



Politechnika Łódzka

Instytut Fizyki

Laboratorium elektroniki

Ćwiczenie E02IN

Diody

Spis treści:

1. Cel ćwiczenia.....	3
2. Zagrożenia	3
3. Wprowadzenie teoretyczne.....	3
4. Dostępna aparatura	5
4.1. Moduł doświadczalny.....	5
4.2. Multimetry.....	6
4.3. Zasilacz laboratoryjny	6
5. Przebieg doświadczenia.....	6
6. Wskazówki do raportu.....	7
7. Literatura	9
7.1. Literatura podstawowa	9
7.2. Literatura uzupełniająca	9
8. Aneksy	10
A1. Metoda wyznaczania napięcia Zenera.....	10
A2. Metoda najmniejszych kwadratów	11

Przed zapoznaniem się z instrukcją i przystąpieniem do wykonywania ćwiczenia należy opanować następujący materiał teoretyczny:

1. Półprzewodniki samoistne i domieszkowane. [1], [2], [3].
2. Złącze p-n. Złącze metal-półprzewodnik (m-s). [2], [3].
3. Przebieg Zenera. Przebieg lawinowe. [1], [2], [3].

1. Cel ćwiczenia

Celem ćwiczenia jest pomiar charakterystyk prądowo-napięciowych wybranych diod półprzewodnikowych oraz doskonalenie techniki pomiarowej. Zakres ćwiczenia obejmuje diodę krzemową oraz diodę Zenera.

2. Zagrożenia

Rodzaj	Brak	Małe	Średnie	Duże
zagrożenie elektryczne		+		
zagrożenie optyczne	+			
zagrożenie mechaniczne (w tym akustyczne, hałas)	+			
zagrożenie polem elektro-magnetycznym (poza widmem optycznym)	+			
zagrożenie biologiczne	+			
zagrożenie radioaktywne (jonizujące)	+			
zagrożenie chemiczne	+			
zagrożenie termiczne (w tym wybuch i pożar)	+			

Przewody z wtykami bananowymi są przeznaczone wyłącznie do użytku w obwodach niskiego napięcia – nie wolno podłączać ich do gniazda sieci zasilającej 230 V.

3. Wprowadzenie teoretyczne

Dioda półprzewodnikowa to zwykle niewielka struktura półprzewodnikowa, zawierająca odpowiednio obudowane złącze p-n wraz z dołączonymi dwoma wyprowadzeniami. W diodach Schottky'ego zamiast złącza p-n wykorzystane jest złącze m-s (metal-półprzewodnik). W praktyce używa się wielu odmian konstrukcyjno-technologicznych diod, niekiedy o bardzo różnym wyglądzie zewnętrznym i różnych parametrach elektrycznych. Charakterystyka prądowo-napięciowa każdej diody wykazuje częściowe podobieństwo do teoretycznej charakterystyki idealnego złącza p-n lub m-s, opisanej wzorem:

$$I = I_s \left[\exp\left(\frac{eU}{mkT}\right) - 1 \right], \quad (1)$$

gdzie: I – natężenie prądu płynącego przez złącze, I_s – natężenie wstecznego prądu nasycenia,
 U – napięcie na złączu (wartościom dodatnim odpowiada kierunek przewodzenia),
 T – temperatura złącza w skali Kelvina, k – stała Boltzmanna ($k = 1,38065 \cdot 10^{-23}$ J/K),
 e – ładunek elektronu ($e = 1,602176 \cdot 10^{-19}$ C),
 m – bezwymiarowy parametr, który przyjmuje wartości z przedziału od 1 do 2.

Analiza rzeczywistego złącza diody powinna uwzględniać takie parametry i zjawiska jak: rezystancja szeregową złącza, rezystancja upływu, procesy generacji i rekombinacji nośników oraz przebicie złącza.

W rzeczywistym złączu p-n poza spadkiem napięcia na obszarze przejściowym należy uwzględnić spadki napięć na rezystancjach obszarów p oraz n złącza a także na rezystancjach styków i doprowadzeń. Sumę tych rezystancji nazywamy rezystancją szeregową. Wpływ rezystancji szeregowej jest widoczny przy dużych prądach płynących przez diodę, czyli dla kierunku przewodzenia i zakresu przebicia złącza.

Rezystancja upływu jest rezystancją równoległą do rezystancji obszaru złącza. Rezystancja ta jest wynikiem występowania zjawisk powierzchniowych i defektów w regularnej strukturze złącza. Prąd płynący przez rezystancję upływu dodaje się do prądu płynącego przez idealne złącze p-n, powiększając jego wartość. Z uwagi na to, że wartość rezystancji upływu jest zazwyczaj bardzo duża, jej wpływ na charakterystykę diody uwidacznia się podczas polaryzacji w kierunku zaporowym.

W obszarze złącza p-n mogą zachodzić zjawiska generacji i rekombinacji nośników ładunku. Generacja odgrywa istotną rolę podczas polaryzacji zaporowej, natomiast rekombinacja podczas polaryzacji diody w kierunku przewodzenia. Rezultatem tych dwu zjawisk jest przepływ dodatkowych składników prądu: generacyjnego podczas polaryzacji zaporowej oraz rekombinacyjnego podczas polaryzacji w kierunku przewodzenia. Prąd generacyjny jest o kilka rzędów większy od wstecznego prądu nasycenia i decyduje o wypadkowym prądzie wstecznym w diodach krzemowych, natomiast jest on zaniedbywalny w diodach germanowych.

Po przekroczeniu pewnej wartości napięcia polaryzującego złącze p-n w kierunku zaporowym prąd wsteczny złącza zaczyna szybko wzrastać. Ten wzrost może być wywołany bądź jonizacją elektrostatyczną (przebiecie Zenera), bądź jonizacją zderzeniową (przebiecie lawinowe). Pierwszy z tych efektów występuje dla złącz silnie domieszkowanych i w konsekwencji dla niskich napięć (w przypadku krzemu poniżej 5 V). Drugi z tych efektów jest typowy dla złącz słabiej domieszkowanych (dla złącz krzemowych przebiecie lawinowe następuje dla napięć większych od ok. 7 V). Możliwe jest też jednoczesne wystąpienie obydwu mechanizmów, co w przypadku krzemu ma miejsce jeśli przebiecie występuje przy napięciu bliskim 6 V.

Właściwości elektryczne złączy metal-półprzewodnik zależą od doboru materiałów tworzących złącze. W niektórych przypadkach złącze m-s dobrze przewodzi prąd w obu kierunkach przy stosunkowo niskich napięciach. Takie złącze nazywa się kontaktem omowym. W innych przypadkach złącze m-s w jednym kierunku przewodzi duże prądy przy niskich napięciach polaryzujących złącze, a dla odwrotnego kierunku polaryzacji prądy są znikomo małe. Takie złącze m-s nazywa się złączem prostującym lub złączem (diodą) Schottky'ego. Spadek napięcia na diodzie Schottky'ego w stanie przewodzenia jest zwykle mniejszy niż w diodach krzemowych (dla diod krzemowych spadek napięcia przyjmuje wartości $\sim 0,7$ V), ale większy niż w germanowych (dla diod germanowych spadek napięcia ma wartość $\sim 0,3$ V). Wadą diod Schottky'ego jest stosunkowo niska wytrzymałość na przebicie w kierunku zaporowym, która zwykle wynosi $30\text{ V} \div 50\text{ V}$ (największe spotykane wartości nie przekraczają 100 V).

Parametry diod ulegają zmianie przy zmianach temperatury. Jak wynika z analizy wzoru (1), dla kierunku przewodzenia, przy wzroście temperatury należy oczekiwać malenia spadku napięcia na złączu diody. Przy temperaturach użytkowych z przedziału $250\text{ K} \div 350\text{ K}$ ($-20\text{ }^\circ\text{C} \div 80\text{ }^\circ\text{C}$) wpływ jest następujący:

- spadek napięcia na złączu w stanie przewodzenia maleje o ok. 20 mV przy wzroście temperatury o 10 K

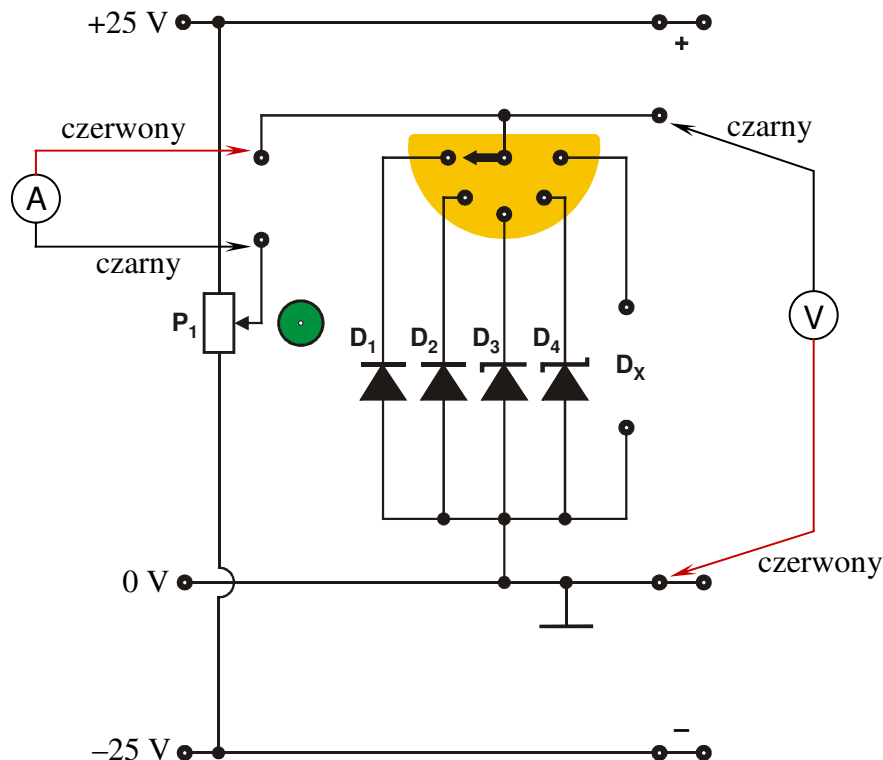
- natężenie wstecznego prądu nasycenia wzrasta blisko dwukrotnie przy wzroście temperatury o 10 K
- przy wzroście temperatury maleje napięcie przebicia Zenera, a intensywność zmian zależy od wartości napięcia przebicia.

Należy pamiętać, że wzrost temperatury złącza diody najczęściej jest wywołany przepływem zbyt dużego prądu przewodzenia lub prądu przebicia (Zenera lub lawinowego). Pomiary należy zatem wykonywać w miarę możliwości sprawnie, aby nie dopuszczać do nadmiernego nagrzewania się badanej diody.

4. Dostępna aparatura

4.1. Moduł doświadczalny

Panel czołowy modułu doświadczalnego przedstawiono na rys. 1. Multymetry podłączono do układu w taki sposób, aby dla kierunku przewodzenia wskazywane przez nie wartości były dodatnie, a dla kierunku zaporowego ujemne.



Rys. 1. Panel czołowy modułu doświadczalnego do pomiaru charakterystyk prądowo-napięciowych czterech diod (D_1 – dioda germanowa; D_2 – dioda krzemowa; D_3 – dioda Zenera; D_4 – dioda Schottky’ego). Na schemacie podano wartości napięć zasilających układ oraz sposób podłączenia multimetrów z zaznaczeniem kolorów przewodów pomiarowych.

4.2. Multimetry

Pomiary napięć i natężeń prądów w module doświadczalnym wykonuje się przy użyciu multimetrów cyfrowych KT890, M-3800, M-4650, UT-804 lub Protek 506 (na danym stanowisku pomiarowym dostępne są tylko wybrane modele) [4]. Przed przystąpieniem do pomiarów należy sprawdzić czy przełącznikiem każdego z multimetrów cyfrowych został wybrany właściwy przyrząd pomiarowy oraz czy końcówki pomiarowe przyłączone zostały do właściwych zacisków wejściowych. Jeden z multimetrów powinien pełnić rolę amperomierza a drugi woltomierza. Należy jednocześnie pamiętać o wyborze odpowiedniego zakresu pomiarowego woltomierza (20 V DC albo 40 V DC w zależności od użytego multimetru) i amperomierza dla kierunku przewodzenia (w przypadku diod D_1 , D_2 oraz D_4 proszę użyć zakresu 200 mA DC albo 400 mA DC w zależności od użytego multimetru). Dla kierunku zaporowego należy użyć najmniejszego dostępnego zakresu, tzn. 20 μ A, 200 μ A albo 400 μ A DC w zależności od użytego multimetru, przy rozdzielczości odczytu nie gorszej niż 0,01 μ A. W przypadku diody Zenera D_3 należy użyć zakresu 200 mA DC albo 400 mA DC niezależnie od kierunku.

Uwaga: użycie multimetrów KT890 oraz Protek 506 do pomiarów prądu wstecznego nie jest zalecane, ponieważ ich rozdzielczość 1 μ A albo 0,1 μ A na najniższym dostępnym zakresie może być niewystarczająca podczas pomiarów prądów wstecznych diody D_2 rzędu kilku dziesiątych mikroampera.

Z rys. 1 wynika, że charakterystyka każdej diody jest mierzona w układzie dokładnego pomiaru napięcia. Wynika to z właściwości wchodzących w skład stanowiska laboratoryjnego przyrządów pomiarowych jak też z faktu, że w metodzie dokładnego pomiaru natężenia prądu rezystancja wewnętrzna amperomierza dodawałaby się do rezystancji szeregowej diody utrudniając przeprowadzenie precyzyjnego pomiaru.

4.3. Zasilacz laboratoryjny

Zasilanie modułu doświadczalnego zrealizowano przy użyciu zasilacza laboratoryjnego SIGLENT model SPD3303D [4]. Przed rozpoczęciem pomiarów zasilacz stabilizowany SPD3303D należy połączyć z badanym układem, pamiętając o wyborze szeregowego trybu pracy zasilacza. Po włączeniu zasilacza należy ustawić symetryczne napięcie zasilające na ± 25 V.

5. Przebieg doświadczenia

Należy wykonać możliwie rzetelne pomiary (kilkadziesiąt punktów pomiarowych) charakterystyk poszczególnych diod. W ćwiczeniu mierzy się statyczne charakterystyki prądowo-napięciowe metodą „punkt po punkcie”, która polega na zmianie żądanego napięcia (za pomocą potencjometru wieloobrotowego P_1) i na odczycie – po każdej zmianie – wartości przyłożonego napięcia U oraz natężenia I płynącego przez diodę prądu. Gęstość punktów pomiarowych należy dobierać odpowiednio do szybkości zmian natężenia prądu przepływającego przez diodę.

1. Połączyć układ pomiarowy do badania diod zgodnie ze schematem przedstawionym na rys. 1. Przy pomocy zwory wybrać drugą diodę zestawu (D_2). Zasilacz połączyć z układem za pomocą trzech różnokolorowych przewodów (np. czerwony: „plus”, czarny: „minus” i niebieski: „zero”).

UWAGA: pamiętać o wyborze szeregowego trybu pracy zasilacza oraz o wstępnym ustawieniu napięcia wyjściowego zasilacza na wartość minimalną/zerową.

- Dokonać odpowiedniego wyboru zakresów pomiarowych multimetrów (patrz rozdział 4.2. **Multimetry**).
 - Poprosić prowadzącego zajęcia o sprawdzenie poprawności połączeń oraz o wyrażenie zgody na załączenie napięcia na wyjściach zasilacza.**
 - Ustawić symetryczne napięcie zasilające na ± 25 V. Sprawdzić czy świecą się kontrolki LED w liniach zasilających układ (czerwona: +25 V; zielona: -25 V).
 - Wykonać pomiary charakterystyki prądowo-napięciowej diody. Charakterystyka powinna się składać z minimum 40 punktów pomiarowych ($20 \div 25$ dla kierunku przewodzenia oraz $15 \div 20$ dla kierunku zaporowego). Gęstość punktów pomiarowych należy dobierać odpowiednio do dynamiki zmian natężenia prądu płynącego przez diodę. Potencjometr wieloobrotowy P_1 umożliwia płynną i precyzyjną zmianę napięcia polaryzującego diodę. Skrajnemu lewemu położeniu potencjometru P_1 odpowiada maksymalne napięcie zaporowe (ok. -9 V). Skrajnemu prawemu położeniu potencjometru P_1 odpowiada maksymalny spadek napięcia na diodzie w kierunku przewodzenia (wartość maksymalnego spadku napięcia jest zależna od diody). Wskazane jest, aby właściwe pomiary dla kierunku zaporowego jak i przewodzenia rozpoczynać od ustawienia potencjometrem P_1 napięcia polaryzującego 0 V. Pomiary należy wykonywać w miarę możliwości sprawnie, aby nie dopuszczać do nagrzewania się diody (szczególnie dla części charakterystyki w kierunku przewodzenia).
- UWAGA:** układ regulacji napięcia polaryzującego diodę ma wbudowane ograniczenie prądowe (75 mA), aby zabezpieczyć badane diody przed uszkodzeniem.
- Pamiętać o zmianie zakresu amperomierza przy przejściu od kierunku zaporowego do kierunku przewodzenia. Podczas zmiany zakresu wyłączać multimetr.**
 - Wyniki notować w tabeli (przy czym ΔU , ΔI oznaczają odpowiednie niepewności graniczne multimetrów użytych podczas pomiarów)

Symbol diody	kierunek zaporowy				kierunek przewodzenia				
	Lp.	U [V]	ΔU [V]	I [μ A]/[mA]	ΔI [μ A]/[mA]	U [V]	ΔU [V]	I [mA]	ΔI [mA]
D ₂	1.								
	2.								

- Powtórzyć wymienione powyżej czynności dla diody Zenera D₃ wchodzącej w skład zestawu.

6. Wskazówki do raportu

Raport powinien zawierać:

- Stronę tytułową.
- Sformułowanie celu ćwiczenia.
- Podstawowe definicje, wzory i opisy użytych symboli i oznaczeń.
- Schematy układów pomiarowych.

W sprawozdaniu należy umieścić schematy tylko takich układów, które były rzeczywiście zestawiane w trakcie wykonywania pomiarów. Każdy schemat powinien być opatrzony numerem kolejnym i zatytułowany. Wszystkie elementy pokazane na schemacie muszą być jednoznacznie opisane i oznaczone za pomocą powszechnie stosowanej symboliki.

5. Wykaz aparatury (nr inwentarzowy, typ, wykorzystywane nastawy i zakresy).
W wykazie aparatury należy jednoznacznie opisać używaną aparaturę pomiarową poprzez podanie numeru inwentarzowego, typu itd. Nadane poszczególnym przyrządom oznaczenia należy konsekwentnie stosować na wszystkich schematach i w opisach.
6. Stabelaryzowane wyniki pomiarów.
Jako wyniki pomiarów należy zamieścić tabele ze zmierzonymi wartościami. Każda tabela powinna posiadać swój numer kolejny i tytuł.
7. Wykresy.
Wszystkie wykresy wykonane na podstawie przeprowadzonych pomiarów powinny mieć numery porządkowe oraz podpisy zawierające informację o tym co dany wykres przedstawia.
Wykorzystując otrzymane dane pomiarowe należy wykonać wykresy przedstawiające charakterystyki prądowo-napięciowe obu zbadanych diod w skali liniowej. Charakterystykę diody D_2 należy zamieścić na dwu wykresach. Na jednym z nich należy przedstawić charakterystykę dla kierunku przewodzenia natomiast na drugim dla kierunku zaporowego. Charakterystykę diody Zenera D_3 należy wykonać na oddzielnym wykresie z zaznaczeniem napięcia Zenera (patrz p.A1 aneksu do niniejszej instrukcji).
Dodatkowo dla diody D_2 należy wykreślić charakterystykę w kierunku przewodzenia w układzie półlogarytmicznym $\ln(I) \sim U$ (patrz wzór (3) oraz rozdz. IV.3.3. pozycji [6]). Na podstawie otrzymanego dla diody D_2 wykresu określić zakres stałego nachylenia charakterystyki i oprócz punktów pomiarowych nanieść na charakterystyce także linię trendu prostoliniowego.
8. Analizę wyników.
Dla określonego w punkcie 7 zakresu stałego nachylenia charakterystyki diody D_2 wyznaczyć wartość współczynnika kierunkowego a prostej $\ln(I) = aU + b$ (wykorzystać metodę najmniejszych kwadratów – patrz aneks A2) oraz odpowiadającą mu wartość bezwymiarowego parametru m . Należy zauważyć, że w przypadku ograniczenia analizy charakterystyki jedynie do kierunku przewodzenia równanie (1) upraszcza się i otrzymujemy:

$$I = I_s \exp\left(\frac{eU}{mkT}\right). \quad (2)$$

Równanie (2) można następnie przepisać w postaci

$$\ln(I) = \frac{eU}{mkT} + \ln(I_s) \quad (3)$$

ułatwiającej analizę wykresu w skali półlogarytmicznej. Jeśli podczas wykonywania pomiarów nie dopuszczano do nadmiernego nagrzewania się diod to można przyjąć, że temperatura T w równaniu (3) równa jest temperaturze pokojowej (ok. 300 K).

9. Rachunek niepewności pomiarowych.
Wyznaczyć złożoną niepewność standardową wartości parametru m obliczonej dla diody D_2 korzystając ze wzoru

$$u(m) = \sqrt{\left(\frac{\partial m}{\partial a}\right)^2 u^2(a) + \left(\frac{\partial m}{\partial T}\right)^2 u^2(T)} = m \sqrt{\left[\frac{u(a)}{a}\right]^2 + \left[\frac{u(T)}{T}\right]^2}, \quad (4)$$

gdzie $u(a)$ jest niepewnością standardową typu A współczynnika kierunkowego a wyznaczonego metodą najmniejszych kwadratów (patrz aneks A2), zaś $u(T)$ jest niepewnością standardową temperatury złącza T . Ponieważ stanowisko pomiarowe nie umożliwia pomiaru temperatury, należy przyjąć, że wartość całkowitej niepewności standardowej $u(T)$ wynika jedynie z niepewności standardowej typu B, którą z kolei można

oszacować na podstawie niepewności odczucia temperatury przez eksperymentatora $\Delta_e T$, przy czym

$$u(T) \approx u_B(T) \approx \frac{\Delta_e T}{\sqrt{3}}. \quad (5)$$

Można założyć, że w typowych warunkach $\Delta_e T = 10$ K, czyli $T = (300 \pm 10)$ K, natomiast w uzasadnionych sytuacjach, np. krótko po włączeniu klimatyzacji lub otwarciu okna w pracowni, przyjęta wartość niepewności $\Delta_e T$ może być większa.

Obliczyć wartość niepewności rozszerzonej $U(m)$ parametru m dla diody D_2 , przyjmując wartość współczynnika rozszerzenia $k = 1,75$.

10. Zapis wyniku końcowego.

Dla diody D_2 zapisać wynik końcowy obliczeń parametru m wraz z jego niepewnością rozszerzoną $U(m)$, zwracając przy tym uwagę na zastosowanie odpowiedniego formatu zapisu (patrz przykłady z aneksie A2).

11. Uwagi końcowe i wnioski.

W uwagach należy zamieścić ocenę dokładności pomiarów oraz własne spostrzeżenia co do przebiegu ćwiczenia. Należy również zidentyfikować badane diody (tzn. określić ich symbole katalogowe) porównując otrzymane wyników z kartami katalogowymi wybranych diod, które można pobrać ze strony internetowej Laboratorium Elektroniki [5]. Ważną część wniosków powinno stanowić wyszczególnienie czynności i procedur pomiarowych, które należałoby przeprowadzić inaczej (jak?), gdyby ćwiczenie miało zostać powtórzone.

W raporcie ocenie podlegać będzie obecność i poprawność wszystkich wymienionych powyżej składników, czytelność prezentacji wyników (w postaci tabel, wykresów) oraz jakość dyskusji i sformułowanych wniosków. Wstęp teoretyczny nie jest wymagany i w przypadku jego zamieszczenia w raporcie nie wpłynie na ocenę.

7. Literatura

7.1. Literatura podstawowa

- [1] F. Przedziecki, A. Opolski, *Elektrotechnika i elektronika*, PWN, Warszawa, 1986.
- [2] A. Rusek, *Podstawy elektroniki*, część pierwsza, WSiP, Warszawa, 1979.
- [3] K. Braclawski, A. Siennicki, *Elementy półprzewodnikowe*, WSiP, Warszawa, 1986.

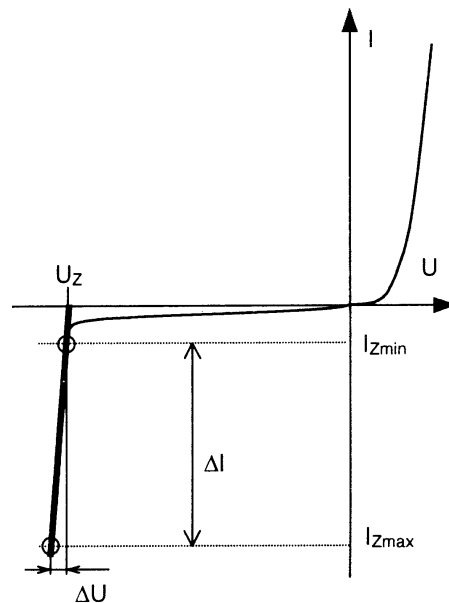
7.2. Literatura uzupełniająca

- [4] Instrukcje obsługi do multimetrów, zasilacza laboratoryjnego, dostępne są na stronie internetowej:
<http://fizyka.p.lodz.pl/pl/dla-studentow/podstawy-elektroniki/peim-zasoby/>
- [5] Karty katalogowe diod – strona internetowa Laboratorium Elektroniki.
- [6] B. Żółtowski, „Skrypt wprowadzenie do zajęć laboratoryjnych”, Skrypt PŁ, Łódź, 2002 – strona internetowa Laboratorium Elektroniki (na serwerze: fizyka.p.lodz.pl).

8. Aneksy

A1. Metoda wyznaczania napięcia Zenera

Napięcie Zenera można wyznaczyć kilkoma metodami. Jedną z najprostszych metod jest określenie napięcia Zenera na podstawie charakterystyki badanej diody Zenera. Przykładowa charakterystyka pokazana jest na rys. A1.1. I_{Zmax} oznacza maksymalną wartość natężenia prądu dopuszczalną dla badanej diody (w przypadku diod badanych w Laboratorium Podstaw Elektroniki maksymalną wartość natężenia prądu ogranicza układ zasilający diody). Przyjmując z kolei, że $I_{Zmin} = 0,05 \cdot I_{Zmax}$ łatwo znaleźć odpowiednie dwa punkty na charakterystyce, a następnie poprowadzić przez nie prostą przecinającą oś napięć. Miejsce przecięcia z osią napięć wyznacza poszukiwane napięcie Zenera U_Z .



Rys. A1.1. Idea metody wyznaczania napięcia Zenera

A2. Metoda najmniejszych kwadratów

Kompletny opis metody najmniejszych kwadratów można znaleźć w skrypcie [6], który jest dostępny w postaci pliku PDF wśród materiałów pomocniczych dla studentów odbywających zajęcia w Laboratorium elektroniki. Zakres prezentacji metody w niniejszym aneksie został dostosowany do zakresu jej wykorzystania w Laboratorium elektroniki.

Typowym zadaniem dla eksperymentatora jest znalezienie analitycznej postaci funkcji $y = f(x)$ na podstawie n par wyników pomiarów (x_i i y_i , gdzie $i = 1, 2, \dots, n$), przeprowadzonych w pewnym zakresie zmian wartości x i y . Proces ten nazywamy dopasowaniem, aproksymacją lub regresją. W przypadku doświadczeń wykonywanych w Laboratorium elektroniki, ogólna postać poszukiwanej funkcji wynika z podanych związków teoretycznych lub przyjętych definicji szukanych wielkości i wówczas celem regresji jest wyznaczenie wartości liczbowych parametrów występujących w założonej ogólnej postaci funkcji. Wśród różnych metod stosowanych w tym celu szczególnie popularna jest tzw. **metoda najmniejszych kwadratów**. Zakładając, że każdy wynik pomiaru y_i podlega rozkładowi Gaussa z taką samą niepewnością oraz że niepewności pomiarów x_i są pomijalnie małe, metoda sprowadza się do minimalizacji sumy kwadratów różnic pomiędzy rzeczywistymi wartościami zmiennej zależnej $y_i(x_i)$ a wartościami przewidzianymi przez model $f(x_i, a, b, c, \dots)$

$$S = \sum_{i=1}^n [y_i - f(x_i, a, b, c, \dots)]^2. \quad (\text{A2.1})$$

Suma S osiąga minimum ze względu na wszystkie parametry funkcji f , czyli a, b, c, \dots gdy $\partial S / \partial a = 0$, $\partial S / \partial b = 0$, $\partial S / \partial c = 0$, \dots . Analityczne rozwiązanie tak otrzymanego układu równań jest znane tylko dla niektórych postaci funkcji f , natomiast minimalizację sumy S dla innych funkcji można zazwyczaj przeprowadzić metodami numerycznymi. W ćwiczeniach przygotowanych w Laboratorium elektroniki Instytutu Fizyki PŁ przewidziano jedynie **regresję liniową**, tzn. dopasowanie współczynników funkcji liniowej.

W przypadku funkcji liniowej $f(\mathbf{x}) = \mathbf{a} \cdot \mathbf{x} + b$ rozwiązanie układu równań $\partial S / \partial a = 0$ i $\partial S / \partial b = 0$ prowadzi do znalezienia współczynników danych wzorami:

$$a = \frac{n \cdot \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \cdot \sum_{i=1}^n y_i}{D}, \quad (\text{A2.2})$$

$$b = \frac{\sum_{i=1}^n x_i^2 \cdot \sum_{i=1}^n y_i - \sum_{i=1}^n x_i \cdot \sum_{i=1}^n x_i y_i}{D}, \quad (\text{A2.3})$$

gdzie

$$D = n \cdot \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2. \quad (\text{A2.4})$$

Precyzja oszacowanych współczynników a i b zależy od ich niepewności standardowych typu A, tzn. odchyłeń standardowych, $u(a)$ i $u(b)$ danych wzorami:

$$u(a) = \sqrt{\frac{n \sum_{i=1}^n (y_i - ax_i - b)^2}{(n-2)D}}, \quad (\text{A2.5})$$

$$u(b) = \sqrt{\frac{\sum_{i=1}^n x_i^2 \cdot \sum_{i=1}^n (y_i - ax_i - b)^2}{(n-2)D}}, \quad (\text{A2.6})$$

gdzie D jest zdefiniowane wzorem (A2.4). Ponadto statystyczne powiązanie (skorelowanie) wartości dwóch zmiennych x_i oraz y_i możemy ocenić obliczając współczynnik korelacji liniowej zdefiniowany jako:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \cdot \sum_{i=1}^n (y_i - \bar{y})^2}}. \quad (\text{A2.7})$$

Współczynnik r przyjmuje wartości z przedziału od -1 do $+1$. Wartości graniczne $r = \pm 1$ oznaczają idealną korelację liniową, natomiast $r = 0$ oznacza brak korelacji i w takim przypadku próby dopasowania funkcji liniowej są nieuzasadnione. Ujemny znak współczynnika korelacji oznacza, że wzrost wartości jednej zmiennej jest związany ze zmniejszeniem wartości innej zmiennej, zaś dodatni znak oznacza zgodny kierunek zmian obu zmiennych.

Jeżeli wykorzystywany związek teoretyczny lub definicja szacowanej wielkości wskazuje na postać funkcji liniowej bez tzw. wyrazu wolnego, $f(\mathbf{x}) = \mathbf{a} \mathbf{x}$, to współczynnik nachylenia prostej należy obliczyć ze wzoru:

$$a = \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2}, \quad (\text{A2.8})$$

zaś jego odchylenie standardowe według wzoru

$$u(a) = \sqrt{\frac{\sum_{i=1}^n (y_i - ax_i)^2}{(n-1) \sum_{i=1}^n x_i^2}}. \quad (\text{A2.9})$$

Autor raportu powinien zawsze samodzielnie rozważyć, który z dwóch wyżej omówionych przypadków jest odpowiedni w rozważanym kontekście.

W praktyce, zamiast pracochłonnego wykonywania kolejnych działań opisanych wzorami (A2.2)-(A2.9), często wykorzystuje się funkcje statystyczne wbudowane w arkusze kalkulacyjne Microsoft Excel oraz Open Office Calc. Szczególnie uniwersalna jest funkcja REGLINP, której składnia jest następująca:

=REGLINP(zakres_y; zakres_x; stała; statystyka)

gdzie

- zakres_y – zakres adresów komórek zawierających wartości y_i ,
- zakres_x – zakres adresów komórek zawierających wartości x_i ,
- stała – wartość logiczna PRAWDA albo FAŁSZ określająca, czy równanie prostej zawiera wyraz wolny b. Wszystkie wartości liczbowe różne od zera, np. 7, są interpretowane jako PRAWDA, zaś wartości 0 odpowiada FAŁSZ,
- statystyka – wartość logiczna określająca czy zostaną obliczone dodatkowe statystyki regresji; w przypadku FAŁSZ zwracane są tylko współczynniki a i b.

Ponieważ funkcja REGLINP zwraca dwuwymiarową tablicę wartości, powinna być wprowadzona jako formuła tablicowa. W tym celu należy najpierw zaznaczyć obszar komórek dla rezultatu funkcji, napisać formułę w pasku edycyjnym nad arkuszem i następnie wprowadzić do arkusza przez naciśnięcie kombinacji klawiszy Ctrl + Shift + Enter.

Założmy, że komórki A1:A5 zawierają znane wartości x_i , komórki B1:B5 zawierają znane wartości y_i , a obszar C1:D5 jest zaznaczony jak w przykładowej Tabeli A2.1 poniżej. Wpisanie w pasku edycyjnym formuły

=REGLINP(B1:B5; A1:A5; 1; 1)

i naciśnięcie kombinacji klawiszy Ctrl + Shift + Enter spowoduje wyświetlenie wyników obliczeń w zaznaczonym uprzednio obszarze. Rozpoczynając od lewego górnego rogu obszaru zaznaczonego kolorem jasnoniebieskim w Tabeli A2.1, kolejność zwracanych wyników jest następująca:

- a – wartość współczynnika nachylenia prostej obliczona według wzoru (A2.2),
 - b – wartość wyrazu wolnego – wzór (A2.3),
 - u(a) – wartość odchylenia standardowego współczynnika a – wzór (A2.5),
 - u(b) – wartość odchylenia standardowego współczynnika b – wzór (A2.6),
 - r^2 – kwadrat współczynnika korelacji r obliczonego według wzoru (A2.7),
- natomiast dalsze wyniki F, d_f , ss_{reg} , ss_{resid} nie mają zastosowania w Laboratorium elektroniki.

	A	B	C	C	D
1	1	110		a	b
2	2	131		u(a)	u(b)
3	3	192		r^2	se_y
4	4	317		F	d_f
5	5	489		ss_{reg}	ss_{resid}

Tabela A2.1. Przykładowe dane do obliczeń (x_i w kolumnie A oraz y_i w kolumnie B) i kolejność wyników regresji zwracanych przez funkcję REGLINP w obszarze zaznaczonym kolorem jasnoniebieskim.

W przypadku gdy argument stała ma wartość FAŁSZ, wyniki są zwracane w podanej wyżej kolejności, z tym że współczynnik b jest wyzerowany, zaś wynik u(b) jest nieobecny.

Jeżeli formuła tablicowa jest trudna do zrozumienia, to można ją zastąpić zbiorem formuł skalarnych z funkcją INDEKS. Wywołanie funkcji z argumentami

=INDEKS(tablica; nr_wiersza; nr_kolumny)

zwraca wartość pojedynczej komórki z argumentu tablica na pozycji o współrzędnych (nr_wiersza; nr_kolumny), które liczone są od (1;1) dla lewego górnego rogu tablicy. Otrzymanie zbioru wyników równoważnych jednemu wywołaniu REGLINP jako formuły tablicowej wymaga zatem użycia kilku formuł skalarnych. Alternatywnie, niektóre wyniki zwracane w tablicy przez funkcję REGLINP, można otrzymać wykorzystując specjalizowane funkcje skalarne NACHYLENIE, ODCIĘTA oraz R.KWADRAT.

a:	=INDEKS(REGLINP(B1:B5; A1:A5; 1; 1); 1; 1)	albo	=NACHYLENIE(B1:B5; A1:A5)
b:	=INDEKS(REGLINP(B1:B5; A1:A5; 1; 1); 1; 2)	albo	=ODCIĘTA(B1:B5; A1:A5)
u(a):	=INDEKS(REGLINP(B1:B5; A1:A5; 1; 1); 2; 1)		
u(b):	=INDEKS(REGLINP(B1:B5; A1:A5; 1; 1); 2; 2)		
r ² :	=INDEKS(REGLINP(B1:B5; A1:A5; 1; 1); 3; 1)	albo	=R.KWADRAT(B1:B5; A1:A5)

Tabela A2.2. Skalarne zamienniki dla funkcji REGLINP wstawianej jako formuła tablicowa.

Należy pamiętać, że format zapisu otrzymanego wyniku wraz z jego niepewnością informuje o rodzaju tej niepewności. Przykładowo, wynik $a = 136,173 \Omega$ razem z jego niepewnością standardową $u(a) = 1,187 \Omega \approx 1,2 \Omega$ zapisujemy w postaci:

$$a = 136,2(1,2) \Omega \quad \text{zapis wyniku końcowego wraz z niepewnością standardową}$$

Jeżeli natomiast została policzona niepewność rozszerzona U – np. dla powyższego wyniku serii $n = 22$ par pomiarów $U(a) = k(P, \nu) u(a) = 3,377 \Omega$, gdzie $k(P, \nu) = 2,845$ jest współczynnikiem rozszerzenia opisanym rozkładem t-Studenta dla przyjętego prawdopodobieństwa objęcia $P = 99\%$ i liczby stopni swobody $\nu = 20$ ($\nu = n - 2$, gdzie 2 to liczba parametrów regresji) – to wynik końcowy należy zapisać w postaci:

$$a = (136,2 \pm 3,4) \Omega \quad \text{zapis wyniku końcowego wraz z niepewnością rozszerzoną}$$

UWAGA! W zależności od wybranego w ustawieniach regionalnych systemu operacyjnego znaku separatora listy, w formułach arkusza kalkulacyjnego Microsoft Excel należy zastosować przecinki albo średniki. W przypadku arkusza kalkulacyjnego Open Office Calc separator listy może być zdeterminowany przez wersję językową programu.

Nazwy opisanych powyżej funkcji i stałych logicznych dotyczą polskich wersji programów. Użytkownicy innych wersji muszą używać nazw specyficznych dla ich języka. Ponadto, nazwy te zostały przetłumaczone tylko dla wybranych języków. Dla pozostałych języków używane są nazwy angielskie, co podczas przygotowywania niniejszego aneksu dotyczyło np. języka ukraińskiego.

Język polski	Język angielski/ukraiński
FAŁSZ	FALSE
INDEKS	INDEX
NACHYLENIE	SLOPE
ODCIĘTA	INTERCEPT
PRAWDA	TRUE
R.KWADRAT	RSQ
REGLINP	LINEST

Tabela A2.3. Nazwy wybranych funkcji i stałych używanych w arkuszach kalkulacyjnych w języku polskim, angielskim i ukraińskim.