Metryka a geometria przestrzeni

Jarosław GÓRNICKI, Michał LORENS, Rzeszów

Fragment matematyki, w którym przeplatają się metody analityczne, topologiczne i algebraiczne nazywamy *analizą funkcjonalną*. Obiektami badań są tu m.in. przestrzenie funkccyjne, operatory i ich własności. Dojrzały wiek tej dyscypliny rozpoczął się w latach trzydziestych obecnego stulecia, gdy ukazała się monografia Stefana Banacha (1892–1945), *Théorie des opérations lineaires*, Monografie Matematyczne t. I, Warszawa 1932.

Oto, co S. Banach pisze w przedmowie (pisownia oryginalna):

Niema prawie dziedziny matematyki, gdzieby teorja powyższa nie wnikła w sposób istotny. Dość wspomnieć, że rachunek wariacyjny i teoria równań całkowych okazały się szczególnymi przypadkami ogólnych działów zawartych w teorji operacyj. Piękność teorji operacyj leży głównie w tem, że w niej łączą się w harmonijną całość metody matematyki klasycznej z metodami matematyki nowożytnej.

Współczesna analiza funkcjonalna to obszerna dyscyplina rozwijająca się w różnych kierunkach. Wiele uzyskiwanych w niej wyników zależy od „geometrii rozważanej przestrzeni". To co kryje się pod tym terminem może być rozumiane w różny sposób. My przez to określenie będziemy rozumieć wpływ metryki (sposobu mierzenia odległości) na własności przestrzeni.

Nasze rozważania ograniczymy głównie do przestrzeni $\mathbb{R}^2$, która jest najbliższa matematyce szkolnej. Jednak siła i znaczenie tego zagadnienia ujawnia się przede wszystkim w przestrzeniach liniowych nieskończenie wymiarowych (liniowość przestrzeni gwarantuje nam to, że wykonywalne są w niej operacje dodawania do siebie elementów i mnożenia ich przez liczby).

1. Metryki, kule sfery i ... deformacje

Do mierzenia odległości między punktami x i y w niepustym zbiorze E możemy używać różnych funkcji zwanych metrykami (pojęcie to wprowadził w 1906 roku M. R. Frechet (1878-1973)). Muszą one jednak spełniać kilka naturalnych warunków, mianowicie:

funkcja $d : E \times E \to [0, +\infty)$ jest *metryką*, gdy dla dowolnych $x, y, z \in E$

1) $d(x, y) = 0 \Leftrightarrow x = y$,

2) $d(x, y) = d(y, x)$ – warunek symetrii,

3) $d(x, y) \leq d(x, z) + d(z, y)$ – nierówność trójkąta.

Nasuwa się pytanie: czy w każdym niepustym zbiorze można określić metrykę? Odpowiedź brzmi – TAK!

Można to uczynić w zbiorze E w następujący sposób:

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

Inne przykłady metryk znamy ze szkoły, np.: naturalna odległość w zbiorze $\mathbb{R}$ liczb rzeczywistych. W kartezjańskim prostokątnym układzie współrzędnych XOY odległość pomiędzy punktami $P_1(x_1, y_1)$ i $P_2(x_2, y_2)$ możemy wyrazić wzorem

$$d_e(P_1, P_2) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}\,.$$

Funkcja d_e oczywiście spełnia warunki 1), 2), 3) (proszę sprawdzić!). Jest to tzw. *odległość euklidesowa* w $\mathbb{R}^2$.

Zauważmy teraz, że dysponując na płaszczyźnie $\mathbb{R}^2$ dwiema metrykami d_{0-1} i d_e, możemy określić rodzinę metryk (rozważania te mogą być prowadzone nie tylko na płaszczyźnie, ale i w dowolnym zbiorze E z dwiema różnymi metrykami – w tym artykule wszystkie przykłady dotyczyć będą dla uproszczenia zbiorów w $\mathbb{R}^n$):

$$(1)\quad d_t(P_1, P_2) := (1-t) \cdot d_{0-1}(P_1, P_2) + t \cdot d_e(P_1, P_2), \qquad 0 \leq t \leq 1\,.$$

Funkcja d_t jest metryką dla $t \in (0, 1)$. Mamy bowiem:

1) $P_1 = P_2 \Rightarrow d_t(P_1, P_1) =$

$$= (1-t) \cdot d_{0-1}(P_1, P_1) + t \cdot d_e(P_1, P_1) = 0\,.$$

$$d_t(P_1, P_2) = 0 \Rightarrow$$
$$\Rightarrow (1-t) \cdot d_{0-1}(P_1, P_2) + t \cdot d_e(P_1, P_2) = 0 \Rightarrow$$
$$\Rightarrow d_{0-1}(P_1, P_2) = 0 \wedge d_e(P_1, P_2) = 0 \Rightarrow$$
$$\Rightarrow P_1 = P_2\,;$$

2) warunek symetrii jest spełniony wprost z definicji;

3)
$$\begin{aligned} d_t(P_1, P_2) &= (1-t)d_{0-1}(P_1, P_2) + t \cdot d_e(P_1, P_2) \\ &\leq (1-t) \cdot [d_{0-1}(P_1, P_3) + d_{0-1}(P_3, P_2)] \\ &\quad + t \cdot [d_e(P_1, P_3) + d_e(P_3, P_2)] \\ &= \{(1-t) \cdot d_{0-1}(P_1, P_3) + t \cdot d_e(P_1, P_3)\} \\ &\quad + \{(1-t) \cdot d_{0-1}(P_3, P_2) + t \cdot d_e(P_3, P_2)\} \\ &= d_t(P_1, P_3) + d_t(P_3, P_2)\,. \end{aligned}$$

Możliwa jest więc ciągła deformacja (ciągłe przejście od) jednej metryki do drugiej!

Spróbujmy teraz zastanowić się nad tym, jaki wpływ ma deformacja metryki na kształt kuli (sfery)?

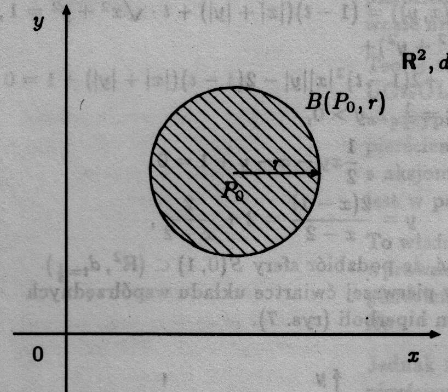
Wcześniej jednak przypomnijmy, że kulą (domkniętą) o środku w punkcie p i o promieniu $r > 0$, w przestrzeni metrycznej (E, d) nazywamy zbiór

$$B(p, r) = \{x \in E : d(x, p) \leq r\}\,.$$

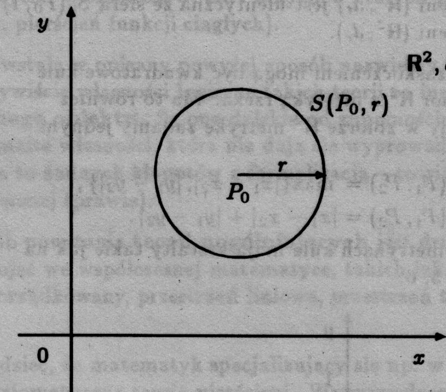
Natomiast zbiór

$$S(p, r) = \{x \in E : d(x, p) = r\}$$

nazywamy sferą.



Rys. 1.



Rys. 2.

Już pierwsza obserwacja może być dla Czytelnika poznającego ten temat zaskakująca: w niektórych metrykach kształt kuli (sfery) zależy od wielkości promienia r. Na przykład w zbiorze R^2 z metryką d_{0-1}:

$$B(p,r) = \begin{cases} \{p\} & \text{dla } 0 < r < 1, \\ R^2 & \text{dla } r \geq 1, \end{cases}$$

$$S(p,r) = \begin{cases} \emptyset & \text{dla } 0 < r < 1, \\ R^2 - \{p\} & \text{dla } r = 1, \\ \emptyset & \text{dla } r > 1. \end{cases}$$

Z tego punktu widzenia nie jest to metryka „stabilna". Podobnie „niestabilne" ale z innego punktu widzenia są kule (sfery) w zbiorze R^2 z tzw. metryką rzeka:

$$d_{rz}(P_1, P_2) = \begin{cases} |y_2 - y_1|, & \text{gdy } x_1 = x_2, \\ |y_2| + |x_2 - x_1| + |y_1|, & \text{gdy } x_1 \neq x_2. \end{cases}$$

Tutaj kształt kul $B(P_0, r)$ (sfer $S(P_0, r)$) zależy nie tylko od wielkości promienia $r > 0$, ale również od położenia punktu P_0 (proszę sporządzić odpowiednie rysunki!).

Inaczej wygląda sytuacja w zbiorze R^2 z metryką euklidesową. Wiadomo, że kule bądź sfery o środku w punkcie P_0 i promieniu $r > 0$ mają tu kształt taki jak na rysunkach 1, 2, i co ważniejsze kształt ten nie zależy ani od wielkości $r > 0$, ani od położenia punktu P_0.

Zobaczmy teraz sfery w zbiorze R^2, gdy odległość mierzymy za pomocą metryki d_t określonej wzorem (1) dla $0 < t < 1$.

Badamy zbiór

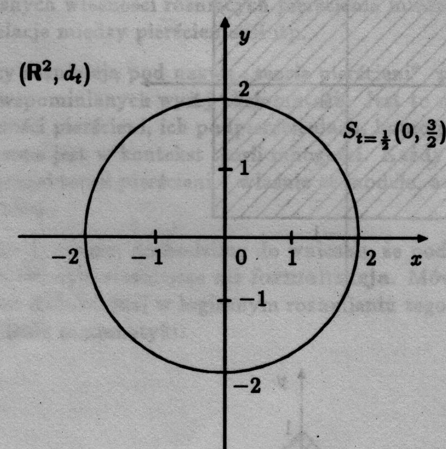
$$S_t(P_0, r) = \{P \in R^2 : d_t(P, P_0) = r\}, \quad r > 0.$$

Gdy $P = P_0$, to $P \notin S_t(P_0, r)$.

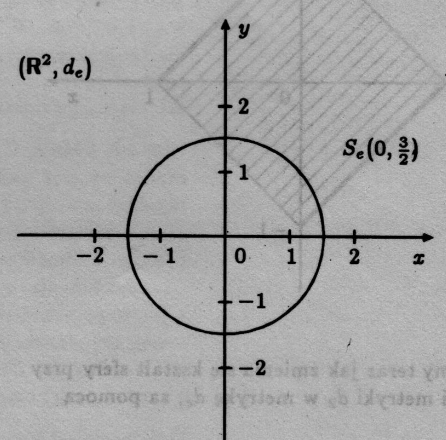
Gdy $P \neq P_0$, to $d_{0-1}(P, P_0) = 1$, i mamy:

$$d_e(P, P_0) = \frac{r + t - 1}{t}, \qquad (t > 0).$$

Ponieważ metryka jest funkcją nieujemną, więc musi



Rys. 3.



Rys. 4.

być $r + t - 1 \geq 0$ czyli $r \geq 1 - t$, a to oznacza, że:

– dla $0 < r \leq 1 - t$, sfera $S_t(P_0, r) = \emptyset$;

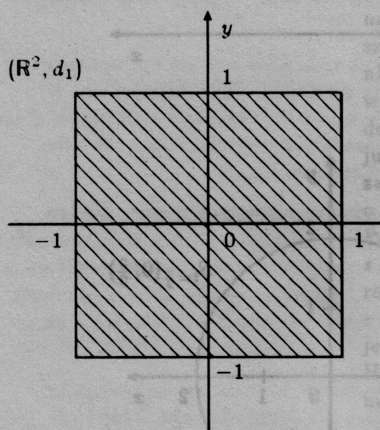
– dla $r > 1 - t$, sfera $S_t(P_0, r)$ w przestrzeni (R^2, d_t) jest taka sama jak sfera $S_e(P_0, R)$ o promieniu $R = 1 + \frac{r-1}{t}$ w przestrzeni (R^2, d_e) – rys. 3, 4.

W szczególności, gdy $r = 1$, to sfera $S_t(P_0, 1)$ w przestrzeni $(\mathbb{R}^2, d_t)$ jest identyczna ze sferą $S_0(P_0, 1)$ w przestrzeni $(\mathbb{R}^2, d_e)$.

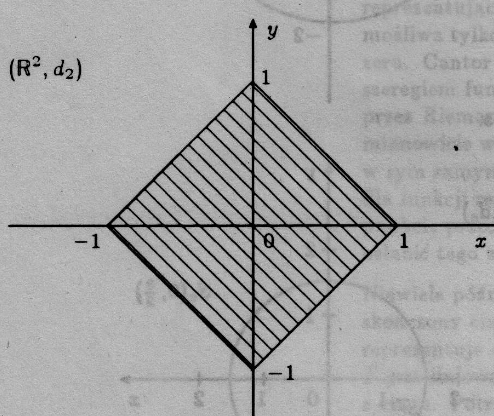
Kolejnym zaskoczeniem mogą być kwadratowe kule – patrz zbiór $\mathbb{R}^2$ z metryką rzeka. Ma to również miejsce, gdy w zbiorze $\mathbb{R}^2$ metrykę zadamy jednym ze wzorów:

$$d_1(P_1, P_2) = \max\{|x_1 - x_2|, |y_1 - y_2|\},$$
$$d_2(P_1, P_2) = |x_1 - x_2| + |y_1 - y_2|.$$

Przy tych metrykach kule mają kształty takie jak na rysunkach 5, 6.



Rys. 5.



Rys. 6.

Prześledźmy teraz jak zmienia się kształt sfery przy deformacji metryki d_2 w metrykę d_e, za pomocą metryk

$$\widehat{d}_t(P_1, P_2) := (1-t)\cdot d_2(P_1, P_2) + t\cdot d_e(P_1, P_2),$$
$$0 \le t \le 1.$$

Zrobimy to na przykładzie sfery o środku w początku układu współrzędnych i o promieniu $r = 1$.
Dla $t = 0$ oraz $t = 1$ otrzymujemy dobrze znane kształty – patrz rys. 6 i 2. Jaki kształt ma sfera, gdy $t \in (0, 1)$?

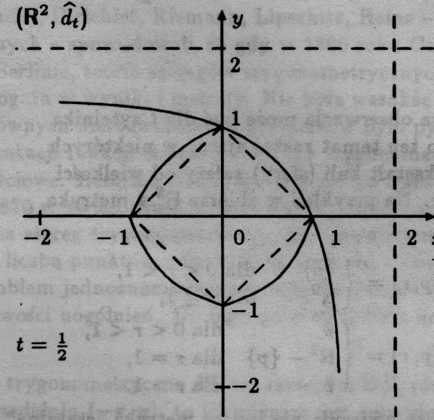
Zauważmy, że

$$\widehat{d}_t((0,0),(x,y)) = (1-t)(|x| + |y|) + t\cdot\sqrt{x^2 + y^2} = 1,$$
$$(1-2t)(x^2 + y^2) +$$
$$+2(1-t)^2|x||y| - 2(1-t)(|x| + |y|) + 1 = 0.$$

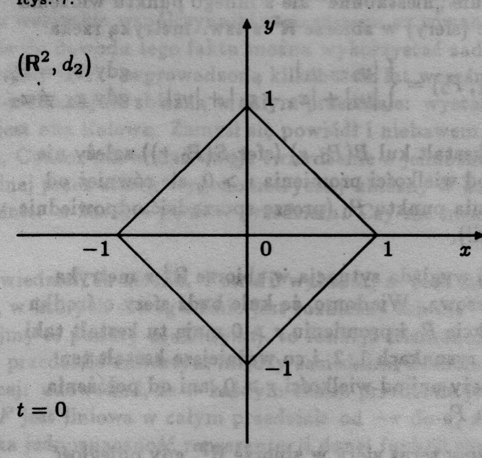
Stąd dla $t = \frac{1}{2}$, $x, y > 0$,

$$\frac{1}{2}xy - x - y + 1 = 0,$$
$$y = \frac{2(x-1)}{x-2} = 2 + \frac{2}{x-2},$$

i już widać, że podzbiór sfery $S(0,1) \subset (\mathbb{R}^2, d_{t=\frac{1}{2}})$ zawarty w pierwszej ćwiartce układu współrzędnych jest łukiem hiperboli (rys. 7).



Rys. 7.



Rys. 8.

Zupełnie podobnie możemy zobaczyć jak zmienia się na płaszczyźnie kształt sfery $S(0,1)$ przy deformacji metryki d_2 w metrykę d_1 za pomocą metryk

$$d_t^*(P_1, P_2) := (1-t)\cdot d_2(P_1, P_2) + t\cdot d_1(P_1, P_2),$$
$$0 \le t \le 1.$$

Obliczamy

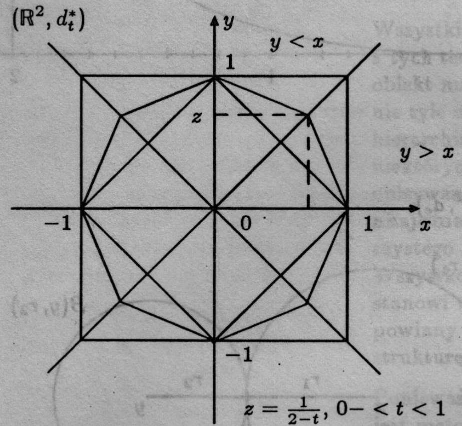
$$d_t^*((0,0),(x,y)) =$$
$$= (1-t)(|x| + |y|) + t\cdot\max(|x|, |y|) = 1,$$

i dla $x, y > 0$, korzystając ze wzoru $\max(a,b) = \frac{1}{2}(|a-b| + a + b)$, rozpatrujemy dwa przypadki (rys. 9):

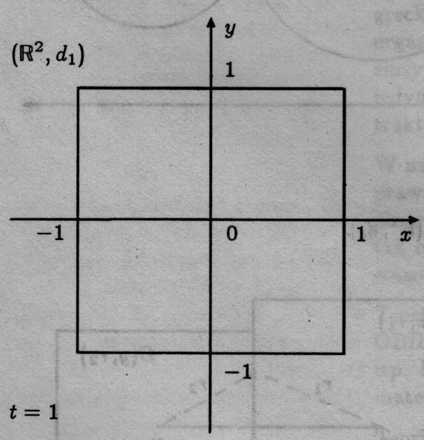
$1^\circ.\ x > y$, $\qquad 2^\circ.\ y > x$,

$$(2)\ y = \frac{1}{1-t}(1-x)\,, \qquad (3)\ y = 1-(1-t)x\,.$$

Zatem odpowiednie fragmenty sfery $S(0,1)$ zawierają się w prostych opisanych równaniami (2) i (3). Oznacza to, że „pośrednie” sfery mają kształt ośmiokątów o równych bokach.



Rys. 9.



Rys. 10.

Reasumując, kształt kuli (sfery) zależy od przyjętej metryki, położenia jej środka i wielkości promienia. Jednak, co ważniejsze, może on być modyfikowany za pomocą stosownych zmian metryki. Dlaczego badanie kształtu kuli (a w rzeczywistości metryk) jest tak istotne? Otóż w wielu sytuacjach metryka ma decydujący wpływ na wiele ważnych własności przestrzeni.

2. Okrągłe, bardziej okrągłe,... czyli jak to mierzyć

W wyniku prowadzonych badań ustalono m.in., że im kula w danej przestrzeni jest „bardziej okrągła”, tym przestrzeń ta ma bardziej „porządne” własności (zgodnie z oczekiwaniami jakich nabywamy badając trójwymiarową przestrzeń euklidesową – bardzo regularną). O innych przestrzeniach, których własności różnią się znacznie od naszych przewidywań, nabytych w wyniku badania zbioru $\mathbb{R}^3$ z metryką euklidesową lub jego naturalnych uogólnień, mówimy, że są one „osobliwe”.

Warto zastanowić się w tym momencie, których przestrzeni jest więcej – tych „osobliwych”, czy tych „porządnych”.

Matematyka szkolna może wytworzyć tu dość niebezpieczne złudzenie. W praktyce szkolnej mamy zwykle do czynienia z „porządnymi” obiektami, np. z funkcjami ciągłymi różniczkowalnymi, zbiorami wypukłymi. Stwarza to wrażenie, że w matematyce istnieją na ogół obiekty „porządne” i tylko wyjątkowo natrafia się na „przypadki zwyrodniałe”, takie jak: funkcje nieciągłe, nieróżniczkowalne, kwadraty sita. Oczywiście tak nie jest! Wystarczy chwila namysłu (a o nią wcale nie tak łatwo) aby dostrzec, że sytuacja jest wręcz przeciwna – otacza nas „grzęzawisko osobliwości”.

Już w zbiorze liczb rzeczywistych większość to liczby niewymierne, a wśród nich – przestępne. Liczby, na których potrafimy wykonać dokładne obliczenia to liczby wymierne – a tych w stosunku do całego zbioru $\mathbb{R}$ jest śmiesznie mało.

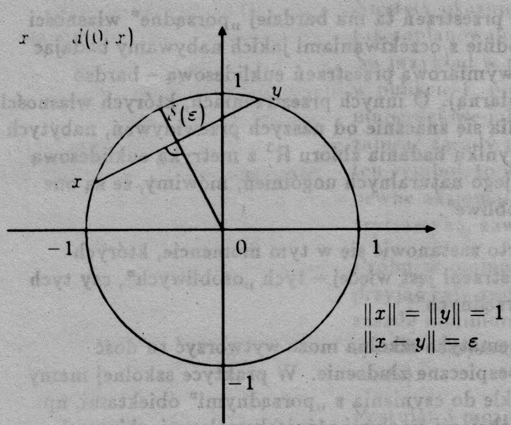
Obiekty „porządne” stanowią w matematyce zjawisko wyjątkowe. Jak zatem wskazać sposób wyszukiwania tych przestrzeni, w których kule są „okrągłe”? Wiąże się to z odpowiedzią na pytanie – okrągłe to znaczy jakie? Można odpowiedzieć na wiele sposobów, ale trzeba wybrać taki, który jest łatwy w sprawdzeniu i ma istotne następstwa.

Możemy na przykład powiedzieć, że kula jest okrągła, gdy jej brzeg (= sfera) nie zawiera żadnego odcinka. Przestrzenie, w których kule mają taką własność nazywamy *ściśle wypukłymi.* Taką jest przestrzeń $(\mathbb{R}^2, d_e)$, a przestrzenie $(\mathbb{R}^2, d_1)$, $(\mathbb{R}^2, d_2)$ tej własności już nie mają. Praktyka matematyczna pokazała, że to określenie „okrągłości kul” jest mało finezyjne. Potrzebne było subtelniejsze określenie, umożliwiające porównywanie „stopnia okrągłości kul”. Jak to zrobić?

Umieśćmy w tym celu w kuli jednostkowej odcinek o długości ε $(0 < \varepsilon \leq 2)$ i badajmy odległość jego środka od brzegu kuli (rys. 11). Precyzyjniej, określamy funkcję $\delta_E : [0,2] \to [0,1]$ wzorem

$$\delta_E(\varepsilon) = \\ = \inf\left\{1 - \frac{\|x+y\|}{2} : \|x\| = \|y\| = 1,\ \|x-y\| \geq \varepsilon\right\}$$

– przez $\|x\|$ można rozumieć w $\mathbb{R}^2$ odległość punktu x od punktu $(0,0)$. Funkcja ta zwana jest *modułem wypukłości*, a przestrzenie spełniające warunek: $\delta_E(\varepsilon) > 0$ dla $\varepsilon > 0$, nazywamy *jednostajnie wypukłymi* (J. Clarkson, 1936).



Rys. 11.

Z prostych rachunków wynika, że w przypadku przestrzeni $(\mathbb{R}^2, d_e)$

$$(4) \qquad \delta(\varepsilon) = 1 - \sqrt{1 - \left(\frac{\varepsilon}{2}\right)^2}\,, \quad 0 \le \varepsilon \le 2\,,$$

a zatem $(\mathbb{R}^2, d_e)$ jest przestrzenią jednostajnie wypukłą. Oczywiście każda przestrzeń jednostajnie wypukła jest ściśle wypukła, ale nie na odwrót. Ponadto ważne w badaniach matematycznych oraz w zastosowaniach przestrzenie Hilberta również są jednostajnie wypukłe, a ich moduł wypukłości wyraża się wzorem (4). Co więcej, G. Nordlander wykazał, że kule w przestrzeniach Hilberta $\mathcal{H}$ są „najbardziej okrągłe" w takim znaczeniu, że dla innych przestrzeni Banacha $(E, \|\cdot\|)$, $\delta_E(\varepsilon) \le \delta_{\mathcal{H}}(\varepsilon)$, gdy $0 \le \varepsilon \le 2$.

3. Co z tego wynika?

Czas teraz powiedzieć o wpływie metryki (= kształtu kuli) na własności przestrzeni. Uczynimy to podając kilka faktów, patrząc w jakim stopniu zależą one od metryki.

Przykład 1.

W liniowej przestrzeni metrycznej (E, d), *środkiem algebraicznym* odcinka $\overline{ab}$ nazywamy punkt $\frac{1}{2}a + \frac{1}{2}b$, zaś *środkiem metrycznym* odcinka $\overline{ab}$ punkt $c \in \overline{ab}$ o tej własności, że $d(a, c) = d(c, b)$.

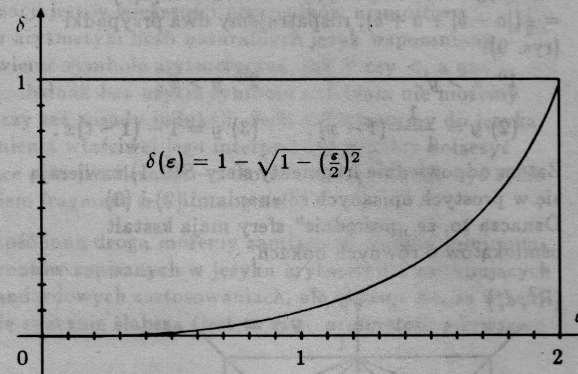
W przypadku przestrzeni metrycznej $(\mathbb{R}^2, d_e)$ pojęcia te pokrywają się – środek algebraiczny dowolnego odcinka jest identyczny z jego środkiem metrycznym.

Spójrzmy teraz na zbiór $\mathbb{R}^2$ z tzw. „odległością wzdłuż cięciwy"

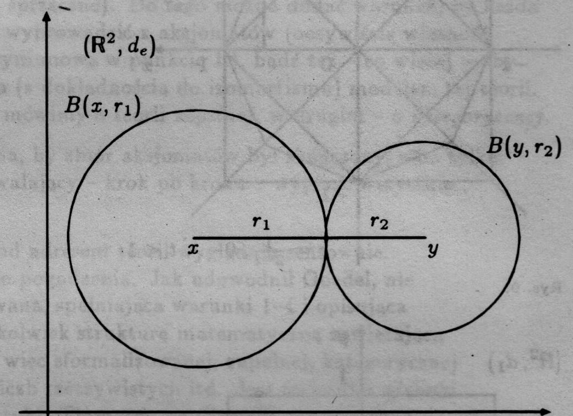
$$d_*(P_1, P_2) = \frac{d_e(P_1, P_2)}{\sqrt{(1 + x_1^2 + y_1^2)(1 + x_2^2 + y_2^2)}}\,,$$

$$P_1 = (x_1, y_1), \quad P_2 = (x_2, y_2)\,.$$

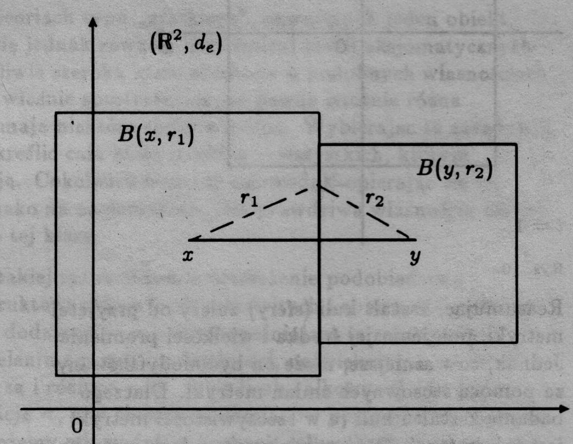
Gdy w przestrzeni $(\mathbb{R}^2, d_*)$ rozważymy odcinek o końcach $P_1 = (0,0)$, $P_2 = (2,0)$, to stwierdzimy, że środkiem algebraicznym odcinka $\overline{P_1P_2}$ jest punkt $(\frac{\sqrt{5}-1}{2}, 0)$.



Rys. 12.



Rys. 13.



Rys. 14.

Przykład 2.

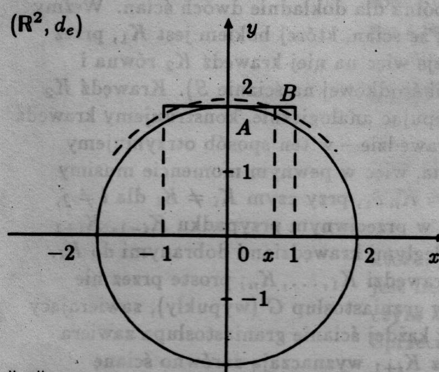
Jeżeli (E, d) jest liniową przestrzenią metryczną ściśle wypukłą, to kule $B(x, r_1)$ i $B(y, r_2)$ takie, że $x \ne y$, $r_1 + r_2 = d(x, y)$, mają dokładnie jeden punkt wspólny (tak jak na płaszczyźnie euklidesowej).

W przypadku przestrzeni metrycznej (R^2, d_e) tę sytuację ilustruje rysunek 13. Stwierdzenie to przestaje być prawdziwe, gdy metrykę d_e zastąpimy metryką d_1 – patrz rysunek 14.

Przykład 3.

W każdym niepustym, wypukłym, domkniętym podzbiorze jednostajnie wypukłej przestrzeni Banacha istnieje dokładnie jeden element o najmniejszej normie.

Dla ilustracji niech M będzie odcinkiem prostoliniowym łączącym punkty $(-1,2), (1,2)$ w zbiorze R^2. Gdy posługujemy się metryką d_e to elementem o najmniejszej normie w zbiorze M jest punkt $(0,2)$ i taki punkt jest dokładnie jeden (rys. 15).

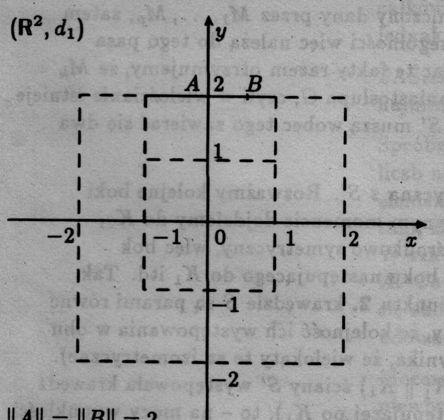


$\|A\| = 2$

$\|B\| = \sqrt{x^2+2^2} > 2$

Rys. 15.

Zmieniając metrykę na d_1 tracimy jednoznaczność – wszystkie punkty zbioru M mają tę samą normę (rys. 16).



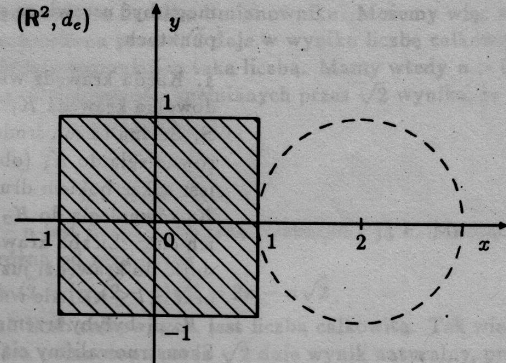
$\|A\| = \|B\| = 2$

Rys. 16.

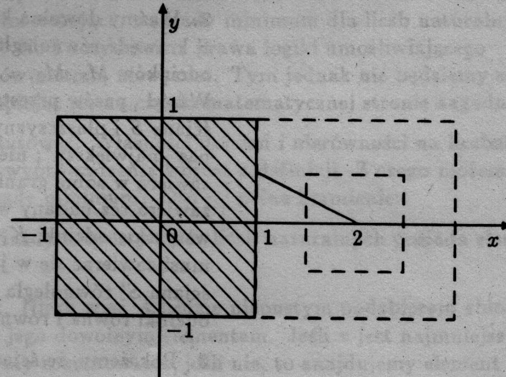
Przykład 4.

Jeżeli C jest niepustym, domkniętym, wypukłym podzbiorem jednostajnie wypukłej przestrzeni Banacha, to dla dowolnego $z \notin C$ istnieje jeden element $x \in C$ taki, że $d(z, C) = d(z, x)$.

Dla ilustracji tego stwierdzenia niech C będzie kwadratem domkniętym o wierzchołkach $(1,1), (1,-1), (-1,-1), (-1,1)$ w zbiorze R^2. Jeżeli $z = (2,0)$, to w metryce d_e istnieje dokładnie jeden element $x = (1,0) \in C$, dla którego $d_e(z, C) = d_e(z, x)$ (rys. 17). W przypadku metryki d_1 istnieje wiele punktów $x \in C$, dla których $d_1(z, C) = d_1(z, x)$ (rys. 18).



Rys. 17.



Rys. 18.

Generalizując, ale jednocześnie upraszczając sytuację, możemy podsumować: odpowiednio wysoki stopień „okrągłości kul" gwarantuje szeroko rozumianą jednoznaczność wielu funkcjonałów, tzn. funkcji rzeczywistych określonych na danej przestrzeni liniowej.