. w punktach $(0, y)$ i przez te punkty nie przechodzi żadna z funkcji $y = y(x)$ obejmowanych przez znalezioną wyżej całkę ogólną. Niemniej, gdy się przepisze równania w postaci $x^2 dy = (-xy + \cos(xy))dx$, to widać, że prosta $x \equiv 0$ spełnia je też. W pozostałych punktach płaszczyzny funkcja po prawej stronie równania $y' = f(x, y)$ jest zupełnie regularna i przez każdy taki punkt przechodzi jedna tylko funkcja (choć może to wymagać innej niż konwencjonalna definicji funkcji $\mathrm{arctg}\,x$).

c) Tu chodzi o zamianę zmiennej niezależnej. Wyobrażamy sobie, że $y(x) = \tilde{y}(t(x))$, gdzie $t(x) = \ln x$. Zatem

$$\begin{aligned} y' &= \frac{d\tilde{y}}{dt}\,\frac{dt}{dx} = \frac{1}{x}\,\frac{d\tilde{y}}{dt} = e^{-t}\,\frac{d\tilde{y}}{dt}\,, \\ y'' &= \frac{d}{dx}\left(\frac{1}{x}\,\frac{d\tilde{y}}{dt}\right) = -\frac{1}{x^2}\,\frac{d\tilde{y}}{dt} + \frac{1}{x^2}\,\frac{d^2\tilde{y}}{dt^2} = e^{-2t}\left(\frac{d^2\tilde{y}}{dt^2} - \frac{d\tilde{y}}{dt}\right). \end{aligned}$$

I teraz te pochodne wstawiamy do równania $x^2y'' + x\,y' + y = 0$:

$$e^{2t}\left[e^{-2t}\left(\frac{d^2\tilde{y}}{dt^2} - \frac{d\tilde{y}}{dt}\right)\right] + e^t\left[e^{-t}\,\frac{d\tilde{y}}{dt}\right] + \tilde{y} = 0\,.$$

Wychodzi więc równanie

$$\frac{d^2\tilde{y}}{dt^2} + \tilde{y} = 0\,,$$

które każdy fizyk już umie rozwiązać: $\tilde{y}(t) = A\cos t + B\sin t$, czyli

$$y(x) = A\cos(\ln x) + B\sin(\ln x)\,.$$

A co, gdy $x < 0$? A nic. Tzn. dokładnie tak samo: można było z mety podstawić $t = \ln|x|$. Można więc napisać

$$y(x) = A\cos(\ln|x|) + B\sin(\ln|x|)\,.$$

Można było się w tym zorientować od razu, bo wyjściowe równanie ma symetrię $x \leftrightarrow -x$. Dokładniej, jeśli $y = f(x)$ jest jego rozwiązaniem, to $y = f(-x)$ też jest: $d^2f(-x)/dx^2 = d^2f(-x)/d(-x)^2$, a $df(-x)/dx = -df(-x)/d(-x)$ ale $x\,df(-x)/dx = (-x)\,df(-x)/d(-x)$. Wiec jak $f(x)$ spełnia równanie, to $f(-x)$ też. To jak z równaniami mechaniki (gdy niema sił tarcia): jeśli $\mathbf{r}(t)$ jest rozwiązaniem, to $\mathbf{r}(-t)$ też: nie sposób powiedzieć, czy film jest puszczony do przodu, czy do tyłu. A co z jajkiem, co spada ze stołu i robi się z niego jajecznica? Też nie można odróżnić czy film był puszczony od tyłu? No i tu wchodzimy w problemy fizyki statystycznej i tego, czy entropia może zmaleć (trzeba tylko jeszcze się umówić, o której entropii mówimy)... Ale kluczem jest to, że jajko to jest *circa* 10^{23} cząstek. Pozostańmy więc przy jednej do kilku najwyżej i wtedy jesteśmy bezpieczni. Tylko $x = 0$ jest trefne, ale tu nie mamy (tzn. MY nie znamy, bo o tym nie było na wykładzie) twierdzeń, ale pewnie jakieś założenia są niespełnione...

d) Znów sobie wyobrażamy, że $y(x) = \tilde{y}(t(x))$ gdzie $t = \arccos x$. No, tu x musi być pomiędzy -1 i 1. Tylko w tym obszarze możemy tak podstawić. Zatem $dt/dx = -1/\sqrt{1-x^2} = -1/|\sin t|$. I jedziemy:

$$y' = \frac{d\tilde{y}}{dt}\,\frac{dt}{dx} = -\frac{1}{\sqrt{1-x^2}}\,\frac{d\tilde{y}}{dt}\,,$$

$$y'' = \frac{d}{dx}\left(-\frac{1}{\sqrt{1-x^2}}\,\frac{d\tilde{y}}{dt}\right) = -\frac{x}{(1-x^2)^{3/2}}\,\frac{d\tilde{y}}{dt} + \frac{1}{1-x^2}\,\frac{d^2\tilde{y}}{dt^2}\,.$$

Podstawiamy

$$(1-x^2)\left[\frac{1}{1-x^2}\,\frac{d^2\tilde{y}}{dt^2} - \frac{x}{(1-x^2)^{3/2}}\,\frac{d\tilde{y}}{dt}\right] - x\left[-\frac{1}{\sqrt{1-x^2}}\,\frac{d\tilde{y}}{dt}\right] + \omega^2\tilde{y} = 0\,.$$

Wychodzi

$$\frac{d^2\tilde{y}}{dt^2} + \omega^2\tilde{y} = 0\,,$$

I to znów umiemy to rozwiązać

$$y(x) = A\cos(\omega \arccos x) + BA\cos(\omega \arccos x)\,.$$

e) Znów sobie wyobrażamy, że $y(x) = \tilde{y}(t(x))$ i teraz $t = 1/x$

$$\begin{aligned}
\frac{dy}{dx} &= -\frac{1}{x^2}\,\frac{d\tilde{y}}{dt}\,,\\
\frac{d^2y}{dx^2} &= \frac{d}{dx}\left(-\frac{1}{x^2}\,\frac{d\tilde{y}}{dt}\right) = \frac{2}{x^3}\,\frac{d\tilde{y}}{dt} + \frac{1}{x^4}\,\frac{d^2\tilde{y}}{dt^2}\,.
\end{aligned}$$

Wstawiamy do równania $x^4y'' + 2x^3y' + x^2y = 0$:

$$x^4\left[\frac{1}{x^4}\,\frac{d^2\tilde{y}}{dt^2} + \frac{2}{x^3}\,\frac{d\tilde{y}}{dt}\right] + 2x^3\left[-\frac{1}{x^2}\,\frac{d\tilde{y}}{dt}\right] + x^2y = 0\,,$$

Czyli

$$\frac{d^2\tilde{y}}{dt^2} + \frac{\tilde{y}}{t^2} = 0\,.$$

Tego to chyba nie umiemy rozwiązać.

f) A to akurat jest proste, choć wygląda na skomplikowane. Podstawić należy (udało mi się to zgadnąć - widać jakąś orientację w terenie mam) $f = e^y$. Wtedy

$$f' = e^y\,y'\,, \qquad f'' = e^y\,(y')^2 + e^y\,y'' = e^y\left[y'' + (y')^2\right].$$

Więc wyjściowe równanie $y'' + (y')^2 = 2\,e^{-y}$, to po prostu $f'' = 2$. "As simple as that" powiedzieliby rodacy pana Pepysa (takie zdanie gdzieś u Herlinga-Grudzińskiego zapamiętałem). Stąd $f = A + B\,x + x^2$ i

$$y = \ln(A + B\,x + x^2)\,.$$

Żeby spełnić warunki początkowe trzeba wziąć $A = 2$ i $B = 2$.

g) Równanie $y'' = 2\,y^3$ wygląda jak równanie ruchu w jednym wymiarze punktu o masie $m = 1$ pod wpływem siły $2\,y^3$, która jest potencjalna, tzn. $2\,y^3 = -dV(y)/dy$, gdzie $V = -y^4/2$ (tylko zamiast czasu t zmienna się nazywa x). Zatem "energia" mechaniczna powinna być zachowana, czyli

$$\frac{d}{dx}\left(\frac{1}{2}\,(y')^2 + \frac{1}{2}\,y^4\right) = 0\,.$$

Formalnie można się o tym przekonać mnożąc równanie $y'' = 2\,y^3$ stronami przez y' i zauważając, że

$$y'y'' = \frac{d}{dx}\left(\frac{1}{2}\,(y')^2\right), \qquad \text{i podobnie} \quad 2\,y'y^3 = \frac{d}{dx}\left(\frac{1}{2}\,y^4\right).$$

Zatem mamy równanie pierwszego rzędu $(y')^2 = C + y^4$, czyli

$$y' = \pm\sqrt{C + y^4}\,,$$

które jest równaniem o zmiennych rozdzielonych. Zanim scałkujemy je, od razu zauważamy, że warunki początkowe $y(0) = 1$, $y'(0) = 1$, czyli $y' = 1$, gdy $y = 1$, są tak dobrane, by $C = 0$. Poza tym trzeba wybrać znak $+$. Zatem

$$\int dx = \int \frac{dy}{y^2} = -\frac{1}{y} + A\,.$$

(Gdyby nie to, że $C = 0$, nie umielibyśmy całki po prawej stronie obliczyć). No i teraz, ponieważ $y = 1/(A - x)$ znów zauważamy, że $A = 1$, by $y(0) = 1$. Rozwiązaniem spełniającym warunki początkowe jest zatem $y = 1/(1 - x)$.

h) Jeśli $u = \mathrm{tg}(y/2)$, czyli $y = 2\,\mathrm{arctg}(u)$, to tak jak przy uniwersalnym podstawieniu w całkach z funkcji trygonometrycznych

$$y' = \frac{2\,u'}{1+u^2}\,, \qquad \sin y = \frac{2u}{1+u^2}\,, \qquad \cos y = \frac{1-u^2}{1+u^2}\,,$$

i gdy to podstawimy do równania $y' + \sin y + x\cos y + x = 0$, to dostaniemy równanie liniowe z niejednorodnością

$$u' + u = -x\,,$$

Równanie jednorodne po rozdzieleniu zmiennych $du/u = -dx$ daje $\ln|u| = C' - x$, czyli $(\pm e^{C'} = C)$

$$u_{\mathrm{hom}} = C\exp(-x)\,.$$

Do niejednorodnego podstawiamy Ansatz $u_{\mathrm{inhom}} = h(x)\,e^{-x}$ i dostajemy

$$h' = -x\,e^{x}\,, \qquad \text{czyli} \quad h = (1 - x)\,e^{x}\,.$$

czyli $u_{\mathrm{inhom}} = 1 - x$, co zresztą można było zgadnąć. Zatem

$$u = u_{\mathrm{hom}} + u_{\mathrm{inhom}} = C\,e^{-x} + 1 - x\,,$$

i $y = 2\,\mathrm{arctg}\,(C\,e^{-x} + 1 - x)$.

i) Ponieważ jest to równanie o prawej stronie będącej funkcją jednorodną stopnia zerowego, podstawiamy $y = x\,u(x)$. Daje to równanie $x\,u' = \sqrt{u^2 - 1}$ o zmiennych rozdzielonych (rozwiązania szukamy oczywiście w obszarach $u \geq 1$ i $u \leq -1$). Całkowanie daje

$$\int \frac{du}{\sqrt{u^2 - 1}} = \int \frac{dx}{x} = \ln|C\,x|\,.$$

Rysunek 35: Funkcje $u(x) = (1 + C^2x^2)/2Cx$, $C = \pm\frac{1}{2}, \pm 2, \pm 3$.

Całka po lewej stronie jest równa[71] $\ln|u+\sqrt{u^2-1}|$. Zatem, zdejmując logarytmy, $Cx - u = \sqrt{u^2-1}$. i podnosząc tę równość stronami do kwadratu otrzymujemy

$$u = \frac{1 + 2C^2x^2}{2Cx}\,.$$

Rodzinę tych funkcji przedstawia rysunek 35. Nie całe jednak krzywe są rozwiązaniami równania $x\,u' = \sqrt{u^2-1}$: lewa strona musi być nieujemna, bo taka jest prawa strona. Oznacza to, że gdy $x > 0$ funkcja $u(x)$ musi być rosnąca, a gdy $x < 0$, malejąca. Zatem tylko kawałki krzywych pokazanych na rysunku 35 spełniają ten warunek. Gdy pytamy o krzywą przechodzącą przez punkt (x_0, u_0) dostajemy na C równanie kwadratowe i trzeba wybrać tę wartość C, która w punkcie (x_0, u_0) daje funkcję rosnącą, jeśli $x_0 > 0$ i malejącą, jeśli $x_0 < 0$. Warunek początkowy $y(1) = 2$ odpowiada $u_0 = 2$, $x_0 = 1$. Równanie $2 = (1 + C^2)/2C$ ma dwa rozwiązania: $C = 2 \pm \sqrt{3}$. Odpowiadające tym dwóm stałym krzywe są pokazane na rysunku 36. Widać, że tylko jedna z nich daje funkcję o dodatniej pochodnej w punkcie $x = 1$. Całką ogólną wyjściowego równania jest $y = x\,u(x) = (1 + C^2x^2)/2C$. Równanie to, jak można bezpośrednio sprawdzić ma jednak także rozwiązania (będące uczciwymi funkcjami) $y = \pm x$ nieobejmowane przez całkę ogólną. Krzywą całkową równania jest też prosta $x = 0$.

Zadanie 40:
a) Jest to równanie liniowe z niejednorodnością. Rozwiązujemy najpierw równanie jednorodne $x\,y' = y$, czyli $y' = y/x$ (szukamy jego całki ogólnej zależnej od dowolnej stałej). Jest to równanie o zmiennych rozdzielonych i to takie, jakie już było w Zadaniu 38c. Zatem piszemy "od ręki"

$$y_{\rm hom}(x, C) = C\,x\,.$$

Teraz uzmienniamy stałą, tj. do równananania nieliniowego $y' = y/x + 2x^2$ podstawiamy Ansatz $y_{\rm inhom} = x\,h(x)$. Daje to na h równanie

$$h' = 2x\,,$$

[71]Podstawienie $u = \operatorname{ch}\theta$ sprowadza całkę do $\int d\theta$; z dwóch rozwiązań na $\mathrm{Arch}(u)$ trzeba wybrać to dodatnie, które jest funkcją rosnącą, bo pochodna, czyli funkcja $1/\sqrt{u^2-1}$ pod całką, była dodatnia.

Rysunek 36: Funkcje $u(x) = (1+C^2x^2)/2Cx$, odpowiadające $C = 2 \pm \sqrt{3}$. Obie spełniają warunek $u(1) = 2$, ale tylko krzywa odpowiadająca $C = 2 + \sqrt{3}$ jest w tym punkcie rosnąca.

czyli $h = x^2$ i stąd, całka ogólna wyjściowego równania ma postać

$$y = y_{\text{inhom}}(x) + y_{\text{hom}}(x, C) = x^3 + C\,x\,.$$

Znów widać, że wszystkie krzywe całkowe obejmowane przez całkę ogólną przechodzą przez punkt $(0, 0)$ i żadna nie przechodzi przez punkty $(0, y)$ o $y \neq 0$. Oczywiście jest to możliwe dzięki temu, że funkcję występującą w równaniu $y' = y/x + 2x^2$ po prawej stronie (traktowanej jak funkcja na $\mathbb{R}^2$) można w punkcie $(0, 0)$ uczynić ciągłą w dowolnie wybranym jednym kierunku, a w punktach $(0, y)$ o $y \neq 0$ nie można. Zatem na funkcję $y = y(x)$ spełniającą równanie $y' = y/x + 2x^2$ nie można narzucić warunku początkowego $y(0) = 0$, bo przez $(0, 0)$ przechodzi nieskończenie wiele rozwiązań, ani warunku $y(0) = y_0 \neq 0$, bo żadna uczciwa funkcja $y = y(x)$ przez te punkty nie przechodzi. Jednak równanie zapisane w postaci $x\,dy = (2x^3 + y)dx$ jest spełniane też przez prostą $x(y) \equiv 0$ i na krzywą całkową (nie funkcję) można narzucać warunek $y(0) = y_0 \neq 0$.

b) Jak wyżej, jest to równanie liniowe z niejednorodnością. Odpowiadające mu równanie jednorodne całkujemy rozdzielając zmienne

$$\int \frac{dy}{y} = -\int \frac{dx\,x}{1+x^2}\,, \qquad \text{co daje} \quad \ln|y| = C' - \frac{1}{2}\ln(1+x^2) = \ln\left|\frac{C}{\sqrt{1+x^2}}\right|.$$

Zatem

$$y_{\text{hom}} = \frac{C}{\sqrt{1+x^2}}\,.$$

Uzmienniamy stałą, tj. do równania $y' = -xy/(1+x^2) - 1/2x(1+x^2)$ podstawiamy Ansatz $y_{\text{inhom}} = h(x)/\sqrt{1+x^2}$. Daje to

$$\frac{h'}{\sqrt{1+x^2}} = -\frac{1}{2(1+x^2)}\,, \qquad \text{czyli} \quad h' = -\frac{1}{2x\sqrt{1+x^2}}\,.$$

Całkujemy więc (podstawiając po drodze $x = 1/t$, $dx/x = -dt/t$

$$h = -\frac{1}{2}\int \frac{dx}{x\sqrt{1+x^2}} = \frac{1}{2}\int \frac{dt}{\sqrt{1+t^2}} = \frac{1}{2}\ln\left(t + \sqrt{1+t^2}\right)$$

$$= \frac{1}{2}\ln\Big(1+\sqrt{1+x^2}\Big) - \frac{1}{2}\ln|x|\,.$$

Dostajemy zatem jako całkę ogólną wyjściowego równania funkcję

$$y = y_{\text{inhom}} + y_{\text{hom}} = \frac{1}{2\sqrt{1+x^2}}\Big(\ln\Big(1+\sqrt{1+x^2}\Big) - \ln|x|\Big) + \frac{C}{\sqrt{1+x^2}}\,.$$

Przez każdy punkt (x, y) o $x \neq 0$ przechodzi dokładnie jedna krzywa bo funkcja $f(x, y) = -xy/(1+x^2) - 1/2x(1+x^2)$ jest we wszystkich takich punktach regularna. W punktach (x, y) o $x \neq 0$ nie jest ona ciągła i nie można jej tam w żaden sposób "uciąglić", więc przez takie punkty żadna funkcja $y = y(x)$ nie przechodzi. Ale jak równanie zapiszemy w postaci $2x(1+x^2)dy = (-1-2x^2y)dx$ to prosta $x = 0$ jest też jego rozwiązaniem.

c) To jest niby równanie drugiego rzędu, ale jak położymy $f = y'$ to się z niego zrobi liniowe równanie pierwszego rzędu z niejednorodnością, tyle, że na funkcję f: $x f' + f = 4x$. Dalej już działamy regulaminowo: rozwiązujemy równanie jednorodne $f' = -f/x$, które jest równaniem o zmiennych rozdzielonych, więc

$$\int\frac{dy}{y} = -\int\frac{dx}{x}\,, \qquad \text{czyli} \quad \ln|f| = C' - \ln|x| = -\ln|C\,x|\,,$$

albo $f_{\text{hom}} = C/x$ po prostu. Teraz do równania niejednorodnego $f' + f/x = 4$ podstawiamy Ansatz $f_{\text{inhom}} = h(x)/x$ i dostajemy

$$-\frac{h}{x^2} + \frac{h'}{x} + \frac{1}{x}\frac{h}{x} = 4\,,$$

czyli po prostu $h' = 4x$. Zatem $h = 2x^2$ i, składając wszystko do kupy,

$$y' = f_{\text{hom}} + f_{\text{inhom}} = \frac{C}{x} + 2x\,.$$

Jeszcze raz to całkujemy, i mamy

$$y = A + C\ln|x| + x^2\,.$$

Rozwiązanie spełniające zadane warunki $y(-1) = 0$, $y'(-1) = 0$ otzymujemy dobierając stałe A i C: z drugiego $-C - 2 = 0$, czyli $C = -2$, a z pierwszego $A = -1$.

d) Całkujemy równanie jednorodne $y'\sin x = -y\cos x$ rozdzielając zmienne

$$\int\frac{dy}{y} = -\int dx\,\frac{\cos x}{\sin x}\,,$$

co daje $y_{\text{hom}} = C/\sin x$. Uzmienniamy stałą: $y_{\text{inhom}} = h(x)/\sin x$ i wstawiamy do równania niejednorodnego, co da

$$h' = \sin 2x\,, \qquad \text{czyli} \;\; h = -\frac{1}{2}\cos 2x\,,$$

Stąd $y = C/\sin x - \cos 2x/2\sin x$. Niby jest osobliwość w mianowniku, gdy $x = 0$, ale jak się weźmie $C = 1$, to

$$y = \frac{2 - \cos 2x}{2\sin x} = \sin x\,.$$

Łatwo zobaczyć, że $y = \sin x$ rzeczywiście spełnia równanie $y' \sin x + u \cos x = \sin 2x$.

e) Podstawiamy oczywiście $y = f^2$, żeby się pozbyć pierwiastka. Na f daje to równanie

$$f' = \frac{2f}{x} + \frac{1}{2}\,x\,,$$

które już jest równaniem liniowym z niejednorodnością. Standardowo: $f_{\mathrm{hom}} = C\,x^2$, a podstawiając $f_{\mathrm{inhom}} = x^2 h(x)$ do powyższego równania otrzymujemy $h' = 1/2x$, czyli $h = \frac{1}{2}\ln|x|$. Zatem

$$y = f^2 = \left(C\,x^2 + \frac{1}{2}\,x^2 \ln|x|\right)^2 .$$

Warunek początkowy $y(1) = 1$ jest spełniony, gdy $C = 1$.

f) Podstawiamy oczywiście $u = \ln|y|$, żeby się pozbyć logarytmu szukanej funkcji. Ponieważ $u' = y'/y$, na u dostaje się równanie

$$u' = \frac{u}{x} - \frac{\ln|x|}{x}\,,$$

które już jest równaniem liniowym z niejednorodnością. Dalej standardowo: $u_{\mathrm{hom}} = C\,x$, a podstawiając $u_{\mathrm{inhom}} = x\,h(x)$ do powyższego równania otrzymujemy

$$x\,h' = -\frac{\ln|x|}{x}\,, \qquad \text{czyli} \quad h = -\int \frac{dx}{x}\,\frac{\ln|x|}{x}\,.$$

Naturalne podstawienie $\xi = \ln|x|$, $d\xi = dx/x$ sprowadza tę całkę do

$$h = -\int d\xi\,\xi\,e^{-\xi} = (1+\xi)\,e^{-\xi} = \frac{1 + \ln|x|}{x}\,.$$

Zatem $u = C\,x + 1 + \ln|x|$ (łatwo sprawdzić, że funkcja ta spełnia równanie na u) i stąd

$$y = \ln|C\,x + 1 + \ln|x||\,.$$

Warunek początkowy $y(1) = \ln 2$ jest spełniony, gdy $C = 1$.

Zadanie 41: Napiszemy najpierw równanie stycznej do krzywej $y = f(x)$ w punkcie x_0. Ma ono szkolną postać $y = ax + b$, przy czym a jest nachyleniem, czyli pochodną w x_0 funkcji f. Prosta

$$y = x\,f'(x_0) + b\,,$$

ma przechodzić przez punkt $(x_0, f(x_0))$ więc $b = -x_0 f'(x_0) + f(x_0)$. Z kolei prosta przechodząca przez punkt $(0, 0)$ i prostopadła do tej stycznej (bo odległość prostej od punktu, to odległość liczona po prostopadłej właśnie) ma równanie

$$y = -\frac{1}{f'(x_0)}\, x\,.$$

Gdzie się te dwie proste przecinaja? Rozwiązujemy układzik $y = ax + b$, $y = -x/a$ i znajdujemy punkt przecięcia $(\tilde{x}, \tilde{y})$

$$\tilde{x} = -\frac{ab}{1+a^2}\,, \qquad \tilde{y} = \frac{b}{1+a^2}\,,$$

przy czym oczywiście $a = f'(x_0)$, $b = f(x_0) - x_0 f'(x_0)$. Odległość d tego punktu przecięcia od punktu $(0, 0)$ (czyli właśnie odległość stycznej od początku układu) jest oczywiście równa

$$d = \sqrt{\tilde{x}^2 + \tilde{y}^2} = \frac{b}{\sqrt{1+a^2}}\,.$$

Odległość ta, zgodnie z żądaniem, ma być równa x_0. Ponieważ ma to być słuszne dla dowolnego punktu krzywej, przeto zmieniamy oznaczenia z x_0 na x, $f(x_0)$ na $y(x)$ i $f'(x_0)$ na $y'(x)$ i piszemy ten warunek, który staje się tym samym równaniem różniczkowym na $y = y(x)$:

$$x^2 = \frac{(y - x\,y')^2}{1 + (y')^2}\,,$$

czyli

$$x^2 = y^2 - 2x\,y\,y'.$$

Jak to rozwiązać (ani to liniowe, ani o zmiennych rozdzielonych)? Ano, pewnie podstawić: samo się narzuca, bo $2yy' = d(y^2)/dx$. Więc $u = y^2$ i mamy

$$x\,u' = u - x^2\,.$$

teraz to jest równanie liniowe z niejednorodnością. Jednorodne to $u' = u/x$ - już dwa razy takie było, więc $u_{\mathrm{hom}} = C\,x$ i podstawiamy Ansatz $u_{\mathrm{inhom}} = x\,h(x)$. Ostatecznie $h' = -1$, czyli $h = -x$ i ($C' = C/2$, żeby było wygodniej) $u = -x^2 + 2C'x$. teraz przypominamy sobie, że $u = y^2$ i mamy z tego okrąg $(x - C')^2 + y^2 = C'^2$ o środku w $(C, 0)$ i promieniu $|C|$. Ale to jeszcze nie wszystko: jak napiszemy otrzymane równanie na u w formie $x\,du = (u - x^2)dx$ (albo $2xy\,dy = (y^2 - x^2)dx$), to widzimy, że $x \equiv 0$ też jest rozwiązaniem. No i wszystko się zgadza: prosta $x = 0$ jest sama do siebie styczna (w każdym swoim punkcie); styczna do niej, czyli ona sama, jest odległa od $(0, 0)$ o zero i odcięta każdego punktu tej prostej też jest równa zero. Więc warunek jest spełniony!

Chociaż samo równanie różniczkowe wyprowadziliśmy przyjmując po cichu, że szukana krzywa jest funkcją $y = y(x)$, to jakoś "wie" ono także o krzywych, które funkcjami nie są. Zapewne przez jakoś rozumianą "ciągłość matematyki".

Zadanie 42: Jeśli w równaniu $y' = y\,a(x) + y^n\,b(x)$ wykładnik $n = 0$, jest to równanie liniowe z niejednorodnością $y' = y\,a(x) + b(x)$, które rozwiązuje się metodą uzmiennienia stałej, by znaleźć $y_{\rm inhom}$, a $y_{\rm hom}$ znajduje się rozdzialając zmienne. Jeśli $n = 1$, rówananie $y' = y\,[a(x) + b(x)]$ jest po prostu równaniem a zmiennych rozdzielonych i całkuje siye je standardowo. Gdy $n \neq 0, 1$, jest to równanie nieliniowe. Jesna podstawienie $z = y^{1-n}$ sprowadza je do równania liniowego z niejednorodnością. Bo istotnie:

$$y' = \frac{z'}{1-n}\,z^{\frac{1}{1-n}-1} = \frac{z'}{1-n}\,z^{\frac{n}{1-n}} = \frac{z'}{1-n}\,y^n\,, \qquad y = z^{\frac{1}{1-n}} = z^{\frac{1-n+n}{1-n}} = z\,y^n\,.$$

Po wstawieniu tych wyrażeń do równania przybiera ono postać

$$z' = (1-n)\,z\,a(x) + (1-n)\,b(x)\,,$$

czyli równania liniowego z niejednorodnością na funkcję $z = z(x)$.

Podane równanie odpowiada $n = -2$ i trzeba dokonać podstawienia $u = y^3$, co i tak można łatwo zgadnąć bez odwoływania się do równań typu Bernoulliego. Podstawienie to prowadzi do równania

$$x\,u' = 2u + x^3\,,$$

które jest równaniem liniowym z niejednorodnością. Całką ogólną równania jednorodnego jest $u_{\rm hom} = C\,x^2$, a podstawienie do równania niejednorodnego Ansatzu $u_{\rm inhom} = x^2 h(x)$ daje na funkcję h równanie $h' = 1$. Stąd,

$$u = C\,x^2 + x^3\,, \qquad \text{i} \quad y = \left(C\,x^2 + x^3\right)^{1/3}\,.$$

Zadanie 43:

a) Jak zawsze rozwiązaniem jest $\mathbf{y}(t) = \exp(tF)\cdot\mathbf{y}(0)$. Cała wiec trudność polega na znalezieniu eksponensu macierzy tF. Równanie charakterystyczne macierzy F

$$W_F(\lambda) = -\lambda^3 - 3\lambda^2 - 3\lambda - 1 = -(\lambda+1)^3 = 0\,,$$

ma jeden pierwiastek potrójny. Szukamy wektorów własnych macierzy F:

$$\begin{pmatrix} 3 & -1 & 2 \\ 5 & -2 & 3 \\ -1 & 0 & -1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Jest tylko jeden

$$\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}.$$

Macierz F nie jest diagonalizowalna. Szukamy zatem jej wektorów pierwiastkowych, na których zeruje się macierz $(F+I)^3$. Ponieważ cała przestrzeń jest trójwymiarowa macierz ta jest po prostu macierzą zerową, a podprzestrzenią pierwiastkową odpowiadającą (jedynej) wartości własnej $\lambda=-1$ jest po prostu cała przestrzeń wektorowa. Zatem za wektory pierwiastkowe można wybrać dowolne dwa wektory, byle z wektorem własnym tworzyły bazę całej przestrzeni. Szczególnie prostym i dogodnym do dalszych rachunków wyborem są wektory

$$\begin{pmatrix}1\\0\\0\end{pmatrix}, \qquad \begin{pmatrix}0\\1\\0\end{pmatrix}.$$

Rozkładamy teraz wektor warunków początkowych na wektor własny i te dwa

$$\begin{pmatrix}1\\0\\1\end{pmatrix}=-\begin{pmatrix}1\\1\\-1\end{pmatrix}+2\begin{pmatrix}1\\0\\0\end{pmatrix}+\begin{pmatrix}0\\1\\0\end{pmatrix},$$

i działamy na obie strony macierzą e^{tF}:

$$e^{tF}\cdot\begin{pmatrix}1\\0\\1\end{pmatrix}=\begin{pmatrix}-1\\-1\\1\end{pmatrix}e^{-t}+e^{-t}\left[I+t(F+I)+\frac{t^2}{2}(F+I)^2\right]\left\{\begin{pmatrix}2\\0\\0\end{pmatrix}+\begin{pmatrix}0\\1\\0\end{pmatrix}\right\},$$

Napisaliśmy tu jak zwykle $e^{tF}=e^{-tI}\,e^{t(F+I)}$ i rozwinęliśmy drugi eksponens w szereg wykorzystując to, że macierz $(F+I)^3$ jest już macierzą zerową. Zatem

$$e^{tF}\cdot\begin{pmatrix}1\\0\\1\end{pmatrix}=\begin{pmatrix}-1\\-1\\1\end{pmatrix}e^{-t}+e^{-t}\left[I+t\begin{pmatrix}3&-1&2\\5&-2&3\\-1&0&-1\end{pmatrix}+\frac{t^2}{2}\begin{pmatrix}2&-1&1\\2&-1&1\\-2&1&-1\end{pmatrix}\right]\begin{pmatrix}2\\1\\0\end{pmatrix}.$$

Składając wszystko razem mamy

$$\begin{pmatrix}y_1(t)\\y_2(t)\\y_3(t)\end{pmatrix}=e^{-t}\begin{pmatrix}1+5t+3t^2/2\\8t+3t^2/2\\1-2t-3t^2/2\end{pmatrix}.$$

Całą macierz e^{tF} można znaleźć albo rozkładając na te same wektory co wyżej ogólny wektor (a,b,c), działają tak jak wyżej macierzę e^{tF} i potem zapisując wynik w postaci pewnej macierzy (którą będzie właśnie szukana macierz e^{tF}) działającej na wektor (a,b,c), albo metodą CH. W tej drugiej metodzie wykorzystujemy to, że $W_F(-1)=0$, $W'(-1)=0$ i $W''(-1)=0$. W rezultacie na współczynniki a_2, a_1 i a_0 wielomianu-reszty $a_2\lambda^2+a_1\lambda+a_0$, które wchodzą w rozkład $e^{tF}=a_2F^2+a_1F+a_0I$ możemy napisać układ równań

$$e^{-t}=a_2-a_1+a_0\,,\quad t\,e^{-t}=-2a_2+a_1\,,\quad t^2\,e^{-t}=2a_2\,.$$

Zatem

$$e^t\, e^{tF} = \frac{t^2}{2}\begin{pmatrix} -3 & 1 & -3 \\ -8 & 4 & -5 \\ 0 & 1 & 2 \end{pmatrix} + (t+t^2)\begin{pmatrix} 2 & -1 & 2 \\ 5 & -3 & 3 \\ -1 & 0 & -2 \end{pmatrix} + \left(1+t+\frac{t^2}{2}\right)\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Ostatecznie więc

$$e^{tF} = e^{-t}\begin{pmatrix} 1+3t+t^2 & -t-t^2/2 & 2t+t^2/2 \\ 5t+t^2 & 1-2t-t^2/2 & 3t+t^2/2 \\ -t-t^2 & t^2/2 & 1-t-t^2/2 \end{pmatrix}.$$

b) Tu też rozwiązaniem jest $\mathbf{y}(t) = \exp(tF)\cdot\mathbf{y}_0$, gdzie $\mathbf{y}_0$ jest wektorem pełniącym rolę stałych dowolnych. Wielomian charakterystyczny macierzy F

$$W_F(\lambda) = \begin{vmatrix} -2-\lambda & 1 & 0 \\ 1 & -2-\lambda & 1 \\ 0 & 1 & -2-\lambda \end{vmatrix} = -\lambda^3 - 6\lambda^2 - 10\lambda - 4\,,$$

ma jako pierwiastek (trzeba trochę popróbować) $\lambda_1 = -2$. Zatem

$$W_F(\lambda) = -(\lambda+2)(\lambda^2 + a\lambda + 2) = -(\lambda+2)(\lambda^2+4\lambda+2),$$

i pozostałymi dwoma pierwiastkami są $\lambda_+ = -2+\sqrt{2}$ oraz $\lambda_- = -2-\sqrt{2}$. Wektory własne odpowiadające tym wartościom własnym spełaniające równania

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} a_1 \\ b_1 \\ c_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \qquad \begin{pmatrix} \mp\sqrt{2} & 1 & 0 \\ 1 & \mp\sqrt{2} & 1 \\ 0 & 1 & \mp\sqrt{2} \end{pmatrix}\begin{pmatrix} a_\pm \\ b_\pm \\ c_\pm \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

też daje się znaleźć, co pozwala rozłożyć na nie dowolny wektor $\mathbf{y}_0$ warunku początkowego,

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = x_1\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + x_+\begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix} + x_-\begin{pmatrix} 1 \\ -\sqrt{2} \\ 1 \end{pmatrix}.$$

Układ ten rozwiązuje się wyjątkowo łatwo i otrzymuje się

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \frac{a-c}{2}\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + \frac{a+\sqrt{2}+c}{4}\begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix} + \frac{a-\sqrt{2}+c}{4}\begin{pmatrix} 1 \\ -\sqrt{2} \\ 1 \end{pmatrix}.$$

Stąd

$$\begin{aligned} \begin{pmatrix} y_1(t) \\ y_2(t) \\ y_3(t) \end{pmatrix} &= \frac{a-c}{2}\, e^{-2t}\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \\ &+ \frac{a+\sqrt{2}+c}{4}\, e^{-(2-\sqrt{2})t}\begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix} + \frac{a-\sqrt{2}+c}{4}\, e^{-(2+\sqrt{2})t}\begin{pmatrix} 1 \\ -\sqrt{2} \\ 1 \end{pmatrix}. \end{aligned}$$

Można też znanym sposobem otczytać z prawej strony powyższego wzoru całą macierz $\exp(tF)$:

$$e^{tF} = \frac{e^{-2t}}{2}\begin{pmatrix} 1+\mathrm{ch}(\sqrt{2}t) & \sqrt{2}\,\mathrm{sh}(\sqrt{2}t) & -1+\mathrm{ch}(\sqrt{2}t) \\ \sqrt{2}\,\mathrm{sh}(\sqrt{2}t) & \mathrm{ch}(\sqrt{2}t) & \sqrt{2}\,\mathrm{sh}(\sqrt{2}t) \\ -1+\mathrm{ch}(\sqrt{2}t) & \sqrt{2}\,\mathrm{sh}(\sqrt{2}t) & 1+\mathrm{ch}(\sqrt{2}t) \end{pmatrix}.$$

c) Najpierw rozwiążemy te równania traktując je traktując je jak układ trzech liniowych równań jednorodnych. Wielomian charakterystyczny macierzy F

$$W_F(\lambda) = \begin{vmatrix} 1-\lambda & 0 & 0 \\ 1 & -1-\lambda & -1 \\ -2 & 4 & 3-\lambda \end{vmatrix} = -(\lambda-1)(\lambda^2-2\lambda+1) = -(\lambda-1)^3\,,$$

ma tylko jeden, zato potrójny, pierwiastek $\lambda = 1$. Wektory własne spełniające równanie

$$\begin{pmatrix} 0 & 0 & 0 \\ 1 & -2 & -1 \\ -2 & 4 & 2 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

są dwa, bo jest tylko jedno równanie $a - 2b - c = 0$ do spełnienia. Ponieważ macierz $(F-I)^2$ jest po prostu macierzą zerową (musi być taką zgodnie z twierdzeniem CH, ale można to bezpośrednio sprawdzić), jako wektor pierwiastkowy odpowiadający wartości własnej $\lambda = 1$ można wziąć jakikolwiek wektor liniowo niezależny od wektorów własnych. Jako bazę całej przestrzeni można więc wziąć dwa wektory własne macierzy F i jakiś liniowo od nich niezależny trzeci wektor i rozłożyć na nie dowolny wektor warunków początkowych:

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \alpha\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + \beta\begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} + \gamma\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Układ równań rozwiązuje się błyskiem: $\alpha = a - 2b$, $\beta = b$, $\gamma = -a + 2b + c$. Działamy teraz macierzą $\exp(tF)$ na obie strony tej równości

$$e^{tF}\cdot\begin{pmatrix} a \\ b \\ c \end{pmatrix} = e^t\left\{(a-2b)\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + b\begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} + (-a+2b+c)\,[I + t\,(F-I)]\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right\}.$$

Działając na wektor pierwiastkowy zapisaliśmy e^{tF} jako $e^{tI}e^{t(F-I)} = e^t[1 + t(F-I) + \ldots]$ i skorzystaliśmy z tego, że wykropkowane wyrazy rozwinięcie są już macierzami zerowymi. Zbierając wszystko razem dostajemy

$$e^{tF}\cdot\begin{pmatrix} a \\ b \\ c \end{pmatrix} = e^t\begin{pmatrix} a \\ b - t(-a+2b+c) \\ (1+2t)(-a+2b+c) \end{pmatrix} = e^t\begin{pmatrix} 1 & 0 & 0 \\ t & 1-2t & -t \\ -2t & 4t & 1+2t \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix}.$$

Macierz stojąca po prawej stronie to e^{tF}. Dla wprawy znajdziemy ją jeszcze metodą CH. Ponieważ istnieje jeden dodatkowy wektor własny macierzy F, należy skorzystać ze zredukowanego wielomianu charakterystycznego, co obniża stopień wielomianu-reszty; poza tym, ponieważ jest tylko jedna wartość własna o krotności trzy, trzeba skorzystać z tricku z pochodną. Zatem równaniami wyznaczającymi współczynniki wielomianu-reszty są

$$e^t = a_1 + a_0 \,, \qquad t\, e^t = a_1 \,.$$

Stąd $e^{tF} = a_1 F + a_0 I = e^t \left[t\, F + (1-t) I \right]$, czyli

$$e^{tF} = e^t \left\{ t \begin{pmatrix} 1 & 0 & 0 \\ 1 & -1 & -1 \\ -2 & 4 & 3 \end{pmatrix} + (1-t) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right\}.$$

Po złożeniu tego w jedną macierz dostaje się to samo co wyżej.

Ponieważ równanie na y_1, które po wypreparowaniu go z zapisu macierzowego ma postać $dy_1/dt = y_1$, jest niezależne od pozostałych, jego jawne rozwiązanie $y_1 = e^t\, a$ można wstawić do równań na y_2 i y_3 i przepisać je w postaci układu równań liniowych pierwszego rzędu z niejednorodnością:

$$\frac{d}{dt} \begin{pmatrix} y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} -1 & -1 \\ 4 & 3 \end{pmatrix} \begin{pmatrix} y_2 \\ y_3 \end{pmatrix} + \begin{pmatrix} e^t\, a \\ -2\, e^t\, a \end{pmatrix}.$$

Wielomianem charakterystycznym stojącej tu macierzy 2×2 (oznaczmy ją $\tilde{F}$) jest $W_{\tilde{F}}(\lambda) = (\lambda - 1)^2$. Ma on podwójny pierwiastek $\lambda = 1$. Macierz $\tilde{F}$ ma tylko jeden wektor własny, więc przy znajdywaniu $\exp(t\tilde{F})$ musimy znów skorzystać z tricku z pochodną; równania na a_1 i a_0 są zresztą te same, co i wyżej. Zatem

$$e^{t\tilde{F}} = e^t \left\{ t \begin{pmatrix} -1 & -1 \\ 4 & 3 \end{pmatrix} + (1-t) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right\} = e^t \begin{pmatrix} 1-2t & -t \\ 4t & 1+2t \end{pmatrix}.$$

(Zabawnie jest sprawdzić, że zamiana $t \to -t$ daje, tak jak być powinno, macierz odwrotną). Podstawiając do równania niejednorodnego Ansatz $\tilde{\mathbf{y}}_{\rm inhom} = \exp(t\tilde{F}) \cdot \mathbf{h}(t)$ otrzymujemy na wektor $\mathbf{h}(t)$ równanie

$$\frac{d}{dt} \begin{pmatrix} h_2 \\ h_3 \end{pmatrix} = e^{-t} \begin{pmatrix} 1+2t & t \\ -4t & 1-2t \end{pmatrix} \begin{pmatrix} e^t\, a \\ -2\, e^t\, a \end{pmatrix} = \begin{pmatrix} a \\ -2a \end{pmatrix}.$$

Możemy zatem złożyć kompletne rozwiązanie na funkcje y_2 i y_3 (C_2 i C_3 są dwiema dowolnymi stałymi w rozwiązaniu równania jadnorodnego):

$$\begin{pmatrix} y_2 \\ y_3 \end{pmatrix} = e^t \begin{pmatrix} 1-2t & -t \\ 4t & 1+2t \end{pmatrix} \begin{pmatrix} C_2 + at \\ C_3 - 2at \end{pmatrix}.$$

Aby tak jak poprzednio $y_2(0) = b$, $y_3(0) = c$, należy położyć $C_2 = b$, $C_3 = c$. Po zadziałaniu macierzą $\exp(t\tilde{F})$ na stojący po prawej stronie wektor otrzyma się (wyrazy z t^2 zredukuja się) na y_2 i y_2 te same rozwiązania, co poprzednią metodą:

$$y_2(t) = e^t \left[(a-c)\, t + (1-2t)\, b \right], \qquad y_3(t) = e^t \left[-2a\, t + 4b\, t + (1+2t)\, c \right].$$

Zadanie 44:

a) Najpierw rozwiązujemy układ macierzowy. (Jest to macierzowe równanie liniowe z wektorową niejednorodnościa). Nietrudno znaleźć (a kto ma w głowie transformację Lorentza, ten to od razu widzi), że

$$e^{xF} = I\,\mathrm{ch}\,x + F\,\mathrm{sh}\,x = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \mathrm{ch}\,x + \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \mathrm{sh}\,x\,.$$

Zatem ogólnym rozwiązaniem równania jednorodnego jest

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}_{\mathrm{hom}} = \begin{pmatrix} C_1 \\ C_2 \end{pmatrix} \mathrm{ch}\,x + \begin{pmatrix} C_2 \\ C_1 \end{pmatrix} \mathrm{sh}\,x\,.$$

Podstawiając do równania niejednorodnego jako Ansatz $\mathbf{y}_{\mathrm{inhom}} = \exp(xF) \cdot \mathbf{h}(x)$, dostajemy na $\mathbf{h}$ równanie $\mathbf{h}' = \exp(-xF) \cdot \mathbf{b}$, gdzie $\mathbf{b}$ jest wektorem-niejednorodnością, czyli

$$\begin{pmatrix} h_1' \\ h_2' \end{pmatrix} = \begin{pmatrix} \sin x \\ 2\cos x \end{pmatrix} \mathrm{ch}\,x - \begin{pmatrix} 2\cos x \\ \sin x \end{pmatrix} \mathrm{sh}\,x\,.$$

Całkowanie jest elementarne (najlepiej wszystkie funkcje pod całkami przerobić na eksponensy) i daje

$$\begin{pmatrix} h_1 \\ h_2 \end{pmatrix} = \begin{pmatrix} -\frac{3}{2}\,\mathrm{ch}\,x\cos x - \frac{1}{2}\,\mathrm{sh}\,x \sin x \\ \frac{1}{2}\,\mathrm{ch}\,x \sin x + \frac{3}{2}\,\mathrm{sh}\,x\cos x \end{pmatrix}.$$

Po zadziałaniu na ten wektor macierzą $\exp(xF)$ dostajemy

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}_{\mathrm{inhom}} = \begin{pmatrix} -\frac{3}{2}\,\mathrm{ch}\,x\cos x - \frac{1}{2}\,\mathrm{sh}\,x \sin x \\ \frac{1}{2}\,\mathrm{ch}\,x \sin x + \frac{3}{2}\,\mathrm{sh}\,x\cos x \end{pmatrix} \mathrm{ch}\,x + \begin{pmatrix} \frac{1}{2}\,\mathrm{ch}\,x \sin x + \frac{3}{2}\,\mathrm{sh}\,x\cos x \\ -\frac{3}{2}\,\mathrm{ch}\,x\cos x - \frac{1}{2}\,\mathrm{sh}\,x \sin x \end{pmatrix} \mathrm{sh}\,x\,.$$

Po skorzystaniu z tożsamości $\mathrm{ch}^2 x - \mathrm{sh}^2 x = 1$ upraszcza się to znacznie i ostatecznie, jako całkę ogólną pełnego równania niejednorodnego otrzymujemy

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}_{\mathrm{hom}} + \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}_{\mathrm{inhom}} = \begin{pmatrix} C_1 \\ C_2 \end{pmatrix} \mathrm{ch}\,x + \begin{pmatrix} C_2 \\ C_1 \end{pmatrix} \mathrm{sh}\,x + \begin{pmatrix} -\frac{3}{2}\cos x \\ \frac{1}{2}\sin x \end{pmatrix}.$$

Można (i należy!) sprawdzić przez bezpośrednie podstawienie części $\mathbf{y}_{\mathrm{inhom}}$ do równania, że jest ono przez tę część rzeczywiście spełniane. Aby spełnić warunki początkowe $y_1(0) = 2$ i $y_2(0) = 0$ należy położyć $C_1 = 7/2$ i $C_2 = 0$.

Ten sam układ dwóch równań

$$\begin{aligned} y_1' &= y_2 + \sin x\,, \\ y_2' &= y_1 + 2\cos x\,, \end{aligned}$$

można (prościej chyba) rozwiązać, dodając i odejmując te dwa równania jedno od drugiego. Otrzymuje się wtedy dwa nizależne równania

$$\begin{aligned} z_1' &= z_1 + \sin x + 2\cos x\,, \\ z_2' &= -z_2 + 2\cos x - \sin x\,, \end{aligned}$$

z których każde z osobna jest liniowym równaniem z niejednorodnością Rozwiązania równań jednorodnych są oczywiste, $z_1^{\mathrm{hom}} = D_1\, e^x$, $z_2^{\mathrm{hom}} = D_2\, e^{-x}$, a standardowe podstawienia $z_1^{\mathrm{inhom}} = e^x\, f_1(x)$, $z_2^{\mathrm{inhom}} = e^{-x}\, f_2(x)$ dają

$$
\begin{aligned}
f_1 &= \int dx\, e^{-x}\, (\sin x + 2\cos x)\,, \\
f_2 &= \int dx\, e^{x}\, (2\cos x - \sin x)\,.
\end{aligned}
$$

Ostatecznie więc w ten sposób dostajemy

$$
\begin{aligned}
z_1 &= D_1\, e^x - \frac{3}{2}\cos x + \frac{1}{2}\sin x\,, \\
z_2 &= D_2\, e^{-x} + \frac{3}{2}\cos x + \frac{1}{2}\sin x\,,
\end{aligned}
$$

i stad

$$
\begin{aligned}
y_1 &= \frac{1}{2}\,(z_1 - z_2) = D_1\, e^x - D_2\, e^{-x} - \frac{3}{2}\cos x\,, \\
y_2 &= \frac{1}{2}\,(z_1 + z_2) = D_1\, e^x + D_2\, e^{-x} + \frac{1}{2}\sin x\,.
\end{aligned}
$$

jest to to samo co poprzednim sposobem jeśli utożsamić D_1 z $(C_1 + C_2)/2$, a D_2 z $(C_2 - C_1)/2$.

b) Wielomian charakterystycznu macierzy F

$$
W_F(\lambda) = \begin{vmatrix} 3-\lambda & -1 & 0 \\ 0 & 3-\lambda & -1 \\ -2 & 5 & -1-\lambda \end{vmatrix} = -\lambda^3 + 5\lambda^2 - 8\lambda + 4 = -(\lambda - 1)(\lambda^2 - 4\lambda + 4)\,,
$$

ma jeden pierwiastek pojedyńczy $\lambda_1 = 1$ i jeden pierwiastek podwójny $\lambda_2 = 2$. Wektory własne spełniające równania

$$
\begin{pmatrix} 2 & -1 & 0 \\ 0 & 2 & -1 \\ -2 & 5 & -2 \end{pmatrix} \begin{pmatrix} a_1 \\ b_1 \\ c_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \qquad \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ -2 & 5 & -3 \end{pmatrix} \begin{pmatrix} a_2 \\ b_2 \\ c_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},
$$

są tylko dwa. Macierz F nie jest więc diagonalizowalna i potrzebną macierz $\exp(tF)$ znajdziemy metodą CH wykorzystując trick z różniczkowaniem. Współczynniki w a_2, a_2 i a_0 w równości $\exp(tF) = a_2 F^2 + a_1 F + a_0 I$ są wyznaczone przez równości

$$
\begin{aligned}
e^t &= a_2 + a_1 + a_0\,, \\
e^{2t} &= 4a_2 + 2a_1 + a_0\,, \\
t\, e^{2t} &= 4a_2 + a_1\,,
\end{aligned}
$$

Znajdujemy: $a_0 = 4e^t - (3-2t)e^{2t}$, $a_1 = -4e^t + (4-3t)e^{2t}$, $a_2 = e^t - (1-t)e^{2t}$. Stąd

$$
\begin{aligned}
e^{tF} &= e^t(1-(1-t)e^t)\begin{pmatrix} 9 & -6 & 1 \\ 2 & 4 & -2 \\ -4 & 12 & 4 \end{pmatrix} \\
&\quad + e^t(-4+(4-3t)e^t)\begin{pmatrix} 3 & -1 & 0 \\ 0 & 3 & -1 \\ -2 & 5 & -1 \end{pmatrix} + e^t(4-(3-2t)e^t)\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
&= e^t \begin{pmatrix} 1+2t\,e^t & -2+(2-3t)e^t & 1-(1-t)e^t \\ 2-2(1-t)e^t & -4+(5-3t)e^t & 2-(2-t)e^t \\ 4-(4-2t)e^t & -8+(8-3t)e^t & 4-(3-t)e^t \end{pmatrix}.
\end{aligned}
$$

Rozwiązanie równania jednorodnego jest dane działaniem tej macierzy na wektor dowolnych stałych (C_1, C_2, C_3). Rozwiązaniem równania niejednorodnego jest ta sama macierz działająca na wektor $\mathbf{h}(t)$, dany przez całkę

$$
\begin{pmatrix} h_1(t) \\ h_2(t) \\ h_3(t) \end{pmatrix} = \int dt\, e^{-tF} \cdot \begin{pmatrix} e^t \\ 0 \\ 0 \end{pmatrix} = \int dt \begin{pmatrix} 1+2t\,e^t \\ 2-2(1-t)e^t \\ 4-(4-2t)e^t \end{pmatrix} = \begin{pmatrix} t-2(1-t)\,e^t \\ 2t-2(2-t)e^t \\ 4t-2(3-t)e^t \end{pmatrix}.
$$

Całka ogólna wyjściowego równania niejednorodnego ma więc postać

$$
\begin{pmatrix} y_1(t) \\ y_2(t) \\ y_3(t) \end{pmatrix} = e^t \begin{pmatrix} 1+2t\,e^t & -2+(2-3t)e^t & 1-(1-t)e^t \\ 2-2(1-t)e^t & -4+(5-3t)e^t & 2-(2-t)e^t \\ 4-(4-2t)e^t & -8+(8-3t)e^t & 4-(3-t)e^t \end{pmatrix} \begin{pmatrix} C_1+t-2(1-t)\,e^t \\ C_2+2t-2(2-t)e^t \\ C_3+4t-2(3-t)e^t \end{pmatrix}.
$$

Ponieważ macierz $\exp(tF)$ jest w $t=0$ macierzą jednostkową, aby spełnić warunek początkowy $y_1(0)=1$, $y_2(0)=1$, $y_3(0)=3$, należy przyjąć $C_1=3$, $C_2=5$, $C_3=9$. Jawne zadziałanie macierzą $\exp(tF)$ na wektor zostawiamy już wytrwałym.

Zadanie 45: Jest to równanie liniowe z niejednorodnością, więc jego całka ogólna ma postać $y = y_{\rm hom}(x, C_1, C_2) + y_{\rm inhom}$. Rozwiązanie $y_{\rm hom}(x, C_1, C_2)$ znajdujemy podstawiając $y = \exp(\lambda x)$ do równania jednorodnego $5\,y'' - 6\,y' + 5\,y = 0$, co da równanie charakterystyczne $5\lambda^2 - 6\lambda + 5 = 0$, o dwóch sprzężonych zespolonych pierwiastkach

$$
\lambda_1 = \frac{3+4i}{5}, \qquad \lambda_2 = \frac{3-4i}{5}.
$$

Zatem

$$
y_{\rm hom}(x, C_1, C_2) = e^{\frac{3}{3}x}\left(C_1 \cos\frac{4}{5}x + C_2 \sin\frac{4}{5}x\right).
$$

Rozwiązania równania niejednorodnego możemy szukać podstawiając Ansatz $y_{\rm inhom} = A\sin\frac{4}{5}x + B\cos\frac{4}{5}x$. Daje to na A i B układ równań

$$
9A + 24B = 5, \qquad -24A + 9B = 0.
$$

Rozwiązaniem jest $A = 5/73$, $B = 40/219$. Zatem całka ogólna równania ma postać

$$y = e^{\frac{3}{3}x}\left(C_1 \cos\frac{4}{5}x + C_2 \sin\frac{4}{5}x\right) + \frac{5}{73}\sin\frac{4}{5}x + \frac{40/}{219}\cos\frac{4}{5}x\,.$$

Zadanie 46: Pierwiastkami równania charakterystycznego są $\lambda_1 = 1$ i $\lambda_2 = 2$. Zatem

$$y_{\rm hom}(x, C_1, C_2) = C_1\, e^{x} + C_2\, e^{2x}\,.$$

Aby znaleźć rozwiązanie równania niejednorodnego można posłużyć się ogólnym wzorem wyprowadzonym w tekście

$$y_{\rm inhom}(x) = \frac{1}{\lambda_2 - \lambda_1}\left\{e^{\lambda_2 x}\int^{x} dx'\, e^{-\lambda_2 x'}\, f(x') - e^{\lambda_1 x}\int^{x} dx'\, e^{-\lambda_1 x'}\, f(x')\right\},$$

który tu daje

$$y_{\rm inhom}(x) = e^{2x}\int^{x} dx'\, e^{-2x'}\sin\!\left(e^{-x'}\right) - e^{x}\int^{x} dx'\, e^{-x'}\sin\!\left(e^{-x'}\right).$$

Naturalne podstawienie $\xi = e^{-x'}$ sprowadza ten wzór do

$$y_{\rm inhom}(x) = -e^{2x}\int^{e^{-x}} d\xi\, \xi\sin\xi + e^{x}\int^{e^{-x}} d\xi\sin\xi\,.$$

Całki są elementarne i dostajemy

$$y_{\rm inhom} = -e^{x}\cos\!\left(e^{-x}\right) + e^{2x}\left\{-\sin\!\left(e^{-x}\right) + e^{-x}\cos\!\left(e^{-x}\right)\right\}.$$

Całkę ogólna ma więc postać

$$y = C_1\, e^{x} + C_2\, e^{2x} - e^{2x}\sin\!\left(e^{-x}\right).$$

Zadanie 47: Podstawienie $\exp(\lambda x)$ daje równanie charakterystycze

$$\lambda^7 - 3\,\lambda^6 + 5\,\lambda^6 - 7\,\lambda^4 + 7\,\lambda^3 - 5\,\lambda^2 + 3\,\lambda + 1 = 0\,.$$

Patrząc na nie przytomnie, od razu można dostrzec, że pierwiastkiem jest $\lambda = 1$. Po napisaniu

$$(\lambda - 1)(\lambda^6 - 2\,\lambda^5 + 3\,\lambda^4 - 4\,\lambda^3 + 3\,\lambda^2 - 2\,\lambda - 1) = 0\,,$$

znów można się zorientować, że $\lambda = 1$ jest pierwiastkiem drugiego nawiasu. Zatem

$$(\lambda - 1)^2(\lambda^5 - \lambda^4 + 2\,\lambda^3 - 2\,\lambda^2 + \lambda - 1) = 0\,,$$

I po raz trzeci $\lambda = 1$ jest pierwiastkiem drugiego nawiasu:

$$(\lambda - 1)^3(\lambda^4 + 2\,\lambda^2 + 1) = (\lambda - 1)^3(\lambda^2 + 1)^2 = 0\,.$$

Ostatecznie więc pierwiastkami są $\lambda_1 = 1$, potrójny oraz $\lambda_2 = i$ i $\lambda_3 = -i$ oba podwójne. Rozwiązaniem równanania różniczkowego, jego całką ogólną, jest więc

$$\begin{aligned} y &= A\,e^x + B\,x\,e^x + C\,x^2\,e^x + \mathbb{D}\,e^{ix} + \mathbb{E}\,x\,e^{ix} + \mathbb{D}^*\,e^{-ix} + \mathbb{E}^*\,x\,e^{-ix} \\ &= (A + B\,x + C\,x^2)\,e^x + D_1\cos x + D_2\sin x + x\,(E_1\cos x + E_2\sin x). \end{aligned}$$

Zadanie 48: Podstawienie $\exp(\lambda x)$ do równania jednorodnego daje równanie charakterystycze

$$\lambda^3 + \lambda^2 + \lambda + 1 = (\lambda + 1)(\lambda^2 + 1) = 0\,,$$

którego pierwiastkami są -1 i $\pm i$. Rozwiązanie równania jednorodnego ma więc ogólną postać

$$y_{\rm hom} = A\,e^{-x} + \mathbb{B}\,e^{ix} + \mathbb{B}^*\,e^{-ix} = A\,e^{-x} + D\cos x + E\sin x\,.$$

Aby znaleźć jakies rozwiązanie równania niejednorodnego będziemy szukać rozwiązania $y^{(1)}_{\rm hom}$ z niejednorodnością $x\,e^{-x}$ i rozwiązania $y^{(2)}_{\rm hom}$ z niejednorodnością $\cos x$. Szukane rozwiązanie będzie wtedy sumą $y^{(1)}_{\rm hom} + y^{(2)}_{\rm hom}$. Spróbujmy najpierw jako $y^{(1)}_{\rm hom}$ podstawić $C\,x\,e^{-x}$. Niestety po lewej stronie wszystkie wyrazy z $x\,e^{-x}$ się zredukują i zostanie tylko stała razy e^{-x} - nie da się więc w ten sposób spełnić równania niejednorodnego. Zgodnie ze wskazówką rozpatrzmy więc równanie $y' + y = x\,e^{-x}$. Tu można zastosować metodę uzmiennienia stałej (w wyjściowym równaniu trzeba by było robić sztuczki z wrońskianami, albo przerobić równanie na równanie macierzowe pierwszego rzędu) i znajdziemy, że szczególnym rozwiązaniem tego równania pierwszego rzędu jest $\frac{1}{2}\,x^2 e^{-x}$. Lekcja, jaką z tego wyciągamy jest taka, że trzeba szukać rozwiązania w postaci wielomianu drugiego stopnia razy e^{-x}. Podstawmy więc $y^{(1)}_{\rm inhom} = (\alpha x^2 + \beta x)\,e^{-x}$. Dostajemy wtedy po lewej stronie

$$\begin{aligned} &[\alpha\,(-x^2 + 6x - 6) + \beta\,(-x + 3)]\,e^{-x} + [\alpha\,(x^2 - 4x + 4) + \beta\,(x - 2)]\,e^{-x} \\ &+[\alpha\,(-x^2 + 2x) + \beta\,(-x + 1)]\,e^{-x} + [\alpha\,x^2 + \beta\,x]\,e^{-x} \\ &\qquad\qquad\qquad\qquad\qquad\qquad = [\alpha\,(4x - 4) + 2\beta]\,e^{-x}\,, \end{aligned}$$

Jest więc jasne, że trzeba przyjąć $\alpha = 1/4$ i $\beta = 1/2$. Podobnie spróbujmy skonstruować $y^{(2)}_{\rm inhom}$ jako $(\alpha' x\cos x + \beta' x\sin x)\,e^{-x}$ (o jedną potęgę x-a więcej niż po prawej stronie). Po lewej stronie dostaniemy wtedy

$$-2\alpha'\,(\cos x + \sin x) + 2\beta'\,(\cos x - \sin x)\,,$$

i widać, że tu trzeba położyć $\alpha' = -1/4$, $\beta' = 1/4$.

Ostatecznie więc całka ogólna wyjściowego równania ma postać

$$y = A\,e^{-x} + D\cos x + E\sin x + \frac{1}{4}\,(x^2+2x)\,e^{-x} - \frac{1}{4}\,x\,(\cos x - \sin x)\,.$$

Zadanie 49: Jest to przypadek z "degeneracją", tj. dwoma liniowo niezależnymi rozwiązaniami równania jednorodnego są

$$y_1 = \exp(-at/2)\,, \qquad y_2 = t\,\exp(-at/2)\,.$$

Rozwiązanie równania niejednorodnego otrzymujemy robiąc sztuczkę z wrońskianem:

$$W(y) \equiv y_1 y_2' - y_1' y_2 = e^{-at}\,.$$

Funkcje $A_1(t)$ i $A_2(t)$ w rozwiązaniu $y_{\rm inhom} = A_1(t)y_1(t) + A_2(t)y_2(t)$ spełniają równania

$$\begin{aligned} A_1' &= -\frac{y_2(y)}{W(y)}\,f(t) = -t\,e^{-at/2}\,f(t)\,,\\ A_2' &= \frac{y_1(y)}{W(y)}\,f(t) \;= e^{-at/2}\,f(t)\,. \end{aligned}$$

Zatem całka ogólna ma postać

$$y = (A + B\,t)\,e^{-at/2} - e^{-at/2}\int^t dt'\,t'\,e^{at'/2}\,f(t') + t\,e^{-at/2}\int^t dt'\,e^{at'/2}\,f(t')\,.$$

Dość łatwo sprawdzić przez bezpośrednie podstawienie pochodnych

$$\begin{aligned} y_{\rm inhom}' &= \frac{a}{2}\,e^{-at/2}\int^t dt'\,t'\,e^{at'/2}\,f(t') - t\,f(t)\\ &\quad + \left(1 - \frac{a}{2}\,t\right)e^{-at/2}\int^t dt'\,e^{at'/2}\,f(t') + t\,f(t)\,,\\ y_{\rm inhom}'' &= -\frac{a^2}{4}\,e^{-at/2}\int^t dt'\,t'\,e^{at'/2}\,f(t') + \frac{a}{2}\,t\,f(t)\\ &\quad + \left(\frac{a^2}{4}\,t - a\right)e^{-at/2}\int^t dt'\,e^{at'/2}\,f(t') + \left(1 - \frac{a}{2}\,t\right)f(t)\,, \end{aligned}$$

i samej funkcji $y_{\rm inhom}$ do równania $y'' + a\,y' + \frac{1}{4}\,a^2\,y = f(t)$, że jest ono spełnione.

To samo rozwiązanie można także dostać przepisując równanie w postaci macierzowej

$$\frac{d}{dt}\begin{pmatrix} y_1\\ y_2\end{pmatrix} = \begin{pmatrix} 0 & 1\\ -\frac{1}{4}a^2 & -a\end{pmatrix}\begin{pmatrix} y_1\\ y_2\end{pmatrix} + \begin{pmatrix} 0\\ f(t)\end{pmatrix},$$

gdzie $y_1 = y$, a $y_2 = y'$. Macierz ma tu tylko jeden wektor własny i jedną wartość własną $\lambda = -a/2$. Macierz $\exp(tF)$ znajdujemy metodą CH korzystając ze sztuczki z różniczkowaniem. W ten sposób układ równań $e^{-ta/2} = a_1(-a/2) + a_0$, $t\, e^{-ta/2} = a_1$ daje $a_0 = (1 + at/2)\, e^{-ta/2}$ i stąd

$$e^{tF} = e^{-ta/2} \begin{pmatrix} 1 + \frac{a}{2}\, t & t \\ -\frac{a^2}{4}\, t & 1 - \frac{a}{2}\, t \end{pmatrix},$$

Rozwiązanie równania niejednorodnego ma postać $\mathbf{y}_{\text{inhom}} = \exp(tF) \cdot \mathbf{h}(t)$, gdzie

$$\begin{pmatrix} h_1' \\ h_2' \end{pmatrix} = e^{ta/2} \begin{pmatrix} 1 - \frac{a}{2}\, t & -t \\ \frac{a^2}{4}\, t & 1 + \frac{a}{2}\, t \end{pmatrix} \begin{pmatrix} 0 \\ f(t) \end{pmatrix}.$$

Stąd szczególnym rozwiaząniem macierzowego równania niejednorodnego jest

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = e^{-ta/2} \begin{pmatrix} 1 + \frac{a}{2}\, t & t \\ -\frac{a^2}{4}\, t & 1 - \frac{a}{2}\, t \end{pmatrix} \begin{pmatrix} -\int dt\, t\, e^{at/2} f(t) \\ \int dt\, (1 + at/2)\, e^{at/2} f(t) \end{pmatrix},$$

a szczególnym rozwiaząniem wyjściowego równania jest y_1:

$$y_1 = e^{-ta/2} \left[-\left(1 + \frac{a}{2}\, t\right) \int dt\, t\, e^{at/2}\, f(t) + t \int dt\, \left(1 + \frac{a}{2}\, t\right) e^{at/2}\, f(t) \right].$$

Widać, że jest to to samo, co metododa z wrońskianem.