Výfučtení: Grafy

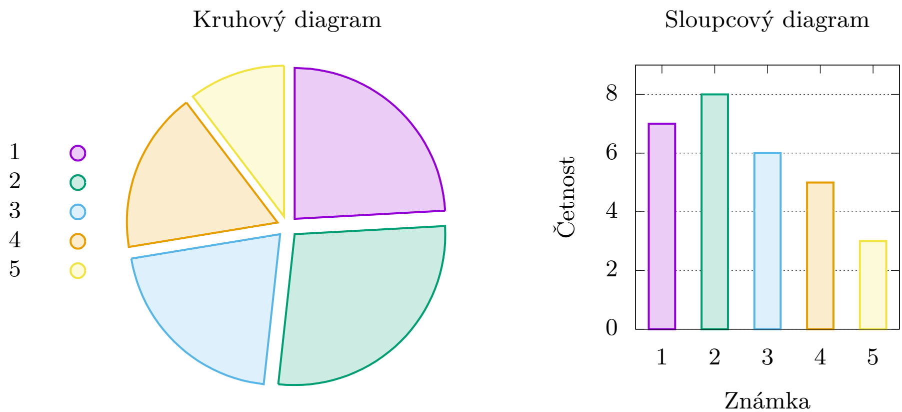
Úvod

Při zkoumání (a nejen tehdy) se potkáváme s velkým množstvím dat. Pokud je používáme k dalším výpočtům, je jejich obyčejná tabulková forma poměrně praktická. Chceme-li ovšem z těchto dat vyvozovat nějaké závěry, dostat reálnou představu o jejich rozložení či vyobrazit jednotlivé datové záznamy v rámci celku, neobejdeme se bez jejich grafického vyobrazení ve formě grafů.

Vzhledem k tomu, že neexistuje pouze jeden typ dat, neexistuje ani jeden typ grafu, ale obecně se vždy jedná o nějaké vizuální vyobrazení. Např. měření proudu, který teče během dne do zásuvky, vyžaduje jinou reprezentaci, než když znázorňujeme výsledky voleb podle věkových kategorií. Proto se v tomto Výfučtení podíváme na různé grafy a pokusíme se vysvětlit, jaké jsou mezi nimi rozdíly.

Diagram

V obyčejném životě se nejčastěji setkáme s diagramy (anglicky *charts*). To jsou například sloupcové nebo kruhové (tzv. koláčové) grafy. Reprezentativně[1] vyobrazují výsledky měření pro dané hodnoty zkoumané veličiny. Pokud navíc používáme sloupcové grafy, většinou vrcholy sloupců nemůžeme propojit (nemáme důvod je propojovat), protože z hodnoty na hodnotu neexistuje přímá závislost či užitečný pojem vývoje (výdělek v jednotlivých měsících, počet jednotlivých známek ve třídě atd.). Na grafech níže např. vidíme dvě znázornění výsledných známek z jednoho testu v jedné třídě.

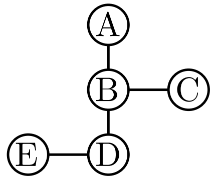


Obr. 1: Porovnání zobrazení dat ve dvou typech diagramů na příkladu známek třídy z testu.

[1] V ideálním případě, avšak v dnešní době bohužel není výjimkou se setkat s grafy poupravenými pro potřebu manipulace nebo dezinformace.

Teorie grafů a diskrétní matematika

Jako „graf" můžeme také označovat množství věcí nebo bodů, které jsou spolu určitým způsobem propojeny. Takové grafy jsou používány v tzv. diskrétní matematice (a poté i v informatice) pro popis prvků v množině a vztahů mezi jednotlivými prvky. Sama diskrétní matematika je průřezem všech oborů matematiky, kde se zabýváme např. vlastnostmi celých čísel (nikoli všech reálných čísel) či konečnými množinami (a nikoli např. nekonečnými množinami bodů na úsečce). Její podstatnou součástí je právě tzv. *teorie grafů*, tj. nauka o grafech, které znázorňují propojení různých prvků.



Jednoduchou analogií pro představu tohoto popisu je vlaková síť. Zmínili jsme se o tom, že jde o grafy vyobrazující propojení prvků v množině, kterými by v naší analogii byly vlakové stanice. Tyto prvky (v grafech také zvané *uzly*) jsou propojeny tzv. *hranami* (kolejemi mezi stanicemi). Ty se dále mohou dělit na *orientované* (jednosměrná kolej) a *neorientované* (obousměrná kolej). V některých případech může hrana mít i určitou hodnotu, jejíž význam záleží na konkrétní situaci, ve které graf používáme (v našem případě by hodnota hrany představovala například délku spojení mezi stanicemi nebo čas, za který se vlak dostane z jedné zastávky do druhé – u grafů diskrétní matematiky totiž nezáleží na grafickém provedení obrázku, délkách čar nebo velikostech bodů...).

Více se o popisované teorii grafů můžete dozvědět v rámci matematických seminářů či na vysoké škole. Rozvoj tohoto oboru přinesl mnohé technické pokroky, např. v hledání ideálního zapojení elektrické či telefonní sítě, v realizaci nejlepších tras rozvozu poštovních zásilek (či pizzy), anebo třeba při modelování struktury mozku.

Funkce nebo data

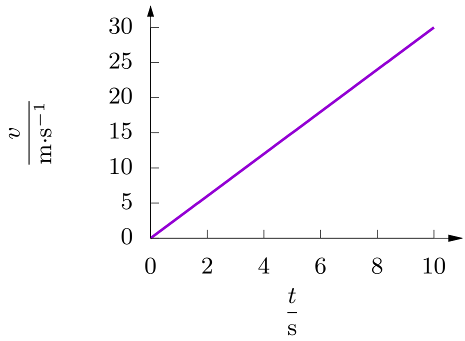
Ve školách a vědě jsou nejběžněji užívané tzv. *grafy funkcí* (obrázek 2), jež vykreslují průběh změn nějaké veličiny nebo naměřená data. Jako „funkci" označujeme to, co takovým grafem zachycujeme – zhruba řečeno jde o závislost nějakého výsledku (tzv. *závislé proměnné*) na vstupním parametru (tzv. *nezávislé proměnné*), např. závislost rychlosti na čase, proudu na napětí atd. Funkce mohou být *spojité* (jejich grafy malujeme nepřerušenou čarou), ale také *diskrétní* (např. naměřená data konečného množství, když známe hodnoty závislé proměnné pouze pro konečný počet hodnot té nezávislé, což je obvyklý případ při experimentech – graf pak zaznamenáme jen jako výběr nespojených bodů).

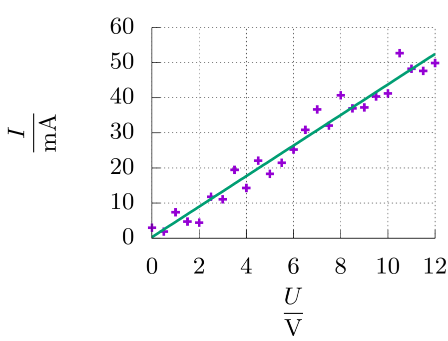
Přesnější matematická definice dává pro závislosti mezi veličinami ještě jednu podmínku, kterou bychom měli mít na paměti, jinak nejde o funkce: ke každé hodnotě nezávislé proměnné může být přiřazena nejvýše jedna hodnota závislé! Hovorově řečeno to znamená, že graf jakékoli funkce může sice vystoupat vícekrát do stejné výšky, ale musí mít stále jen jednu hodnotu (čáry či body grafu nesmějí ležet nikdy přímo nad sebou).

Složení grafu

Osy

Osy určují hodnoty vyobrazených bodů, a jsou tak základními součástmi grafu. Obvykle bývají okolo zakreslených dat, ale v některých případech je můžeme najít i přímo mezi daty. Protnutí os určuje počátek grafu.





Obr. 2: Příklady grafů jednoduchých funkcí. Vlevo je znázorněna lineární závislost rychlosti na čase. Napravo jde vlastně o proložení dvou grafů v jednom: jeden může zobrazovat diskrétní průběh proměnlivého měření proudu v závislosti na napětí, zatímco druhý je teoretickým spojitým průběhem, který se nejvíce blíží onomu měření, avšak předpokládá dokonalou přímou úměru.

Osy jsou označeny pomocí popisků, které se nachází pod/vedle/nad osami, a říkají nám, co daná osa vyjadřuje a v jakých jednotkách.[2] (Samozřejmě v případě, že zobrazované veličiny skutečně nějaký rozměr mají. V matematice obvykle pracujeme s bezrozměrnými veličinami.)

Stejně tak je důležité na osách zvolit správné *měřítko* (vhodnou velikost dílků souřadnicové sítě a jejich popisků). Bez něj by totiž data o ničem nevypovídala, protože bychom nebyli schopni určit, jakou hodnotu má daný bod vyobrazený v grafu. S měřítkem se pojí také *meze grafu*, které jsou důležité hlavně při relativním porovnávání vyobrazených údajů.[3] V neposlední řadě můžeme do grafu pro lepší čitelnost vynést čtvercovou mřížku.

Počátek

Počátek grafu je definován jako bod $[0; 0]$ nebo bod protnutí os (obě definice jsou stejně smysluplné a v mnoha případech se setkáte se situací, kdy se osy protínají právě v bodě $[0; 0]$). Je důležitý hlavně z důvodu popisu vlastností funkcí.

Kvadranty

Celý graf můžeme rozdělit na 4 menší části zvané *kvadranty*. Ty jsou vymezeny nulovými hodnotami na ose x a y. Každý kvadrant nazýváme pořadovým číslem, přičemž první kvadrant se nachází vpravo nahoře (kladné x i y) a jejich číslování pokračuje v protisměru hodinových ručiček.

Dělení grafů na kvadranty je také důležité při popisů funkcí. Funkce totiž nemusí procházet všemi kvadranty, ale díky popisu, ve kterých se nachází, můžeme odvodit, jakých hodnot funkce nabývá.

[2] Tato informace je poměrně důležitá. Často se při vytváření grafů stává, že se na jednotky zapomene, ovšem bez nich nemá graf žádnou vypovídající hodnotu.

[3] Špatného vymezení mezí, respektive zavádějícího vymezení, se využívá hlavně v médiích, kupříkladu v momentě, kdy chceme něco ukázat v lepším nebo horším světle. Pak například nezvolíme spodní mez nulovou, čímž dojde ke grafickému zvýraznění rozdílů všech hodnot.

kvadrant	I	II	III	IV
osa x	> 0	< 0	< 0	> 0
osa y	> 0	> 0	< 0	< 0

Tab. 1: Tabulka hodnot pro jednotlivé kvadranty.

Čtení z grafu

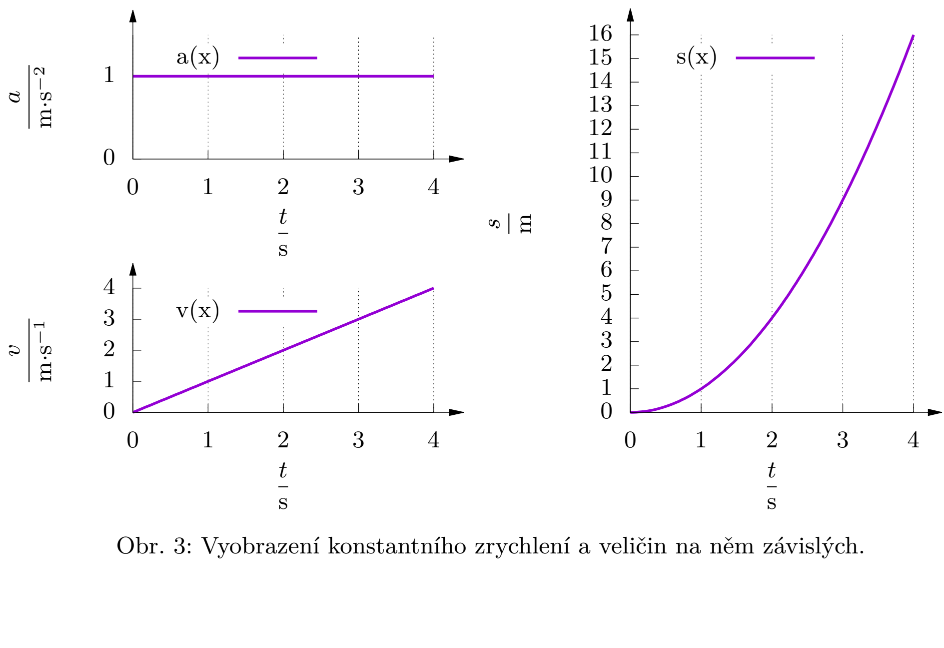
Správné čtení z grafu je základem jeho používání.

Pro každý bod platí, že když sestrojíme kolmici na osu, která bude procházet daným bodem, bude pata kolmice určovat hodnotu bodu na dané ose.

Vezměme si příklad, kdy máme graf závislosti rychlosti na čase a chceme zjistit, v jakém čase se těleso pohybovalo rychlostí v. Najdeme si hodnotu v na ose zobrazující rychlost a sestrojíme kolmici. Ta se nám někde protne s grafem funkce. Tam sestrojíme další kolmici na osu času. Nová kolmice a osa času se protnou také v nějakém bodě, čímž budeme mít určen časový údaj.

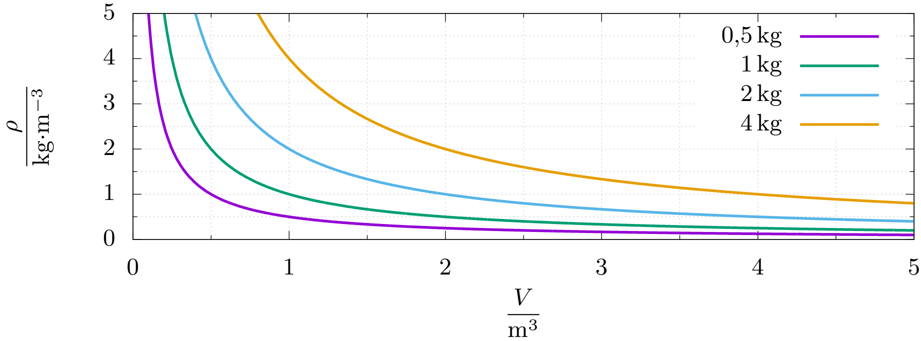
Grafy ve fyzice

Grafy jsou ve fyzice hojně využívány. Můžeme jimi popsat různé fyzikální principy a závislosti, což nám pomáhá v odvození dalších vztahů. Příkladem takových grafů, se kterými jste se mohli hojně setkat například ve Fyzikální olympiádě, jsou grafy *kinematické*. Ty nám popisují kinematiku tělesa, neboli to, jak se těleso pohybuje. Příkladem takového grafu je graf na obr. 3, kde máme vyobrazené konstantní zrychlení, které vede k lineárnímu růstu rychlosti (grafem je přímka) a kvadratickému růstu dráhy (grafem je parabola).



Obr. 3: Vyobrazení konstantního zrychlení a veličin na něm závislých.

Ale kinematika není jediným oborem, kde se grafy využívají. Na ukázku zde uvádíme porovnání závislostí hustoty na objemu pro různě hmotná tělesa (obr. 4). Vzhledem k nepřímé úměře mezi hustotou a objemem je grafem křivka zvaná hyperbola. Na tomto příkladu vidíme, že grafy zdaleka nemusejí ukazovat pouze funkce závislé na čase.



Obr. 4: Závislost hustoty tělesa na objemu pro čtyři různé hmotnosti. Všimněte si, že grafy nemusíme vytvářet pouze k záznamu časových průběhů.

Plocha pod křivkou

Jistě jste si někdy všimli, že se ve fyzice vyskytuje hodně vztahů, kde se násobí. Vezměme si například závislost rychlosti na čase při zrychleném pohybu:

$$v = at\,.$$

V našem příkladu předpokládejme, že je zrychlení konstantní. Když jej vyneseme do grafu, zjistíme, že se nám zde objevuje obdélník se stranami a a t. Předešlý vztah nám proto vlastně jen počítá obsah pod křivkou (v tomto případě obsah obdélníku), který je roven rychlosti.

Podobnou ukázkou důležitosti plochy pod křivkou je počítání dráhy z rychlosti zrychleného pohybu. Při vynesení rychlosti do grafu dostáváme pravoúhlý trojúhelník o stranách v a t. Pro takový trojúhelník platí vzoreček pro obsah

$$S = \frac{1}{2}ab \quad \Rightarrow \quad s = \frac{1}{2}vt\,.$$

Tím jsme opět ukázali, že fyzikální veličiny můžeme počítat pomocí plochy pod grafem, což se nám hodí vědět například tehdy, kdy je rychlost či jiná veličina nelineární, ale plocha v jejím grafu může být vyjádřena pomocí součtu ploch jednoduchých geometrických útvarů (obdélníků, trojúhelníků atd.).[4]

Plochy pod grafem hrají rovněž roli při určování *průměrných veličin*. Jako příklad si můžeme vzít opět rychlost. Víme, že pokud chceme zjistit průměrnou rychlost tělesa, nemůžeme

[4] Těm, které by toto téma více zajímalo, doporučujeme podívat se blíže na pojem *integrálu*. Tomuto tématu jsme se zároveň věnovali v páté úloze páté série devátého ročníku: `https://vyfuk.mff.cuni.cz/_media/ulohy/r09/s5/priklad5-5.pdf`.

jednoduše numericky zprůměrovat jednotlivé rychlosti, ale musíme si vyjádřit celkovou dráhu a tu podělit celkovým časem. Děláme tím vlastně to, že zjišťujeme plochu pod grafem rychlosti (což je dráha), a následně ji dělíme časem, čímž dostáváme průměrnou hodnotu rychlosti.[5]

Jednoduše řečeno: průměrnou hodnotu závislé veličiny na nějakém intervalu veličiny nezávislé určíme jako plochu pod grafem podělenou délkou samotného intervalu. Dostaneme tak hodnotu, kterou by měla závislá veličina mít, pokud by byla konstantní a zároveň měla na stejně dlouhém intervalu utvořit stejnou plochu pod grafem.

Vytváření grafů

Velkou součástí práce s grafy je jejich samotné vytváření a zakreslování. To můžeme provádět dvěma hlavními způsoby: ručně, anebo pomocí počítače. Každá metoda má své výhody i nevýhody, a proto se na různé možnosti podívejme.

Ručně

S metodou ručního zakreslování grafů jste se již s velkou pravděpodobností setkali ve škole nebo na Fyzikální olympiádě. Není na ni totiž potřeba nic více než tužka, papír (ideálně čtverečkovaný) a pravítko nebo trojúhelník.

Postup kreslení je také poměrně přímočarý: zvolíme si vhodné měřítko a meze, nakreslíme osy na sebe kolmé, vyneseme na ně stupnici a do grafu vyneseme data.

Výhodou takovéto metody je, že je poměrně jednoduchá na naučení a víme přesně, co děláme. Také máme plnou kontrolu nad tím, jak graf vypadá, protože se jedná o papír, kam můžeme cokoliv nakreslit či dopsat (např. když chceme napsat popisky hodnot, vyznačit na grafu specifické místo atd.).

Výhoda této metody je ovšem zároveň i její nevýhodou. Sice máme naprostou kontrolu nad tím, co v grafu je a není, ale musíme jej celý kreslit ručně. To často není takový problém, pokud se jedná o vynesení pouze pár bodů, ovšem začíná být komplikací v momentu, kdy potřebujeme vynést nelineární funkci, vynést stovky datových bodů či proložit body jinou funkcí.

Tabulkové procesory

Nyní se již dostáváme k vytváření grafů pomocí počítačů a nezačneme ničím jiným než tabulkovými procesory, jakými jsou například Excel, LibreOffice Calc, OpenOffice Calc a další. V nich máme možnost data spravovat v jedné velké tabulce a poté z nich vytvářet grafy.

Vytvoření jednoduchého grafu zde nezabere moc času, vzhledem k tomu, že ve většině případů stačí pouze vybrat data a položku vložení grafu. Grafy můžeme libovolně zvětšovat, měnit barvy, popisky atd. a celkově si vyhrát s tím, aby výsledek vypadal dobře.

Všechny možnosti a nastavení jsou rozřazeny v různých menu a podmenu, takže je občas potřeba déle hledat, než najdeme to jedno nastavení, které potřebujeme změnit, a mnohdy narazíme i na to, že danou věc ani moc jednoduše nezměníme.

Obecně řečeno, vytváření grafů v tabulkových procesorech je dobré pro rychlé vynesení několika datových bodů, ovšem už se moc nehodí pro kreslení funkcí či jednoduché vytváření elegantních vědeckých grafů.

---

[5] V tomto případě také můžeme říct, že počítáme tzv. *vážený průměr*.

Gnuplot

Gnuplot je program pro vykreslování funkcí a dat ve 2D a 3D prostoru, založený na příkazové řádce. To znamená, že zde nenajdete pěkná tlačítka pro nastavení všeho možného, ale všechny vlastnosti grafů definujete pomocí příkazů. To je také důvod, proč se někteří takovýmto programům vyhýbají.

Gnuplot je zpočátku těžký na naučení, pokud se s příkazovou řádkou setkáváte poprvé. Ovšem jakmile si ho člověk trochu osvojí, je v něm schopný dělat kvalitnější grafy[6] za kratší čas, protože napsat pár příkazů, které máte zaryty pod kůží (nebo uloženy na disku), vám zabere méně času než nastavovat jednotlivé vlastnosti v tabulkových procesorech.

Pokud byste chtěli Gnuplot vyzkoušet, můžete začít třeba s tímto příkladem:

```
set xrange [0:10] # Nastavení rozmezí dat na ose x od 0 do 10
set yrange [0:*] # Nastavení rozmezí dat na ose x od 0

set xlabel "Čas (s)" # Nastavení popisku pro osu x
set ylabel "Dráha (m)" # Nastavení popisku pro osu y

rychlost = 4  # Definice proměnné rychlosti

plot rychlost*x title "Graf rychlosti" # Vykreslení samotného grafu s popiskem
```

a pokud vás zajímá více, můžete si podrobnější návod na Gnuplot přečíst na našem webu.[7]

Další metody

Samozřejmě, již zmíněné programy nejsou jedinou cestou, jak vytvářet grafy. Dalšími jsou například programovací jazyky *Python* a *R*, z nichž Python má velkou spoustu knihoven zaměřených na zpracování dat a kreslení grafů (nejoblíbenější se nazývá *Matplotlib*) a R je pro tento účel přímo navrženo, proto je oblíbené ve statistických aplikacích (ekonomie, ekologie, sociologie a další).

Podobný Gnuplotu je také *MetaPost* (či novější *Asymptote*), který využívá textového vstupu pro kreslení obrázků (často odborných ilustrací). Oproti tomu na pomezí tabulkových procesorů a Gnuplotu je (placený) software *Origin*, vybavený klasickým rozhraním v podobě oken (i když plně podporuje i některé zde zmíněné jazyky).

V neposlední řadě nesmíme zapomenout na *Wolfram Mathematicu* a velmi zjednodušenou výpočetně-vyhledávací webovou aplikaci *Wolfram Alpha*, pomocí kterých si můžete (i online) vykreslit funkci či počítat řadu jiných úloh.

Nakonec několik zásad

Když budete sami vytvářet grafy, musíte se zamyslet nad všemi jejich výše zmíněnými prvky. Zde předkládáme několik zásad převzatých z našeho textu Hokus Pokus.[8]

- Musí být jasné, které ose odpovídá která veličina a také v jakých jednotkách je každá veličina vyjádřena. Nezávislá proměnná je vždy na vodorovné ose.

[6] Pokud byste chtěli představu, jak takové grafy vypadají, můžete se podívat do řešení úloh Výfuku či na grafy v tomto Výfučtení.

[7] `https://vyfuk.mff.cuni.cz/rady_a_tipy/gnuplot`

[8] `https://vyfuk.mff.cuni.cz/rady_a_tipy/hokus_pokus`

- Popisky os grafu by měly být jednoznačně umístěny tak, aby bylo zřejmé, ke které z os se vztahují (například v levém dolním rohu obvykle začínají obě osy, proto jde o nevhodné místo).

- Hustota dílků (a jejich popisků) by měla být přiměřená, stejně jako zvolené meze. Pokud například zapisujeme hodnoty po stále stejném narůstajícím množství nezávislé proměnné, není často důvod na vodorovné ose dělat více značek, než je bodů grafu. Stejně tak by bylo nevhodné osy prodloužit výrazně přes meze, které v závislé i nezávislé proměnné zabírají vložené hodnoty. Měly by rovnoměrně pokrývat plochu grafu a nevytvářet nevyužité místo či tak zbytečně zhoršovat čitelnost.

- Datové body či křivka by měly být správně výrazné, nejlépe barevně odlišené.

- Experimentálně naměřená data nespojujeme plnou čarou. Značila by totiž závislost, která se v datech nevyskytuje.

Adam Krška
`adam@vyfuk.mff.cuni.cz`

Daniel Slezák
`dans@vyfuk.mff.cuni.cz`

Korespondenční seminář Výfuk je organizován studenty a přáteli MFF UK. Je zastřešen Oddělením propagace a mediální komunikace MFF UK a podporován Katedrou didaktiky fyziky MFF UK, jejími zaměstnanci a Jednotou českých matematiků a fyziků.