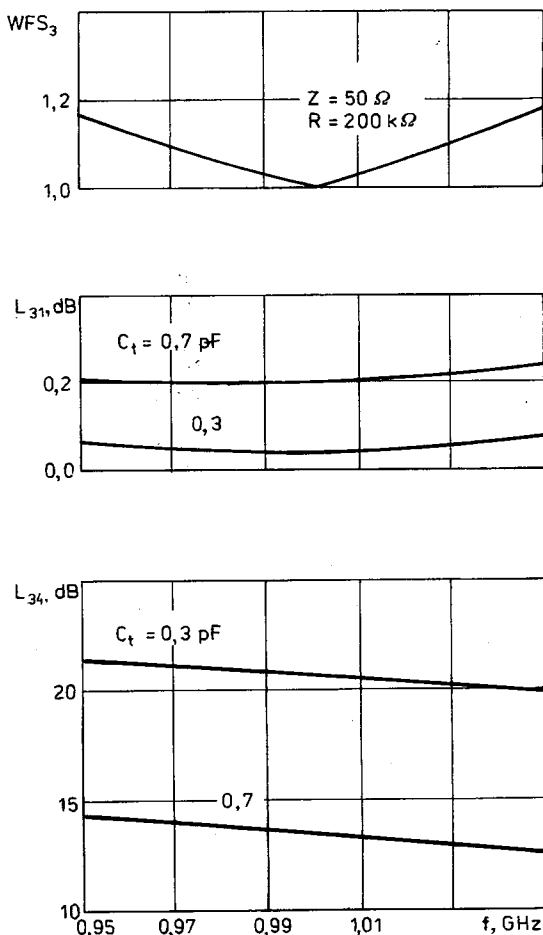
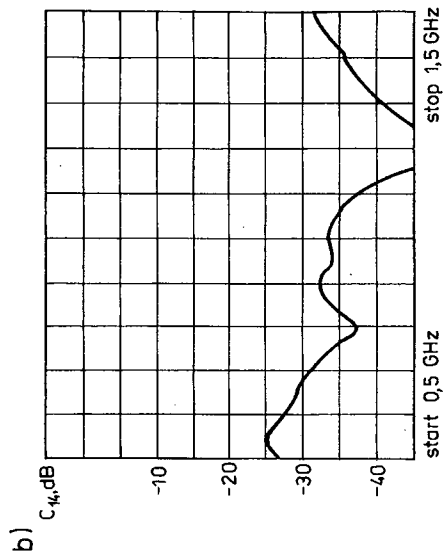
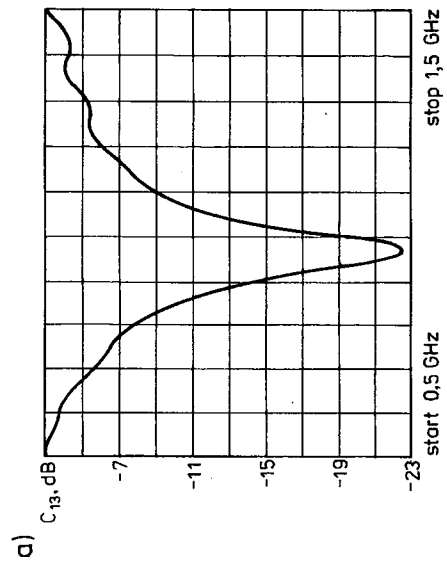
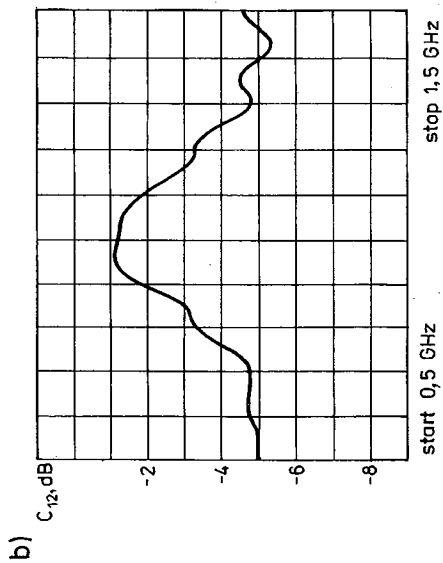
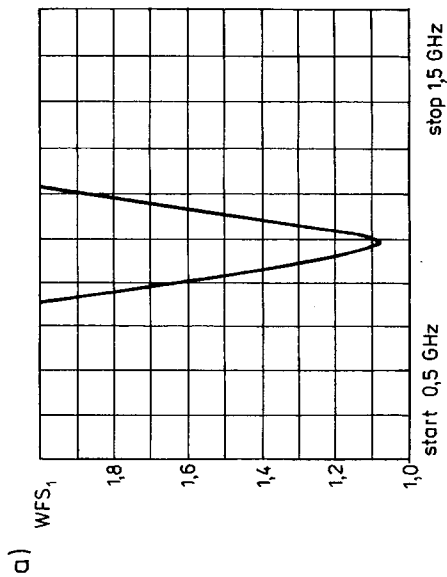




Rys. 7. Charakterystyki $WFS_1(f)$, $L_{12}(f)$ i $L_{13}(f)$ obliczone dla filtra z diodami PIN typu MA4P7000 spolaryzowanymi zaporowo

3. WYNIKI BADAŃ EKSPERYMENTALNYCH

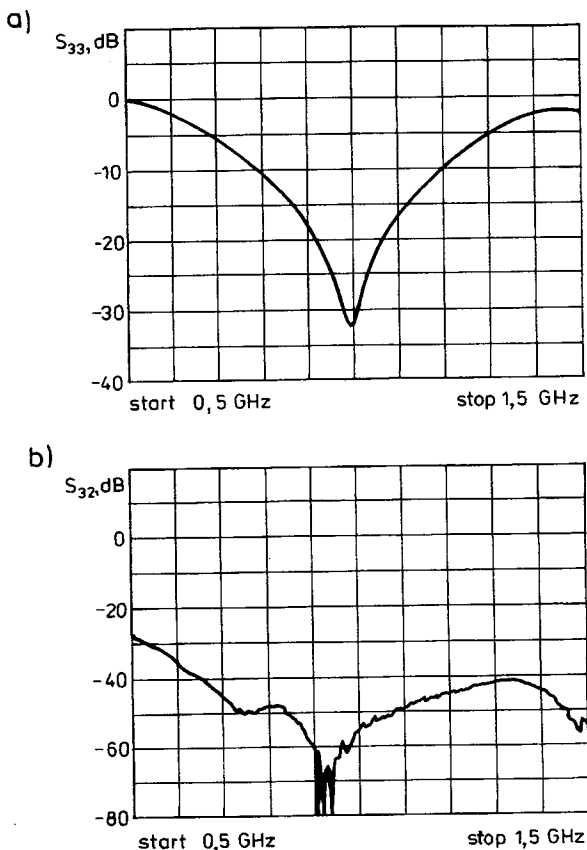
W celu potwierdzenia słuszności przedstawionych wyżej zależności teoretycznych zrealizowano analizowany filtr w technice niesymetrycznej linii paskowej. Podłoże tej linii stanowi laminat epoksydowo-szkłany o grubości dielektryku $h=1.76$ mm i względnej przenikalności dielektrycznej $\epsilon_r=4.23$. Filtr ten zaprojektowano dla następujących danych: $Z_0=50 \Omega$, $Z=50 \Omega$, $f_0=1$ GHz i $C_s=2.2$ nF. W trakcie eksperymentu diody PIN spolaryzowane w kierunku przewodzenia (stan włączenia filtra) zastąpiono krótkimi (bezindukcyjnymi), miedzianymi mostkami. Uzyskane dla takiego modelu filtra charakterystyki: $WFS_1(f)=[1+|S_{11}(f)|]/[1-|S_{11}(f)|]$, $C_{12}(f)[dB]=20\log|S_{12}(f)|$, $C_{13}(f)[dB]=20\log|S_{13}(f)|$ i $C_{14}(f)[dB]=20\log|S_{14}(f)|$



Rys. 8. Charakterystyki: a) $WFS_1(f)$ i b) $C_{12}(f)$ wyznaczone eksperymentalnie dla filtra z idealnymi diodami PIN będącymi w stanie włączenia

Rys. 9. Charakterystyki: a) $C_{13}(f)$ i b) $C_{14}(f)$ wyznaczone eksperymentalnie dla filtra z idealnymi diodami PIN będącymi w stanie włączenia

przedstawiono kolejno na rys. 8 i 9. Poprzez odlutowanie wspomnianych wyżej miedzianych mostków uzyskuje się stan wyłączenia filtra. Wyznaczone w tych warunkach charakterystyki strat odbiciowych $S_{33}(f)[\text{dB}] = 20\log |S_{33}(f)|$ i sprzężenia $C_{32}(f)[\text{dB}] = 20\log |S_{32}(f)|$ pokazano na rys. 10.



Rys. 10. Charakterystyki: a) $S_{33}(f)$ i b) $C_{32}(f)$ wyznaczone eksperymentalnie dla filtra z idealnymi diodami PIN będącymi w stanie włączenia

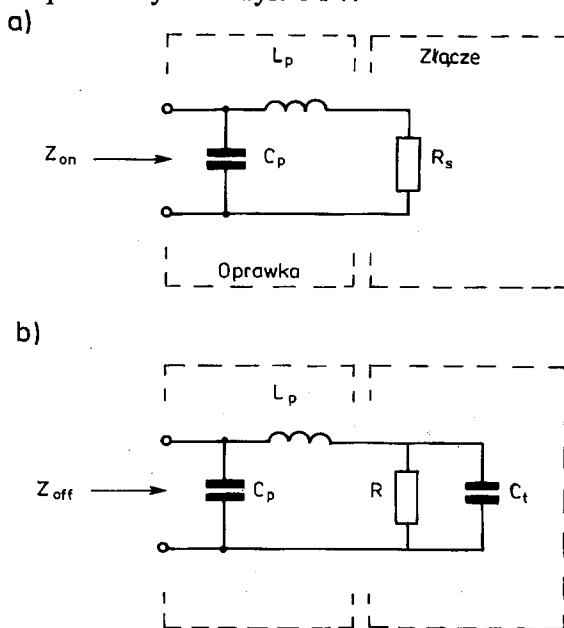
WNIOSKI

W pracy przedstawiono nową, oryginalną strukturę mikrofalowego filtra kierunkowego z dwiema sterowanymi prądowo diodami PIN. Przy spolaryzowaniu tych diod w kierunku zaporowym (stan wyłączenia) tłumienie w linii głównej filtra jest pomijalnie małe w zadanym, względnie wąskim, pasmie częstotliwości. Jednocześnie zapewniona jest duża (teoretycznie nieskończona) szerokopasmowa separacja pomiędzy przeciwnymi wrotami linii głównej i sprzężonej. Właściwość ta zachowuje się także w przypadku spolaryzowania diod w kierunku przewodzenia (stan włączenia), o ile zastosowane diody są identyczne. Dzięki tej, unikalnej właściwości omawiany

filtr może być stosowany jako zwrotnica antenowa wielu urządzeń radiokomunikacyjnych i radiolokacyjnych wykorzystujących dla nadawania i odbioru te same częstotliwości. Słuszność prezentowanych w artykule charakterystyk teoretycznych potwierdzają wyniki eksperymentu przeprowadzonego w pasmie L.

DODATEK

Na rys. D1 przedstawiono schematy zastępcze diod PIN spolaryzowanych w kierunkach przewodzenia i zaporowym. Zgodnie z [10] diody typu MA4P7000 dla $U_R = 100 \text{ V}$ i $I_0 = 100 \text{ mA}$ charakteryzują się następującymi parametrami: $C_t(U_R) = 0.7 \text{ pF}$, $R_s(I_0) = 0.8 \Omega$ i $R(U_R) = 200 \text{ k}\Omega$. Parametry te wykorzystano przy obliczaniu charakterystyk filtru pokazanych na rys. 6 i 7.



Rys. D1. Schematy zastępcze diod PIN a) spolaryzowanych w kierunku przewodzenia i b) spolaryzowanych zaporowo

BIBLIOGRAFIA

1. S.B. C o h n, F.S. C o a l e: *Directional channel separation filters*. IRE Proc. v-44, August 1956, pp. 1018–1024
2. R.D. W a n s e l o w, L.P. T u t t l e Jr.: *Practical design of strip transmission line half-wavelength resonator directional filters*, IRE Trans. MTT-7, Microwave Theory and Techniques, January 1959, pp. 168–173
3. G.L. M a t t h a e i, L. Y o u n g, E.M.T. J o n e s: *Microwave filters, impedance matching networks and coupling structures*, McGraw Hill, New York 1964, Chapter 14

4. J. Reed, G. Wheeler: *A method of analysis of symmetrical four-port networks*. IRE Trans., MTT-4, Microwave Theory and Techniques, October 1956, pp. 246—252
5. S. Rosłonec: *Metody matematyczne w projektowaniu układów elektronicznych o parametrach rozłożonych*, WNT, Warszawa 1988
6. S. Rosłonec, P. Łochowicz: *Split-tee power divider*. RF Design, Vol. 13, February 1990, pp. 52—56
7. M.W. Medley Jr.: *Microwave and RF circuits*. Artech House Inc., Boston—London 1993.
8. T.Q. Ho, S.M. Hart: *A broad-band complanar waveguide to slotline transition*. IEEE Microwave and Guided Wave Letters V-2, October 1992, pp. 415—416
9. V.F. Fusco, Q. Chen: *Slotline short and open circuit analysis by the finite-difference time-domain method*. Proc. of the 24 European Microwave Conference, Cannes 1994, vol. 2, pp. 1720—1726
10. M/A COM Catalogue: *rf and microwave semiconductors*. Burlington, MA 1994, pp. 19—24

S. ROSŁONIEC, T. HABIB

TWO PIN DIODE MICROWAVE DIRECTIONAL FILTER

S u m m a r y

This contribution presents a new microwave directional filter incorporating two current controlled PIN diodes. An unique feature of this filter is that it offers perfect, independent of frequency, isolation for two pair of opposite ports over the relatively wide frequency range. These isolation characteristics are also independent of the PIN diode parametres if used diodes are identical. Above conclusions relate to the filter operating in two modes (reception and transmission), i.e. when PIN diodes are forward or reverse biased. Thus, the proposed filter seems to be suitable for practice, especially for some pulse radar applications.

Key words: telecommunication, microwave technique, filters.

Rozwój konstrukcji tranzystorów IGBT

ZBIGNIEW LISIK

Instytut Elektroniki, Politechnika Łódzka

Otrzymano 1996.03.21

Autoryzowano do druku 1996.04.18

W pracy przedstawiono genezę tranzystora IGBT oraz omówiono oddzielne drogi rozwojowe dwóch odmian tego przyrządu — tranzystora V-IGBT (vertical IGBT), o kontaktach emitera i kolektora ułożonych po przeciwległych płaszczyznach płytki półprzewodnikowej oraz tranzystora L-IGBT (lateral IGBT), o kontaktach emitera i kolektora umieszczonych po tej samej stronie pastylki półprzewodnikowej. Tranzystory te utworzono z odpowiednich wersji tranzystorów polowych DMOS w wyniku zmiany typu materiału podłożowego. W efekcie do struktury został wprowadzony tranzystor bipolarny a przyrząd stał się przyrządem BiMOS. W kolejnych rozdziałach pracy omówiono modyfikacje zmierzające do polepszenia parametrów tych przyrządów.

Słowa kluczowe: tranzystor V-IGBT, tranzystor L-IGBT, tranzystory DMOS, przyrządy BiMOS.

1. WSTĘP

Mimo ogromnej ekspansji przyrządów półprzewodnikowych sterowanych polowo, podstawową grupą przyrządów półprzewodnikowych mocy były i jeszcze w dalszym ciągu pozostają przyrządy bipolarne, do których należą diody mocy, tyrystory mocy oraz bipolarne tranzystory mocy. Ich podstawową cechą jest występowanie wysokooporowego obszaru, który w stanie zaporowym zapewnia odpowiednio wysokie napięcia blokowania, a w stanie przewodzenia, dzięki modulacji jego przewodności w wyniku wprowadzenia nośników nadmiarowych, staje się obszarem niskooporowym, zapewniającym mały spadek napięcia na załączonym przyrządzie. Przyrządy te są sterowane ładunkowo. Oznacza to, że o ich zachowaniu decydują aktualne rozkłady koncentracji nośników nadmiarowych, a każda zmiana punktu pracy przyrządu wymaga doprowadzenia lub usunięcia odpowiedniego ładunku nośników nadmiarowych.

To ładunkowe sterowanie pracą przyrządów bipolarnych mocy jest powodem ich głównych wad, do których należą m.in. ograniczona szybkość przełączania, zdeteminowana przez czas niezbędny do zmiany koncentracji nośników nadmiarowych,

oraz konieczność stosowania do sterowania odpowiednio wydajnych źródeł prądowych, co w szczególnie drastycznej formie występuje np. w przypadku bramkowego wyłączania tyrystorów GTO. Wad tych nie wykazują przyrządy unipolarne. W przyrządach tych prąd płynie zawsze poprzez obszar jednego typu, nie występują w nich nadmiarowe koncentracje nośników, a efekt zmiany konduktancji uzyskuje się poprzez zmianę geometrii kanału przewodzącego prąd. Zmiana ta jest wywołana sygnałem napięciowym, który w przypadku złączowych przyrządów polowych jest przyłożony do wstecznie spolaryzowanego złącza p-n, a w przypadku przyrządów z izolowaną bramką, do okładek kondensatora MIS. Oba te rozwiązania, zastosowane początkowo w małych elementach, doczekały się w latach 80 i 90 swoich odpowiedników w dziedzinie przyrządów średniej i dużej mocy. Typowymi przyrządami mocy, w których zastosowano to pierwsze rozwiązanie są tranzystory SIT i tyrystory FCT [1–3], natomiast do najbardziej rozpowszechnionych przyrządów, w których zastosowano drugie rozwiązanie należą tranzystory VMOS i VDMOS [4–7].

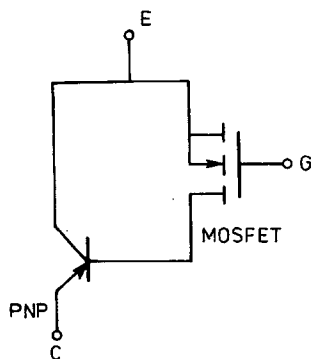
Dyskretne tranzystory mocy VMOS i VDMOS, dzięki swojej dużej rezystancji wejściowej i małych czasach przełączania, skutecznie wyparły elementy bipolarne z szeregu zastosowań. Zakres tych zastosowań był jednak w istotny sposób limitowany przez wzrost spadku napięcia dren–źródło w stanie przewodzenia wraz ze wzrostem maksymalnych napięć blokowania tych tranzystorów, co powodowało, że napięcia te nie przekraczały w praktyce kilkuset woltów. W celu usunięcia tej istotnej wady zaproponowano w roku 1983 modyfikację tradycyjnej struktury tych tranzystorów polegającą na wprowadzeniu do niej tranzystora bipolarnego odpowiedzialnego zarówno za wielkość blokowanego napięcia jak i spadek napięcia na przewodzącym przyrządzie. W efekcie, przy zachowaniu tego samego napięcia blokowania, udało się w nich osiągnąć około 10-krotnie niższe napięcia przewodzenia, dzięki modulacji przewodności wysokooporowej warstwy w wyniku wstrzykiwania do niej nośników nadmiarowych. Przyrządy te należą do rodziny przyrządów Bi-MOS i nadano im pierwotnie nazwę COMFET (CONductivity Modulated FET), dla podkreślenia istoty ich działania. Obecnie występują one głównie pod nazwą tranzystorów IGBT (Isolated Gate Bipolar Transistor), a ich pozycja na rynku przyrządów półprzewodnikowych mocy nieustannie rośnie.

Tranzystory IGBT, analogicznie jak polowe tranzystory mocy, są wytwarzane jako układ scalony utworzony z dużej ilości połączonych równolegle identycznych komórek, z których każda jest pojedynczym tranzystorem. Pierwsze tranzystory IGBT powstały na bazie technologii tranzystorów DMOS, w której zmieniono jedynie typ materiału podłożowego. Postępowanie takie było racjonalne z ekonomicznego punktu widzenia chociaż zarówno możliwy zakres zastosowań jak i warunki pracy obu typów przyrządów znacznie się różnią. Od tego czasu dokonano szeregu zmian konstrukcyjnych zmierzających do polepszenia parametrów produkowanych przyrządów. Postęp ten jest tak szybki, że przyjmuje się iż mniej więcej co trzy lata pojawia się nowa generacja tranzystorów IGBT. Obecnie na rynku występują przyrządy zaliczane umownie do II [8, 9] i III [9, 10] generacji, a wkrótce należy oczekiwać kolejnych generacji. W pracy przedstawiono genezę i drogi rozwoju podstawowych

konstrukcji tranzystorów IGBT. Szczególny nacisk położono na przedstawienie pożądanych kierunków zmian konstrukcyjnych oraz konkretnych rozwiązań wprowadzonych do tych przyrządów przy tworzeniu kolejnych ich generacji.

2. ZASADA DZIAŁANIA TRANZYSTORA IGBT

Tranzystor IGBT należy do rodziny przyrządów Bi-MOS. Istota tego rozwiązania polega na scaleniu w jednej strukturze półprzewodnikowej połączenia tranzystora bipolarnego i tranzystora MOS z kanałem wzbogacanym w konfiguracji przedstawionej na rys. 1. W układzie tym tranzystor bipolarny jest elementem mocy, przez który powinna płynąć główna część prądu obciążenia, natomiast tranzystor polowy pełni zasadniczo rolę elementu sterującego pracą tranzystora bipolarnego.



Rys. 1. Schemat zastępczy tranzystora IGBT

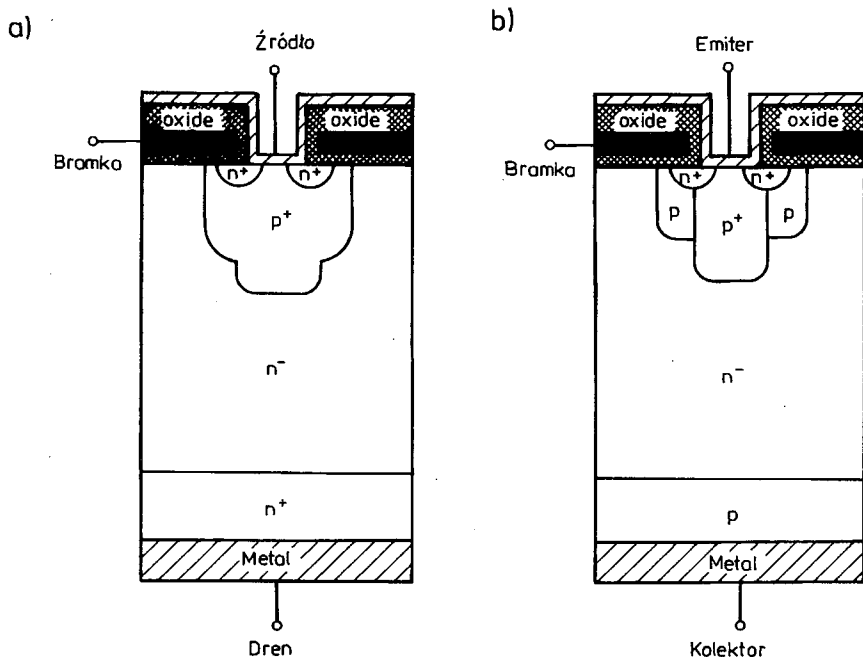
Zaletą tej konfiguracji jest możliwość uzyskiwania takich napięć na przewodzącym elemencie jakie wynikają z parametrów tranzystora bipolarnego przy jednoczesnym zachowaniu wszystkich zalet wynikających ze sterowania pracą przyrządu w sposób właściwy dla tranzystorów polowych. Zastosowanie tranzystora bipolarnego jako elementu mocy powoduje jednak, że dynamika całego układu, a w szczególności czasy przełączania, jest zdeterminowana przez szybkość zmian koncentracji nośników nadmiarowych. W efekcie, układ ten jako całość może pracować przy znacznie mniejszych częstotliwościach niż równoważny mu tranzystor polowy mocy.

Z uwagi na uwarunkowania wynikające z zastosowanych rozwiązań konstrukcyjnych i technologicznych, w rzeczywistych realizacjach koncepcji przedstawionej na rys. 1 można wyodrębnić obok podstawowych elementów uwidocznionych na rysunku również elementy dodatkowe. Elementy te mogą mieć istotny wpływ na własności powstałego układu scalonego Bi-MOS i muszą być brane pod uwagę przy optymalizacji jego konstrukcji. W literaturze przedstawiono szereg konkretnych rozwiązań takiej struktury Bi-MOS [11–18], praktycznie znaczenie uzyskały jednak jedynie dwa podstawowe rozwiązania, omówione w dalszej części rozdziału.

2.1. WERSJA Z DWUSTRONNYMI KONTAKTAMI (VERTICAL IGBT)

Rozwiązania podstawowe

Koncepcja tranzystora IGBT z dwustronnymi kontaktami została po raz pierwszy przedstawiona w roku 1983 [11, 12], jako rozwinięcie konstrukcji tranzystora polowego VDMOS. Wprowadzona modyfikacja polegała na zastąpieniu wyjściowej warstwy podłożowej n^+ warstwą p , jak to pokazano na rys. 2. Dzięki temu uzyskano w przyrządzie emiterowe złącze p - n^- wstrzykujące nadmiarowe dziury do niskodmieszkowego obszaru n^- . Jak poprzednio, przyrząd jest sterowany napięciowo napięciem bramka – katoda (źródło), ale teraz prąd elektronowy wpływający poprzez zaindukowany kanał do obszaru n^- wywołuje po osiągnięciu złącza emiterowego wstrzykiwanie nadmiarowych dziur. W efekcie następuje zwiększenie konduktywności obszaru n^- o kilka rzędów.

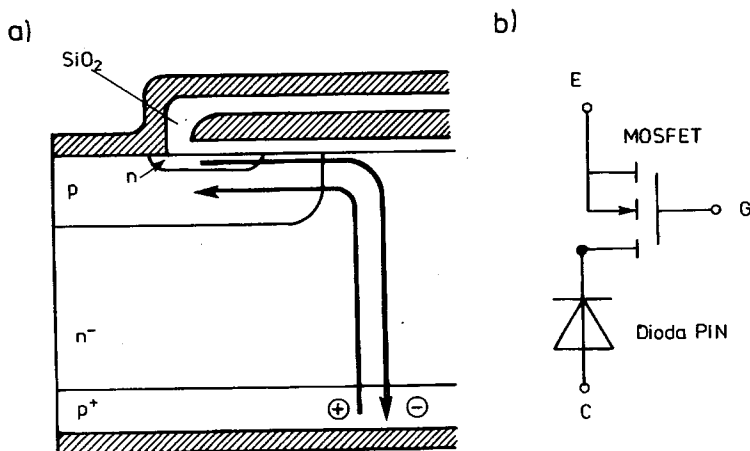


Rys. 2. Komórka tranzystora VDMOS (a) oraz odpowiadająca jej komórka tranzystora V-IGBT (b)

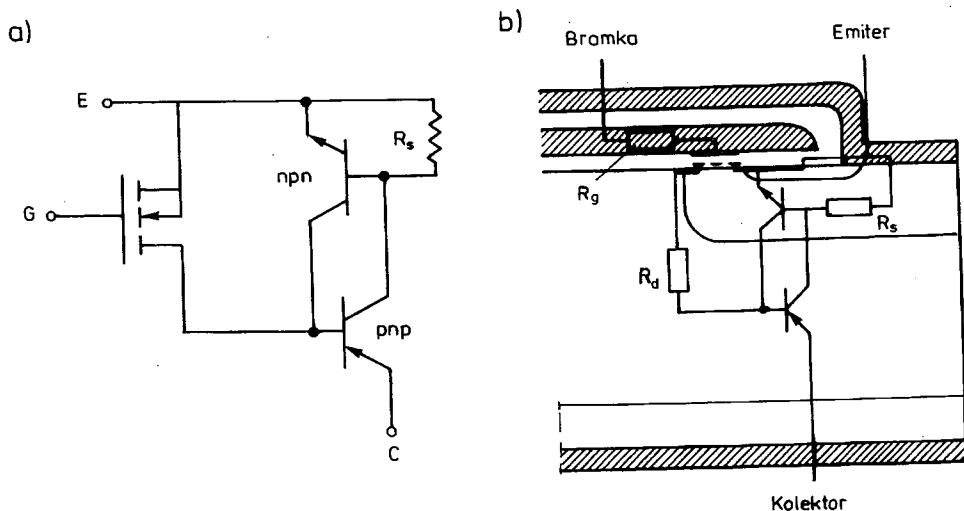
Wprowadzenie do konstrukcji tranzystora V-IGBT złącza emiterowego powoduje, że niskodmieszkowy obszar n^- pracuje w warunkach analogicznych jak niskodmieszkowana warstwa v diody p - v - n . W stanie przewodzenia obszar ten ulega „zalaniu” przez nośniki nadmiarowe dzięki czemu tworzy się niskooporowa ścieżka przepływu prądu obejmująca zaindukowany kanał i „zalany” obszar n^- . Pozwala to na potraktowanie przewodzącego tranzystora IGBT jako równoległego połączenia tranzystora MOS i diody PIN, jak to pokazano schematycznie na rys. 3. Spadek napięcia na przyrządzie jest w tym przypadku równy sumie spadków napięć na obu

elementach składowych i z uwagi na dużą gęstość prądu w kanale jest zwykle rzędu kilku woltów.

Struktura wewnętrzna komórki tranzystora V-IGBT jest dość złożona i pełny schemat zastępczy tego tranzystora musi to uwzględniać. Na rys. 4 przedstawiono taki schemat wraz z widokiem komórki, na którym zaznaczono lokalizację poszczególnych elementów tego schematu. Obok podstawowych elementów struktury Bi-MOS z rys. 1, tzn. sterującego tranzystora polowego oraz wyjściowego tranzystora bipolarnego, zawiera on szereg dodatkowych elementów, które w sposób istotny wpływają na warunki pracy przyrządu. Do najważniejszych z nich należy układ



Rys. 3. Rozkład prądu w komórce tranzystora V-IGBT w stanie przewodzenia (a) i odpowiadający mu podstawowy schemat zastępczy (b)



Rys. 4. Schemat zastępczy komórki tranzystora V-IGBT (a) oraz lokalizacja poszczególnych elementów schematu w obszarze komórki (b)

TT utworzony z dwóch tranzystorów bipolarnych, tranzystora wyjściowego pnp i tranzystora pasożytniczego npn obejmującego obszar n^+ źródła tranzystora polowego oraz obszary p i n tranzystora wyjściowego. Jest to więc typowa czterowarstwowa struktura tyrystorowa $n^+ - p - n - p^+$. Struktura ta może zostać załączona w sposób niekontrolowany o ile w wyniku wzrostu gęstości prądów emiterowych tranzystorów składowych zostanie lokalnie spełniony warunek $\alpha_{npn} + \alpha_{pnp} \geq 1$. Efekt ten, nazywany efektem zatrasku (*latch-up effect*), prowadzi do utraty sterowalności przez przyrząd i z tego względu jego wystąpienie w normalnych warunkach pracy jest niedopuszczalne.

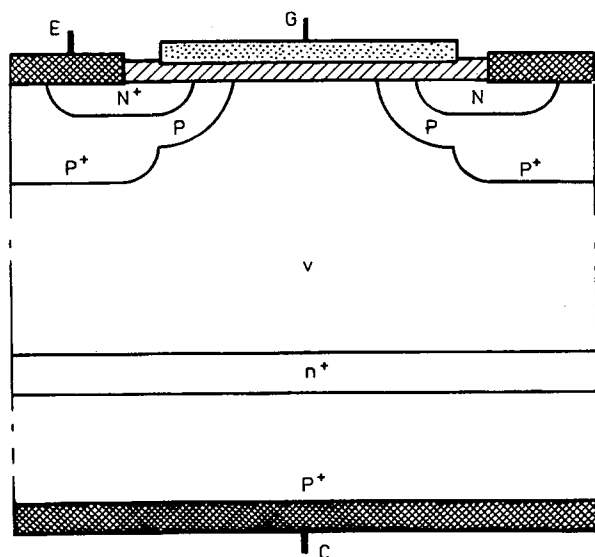
Zachowanie pełnej sterowalności jest podstawowym warunkiem poprawnej pracy tranzystora V-IGBT. Jak wynika z rys. 4, przepływ poprzecznego prądu kolektorowego tranzystora wyjściowego w obszarze kolektora pod wysepką n^+ wywołuje spadek napięcia na poprzecznej rezystancji R_s . Napięcie to polaryzuje w kierunku przewodzenia złącze emiterowe tranzystora pasożytniczego, które pełni rolę złącza bramkowego struktury tyrystorowej. W efekcie, przy odpowiednio dużym spadku napięcia na rezystancji poprzecznej pod obszarem n^+ , może dojść do załączenia tyrystora. Konieczność uniknięcia tego efektu prowadzi do szeregu istotnych ograniczeń w wykorzystaniu pełnych możliwości tej struktury Bi-MOS (gęstość prądu musi być odpowiednio mała). Jest ona także głównym impulsem wprowadzenia szeregu rozwiązań konstrukcyjnych, o których będzie mowa w następnym rozdziale.

Rozwój konstrukcji

Pierwsze tranzystory V-IGBT powstały w połowie lat 80-tych na bazie technologii tranzystorów VDMOS, w której zmieniono jedynie typ materiału podłożowego (rys. 2). Od tego czasu dokonano szeregu zmian konstrukcyjnych zmierzających do polepszenia parametrów produkowanych przyrządów, a postęp ten jest tak szybki, że przyjmuje się iż mniej więcej co trzy lata pojawia się nowa generacja tranzystorów IGBT. Zmiany konstrukcyjne zmierzają generalnie do otrzymania przyrządu szybkiego, o krótkich czasach przełączania, o małym spadku napięcia w stanie przewodzenia, dużym napięciu blokowania, a ponadto niewrażliwego na efekt *latch-up*. Jest oczywiste, że nie można uzyskać przyrządu zawierającego w sobie wszystkie te cechy, szczególnie że ich uzyskanie wymaga często podjęcia przeciwnych działań. Z tego względu dąży się do wytwarzania przyrządów konstruowanych pod kątem konkretnych zastosowań. Można tu wyróżnić trzy podstawowe grupy tranzystorów IGBT przeznaczonych kolejno do zastosowań jako: statyczne klucze, elementy układów wysokiej częstotliwości oraz elementy układów napędowych [19].

Dokładne omówienie zasad optymalizacji konstrukcji wykracza poza ramy tej pracy, poniżej przedstawiono natomiast typowe rozwiązania konstrukcyjne wprowadzane w celu uzyskania pożądaných wartości wymienionych wyżej parametrów:

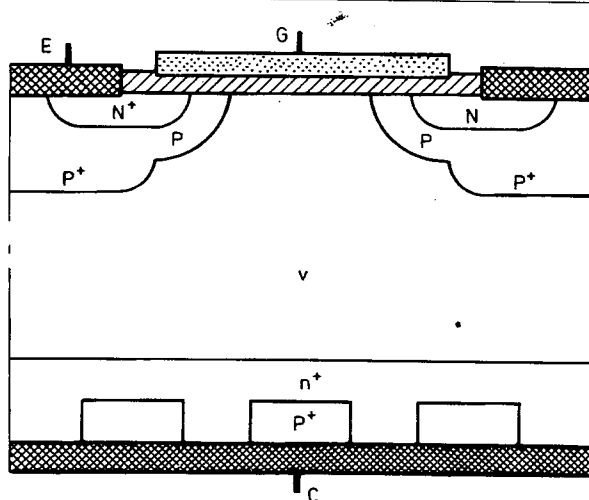
- **dodatkowa warstwa buforowa n^+** — rozwiązanie to, przedstawione na rys. 5, jest analogiczne do rozwiązania stosowanego w tyrystorach asymetrycznych ASCR i ma podobne przeznaczenie. Pozwala ono na uzyskanie żądanych napięć blokowa-



Rys. 5. Komórka tranzystora NPT IGBT z buforową warstwą n^+

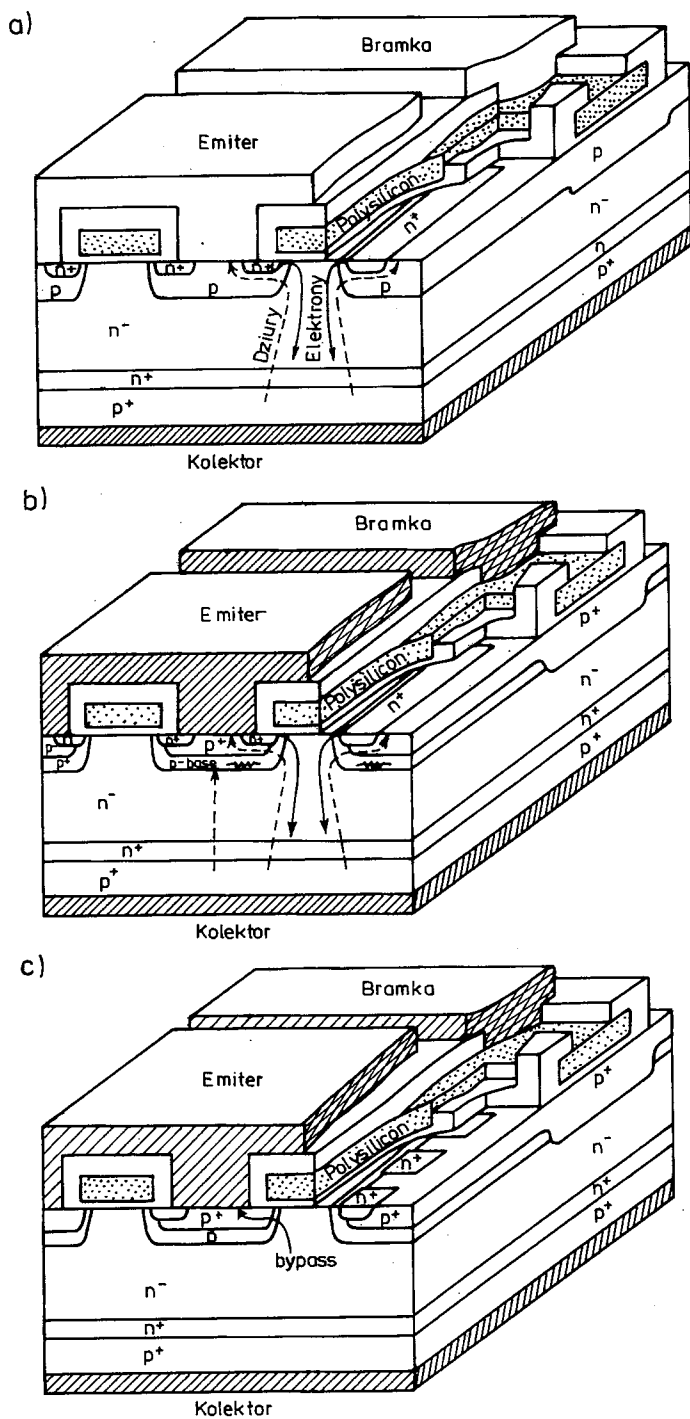
nia przy zredukowanej grubości n-bazy, a więc i niższym napięciu przewodzenia oraz krótszym czasie wyłączania, bez ryzyka wystąpienia przebiccia skrośnego przyrządu (ang. *punch through*). Stąd podział tranzystorów IGBT na tranzystory NPT IGBT (Non Punch Through) z warstwą buforową oraz tranzystory PT IGBT bez tej warstwy. Warstwa ta dodatkowo zmniejsza współczynnik wzmocnienia α_{pnp} tranzystora bipolarnego wpływając na zwiększenie prądu tranzystora IGBT, przy którym wystąpi załączenie tyrystora pasożytniczego (efekt *latch-up*).

- **zmiany grubości warstwy emiterowej p^+** — wprowadzenie odpowiednio cienkiej warstwy anodowej pozwala na zmniejszenie efektywności wstrzykiwania emitery anodowego i tym samym zwiększenie prądu przy którym następuje niekontrolowane załączenie pasożytniczego tyrystora.
- **wprowadzenie zwarć emiterowych do warstwy p^+** — rozwiązanie analogiczne do stosowanego w tyrystorach. Ma ono na celu zmniejszenie sumarycznej efektywności wstrzykiwania dziur przez p^+ -emiter tranzystora pnp i tym samym zmniejszenie współczynnika wzmocnienia prądowego α_{pnp} tego tranzystora. W efekcie wzrasta wartość prądu obciążenia, I_L , przy której występuje efekt *latch-up*. Istota tego rozwiązania jest przedstawiona na rys. 6.
- **wprowadzenie centrów rekombinacyjnych do n-bazy** — ma na celu zmniejszenie czasu życia nośników nadmiarowych i zmniejszenie czasu wyłączania, zwykle w wyniku napromieniowania elektronami [23]. Ubocznym skutkiem jest zwiększenie napięcia przewodzenia.

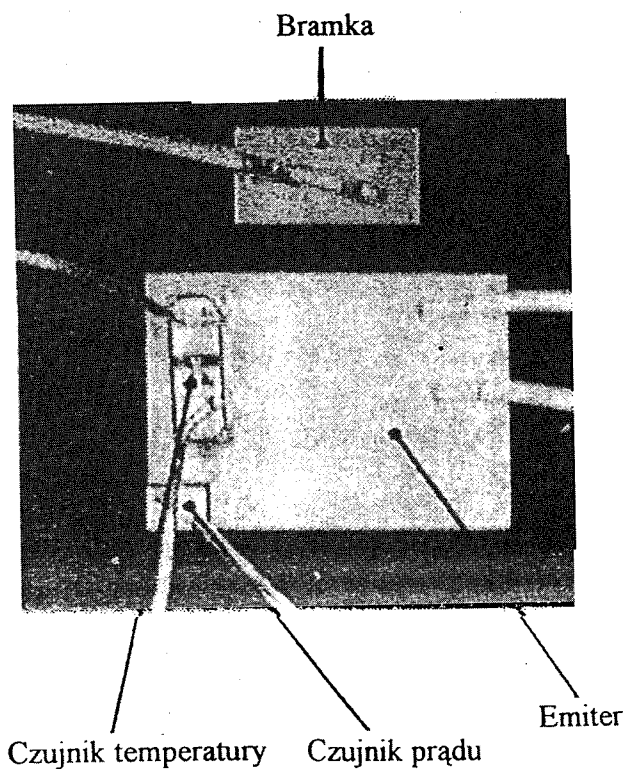
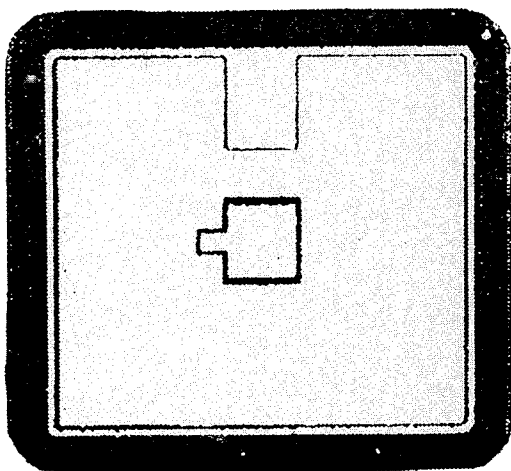


Rys. 6. Tranzystor IGBT z konstrukcją zwartego emitera zastosowaną do warstwy p^+

- **wprowadzenie dodatkowej warstwy p^+ w obrębie wysepki katodowej** — rozwiązanie to jest przedstawione na rys. 7b. Celem jego jest zmniejszenie rezystancji poprzecznej R_s obszaru p-bazy tranzystora pasożytniczego oraz zmniejszenie efektywności wstrzykiwania emitera n^+ tego tranzystora. Daje to w efekcie wzrost prądu I_L , przy którym wystąpi efekt *latch-up*. W rozwiązaniu tym jest istotne aby warstwa silnie domieszkowana p^+ nie sięgała obszaru, w którym jest tworzony kanał tranzystora MOS. Obszar ten musi być niskodomieszkowany aby zapewnić odpowiednio niskie napięcie progowe tranzystora MOS oraz niską oporność kanału w stanie załączonym.
- **konstrukcja zbocznikowanej warstwy n^+ emitera** — rozwiązanie to jest przedstawione na rys. 7c. Polega ono na wprowadzeniu przerw typu p^+ w warstwie n^+ emitera w celu zwiększenia przekroju poprzecznego niskooporowej ścieżki dla prądu dziuowego w p-bazie tranzystora pasożytniczego. W efekcie rezystancja poprzeczna R_s ulega dalszemu zmniejszeniu i maleje zagrożenie efektem *latch-up*.
- **pierścienie ograniczające** — konstrukcja V-IGBT powstała jako rozwinięcie konstrukcji VDMOS i złącze odpowiedzialne za wartość blokowanego napięcia jest w nich zwykle wykonywane jako złącze planarne, co pociąga za sobą ograniczenia maksymalnej wartości tego napięcia. Aby uzyskać większe wartości napięcia blokowania przy zachowaniu technologii planarnej stosuje się na obrzeżu struktury tzw. pierścienie ograniczające. Wpływają one na rozkład natężenia pola elektrycznego przy powierzchni i pozwalają na takie jego uformowanie aby uniknąć powierzchniowego przebicia złącza także przy dużych wartościach blokowanego napięcia [24, 25].



Rys. 7. Przekrój przez strukturę tranzystora IGBT firmy TOSHIBA: (a) wykonanego standardowo [20], (b) wykonanego z podwójną dyfuzją p-bazy [21], (c) wykonanego z podwójną dyfuzją p-bazy i zbocznikowaną warstwą emiterową [22]



Rys. 9. Struktury tranzystora IGBT z wbudowanymi „inteligentnymi” układami zabezpieczającymi : (a) widok całej struktury z czujnikiem przeciążenia prądowego w środku struktury [30], (b) widok fragmentu z kontaktem emitera i bramki oraz wbudowanym czujnikiem przeciążenia prądowego i czujnikiem temperatury naklejonym na elektrodzie emitera [42]

kami. Zbyt mały dystans pomiędzy wysepkami wywołuje wystąpienia efektu pasożytniczego złączowego tranzystora polowego JFET, który prowadzi do zmniejszenia koncentracji nośników w n-bazie i zwiększenia spadku napięcia na przyrządzie w stanie przewodzenia. Efektowi temu zapobiega się utrzymując parametr b na odpowiednio dużym poziomie (ok. $16\text{--}20\text{ }\mu\text{m}$) lub wprowadzając w przestrzeń pomiędzy wysepkami p dodatkową wysepkę n^+ . Utrzymanie odpowiednio niskiego spadku napięcia w stanie przewodzenia wymaga również aby szerokość wysepki, a , oraz odstęp pomiędzy nimi, b , były w przybliżeniu równe sobie.

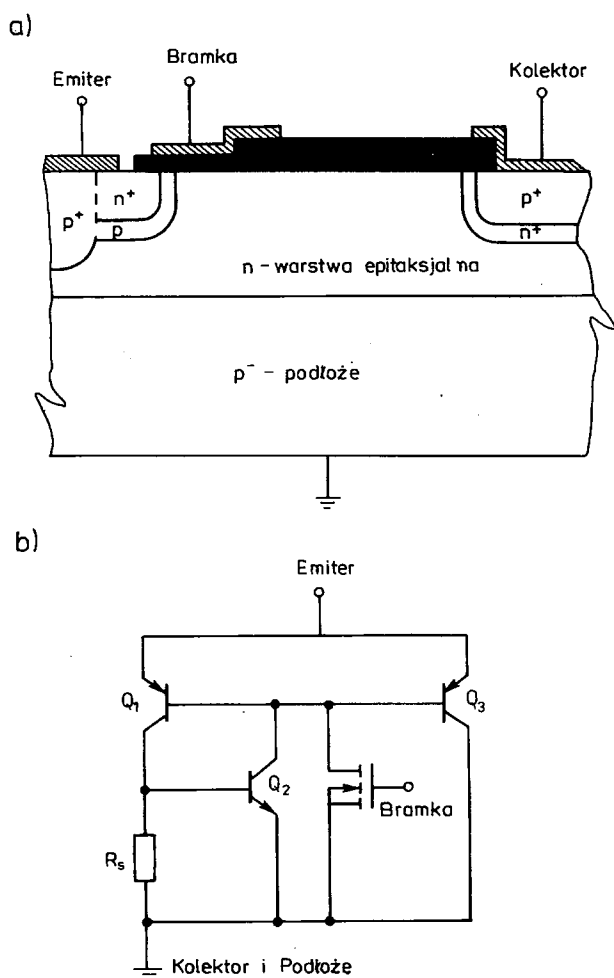
- **wbudowane układy zabezpieczające** — możliwość łatwego wyłączenia tranzystora IGBT poprzez likwidację kanału w wyniku usunięcia napięcia polaryzującego bramkę było podstawą wprowadzenia tzw. „inteligentnych” tranzystorów IGBT. Zawierają one wbudowane układy zabezpieczające reagujące na nadmierną wartość prądu (zabezpieczenie przeciwzwarciowe i przeciw efektowi *latch-up*) lub temperatury (zabezpieczenie przeciążeniowe) zwarcie elektrod bramki i kolektora powodującym wyłączenie przyrządu. Początkowo układy zabezpieczające były wprowadzone do modułów z tranzystorami IGBT, jako odrębne układy scalone, a obecnie pojawiają się już rozwiązania, w których układ zabezpieczający znajduje się na tym samym chipie co tranzystor [27 – 30]. Przykład takiego rozwiązania jest pokazany na rys. 9.

2.2. WERSJA Z KONTAKTAMI JEDNOSTRONNYMI (LATERAL IGBT)

Rozwiązanie podstawowe

Koncepcja tranzystora IGBT z kontaktami jednostronnymi, określana mianem tranzystora L-IGBT, została po raz pierwszy przedstawiona w roku 1984 [13]. Jest ona rozwinięciem konstrukcji lateralnego tranzystora LDMOS. Podobnie jak to miało miejsce w przypadku tranzystora V-IGBT, podstawowa modyfikacja tej struktury polegała na wprowadzeniu dodatkowej warstwy p^+ przy kontakcie anodowym, której zadaniem jest wstrzykiwanie dziur do niskodmieszkowanego obszaru n i modulacja przewodności tego obszaru. Na rys. 10 przedstawiono rozkład domieszkowania w komórce tranzystora L-IGBT oraz odpowiadający mu elektryczny schemat zastępczy. Chociaż rozkłady domieszek w tranzystorze V-IGBT (rys. 2b) oraz L-IGBT (rys. 10a) są niemal identyczne, ich elektryczne schematy zastępcze przedstawione, odpowiednio, na rys. 4a i 8b są istotnie różne. W tranzystorze lateralnym L-IGBT pojawił się dodatkowy tranzystor p^+np^- , Q_3 , utworzony przez warstwę p^+ anody, niskodmieszkowaną warstwę n oraz warstwę podłożową p^- . Tranzystor ten ma istotny wpływ na własności całego przyrządu.

Pojawienie się dodatkowej struktury tranzystorowej (tranzystor Q_3 w schemacie zastępczym) w sposób istotny wpływa na właściwości tranzystora L-IGBT. Przejmuje ona część prądu obciążenia płynącego pomiędzy anodą i katodą, a tym samym prowadzi do zmniejszenia spadku napięcia na rezystancji R_s . W efekcie zmniejsza się zagrożenie niekontrolowanego załączenia pasożytniczej struktury tyrystorowej i wartość prądu obciążenia, I_L , przy której występuje efekt *latch-up* wzrasta. Wzrost ten może być nawet kilkukrotny.



Rys. 10. Widok komórki tranzystora IGBT (a) oraz odpowiadający jej elektryczny schemat zastępczy (b)

Zmiany konstrukcyjne wprowadzone w tranzystorze L-IGBT prowadzą także do innych zmian w warunkach pracy przyrządu. Mamy w nim przede wszystkim do czynienia z poprzecznym polem elektrycznym w warstwie n wzdłuż górnej powierzchni przyrządu. Natężenie tego pola jest jednak zredukowane przez obecność niskodominowanej warstwy podłożowej p^- . Warstwa ta jest połączona elektrycznie z kontaktem katodowym. Przy odpowiedniej różnicy w domieszkowaniu warstw n i p^- powoduje to, że znaczna część ładunku przestrzennego wystąpi w warstwie p^- i pole elektryczne będzie miało w znacznym stopniu charakter wertykalny a nie lateralny. Dodatkowo, warstwa n, która nie jest teraz wyłącznie odpowiedzialna za ostateczną wartość napięcia blokowania, może być wyżej domieszkowana i cieńsza niż to ma miejsce w tranzystorze V-IGBT. W efekcie można oczekiwać, że czas wyłączenia tranzystora L-IGBT będzie mniejszy.

Modyfikacje konstrukcji

Tranzystor L-IGBT może być wykonany zarówno jako przyrząd dyskretny jak i jako fragment większego układu klasy „*smart power*” wykonanego w technologii hybrydowej lub technologii układów scalonych mocy PIC (*Power Integrated Circuit*). Z uwagi na jego planarno-lateralną strukturę jak i identyczną z elementami L-VDMOS technologię wytwarzania (konwencjonalny proces CMOS), to drugie zastosowanie L-IGBT jest obecnie dominujące i o takich zastosowaniach tych elementów mówi się głównie w literaturze [31–34]. Praktycznie wszystkie rozwiązania zmierzające do polepszenia parametrów tranzystorów V-IGBT, które przedstawiono w rozdz. 2.1.2, posiadają swoje odpowiedniki w modyfikacjach wprowadzanych do konstrukcji tranzystorów L-IGBT. Istnieją jednak również rozwiązania specyficzne tylko dla tego typu tranzystora i one właśnie zostaną przedstawione w dalszej części tego rozdziału.

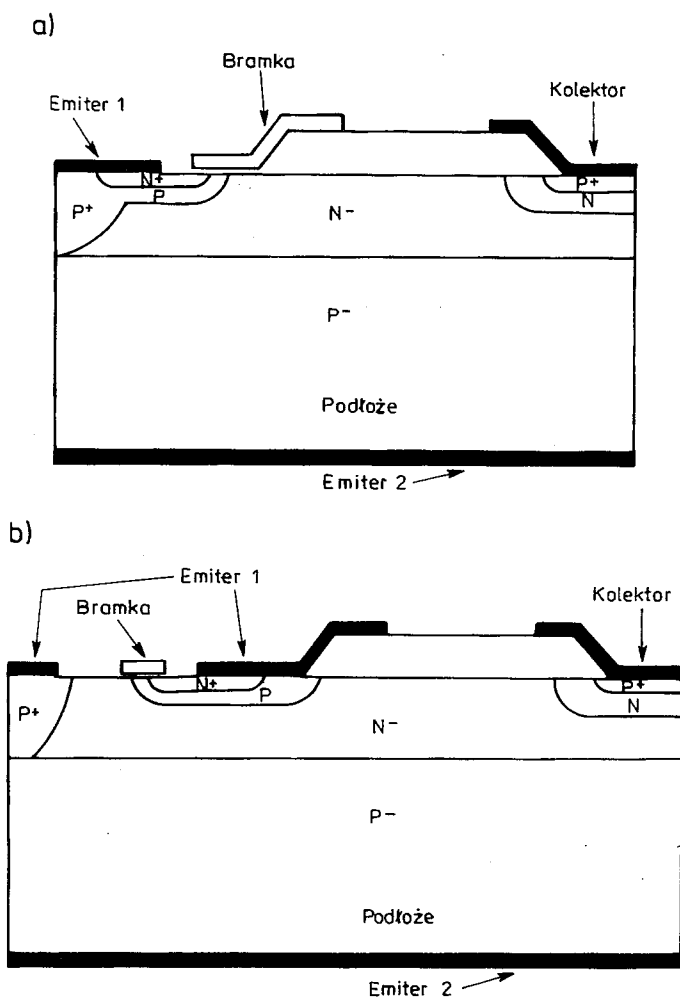
- **struktura z oddaloną bramką**

Konfiguracja lateralna tranzystora L-IGBT pozwala na wprowadzenie jeszcze jednej modyfikacji zmierzającej do ograniczenia wpływu efektu *latch-up*, która jest możliwa jedynie w tej strukturze. Istota tej modyfikacji, pokazana na rys. 11, polega na zastąpieniu miejscami kontaktu bramki i kontaktu emitera, tak aby kontakt emitera sąsiadował z kontaktem kolektora. W efekcie strumień dziur wstrzykiwany przez złącze anodowe (kolektor) może dotrzeć bezpośrednio do katody (emitera) nie wywołując spadku napięcia na rezystancji szeregowej pod wysepką n^+ źródła wejściowego tranzystora MOS. Rozwiązanie to daje kilkakrotny wzrost wartości prądu I_L , przy której wystąpi efekt *latch-up*.

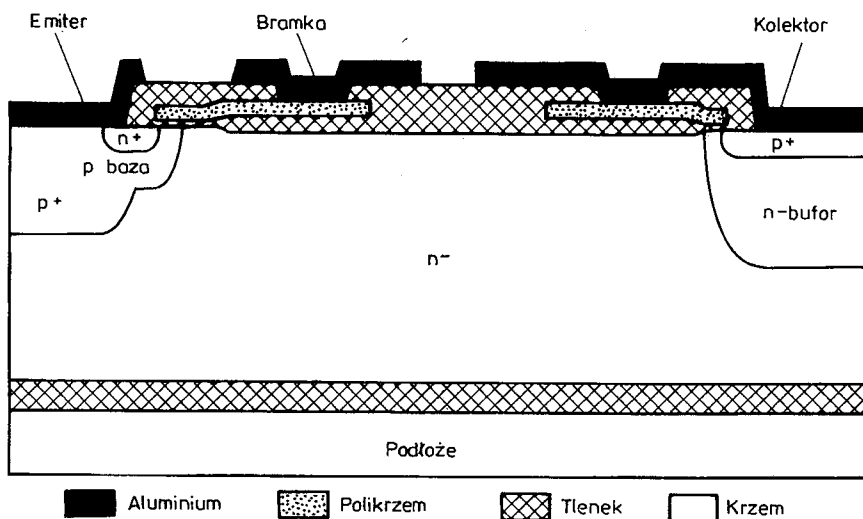
- **struktura SOI**

Technologia SOI (*Silicon On Insulator*) jest obecnie traktowana jako jedna z najbardziej obiecujących metod wytwarzania zintegrowanych układów mocy PIC zawierających w sobie zarówno niskonapięciowe elementy logiczne jak i wysokonapięciowe elementy wykonawcze mocy. Dielektryczna izolacja cienkiej aktywnej warstwy półprzewodnikowej od podłoża pozwala na łatwe oddzielenie od siebie poszczególnych elementów układu, eliminuje pasożytnicze struktury bipolarne i pozwala na pracę tych układów w wyższych temperaturach.

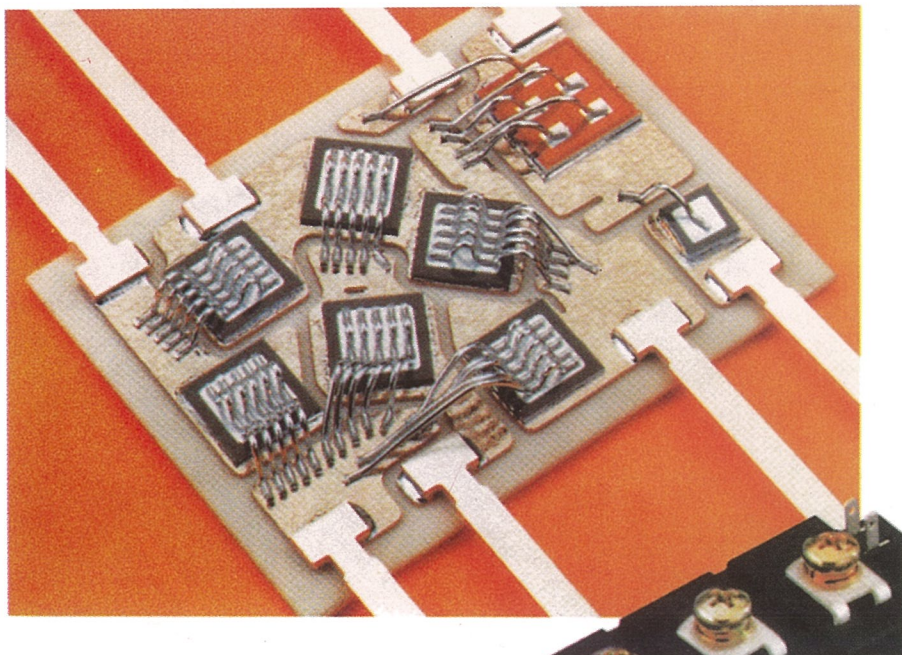
Z samej definicji układów SOI, w technologii tej mogą być praktycznie wykonywane jedynie lateralne przyrządy półprzewodnikowe. Parametry tych przyrządów często różnią się jednak od parametrów analogicznych przyrządów wykonywanych w innych technologiach. Sytuacja taka występuje także w przypadku tranzystora L-IGBT wykonanego w technologii SOI. Na rys. 12 przedstawiono schematycznie przekrój przez strukturę tego tranzystora. Najważniejszą różnicą jest brak warstwy p^- tworzącej kolektor tranzystora Q_3 występującego na schemacie zastępczym przedstawionego na rys. 10b.



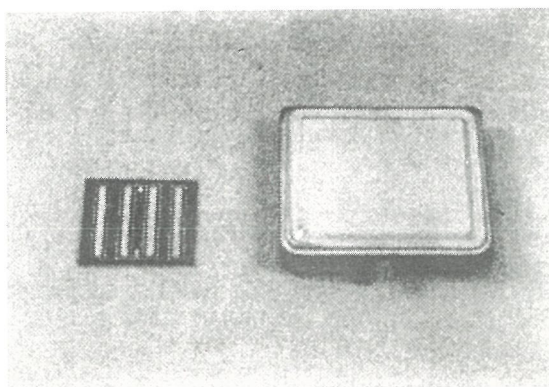
Rys. 11. Struktura konwencjonalna tranzystora L-IGBT (a) i odpowiadająca jej struktura tranzystora L-IGBT z oddaloną bramką (b) [36]



Rys. 12. Struktura tranzystora L-IGBT wykonana w technologii SOI [37]



Rys. 13. Widok połączeń lutowanych w typowym tranzystorze IGBT



Rys. 14. Tranzystor IGBT firmy FUJI w obudowie „press pack” — pastylka półprzewodnikowa oraz tranzystor w obudowie [38]

3. KONSTRUKCJA MECHANICZNA

Sposób zamykania tranzystorów IGBT w obudowy jest na tyle istotny dla ich własności, że wymaga odrębnego omówienia. Tranzystory IGBT mocy są wykonywane w postaci modułów elektroizolowanych lub jako elementy dyskretne. W obu przypadkach stosowana jest obecnie ta sama technologia łączenia wywodząca się z praktyki montażu elementów polowych. Pastylka półprzewodnikowa jest przytwierdzana bezpośrednio do podłoża w procesie „*direct copper bonding*”, natomiast pola kontaktowe pastylki są łączone z polami kontaktowymi obudowy równoległymi wiązkami przewodów w procesie termokompresji lub lutowania ultradźwiękowego.

Z uwagi na „subtelną” strukturę emitera nie jest możliwe wykonywanie tych tranzystorów jako przyrządów o dużej powierzchni. Pojedyncze tranzystory IGBT są wykonywane zwykle w postaci prostokątnych „*chipów*” o maksymalnych wymiarach nie przekraczających 13×13 mm. Nominalny prąd przewodzenia takiego pojedynczego tranzystora nie przekracza 75–100 A, a dla sterowania większymi prądami trzeba je łączyć równolegle budując odpowiednie moduły. Do każdego pojedynczego tranzystora dochodzi około 8–12 aluminiowych przewodów dużej średnicy, rzędu $0.3 \div 0.7$ mm, łączących jego pola kontaktowe z polami kontaktowymi obudowy. Powoduje to, że np. w module zaprojektowanym na 400 A, który zawiera co najmniej 4 połączone równolegle tranzystory, całkowita liczba połączeń przewodowych przekracza setkę. Tak duża ilość połączeń lutowanych wykonanych z grubego drutu aluminiowego i przewodzących duże gęstości prądu (rzędu $50 - 100$ A/mm²) stanowi nową jakość sama w sobie, a ich prawidłowe wykonanie ma istotne znaczenie dla niezawodności przyrządu. Na rys. 13 pokazano widok wewnętrznych połączeń lutowanych w typowym tranzystorze IGBT.

Rozwiązanie przedstawione na rys. 13 jest rozwinięciem koncepcji stosowanych przy produkcji tranzystorów polowych mocy posiada szereg wad, istotnych w przypadku przyrządów przewodzących duże prądy. Podstawowe z nich to jednostronne chłodzenie oraz konieczność wykonywania złożonej sieci połączeń przewodowych w celu wyprowadzenia prądu z pastylki półprzewodnikowej do zewnętrznej elektrody przyrządu. Połączenia te najszybciej ulegają uszkodzeniu z powodu „zmęczenia termicznego”, co w sposób istotny wpływa na niezawodność całego przyrządu [35].

Problemy te zostały już dawno rozwiązane z powodzeniem w tradycyjnych przyrządach bipolarnych mocy poprzez wprowadzenie typowej dla nich obudowy „*press pack*”. Ostatnio podjęto próby zaadoptowania tego typu obudów do tranzystorów IGBT w wersji czysto tranzystorowej [38] oraz w wersji tranzystora wstecznie przewodzącego RC-IGBT, zawierającej zestaw połączonych równolegle struktur tranzystorów IGBT i diod zwrotnych w typowej obudowie pastylkowej [39] lub w specjalnie opracowanej obudowie prostokątnej [40]. Na rys. 14 pokazano widok zewnętrzny takiego tranzystora wykonanego przez firmę FUJI. Według zapowiedzi producentów, należy oczekiwać pojawienia się na rynku tak wykonanych tranzystorów IGBT w pierwszej połowie roku 1996.

4. PODSUMOWANIE

Tranzystory IGBT są obecnie najbardziej obiecującymi i najintensywniej rozwijanymi przyrządami półprzewodnikowymi mocy. W przeciwieństwie do poprzedzających je przyrządów półprzewodnikowych mocy nie są one „powiększoną” wersją przyrządu, który wcześniej powstał jako przyrząd małej mocy. Pierwsze tranzystory IGBT były od razu budowane jako elementy mocy i stanowiły zmodyfikowaną wersję polowych tranzystorów mocy VDMOS. W tranzystorach tych, w wyniku prostego zabiegu zmiany typu materiału podłożowego, wprowadzono dodatkowy tranzystor bipolarny tworząc z przyrządu czysto polowego przyrząd Bi-MOS. Postępowanie takie było racjonalne z ekonomicznego punktu widzenia chociaż zarówno możliwy zakres zastosowań jak i warunki pracy obu typów przyrządów znacznie się różnią. Ten rodowód jest zapewne jedną z przyczyn, że niemal natychmiast po pojawieniu się pierwszych konstrukcji, rozpoczął się proces ich modyfikacji. Proces ten trwa do tej pory i jest tak intensywny, że praktycznie co 3–4 lata pojawia się nowa generacja tranzystorów IGBT. W Polsce przyrządu tego typu nie są obecnie produkowane, aczkolwiek są czynione w tym kierunku pewne przymiarki [41].

W pracy przedstawiono genezę tranzystora IGBT oraz omówiono oddzielnie drogi rozwojowe dwóch odmian tego przyrządu — tranzystora V-IGBT (vertical IGBT), o kontaktach emitera i kolektora ulokowanych po przeciwległych płaszczyznach płytki półprzewodnikowej, oraz tranzystora L-IGBT (lateral IGBT), o kontaktach emitera i kolektora umieszczonych po tej samej stronie pastylki półprzewodnikowej.

Tranzystory V-IGBT pochodzą od tranzystora mocy VDMOS, w którym elektrody drenu i źródła są także ulokowane po przeciwległych stronach pastylki półprzewodnikowej. Podstawowymi problemami jakie należy rozwiązać modyfikując jego konstrukcję jest uniknięcie efektu *latch-up* związanego z niekontrolowanym załączeniem pasożytniczej struktury tyrystorowej oraz zapewnienie odpowiednio dużego napięcia blokowania i małego napięcia przewodzenia. Większość rozwiązań szczegółowych, opisanych w rozdziale „Rozwój konstrukcji” służy właśnie temu celowi. Istotną cechą tranzystorów V-IGBT, stosowanych jako elementy mocy, jest konieczność łączenia równoległego szeregu struktur w jednym module, który może zawierać nawet kilkaset połączeń lutowanych. Ma to istotny wpływ na jego niezawodność i dlatego prowadzone są prace nad wprowadzeniem modułów zamykanych w obudowach typu „*press pack*” z dwustronnym dociskowym kontaktem zewnętrznym.

Tranzystory L-IGBT pochodzą od tranzystora LDMOS, w którym elektrody drenu i źródła znajdują się na górnej powierzchni, a dolna powierzchnia jest połączona z podłożem. Tranzystory te mogą być wykonane zarówno jako przyrządy dyskretne jak i jako fragmenty większego układu klasy „*smart power*” wykonanego w technologii hybrydowej lub technologii układów scalonych mocy PIC (*Power Integrated Circuit*). Obecnie największe znaczenie mają te dwa ostatnie zastosowania. Szczególne znaczenie mogą tu mieć tranzystory L-IGBT wykonane w technologii SOI (*Silicon On*

Insulator). Technologia ta jest obecnie traktowana jako jedna z najbardziej obiecujących metod wytwarzania zintegrowanych układów mocy PIC zawierających w sobie zarówno niskonapięciowe elementy logiczne jak i wysokonapięciowe elementy wykonawcze mocy.

BIBLIOGRAFIA

1. J. Nishizawa, T. Terasaki, J. Shibata: *Field effect transistor versus analog transistor (static induction transistor)*. IEEE Trans. Electron Dev., vol. ED-22, no. 4, 1975, ss. 185—195
2. J. Nishizawa: *Application of the power state induction (SI) devices*. Proc. PCIM'88, pp. 1—12.
3. B.W. Wessels, B.J. Baliga: *Vertical channel field control thyristors with high gain and fast switching speeds*. IEEE Trans. Electron Dev., vol. ED-25, no. 10, 1978, ss. 1261—1265
4. B.J. Baliga, M.S. Adler, R.P. Love, P.V. Gray, N.D. Zommer: *The insulated gate transistor. A new three terminal MOS-controlled bipolar device*. IEEE Trans. Electron Dev., vol. ED-31, no. 6, 1984, ss. 821—828
5. R. Blanchard, S. Hayne: *Power MOS transistors: structure & performance*. Power Conv. Int., 1980, vol. 6, no. 2, ss. 43—49
6. A. Nakagawa, J. Yoshida, T. Utagawa, T. Tsukakoshi, H. Tanabe, T. Kuramoto: *High voltage low on resistance VDMOS FET*. Jap. J. Appl. Phys., 1982, vol. 21, Suppl. 21—1, pp. 97—101
7. T. Sakai, N. Murakami: *A new VDMOSFET structure with reduced reverse transfer capacitance*. IEEE Trans. Electron Dev., vol. 36, 1989, ss. 1381—1386
8. Y. Sasada, S. Matsubagashi: *New second-generation IGBT*. Proc. Power Conversion, 1993, ss. 107—116
9. F. Blaabjerg, J. Pedersen, U. Jaeger: *A critical evaluation of modern IGBT-modules*. Proc. EPE'95 Conf., Sevilla 1995, pp. 1.594—1.601
10. Y. Nakaguni, T. Otami: *The third-generation IGBT and its model for circuit simulation*. Proc. Power Conversion, 1993, pp. 223—227
11. J.P. Russell, A.M. Goodman, L.A. Goodman, J.M. Neilson: *the COMFET — a new high conductance MOS-gated device*. IEEE Electron. Dev. Lett. 1983, vol. EDL-4, pp. 63—65
12. B.J. Baliga: *Fast-switching insulated gate transistors*. IEEE Electron. Dev. Lett., 1983, vol. EDL-4, ss. 452—454
13. M. Darwish, K. Board: *Lateral resurfed COMFET*. Electron. Lett., 1984, vol. 20, pp. 519—520.
14. H.R. Chang, B.J. Baliga: *500-V n-channel insulated-gate bipolar transistor with a trench gate structure*. IEEE Trans. Electron Dev., 1989, vol. 36, pp. 1824—1828
15. B.R. Pelly: *Power semiconductor devices — a status review*. Int. Semicond. Power Converter Conf., 1982
16. J.S. Ajit, B.J. Baliga, S. Tandon, A. Reisman: *Comparison of MOS-gated bipolar transistor structures*. EPE-MADEP Conf., Firenze 1991, pp. 148—151
17. T. Tanaka, Y. Yasuda, M. Ohayashi: *A new MOS-gate bipolar transistor for power switches*. IEEE Trans. Electron. Dev., 1986, vol. ED-33, pp. 2041—2045
18. P. Godignon, J. Fernandez, S. Hidalgo, F. Bertau, J. Rebollo, J. Millan: *Analysis of vertical insulated base transistor*. EPE-MADEP Conf. Firenze 1991, ss. 215—219
29. P. Aloisi: *Insulated gate bipolar transistor family*. Proc. EPE'95, Conf., Sevilla 1995, pp. 1.608—1.613
20. A. Nakagawa, H. Ohashi, K. Muramoto: *Non-latchup bipolar mode MOSFETs with large current capability*. Toshiba Rev., 1985, no. 152, pp. 27—29

21. A. Nakagawa, H. Ohashi: *600- and 1200-V bipolar-mode MOSFET's with high current capability*. IEEE Electron. Dev. Lett., 1985, vol. EDL-6, pp. 378–380
22. A. Nakagawa, Y. Yamaguchi, K. Watanabe: *Experimental and numerical study of nonlatch-up bipolar-mode MOSFET characteristics*. Proc. IDEM 1985, Washington 1985, pp. 150–153
23. W.A. Striffler: *Comparison of neutron and electron irradiation for control IGT switching speed*. IEEE Trans. Electron Dev., 1985, vol. ED-32, pp. 1629–1632
24. M.K. Johnson, A.D. Annis, J.N. Sandoe, D. Coe: *An analysis of the dynamic behaviour of field-limiting ring-passivation systems*. IEEE Trans. Electron Devices, 1989, vol. 36, ss. 1203–1211
25. D. Jaume, G. Charitat, J.M. Reynes, P. Rossel: *High-voltage planar devices using field plate and semi-resistive layers*. IEEE Trans. Electron Dev., 1991, vol. 38, pp. 1681–1684
26. H. Brunner, T. Laska, A. Prost: *IGBT cell optimization with simulation and comparison with measurement*. Proc. EPE'95, Sevilla, 1995, pp. 1.167–1.170
27. Y. Shimizu, Y. Nakano, N. Sakurai, Y. Sugawara, S. Otaka: *A high performance intelligent IGBT with overcurrent protection*. Proc. ISPDSD, 94, Davos 1994, pp. 37–44
28. Y. Seki, Y. Harada, N. Iwamuro, N. Kumagai: *A new IGBT with a monolithic over-current protection circuit*. Proc. ISPD'94, Davos 1994, pp. 31–35
29. Z. Shen, K.C. So, T.P. Chow: *Comparative study of integrated current sensors in n-channel IGBTs*. Proc. ISPD'94, Davos 1994, pp. 75–80
30. S.P. Robb, A. Taomoto, S.L. Tu: *Current sensing in IGBT's for short-circuit protection*. Proc. ISPD'94, Davos 1994, pp. 81–85
31. T.P. Chow, D.N. Pattanayak, B.J. Baliga, M.S. Adler, W.A. Hennessy, C.E. Logan: *Interaction between monolithic, junction-isolated lateral insulated-gate bipolar transistors*. IEEE Trans. Electron Dev., 1991, vol. 38, pp. 310–315
32. G.Y. Tsui, P.V. Gilbert, S.W. Sun: *A versatile half-micron complementary BiCMOS technology for microprocessor-based smart power applications*. IEEE Trans. Electron Dev., 1995, vol. 42, pp. 564–570
33. K. Endo, Y. Baba, Y. Udo, M. Yasui, Y. Sano: *A 500 V 1 A 1-chip inverter IC with a new electric field reduction structure*. ISPD'94, Davos 1994, pp. 379–383
34. S.L. Wong, N. Majid: *A single-chip pre-regulator circuit using IGBT and current mode sensing*. Proc. ISPD'94, Davos 1994, pp. 385–389
35. G. Coqueny, R. Lallemand, D. Wagner, P. Gibard: *Reliability of the 400 A IGBT modules for traction converters contribution on the power thermal fatigue influence*. Proc. EPE'95, Sevilla 1995, pp. 1.060–1.065
36. T.P. Chow: *A reverse-channel, high-voltage lateral IGBT*. Proc. ISPDSD'94, Davos 1994, pp. 57–61.
37. R. Constapel, J. Korec: *Forward blocking characteristics of SOI power devices at high temperatures*. Proc. ISPDSD'94, Davos 1994, ss. 117–121
38. Y. Takahashi, T. Koga, H. Kirihata, Y. Seki: *2.5 kV 100 A μ -stack IGBT*. Proc. ISPDSD'94, Davos 1994, pp. 25–29
39. M. Hiyoshi, S. Yanagisawa, K. Nishtani, K. Kotake, H. Matsuda, S. Teramae, Y. Baba: *A 1000 A 2500 V pressure mount RC-IGBT*. Proc. EPE'95, Sevilla 1995, pp. 1.051–1.055
40. Y. Seki, Y. Takahashi, T. Koga, M. Ichijo, H. Kirihata: *Power pack IGBT: high power (2.5 kV, 1 kA) RC-IGBT with highly reliable flat package*. Proc. EPE'95, Sevilla 1995, (late paper)
41. Z. Lisik: *Możliwości uruchomienia produkcji tranzystorów IGBT w ZE LAMINA*. Raport opracowany na zamówienie ZE LAMINA, 1995
42. K. Reinmuth, H. Stut, L. Lorenz: *Intelligent power modules for driving systems*. Proc. ISPDSD'94, Davos 1994, ss. 93–97

Z. LISIK

DEVELOPMENT OF IGBT TRANSISTORS DESIGN

S u m m a r y

In the paper, the origin of IGBT transistors as well as the ways for development of the IGBT transistor design are presented. Two basic types of these transistors are distinguished and discussed separately — vertical IGBT (V-IGBT) in which emitter and collector contacts are on opposite sides of the chip, and lateral IGBT (L-IGBT) in which the both contacts are on the same side of the chip. Both the IGBT transistors are made from appropriate types of DMOS transistors by changing the type of substrate material. As result, a bipolar transistor is introduced into the structure and the devices become the Bi—MOS ones. The modifications leading to improve the features of these devices are described in paper.

Key words: transistor V-IGBT, transistor L-IGBT, transistors DMOS, devices Bi-MOS.



Matematyczny model wielowarstwowego czujnika piezoceramicznego

DIONIZY CZEKAJ, ZYGMUNT SUROWIAK

Instytut Problemów Techniki, Uniwersytet Śląski

ALEKSANDER V. GORISH, ANATOL A. KUPRIENKO, ANATOL E. PANICH

Zakład „PIEZOPRIBOR”, Uniwersytet Rostowski, Rostów nad Donem

Otrzymano 1996.02.20

Autoryzowano do druku 1996.06.21

Ferroelektryczne materiały ceramiczne charakteryzują się silnym efektem piezoelektrycznym i dlatego znajdują szerokie zastosowanie w elektronice jako czujniki (sensory) i przetworniki elektromechaniczne. Czujniki piezoceramiczne mają najczęściej postać układów typu „sandwich”, w których przeplatają się warstwy piezoelektryka oraz metalicznych elektrod. W niniejszej pracy zaproponowano matematyczny model takiego sensora. Na podstawie tego modelu można oceniać i przewidywać zachowanie się sensorów piezoceramicznych w różnych warunkach eksploatacyjnych, w tym również w obecności działających na niego pól elektrycznych, mechanicznych i cieplnych.

Słowa kluczowe: ferroelektryki, czujniki piezoceramiczne, efekt piezoelektryczny, układy wielowarstwowe.

1. WPROWADZENIE

W układach pomiarowych i kontrolnych stosowanych w nowoczesnej technice (w przemyśle lotniczym, samochodowym, raketowym itd.) szczególne miejsce zajmują piezoelektryczne czujniki (sensory) ceramiczne i cienkowarstwowe [np. 1–7]. Rezultaty analizy różnorodnych programów badawczych z zakresu metrologii oraz programów modernizacji różnych dziedzin techniki wskazują, że piezoelektryczne czujniki na bazie materiałów ferroelektrycznych stanowią ponad 30% ogólnej liczby stosowanych sensorów i udział ich stale wzrasta [8].

Rodzaje parametrów fizycznych mierzonych za pomocą czujników piezoelektrycznych obejmują szeroki zakres wielkości mechanicznych i elektrycznych [1–2]. Stosuje się je między innymi do pomiaru szybkozmiennych ciśnień, odkształceń, przyspieszeń i wibracji oraz charakterystyk uderu, siły i przepływu. Czujniki piezoceramiczne posiadają wyjątkowo doskonale charakterystyki eksploatacyjne, są wytrzymałe na

cykliczne obciążenia i deformacje, cechują się prostą budową, są bezawaryjne i ekonomiczne (tanie), mają małe wymiary oraz nie wymagają zewnętrznych źródeł energii (źródeł zasilania).

Czujniki takie mogą pracować w bardzo trudnych warunkach eksploatacyjnych (w szerokich zakresach temperatur, wibracji, szumów akustycznych, liniowych przyspieszeń, mechanicznych i hydraulicznych uderów przy dużych gradientach temperatury oraz w agresywnych i kriogenicznych środowiskach), przy czym najczęściej działa na nie równocześnie kilka fizycznych czynników destabilizujących. Z zasady czujnik piezoceramiczny powinien charakteryzować się wytrzymałością mechaniczną większą niż wytrzymałość mechaniczna węzła, w którym został on zamocowany. W celu znalezienia granicznych obciążeń w anomalnych warunkach czujnik piezoceramiczny powinien „żyć” dłużej (być trwalszy) niż konstrukcja danego urządzenia (maszyny, agregatu, rakiety itp.). Ponadto czujnik piezoceramiczny charakteryzować się musi stabilnością właściwości w czasie całego okresu eksploatacji. Spełnienie tych wymagań wiąże się przede wszystkim z bezawaryjną pracą piezoelektrycznego elementu sensora. Element piezoelektryczny sensora stanowi wielowarstwowy monolityczny moduł ceramiczny składający się z kolejno przeplatających się warstw piezoaktywnych z cienkimi warstwami przewodzącymi (elektrodami platynowymi) [8]. Biorąc pod uwagę zarówno szerokie spektrum możliwych oddziaływań zewnętrznych na czujnik piezoceramiczny jak i możliwych reakcji elementu piezoelektrycznego na dowolny typ wymienionych wyżej oddziaływań, podstawowym problemem dla konstruktorów (elektroników, fizyków) jest wydzielenie sygnału użytecznego i zmniejszenie błędu pomiarów.

Autorzy przedstawili już wcześniej konstrukcję i własności elektrofizyczne izotropowych i anizotropowych sensorów piezoelektrycznych na bazie ferroelektrycznych cienkich warstw BaTiO_3 i LiNbO_3 [3] oraz $\text{Pb}(\text{Zr}, \text{Ti})\text{O}_3$ czyli PZT [1–2] a także sensorów piezoceramicznych na bazie PZT [5, 7]. Charakterystyki eksploatacyjne skonstruowanych sensorów badano zarówno w warunkach laboratoryjnych jak i przemysłowych.

Celem niniejszej pracy jest opracowanie matematycznego modelu ceramicznego elementu piezoelektrycznego i otrzymanie na jego podstawie analitycznego rozwiązania dla opisu elektrofizycznych charakterystyk wielowarstwowego czujnika piezoelektrycznego. Model ten uwzględnia wszystkie przedstawione wyżej zasadnicze wymagania stawiane tego typu sensorom. Opisuje z dużą dokładnością wszystkie istotne charakterystyki ceramicznego elementu piezoelektrycznego, pozwala określić wkład najważniejszych zjawisk fizycznych w funkcję przetwarzania elektromechanicznego, zapewnić możliwość badania różnych trybów pracy sensora i przewidywania najbardziej optymalnych warunków jego eksploatacji.

2. MODEL TEORETYCZNY

Rozpatrzony zostanie model piezoceramicznego czujnika wielowarstwowego w postaci systemu naprzemiennych warstw piezoelektrycznych i metalicznych o grubościach odpowiednio l_1 i l_2 . W dalszej części pracy wygodniej będzie operować pojęciami względnej grubości warstw oraz ich objętościową koncentracją (m_i):

$$m_i = l_i / 2l, \quad (1)$$

gdzie:

$2l$ — okres układu warstw ($2l = l_1 + l_2$).

Założono, że częstotliwość zewnętrznej siły działającej na czujnik jest mniejsza od częstotliwości podstawowego rezonansu, co oznacza że pracuje on w reżimie quasi-statycznym. Ponieważ prędkość dźwięku (v) jest znacznie mniejsza od prędkości światła (c) więc elektryczny podukład podczas drgań elementu piezoelektrycznego zdąży dostroić się do sprężystego podukładu i uwzględnienie dynamiki elektrycznego podukładu prowadzi tylko do względnej zmiany prędkości dźwięku w piezoceramice o wartości v^2/c^2 [9]. Uwzględniając tę wartość przybliżenie elektrostatyczne można uznać za dopuszczalne z dokładnością do $10^{-4}\%$. W konsekwencji jako główne i jedyne założenie modelu przyjęto, że okres układu warstw ($2l$) jest znacznie mniejszy od poprzecznych rozmiarów czujnika w płaszczyźnie rozdziału warstw. Jeśli oznaczyć przez R_0 promień elementu piezoceramicznego w płaszczyźnie rozdziału warstw, to warunek jednowymiarowości zadania przyjmie postać następującą:

$$2l/R_0 \ll 1. \quad (2)$$

Ponieważ rozpatrywany element piezoceramiczny przeznaczony jest do pomiaru siły, ciśnienia, przyspieszenia itd. przy jednoczesnym oddziaływaniu na niego pola temperatur, więc warstwy piezoelektryczne powinny być opisane równaniami stanu dla termoelektrosprężystego układu anizotropowego [10]:

$$\begin{aligned} D_i &= \xi_{ij}^{\sigma, \theta} E_j + d_{i\alpha}^{\theta} \sigma_{\alpha} + p_i^{\sigma} \theta, \\ \xi_{\beta} &= d_{k\beta}^{\theta} E_k + S_{\alpha\beta}^E \sigma_{\alpha} + \alpha_{\beta}^E \theta, \end{aligned} \quad (3)$$

$$\Sigma = p_i^{\sigma} E_i + \alpha_{\beta}^E \sigma_{\beta} + \frac{\rho C_p^E}{T} \theta,$$

gdzie:

$i, j = 1, 2, 3; \alpha, \beta = 1, 2, \dots, 6;$

D_i — składowe wektora indukcji elektrycznej;

E_j — składowe wektora natężenia pola elektrycznego;

θ — zmiana temperatury ($\theta = T - T_0$, gdzie T — bieżąca temperatury elementu piezoceramicznego, zaś T_0 — temperatura początkowa);

Σ — zmiana entropii ($\Sigma = S - S_0$, gdzie S — bieżąca entropia układu, zaś S_0 — entropia stanu początkowego);

σ_{α} — składowe sensora naprężeń mechanicznych (zapis macierzowy: $\alpha = 1, \dots, 6$);

ξ_{β} — składowe tensora deformacji sprężystych (zapis macierzowy: $\beta = 1 \dots 6$);

$\varepsilon_{ij}^{\sigma, \theta}$ — składowe tensora przenikalności elektrycznej, mierzonej przy $\sigma = \text{const}$ i $\theta = 0$;

$d_{i\alpha}^{\theta}$ — składowe tensora modułów piezoelektrycznych mierzonych w warunkach izotermicznych ($\theta = 0$);

$S_{\alpha\beta}^E$ — składowe tensora podatności sprężystej (stałych sprężystości), mierzonej przy $E = \text{const}$ i $\theta = 0$;

- p_i^σ — współczynnik piezoelektryczny mechanicznie swobodnego kryształu, opisujący całkowity efekt piezoelektryczny badany przy $\sigma=0$;
 α_p^E — składowe tensora współczynników rozszerzalności cieplnej, mierzonej przy $E=\text{const}$;
 C_p^E — ciepło właściwe mierzone przy stałym ciśnieniu ($p=\text{const}$) i w stałym polu elektrycznym ($E=\text{const}$);
 ρ — gęstość materiału piezoelektrycznego.

Odpowiednie równania stanu opisujące środowisko izotropowe (elektrody metaliczne) mają postać następującą [11]:

$$\begin{aligned}
 D_m &= \varepsilon E_m, \\
 \xi_{ij} &= [(1+\nu)\sigma_{ij} - \nu\sigma_{kk}\delta_{ij}]/E^{-1} + \alpha\delta_{ij}\theta, \\
 \Sigma &= \alpha\delta_{ij}\sigma_{ij} + \frac{\rho_m C_p}{T}\theta,
 \end{aligned} \tag{4}$$

gdzie: $m, i, j = 1, 2, 3$;

- ε — przenikalność elektryczna;
 σ_{ij} — składowe tensora naprężeń mechanicznych (zapis tensorowy);
 ξ_{ij} — składowe tensora deformacji sprężystych (zapis tensorowy);
 ν — współczynnik Poissona;
 E — moduł Younga dla metalu (środowiska izotropowego);
 α — współczynnik rozszerzalności cieplnej metalu;
 ρ_m — gęstość metalu;
 C_p — ciepło właściwe metalu;
 δ_{ij} — symbol Kroneckera (funkcja dwóch całkowitoliczbowych i oraz j , określona następująco: $\delta_{ij}=1$, gdy $i=j$; $\delta_{ij}=0$, gdy $i \neq j$).

Rozpatrzony zostanie teraz wielowarstwowy czujnik piezoceramiczny w układzie współrzędnych kartezjańskich z osią Z prostopadłą do płaszczyzny rozdziału warstw i równoległą do wektora polaryzacji spontanicznej $\vec{P}_s$; osie X i Y leżą w płaszczyźnie rozdziału warstw. Taka polaryzacja elementów piezoelektrycznych określa makroskopową symetrię układu $\infty \cdot m$ i odpowiadające jej równania stanu polarnej tekstury przy uśrednionych wartościach pól elektrycznych, mechanicznych i temperaturowych. Sposób polaryzowania wielowarstwowego czujnika piezoceramicznego może prowadzić do tego, że w sąsiadujących ze sobą warstwach wektory polaryzacji będą skierowane antyrównoległe i w związku z tym makroskopowa symetria całego układu będzie niepolarna. Jednakże odpowiednio skoordynowane podłączenie elektrod, zapewniające równoległe połączenie warstw, stwarza możliwość opisu piezoceramicznego układu w postaci polarnej tekstury o symetrii $\infty \cdot m$. Taki układ pomiarowy najczęściej musi być eksploatowany przy jednoczesnym oddziaływaniu na niego pól elektrycznych, mechanicznych i cieplnych. Zachowanie niejednorodnego wielowarstwowego sensora piezoceramicznego w takich warunkach można opisać za pomocą uśrednionych stałych fizycznych, których wartość zależy od własności warstw składowych i ich grubości (koncentracji).

Uwzględniając prawa dodawania symetrii a także symetrię spolaryzowanej ceramiki $\infty \cdot m$ [10] można zgodnie z (3) przedstawić poszukiwany układ równań stanu termoelektrosprężystości, łączący średnie (tzn. pojawiające się w eksperymencie) mechaniczne naprężenie $\bar{\sigma}_\alpha$, natężenie pola elektrycznego $\bar{E}_k$ i temperaturę $\bar{\theta}$ ze średnią mechaniczną deformacją $\bar{\xi}_\alpha$, elektryczną indukcją $\bar{D}_k$ i entropią $\bar{S}$:

$$\begin{aligned}\bar{D}_1 &= d_{15}^{*\theta} \bar{\sigma}_5 + \varepsilon_{11}^{*\sigma, \theta} \bar{E}_1, \\ \bar{D}_2 &= d_{15}^{*\theta} \bar{\sigma}_4 + \varepsilon_{11}^{*\sigma, \theta} \bar{E}_2, \\ \bar{D}_3 &= d_{31}^{*\theta} (\bar{\sigma}_1 - \bar{\sigma}_2) + d_{33}^{*\theta} \bar{\sigma}_3 + \varepsilon_{33}^{*\sigma, \theta} \bar{E}_3 + p_3^{*\sigma} \bar{\theta}, \\ \bar{\xi}_1 &= S_{11}^{*E, \theta} \bar{\sigma}_1 + S_{12}^{*E, \theta} \bar{\sigma}_2 + S_{13}^{*E, \theta} \bar{\sigma}_3 + d_{13}^{*\theta} \bar{E}_3 + \alpha_1^{*E} \bar{\theta}, \\ \bar{\xi}_2 &= S_{12}^{*E, \theta} \bar{\sigma}_1 + S_{12}^{*E, \theta} \bar{\sigma}_2 + S_{13}^{*E, \theta} \bar{\sigma}_3 + d_{13}^{*\theta} \bar{E}_3 + \alpha_1^{*E} \bar{\theta}, \\ \bar{\xi}_3 &= S_{13}^{*E, \theta} (\bar{\sigma}_1 + \bar{\sigma}_2) + S_{33}^{*E, \theta} \bar{\sigma}_3 + d_{33}^{*\theta} \bar{E}_3 + \alpha_3^{*E} \bar{\theta}, \\ \bar{\xi}_4 &= S_{44}^{*E, \theta} \bar{\sigma}_4 + d_{15}^{*\theta} \bar{E}_2, \\ \bar{\xi}_5 &= S_{44}^{*E, \theta} \bar{\sigma}_4 + d_{15}^{*\theta} \bar{E}_1, \\ \bar{\xi}_6 &= S_{66}^{*E, \theta} \bar{\sigma}_6, \\ \bar{S} &= \alpha_1^{*E} (\bar{\sigma}_1 + \bar{\sigma}_2) + \sigma_3^{*E} \bar{\sigma}_3 + p_3^{*\sigma} \bar{E}_3 + \frac{\rho^* C_p^{*E}}{T^*} \bar{\theta},\end{aligned}\tag{5}$$

gdzie:

gwiazdka (*) — efektywne wartości parametrów materiałowych,

kreska (—) — średnie wartości zmiennych termodynamicznych.

Efektywne wartości składowych tensorów stałych dielektrycznych $\varepsilon_{mk}^{*\sigma, \theta}$, piezoelektrycznych $d_{ia}^{*\theta}$, sprężystych $S_{\alpha\beta}^{*E, \theta}$, piroelektrycznych $p_3^{*\sigma}$, termodynamicznych C_p^{*E} i termosprężystych α_m^{*E} są poszukiwanymi charakterystykami niejednorodnego warstwowego sensora piezoceramicznego. Znając zbiór tych tensorów można opisać zachowanie takiego sensora przy różnego typu warunkach brzegowych, a także przy równoczesnym oddziaływaniu na niego pól elektrycznych, sprężystych i cieplnych. Układ równań (5) może opisywać odpowiedź czujnika na oddziaływanie różnego typu: naprężenia rozciągająco-ściskające (osiowe i polarne), naprężenia ścinające, pole elektryczne i temperatura.

Z ogólnej teorii ośrodków heterogenicznych wiadomo, że w przypadku ośrodków z jednowymiarową niejednorodnością właściwości fizycznych, bardzo dokładnie można określić efektywne charakterystyki [12].

Układ równań opisujących ośrodek termoelektrosprężysty składa się z równań elektrostatyki [13], równań równowagi i warunków deformacji [11, 13], drugiej zasady termodynamiki i równania przepływu ciepła w wyniku przewodnictwa cieplnego [13—14].

Równanie dla wektora strumienia cieplnego $\vec{q}$ związanego z gradientem temperatury (grad θ) poprzez tensor przewodnictwa cieplnego κ_{ik} ma postać [13]:

$$\vec{q} = -\kappa \cdot \text{grad} \theta. \quad (6)$$

Ponieważ równania stanu ośrodka termoelektrosprężystego są słuszne dla procesów odwracalnych [10] niezbędne jest w równaniu przepływu ciepła przyjąć $\partial \vec{q} / \partial t = 0$. Odpowiadać to będzie warunkom „termostatyki” [15].

Z nierówności (2) i odpowiadającej jej jednowymiarowości zadania można wnioskować, że wszystkie zmienne termodynamiczne w równaniach definiujących i w równaniach stanu zależą tylko od współrzędnej z (niejednorodność jednowymiarowa). Oznacza to, że w operatorach różniczkowych układu równań definiujących przyjąć należy:

$$\frac{\partial}{\partial x} = \frac{\partial}{\partial y} = 0; \quad \frac{\partial}{\partial z} \neq 0. \quad (7)$$

Po uwzględnieniu warunków (7) rozwiązanie układu równań definiujących dla jednowymiarowego ośrodka sprowadza się do tego, że wielkości $D_3, E_1, E_2, \sigma_3, \sigma_4, \sigma_5, \xi_1, \xi_2, \xi_6$ i θ nie zależą od współrzędnej z i dlatego są równe swoim wartościom średnim:

$$\begin{aligned} D_3 &= \bar{D}_3, & E_1 &= \bar{E}_1, & E_2 &= \bar{E}_2, & \theta &= \bar{\theta}, & \sigma_3 &= \bar{\sigma}_3, \\ \sigma_4 &= \bar{\sigma}_4, & \sigma_5 &= \bar{\sigma}_5, & \xi_1 &= \bar{\xi}_1, & \xi_2 &= \bar{\xi}_2, & \xi_6 &= \bar{\xi}_6. \end{aligned} \quad (8)$$

Układ równań stanu składa się z 10 równań obejmujących 20 niewiadomych, z których 10 jest już określonych zależnościami (6). Pozostałe 10 niewiadomych składowych pól elektrycznych, sprężystych i cieplnych określone są z układu równań stanu po uwzględnieniu zależności (6). W rezultacie uśrednienia wzdłuż grubości czujnika (okresu $2l$ wielowarstwowego układu) otrzymuje się następujący zestaw efektywnych składowych tensorów fizycznych własności piezoceramicznego elementu

$$\varepsilon_{33}^{*\sigma, \theta} = \frac{\varepsilon_{33}^{\sigma, \theta}}{m_1} \left[1 - \frac{2m_2 k_{31}^2}{m_1(1-\nu)/ES_{11}^{E, \theta} + m_2(1-\nu)S_{12}^E} \right];$$

$$\varepsilon_{11}^* = \infty;$$

$$d_{31}^{*\theta} = d_{31} / [m_1 + m_2(1-\nu)S_{11}^{E, \theta}E / (1-\nu)];$$

$$d_{15}^* = m_1 d_{15};$$

$$d_{33}^* = d_{33} - \frac{2d_{31}(\nu + S_{13}^E E)}{(m_1/m_2)(1-\nu) + (1-\nu)S_{12}^E E};$$

$$S_{11}^{E, \theta} = \frac{S_{11}^E [m_1 + m_2 E (S_{11}^2 - S_{12}^2) / S_{11}^E (1-\nu^2)]}{[m_1 + m_2 (S_{11}^E - S_{12}^E) E / (1-\nu)] \cdot [m_1 + m_2 (S_{11}^E + S_{12}^E) / (1-\nu)]};$$

$$\begin{aligned}
 S_{12}^{*E,\theta} &= \frac{S_{12}^E [m_1 + m_2 \nu E (S_{11}^2 - S_{12}^2) / S_{12}^E (1 - \nu^2)]}{[m_1 + m_2 (S_{11}^E - S_{12}^E) E / (1 + \nu)] \cdot [m_1 + m_2 (S_{11}^E + S_{12}^E) E / (1 - \nu)]}; \\
 S_{33}^{*E,\theta} &= \\
 &= m_1 S_{33}^E + m_2 / E + \frac{2m_1 m_2 (\nu / E + S_{13}^E) [2k_{31}^2 (\nu / E + S_{13}^E) / (1 - \nu_{12}^E) - (\nu / E - S_{13}^E) - 2d_{31} d_{33} / \epsilon_{33}^\sigma]}{(1 - \nu) [m_1 + m_2 (1 - \nu_{12}) S_{11}^E E / (1 - \nu)] \cdot [1 - 2k_{31}^2 / (1 - \nu_{12}^E)]}, \\
 S_{44}^{*E,\theta} &= m_1 S_{44}^E + m_2 / G; \\
 S_{66}^{*E,\theta} &= (S_{66}^E / G) / (m_1 / G + m_2 S_{66}^E); \\
 p_3^* &= p_3^\sigma - \frac{2m_2 d_{31} (\alpha_1 - \alpha) E (1 - \nu)}{m_1 + m_2 (S_{11}^E + S_{12}^E) E / (1 - \nu)}; \quad (9)
 \end{aligned}$$

gdzie:

ν_{12} — współczynnik Poissona piezoceramiki;

$$\nu_{12} = - \frac{S_{12}^{E,\theta}}{S_{11}^{E,\theta}}; \quad (10)$$

k_{31} — współczynnik elektromechaniczny;

G — moduł poprzecznej sprężystości warstwy metalicznej.

PODSUMOWANIE

Otrzymany zestaw efektywnych tensorów materiałowych wielowarstwowego czujnika piezoceramicznego pozwala przeprowadzić obliczenia i optymalizację jego charakterystyk makroskopowych dla różnych par piezoceramika–metal (różne materiały i grubości względne). Rezultaty niniejszej pracy mogą być wykorzystane do oceny możliwości wykorzystania takich sensorów w różnych warunkach eksploatacyjnych (np. w różnych mechanicznych warunkach przymocowania sensora do badanego elementu; przy równoczesnym oddziaływaniu pól mechanicznych, elektrycznych i cieplnych).

Niniejsza praca stanowi teoretyczne ujęcie problemu, a zaproponowany model matematyczny został tylko wstępnie sprawdzony eksperymentalnie na prostym układzie warstwowym. Wynikow tej wstępnej weryfikacji nie zamieszczono jednak w artykule, gdyż będzie ona przedmiotem oddzielnego artykułu. Zaprezentowane w nim będą rozwiązania konstrukcyjne wykorzystujące nieporowatą, drobnoziarnistą ceramikę piezoelektryczną typu PZT.

BIBLIOGRAFIA

1. Z. Surowiak, D. Czekaj, V.P. Dudkevich, A.A. Bakirov: Elektronika, 1994, t. 1, s. 12
2. Z. Surowiak, D. Czekaj, A.A. Bakirov, V.P. Dudkevich: Thin Solid Films, 1995, t. 256, s. 226

3. D. Czekaj, Z. Surowiak, V.P. Dudkevich, A.A. Bakirov: *Akustyka Molekularna i Kwantowa*, 1994, t. 15, s. 43
4. Z. Surowiak, D. Czekaj, A.A. Bakirov, V.P. Dudkevich: Piezoelectric properties of the thin PZT-type films obtained by r.f. sputtering method. In: *Proceedings „8th Piezoelectric Conference PIEZO'94”*, 5–7 October 1994, Zakopane. Tele and Radio Research Institute, Warszawa 1995, p. 289
5. Z. Surowiak, J. Dudek, M. Łoposzko: *Inżynieria Materiałowa*, 1992, t. 6, s. 123
6. Z. Surowiak, D. Czekaj, A.M. Margolin, E.V. Sviridov, V.A. Aleshin, V.P. Dudkevich: *Thin Solid Films*, 1992, t. 214, p. 78
7. Z. Surowiak, J. Dudek, Yu.I. Goltzov, I.A. Bugayan, V.E. Yurkevich: *J. Mater. Sci.*, 1991, t. 26, p. 4407
8. *Sredstva piezoelektricheskovo priborostroeniya. Katalog 90*. RGU, Rostov-na-Donu 1990, s. 12
9. E.Z. Govorova, F.D. Kiselev, M.M. Frisova: *Fiz. Tverd. Tela*, 1995, t. 17, 5, s. 1280
10. D. Berlinkur, D. Kerran, G. Zhaffe: *Piezoelektricheskie i piezomagnitnye materialy i ikh primeneniye v preobrazovatelakh*. W: *Fizicheskaya akustika*, t. 1 (A). Red. U. Mezon, Mir, Moskva 1966, s. 204
11. L.D. Landau, E.M. Lifshic: *Teoreticheskaya fizika*, t. 7. *Teoriya uprugosti*. Nauka, Moskva 1987, s. 248
12. T.D. Shermegor: *Teoriya uprugosti mikroneodnorodnykh sred*. Nauka, Moskva, 1977, s. 400
13. Yu.I. Siroton, M.P. Shaskolskaya: *Osnovy kristalofiziki*. Nauka, Moskva 1979, s. 640
14. A.V. Lykov: *Teoriya teploprovodnosti. Vysshaya shkola*, Moskva 1967, s. 600
15. D. Nai: *Fizicheskie svoistva kristallov*. Mir, Moskva 1967, s. 388

D. CZEKAJ, Z. SUROWIAK, A.V. GORISH, A. KUPRIENKO, A.E. PANICH

MATHEMATICAL MODEL OF THE MULTILAYERED PIEZOELECTRIC SENSOR

S u m m a r y

Ferroelectric ceramic materials are characterized by strong piezoelectric effect. Therefore they have gained widespread application in electronics as sensors and electromechanical transducers. The piezoceramic sensors are most often the „sandwich” type structures, where piezoelectric layers and metallic electrodes alternate each other. In the present paper the mathematical model of such a sensor is proposed. On the basis of the model one can estimate and foresee behaviour of the piezoceramic sensors under different operating conditions, including influence of electrical, mechanical and thermal fields.

Key words: sensor, piezoelectric effect, ceramics, multilayered system.

Implementacja współbieżnych kontrolerów cyfrowych z wykorzystaniem PLD

MAREK WĘGRZYN

*Instytut Informatyki i Elektroniki, Wyższa Szkoła Inżynierska
w Zielonej Górze*

Otrzymano 1996.02.09

Autoryzowano do druku 1996.06.21

Działanie układów cyfrowych, a szczególnie układów sterowania logicznego coraz częściej opisywane jest za pomocą interpretowanych (etykietowanych) sieci Petriego. Umożliwia to przedstawienie współdziałania procesów współbieżnych w bardzo zwarty i przejrzysty sposób. Ze względu na późniejszą implementację układu sterowania korzystne jest rozłożenie (dekompozycja) sieci Petri na moduły (podsieci) składowe o rozmiarach umożliwiających realizację każdej z części w oddzielnych układach scalonych lub określonych elementach konstrukcyjnych (szczególnie układów PLD). Dodatkową korzyścią dekompozycji jest rozbicie złożonego problemu syntezy na kilka łatwiejszych zadań, co prowadzi do znacznego uproszczenia procesu projektowania, przy użyciu znanych, klasycznych już metod syntezy układów sekwencyjnych. W pracy przedstawiono metodę, wykorzystującą kolorowanie sieci Petriego, pozwalającą w stosunkowo prosty sposób uzyskać dekompozycję sieci Petriego. Kolorowanie, opisane w pracy, jest nie tylko narzędziem dekompozycji, ale również analizy i implementacji sieci. Omówiono metodologię syntezy i symulacji automatów współbieżnych z wykorzystaniem standardowych języków takich, jak CUPL, PALASM i VHDL. Uzyskano rezultaty o znacznej wartości teoretycznej i praktycznej (implementacja programowa).

Słowa kluczowe: układy cyfrowe, sieci Petriego, układy PLD, automaty współbieżne.

1. WSTĘP

Działanie układów sterowania binarnego (układów sekwencyjnego sterowania logicznego) coraz częściej opisywane jest za pomocą interpretowanych (etykietowanych) sieci Petriego [1, 2, 8, 10, 15, 18, 21], umożliwiających w zwarty i przejrzysty sposób przedstawienie współdziałania procesów współbieżnych.

Ze względu na późniejszą implementację układu cyfrowego, a w szczególności jego części sterującej korzystna jest dekompozycja sieci Petriego na moduły (podsieci). Mniejsze podsieci umożliwiają realizację każdej z jej części w oddzielnych układach

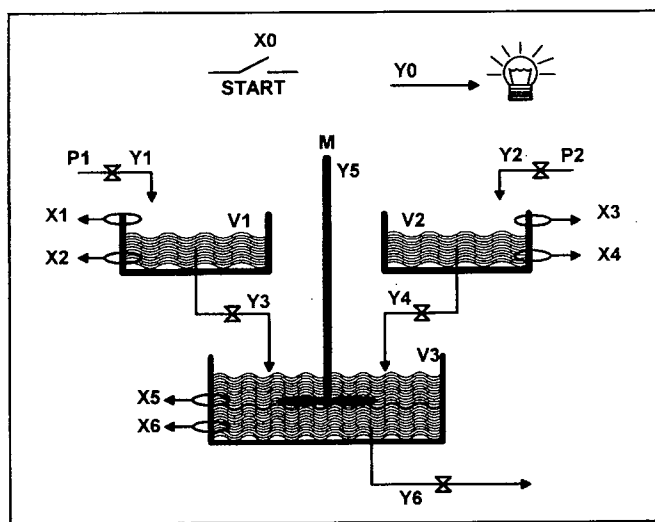
scalonych lub określonych elementach konstrukcyjnych. Dotyczy to szczególnie programowanych struktur logicznych FPLD lub połączonych mikrokontrolerów, albo układów hybrydowych. Dodatkową korzyścią płynącą z dekompozycji jest rozbięcie złożonego problemu syntezy na kilka łatwiejszych zadań. Prowadzi to do znacznego uproszczenia procesu projektowania, często przy użyciu znanych, klasycznych i zalgorytmizowanych metod syntezy układów sekwencyjnych [16]. Stąd duża część znanych z literatury metod dekompozycji prowadzi do rozłożenia sieci Petriego na podsieci składowe, będące odpowiednikami automatów sekwencyjnych.

W pracy przedstawiono sposób dekompozycji sieci Petriego, oparty na kolorowaniu graficznej reprezentacji sieci. Implementacja poszczególnych podsieci automatowych przebiega już automatycznie, np. z wykorzystaniem standardowych systemów automatycznego projektowania elementów PLD takich, jak np. PALASM oraz CUPL [16]. Lepsze możliwości stwarzają uniwersalne systemy CAD z wejściem w języku VHDL, jak np. Cypress albo Altera.

2. MODELOWANIE KONTROLERÓW WSPÓLBIEŻNYCH

Sieć Petriego stanowi dogodne narzędzie modelowania układów, w których pewne zmiany mogą zachodzić równolegle, lecz istnieją pewne ograniczenia co do równoczesności, kolejności i częstotliwości tych zmian. Sieć Petriego opisująca działanie pewnego układu sterującego sekwencją czynności, wykonywanych równolegle jest znacznie prostsza niż opis funkcjonowania tego samego układu za pomocą np. sieci działań, diagramu przejść itp.

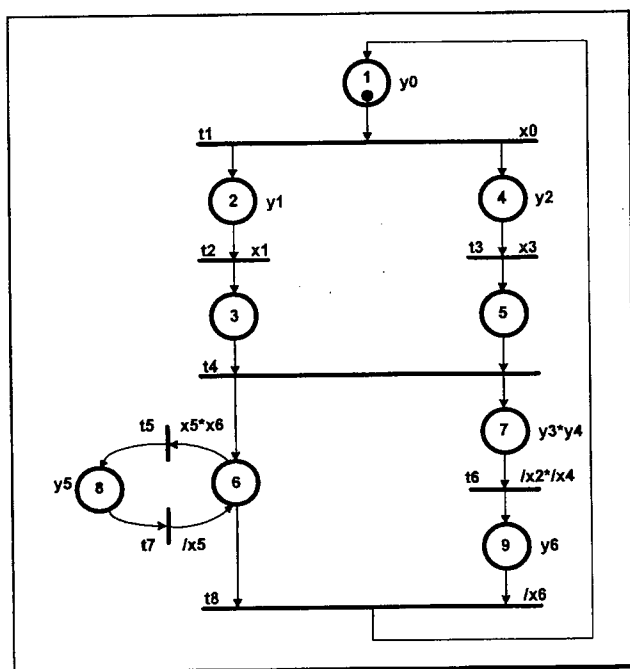
W artykule omawia się interpretowane, warunkowe sieci Petriego [1, 2, 10, 12, 13] opisujące układ sterowania binarnego, posługując się przykładem sterowania zbior-



Rys. 1. Przykład układu sterowania

nikami płynów zaczerpniętym z prac [4, 8]. Wersję sekwencyjną realizacji części sterującej można znaleźć w wielu podręcznikach akademickich, np. w [17]. Schemat blokowy układu przedstawiono na rys. 1. Jest to system trzech zbiorników, w którym zbiorniki V1 i V2 zapelniane są płynnymi substratami, które następnie są mieszane w zbiorniku V3. Odpowiednie sygnały wejściowe ($X_1, \dots, X_6$) reprezentują stany czujników wypełnienia poszczególnych zbiorników. Sygnały wyjściowe ($Y_1, \dots, Y_4, Y_6$) sterują odpowiednimi zaworami, natomiast sygnał wyjściowy Y_5 steruje mieszaniem cieczy w zbiorniku V3. Sygnał Y_0 wskazuje gotowość urządzenia do pracy.

Sieć Petriego opisująca funkcjonowanie układu sterującego została przedstawiona na rys. 2.



Rys. 2. Interpretowana sieć Petriego

2.1. REGUŁOWY OPIS SIECI PETRIEGO

Do symbolicznego opisu sieci Petriego wykorzystywany jest opis regułowy oparty na logice sekwentów Gentzena, z elementami logiki temporalnej [1, 8]. Każda tranzycja sieci jest przedstawiona jako reguła IF...THEN... zapisana w postaci sekwentu. Po lewej stronie znaku wynikania $\vdash$ podane są warunki realizacji (zapalenia) danej tranzycji, natomiast po prawej skutki jej zapalenia.

Przykładowo zapis

$$T1: P1 * X0 \vdash P2 * P4;$$

oznacza, że warunkiem realizacji tranzycji T1 jest oznaczenie miejsca P1 oraz pojawienie się sygnału wejściowego X0 równego 1 (przycisk startowy włączony), natomiast skutkiem realizacji tej tranzycji jest oznakowanie miejsc P2 i P4.

W realizacji komputerowej opis automatu współbieżnego zawiera listy wszystkich sygnałów wejściowych (INPUTS), sygnałów wyjściowych kombinacyjnych (COMB_OUTS), wyjściowych rejestrowych (REG_OUTS), miejsc sieci (PLACES) oraz miejsc w oznakowaniu początkowym (MARKING). W rozpatrywanym przykładzie rozpatrywane są wyłącznie wyjścia kombinacyjne typu Moore'a. Pełny opis regułowy sieci składający się z reguł przejść (TRANSITIONS) i reguł wyjść (OUTPUTS) zamieszczono na rys. 3.

```
PLACES P1,P2,P3,P4,P5,P6,P7,P8,P9
INPUTS X0,X1,X2,X3,X4,X5,X6
COMB_OUTS Y0,Y1,Y2,Y3,Y4,Y5,Y6
MARKING P1

TRANSITIONS
T1:  P1 * X0 |- @P2 * @P4;
T2:  P2 * X1 |- @P3;
T3:  P4 * X3 |- @P5;
T4:  P3 * P5 |- @P6 * @P7;
T5:  P6 * X5 * X6 |- @P8;
T6:  P7 * /X2 * /X4 |- @P9;
T7:  P8 * /X5 |- @P6;
T8:  P6 * P9 * /X6 |- @P1;

OUTPUTS
P1 |- Y0;
P2 |- Y1;
P4 |- Y2;
P7 |- Y3*Y4;
P8 |- Y5;
P9 |- Y6.
```

Rys. 3. Regułowy opis sieci Petriego

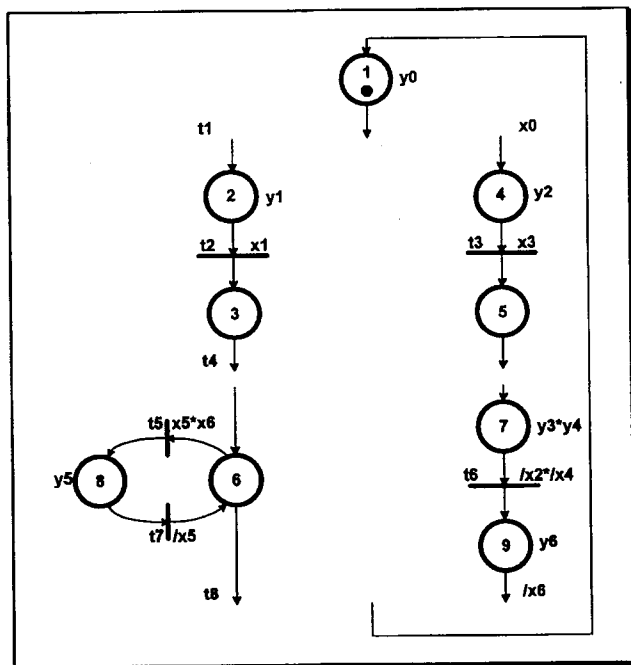
3. AUTOMATOWA DEKOMPOZYCJA SIECI PETRIEGO

W wielu przypadkach rozmiary realizowanego układu wykraczają poza ramy narzucone przez dostępny element scalony. Z tego względu konieczny jest podział układu na szereg współpracujących ze sobą podukładów. Z drugiej strony zdekomponowanie bardziej złożonego układu ułatwia proces projektowania. Rezultaty dekompozycji mogą być również wykorzystane do efektywnego kodowania lokalnych stanów wewnętrznych układu sterującego (miejsc sieci Petriego). Dodatkowo, zdekomponowany układ może być łatwo opisany nawet przy użyciu prostych, popularnych języków automatowych, wykorzystywanych do syntezy układów PLD.

3.1. DEKOMPOZYCJA SIECI PETRIEGO NA PODSIECI TYPU AUTOMATOWEGO

Częstym, stosowanym w praktyce rozwiązaniem umożliwiającym łatwy sposób syntezy układu cyfrowego jest dekompozycja automatu współbieżnego na sekwencyjne automaty składowe, rozpatrywane w sposób odrębny z wykorzystaniem metod klasycznych. Problem ten sprowadza się do dekompozycji sieci Petriego na podsieci automatowe, reprezentowane grafem bez tranzycji wielowyjściowych i zawierające tylko jeden znak (marker).

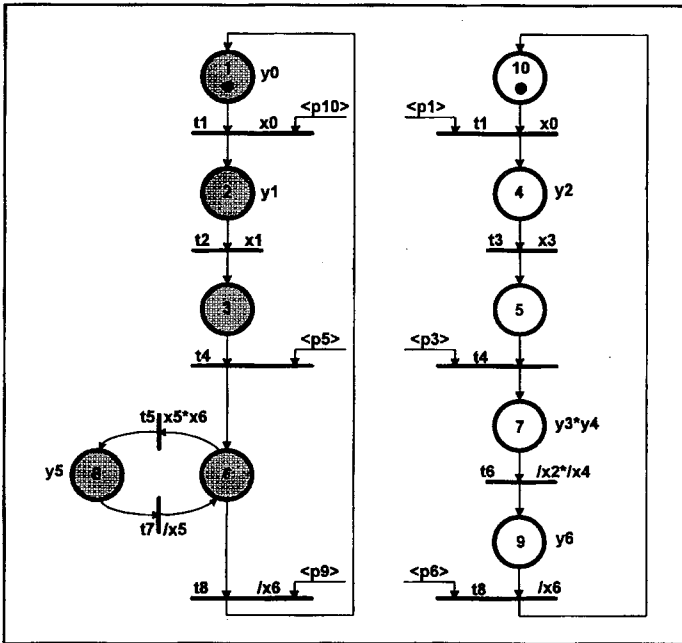
Najczęściej sieci Petriego rozbija się na podsieci automatowe stosując metodę zapoczątkowaną przez Andre'a [5]. Metoda Andre'a polega na wykorzystaniu makrosieci, zawierającej wyłącznie tranzycje wielowejściowe i wielowyjściowe oraz makromiejsca typu automatowego. Makromiejsca te reprezentują sekwencyjne części sieci, przedstawione przez składowe spójne grafy, powstałe po usunięciu tych tranzycji (rys. 4). Jeśli poszczególne części sieci nie zawierają początkowo oznakowanego miejsca sieci, to dodaje się do niego dodatkowe miejsce spoczynkowe, zawierające znak. Miejsce spoczynkowe bez znaku wprowadza się do podsieci zawierającej już znak, gdy nie jest ona spójna. Zadaniem miejsc spoczynkowych (np. miejsca P10 na rys. 5) jest przechowywanie znaku w tych sytuacjach, w których żadne z pozostałych miejsc rozpatrywanego pokrycia nie jest znakowane. Po ponownej agregacji (złożeniu) otrzymana w ten sposób sieć jest siecią konserwatywną, ze stałą liczbą znaków o różnych kolorach (rozdz. 4).



Rys. 4. Metody Andre'a dekompozycji sieci Petriego

Metoda Andre'a wyznaczania makromiejsc typu automatowego została zaadaptowana w wielu pracach, np. [8, 12, 15, 21]. W praktyce scala się kolejne miejsca sieci połączone tranzycjami o jednym wejściu i jednym wyjściu stopniowo tworząc makromiejsca.

Dekompozycje sieci na składowe automaty [1, 8] przeprowadza się na podstawie jej makrosieci. Znanymi metodami tworzy się jej graf znakowań, po czym na jego podstawie określa relację współbieżności między makromiejscami. Relacja ta jest przedstawiona w formie grafu nieorientowanego lub jego tablicowego zapisu (tablica współbieżności).



Rys. 5. Zdekomponowana sieć Petriego

Kolorując (w sensie topologicznym) graf współbieżności metodami znanymi z teorii grafów, określa się zbiory makromiejsc jednakowo pokolorowanych. Wyznaczone podzbiory miejsc tworzą pokrycie rozpatrywanej sieci Petriego podsieciami typu automatowego. Inna, równoważna metoda polega na wyszukiwaniu klik w dopełnieniu grafu współbieżności, czyli grafie niewspółbieżności. Wierzchołki należące do klik mają ten sam kolor, a klikom przypisuje się różne kolory. Zbiór miejsc o tym samym kolorze dalej będzie nazywany krótko blokiem. Pokrycie miejsc sieci blokami nazywane jest nieredundantnym (nienadmiarowym), jeśli z niego nie można usunąć żadnego z bloków. Mając do dyspozycji pokrycie nieredundantne można wygenerować szereg podziałów zbioru miejsc, pozwalających na dogodną implementację sieci Petriego. W prezentowanym dalej algorytmie dąży się do minimalizacji łącznej liczby symboli (przerzutników) potrzebnych do zakodowania stanów wewnętrznych w automatach sekwencyjnych, odpowiadających podsieciami zdekomponowanej sieci.

W celu kolorowania topologicznego grafu współbieżności zaadaptowano metodę Jerszowa [11]. Obiecujące wyniki otrzymano również za pomocą metod symbolicznych, których efektywność może być ewentualnie poprawiona przez zastosowanie grafów OBDD i komercyjnego oprogramowania. Otrzymane w ten sposób kolory miejsc przenosi się na graf sieci Petriego. W celu syntezy rozpatruje się odrębnie poszczególne podsieci, powiązane ze sobą, lecz identyfikowane swoimi kolorami. Opisując tranzycje wielowejsciowe podaje się jedno centralne miejsce wejściowe z danej podsieci oraz wszystkie miejsca poboczne, o innych kolorach, pochodzące z innych podsieci i wchodzące do jej tranzycji.

W zapisie symbolicznym (plik wejściowy) miejsca poboczne wyróżnione są przez ujęcie ich nazw między znakami $< >$. Obok każdego miejsca pochodzącego z innej

```

PLACES P1,P2,P3,P4,P5,P6,P7,P8,P9,P10
INPUTS X0,X1,X2,X3,X4,X5,X6
COMB_OUTS Y0,Y1,Y2,Y3,Y4,Y5,Y6
MARKING P1

AUTOMATA0
PLACES P1,P2,P3,P6,P8
INPUTS X0,X1,X5,X6
COMB_OUTS Y0,Y1,Y5
MARKING P1

TRANSITIONS
T1: P1 * <P10[1]> * X0 |- @P2;
T2: P2 * X1 |- @P3;
T4: P3 * <P5[1]> |- @P6;
T5: P6 * X5 * X6 |- @P8;
T7: P8 * /X5 |- @P6;
T8: P6 * <P9[1]> * /X6 |- @P1;

OUTPUTS
P1 |- Y0;
P2 |- Y1;
P8 |- Y5;

AUTOMATA1
PLACES P10,P4,P5,P7,P9
INPUTS X0,X2,X3,X4,X6
COMB_OUTS Y2,Y3,Y4
MARKING P10

TRANSITIONS
T1: P10 * <P1[0]> * X0 |- @P4;
T3: P4 * X3 |- @P5;
T4: P5 * <P3[0]> |- @P7;
T6: P7 * /X2 * /X4 |- @P9;
T8: P9 * <P6[0]> * /X6 |- @P10;

OUTPUTS
P4 |- Y2;
P7 |- Y3*Y4;
P9 |- Y6.

```

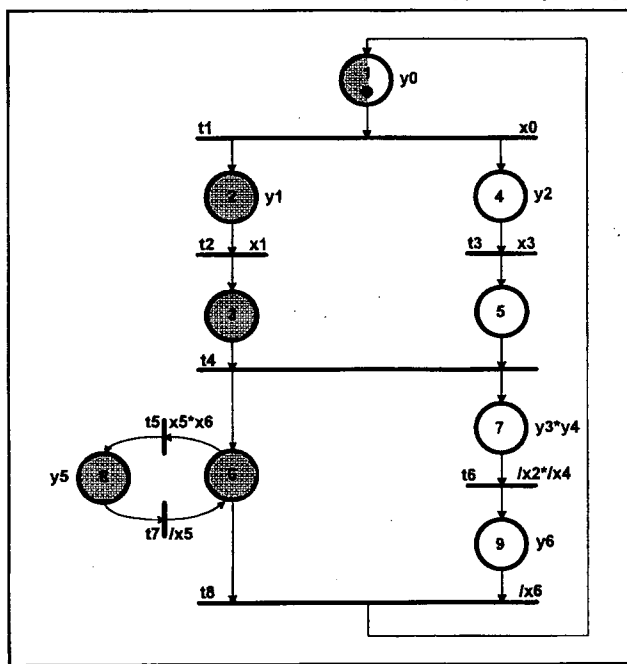
Rys. 6. Regułowy opis podsieci automatowych

podsieci podany jest w nawiasach [] kolor (cecha) podsieci (podautomatu), z której one pochodzą (rys. 6).

4. KOLOROWALNE SIECI PETRIEGO

Jak wynika z poprzednich rozdziałów, kolorowanie sieci Petriego polega na przypisaniu miejscom i tranzycjom sieci określonego koloru, zgodnie z następującymi zasadami [3, 8, 9]:

- wszystkie miejsca początkowo oznakowane (znakowanie początkowe) różnią się kolorami;
- jeżeli miejsce ma określony kolor, to jego wszystkie tranzycje wejściowe i wyjściowe zawierają dany kolor;
- miejsce może posiadać jednocześnie kilka kolorów;
- różne miejsca wejściowe tej samej tranzycji mają różne kolory;
- różne miejsca wyjściowe tej samej tranzycji mają różne kolory;
- podzbiory kolorów dla miejsc wejściowych i wyjściowych są sobie równe.



Rys. 7. Kolorowana sieć Petriego

Kolorowanie nazywane jest minimalnym, jeśli wykorzystano minimalną liczbę kolorów.

Sieć Petriego dającą się pokolorować zgodnie z wyżej podanymi zasadami nazywana jest siecią kolorowalną (albo cechowalną).

Sieć Petriego z rys. 2 może być pokolorowana jak podano na rys. 5.

4.1. INTERPRETACJA KOLOROWALNEJ INTERPRETOWANEJ SIECI PETRIEGO

Dokładną interpretację kolorowanej sieci Petriego reprezentującej kontroler współbieżny przedstawia Tabela 1.

Tabela 1. Interpretacja kolorowanej sieci Petriego

Sieć Petriego	Kontroler współbieżny
Miejsce	Stan lokalny
Tranzycja	Zmiana stanu lokalnego
Znakowanie	Stan globalny
Znakowanie początkowe	Początkowy stan globalny
KOLOR	PROCES SEKWENCYJNY
Etykieta tranzycji	Sygnały wejściowe
Etykieta miejsc	Sygnały wyjściowe

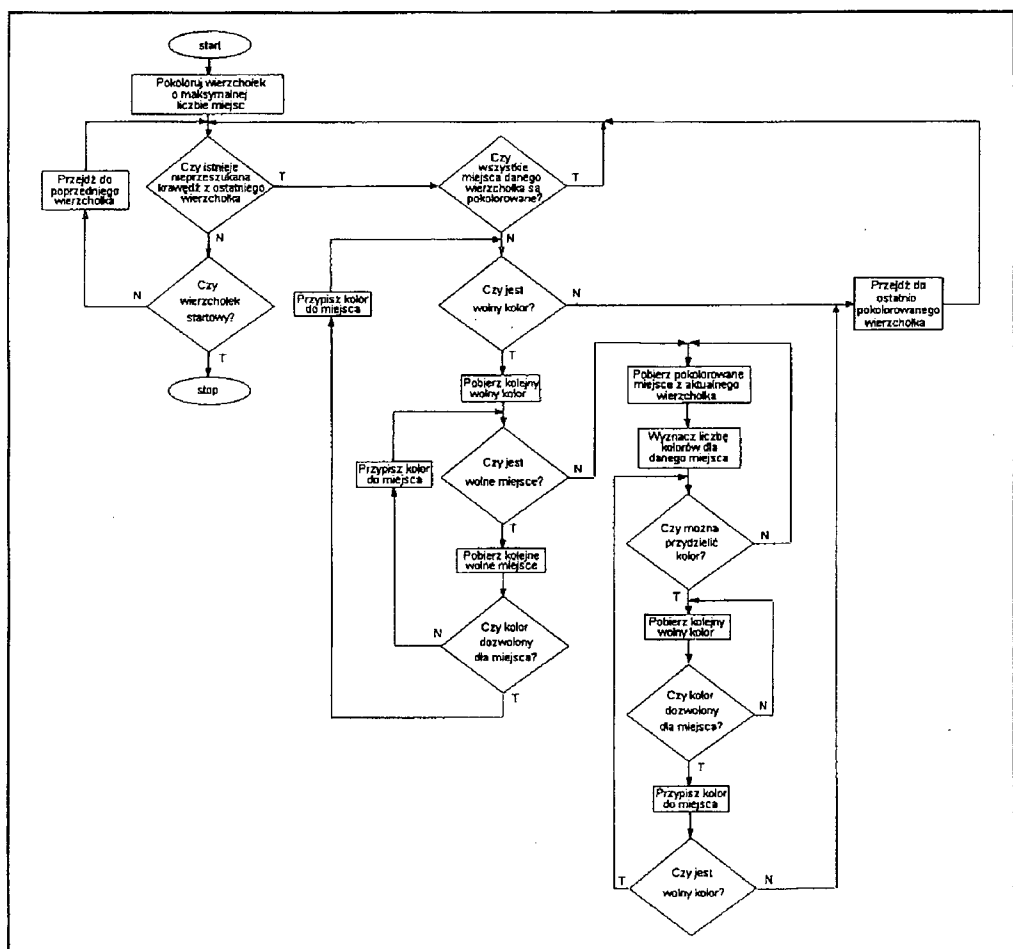
4.2. ALGORYTM KOLOROWANIA SIECI PETRIEGO

Celem kolorowania sieci Petriego jest pokrycie sieci składowymi automatowymi. Każda z nich to zbiór miejsc połączonych szeregowo za pośrednictwem tranzycji. Niektóre miejsca sieci należą jednocześnie do kilku składowych automatowych (podsieci). W dowolnym stanie każda podsieć zawiera dokładnie jeden marker.

Poniżej zostanie przedstawiony efektywny, heurystyczny algorytm kolorowania sieci Petriego, bazujący na grafie znakowań sieci. Wykorzystuje on i rozszerza rezultaty przedstawione w pracy [14].

W algorytmie kolorowania sieci wykorzystuje się następujące własności grafu znakowań:

- każdy wierzchołek grafu znakowań reprezentuje stan globalny automatu i musi zawierać wszystkie kolory;
- miejsca zawarte w jednym wierzchołku należą (każde oddzielnie) do wszystkich podsieci;
- miejsca należące do jednego wierzchołka grafu znakowań są współbieżne, natomiast miejsca należące do jednej podsieci są sekwencyjne (nigdy nie są oznakowane jednocześnie);
- w każdym wierzchołku występuje dokładnie jedno miejsce z każdej podsieci (niektóre miejsca reprezentowane są w kilku podsieciach). Wynika stąd wniosek, że miejsca należące do jednego wierzchołka grafu znakowań muszą być pokolorowane każde innym kolorem, niektóre nawet kilkoma kolorami;
- ponieważ każde miejsce zawarte w wierzchołku grafu znakowań musi posiadać co najmniej jeden kolor, zatem minimalna liczba kolorów potrzebna do pokolorowania sieci jest równa maksymalnej liczbie miejsc należących do tego samego wierzchołka grafu znakowań.



Rys. 8. Schemat blokowy algorytmu kolorowania sieci Petriego

Algorytm kolorowania sieci na podstawie grafu znakowań jest następujący:

1. Wybieramy wierzchołek grafu znakowań zawierający maksymalną liczbę miejsc.
2. Każdemu miejscu z tego wierzchołka przydzielamy jeden kolor.
3. Wybieramy jedną krawędź wychodzącą z tego wierzchołka, aby za jej pośrednictwem pokolorować miejsca z następnego wierzchołka.
4. Każdemu wolnemu miejscu przydzielamy jeden wolny kolor, zachowując zasadę różnych kolorów dla miejsc współbieżnych.
5. Jeżeli po początkowym pokolorowaniu miejsc zostaną wolne kolory, to:
 - a) wybieramy ostatnio pokolorowane miejsce,
 - b) określamy (na podstawie maksymalnej liczby kolorów i liczby miejsc już pokolorowanych w rozpatrywanym znakowaniu sieci) dopuszczalne kolory dla danego miejsca,
 - c) przydzielamy odpowiednią liczbę kolorów rozpatrywanemu miejscu,
 - d) o ile zostały wolne kolory — wracamy do a).

6. Sprawdzamy, czy w następnych wierzchołkach wszystkie miejsca są pokolorowane:
 - a) jeżeli tak, sprawdzamy, czy jesteśmy w wierzchołku, od którego zatrzymaliśmy kolorowanie:
 - a)' jeżeli tak, kończymy kolorowanie,
 - a)" jeżeli nie, wracamy do wcześniej rozpatrywanego wierzchołka (w górę grafu znakowań) i ponownie wykonujemy punkt 6).
 - b) jeżeli nie, przystępujemy do czynności opisane w punkcie 3. Kolorowanie zostaje zakończone gdy miejsca zawarte we wszystkich wierzchołku otrzymają kolory.

Omówiony algorytm kolorowania miejsc sieci Petriego przedstawiono na schemacie blokowym (rys. 8).

Inny algorytm kolorowania związany z wyznaczeniem cykli w grafie sieci Petriego omówiono w pracy [9].

5. SPOSOBY MODELOWANIA SIECI AUTOMATOWYCH

W dalszej części w celu zaprezentowania danego sposobu opisu będą podawane reprezentacje sieci Petriego w wybranym języku kompilatora PLD. Z braku miejsca ograniczono się do opisu pojedynczej podsieci w kolorze [0], nazwanej Automata0. Nie zamieszczono również specyficznych cech wybranego elementu PLD oraz koniecznych w tym przypadku deklaracji dodatkowych.

5.1. SPECYFIKACJA W JĘZYKU CUPL

System CUPL firmy Logical Devices należy do grupy chętnie stosowanych komercyjnych kompilatorów układów programowalnych rozbudowanych w kierunku współpracy z użytkownikiem. Opis automatowy języka CUPL został wykorzystany do modelowania i syntezy zdekompowanej sieci Petriego. Należy tu zwrócić uwagę, że umożliwia on łączną syntezę kilku współpracujących podsieci w jednym elemencie scalonym.

Najpierw należy zdefiniować wektor zmiennych stanu (Field Automata0=[QA2..0]), następnie zakodować kolejne stany (miejsca). Całą sieć automatową można opisać wykorzystując instrukcję SEQUENCE. Każdej tranzycji odpowiada pojedynczy moduł PRESENT przedstawiający aktualny stan automatu. Przejście do następnego miejsca jest możliwe tylko pod podanymi warunkami. Przejście DEFAULT utrzymuje automat w aktualnym stanie. Wyjścia określone są poprzez OUT w danym stanie. Przykładowy opis dla podsieci automatowej Automata0 podano na rys. 9.

5.2. SPECYFIKACJA W JĘZYKU PALASM

Język PALASM dostarcza kilku kolejnych możliwości opisu pojedynczych automatów. Wygodną i zrozumiałą do tego celu formą jest instrukcja IF...THEN...ELSE...

```

Field Automata0 = [QA2..0];

$define P1 'b'000
$define P2 'b'001
$define P3 'b'010
$define P6 'b'011
$define P8 'b'100

Sequence Automata0 {
    Present P1
        if Automata1:P10 & X0 Next P2;
        default Next P1;
        Out Y0;
    Present P2
        if X1 Next P3;
        default Next P2;
        Out Y1;
    Present P3
        if Automata1:P5 Next P6;
        default Next P3;
    Present P6
        if X5 & X6 Next P8;
        if Automata1:P9 & !X6 Next P1;
        default Next P6;
    Present P8
        if !X5 Next P6;
        default Next P8;
        Out Y5;
}

```

Rys. 9. Specyfikacja automatu w języku CUPL

Opis regułowy, w odróżnieniu od wycofywanego już dedykowanego opisu automatowego (wariant Palasm – MachPLUS) umożliwia umieszczenie kilku współpracujących podsieci w jednym module FPLD.

Tranzycję (w zależności od jej rodzaju) przedstawia się w następujący sposób:

a) tranzycja bezwarunkowa

```

IF DaneMiejsce THEN BEGIN NastepneMiejsce END
                        ELSE BEGIN DaneMiejsce      END

```

b) tranzycja warunkowa

```

IF DaneMiejsce THEN IF Warunek THEN BEGIN NastepneMiejsce END
                        ELSE BEGIN DaneMiejsce      END

```

Dla warunków zagnieżdżających się można wprowadzać kolejne zagnieżdżenia lub tworzyć iloczyn warunków.

Dla rozpatrywanych tutaj wyjść typu Moore'a wystarczy w odpowiedniej regule po lewej stronie określić wszystkie miejsca, w których ustawiane jest odpowiednie wyjście, jako sumę logiczną symboli tych miejsc. Na przykład jeżeli wyjście Y10 jest aktywne gdy miejsca P1, P2 lub P3 są oznakowane, zapisujemy to następująco:

$$Y10 = P1 + P2 + P3$$

Rys. 10 przedstawia ponownie fragment podsieci Automatu w języku regułowym. Inny sposób modelowania sieci Petriego w języku PALASM4 można znaleźć w pracach [4, 18].

```

;Kodowanie stanów dla automatu 0
STRING P1  ' (/Q2*/Q1*/Q0) '
STRING P2  ' (/Q2*/Q1* Q0) '
STRING P3  ' (/Q2* Q1*/Q0) '
STRING P6  ' (/Q2* Q1* Q0) '
STRING P8  ' ( Q2*/Q1*/Q0) '

STRING @P1 'Q2:=0 Q1:=0 Q0:=0 '
STRING @P2 'Q2:=0 Q1:=0 Q0:=1 '
STRING @P3 'Q2:=0 Q1:=1 Q0:=0 '
STRING @P6 'Q2:=0 Q1:=1 Q0:=1 '
STRING @P8 'Q2:=1 Q1:=0 Q0:=0 '

IF P1 THEN BEGIN
    IF P10*X0      THEN BEGIN @P2 END
                  ELSE BEGIN @P1 END
END

IF P2 THEN BEGIN
    IF X1          THEN BEGIN @P3 END
                  ELSE BEGIN @P2 END
END

IF P3 THEN BEGIN
    IF P5          THEN BEGIN @P6 END
                  ELSE BEGIN @P3 END
END

IF P6 THEN BEGIN
    IF P9*/X6      THEN BEGIN @P1 END
                  ELSE BEGIN
                        IF X5*X6 THEN BEGIN
@P8 END
                        ELSE BEGIN
@P6 END
                        END
END

IF P8 THEN BEGIN
    IF /X5 THEN BEGIN @P6 END
    ELSE BEGIN @P8 END
END

Y0 = P1
Y1 = P2
Y5 = P8

```

Rys. 10. Specyfikacja automatu w języku PALASM

5.3. SPECYFIKACJA W JĘZYKU VHDL

Język VHDL jest coraz częściej wykorzystywany do syntezy układów FPLD i FPGA (np. system MINC, ALTERA). Istnieje wiele metod modelowania sieci Petriego w języku VHDL, ukierunkowanych głównie na jej symulację. Większość z nich jednak wykorzystuje w miarę złożone struktury języka, aby przedstawić współbieżność poszczególnych procesów. W prezentowanej metodzie zdecydowano się na znaczne uproszczenie opisu sieci Petriego, pod kątem jej szybkiej i efektywnej syntezy.

```

use std.std_logic_1164.all;
PACKAGE Pakiet Przyklad is
    constant NUMBER_INPUTS      :      integer :=7;
    constant NUMBER_OUTPUTS     :      integer :=7;
    constant NUMBER_STATES0     :      integer :=5;
    subtype T_INPUTS is std_logic_vector(NUMBER_INPUTS-1
downto 0);
    subtype T_OUTPUTS is std_logic_vector(NUMBER_OUTPUTS-1
downto 0);
    type T_STATES0 is (P1, P2, P3, P6, P8);
end Pakiet Przyklad;

use std.std_logic_1164.all;
use work.Pakiet Przyklad.all;
entity Przyklad is
    port ( CLK      : in std_logic;
           X        : in T_INPUTS := (others => '0');
           Y        : out T_OUTPUTS := (others => '0') );
end Przyklad;

architecture ARCH1 of Przyklad is
    signal CURRENT_STATE0 : T_STATES0;
begin
    SET_STATE0 : process (CLK)
        variable NEXT_STATE0 : T_STATES0 := CURRENT_STATE0;
        begin
            case CURRENT_STATE0 is
                when P1 => if (X(0)='1') and (CURRENT_STATE1=P10)
                           then NEXT_STATE0:=P2;
                           end if;
                when P2 => if (X(1)='1') then NEXT_STATE0:=P3;
                           end if;
                when P3 => if (CURRENT_STATE1=P5) then
NEXT_STATE0:=P6;
                           end if;
                when P6 => if (X(5)='1') and (X(6)='1') then
NEXT_STATE0:=P8;
                           elsif (X(6)='0') then NEXT_STATE0:=P1;
                           end if;
                when P8 => if (X(5)='0') then NEXT_STATE0:=P6;
                           end if;
            end case;
            CURRENT_STATE0 <= NEXT_STATE0;
        end process SET_STATE0;
    Y(0) <= '1' when CURRENT_STATE0=P1 else '0';
    Y(1) <= '1' when CURRENT_STATE0=P2 else '0';
    Y(5) <= '1' when CURRENT_STATE0=P8 else '0';
end ARCH1;

```

Rys. 11. Specyfikacja podsięci automatowych w języku VHDL

Każdy automat może być przedstawiony jako oddzielny proces z wykorzystaniem niezależnych zmiennych i sygnałów. Opisując kolejne automaty można zdefiniować własne typy wyliczeniowe, które symbolicznie przedstawiają poszczególne stany automatu. W rozpatrywanym przypadku są to symboliczne nazwy miejsc.

Sygnały wejściowe i wyjściowe są przedstawione jako wektory binarne o odpowiedniej długości. Aby uprościć automatyczne wygenerowanie modelu można dodatkowo przygotowywać moduł biblioteki (package), w której zdefiniowane są niezbędne typy i stałe.

Dodatkową zaletą opisu automatu współbieżnego w języku VHDL jest możliwość łącznej symulacji całego urządzenia cyfrowego [22].

Rys. 11 przedstawia sposób modelowania automatu 0 z wykorzystaniem instrukcji warunkowych języka VHDL.

PODSUMOWANIE

W pracy przedstawiono metodę pozwalającą w stosunkowo prosty sposób dekomponować sieci Petriego metodą kolorowania jej miejsc. Szczegółowo omówiono jedną z bardzo efektywnych metod heurystycznych, stosowaną wówczas, gdy dany jest graf znakowań makrosieci. Kolorowanie, opisane w pracy, jest nie tylko narzędziem dekompozycji. Z braku miejsca pominięto inne aspekty jego zastosowania, np. do kodowania miejsc sieci [3, 4, 9] i jej analizy. Omówiono metodologię syntezy i symulacji automatów współbieżnych z wykorzystaniem standardowych języków takich jak CUPL, PALASM i VHDL. Krótko przedstawiono trzy wersje programów przygotowujących pliki dla popularnych kompilatorów FPLD i FPGA.

Praca powstała w ramach grantu KBN nr 8 T11C 051 08 pod kierunkiem naukowym prof. dr hab. inż. Mariana Adamskiego.

Autor składa szczególne podziękowania Panu Profesorowi Marianowi Adamskiemu za udzieloną pomoc podczas realizacji projektu.

BIBLIOGRAFIA

1. M. A d a m s k i: *Projektowanie układów cyfrowych systematyczną metodą strukturalną*. Wydawnictwo Wyższej Szkoły Inżynierskiej w Zielonej Górze, Zielona Góra 1990
2. M. A d a m s k i: *Parallel Controller Implementation using Standard PLD Software*. International Workshop on Field Programmable Logic and Applications, Oxford, UK, 4–6.09.91, s. K5.1–9, 1991.
3. M. A d a m s k i, M. W ę g r z y n: *Hierarchically structured coloured Petri net specification and validation of concurrent controllers*. 39 Internationales Wissenschaftliches Kolloquium IWK'94, Technische Universität in Ilmenau (Niemcy), 1994, vol. 1, s. 517–522
4. M. A d a m s k i, J. L. M o n t e i r o: *PLD Implementation of Logic Controllers*. Proceedings of IEEE International Symposium on Industrial Electronics, ISIE'95 — Athens, Greece, July 10–14, 1995, vol. 2, s. 706–711

5. C. Andre, F. Boeri, J. Marin: *Synthese et realisation des systemes logiques a evolutions simultanees*. R.A.I.R.O 1976, vol. 10, no. 4
6. M. Auguin, F. Boeri, C. Andre: *New design using PLAs and Petri nets*. International Symposium MECO'78, Athens, June 1978
7. M. Auguin, F. Boeri, C. Andre: *Systematic method of realisation of interpreted Petri nets*. Digital Processes, 1980, vol. 6, no. 1
8. Z. Banaszak, J. Kuś, M. Adamski: *Sieci Petriego. Modelowanie, sterowanie i synteza systemów dyskretnych*. Wydawnictwo Wyższej Szkoły Inżynierskiej w Zielonej Górze, Zielona Góra 1993
9. K. Biliński, M. Adamski, J.M. Saul, E.L. Dagless: *Petri net based algorithms for parallel controller synthesis*. IEE Proceedings — E, Computers and Digital Techniques, Nov. 1994, vol. 141, no. 6, s. 405–412
10. R. David, H. Alla: *Petri Nets for Modelling of Dynamic Systems — A survey*. Automatica, 1994, vol. 30, no. 2, s. 175–202
11. A.P. Jerszow: *Wprowadzenie do teorii programowania*. WNT, Warszawa 1981
12. J. Kalinowski, K. Jasiński: *Synteza układów cyfrowych z wykorzystaniem sieci Petriego*. Rozprawy Elektrotechniczne, 1986, t. 32, z. 2, s. 339–351
13. J. Kalinowski, T. Łuba: *Metoda syntezy logicznej układów cyfrowych opisywanych sieciami Petriego*, Rozprawy Elektrotechniczne, 1986, t. 32, z. 4, s. 1253–1263
14. R. Kłosiński: *Program kolorowania sieci Petriego na podstawie grafu znakowań makrosieci*. Praca magisterska. Wyższa Szkoła Inżynierska w Zielonej Górze, Zielona Góra 1988
15. T. Kozłowski, E.L. Dagless, J.M. Saul, M. Adamski, J. Szajna: *Parallel controller synthesis using Petri nets*. IEE Proceedings — E, Computers and Digital Techniques, July 1995, vol. 142, no. 4, s. 263–271
16. W. Majewski, T. Łuba, K. Jasiński, B. Zbierzchowski: *Programowalne moduły logiczne w syntezie układów cyfrowych*. WKŁ, Warszawa 1992
17. M. Molski: *Modułowe i mikroprogramowane układy cyfrowe*. WKŁ, Warszawa 1986
18. J. Pardey, A. Amroun, M. Bolton, M. Adamski: *Parallel Controller Synthesis for Programmable Logic Devices*. Microprocessors and Microsystems, Oct. 1994, vol. 18, n. 8, s. 451–458
19. T. Murata: *Petri Nets: properties, analysis and applications*. Proceedings of the IEEE, April 1989, vol. 77, no. 4, pp. 541–580
20. J.M. Toulotte, J.P. Parsy: *A method for decomposing interpreted Petri Nets and its utilisation*. Digital Processes, 1979, vol. 5, no. 5, s. 223–234
21. F. Wagner, M. Rakowski: *Zastosowanie sieci Petri do projektowania mikroprogramowanych układów sekwencyjnych*. Archiwum Automatyki i Telemekhaniki, 1981, tom XXVI, zeszyt 3
22. M. Węgrzyn, M. Oleszczuk: *Zastosowanie języka VHDL do modelowania i testowania programowalnych struktur logicznych PLD*, 16 Międzynarodowe Sympozium Naukowe, t. 2, Wyższa Szkoła Inżynierska w Zielonej Górze, 1994, s. 75–78
23. P. Wolański: *Specyfikacja i symulacja ASM w języku VHDL*. Kwartalnik Elektroniki i Telekomunikacji, 1996, t. 42, z. 2, s. 129–144

M. WĘGRZYN

PLD-BASED IMPLEMENTATION OF CONCURRENT CONTROLLERS

Summary

Design methods for concurrent controllers play a very important role in advanced digital system design. Behaviour of the controller can be described by means of using interpreted Petri nets. For the reason of effective implementation, the net that describes concurrent controller, may be divided (decomposed) into

the separate parts, and be coloured. In the paper a Petri net colouring-based method of parallel decomposition is presented. In addition the simple design methods for synthesis and simulation of concurrent controllers are described. The standard synthesis tools, such as PALASM, CUPL and VHDL are used for the final logic design.

Key words: digital systems, Petri nets, PLD circuits, concurrent controllers.

Optical frequency modulation for coherent optical subcarrier multiplexing transmission

MIKHAIEL SALIM MAKHOUL

Instytut Telekomunikacji, Politechnika Warszawska

Received 1996.03.27

Authorized 1996.05.17

This paper describes the coherent binary frequency shift keying-subcarrier multiplexing (BFSK-SCM) system with a bandwidth efficiency of 1 bps/Hz, using frequency optical modulation (FM) in the transmitter and phase noise cancelling (PNC) circuit in the receiver. We will give attention on the description and analysis of these techniques. We shall also see the impact of various noise components on the general performance of the system, as well as how the PNC circuit improves the general performance of the system.

Key words: frequency modulation, frequency BFSK-SCM system, lasery DFB, PNC circuits.

1. INTRODUCTION

The use of communication systems is increasing continually, and people need more communication in everyday life. World population increases, and communication among people also increases, that causes more demand of communication systems. It should be done something to fulfill this demand. This motive of satisfying the demand makes researchers give more of their attention and interests in this matter. It makes them exploit new systems and techniques, to achieve this requirement. That is why one can think the possibility of using one communication line for carrying more than one channel.

To solve these problems, multiplexing channels technique has been found. These techniques include: time division multiplexing (TDM), frequency division multiplexing (FDM), pulse code multiplexing (PCM) and subcarrier multiplexing (SCM) system. This will be our main subject: “*coherent subcarrier multiplexing system (CSCMS)*”, which will be discussed later. CSCMS and other communication systems have their advantages and disadvantages. But usually systems are evaluated on the base of total cost and general performance. The performance of a system is evaluated by carrier to noise rate (CNR) or bit error rate (BER) which are nearly about 17 dB

and 10^{-9} respectively. The use of their CNR or BER depends on whether the system is analogue or digital. Generally all communication systems are influenced by noise during their performance in various degrees. And therefore we will take a concrete analysis of this matter. As we said we will analyze the coherent subcarrier multiplexing system using frequency shift keying, optical frequency modulation in the transmitter and phase noise cancelling in the receiver. Thus, we will also get an idea about angular modulation which can be used in this system.

The advantage of this system is that optical transmission requires only a single optical carrier for N channel's transmission, also can increase the number of multiplexing channels. Its disadvantage is that optical modulator in SCM system requires wide bandwidth. Therefore the transmission capacity of this system may be small. But a coherent SCM system can be used with an optical frequency division multiplexing (OFDM) which includes a very finely tunable heterodyne receiver.

We are now through with the introductory part of this paper, and in the remaining parts we are going to see in details the following sections: angular modulation, distributed feedback-light diode frequency modulation response, system configuration, analyses of the optical frequency modulation for coherent transmission optical SCM-FSK, and in the end we will discuss about the effects of noise analysis.

2. ANGULAR MODULATION

Fig. 1 shows the schematic diagram of the angular modulation, frequency modulation, and phase modulation (FM, PM). In the direct frequency modulation, as we see in Fig. 1(a), the output signals of the subcarrier multiplexing (SCM) is the input to the light diode (LD), and its output signal is modulated directly by FM. Therefore,

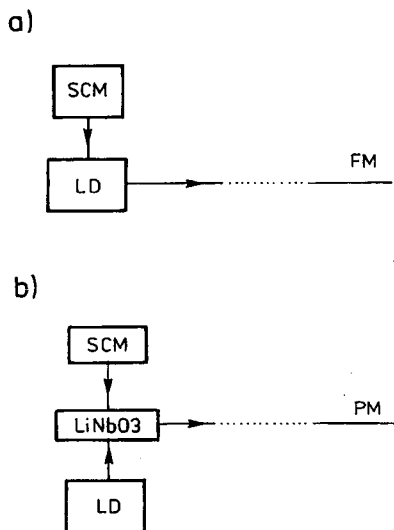


Fig. 1. The two angular modulation: direct frequency modulation (a), external modulation (b)

in the external modulation, see Fig. 1(b), the output signal of the SCM stage is the input to LiNbO_3 modulator laser, also, with the signal output of the LD, the output signal of the LiNbO_3 is modulated by PM. Among the advantages of the optical angular modulation (FM, PM) are as follows: in the first case the number of the multiplexed channels can be increased, because it uses the wide flat response frequency characteristics of the angular modulation. The second is, it is easy to apply to highly sensitive heterodyne detection and this allows to increase the number of subcarrier distribution. The third advantage is optical frequency modulation is easily done by directly modulation laser diodes.

Distributed feedback-light diode (DFB-LD) with a very wide flat FM response, a narrow linewidth and high output power, has been developed. Because of the advantages and the possible high output power optical FM by direct modulation of a local oscillator light (LO) could be well suited to optical SCM system.

3. DFB FREQUENCY MODULATION (FM) RESPONSE

As we can see in the Fig. 2, the FM sensitivity of DFB laser is much higher than the AM sensitivity, typical values are $500 \text{ MHz/mA} \div 1 \text{ GHz/mA}$ FM for DFB laser, and the typical frequency deviation of few GHz are used in FM or FSK coherent optical systems. That is why usually the incidental AM associated with the direct frequency modulation, is neglected. Fortunately, direct frequency modulation of laser diode seems simple.

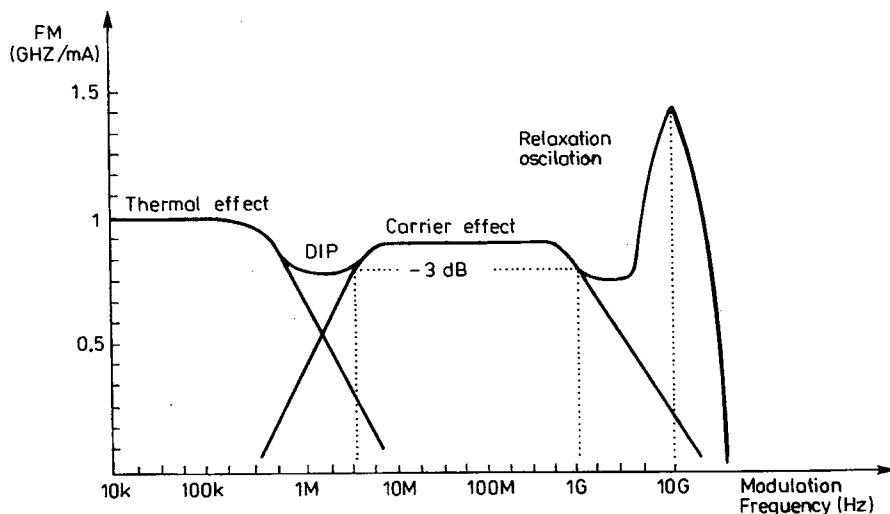


Fig. 2. FM response of a DFB-LD

The relaxation oscillation can be shifted to higher frequency, usually a few GHz, by increasing the bias current of the laser. The figure clearly shows that. In this method its effect on the FM can be neglected, and the maximum bit rate can be used, for direct

FM of the laser is limited by (-3 dB) carrier frequency. For low-frequencies the complete laser chip follows the current variations by the thermal heating and cooling, which dominates over the carrier density effect. We can say, for a very low-frequency current increases, the laser exhibits a change to larger wavelengths and lower optical frequency, while for a high-frequency current increase, the optical frequency increases too. Fig. 2 shows that, at some frequencies, these two effects cancel each other, resulting in a "dip" in the FM curve. This "dip" is usually located between $100 \text{ KHz} \div 10 \text{ MHz}$.

4. SYSTEM CONFIGURATION

The configuration for coherent optical transmission frequency shift keying subcarrier multiplexing frequency modulation (FSK-SCM-FM) can be seen in Fig. 3. The microwave subcarriers are modulated by digital data signals. Frequency shift keying subcarrier multiplexing (FSK-SCM) signals are combined, and the output of the combiner (C) is the input to an optical frequency modulator (FM).

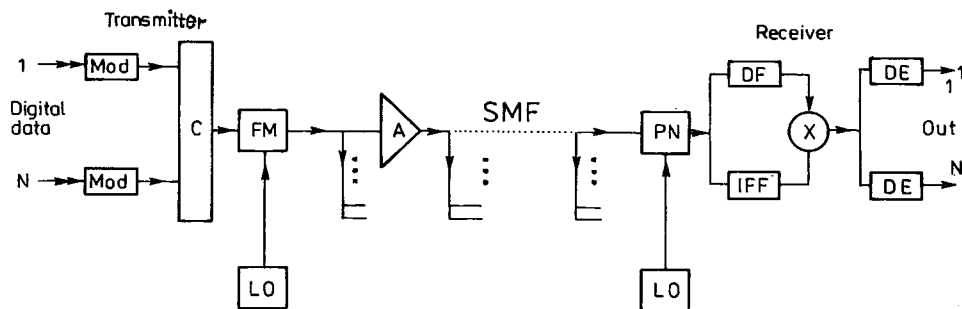


Fig. 3. Schematic diagram of the coherent optical transmission FSK-SCM-FM system

As we can see in detail, in Fig. 3, a local laser is set up in the transmitter, and the modulated optical signal is mixed before transmission with the local laser light. We assumed that before the local laser light (LO), and the subcarrier multiplexing (SCM) signal has the same state of polarization. Therefore, the coherent system (FSK-SCM-FM) is unaffected by the state of the polarization in the transmission fiber. The optical frequency modulation-subcarrier multiplexing (FM-SCM) signal is transmitted and distributed through a single-mode fiber (SMF). Direct optical amplification by an optical amplifier (A) before or after the splatter should therefore be effective. The (FSK-SCM-FM) signal is detected and demodulated by the receiver. Receiver detects the transmitted optical signal using a photodetector (PN), PNC circuit and microwave radio frequency (RF) conventional receiver.

It is of interest to be able to do PNC circuit in the coherent (FSK-SCM-FM) system, as we can see in Fig. 5, the photocurrent is divided and filtered such that the one branch of the PNC circuit pass the data channels, and we are using for that data filter (DF). In the second branch passes the intermediate frequency (IF) signal, and is used for that IF-filter (IFF). The photocurrent in the data branch and in the IF branch

are mixed in the output of the PNC circuit. Because the phase noise is common to both signals the noise can be cancelled, and as we can see, after mixing the signals in the two branches at the output of the PNC circuit, the subcarriers are now clearly separated because the phase noise has been properly cancelled — Fig. 6. On the other hand, the degree of the noise cancellation depends on the linewidth of the laser and the delay time (DEL) between the two branches. The phase noise cancelling circuit also performs the function of the price automatic frequency-control loop. In addition, any frequency of the laser will cause the IF and the data signals to drift together, and this drift can be cancelled by the PNC circuit. In the end of this section, we can see that clearly, the PNC circuit will enable coherent FSK-SCM system to be built using conventional FM with DFB.

In the receiver, most, the heterodyne detection, if used, is the most attractive and practical method. This is supported by the advantages which follows. It can detect all SCM signals in the intermediate frequency range using wide-band optical receiver and microwave receiver. Apart from this, the system gives a high receiver sensitivity than the direct detection. It is easily upgrade using OFDM techniques.

Now if we want to realize heterodyne reception, some problems must be solved. These problems include the need of a frequency controlled local laser for heterodyne detection, overcoming the signal fading caused by polarization fluctuation and the overcome of degradation due to the phase noise.

Solving these problems, will make the total receiver cost become much higher than when using direct detection. This approach does not seem practical for single optical-carrier transmission.

To solve the problem, however, configuration with the local oscillator light diode in the transmitter, with the splitting terminal is proposed.

5. ANALYSIS OF THE SYSTEM OPTICAL MODULATION FOR COHERENT TRANSMISSION OPTICAL SCM

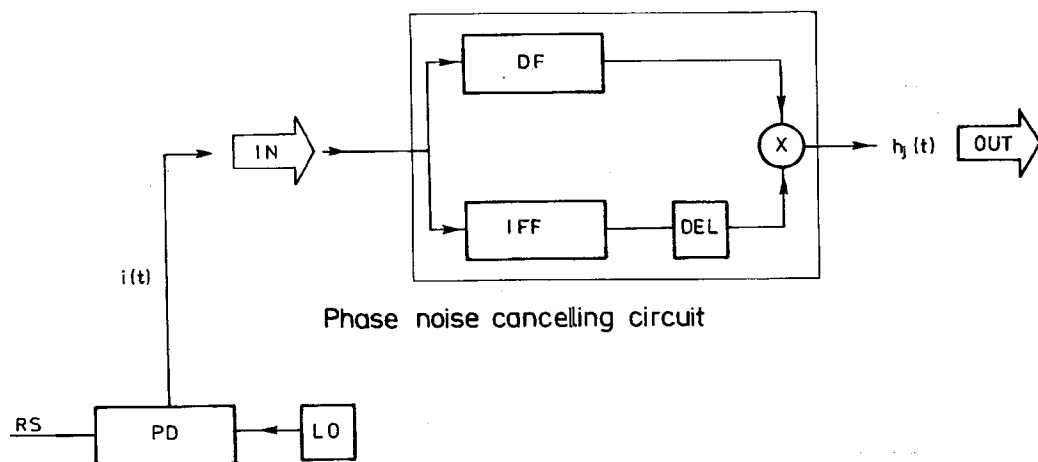


Fig. 4. The schematic diagram of the coherent receiver FSK-SCM-FM system using PNC circuit

Fig. 4 shows the receiver. The photodiode (PD) detects the photocurrent, and the detected photocurrent at the receiver is written as:

$$i(t) = 2R(P_S P_{LO})^{1/2} \cos[2\pi f_{IF} t + \theta(t) + \phi(t)] + n(t) \quad (1)$$

and thereof can be express the angular modulation SCM signal as:

$$\theta(t) = 2\pi f_a \int_0^t \sum_{j=1}^N \beta_j(t) \cos[2\pi f_j t + \alpha_j(t)], \quad (2)$$

where N : is the number of data channels, R : Photodiode (PD) responsivity, P_S : is the message signal power incident on the PD, P_{LO} : is the signal power incident on the PD, f_{IF} : is the intermediate frequency, $\theta(t)$: is the angular modulation SCM signal, $\phi(t)$: is the phase noise, $n(t)$: is the noise, $\alpha_j(t)$: is the angular modulation data signal, also, $\beta_j(t)$: is the PM index for channel j , f_j : is the microwave subcarrier frequency, f_a : the peak frequency deviation in the optical FM. As we know usually subcarrier frequencies are set much higher than data rates, and can approximated (2) as:

$$\theta(t) = \sum_{j=1}^N m_j(t) \sin[2\pi f_j t + \alpha_j(t)] \quad (3)$$

$$m_j(t) = \left(\frac{f_a}{f_j} \right) \beta_j(t), \quad (4)$$

where $m_j(t)$ is the FM index for channel j . Therefore, the frequency shift keying information for each data channels is given by

$$\alpha_j(t) = 2\pi f_a \int_{-\infty}^t \sum_i a_i g(t - iT) dt, \quad (5)$$

$$a_i = \pm 1$$

where a_i : the binary data, $g(t)$: a rectangular pulse, and T — the bit period. To characterize specific FSK, it is convenient to use the normalized frequency deviation defined as:

$$\Delta f = 2f_a T. \quad (6)$$

Thereout, to add some insight in to spectrum characteristics, and as matter-of-fact, if the modulation index m is the same for all channels, we can express the photocurrent at the detector as [from (1), (3) and (4)]

$$i(t) = 2R(P_S P_{LO})^{1/2} \cdot \sum_{n_1=-\infty}^{+\infty} \dots \sum_{n_N=-\infty}^{+\infty} j_{n_1}(m) \dots j_{n_N}(m) \cos \left[2\pi f_{IF} t + \sum_{k=1}^N n_k (2\pi f_k t + \alpha_k(t)) + \phi(t) \right], \quad (7)$$

where $j_n(t)$ is an n th-order Bessel function of the first kind with an frequency modulation index of m . Since the main information is contained in the first order terms of the Bessel series, the heterodyne detection IF current for channel J can be written as

$$i_j(t) = 2R\sqrt{P_s P_{LO}} j_1(m) [j_0(m)]^{N-1} \cos[2\pi(f_{IF} - f_j)t - \alpha_j(t) - \phi(t)]. \quad (8)$$

Hence, the output of PNC becomes:

$$h_j(t) = 2R\sqrt{P_s P_{LO}} j_1(m) [j_0(m)]^{N-1} \cos[2\pi f_j t + 2\pi f_{IF} d + \alpha_j(t) + \Delta\phi(t)] \quad (9)$$

$$\Delta\phi(t) = \phi(t) - \phi(t-d),$$

where $\Delta\phi(t)$ is the phase difference between the two branches of the PNC circuit. Definitions of the parameters in these equation are listed in Table (A). We can see that, form (9), it is not important to set intermediate frequency (IF) to a specific value, and the power penalty caused by IF center frequency between deviation can be suppressed, if the delay time (d) is made small enough. Since the subcarrier frequencies can be set very precisely, no automatic frequency control (AFC) loop is needed in the heterodyne receiver.

Definition of parameters with specific values

Table A

R	Responsivity $\eta=0.8$
P_s	Received signal power
P_{LO}	Received local oscillator power
RIN	Relative intensity noise light diode (-155 dB/Hz)
e	Electron charge
Γ	Transmission loss between optical amplifier and receiver (0.22 dB/Km)
m	Optical modulation index
B	Intermediate frequency bandwidth of the receiver (30 MHz for FM)
$h\nu$	Photon energy (1.28×10^{-19} j)
n_{sp}	Spontaneous emission factor (2.5)
I_n	Equivalent input noise current of photodetector [15 pA/(Hz) $^{1/2}$]
k_2	Number of second order IMP,s
k_3	Number of third order IMP,s
T	Absolute temperature (300 K)
NF	Figure noise (3 dB)

Fig. 5 and Fig. 6 show spectrum of the N unmodulated microwave subcarriers before phase noise cancellation, and after cancellation.

The subcarrier multiplexing (SCM) signal is modulated by an optical frequency modulator (FM), the required bandwidth of the optical modulator is nearly $4NB_r$. If we assume each channel spacing to be $D=2B_r$, where B_r is the bit rate, the optical signal which is transmitted is detected by the receiver whose bandwidth is about $6NB_r$, without phase noise cancelling (PNC) circuit or about $8NB_r$, with a phase noise cancelling circuit, where N is the number of channels. Each channel is selected by an optical fiber with a bandwidth of about $2B_r$. In this technique, a bandwidth nearly $8B_r$, is used in the optical fiber, so a bandwidth of $8B_r$ is needed to transmit each channel.

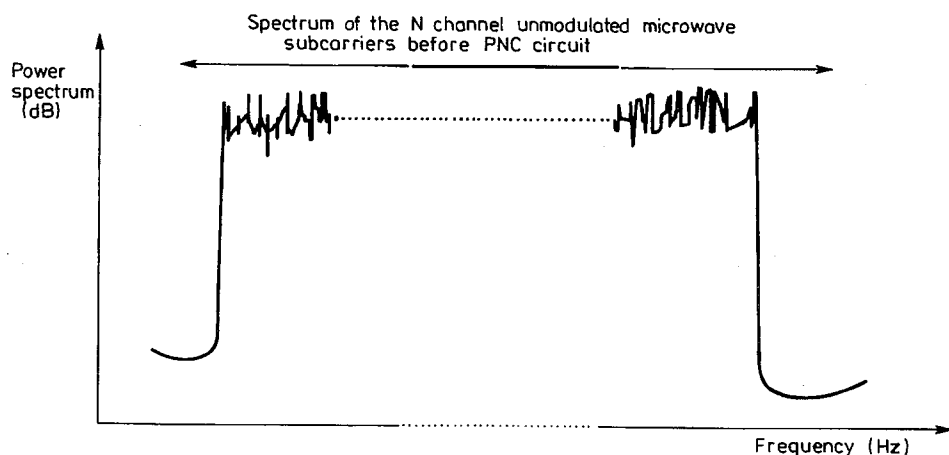


Fig. 5. General spectrum of the N channel unmodulated microwave subcarriers before the PNC circuit

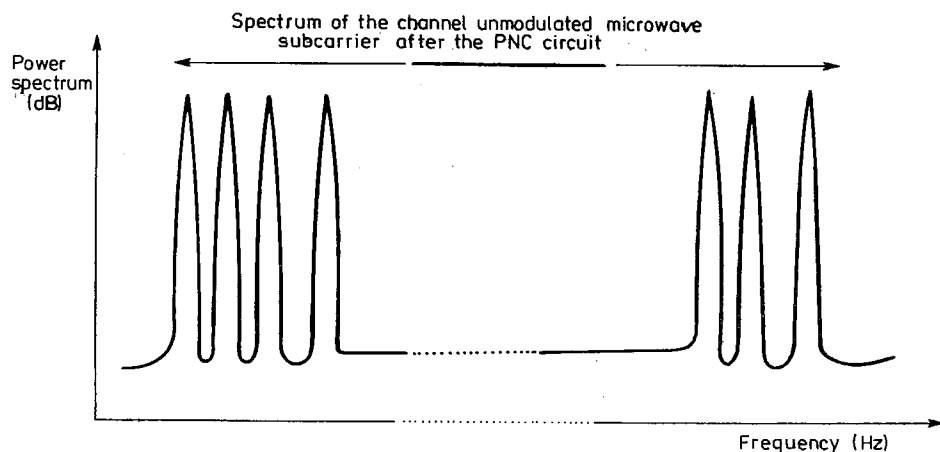


Fig. 6. General spectrum of the N channel unmodulated microwave subcarriers after the PNC circuit

Therefore, we can see in Fig. 7 two examples. First, which if we have a transmitter with an output power of 0 dBm in configuration (1), we can distribute a 100 channel FM signal to about 600 subscribers without any optical amplifier. And, if we can use single stage optical amplifiers with a signal gain of about 15 dB, the system can be up-graded to a more than 100 channels FM signal distribution to 10000 subscribers. Second, if we transmit with an output power of 0 dBm in configuration (2), the distribution number for a 100 channel FM signal is limited to about 20 subscribers, but it can be increased to about 10000, using double stage amplifiers with the signal gain of about 17 dB.

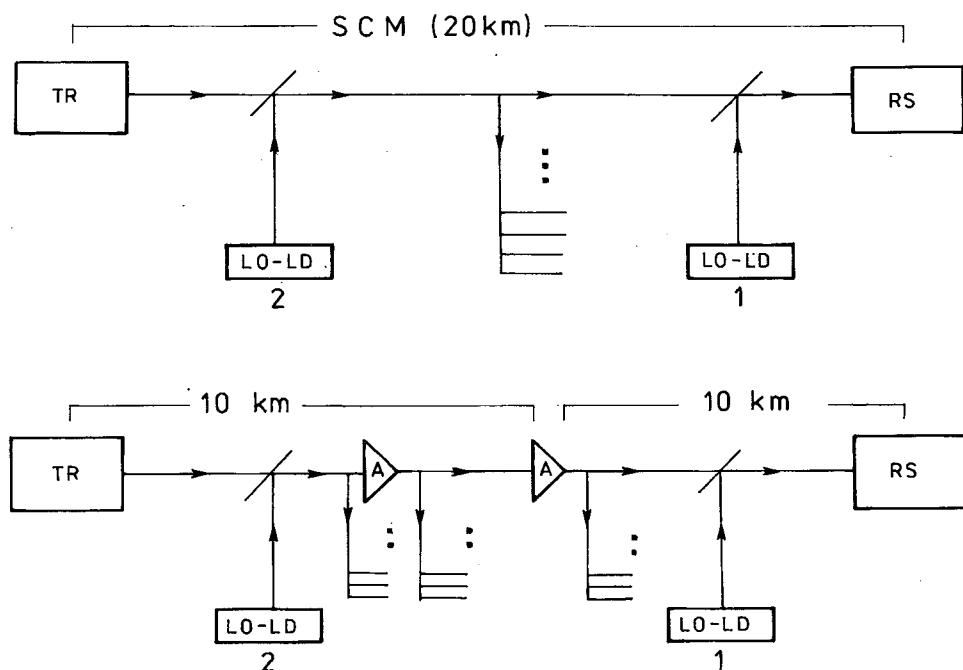


Fig. 7. Model optical coherent SCM system

6. EFFECTS OF NOISE

6.1. ADJACENT CHANNEL CROSSTALK

The required channel spacing, for such system depends on the IF, bit rate, modulation format, and bandpass filter. For FSK systems channel spacing 4 to 5 times the bit rate is required. This large spacing is necessary to avoid interference from the image spectrum of the adjacent channels, which is present when the heterodyne signal is converted from the optical domain to the electrical domain with multichannel coherent SCM systems. When we use a single optical carrier, there is no difference between the channel spacing in the optical and electrical domain, this means that values nearly $D=2B_r$ are sufficient. The level of the crosstalk is evaluated to decide the necessary channel spacing by measuring the signal-to-interference ratio (SIR), and usually about $(-150 \div -160 \text{ dB/Hz})$. If we employ PNC circuit, therefore, the system does not suffer degradation from channels' interferences. Therefore, we can consider this crosstalk as noise. Clearly, Fig. 8 shows the interference related to the tail of the power spectrum of adjacent channels. As a matter-of-fact, the crosstalk noise variance of channels, from (2) to $(N-1)$, [2 times the variance noise crasstalk] results from the neighbourhood channels. On the other hand, the crosstalk noise variance of channel 1 and N is equal to the variance crosstalk noise.

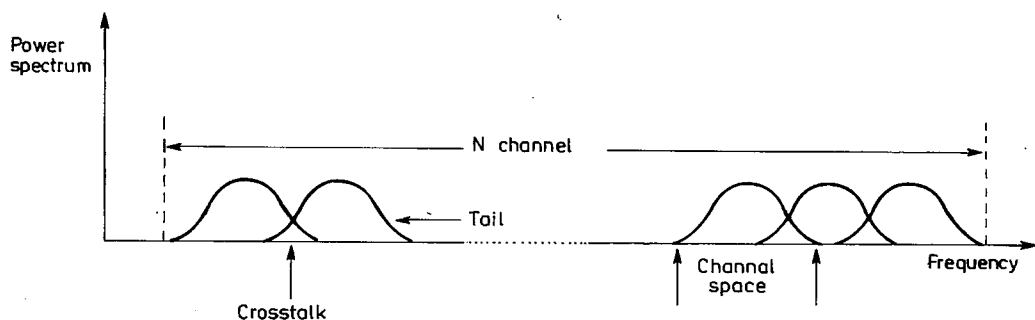


Fig. 8. The crosstalk among adjacent channels

6.2. INTERMODULATION DISTORTION (IMD)

Multichannel SCM systems are degraded by IMD, this is true for both direct detection and coherent systems. IMD results primarily from the nonlinear laser characteristics, with coherent SCM systems, the IMD is a fundamental characteristics of the nonlinear modulation format and the coherent detection process. When limiting the transmission band on a signal octave, only third-order IMD is important, in contrast to a multioctave form, both second and third-order intermodulation product (IMP's) are significant to the IMD. Clearly, third-order IMP's has been shown that products of the form $(f_h \pm f_i - f_j)$, where f is the frequency of the channel, and has varying value for every channel h, i, j , a thid-order IMP that falls directly within the frequency band of the j th in the N -channel coherent SCM system, Fig. 9 describes that. The number of third-order IMP, s denoted as k_3 , varies from channel

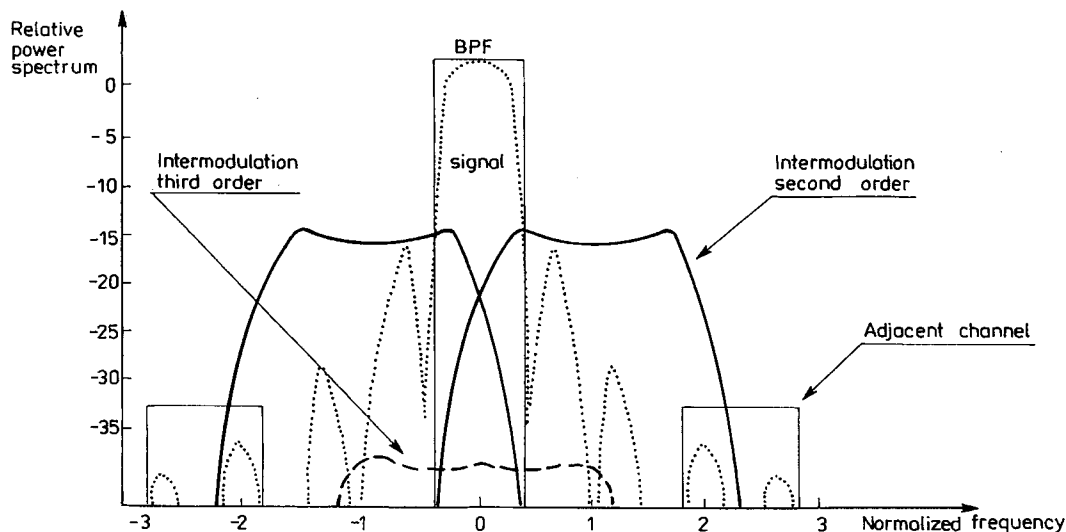


Fig. 9. The effect intermodulation second and third order

to channel with in a given system and is proportional to $[N^2]$. Therefore, the center channel ($j = N/2$) has the most IMP, s and the last channel ($j = N$) has the least IMP's, the maximum number of the third-order products is bounded by $(3N^2)/8$. As it is seen from the IMP spectrum, the convoluted version of the individual signal spectra causes a wider bandwidth. Some power is filtered out by the bandpass filter (BPF), as shown in Fig. 9. If we assume that all channels have the same magnitude and spectral shape, k_3 becomes a constant for a given system. It is seen clearly that as N increases, the number of IMP's increases, and the total IMD becomes greater.

The analysis for second-order IMP's, is similar to the approach described for third-order, a second-order term is of the form $(f_i \pm f_j)$, also, f is the frequency of channel, and has vary value for every channel, which falls directly in the desired signal, noted that in the Fig. 9. By choosing the microwave subcarrier frequency, all second-order terms fall between adjacent channels. This minimizes the degradation but does not eliminate it, the number of second-order (k_2) terms increases linearly with the channel number N and is again different for individual channels. In contrast to the simulation for third-order products, the first channel ($j = 1$) now has the most second-order IMP's, on the other side the number of second-order has varied value for every channel.

In this discussion, we have been concerned only the optical IMD that is intrinsic to coherent SCM system. IMD results from the electrical amplification of the composite microwave signal, this electrical IMD can make same degradation, but in some conditions, the electrical IMD is negligible. The IMD is signal dependent and will also increases with the detect power, when the IMD significant the CNR does not increase as rapidly with increasing signal power. To eliminate the effects of second-order IMP's the system could operate over the $(4 \div 8)$ GHz octave or in the other sentence, the system could operates over the single octave. The values for k_2 and k_3 are determined by the total channel number N and the specific channel of interest.

Now at the end of this section, we can say that in the single octave system, the center channel has most the third-order IMP's. In contrast to that the multioctave system, channel (1) has the most second-order IMP's. In the single octave system, as matter-of-fact, channel (1) requires less power but in the multioctave system requires considerably more optical power especially for the worst case channel ($j = 1$). It has been seen that the penalty due to the second-order distortion can be very significant. For example, if was $N = 40$ channel, multioctave system requires about 4.25 microwatt, in contrast to, single octave system needs about 0.5 microwatt a difference 9.3 dB.

6.3. LASER PHASE NOISE

Lase phase noise is caused by randomly occurring spontaneous emission, which is an inevitable aspect of laser operation. Each event causes a sudden jump in the phase of the electromagnetic field generated by the device. In the coherent optical SCM system, appear clearly, the laser phase noise may cause signal spectrum brooding and

distort the performance BFSK-SCM-FM system seriously. Fig. 10 shows the effect of the phase noise on the detected signal. Also, as we can see, the tail causes interference between adjacent channels. The power spectrum of a signal contaminated by the phase noise, can be expressed as convolution of the power spectra of the signal and the phase noises. We can write that as $[S(f) = S_{NPN}(f) * G_{NP}(f)]$ where $*$ denotes convolution, $S(f)$, $S_{NPN}(f)$, $G_{NP}(f)$ represent the power spectra of the signal contaminated by the phase noise, J th channel signal without the phase noise, and the phase noise, respectively. Because of spectrum brooding, the bandwidth of the bandpass filter (BPF) must be increased, to avoid signal distortion caused by the filter, as seen in the Fig. 10. We can say that, the IF bandwidth required to accommodate nearly 95 percent of the signal power in the presence of phase noise, therefore, the value of B_{IF} depends on linewidth of the laser. Related to the power spectra of the second-order and the third-order (see Fig. 9) IMP's in the neighborhood of the signal channel band, are the fractions of their power within the passband of the IF filter taking it within the passband of the BPF which is shown as a function of linewidth of the laser. Again we can say, the phase noise caused the crosstalk to adjacent channels.

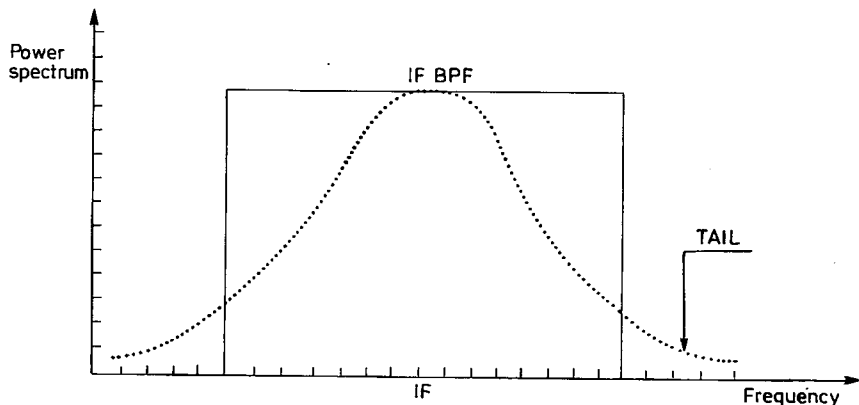


Fig. 10. The effect of the noise on the signal detection

If we want to summarize this section, we can see, to improve the performance, a very simple phase-noise-cancelling circuit can be done in the coherent FSK-SCM-FM system. The signals will be received and demodulated giving a significant output despite the presence of phase noises. This circuit also cancels the frequency drift in the local oscillator and transmitter. The PNC circuit will enable coherent SCM system to be built using conventional DFB laser.

6.4. THERMAL AND SHOT NOISE

The thermal and shot noises have a big effect on the performance of the coherent FSK-SCM-FM system. Equivalent input thermal noise current of the photodetector is about $15 \text{ PA(Hz)}^{1/2}$. Clearly, the variance functions of the shot and thermal noise

depends on, responsivity of the diode (R), bandpass intermediate frequency (B_{IF}), local oscillator power (P_{LO}), the circuit noise figure (NF), the Boltzmann constant (K), absolute temperature (T), and load resistance (r). In this section it is not important to discuss it deeply.

SUMMARY

As we can see the BFSK-SCM-FM system has improved the performance of system a better than the coherent direct detection system and the coherent SCM intensity modulation system. Clearly, the coherent FSK-SCM-FM system suffers from the phase noise, and fortunately we can use the phase noise cancelling (PNC) circuit, which improves the overall performance of the system. That means, it must use the (PNC) circuit. We have seen also that when limiting the transmission band on the single-octave, only third order IMD is of importance. On the other hand, when limiting transmission on the multilevel-octave, the second order and the third order are of importance.

In the end, as we expect in future, if we use the multilevel digital modulation in the transmitters, and amplifiers between transmitter and the receiver, the capability and capacity of the new coherent system (QFSK-SCM-FM) will become greater than the coherent BFSK-SCM-FM system. More informations about phase noise cancelling circuit are in the references [3] and [16], the example of the mathematical analysis is in 10. More informations in this domain was shown also in [2, 9, 11, 13].

REFERENCES

1. S. Watanabe, I. Yokota, T. Naito, T. Chikama, H. Kuwahara, T. Togue: *Polarisation-intensitive optical coherent SCM system using mixed lightwave transmission of signal and local laser light*. Electronics Letters 14th February 1991, Vol. 27, No. 4
2. R. Gross, M. Schmidt, R. Olshansky, V. Lanzisera: *Coherent transmission of 60 FM-SCM video channels*. IEEE Photonics Technology Letters, Vol. 2, No. 4, April 1990
3. R. Gross, R. Olshansky, M. Schmidt: *Coherent FM-SCM system using DFB laser and phase noise cancellation circuit*. IEEE Photonics Technology Letters, Vol. 2, No. 1, January 1990
4. R. Gross, R. Olshansky: *Performance evaluation of digital and analogue video distribution systems using coherent SCM*. Electronics Letters, 7th December 1989, Vol. 25, No. 25
5. R. Gross, R. Olshansky, P. Hill: *20 Channel Coherent FSK system using SCM*. IEEE Photonics Technology Letters, Vol. 1, No. 8, August 1989
6. Y.H. Lee, J. Wu, H.W. Tsa o: *Uniform CNR design rules for coherent SCM system with multioctave frequency allocation*. Electronics letters 28th March 1991, vol. 27, No. 7
7. R. Gross, R. Olshansky: *Multichannel coherent FSK experiments using SCM techniques*. Journal of Lightwave Technology, Vol. 8, No. 3, March 1990
8. R. Gross, W. Rideout, R. Olshansky, G. Joyce: *Heterodyne video distribution system sharing transmitter and local oscillator lasers*. Journal of lightwave technology, Vol. 9, No. 4, April 1991
9. Y.H. Lee, J. Wu, H.W. Tsa o: *The impact of laser phase noise on the coherent SCM system*. Journal of lightwave technology, Vol. 9, No. 3, March 1991

10. S. Watanabe, T. Terahara, I. Yokota, T. Naito, T. Chikama, H. Kuwahara: *Optical coherent broad-band transmission for long-haul and distribution system using SCM*. Journal of lightwave technology, Vol. 11, No. 1, January 1993
11. Y.H. Lee, H.W. Tsao, J. Wu: *Analysis of the multichannel coherent FSK-SCM system with pilot carrier and PNC cancelling scheme*. 140/Spie Vol. 1372 Coherent lightwave communications, 1990
12. J. Wu, H.W. Tsao, K.P. Ho, Y.H. Lee: *Analysis of simultaneous digital and analogue transmission in a coherent optical SCM system*. Electronics Letters, 11th April 1991, Vol. 27, No. 8
13. R. Gross, R. Olshansky: *Third-order intermodulation distortion in coherent SCM systems*. April 1989, No. 4
14. R. Gross, R. Olshansky, P. Hill: *Five channel coherent heterodyne SCM system*. IEEE Photonics technology letters, July 1989, Vol. 1, No. 7
15. R. Gross, R. Olshansky, M. Schmidt: *SCM coherent lightwave systems for video distribution*. IEEE Journal on selected area in communication, September 1990, Vol. 8, No. 7
16. S. Watanabe, T. Naito, T. Chikama, H. Kuwahara: *Phase noise cancellation in a carrier recovered optical heterodyne receiver*. Journal of lightwave technology, December 1992, Vol. 10, No. 2
17. Pieter W. Hoijmans: *Coherent optical system design*. John Wiley & Sons, U.K. 1994

MIKHAIEL SALIM MAKHOUL

MODULACJA CZĘSTOTLIWOŚCI OPTYCZNEJ DLA OPTYCZNYCH SYSTEMÓW Z MODULACJĄ PODNOŚNYCH

S u m m a r y

W artykule podano opis system spójnej binarnej modulacji z kluczowaniem częstotliwościowym podnośnej mikrofalowej. Pokazano możliwość budowy koherentnego optycznego układu, który umożliwia transmisję sygnałów w szerokim pasmie modulacji częstotliwości przy utrzymaniu szumów na względnie niskim poziomie. Dzięki wprowadzeniu układu dławienia szumów otrzymano szerokopasmową transmisję przy małych szumach odbiornika.

Słowa kluczowe: układy optyczne, transmisja sygnałów, laser DFB, modulacja częstotliwości.

SPIS TREŚCI

J. Zarębski: Elektrotermiczne wielkosygnałowe modele wybranych elementów półprzewodnikowych	93
P. Wolański: Specyfikacja i symulacja ASM w języku VHDL	129
J. Zarębski: Analiza elektrotermicznych stanów przejściowych w układach elektronicznych	145
Z. Kozieł: Interpolacja temperatury w modelach wielowymiarowych metodą odwzorowania funkcji wielu zmiennych w odcinek prostej	171
S. Rosłonec, Tahar Habib: Dwudiodowy, mikrofalowy filtr kierunkowy	195
Z. Lisik: Rozwój konstrukcji tranzystorów IGBT	207
D. Czekaj, Z. Surowiak, A.V. Gorish, A.A. Kurpienko, A.E. Panich: Matematyczny model wielowarstwowego czujnika piezoceramicznego	227
M. Węgrzyn: Implementacja współbieżnych kontrolerów cyfrowych z wykorzystaniem PLD ...	235
M. S. Makhoul: Modulacja częstotliwości optycznej dla optycznych systemów z modulacją podnośnych	253

CONTENTS

J. Zarębski: Electrothermal large-signal models of selected semiconductor devices	93
P. Wolański: VHDL specification and simulation of algorithmic state machine	129
J. Zarębski: Analysis of electrothermal transients in electronics circuits	145
S. Kozieł: Interpolation of temperature in multidimensional models by transformation of multi-variable functions into line segment	171
S. Rosłonec, T. Habib: Two PIN diode microwave directional filter	195
Z. Lisik: Development of IGBT transistors design	207
D. Czekaj, Z. Surowiak, A.V. Golrish, A. Kurpienko, A.E. Panich: Mathematical model of the multilayered piezoelectric sensor	227
M. Węgrzyn: PLD-based implementation of concurrent controllers	235
M. S. Makhoul: Optical frequency modulation for coherent optical subcarrier multiplexing transmission	253

Subscription for external subscribers:

The promotional subscription price in 1996 is \$ 90 including postage for institutions. Subscriptions should be sent to the publisher, Polish Scientific Publishers PWN Ltd, Journal Division, Miodowa 10, 00-250 Warsaw, POLAND, fax (48) (22) 26 09 50, (48) (22) 26 71 63, with a copy by fast to Electronics and Telecommunications Quarterly 00-665 Warsaw, ul. Nowowiejska 15/19, p. 470. Subscription is occupied after showing a checque or transfer documents. Our bank account is as follows: Bank account: PBK VIII O/Warszawa nr 370028-1052. At subscriber's request this journal will be air mailed at additional postage to 50% dross price to European countries and 65% overseas.

Subscription orders available through the local press distribution or through the Foreign Trade Enterprise ARS Polona, 00-068 Warszawa, Krakowskie Przedmieście 7, Poland.

Kwartalnik Elektroniki i Telekomunikacji

WARUNKI PRENUMERATY

Wpłaty na prenumeratę przyjmowane są na okresy roczne

Na teren kraju prenumeratę przyjmują jednostki kolportażowe RUCH S.A., urzędy pocztowe oddawcze właściwe dla miejsca zamieszkania lub siedziby prenumeratora oraz doręczyciele w miejscowościach, gdzie dostęp do urzędu jest utrudniony, a także Zakład Kolportażu Wydawnictwa SIGMA-NOT

Termin przyjmowania przez RUCH S.A. prenumeraty krajowej i na zagranicę oraz przez Poczta Polską (tylko prenumerata krajowa) są kwartalne: do 20.XI na I kw., do 20.II na II kw., do 20.V na III kw., i do 20.VIII na IV kw.

Zakład Kolportażu Wydawnictwa SIGMA-NOT przyjmuje prenumeratę do 5 XII roku poprzedzającego

Dostawa zamówionej prasy następuje w sposób uzgodniony z zamawiającym, pod wskazanym adresem, w ramach opłaconej prenumeraty.

Adresy przedsiębiorstw kolportażowych i ich konta bankowe:

Zakład Kolportażu Wydawnictwa SIGMA-NOT, 00-716 Warszawa, ul. Bartycka 20, skr. poczt. 1004

Konto PBK III Oddział Warszawa nr 370015-1573-139-11

Prenumerata ze zleceniem wysyłki za granicę jest dwukrotnie wyższa od ceny krajowej. Zlecający powinien podać dokładny adres odbiorcy. Dodatkowe informacje można otrzymać pod nr telefonu 49-30-86 lub 40-00-21 wew. 249, 295, 299.

Ruch S.A. Oddział Krajowej Dystrybucji Prasy, 00-958 Warszawa, ul. Towarowa 28
Konto PBK XIII Oddział Warszawa nr 370044-16551

Dostawa odbywa się pocztą zwykłą w ramach opłaconej prenumeraty, z wyjątkiem zlecenia dostawy pocztą lotniczą, której koszt w pełni pokrywa zleceniodawca.

Prenumerata ze zleceniem dostawy za granicę jest o 100% wyższa od krajowej. Od osób lub instytucji, zamieszkających lub mieszcących się w miejscowościach, w których nie ma jednostek kolportażowych RUCH prenumeratę Kwartalnika można zamówić w Oddziale Warszawskim RUCH na konto: PBK XIII Oddział Warszawa, nr 370044-16551.

Bieżące numery można nabyć w Księgarni Wydawnictwa Naukowego PWN, ul. Miodowa 10, 00-251 Warszawa. Również można je nabyć, a także zamówić (przesyłka za zaliczeniem pocztowym) we Wzorcowni Ośrodka Rozpowszechniania Wydawnictw Naukowych PAN, Pałac Kultury i Nauki, 00-901 Warszawa w cenie podanej na okładce kwartalnika.

W roku 1996

przewidywana cena 1 egz. wyniesie	7,00 zł (70.000 zł)
prenumerata roczna w kraju	28,00 zł (280.000 zł)
prenumerata roczna za granicę	90 \$ USA