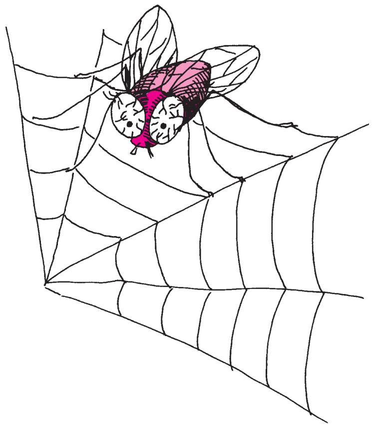
O zdolności rozdzielczej

Marek W. GUTOWSKI



Termin **zdolność rozdzielcza** pojawia się w fizyce dość często. Gdyby kogoś zapytać, co on właściwie oznacza, to najprawdopodobniej ten ktoś rozpocząłby wyjaśnienia od zjawisk związanych z optyką. Hobbystom i osobom związanym zawodowo z przetwarzaniem obrazów: fotografom, kamerzystom, astronomom – termin ten jest z całą pewnością dobrze znany. Dla tych, którzy jeszcze nie znają tego pojęcia, również i w tym artykule rozpoczniemy od zjawisk optycznych. Prawdziwym naszym celem jest jednak zupełnie co innego: pragniemy mianowicie omówić związki między zdolnością rozdzielczą, zwaną czasem krótko **rozdzielczością**, a pomiarami wielkości fizycznych bardzo różnej natury, nie tylko związanych z optyką.

Jak to jest w optyce?

Czytelnik nie wyposażony w skomplikowany sprzęt optyczny zapewne ogląda czasem programy telewizyjne. Niektóre stacje nadawcze wciąż jeszcze mają zwyczaj nadawać, przed lub po właściwym programie, tzw. *obraz kontrolny*. Nie chodzi nam o ten ze skalą barw, ale o ten bardziej skomplikowany, przedstawiający rozmaite linie, okręgi i inne figury geometryczne. Służy on do regulacji odbiornika oraz do oceny jakości obrazu – a to zapewne interesuje potencjalnego nabywcę, względnie posiadacza.

Na obrazie kontrolnym widzimy m.in. pęki linii prostych zbiegających się w jednym punkcie. W tych miejscach, gdzie linie są dość daleko, nietrudno jest policzyć, *ile* ich jest. W miarę, gdy linie są coraz bliżej, ich rozróżnienie staje się coraz trudniejsze i w końcu zlewają się w *jedną* grubą krechę. Odbiornik jest tym lepszy, im więcej linii leżących blisko widzimy wyraźnie jako oddzielne obiekty. Nic dziwnego, że cechę tę charakteryzuje się ilościowo, czyli w sposób mierzalny, poprzez podanie, ile linii na centymetr jest jeszcze dobrze widocznych. Podobnie można porównywać jakość rozmaitych przyrządów optycznych, poczynając od pojedynczej soczewki (lupy) czy zwierciadła, poprzez mikroskopy, lornetki, dalmierze aż do wielkich lunet i teleskopów używanych w obserwacjach astronomicznych.

Zwłaszcza te ostatnie przyrządy zostały kiedyś starannie „przeegzaminowane". Gwiazdy są tak daleko od nas, że można je uważać za punktowe źródła światła. Jeśli więc któraś z nich jest układem podwójnym, to powinno być możliwe zarejestrowanie oddzielnych obrazów obu składników, jeśli tylko luneta czy teleskop, ma dostatecznie duże powiększenie. Tymczasem, nawet jeśli przyrząd zostanie wykonany niezwykle starannie, to i tak nie rozróżnimy za jego pomocą dwóch obiektów, których obrazy będą położone bliżej niż $\lambda/1{,}22$, gdzie λ jest długością fali, na której prowadzimy obserwacje. Wiąże się to, oczywiście, z falową naturą światła. Wielkość $R = 1{,}22/\lambda$ nazywa się z tego powodu (maksymalną teoretyczną) *zdolnością rozdzielczą* instrumentu optycznego. Jak widzimy, wyraża się ona w tych samych jednostkach co parametr opisujący jakość obrazu telewizyjnego, tyle że dla teleskopu osiąga on wartość $\sim 10^3$/mm (światło widzialne, o długości fali $\lambda \sim 1\mu\mathrm{m}$), podczas gdy w najlepszych telewizorach jest ponad tysiąc razy mniejszy. Specyficzny przyrząd optyczny, jakim jest ludzkie oko, jest pod tym względem również daleki od ideału: skoro widzimy na wydruku z drukarki o rozdzielczości 300 dpi (ang. *dots per inch*, kropek na cal), literę l jako krótki odcinek, a nie zbiór oddzielnych kropek, to oznacza, że w najlepszym razie jesteśmy w stanie odróżnić $\sim$12 linii/mm.

Rozdzielczość spektrometrów

Jak działa spektrometr? Światło z badanego obiektu (wytworzone przez obiekt albo odbite od niego, albo przechodzące) dociera do instrumentu przez wąską *szczelinę wejściową*, po czym zostaje rozszczepione za pomocą pryzmatu albo siatki dyfrakcyjnej. Powstałe w ten sposób *widmo* można obserwować gołym okiem na odpowiednim ekranie. Zamiast oka wolimy użyć odpowiedniego *detektora*, który umieszczamy w coraz to innym miejscu widma. Detektor „obserwuje" widmo przez *szczelinę wyjściową* przyrządu. Typowe warunki pomiaru są takie, że szczelina wejściowa jest możliwie wąska – w ten sposób badany obiekt można traktować jak punktowe źródło światła. Szerokość szczeliny wyjściowej daje się regulować w szerokich granicach. Eksperymentator stara się korzystać z możliwie wąskiej szczeliny wyjściowej, nie może jej jednak zmniejszać całkiem dowolnie, gdyż węższa szczelina oznacza jednocześnie mniejszą ilość światła docierającą do detektora, a więc gorszy stosunek sygnału do szumu.

Od czego zależy sygnał produkowany przez detektor? Nie wchodząc w szczegóły, możemy powiedzieć, że sygnał użyteczny (wszelkie szumy pomijamy w naszych rozważaniach) powinien być proporcjonalny do liczby fotonów docierających do detektora w jednostce czasu, powiedzmy, w ciągu sekundy. Możemy to wyrazić następującym wzorem:

$$S(E_0) = C \int_{E_0-\frac{R}{2}}^{E_0+\frac{R}{2}} I(E)\ dE, \tag{1}$$

gdzie $S(\cdot)$ to obserwowany sygnał z detektora, będący funkcją energii, C jest odpowiednią stałą proporcjonalności, a $I(E)$ to prawdziwe widmo. Granice całkowania odzwierciedlają fakt, że do detektora docierają i są rejestrowane jednocześnie fale elektromagnetyczne o różnych energiach zbliżonych do E_0. Tak więc do instrumentu dociera światło, którego widmo opisuje funkcja $I(E)$, natomiast my rejestrujemy je jako funkcję $S(E)$. To samo możemy zapisać inaczej:

$$S(E_0) = C \int_{-\infty}^{+\infty} I(E)\ g(E - E_0)\ dE, \tag{2}$$

gdzie wprowadziliśmy funkcję $g(\cdot)$, którą nazywa się *funkcją aparaturową* albo *okienkową*. Jest ona różna od zera (a ściślej: dodatnia) tylko wtedy, gdy

$$|E - E_0| \leqslant \frac{R}{2}.$$

W ogólności, funkcja $g(\cdot)$ może zależeć od energii (detektor ma różną czułość w rozmaitych fragmentach widma), my jednak zaniedbajmy tę zależność. Myślmy raczej o g jako o funkcji, która ma wartość 1 w okolicy E_0 oraz wartość 0 poza tym. Długość przedziału, w którym $g(\cdot) \neq 0$, to liczba R – zdolność rozdzielcza. Dlatego $g(\cdot)$ nazywamy *funkcją okienkową*: ma ona tę właściwość, że zamienia całkowanie względem całej osi rzeczywistej na całkowanie tylko względem tego obszaru widmowego, który aktualnie „widzi” nasz detektor. (Tym, którym „nie zgadzają się jednostki”, spieszymy z wyjaśnieniem, że stała C jest mianowana).

Forma zapisu (2) opisuje definicję operacji matematycznej zwanej *splotem* dwóch funkcji i oznaczanej zwykle znakiem „$*$”. Tak więc

$$S(E_0) = (I * g)(E_0), \tag{3}$$

co wyrażamy słowami: *widmo obserwowane jest splotem widma oryginalnego z funkcją okienkową*.

Funkcje aparaturowe prawdziwych spektrometrów raczej rzadko dają się przybliżać w tak drastycznie uproszczony sposób (dyfrakcja na wąskich szczelinach!), jednakże przy szerokich szczelinach wyjściowych, oraz do naszych celów, przybliżenie to jest zupełnie poprawne i wystarczające. Dla porządku wspomnimy jedynie, że do celów praktycznych najchętniej posługujemy się *znormalizowanymi* funkcjami okienkowymi, tj. przemnożonymi przez taką liczbę, aby prawdziwa była równość

$$\int_{-\infty}^{+\infty} g(x)\ dx = 1$$

(oczywiście, że tak naprawdę granice całkowania są skończone). Odpowiedni współczynnik skalujący zawsze przecież możemy zawrzeć w stałej C.

Tym, którzy nie boją się całkowania, proponujemy samodzielne przećwiczenie kilku prostych przypadków: jak powinno wyglądać widmo obserwowane, kiedy widmo oryginalne składa się – na początek – z pojedynczej linii widmowej o kształcie prostokąta (*pasmo*) lub trójkąta? Próby te z pewnością wykażą różnice między oryginałem a tym, co powinniśmy obserwować. Generalnie, widmo splecione z funkcją okienkową wygląda na „wygładzone” czy wręcz „rozmyte” w porównaniu do oryginału. Na kolejne pytania znamy odpowiedzi, ale wolimy, aby Czytelnik sam do nich doszedł:

∗ jak mają się do siebie położenia centrów linii widmowych przed i po splocie z funkcją okienkową?

∗ a amplitudy?

∗ i pola pod tymi krzywymi?

∗ jak wygląda splot, kiedy widmo oryginalne składa się z dwóch linii odległych wzajemnie mniej niż wynosi szerokość „okienka”?

Ci, którzy mają dostęp do pracowni optycznej, mogą sprawdzić doświadczalnie, jak zmienia się wygląd widma rejestrowanego przy różnych szerokościach szczeliny wyjściowej. Szczególnie dobrym obiektem do obserwacji powinny być wyładowania w tzw. rurkach Geislera wypełnionych rozmaitymi gazami pod zmniejszonym ciśnieniem.

Inne pomiary a operacja splotu

Przyroda wykonuje operację splotu o wiele częściej, niż mogłoby się wydawać. Damy tu przykład z miernictwa elektrycznego. Jeśli dołączymy woltomierz (wszystko jedno – tradycyjny wychyłowy czy cyfrowy) do źródła napięcia, które nie jest stałe w czasie, to co zaobserwujemy? Przyrząd pomiarowy ma pewną bezwładność, czysto mechaniczną (woltomierz wskazówkowy) lub wynikającą z obecności rozmaitych filtrów przeciwzakłóceniowych, niepożądanych pojemności itp. w przypadku woltomierza elektronicznego. W rezultacie, przyrząd nigdy nie wskazuje prawdziwej chwilowej wielkości mierzonego napięcia, z wyjątkiem sytuacji, gdy to napięcie od dawna ma ustabilizowaną wartość lub zmienia się niezwykle powoli. To, co widzimy, to coś w rodzaju uśrednionej historii, przy czym w tej średniej największą wagę mają zdarzenia najświeższe. Przyrząd stopniowo „zapomina”, jakie wartości napięcia występowały wcześniej. Szybkość tego zapominania można scharakteryzować tzw. stałą czasu τ – im jest ona mniejsza, tym zapominanie jest skuteczniejsze (szybsze). W tym przypadku napięcie wskazywane przez miernik jest splotem napięcia wejściowego z funkcją aparaturową postaci (czynnik skalujący pominięto)

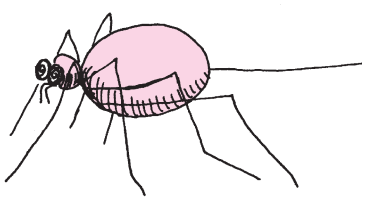
$$g(t) = \begin{cases} e^{t/\tau} & \text{dla } t < 0, \\ 0 & \text{gdy } t \geqslant 0. \end{cases} \tag{4}$$

Można to odczytać jako stwierdzenie, że wpływ na bieżące wskazania woltomierza ma cała historia zmian napięcia wejściowego; zdarzenia, które dopiero mają nastąpić – zgodnie ze zdrowym rozsądkiem – takiego wpływu nie mają.

Trochę pechowo się składa, że w tym przypadku zbiór, na którym funkcja okienkowa jest różna od zera, jest nieograniczony, gdyż jest to półprosta. Nie popełnimy jednak istotnego błędu, jeśli przyjmiemy, że $g(t) \equiv 0$ dla $t < -k\tau$, gdzie dla k przyjmuje się zwykle wartość pomiędzy 5 a 6. Czasowa zdolność rozdzielcza woltomierza, rzędu τ, uniemożliwi nam spostrzeżenie impulsów napięcia o *czasie trwania* mniejszym niż τ, praktycznie niezależnie od ich amplitudy. Z drugiej strony, impuls taki może się okazać wystarczający do tego, aby trwale uszkodzić woltomierz – nie dlatego, że jest krótkotrwały, lecz z powodu dużej amplitudy.

Jeszcze inny przykład to pomiary czegokolwiek w zależności od temperatury. Może to być np. objętość gazu (niedoskonałego) czy ciepło właściwe. Wykonując badania tego rodzaju, postępujemy najczęściej tak: ustalamy temperaturę układu, po czym wykonujemy odpowiedni pomiar. Potem zmieniamy temperaturę i wykonujemy następny pomiar. Czas trwania pomiaru nie jest zerowy; kto nam zaręczy, że podczas pomiaru temperatura była rzeczywiście stała? Z pewnością nieco się zmieniła, może tylko „płynęła” w jedną stronę, a może fluktuowała w jakichś granicach? Wynik pomiaru nie dotyczy zatem konkretnej, precyzyjnie określonej temperatury, jest raczej pewnym uśrednieniem badanej wielkości fizycznej w pewnym, miejmy nadzieję wąskim, przedziale temperatury. Znowu natykamy się na operację splotu, tyle że tym razem wszystko, co możemy powiedzieć o funkcji okienkowej, to jej „rozciągłość” na skali temperatury, czyli maksymalna *niepewność* określenia temperatury, nominalnie ustabilizowanej.

Nie sposób powstrzymać się od generalnej uwagi związanej z cyfrowymi przyrządami pomiarowymi. Ich działanie jest takie, że „muszą” one utożsamiać wyniki pomiarów, jeśli są one zbyt bliskie. Niepewność pomiarowa wnoszona przez mierniki cyfrowe jest więc co najmniej równa ich rozdzielczości, czyli jest na poziomie ostatniej wyświetlanej cyfry. Konkretna zasada działania może jednakże być powodem, że rzeczywista niepewność bywa nawet kilkakrotnie większa.



Na koniec zachowaliśmy dwie wiadomości: jedną dobrą i jedną złą. Zła jest taka, że nie istnieje prosty ogólny wzór na „odwrócenie” skutków operacji splotu, nawet gdy funkcja okienkowa jest doskonale znana.

Wiadomość dobra to to, że prawdziwe jest twierdzenie, które tu sformułujemy niezbyt precyzyjnie: jeśli możliwe są pomiary z coraz lepszą zdolnością rozdzielczą, to otrzymywane wyniki coraz lepiej odtwarzają badany oryginalny sygnał, przy czym zbieżność ta jest jednostajna. Ten, kto nie bardzo wie, o co chodzi w końcówce zdania – niech się nie martwi. To tylko dodatkowy argument za tym, że wiadomość jest istotnie dobra.

W przykładzie ze spektrometrem oznacza to pomiary z jak najmniejszymi szczelinami wejściową i wyjściową, w przypadku woltomierza – należy używać przyrządu o możliwie małej stałej czasu. Matematyk byłby zadowolony. Dla fizyka jednak oznacza to kłopoty: coraz gorszy stosunek sygnału badanego do szumów, wzrost podatności na zakłócenia, wzrost kosztów aparatury i tym podobne plagi. A to już temat na całkiem inne opowiadanie.

Warto jeszcze zapamiętać: przy prezentacji wyników pomiarów, w których w jakiś sposób zaangażowana jest operacja splotu, niezbędne jest podawanie zdolności rozdzielczej oraz niepewności wartości parametrów, które miały być ustalone. Bez tych informacji porównywanie wyników otrzymanych przy użyciu różnych przyrządów, albo przez różne grupy badawcze, zwyczajnie nie ma sensu.

Rozwiązanie zadania M 1006. Ponumerujmy wierzchołki n-kąta foremnego liczbami $0, 1, 2, \ldots, n-1$ zgodnie z ruchem wskazówek zegara. Zauważmy, że odcinek o końcach a_1, b_1 jest równoległy do odcinka o końcach a_2, b_2 wtedy i tylko wtedy, gdy

$$a_1 + b_1 \equiv a_2 + b_2 \pmod{n}. \tag{$*$}$$

Z zasady szufladkowej wynika, że dla co najmniej dwóch odcinków zachodzi $(*)$.