

IL DIODO TUNNEL DI REONA ESAKI

I diodi Tunnel (TD) e i diodi Resonant Tunnel (RTD)

di Carlo Mozetic

Pubblicazione gratuita per la libera distribuzione offerta da:

© OPTELEX - Optical and RF Engineering – Rome (IT)

Diritti riservati – All Rights Reserved – 04/2026

INTRODUZIONE.....	2
PARTE I : il diodo ad effetto tunnel (o diodo di Esaki).....	4
1.1 Amplificatori con i diodi tunnel.....	17
PARTE II : il diodo RTD (Resonant Tunnel Diode).....	21
2.1 Diodi tunnel risonanti a barriera tripla (TBRTD o DWRTD).....	65
RIFERIMENTI E BIBLIOGRAFIA.....	74

INTRODUZIONE

Reona è un antico nome giapponese che deriva dal latino e che vuol dire “Leona”, ovvero forte come un leone (o meglio come una leonessa), un nome frequentemente utilizzato che di solito si attribuisce a una donna ma che talvolta viene anche utilizzato al maschile, un po' come da noi il nome “Andrea”; per questo motivo il premio Nobel Reona Esaki veniva amichevolmente chiamato “Leo” da tutti i suoi collaboratori e dalla maggior parte dei suoi studenti. Esaki oggi ha già compiuto 101 anni, essendo nato a Osaka il 12 marzo 1925 ed è un noto fisico tecnologo dello stato solido, ed è proprio in suo onore che ho deciso di pubblicare questo articolo estratto in parte dal libro “Le resistenze negative” un libro piuttosto articolato di cui sono molto orgoglioso di esserne l'autore. In questo libro si affronta l'argomento del famoso “diodo tunnel”, detto anche “diodo di Esaki”, e dei suoi diretti derivati, come il diodo RTD; naturalmente trattasi di un argomento complesso, che nel libro è stato sviscerato in modo più completo con l'aggiunta di numerosi particolari costruttivi e di alcuni esempi di applicazioni tipiche, anche nel campo delle microonde.

Esaki fu il primo scienziato a trasferire le teorie quantistiche di Bohr e di Schrödinger all'interno di un semiconduttore, tanto semplice quanto utile per le applicazioni più disparate. La scoperta fu fatta per caso, come accade quasi sempre, tuttavia la sua spiegazione “quantistica” avvenne successivamente, quando vennero introdotti i modelli a bande di energia per i quali anche il nostro Enrico Fermi diede il suo prezioso contributo. Per questo motivo Esaki si guadagnò il premio Nobel per la fisica nel 1973 quando ancora lavorava per la società Tokyo Tsushin Kogyo, la Sony di oggi. In seguito per l'IBM sviluppò tutte le teorie pionieristiche sui “superlattice” ovvero i superreticoli di semiconduttori creati in laboratorio artificialmente con tecniche che ancora oggi potremo definire “spaziali” per quanto sia difficile poterli immaginare; dobbiamo a lui l'esistenza delle tecnologie epitassiali MBE (Molecular Beam Epitaxy), e quelle per la realizzazione di strutture complesse depositate in fase vapore, come la CVD (Chemical Vapor Deposition), utilizzata per far crescere i diamanti sintetici ma anche per realizzare alcuni potenti semiconduttori per microonde (GaN), come anche la MOCVD, ossia

la “Metal-Organic Chemical Vapor Deposition “, detta anche MOVPE, cioè “Metal-Organic Vapor Phase Epitaxy”, una tecnica piuttosto recente per far crescere strati sottilissimi di materiali cristallini sovrapposti per semiconduttori miniaturizzati che vengono direttamente depositati sul substrato in cui sono presenti altri componenti discreti di un circuito elettronico.

A Leo Esaki, oltre alla scoperta del diodo tunnel originale, bisogna attribuire anche le scoperte del diodo tunnel risonante a doppia barriera (RTD, Resonant Tunnel Diode o DBRTD cioè Double Barrier RTD), quello risonante a tripla barriera (TBRTD ovvero Triple Barrier RTD o anche DWRTD cioè Double Well RTD), del diodo TUNNETT e del diodo QWITT, ossia del diodo “Quantum Well Injection Transit Time”, cioè un diodo che utilizza un RTD come elemento di iniezione delle cariche lungo la regione di drift. Però Esaki non si è fermato nello studio di dispositivi utili solo nel campo dell'elettronica e delle microonde, a lui infatti si attribuisce anche l'intuizione nel recente sviluppo delle memorie T-SRAM (Tunnelling-based Static Random Access Memories), cioè memorie miniaturizzate ad alta densità e ultra-veloci costituite da semplici matrici di diodi RTD utilizzati come strutture “bi-stabili” riprogrammabili, realizzate anche per lavorare immerse nell'azoto liquido a -196°C (77°K).

Le culture orientali hanno sempre avuto un loro fascino, ma le tecnologie made-in-Japan sono quelle che più di tutte meritano una nota di riguardo; nel Paese dei ciliegi in fiore da sempre sono esistite figure storiche di rilievo e Reona Esaki non fa eccezione. Sony, Anritsu, NEC, JVC, Kenwood, Fujitsu, Hitachi, Mitsubishi, Murata, TDK, Toshiba, Renesas, Kyocera, Panasonic (ex Matsushita), Nintendo, Sharp, Sumitomo, Canon e Casio, sono solo alcune delle società giapponesi leader nella ricerca e nella fabbricazione di componenti e di dispositivi elettronici all'avanguardia che non hanno nulla da invidiare alle più famose e blasonate case costruttrici europee ed americane, senza contare il settore automotive in cui al momento il Giappone è considerato il leader mondiale. Per me è un vero onore pubblicare questo articolo per ricordare sempre e comunque le imprese e le capacità straordinarie di una delle migliori menti di questo secolo e del secolo scorso, spero che i lettori sapranno apprezzare e condividere questo mio pensiero.

PARTE I : il diodo ad effetto tunnel (o diodo di Esaki)

Per presentare correttamente il diodo tunnel sono costretto a rispolverare alcuni concetti di base di fisica dei semiconduttori, altrimenti correrei il rischio di annoiare i lettori meno preparati su questa materia così importante e articolata. Non me ne vogliano gli esperti che mi auguro comprendano la necessità di inserire questa mia breve introduzione.

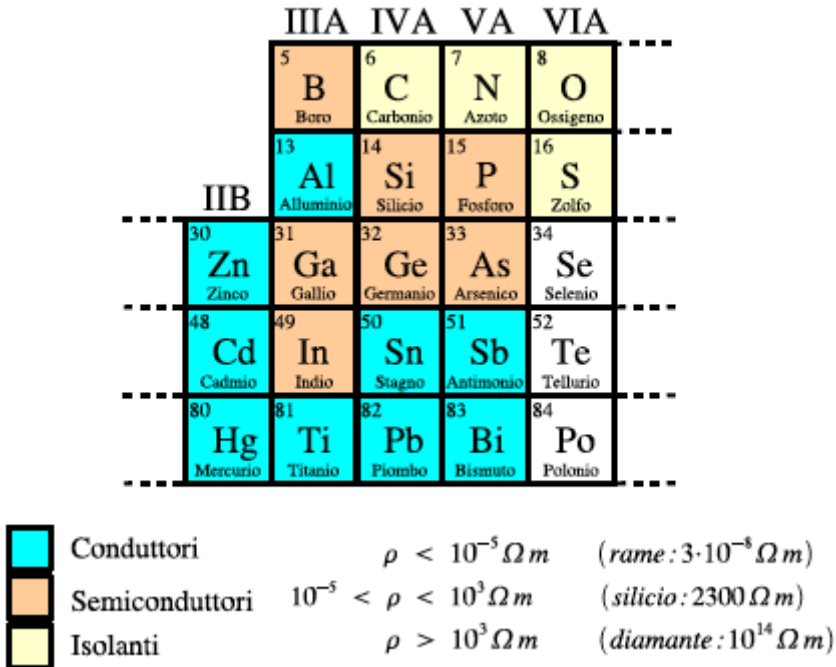


Fig.1.1 Principali materiali semiconduttori utilizzati nell'industria

In fig.1.1 è rappresentata la distinzione esatta fra materiali conduttori, isolanti e semiconduttori, inclusa la loro posizione nella tavola periodica. In realtà come vedremo esistono anche dei materiali "composti" da due o più elementi che presentano proprietà elettriche simili a quella dei principali semiconduttori, talvolta con caratteristiche elettriche molto interessanti, come ad esempio l'arseniuro di gallio (GaAs), il fosfuro di indio (InP), il carburo di silicio (SiC) o il nitruro di gallio (GaN).

In natura anche i semiconduttori non si trovano mai allo stato puro, essi sono sempre mescolati ad altre sostanze e pertanto necessitano di una serie di trattamenti per ridurre più possibile tali impurità. Quando la sporcizia residua offre un contributo irrilevante alle sue proprietà elettriche il semiconduttore si definisce “puro” o “intrinseco”, esso è formato da una struttura atomica regolare identificabile con una maglia omogenea di nuclei atomici unita insieme da legami covalenti. Nel caso del silicio che ha quattro elettroni di valenza la maglia atomica appare come nella figura 1.2(a).

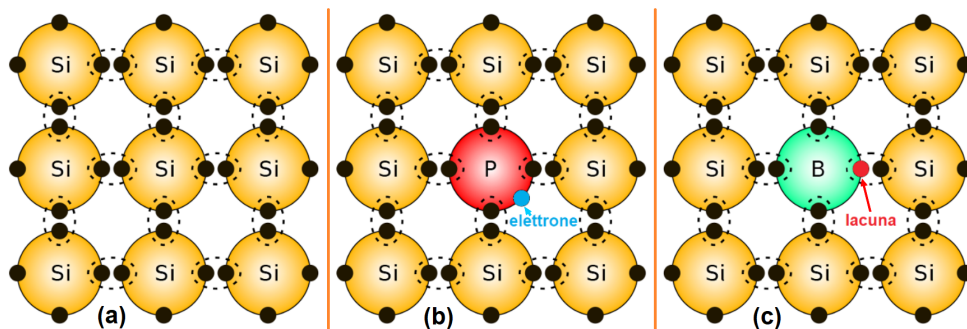


Fig.1.2 Struttura atomica del silicio: (a) regolare o “intrinseco”; (b) drogato con fosforo [donatore, N]; (c) drogato con boro [accettore, P]

In questo caso ciascun atomo di silicio condivide n.4 elettroni di valenza con gli atomi adiacenti, l'energia intrinseca si bilancia e il semiconduttore offre semplicemente una resistenza elettrica al passaggio della corrente.

Per trasformare un semiconduttore intrinseco in uno “estrinseco”, cioè un materiale avente una energia intrinseca sbilanciata verso una polarità qualsiasi è necessario “doparlo” o “drogarlo” con impurità ben definite: nel caso di un semiconduttore drogato “N” si usa un'impurità pentavalente, ad esempio il fosforo o l'arsenico; nel caso di un semiconduttore drogato “P” si usa un'impurità trivalente, ad esempio il boro o il gallio. Il tipo “N” presenta un eccesso di elettroni perché ciascun atomo dopante mette a disposizione un elettrone libero in più, vedi figura 1.2(b), mentre il tipo “P” presenta un difetto di elettroni perché ciascun atomo dopante mette a disposizione una lacuna libera in più, vedi figura 1.2(c). Considerando la distanza fra gli atomi alla temperatura ambiente la concentrazione degli atomi in un cm^3 di silicio

è orientativamente di 5×10^{22} ; invece il livello di concentrazione delle impurità dopanti di solito è compresa fra i 10^{13} atomi per cm^3 e i 10^{18} atomi per cm^3 , oltre questo limite il semiconduttore si definisce “degenerato”, di seguito chiarirò anche il significato di questo strano termine.

All values at 300°K	E_g (eV)	μ_n ($\text{cm}^2/\text{V}\cdot\text{s}$)	μ_p ($\text{cm}^2/\text{V}\cdot\text{s}$)	*Vaporizes		a (Å)	ϵ_r	Density (g/cm^3)	Melting point (°C)	
				m_n^*/m_0 (m_l, m_t)	m_p^*/m_0 (m_{lh}, m_{hh})					
Si	(i/D)	1.11	1350	480	0.98, 0.19	0.16, 0.49	5.43	11.8	2.33	1415
Ge	(i/D)	0.67	3900	1900	1.64, 0.082	0.04, 0.28	5.65	16	5.32	936
SiC (α)	(i/W)	2.36	900	—	0.6	1.0	3.08	10.2	3.21	2830
AlP	(i/Z)	2.45	80	—	—	0.2, 0.63	5.46	9.8	2.40	2000
AlAs	(i/Z)	2.16	1200	420	2.0	0.15, 0.76	5.66	10.9	3.60	1740
AlSb	(i/Z)	1.6	200	300	0.12	0.98	6.14	11.0	4.26	1080
GaP	(i/Z)	2.26	300	150	1.12, 0.22	0.14, 0.79	5.45	11.1	4.13	1467
GaAs	(d/Z)	1.43	8500	400	0.067	0.074, 0.50	5.65	13.2	5.31	1238
GaN	(d/Z, W)	3.4	1100	—	0.19	0.60	4.5	12.2	6.1	2530
GaSb	(d/Z)	0.7	5000	1000	0.042	0.06, 0.23	6.09	15.7	5.61	712
InP	(d/Z)	1.35	5400	100	0.077	0.089, 0.85	5.87	12.4	4.79	1070
InAs	(d/Z)	0.36	22600	200	0.023	0.025, 0.41	6.06	14.6	5.67	943
InSb	(d/Z)	0.18	10^5	1700	0.014	0.015, 0.40	6.48	17.7	5.78	525
ZnS	(d/Z, W)	3.6	180	10	0.28	—	5.409	8.9	4.09	1650*
ZnSe	(d/Z)	2.7	600	28	0.14	0.60	5.671	9.2	5.65	1100*
ZnTe	(d/Z)	2.25	530	100	0.18	0.65	6.101	10.4	5.51	1238*
CdS	(d/W, Z)	2.42	250	15	0.21	0.80	4.137	8.9	4.82	1475
CdSe	(d/W)	1.73	800	—	0.13	0.45	4.30	10.2	5.81	1258
CdTe	(d/Z)	1.58	1050	100	0.10	0.37	6.482	10.2	6.20	1098
PbS	(i/H)	0.37	575	200	0.22	0.29	5.936	17.0	7.6	1119
PbSe	(i/H)	0.27	1500	1500	—	—	6.147	23.6	8.73	1081
PbTe	(i/H)	0.29	6000	4000	0.17	0.20	6.452	30	8.16	925

Tabella 1.1 Proprietà fisiche dei materiali semiconduttori

Ma adesso fermiamoci un attimo per commentare le proprietà fisiche dei vari semiconduttori riportate in tabella 1.1. In particolare si noti il differente punto di fusione dei vari materiali, quelli che sopportano temperature più elevate sono il carburo di silicio (SiC) e il nitruro di gallio (GaN), e in effetti quelli sono i semiconduttori più promettenti perché più robusti per realizzare le nuove famiglie di dispositivi elettronici di elevata potenza. Per quanto riguarda invece l'elevata mobilità elettronica, essenziale nei dispositivi per onde millimetriche, osservate i valori μ_n dell'arseniuro di gallio (GaAs), del fosfuro di indio (InP) e dell'arseniuro di indio (InAs) e poi confrontateli con quelli dell'antimoniuro di indio (InSb), che tipicamente viene utilizzato invece nelle termo-camere a infrarossi per le sue note proprietà

fotovoltaiche a causa del band-gap estremamente ridotto, ovvero lo spessore della banda proibita. Ma si noti anche il suo punto di fusione estremamente basso che lo rende meno adatto agli impieghi più gravosi. Esisterebbe anche un semiconduttore composto a lega ternaria chiamato arseniuro di gallio-indio, o anche indio di arseniuro di gallio (InGaAs o GaInAs), che non appare in tabella ma che ha eccellenti caratteristiche per applicazioni fotoniche ad alta velocità, ad esempio i laser all'infrarosso e i fotorivelatori in fibra ottica per telecomunicazioni.

Il processo di doping di un semiconduttore avviene di solito dopo la crescita del cristallo intrinseco (o già dopato) per “diffusione” o per “impiantazione ionica”, il primo è un processo termico, il secondo invece è di tipo elettromeccanico ma è decisamente più complesso e costoso. Rimando al lettore l'onere di un eventuale approfondimento di questi due processi separatamente, in autonomia. I semiconduttori per microonde, invece, vengono tutti realizzati per epitassia, ovvero la sovrapposizione di vari strati di semiconduttore già dopati in fase vapore (VPE, Vapor Phase Epitaxy) che vengono fatti crescere verticalmente su un supporto metallico che funge da contatto elettrico e da principale diffusore termico. Questo processo è estremamente costoso, ma consente di ottenere dispositivi di altissima qualità con superfici di giunzione perfette e soprattutto ripetitive nel caso di produzioni di massa. L'elevato costo dei diodi tunnel e degli altri diodi che tratteremo in questo libro dipende soprattutto da questo aspetto che non è affatto da sottovalutare. Tuttavia esistono sul mercato anche dei transistor planari epitassiali a basso costo che offrono caratteristiche elettriche particolari e che pertanto la loro realizzazione diventa incompatibile con altri processi costruttivi.

Il termine “giunzione” si applica per qualsiasi semiconduttore utilizzato nell'industria, nessuno escluso; due elementi semiconduttori dopati in modo differente che vengono in contatto fra loro realizzano una giunzione sulla superficie di contatto. La giunzione si dice omogenea qualora le parti che vanno in contatto sono realizzate col medesimo materiale semiconduttore, indipendentemente dal tipo di agente dopante utilizzato, altrimenti la giunzione si definisce eterogenea, cioè composta da due distinti materiali semiconduttori. Le giunzioni eterogenee costringono l'uso dell'epitassia

come unico metodo di costruzione e pertanto tale tipo di giunzione si usa esclusivamente per produrre dispositivi semiconduttori speciali, di norma quasi tutte le giunzioni sono sempre omogenee.

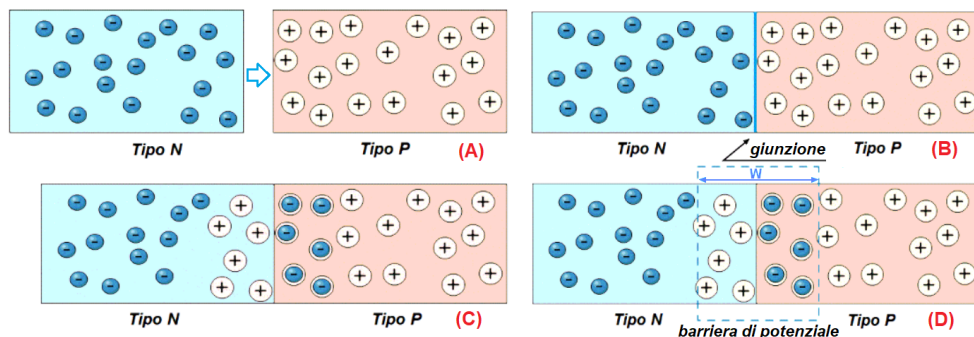


Fig1.3 Barriera di potenziale causata da una giunzione PN

Nel momento in cui due materiali semiconduttori dopati con polarità opposta vengono in contatto fra loro si forma una barriera di potenziale, cioè una parte degli elettroni in eccesso del semiconduttore dopato “N” si trasferiscono e vanno a riempire una parte delle lacune in eccesso del semiconduttore dopato “P” (vedi fig.1.3). Lo spessore “W” di tale barriera è inversamente proporzionale alla concentrazione di doping dei due elementi; essa agisce come un piccolo generatore di tensione interno avente una polarità opposta rispetto allo sbilanciamento energetico causato dal doping; in questo modo, in assenza di perturbazioni, non avverrebbe alcun passaggio di corrente qualora i contatti esterni della giunzione PN fossero chiusi in cortocircuito. Nelle stesse condizioni ma in presenza di stimoli esterni (luce, radiazioni, campi magnetici, ecc.) il semiconduttore produce una debole corrente proporzionale all'intensità dello stimolo.

Applicando ai contatti esterni alla giunzione PN una differenza di potenziale elettrico di segno concorde con la polarità indotta dal doping lo spessore della barriera diminuisce fino quasi a sparire, in questo caso una volta superato il potenziale della barriera la corrente può circolare liberamente attraverso il semiconduttore. Viceversa se la d.d.p. applicata esternamente ha un segno discorde con la polarità indotta dal doping lo spessore della barriera aumenta in proporzione al valore di tensione applicata opponen-

dosi al passaggio di corrente. La rigidità dielettrica offerta dalla barriera non è infinita, pertanto quando il campo elettrico in polarità inversa supera un certo limite la giunzione va in breakdown, cioè si distrugge. Tale limite è proporzionale alle dimensioni fisiche del semiconduttore tuttavia anche il livello di concentrazione delle impurità dopanti hanno un ruolo per stabilire questo limite visto che da esse dipende il valore di una seconda soglia che anticipa quella di breakdown e che non è altrettanto distruttiva: si tratta della famosa “valanga zener” o “breakdown zener”, un fenomeno elettrico attribuito all'omonimo fisico statunitense che per primo nel 1934 pubblicò uno studio importante col quale analizzava e descriveva nel dettaglio il breakdown nei materiali isolanti.

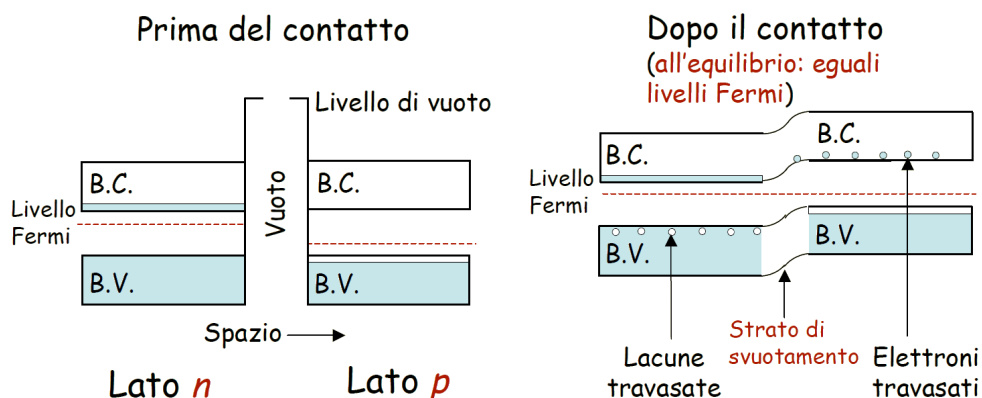


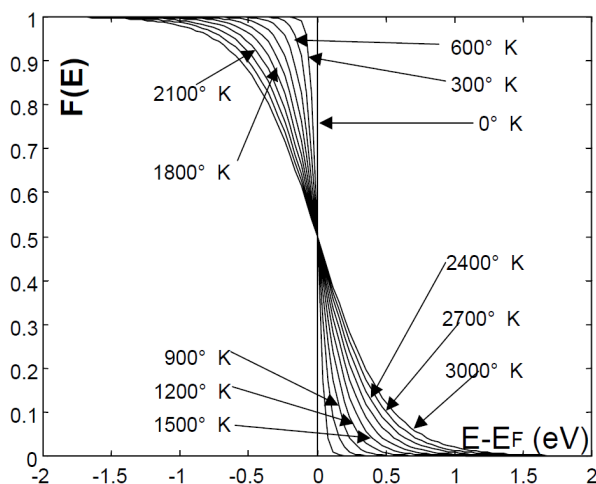
Fig.1.4 Effetto di una giunzione PN nel modello a bande

Si può dire che il lettore a questo punto ha potuto acquisire una preparazione sufficiente sui diodi a semiconduttore per iniziare a studiare cos'è un diodo tunnel, ma per fare questo è necessario introdurre un nuovo modello di studio sicuramente più adatto per focalizzare il meccanismo col quale avvengono alcuni fenomeni fisici ormai noti: si tratta del modello a bande rappresentato in fig.1.4. Sull'asse orizzontale è rappresentato lo spazio, o meglio la distanza dalla giunzione, su quello verticale c'è l'energia, di solito indicata in elettron-Volt (eV); verticalmente distinguiamo tre bande distinte, dal basso verso l'alto abbiamo: la banda di valenza, la banda proibita e infine la banda di conduzione. Osserviamo che all'interno della banda proibita c'è il livello Fermi che, nel caso di assenza del contatto di giunzione è

leggermente fuori asse, in particolare esso si avvicina di più alla banda di conduzione in un elemento semiconduttore dopato "N", al contrario esso si avvicina alla banda di valenza in un elemento dopato "P". Nel caso di un semiconduttore intrinseco il livello Fermi si trova esattamente al centro della banda proibita, esso rappresenta la probabilità che ha un elettrone di occupare uno stato con livello energetico E, la funzione di Fermi infatti vale:

$$F(E) = [1 + e^{(E-E_F)/k_B T}]^{-1} = \frac{1}{1 + e^{(E-E_F)/k_B T}} \quad [1]$$

dove k_B è la costante di Boltzmann, T è la temperatura assoluta, mentre E_F è il livello di Fermi. Poiché per $E = E_F$, $F(E) = 1/2$, indipendentemente dalla temperatura, E_F rappresenta quindi l'energia per la quale l'occupazione di quello stato è 0,5.



Rappresentazione della funzione di Fermi

Il livello Fermi non è una prerogativa dei soli semiconduttori, esso vale anche per i conduttori e per gli isolanti. Allo zero assoluto esso si riduce a una linea spezzata (fig.1.5-a) perché la probabilità di avere degli elettroni liberi aventi una energia superiore al livello Fermi è praticamente nulla, così come diventa unitaria la probabilità di averne con energia inferiore. Invece

nei metalli E_F rappresenta la massima energia che un elettrone può avere allo zero assoluto.

Nei conduttori la banda proibita non esiste perché la banda di valenza invade quella di conduzione; al contrario, negli isolanti la banda proibita è enorme per cui è necessaria una grande energia per superarla.

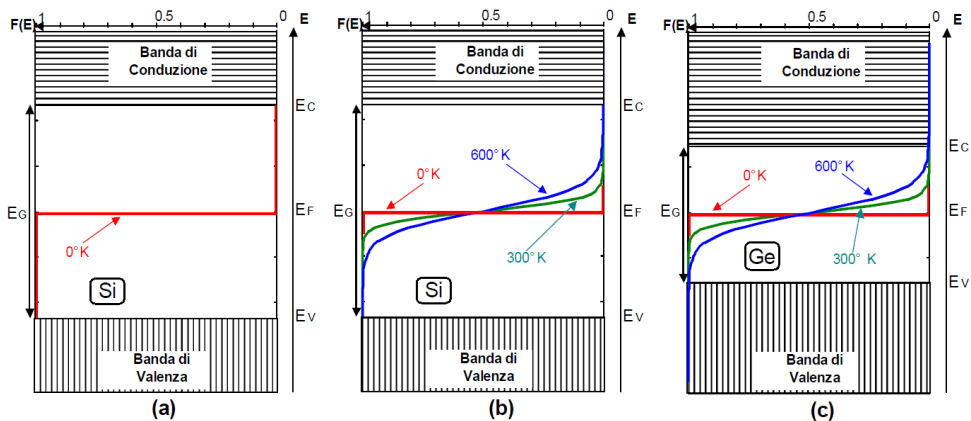


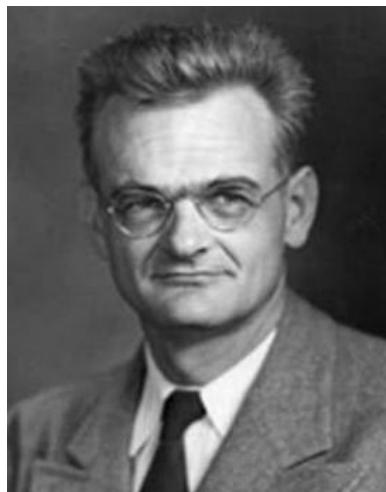
Fig.1.5 Livelli di Fermi e "gap" fra le bande di valenza e di conduzione su semiconduttori intrinseci diversi e a temperature diverse

Ma torniamo ai semiconduttori, si osservi di nuovo la fig.1.4: nel caso di una giunzione fra due elementi dopati diversamente le bande di valenza e di conduzione subiscono uno slittamento fra di loro perché il livello Fermi deve necessariamente essere comune a entrambi elementi. Si viene a creare uno strato di svuotamento attraverso il quale avviene uno scambio di elettroni e di lacune con la formazione della famosa barriera di potenziale che abbiamo visto in precedenza. In condizioni di equilibrio, ovvero senza applicare stimoli esterni di alcun tipo, il livello Fermi taglia la banda proibita esattamente al centro, ma soltanto all'interno strato di svuotamento; tuttavia questo è solo un caso limite perché rappresenta una situazione in cui entrambi gli elementi siano stati dopati con un numero uguale di impurità (uguale numero di elettroni e di lacune). Nella maggior parte dei casi, invece, il livello Fermi si trova leggermente spostato ma questo non costituisce un problema, la cosa importante da capire è che in una giunzione PN al variare della temperatura sono le bande a spostarsi e non il livello Fermi.

Cosa succede aumentando la percentuale di doping dei due elementi? Innanzi tutto il livello Fermi si sposta ancora di più dal centro avvicinandosi alle bande di valenza e di conduzione, e ciò aumenta il loro disassamento in caso una giunzione, ma contemporaneamente riduce anche la dimensione della barriera di potenziale determinando così la famosa "valanga zener",

nel caso in cui si venga ad aggiungere un campo elettrico esterno sufficiente a polarità invertita.

Aumentando ancora di più il doping dei due elementi il livello Fermi può raggiungere e addirittura superare i livelli limite delle due bande, in questo caso si dice che il semiconduttore è “degenerato”. In tale situazione la barriera di potenziale non rappresenta più un ostacolo e l'eventuale travaso di elettroni inizia ad avvenire per effetto tunnel attraverso la banda proibita, vedi fig.1.6. In particolare ciò avviene quando la concentrazione di impurità dopanti arriva ad essere 10^3 volte più elevata rispetto a quella tipica di un diodo normale, a parità di semiconduttore.



Clarence M. Zener 1905-1993

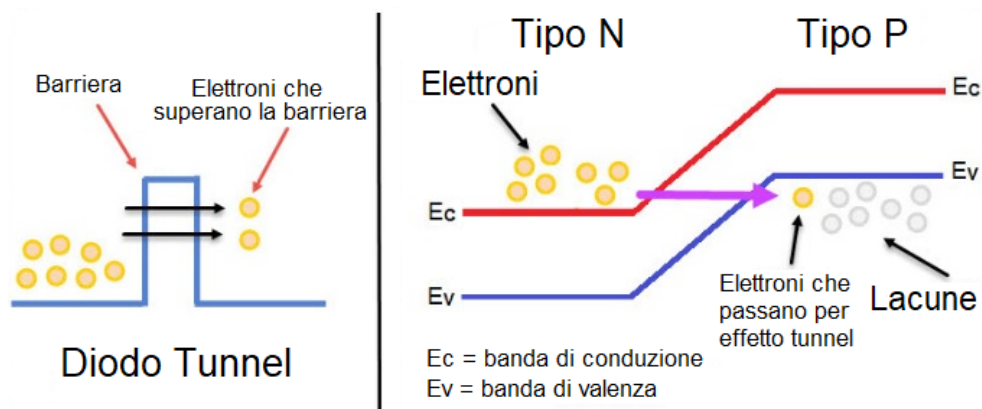


Fig.1.6 Effetto di una giunzione PN in un diodo tunnel

Tale valore consente di avere un picco di corrente sufficientemente elevato da garantire una corretta estensione della curva V-I nel tratto in cui si verifica la condizione di resistenza differenziale negativa.

Le varie possibilità di polarizzazione di un diodo tunnel sono state tutte raccolte in fig.1.7. In basso è indicato il punto di polarizzazione su un diagramma V-I, in alto invece c'è la rappresentazione corrispondente di ciò che accade nel diodo attraverso il modello a bande. Come al solito sopra c'è la banda di conduzione e sotto quella di valenza, nel mezzo la banda proibita; si osservi quanto è ampio il disassamento fra le bande.

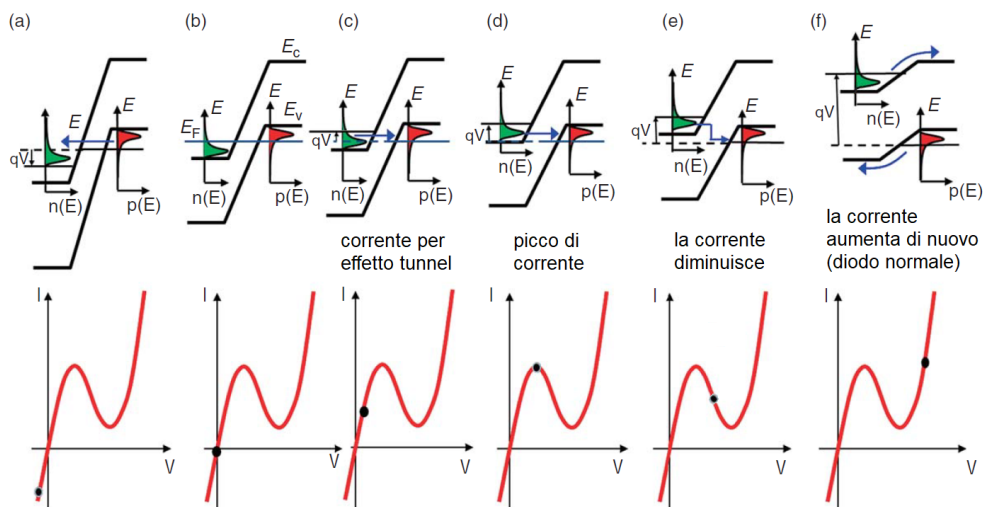


Fig.1.7 Dinamica dei principali stati di polarizzazione di un diodo tunnel

Nel caso (b) il diodo non è polarizzato, pertanto non scorre alcuna corrente e il livello Fermi E_F taglia tutte e tre le bande posizionandosi esattamente al centro. Nel caso (a), invece, il diodo è polarizzato inversamente e gli elettroni attraversano la banda proibita senza ostacoli perché l'energia nella banda di conduzione ($E_C = qV$) ha verso opposto rispetto a E_V cioè quella resa disponibile dalla banda di valenza e quindi il diodo si comporta come un conduttore elettrico. Ma ciò avviene anche nel caso (c), cioè nel caso di una leggera polarizzazione diretta applicata al diodo, con la differenza che l'energia della banda di conduzione qV questa volta ha verso concorde con E_V , tuttavia gli elettroni riescono ad attraversare comunque la banda proibita proprio a causa dell'effetto tunnel, "bucandola" ma senza danneggiare la giunzione.

Questo fenomeno si mantiene fino a quando i livelli energetici della banda di conduzione si bilanciano esattamente con quelli della banda di valenza e si portano sullo stesso piano, vedi il caso (d). Questo è il punto in cui la corrente raggiunge il suo massimo valore (picco di corrente), dopodiché la corrente inizia nuovamente a diminuire a causa del nuovo disassamento che si viene a creare fra i due livelli energetici, qV ed E_v , vedi il caso (e). Questa è la regione dove si viene a creare la famosa resistenza differenziale negativa, in cui all'aumentare del campo elettrico (la tensione applicata al diodo) la corrente invece di aumentare diminuisce. In questa condizione però, se osservate bene la figura, il limite minimo della banda di conduzione è ancora più basso rispetto a quello massimo della banda di valenza. Quando questi due limiti si equivalgono la corrente che circola nel diodo è minima, dopodiché riprende ad aumentare perché gli elettroni iniziano di nuovo a superare la banda proibita ma senza "bucarla" perché si ritrovano ad avere una energia sufficiente per "saltare" dalla banda di valenza alla banda di conduzione, esattamente come accade in un diodo normale polarizzato direttamente da un generatore esterno, vedi il caso (f).

Pertanto, l'effetto tunnel nell'omonimo diodo si verifica in modo evidente soltanto nei casi (c) e (d), mentre la resistenza negativa si produce solo quando tale effetto tende a diminuire fino ad annullarsi, cioè soltanto nel caso (e).

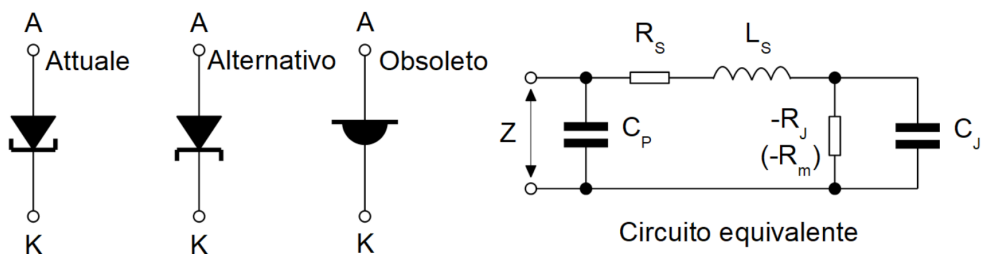


Fig.1.8 Simboli e circuito equivalente di un diodo tunnel

Tutte queste condizioni raccolte in fig.1.7 si verificano per variazioni di tensione di polarizzazione che di solito sono comprese entro 1V al massimo; quindi, con un campo elettrico esterno di verso concorde alla polarità

del diodo e avente un'intensità che corrisponde a tale valore assoluto è molto probabile trovarsi già nella condizione (f), tutte le altre si presentano con valori di polarizzazione inferiori.

In fig.1.8 oltre ai simboli del diodo tunnel troverete anche il suo circuito equivalente contemplato nella regione in cui esso presenta una NDR. La resistenza $-R_j$ rappresenta l'inclinazione della retta tangente alla curva in tale regione mentre $-R_m$ rappresenta il valore resistivo minimo riscontrabile sulla curva stessa. Il condensatore C_p rappresenta la capacità distribuita del contenitore del semiconduttore, esso di solito non viene contemplato nei calcoli di impedenza perché assorbito dall'induttanza di shunt della rete di matching che vedremo successivamente. R_s rappresenta la somma delle resistenze offerte dal conduttore metallico che contiene il diodo e delle resistenze di tutti i contatti elettrici fino al punto di giunzione; questo valore dipende dalla frequenza a causa dell'effetto pelle. L_s invece è la somma di tutte le induttanze parassitarie offerte da tutti i conduttori elettrici fino al punto di giunzione. Infine C_j è la capacità intrinseca della giunzione, tale capacità è quella che maggiormente limita la frequenza operativa del diodo.

Volendo calcolare l'impedenza di questa rete nel punto di resistività minima (ovvero nel punto di incrocio con la retta di carico ideale) e al netto della capacità C_p abbiamo:

$$Z = R_s - \frac{R_m}{1 + (\omega R_m C_j)^2} + j\omega \left[L_s - \frac{R_m^2 C_j}{1 + (\omega R_m C_j)^2} \right] \quad [2]$$

ovvero $Z = R_E + j I_M$ da cui ponendo $\begin{cases} R_E = 0 \rightarrow f_{r0} \\ I_M = 0 \rightarrow f_{x0} \end{cases}$

in cui f_{r0} rappresenta la frequenza di taglio resistiva che si ottiene annullando R_E , cioè la parte reale dell'impedenza, pertanto avremo:

$$f_{r0} = \frac{1}{2\pi R_m C_j} \cdot \sqrt{\frac{R_m}{R_s} - 1} \quad [3]$$

e avremo anche f_{x0} che rappresenta la frequenza di taglio reattiva, detta anche frequenza di auto-risonanza, che si ottiene annullando la parte immaginaria I_M e cioè:

$$f_{x0} = \frac{1}{2\pi R_m C_j} \sqrt{\frac{R_m^2 C_j}{L_S} - 1} = \frac{1}{2\pi} \sqrt{\frac{1}{L_S C_j} - \frac{1}{(R_m C_j)^2}} \quad [4]$$

Il fattore di rumore del diodo tunnel è definita invece da tre rapporti: il rapporto fra la frequenza operativa e la frequenza di taglio resistiva f/f_{r0} , il rapporto $|R_j|/R_{eq}$, cioè un parametro che determina il peso del rumore “shot” rispetto alla resistenza negativa media del diodo, e dal rapporto $R_s/|R_j|$ che identifica il rumore “termico” determinato dalla piccola ma inevitabile resistenza serie R_s dovuta alla resistività dei materiali conduttori, con i quali è costruito il diodo, rispetto alla resistenza negativa del diodo stesso $|R_j|$. Tali rapporti vengono messi insieme dalla seguente relazione:

$$F = \frac{1 + Ka}{\left(1 - \frac{R_s}{|R_j|}\right) \left[1 - \left(\frac{f}{f_{r0}}\right)^2\right]} \quad \text{dove} \quad Ka = \frac{|R_j|}{R_{eq}} \quad [5]$$

in cui Ka viene definita la “costante di rumore” che rappresenta il rapporto fra il modulo della resistenza di giunzione e la resistenza di rumore “shot” equivalente; ad esempio nei diodi tunnel al germanio la costante di rumore di solito è compresa fra 1,2 e 1,4. Alle frequenze più basse il rapporto f/f_{r0} è trascurabile, ma anche $R_s/|R_j|$ lo è (nei diodi al germanio il valore tipico è 0,005), per cui la [5] si può così approssimare:

$$F \cong 1 + Ka \cong 1 + \frac{|R_j|}{R_{eq}} \quad [6]$$

Pertanto alle frequenze più basse il maggior contributo lo dà il rumore “shot”, infatti la figura di rumore mediamente si attesta fra 10 e 20dB a frequenze intorno a 1KHz ($NF_{(dB)} = 10\log_{10}F$) mentre scende a 3 o 4dB a frequenze intorno ai 30KHz, cioè migliore rispetto a quella offerta da un tubo o da un transistor. Alle frequenze più elevate, invece, quando f è inferiore

alla metà di f_{r0} , le prestazioni legate al rumore sono davvero eccezionali perché il rapporto $|R_j|/R_{eq}$ è approssimativamente uguale all'unità mentre $R_s/|R_j|$ è pari a circa 0,01. Tuttavia all'aumentare ancora della frequenza la figura di rumore tende comunque a risalire per l'aumento di R_s provocato dall'effetto pelle; ma anche la resistenza di giunzione effettiva R_j' tende a diminuire a causa della presenza della capacità C_j la cui reattanza alle frequenze più elevate non può più essere trascurata.

$$R_j' = \frac{|R_j|}{(2\pi f C_j |R_j|)^2 + 1} \quad [7]$$

Il rumore termico appare sempre più dominante all'aumentare della frequenza, al punto che quando f comincia ad avvicinarsi a f_{x0} (che è inferiore rispetto a f_{r0}) esso diventa eccessivo e quel diodo sarà sempre più sconsigliato. Per evitare questa situazione è importante che $f < f_{x0}/\sqrt{2}$, decisamente molto meglio se $f < f_{x0}/2$. Naturalmente tutte queste considerazioni sono valide soltanto nel caso in cui il diodo tunnel fosse utilizzato come "amplificatore", e non come oscillatore.

1.1 Amplificatori con i diodi tunnel

Fra le molteplici applicazioni che prevedono l'impiego dei diodi tunnel ci sono sicuramente gli amplificatori per i piccoli segnali, piccoli perché la potenza massima ricavabile all'uscita di un diodo tunnel è dell'ordine del mW o di sue frazioni, certamente non di più. Tuttavia il diodo tunnel ha il grande vantaggio di offrire una figura di rumore molto bassa, per questo motivo ancora oggi viene utilizzato spesso come front-end di ricevitori speciali per i quali è richiesta una particolare sensibilità. Ma questo componente è semplicemente un diodo, cioè ha solo due contatti, un anodo e un catodo, non è quindi adatto per realizzare amplificatori ad alto isolamento fra la porta di uscita e quella di ingresso e questo è il suo vero limite che non è affatto da sottovalutare. In letteratura esistono infatti numerose configurazioni di amplificatori più o meno lineari che impiegano i diodi tunnel ma in questo contesto saranno analizzate soltanto strutture di tipo parallelo e a singola porta, cioè quelle utilizzate più comunemente per realizzare amplificatori per

segnali RF-MW. La domanda sorge spontanea: ma come può funzionare un amplificatore se ha una sola porta? Un amplificatore dovrebbe avere quantomeno una porta di ingresso e una porta di uscita. Beh, la risposta è più semplice del previsto: l'unica porta serve sia per l'ingresso che per l'uscita, il segnale “entra” da quell'unica porta e ne riesce amplificato. Naturalmente quella porta dovrà essere collegata a un altro circuito che riesce a separare questi due segnali, quello che entra e quello che esce, tutto qui. In seguito svelerò come funziona questo magico “trucco” ma adesso andiamo per ordine e cominciamo dalle basi.

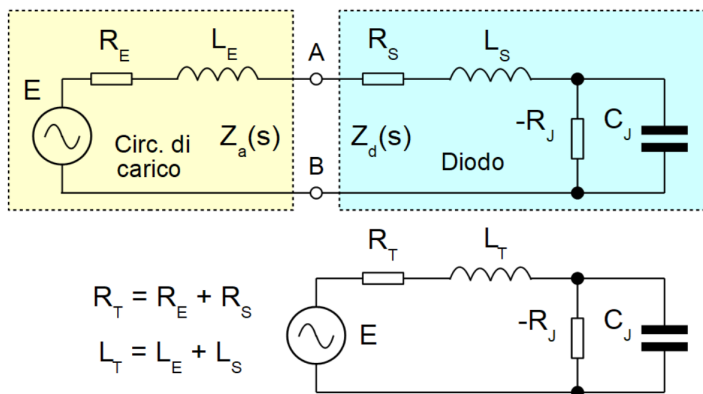


Fig.1.9 Rete di adattamento per stabilizzare un diodo tunnel

Un amplificatore è sostanzialmente un circuito “stabile” e che dovrebbe rimanere tale con qualsiasi valore di impedenza complessa, sia alla sorgente che al carico; in questo caso si dice che un amplificatore è “incondizionatamente stabile”. Per raggiungere questa condizione di solito servono delle reti di adattamento all'ingresso e/o all'uscita per escludere qualsiasi fenomeno di instabilità che porterebbe l'amplificatore ad oscillare.

Avevamo visto che un amplificatore con diodo tunnel opera sempre su frequenze inferiori alla f_{x0} , cioè a una frequenza inferiore alla frequenza di taglio reattiva; in questa situazione il diodo presenta una impedenza prevalentemente capacitiva per la quale al fine di raggiungere la stabilità è necessaria una compensazione di tipo induttivo (vedi fig.1.9). Sommando gli effetti resistivi e induttivi della sorgente (o carico) con quelli del diodo si ottiene un nuovo circuito equivalente che ora andremo ad analizzare.

Volendo distinguere i due circuiti secondo Laplace in funzione della variabile complessa (s) avremo che:

$$Z_a(s) + Z_d(s) = 0 \quad [8]$$

e la frequenza di taglio resistiva del circuito sarà la seguente:

$$f_{CR0} = \frac{1}{2\pi |R_j| C_j} \cdot \sqrt{\frac{|R_j|}{R_T} - 1} \quad [9]$$

mentre la frequenza di taglio reattiva del circuito sarà:

$$f_{CX0} = \frac{1}{2\pi |R_j| C_j} \cdot \sqrt{\frac{R_j^2 C_j}{L_T} - 1} \quad [10]$$

Si noti che la [9] e la [10], ovvero le due frequenze di taglio che si riferiscono al circuito adattato, sono diverse rispetto alla [3] e alla [4] che si riferivano invece alle due frequenze di taglio del solo diodo (vedi par. precedente). La condizione di stabilità del circuito sarà quindi soddisfatta solo se:

$$\frac{L_T}{|R_j| C_j} < R_T < |R_j| \quad [11]$$

Dalla [11] si evince che la stabilità incondizionata si ottiene per un campo di valori limitato di R_T fuori dal quale il circuito diventa instabile, oppure bistabile, vedi la tavola di riferimento di fig.1.10. Inoltre, se la relazione fra le due frequenze di taglio del diodo è sempre nota (in pratica $f_{r0} > f_{x0}$) lo stesso non si può dire per le due frequenze di taglio del circuito, cioè f_{CR0} e f_{CX0} che dipendono anche dai valori della rete di adattamento esterna, cioè da R_E e da L_E . Considerando i bassi valori della resistenza e dell'induttanza serie del diodo, i valori della rete di adattamento esterna diventano determinanti ai fini della stabilità, tuttavia essi determinano anche il guadagno del circuito perché individuano l'inclinazione della retta di carico.

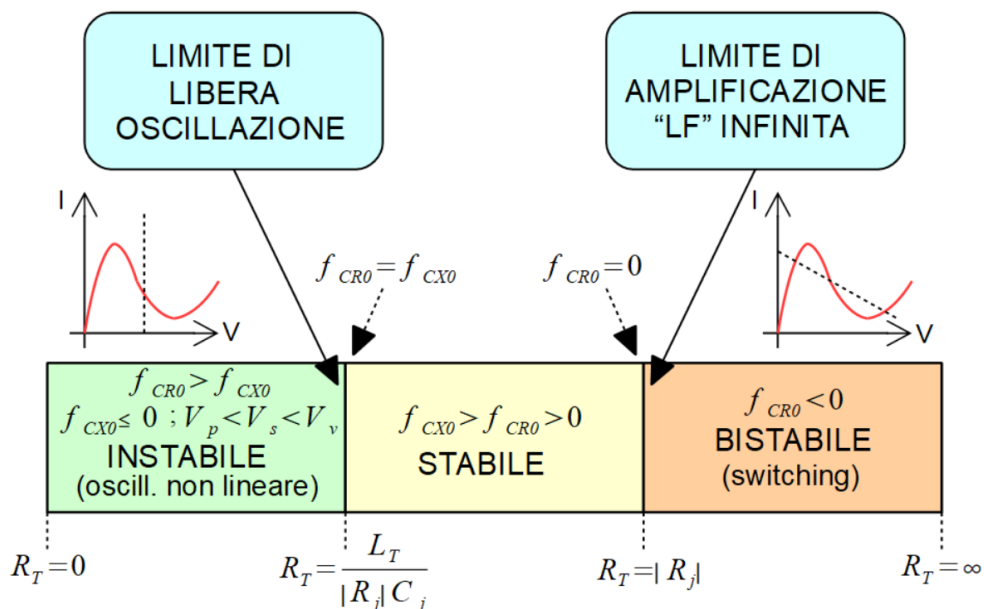


Fig.1.10 Tavola di riferimento per determinare le condizioni di stabilità di un diodo tunnel rispetto alla sua retta di carico

Osservando la fig.1.10 si intuisce che la zona operativa di un amplificatore con diodo tunnel è quella che si trova al centro, quella evidenziata in giallo, una zona che è comunque piuttosto ampia. All'interno di quella zona è possibile settare guadagni talvolta anche elevati, tuttavia in fase di progetto è consigliabile limitare sempre il guadagno al di sotto dei 20dB per scongiurare la possibilità che il circuito diventi troppo sensibile alle variazioni dell'impedenza di carico, e questo vale soprattutto nel caso degli amplificatori a riflessione. La zona evidenziata in verde sulla sinistra è la regione in cui il diodo tunnel diventa instabile e favorisce l'oscillazione spontanea nel circuito in cui esso si trova montato; in queste condizioni è possibile realizzare degli ottimi oscillatori in banda VHF e UHF, per le SHF di solito si utilizzano diodi tunnel montati all'interno di cavità risonanti. La zona evidenziata in arancio sulla destra identifica la modalità in cui il diodo tunnel può essere utilizzato come commutatore bistabile ad alta velocità di commutazione.

PARTE II : il diodo RTD (Resonant Tunnel Diode)

Con questo paragrafo introdurremo il primo di alcuni diodi particolari in grado di produrre o di amplificare segnali a frequenze elevatissime, appartenenti alle bande delle onde millimetriche o di quelle sub-millimetriche; le prime sono frequenze che si trovano sotto i 300GHz, le seconde superano questo limite e lo superano anche abbondantemente, come nel caso del diodo RTD, ovvero il Resonant Tunnel Diode, talvolta chiamato anche DBRTD, cioè Double Barrier Resonant Tunnel Diode. Di questo diodo esistono anche alcune varianti teorizzate più recentemente, come il diodo RITD (Resonant Interband Tunnel Diode) oppure il diodo TBRTD (Triple Barrier Resonant ecc.), definito anche come DWRTD (Double Well Resonant ecc.), tuttavia queste varianti non verranno analizzate in questo contesto perché sarebbe superfluo approfondire anche questi dettagli, servirebbe infatti scrivere un volume a parte soltanto per dare spazio anche a questi semiconduttori nuovissimi che stanno letteralmente rivoluzionando anche il mondo dei sensori e delle memorie statiche (T-SRAM).

La nascita del diodo RTD è avvenuta molti anni dopo la scoperta del diodo tunnel e fu originata dal medesimo scienziato giapponese, Reona Esaki, che ebbe una geniale intuizione nei primi anni '70. In realtà Esaki fu uno dei pionieri della fisica quantistica applicata ai semiconduttori, ma sapeva bene che il suo primo diodo era affetto da un problema abbastanza serio che ne limitava l'utilizzo a frequenze piuttosto "modeste", anche per quel periodo: si tratta della capacità di giunzione C_j , che era troppo elevata e come avevamo studiato in precedenza era dovuta sia al materiale impiegato che alla sezione del diodo. Tale capacità era diventata una "nota dolente" del diodo tunnel e bisognava assolutamente trovare il modo di ridurla, mantenendo però l'operatività del diodo. Cosicché nel 1973, durante un esperimento condotto insieme a un certo Raphael Tsu, Esaki ha individuato un forte picco di corrente attraverso la struttura di un diodo quando i portatori iniettati raggiungevano una certa energia in condizioni di risonanza. Da qui il passo verso l'RTD fu breve, tant'è che nel 1974 fu dimostrato il funzionamento di questo nuovo diodo da Esaki, Leroy Chang e Raphael Tsu su Applied Physics Letters Vol.24 a partire dalla pag.593. L'idea originale era quella di realizzare due sottili barriere di potenziale affiancate al posto di una

sola utilizzando materiali diversi e intercalati da uno strato di semiconduttore intrinseco altrettanto sottile. Le differenze rispetto al diodo tunnel originale erano molteplici: innanzi tutto le barriere sono diventate due e sono state poste in serie fra di loro; inoltre tali barriere non erano più costituite da una giunzione fra semiconduttori pesantemente dopati, erano semplicemente delle barriere realizzate con materiali semiconduttori aventi diversi reticoli cristallini che necessitano di livelli energetici nettamente più elevati per essere superate; infine in mezzo a tali barriere era stato posto un ulteriore strato semiconduttore privo di doping che costituisce il cosiddetto “pozzo quantico”, che tradotto in inglese diventa “Quantum Well”. Attenzione a questo dettaglio perché questa espressione la ritroveremo anche successivamente.

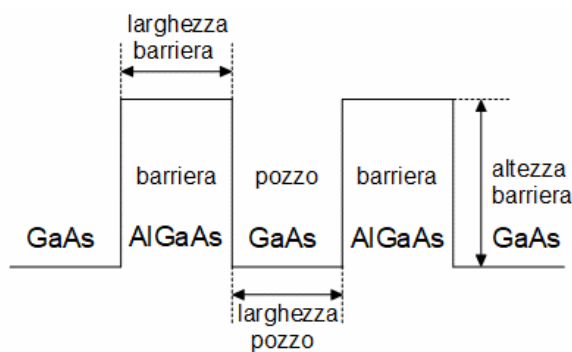


Fig.2.1 Posizionamento delle barriere in un diodo RTD o DBRTD

Ci sono n.5 parametri fisici che possono essere variati in un RTD: l'altezza delle barriere, la loro larghezza, la profondità del pozzo quantico, la sua larghezza e infine la simmetria della struttura; di solito la struttura è sempre abbastanza simmetrica, vedi fig.2.1, nel senso che le due barriere in genere hanno le stesse dimensioni fisiche. Il vantaggio di questa soluzione rispetto a quella di un normale diodo tunnel è quella di avere una capacità intrinseca di almeno 100 volte inferiore, e spesso si riesce a ridurla anche di 2-300 volte a parità di sezione del diodo. Con questi valori si possono raggiungere anche frequenze dell'ordine delle centinaia di GHz; già nel settembre 1983 facendo risuonare un diodo con le sole suscettanze che costituiscono la sua stessa struttura, T. C. L. G. Sollner, W. D. Goodhue, P. E. Tannenwald, C.D. Parker e D.D. Peck hanno potuto teorizzare una sua risonanza a frequenze

fino a 2,5THz (2500GHz), un limite mai contemplato prima di allora (App. Phis. Lett. Vol.43 pp. 588-590). Il diodo RTD infatti non prevede l'utilizzo di un contenitore, esso è un chip realizzato direttamente su un substrato dove è già presente il circuito elettrico in grado di farlo funzionare in base alle esigenze, come oscillatore, come amplificatore oppure come dispositivo bi-stabile, esattamente come accade per un diodo tunnel normale che però, invece, contempla l'utilizzo di un contenitore che gli consente di essere assemblato insieme agli altri elementi del circuito.

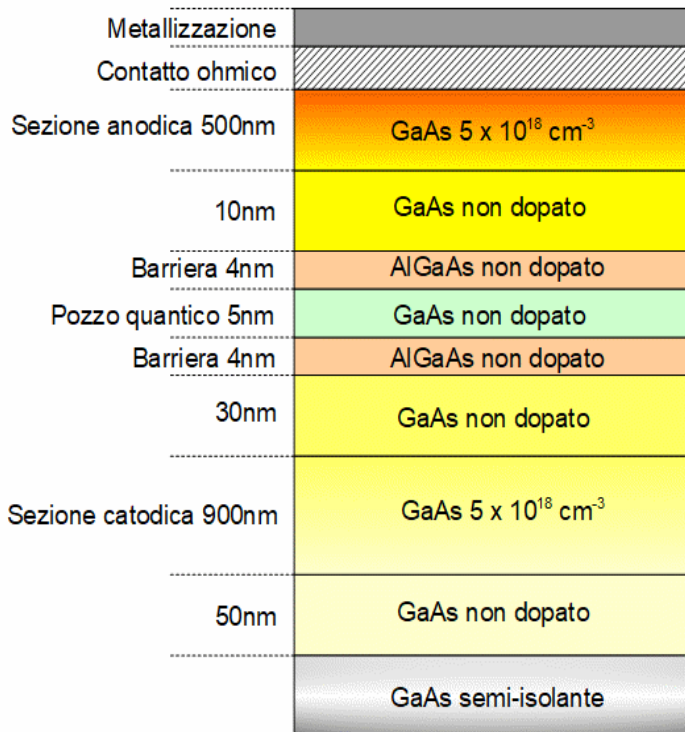


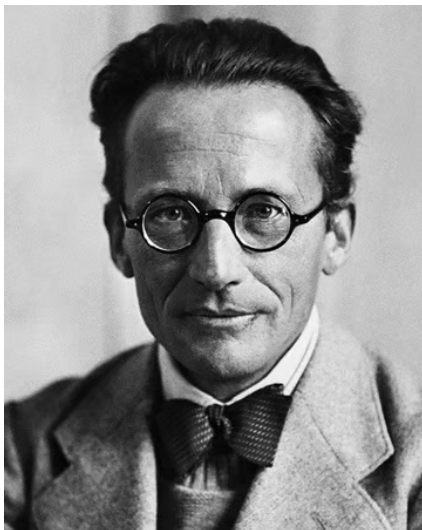
Fig.2.2 Struttura tipica di un diodo RTD o DBRTD

Inoltre, una volta montato il diodo RTD diventa parte integrante del circuito ed è impossibile smontarlo senza provocare danni irreversibili. In fig.2.2 è rappresentata una tipica struttura di un diodo RTD a doppia barriera; talvolta la sezione catodica viene definita anche “emettitore” perché emette elettroni verso il “collettore”, ovvero la definizione alternativa della

sezione anodica. Invito il lettore ad osservare che se nei dispositivi precedenti eravamo abituati a ragionare col micron come unità di misura, con questo diodo dobbiamo scendere orientativamente di tre ordini di grandezza perché parliamo di “nanometri”, ovvero la millesima parte di un micron, cioè un miliardesimo di metro. È difficile trovare nei testi tale evidenza quando si parla di dispositivi a semiconduttore per microonde e/o per onde millimetriche, ma questa differenza è fondamentale anche per capire il differente livello tecnologico che separa un diodo tunnel normale da un RTD. La figura mostra una serie di sottili strati di semiconduttore sovrapposti e separati da due barriere costituite da un materiale particolare di cui si parlerà fra poco e da un “pozzo quantico” che altro non è che un ulteriore sottile strato di semiconduttore intrinseco, cioè privo di impurità dopanti, interposto fra le due barriere.

Perché ho fatto vedere prima la struttura? L'ho fatto perché i diodi RTD, così come tutti gli altri semiconduttori per microonde e per onde millimetriche non sono costruiti così come viene costruito un diodo normale, essi si realizzano con un processo particolare che si chiama MBE, ovvero “Molecular Beam Epitaxy” col quale è possibile realizzare anche etero-strutture complesse con materiali semiconduttori che per loro natura hanno qualche difficoltà ad essere stratificati con altri metodi a causa della loro sottigliezza e/o della loro incompatibilità chimica. Per etero-strutture intendo quelle strutture composte da materiali semiconduttori con reticolo cristallino composto da due o più materiali appartenenti ai gruppi III e V della tavola periodica degli elementi. Ad esempio nel nostro caso si sta utilizzando per le due barriere l'arseniuro di gallio e alluminio (AlGaAs), la cui formula sarebbe in realtà $Al_yGa_{1-y}As$, ovvero un reticolo cristallino particolare composto da due semiconduttori, l'arseniuro di alluminio AlAs e l'arseniuro di gallio GaAs per i quali la “y” rappresenta il numero di atomi di alluminio che andranno a sostituire quelli di gallio nel reticolo cristallino. Nei diodi RTD si usano anche altri materiali semiconduttori composti, ad esempio l'arseniuro di alluminio con l'arseniuro di gallio e indio (AlAs/InGaAs) oppure l'arseniuro di alluminio e indio con l'arseniuro di gallio e indio (InAlAs/InGaAs), di solito il primo dei due materiali costituisce le due barriere dove avviene l'effetto di “Quantum Tunneling”, il secondo invece realizza il vero e proprio “pozzo quantico”.

Sì, ma che cos'è questo “pozzo quantico”, o “Quantum Well” che dir si voglia, e perché si definisce così? Ebbene il “Quantum Well” è un sottile strato di semiconduttore intrinseco, quindi non dopato, che acquisisce un livello energetico più alto rispetto al livello Fermi (per chi non conoscesse questo livello deve tornare a rivedere il paragrafo dedicato al diodo tunnel) proprio per la presenza di due barriere affiancate aventi un livello di potenziale nettamente più alto. Pertanto, senza la necessità di dover stressare il semiconduttore con un doping elevato per ottenere l'effetto tunnel come accade per l'omonimo diodo, la presenza di due sottili barriere affiancate è sufficiente per ricreare un effetto simile anche su un semiconduttore privo di doping; sarebbe questa poi in fondo la scoperta di Esaki che ha portato a realizzare i primi diodi RTD. Ma la vera differenza non è tanto nell'esistenza di questo pozzo quantico, ma si trova invece proprio nella presenza di due barriere, non soltanto di una semplice barriera di potenziale come è previsto per tutti i diodi in generale, esse sono delle vere e proprie barriere fisiche realizzate appositamente con due sottili strati di semiconduttore il cui comportamento nelle varie condizioni di potenziale elettrico applicato al diodo diventa propedeutico un approccio alla meccanica quantistica.



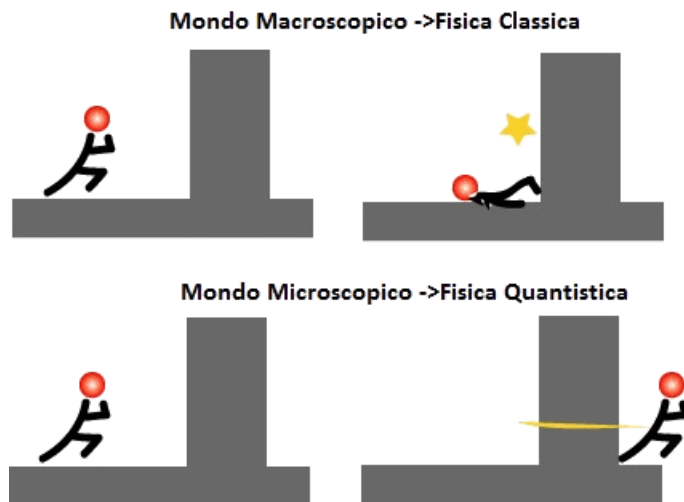
Erwin Schrödinger 1887-1961

Quest'uomo che vedete qui sulla sinistra è Erwin Schrödinger, un fisico austriaco insignito del Premio Nobel nel 1933 e famoso per i suoi contributi alla meccanica quantistica, in particolare all'equazione che porta il suo nome che lui stesso presentò nel 1926. Si tratta di una funzione differenziale a derivate parziali complessa e non relativistica che ha come incognita la funzione d'onda “ ψ ”. Schrödinger però specificò molto bene un concetto: tale relazione aveva valore solo nel mondo delle particelle atomiche e sub-atomiche, non poteva cioè essere applicato all'universo macroscopico che tutti noi conosciamo e che è governato

dalla meccanica classica. Tuttavia, a fronte di questa sua idea della meccanica quantistica, che fra l'altro permetteva di comprendere il motivo per cui soltanto alcuni valori discreti di energia sono ammessi per l'elettrone in un atomo di idrogeno, un altro gruppo di ricercatori fisici che lavoravano a Copenhagen, un gruppo di cui faceva parte anche Niels Bohr e Werner Heisenberg, svilupparono una teoria parallela, basata sull'interpretazione probabilistica della funzione d'onda di Max Born, in cui veniva considerato privo di senso definire preventivamente l'estensione del sistema fisico in cui tale teoria avrebbe potuto o dovuto essere applicata. Inoltre, per questa teoria non sarebbe stato possibile conoscere l'evoluzione di un fenomeno se esso non poteva essere "osservato", nel senso di "misurato". Questa "novità" ha creato un certo subbuglio fra i fisici di quel tempo, e anche Einstein non era assolutamente d'accordo con quest'ultima teoria tanto è vero che in questa occasione aveva pronunciato le famose frasi: "Dio non gioca a dadi col mondo" e "Credi davvero che la luna non sia lì se non la guardi?". Schrödinger invece contestò le teorie di Copenhagen attraverso il famoso paradosso del "gatto di Schrödinger", trattasi infatti di una metafora molto sottile e ingegnosa che a tutt'oggi nessuno è riuscito a confutare.

Si prende un gatto e lo si mette dentro una scatola in modo da non capire cosa stia succedendo là dentro; insieme al gatto ci si mette anche un atomo di un elemento radioattivo qualsiasi che ha il 50% di probabilità di decadere entro un'ora e magari ci si mette pure un meccanismo azionato da un contatore geiger in grado di liberare del cianuro nel caso in cui l'atomo dovesse decadere. La domanda è: dopo un'ora il gatto sarà ancora vivo o sarà morto, avvelenato dal cianuro? Secondo le teorie di Schrödinger non si potrà prevedere quale sia la condizione del gatto dopo un'ora; secondo le teorie di Bohr, invece, il gatto può essere sia vivo che morto, ma nessuno potrà mai saperlo fino a quando qualcuno non apre la scatola per vedere cosa è accaduto. Ma questa condizione vivo-morto secondo Bohr può coesistere ed è questa la stranezza che Schrödinger ha contestato, cioè in un mondo macroscopico non è possibile far coesistere un livello di incertezza così determinante alla pari di ciò che invece può accadere nel mondo delle particelle atomiche o sub-atomiche. Non solo: Schrödinger con il suo paradosso mette in discussione sia il livello di incertezza dovuto al decadimento dell'atomo che quello dell'osservatore, ovvero quello del contatore Geiger che

potrebbe non essersi accorto del decadimento, ma anche quello dell'attuatore, ovvero della funzionalità del meccanismo che libera il cianuro, e magari pure il fatto che il gatto potrebbe non essere morto dopo aver assimilato il cianuro, e forse anche lo scrutatore potrebbe essere stato tratto in inganno quando apre la scatola, magari si scopre che il gatto era solo addormentato perché si era annoiato per essere rimasto solo al buio per un'ora. Tutti questi livelli di incertezza fanno parte di un mondo macroscopico sul quale non è possibile applicare le teorie della meccanica quantistica perché non avrebbe alcun senso. Allo stesso modo se la probabilità che ha un elettrone di attraversare una barriera è diversa da zero, quindi esiste questa possibilità in base alle teorie quantistiche, una moto che impatta contro un muro di cemento armato ha zero probabilità di attraversarlo, secondo Schrödinger nemmeno un pezzo di quella moto lo può attraversare anche se gli andasse addosso con una velocità di 200Km/h.



Ma l'equazione di Schrödinger non mette in discussione il principio di conservazione dell'energia meccanica, anzi, egli si basa proprio su questo principio per sviluppare tutta la sua teoria, infatti dalla meccanica classica sappiamo che $E = K + U$, cioè l'energia totale E è la somma dell'energia cinetica K con l'energia potenziale U . Questi tre elementi li elabora in modo da legarli a una variabile temporale "t", a una variabile dimensionale "x" e a una funzione d'onda qualsiasi che lui rappresenta con la lettera greca " ψ ".

Una volta reso esplicito l'operatore hamiltoniano “ \hat{H} ” ha potuto scrivere la seguente relazione:

$$i\hbar \frac{\partial \psi(x,t)}{\partial t} = \frac{-\hbar^2}{2m} \frac{\partial^2 \psi(x,t)}{\partial x^2} + U(x)\psi(x,t) \quad \text{dove} \quad \hbar = \frac{h}{2\pi}$$

in cui “ i ” è l'unità immaginaria, “ h ” è la costante di Planck, “ \hbar ” è la costante di Planck ridotta, “ m ” è la massa e “ U ” è l'energia potenziale. Il primo termine, cioè quello che precede il segno “ $=$ ” rappresenta l'energia totale del sistema, il secondo termine, quello prima del segno “ $+$ ” è l'energia cinetica e infine il terzo termine rappresenta l'energia potenziale. Per capire meglio di che si tratta si può rappresentare la variabile che compongono questa equazione ipotizzando un'onda sinusoidale “ $\psi(x,t)$ ” di valore medio “ E ” che deve attraversare una barriera di lunghezza “ L ” avente un'energia potenziale U_0 , vedi fig.2.3:

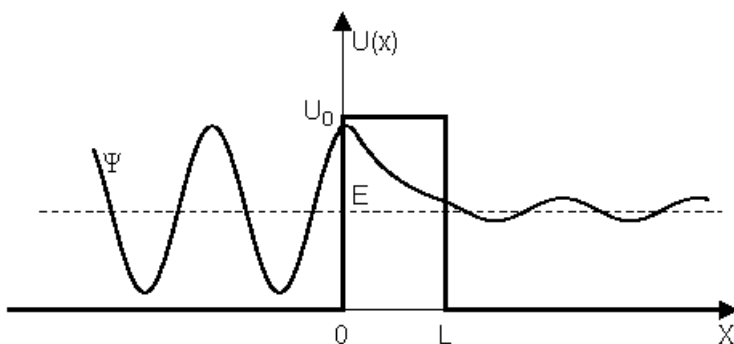


Fig.2.3 Effetto di una barriera di potenziale su una funzione d'onda

La barriera riduce l'ampiezza della funzione ma non ha alcun effetto su altri parametri, come ad esempio la frequenza e il suo livello medio. Ma Schrödinger ovviamente non si limita a formulare le sue teorie attraverso una unica equazione, la risolve per sapere qual'è la probabilità di trovare una particella lungo l'asse X fra due posizioni qualsiasi, X_1 e X_2 ; essa è:

$$P(x,t) = \int_{x_1}^{x_2} |\psi(x,t)|^2 dx \quad \text{normalizzandola con:} \quad \int_{-\infty}^{\infty} |\psi(x,t)|^2 dx = 1$$

Conoscere la funzione d'onda è quindi determinante per Schrödinger anche nel calcolo della probabilità di trovare una particella in uno spazio monodimensionale, ma naturalmente il concetto rimane valido anche in uno spazio a tre dimensioni sostituendo il valore "x" con le coordinate di una ipotetica regione "r(x,y,z)" di uno spazio tridimensionale. La relazione è legata al tempo perché la funzione d'onda " ψ " lo è, tuttavia esiste pure la possibilità di rendere l'equazione indipendente dal tempo (ESIT) che però al momento non ci interessa e quindi la accantoniamo.

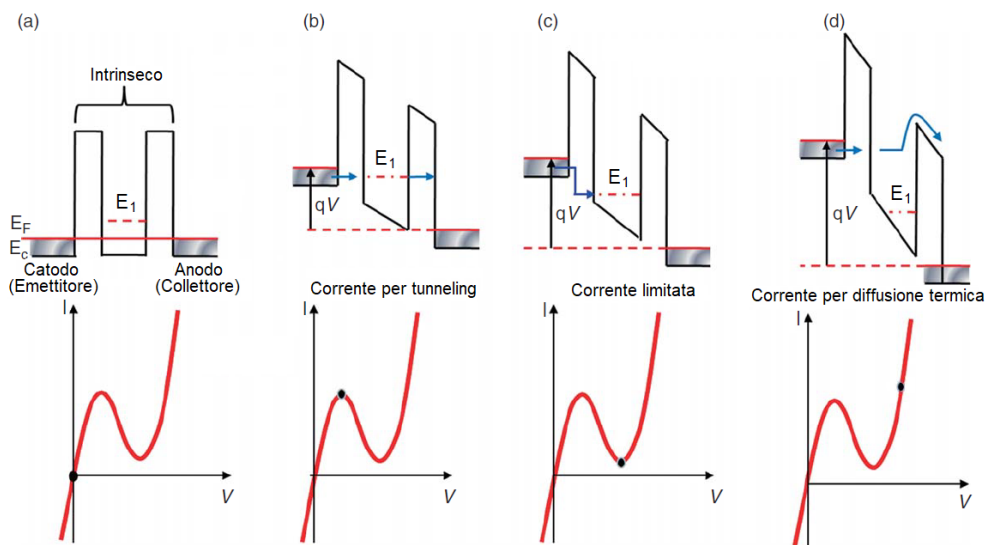


Fig.2.4 Dinamica dei principali stati di polarizzazione di un diodo RTD

Appurato che nel mondo delle particelle una funzione d'onda sarebbe in grado comunque di attraversare una barriera di potenziale è lecito attendersi che tale funzione sia in grado di superare anche più di una barriera in certe condizioni. Ecco, nel diodo RTD accade proprio questo, in condizioni di risonanza la funzione d'onda è in grado di superare entrambe le barriere ripresentandosi dall'altro lato del diodo per effetto tunnel.

Per capire come funziona questa "magia" bisogna analizzare i possibili stati di polarizzazione rappresentati in fig.2.4. Per semplicità ne identificheremo solo quattro: nel primo stato (a) il diodo è privo di polarizzazione e si

riescono a distinguere tre diversi livelli energetici, quello della banda di conduzione E_C , il livello Fermi E_F e un terzo livello E_1 che appartiene solo al pozzo quantico, l'elemento interposto fra le due barriere. Il livello Fermi si trova sopra quello della banda di conduzione esattamente come accade per il diodo tunnel convenzionale, e ciò accade a causa della presenza dei due strati fortemente dopati che avevamo identificato come "sezione catodica" e "sezione anodica", vedi fig.2.2. La larghezza di quei due strati impatta molto sulla traslazione del livello Fermi, al punto da fargli superare quello della banda di conduzione, anche se di poco. Il livello E_1 , invece, riesce a superare anche il livello Fermi a causa della presenza delle due barriere adiacenti aventi un livello energetico decisamente più elevato. In queste condizioni la funzione d'onda " ψ " non è in grado di superare le due barriere, e se ci riuscisse lo farebbe con una forte attenuazione.

Applicando un campo elettrico $E = qV$ ai capi del diodo le bande energetiche si spostano, esattamente come accade per il diodo tunnel fino a quando si raggiunge lo stato (b), detto anche impropriamente stato di "risonanza", una condizione particolare in cui la funzione d'onda " ψ " attraversa le due barriere con la minima attenuazione e lo fa proprio a causa dell'effetto tunnel; in realtà l'appellativo di "risonanza" ha ben poco a che fare con quella che noi tutti conosciamo, perché il diodo in questo caso non sta oscillando, a meno che non si verificano esternamente altre condizioni che per il momento non ci interessano. Superato questo limite si manifesta la famosa "resistenza negativa", ovvero all'aumentare della tensione la corrente diminuisce invece che aumentare; nel diodo RTD questo passaggio avviene in un modo molto brusco a differenza del diodo tunnel e ciò accade proprio per la presenza di due barriere che, quando l'effetto tunnel viene a mancare fanno presto a richiudere il passaggio di corrente fino a riportare il diodo nello stato (c), ovvero nel punto in cui avviene la massima limitazione della corrente e quindi nuovamente la massima attenuazione per la funzione d'onda.

Dopodiché la corrente riprende a salire e questo lo fa non più per l'effetto tunnel ma per diffusione termica; da questo momento la resistenza che offre il diodo torna ad essere positiva e diminuisce dinamicamente all'aumentare del campo elettrico, esattamente come accade in un diodo tunnel,

ma anche come accade in un diodo normale, una volta superata la barriera di potenziale non esistono più altri ostacoli. Nello stato (d), pertanto, gli elettroni riescono a superare le due barriere con estrema facilità perché il campo elettrico esterno glielo consente aumentando sensibilmente la loro energia cinetica. Anche nel diodo RTD tutta questa evoluzione di stati avviene applicando campi elettrici molto contenuti, parliamo di differenze di potenziale elettrico inferiori al volt, fra lo stato (a) e lo stato (d).

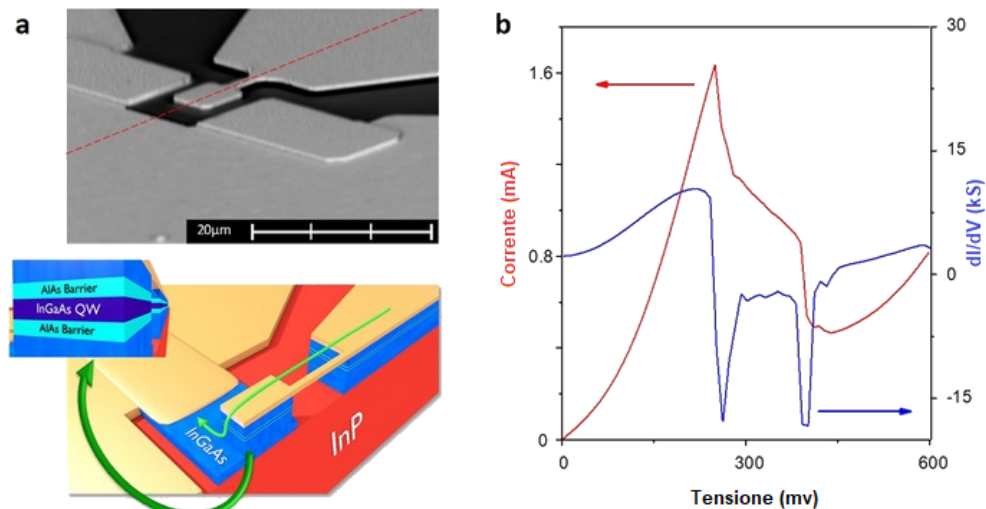
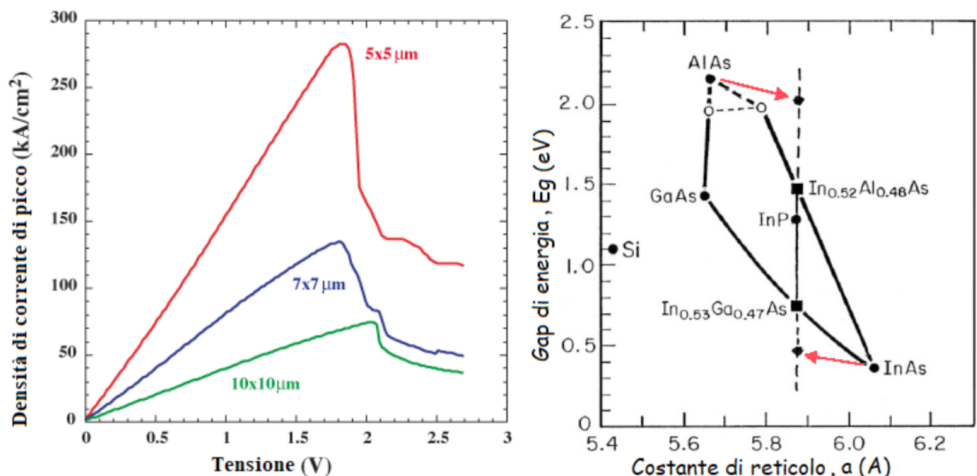


Fig.2.5 A sinistra: struttura fisica tipica di un diodo RTD; a destra: caratteristica tensione corrente e la derivata dI/dV

Per capire quanto può essere piccolo un diodo RTD montato su un substrato al fosforo di indio è sufficiente osservare la fig.2.5 a destra; a sinistra invece è riportata la sua caratteristica tensione/corrente che mette in evidenza uno strano comportamento nella regione dove il diodo presenta la sua resistenza negativa. Di solito la resistenza negativa di un diodo tunnel comune è abbastanza lineare, quella di un diodo RTD non lo è affatto ma questo non lo rende meno interessante, al contrario ne esalta le sue prerogative di essere un componente bistabile, quindi essenzialmente una “memoria” statica allo stato solido; questo aspetto è molto importante ma verrà trattato di seguito. Invece il parametro più importante di un diodo RTD è quello che viene definito PVCR, cioè il Peak to Valley Current Ratio, tradotto

sarebbe il rapporto fra la corrente di picco e quella di valle. Questo rapporto varia in funzione della temperatura, ma soprattutto varia in funzione dei materiali con cui viene realizzato il diodo. In fig.2.2 avevamo già visto una struttura realizzata con materiali del gruppo III e V della tavola degli elementi, e cioè la struttura classica AlGaAs/GaAs/AlGaAs che appartiene alla famiglia di RTD denominata “GaAs”.



Materiale	InGaAs/ AlAs	InAs/ AlSb	Si/SiGe	GaAs/ AlAs	SiGe Esaki
J_p (kA/cm ²)	460	370	282	250	30
PVCR	4	3.2	1.2/2.43	1.8	4.8
$\Delta I \Delta V$ (mW)	5.4	9.4	4.30	4.0	2.5
R_D (Ω)	1.5	14.0	12.5	31.8	16
Area (μm^2)	16	1	25	5	200
Gruppo	MIT	CalTech / MIT	Cambridge	Stanford	Max Planck Stuttgart

Fig.2.6 In alto a sinistra parametri di un RTD Si/SiGe con superfici di sezione differenti; in basso tabella comparativa con diodi di altri materiali

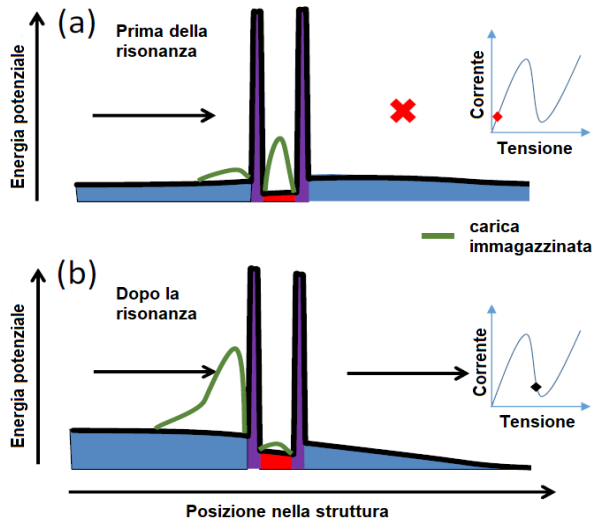
Sempre per quanto riguarda i materiali del medesimo gruppo abbiamo anche la famiglia “InP”, con la quale invece abbiamo una struttura più complessa del tipo InGaAs/AlAs/InAs, e con essa però si riesce ad ottenere una

densità di corrente di picco J_p di quasi 500kA/cm^2 , un valore enorme rispetto a quello di un RTD GaAs, dovuto principalmente al salto energetico che si viene a creare nel reticolo cristallino rappresentato in fig.2.6, in alto a destra. Tuttavia, la struttura più interessante rimane quella evidenziata in rosso nella tabella riportata nella stessa figura, cioè quella dei diodi realizzati con materiali del gruppo IV della tavola degli elementi, in particolare quelli al Silicio-Germanio la cui densità di corrente massima è rappresentata nel grafico in alto a sinistra. Il motivo principale di tutto questo interesse è legato alla disponibilità in natura di questi due materiali semiconduttori, e alla loro compatibilità con i substrati e con i processi industriali già esistenti e ottimizzati nel corso degli anni; in seguito riprenderemo questo discorso perché trattasi di un argomento in avanzata fase di sviluppo.

In fig.2.6 i valori di PVCR del diodo Si/SiGe sono due: il primo si riferisce al valore originale a temperatura ambiente, che era così basso a causa degli scarsi offset ottenuti rispetto alla banda di conduzione nelle prove effettuate sui primi campioni; il secondo è invece il risultato di numerosi esperimenti e di affinamenti del processo produttivo condotti nei Laboratori Cavendish di Cambridge e riportati da D. J. Paul et al su IEEE Electron Device Letters, Vol.22, No.4, Aprile 2001. Per fare un confronto i parametri di un comune diodo tunnel SiGe sono invece riportati nella colonna posta all'estrema destra della stessa tabella.

Oltre ai materiali già raccolti ve ne sono degli altri che sono stati valutati nel corso degli anni, ad esempio ci sono quelli appartenenti ai gruppi II-VI con strutture tipo HgCdTe/HgTe che presentano un PVCR di 1,4 a 300°K ; oppure ci sono quelli con strutture cristalline miscelate, ad esempio quelli MnTe/InSb con un PVCR di 1,7 a 77°K e anche quelli $\text{CaF}_2/\text{CoSi}_2$ con un PVCR di 2 e pure quelli AlAs/ErAs che sarebbero compatibili con substrati GaAs. Infine ci sono le strutture amorfe, tipo SiO_2/Si , $\text{Si}_3\text{N}_4/\text{Si}$ e il SiC/Si con un PVCR che addirittura arriva a 9,4. Ricordo che in tutte queste definizioni il primo materiale rappresenta quello con cui sono realizzate le barriere, il secondo invece è quello con cui si realizza il pozzo quantico. Nelle strutture complesse in cui appaiono tre elementi le barriere sono realizzate col secondo elemento mentre il pozzo è realizzato col terzo, quest'ultimo però di solito si affaccia alle barriere con due strati realizzati col primo elemento,

quello con cui si costruiscono anche le parti esterne e che vanno a contatto con le due barriere, questo per motivi di compatibilità nella stratificazione.



Ricordo che il livello di Fermi rappresenta la probabilità che ha un elettrone di occupare uno stato con livello energetico E , la distribuzione Fermi Dirac infatti è la seguente:

$$F(E) = [1 + e^{(E-E_F)/k_B T}]^{-1} = \frac{1}{1 + e^{(E-E_F)/k_B T}} \quad [1]$$

dove e è il numero di Eulero, k_B è la costante di Boltzmann, T è la temperatura assoluta, mentre E_F è il livello di Fermi. Nel caso in cui $E \gg E_F$ il livello di Fermi diventa trascurabile ai fini del calcolo. La densità degli stati in un semiconduttore è uguale alla densità per unità di volume ed energia del numero di soluzioni dell'equazione di Schrödinger. Assumeremo che il semiconduttore si possa modellare come se fosse un infinito pozzo quantico in cui gli elettroni con massa effettiva, m^* , sono liberi di muoversi; in questo caso la densità degli stati per unità di volume e per unità di energia è:

$$g(E) = \frac{8\pi\sqrt{2}}{h^3} m^{*3/2} \sqrt{E} \quad \text{per } E > 0 \quad [2]$$

l'energia minima degli elettroni si trova nella parte bassa della banda di conduzione, E_c , pertanto la densità degli stati nella banda di conduzione è:

$$g_c(E) = \frac{8\pi\sqrt{2}}{h^3} m^{*3/2} \sqrt{E - E_c} \quad \text{per } E \geq E_c \quad [3]$$

mentre *per* $E < E_c \rightarrow g(E) = 0$ e lo stesso vale per la parte bassa del pozzo quantico, per valori di E negativi la densità degli stati è 0.

Facciamo un esempio: calcolare il numero degli stati per unità di energia in un pezzo di silicio da $100 \times 100 \times 10$ nm ($m^* = 1,08 \times m_0$), 100meV sopra la banda di conduzione; scrivere la soluzione in eV^{-1} .

In base alla [2] calcoliamo la densità degli stati:

$$g(E) = \frac{8\pi\sqrt{2}(1,08 \times 9,1 \times 10^{-31})^{3/2}}{(6,626 \times 10^{-34})^3} \sqrt{0,1 \times 1,6 \times 10^{-19}} = 1,51 \times 10^{56} m^{-3} J^{-1}$$

cosicché il numero totale degli stati per unità di energia sarà uguale a:

$$g(E)V = 1,51 \times 10^{56} \times 10^{-22} J^{-1} = 2,41 \times 10^5 eV^{-1}$$

Il numero degli elettroni per unità di energia è il prodotto fra la densità degli stati e l'equazione statistica Fermi Dirac, e cioè:

$$\frac{dn}{dE} = g(E)F(E) \quad [4]$$

a questo punto volendo integrare tale funzione entro i limiti energetici (E) ed $(E + \Delta E)$ otteniamo:

$$\int_E^{E+\Delta E} g(E)F(E)d(E) = n(E + \Delta E) - n(E) \quad [5]$$

dove "n" rappresenta il numero di elettroni che statisticamente vengono trasferiti a causa del salto energetico, da (E) a $(E + \Delta E)$.

Per modellare la caratteristica V-I dell'RTD si ipotizza che uno spettro di autostati elettronici vadano in contatto fra loro, e le onde di elettroni così prodotte sono quindi soggette a diffusione elastica e anelastica dei vari potenziali di diffusione. Inoltre, si assume che gli autostati elettronici in tali eventi siano delle onde piane e le funzioni d'onda, diretta e riflessa, siano le soluzioni dell'equazione di Schrödinger per un particolare profilo di un RTD, includendo i termini del potenziale di diffusione e le masse efficaci. Una volta trovate le funzioni d'onda, è possibile calcolare la densità di corrente del dispositivo:

$$J_{tot} = -q_e \hbar \sum_k W(k) \operatorname{Im} \left[\psi_k^*(x) \frac{1}{m^*(x)} \frac{\partial \psi_k(x)}{\partial x} \right] \quad [6]$$

dove è stata eseguita una sovrapposizione su stati quantistici opportunamente normalizzati. Allo stesso modo, la densità di corrente può essere valutata utilizzando la formula Tsu-Esaki (vedi R. Tsu and L. Esaki, "Tunneling in a finite superlattice," Appl. Phys. Lett., vol. 22, pp. 562–564, giugno 1973), la densità di corrente totale si ottiene sommando la densità di corrente di ogni stato, soltanto di quelli occupati, e moltiplicandola per la loro probabilità di trasmissione:

$$J_{tot} = \frac{q_e m^* k_B T}{2\pi^2 \hbar^3} \int_{E_c}^{\infty} T(E_x) \log \left[\frac{1 + \exp\left(\frac{E_F - E_x}{kT}\right)}{1 + \exp\left(\frac{E_F - E_x - qV_A}{kT}\right)} \right] dE_x \quad [7]$$

dove $T(E_x)$ può essere calcolato dalle funzioni d'onda disponibili nella soluzione dell'equazione di Schrödinger; anche la risposta impulsiva per gli elettroni incidenti è stata incorporata, una funzione che si integra efficacemente sui vettori d'onda trasversali e che si ottiene integrando $k_{||}$. Gli altri simboli nella [7] sono tutti quelli convenzionali che abbiamo già incontrato nelle precedenti trattazioni. La densità di corrente ricavata con la [6] e la [7] rappresenta il modello fisico di questa struttura che però viene utilizzato raramente nella simulazione del diodo RTD, soprattutto perché si basa su un

calcolo statistico. Spesso è molto più comodo utilizzare un modello virtuale analitico con cui trattare certi argomenti perché con quest'ultimo diventa possibile modificare rapidamente i parametri che ci interessano per adattare il diodo alle nostre esigenze. E quindi, come al solito, si preferisce disegnare un circuito equivalente che simula il comportamento del semiconduttore e lo si analizza come modello elettrico, esattamente come è già stato fatto in precedenza per il diodo tunnel.

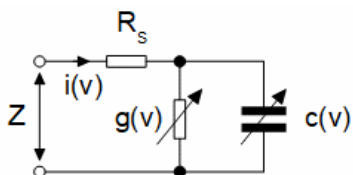


Fig.2.7

Circuito equivalente di un RTD

La fig.2.7 riportata qui accanto rappresenta il circuito equivalente di un RTD che è molto simile a quello di un diodo tunnel; tuttavia alcune differenze rimangono: ad esempio la resistenza R_s non è più considerata come una variabile a causa dell'effetto pelle perché non corrisponde più alle resistenze dei conduttori e dei contatti elettrici del contenitore, visto che l'RTD

non lo prevede. R_s è intesa come la resistività del semiconduttore intrinseco negli strati privi di doping, che è indipendente dalla frequenza.

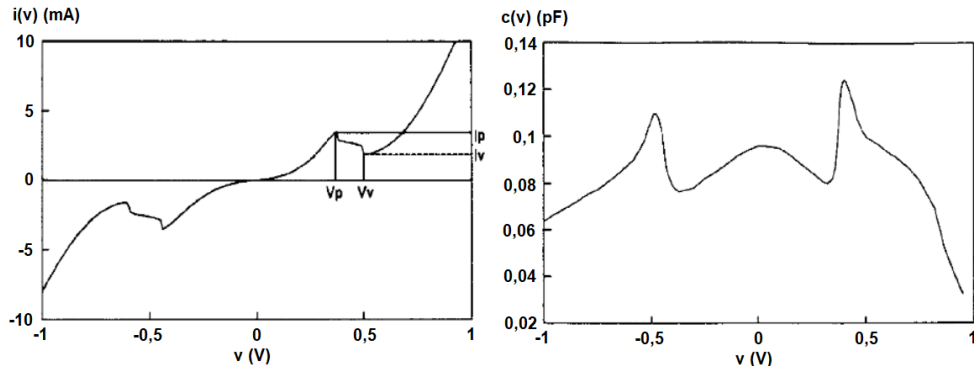


Fig.2.8 Caratteristiche $i(v)$ (a sinistra) e $c(v)$ (a destra) di un diodo RTD

Allo stesso modo notiamo anche l'assenza di una induttanza serie che è trascurabile rispetto a quella di un diodo tunnel, sempre a causa della totale mancanza del contenitore e del relativo bonding per i contatti. Rimangono tre variabili dipendenti su cui concentrare l'attenzione: la conduttanza $g(v)$, la capacità posta in parallelo $c(v)$ e la corrente $i(v)$; tutte e tre queste

variabili dipendono dalla tensione applicata al diodo. In fig.2.8 ne vengono rappresentate solo due, entrambe misurate sperimentalmente su un diodo campione, ma alla fine ricaveremo anche la terza, ovvero la conduttanza "g" che come vedremo in alcuni tratti diventa negativa. Invece la cosa interessante da osservare è la caratteristica $i(v)$ con la polarità inversa del diodo RTD (vedi il terzo quadrante della figura di sinistra), che rispecchia molto quella della polarità diretta, a differenza di ciò che accade per un diodo tunnel. In effetti osservando la struttura di un RTD riportata in fig.2.2 esiste una certa simmetria fra la sezione anodica (collettore) e quella catodica (emettitore). Questa caratteristica semplifica notevolmente la formulazione di un modello analitico, soprattutto volendo considerare simmetrica anche la caratteristica $c(v)$ che invece qualcosa di diverso ce l'ha, se volessimo essere precisi. Pertanto, per descrivere analiticamente un diodo RTD è sufficiente utilizzare tre equazioni diverse ciascuna delle quali individua il profilo di una variabile dipendente dalla tensione "v", ovvero la corrente "i" la conduttanza "g" e infine la capacità "c". L'insieme di queste tre equazioni descrivono analiticamente un diodo RTD, esse possono essere usate anche singolarmente per simularne il comportamento. A tale proposito saranno mostrati due esempi pratici: il primo è un modello "standard" realizzato mediante equazioni esponenziali e gaussiane, il secondo è un modello molto più preciso che però riguarda solo la la caratteristica $i(v)$.

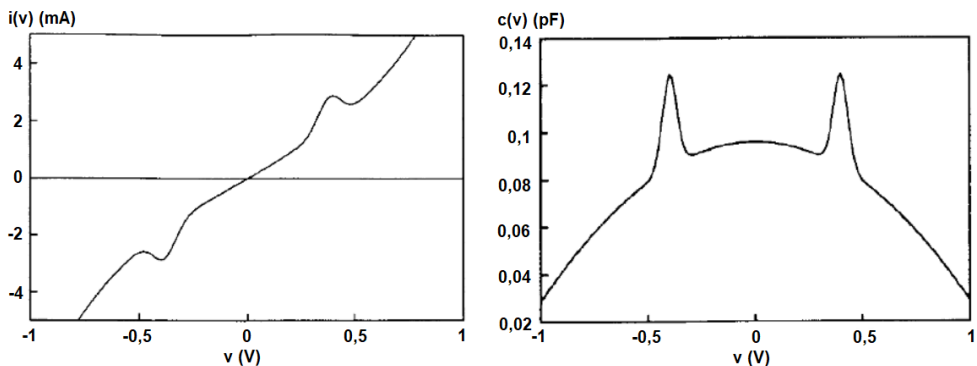


Fig.2.9 Caratteristiche $i(v)$ (a sinistra) e $c(v)$ (a destra) di un modello analitico del diodo RTD di fig.2.8

Non fatevi ingannare dall'aspetto delle curve di fig.2.9 perché la scala del grafico a sinistra è leggermente diversa da quella originale, relativa alla cor-

rente misurata direttamente sul diodo, mentre quella di destra rappresenta semplicemente la caratteristica capacitiva con un grado di approssimazione appena più elevato, ma in realtà numericamente non è molto distante dalla caratteristica originale fra 0V e +0,65V, cioè i due valori di tensione estremi in cui il diodo dovrà funzionare. D'altra parte non ci interessa più di tanto questa asimmetria tipica che presenta un diodo RTD l'importante è trovare un algoritmo che consenta di ottenere un modello con dei valori più vicini possibile a quelli misurabili sperimentalmente su uno o più campioni all'interno delle condizioni tipiche di funzionamento. Vediamo ora come sono fatti questi algoritmi, prendiamo ad esempio in considerazione la caratteristica $i(v)$ del diodo: questo modello viene rappresentato da una curva asimmetrica che consiste in due curve esponenziali e due curve Gaussiane che interagiscono fra loro.

$$i(v) = k_3 (e^{k_0 v} - e^{-k_0 v}) + k_1 \left[e^{-\frac{(v-x_1)^2}{k_2}} - e^{-\frac{(v+x_1)^2}{k_2}} \right] \quad [8]$$

dove: x_1 controlla la posizione del picco di corrente lungo la curva;
 k_0 controlla la pendenza della curva;
 k_1 controlla il livello del picco di corrente I_p ;
 k_2 controlla la posizione della valle della corrente lungo la curva;
 k_3 è semplicemente una costante di trasferimento unitario.

Naturalmente come avevamo già visto nel diodo tunnel I_p e V_p sono i picchi di corrente e di tensione mentre I_v e V_v sono rispettivamente le loro valli; per ottenere la curva a sinistra di fig.2.9 i valori impostati sono:

$$x_1 = 0,38315, k_0 = 2,11303, k_1 = 0,00102, k_2 = 0,0056, k_3 = 0,001;$$

in seguito vedremo anche il significato fisico di queste variabili.

Gli altri parametri legati alla caratteristica $i(v)$ di un diodo RTD sono: la conduttanza della regione positiva G_p , la conduttanza differenziale della regione negativa G_D e il rapporto fra la corrente di picco e quella di valle che avevamo già definito col termine PVCR, cioè Peak to Valley Current Ratio.

Volendo mettere queste tre grandezze in termini analitici abbiamo:

$$G_P = \frac{I_P}{V_P} ; G_D = \frac{I_P - I_V}{V_V - V_P} ; PVCR = \frac{I_P}{I_V}$$

Pertanto anche fornendo l'indicazione di V_P , insieme a quelle di G_P , G_D e $PVCR$, tutti punti estremi della caratteristica $i(v)$ di un diodo RTD possono essere individuati; la conduttanza variabile $g(v)$ di un RTD può quindi essere calcolata derivando la caratteristica $i(v)$, e cioè:

$$g(v) = k_0 k_3 (e^{k_0 v} + e^{-k_0 v}) - \frac{2k_1}{k_2} \left[(v - x_1) e^{-\frac{(v-x_1)^2}{k_2}} - (v + x_1) e^{-\frac{(v+x_1)^2}{k_2}} \right] \quad [9]$$

e anche in questo caso la curva della conduttanza variabile del modello sarà simmetrica, come quella rappresentata in nero nella fig.2.10.

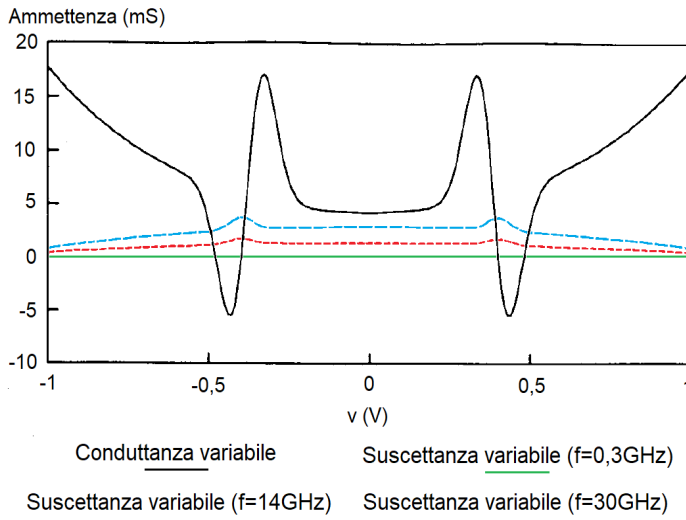


Fig.2.10 Conduttanza e suscettanze variabili di un modello del diodo RTD

Quando la conduttanza variabile $g(v)$ diventa negativa siamo in presenza di una potenziale instabilità e in alcune condizioni il diodo potrebbe oscillare perché siamo in presenza di una resistenza differenziale negativa. A tale

proposito suggerisco il lettore di rivedere le condizioni di stabilità contemplate per un diodo tunnel a partire dal paragrafo 1.1, esse infatti rispecchiano quelle di un diodo RTD. Seguendo quelle indicazioni è possibile far funzionare un RTD come oscillatore, come amplificatore oppure come dispositivo bistabile, oggi quest'ultima possibilità è forse quella più utilizzata.

Vediamo ora il significato fisico delle variabili che abbiamo utilizzato per formulare la [8] e la [9].

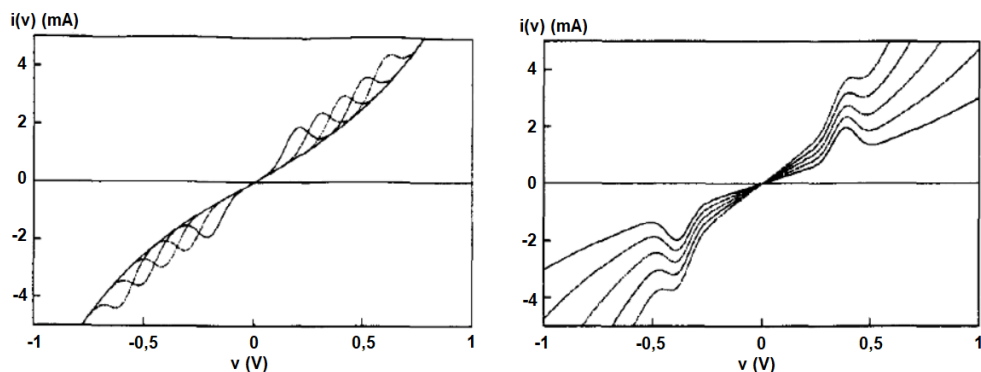


Fig.2.11 Variazioni della curva $i(v)$ al variare dei parametri sopra indicati: a sinistra varia x_1 , a destra varia k_0 ; gli altri parametri rimangono fissi

In fig.2.11 sono rappresentati gli effetti che si possono ottenere sulla curva $i(v)$ al variare dei parametri x_1 e k_0 : in particolare a sinistra troveremo la variazioni di x_1 con i seguenti valori crescenti:

$$x_1 = 0,2; x_1 = 0,3; x_1 = 0,4; x_1 = 0,5; x_1 = 0,6;$$

sulla destra invece troveremo le variazioni che subisce la curva con valori crescenti di k_0 , e cioè:

$$k_0 = 1,2; k_0 = 1,6; k_0 = 2,0; k_0 = 2,4; k_0 = 2,8.$$

In pratica incrementando x_1 varia la posizione dei picchi di tensione e di corrente V_p e I_p lungo la curva, la resistenza differenziale negativa viene semplicemente traslata; G_p , G_D e PVCR invece decrementano. Il controllo

della pendenza invece dipende da k_0 , quando incrementa tale parametro incrementa anche G_p , invece tutti gli altri parametri decrementano, intendo V_p , G_D e PVCR.

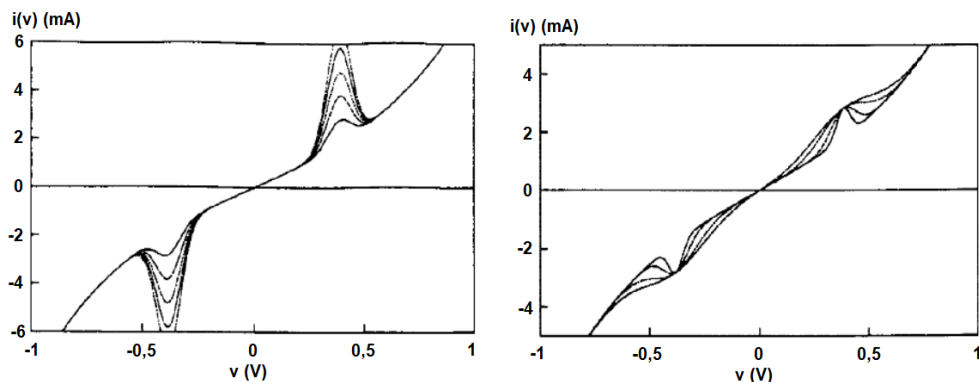


Fig.2.12 Variazioni della curva $i(v)$ al variare dei parametri sopra indicati: a sinistra varia k_1 , a destra varia k_2 ; gli altri parametri rimangono fissi

In fig.2.12 sono rappresentati invece gli effetti che si possono ottenere sulla curva $i(v)$ al variare dei parametri k_1 e k_2 : in particolare a sinistra troveremo la variazioni di k_1 con i seguenti valori crescenti:

$$k_1 = 0,001; k_1 = 0,002; k_1 = 0,003; k_1 = 0,004; k_1 = 0,005;$$

sulla destra invece troveremo le variazioni che subisce la curva con valori crescenti di k_2 , e cioè:

$$k_2 = 0,002; k_2 = 0,006; k_2 = 0,02; k_2 = 0,04.$$

In pratica k_1 controlla il valore di picco della corrente, cioè I_p ; incrementando k_1 V_p , V_V e I_V rimangono invariati, mentre incrementano G_p , G_D e PVCR. Il parametro k_2 invece controlla la posizione e la profondità dei punti di valle (V_V e I_V); incrementando k_2 i punti di picco di corrente (V_p e I_p) rimangono nella stessa posizione ma i punti di valle della corrente e della tensione si spostano lungo la curva riducendone la profondità; G_p rimane invariato mentre G_D e PVCR decrementano.

Per quanto riguarda invece la curva rappresentata a destra di fig.2.8, quella relativa al modello della caratteristica capacitiva di un diodo RTD, essa è stata ottenuta sovrapponendo la curva di una parabola con due curve Gaussiane con la seguente relazione:

$$c(v) = C_0(1 - \Phi v^2) + \Delta C \left[e^{-\frac{(v-V_0^2)}{\Delta V^2} 4 \ln(2)} + e^{-\frac{(v+V_0^2)}{\Delta V^2} 4 \ln(2)} \right] \quad [10]$$

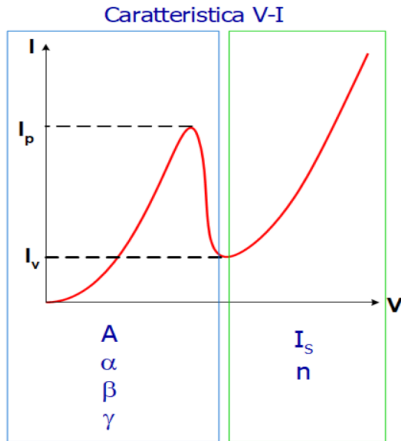
dove: $\Delta C = C_1 - C_0(1 - \Phi V_0^2)$

- e: C_0 è la capacità con 0V di tensione di polarizzazione;
 C_1 è il valore di picco della capacità;
 V_0 è la tensione di polarizzazione in cui C_1 è massimo;
 Φ è il potenziale di tensione accumulato ($\approx 0,7V$);
 ΔV è la differenza di tensione fra i due valori dove le curve Gaussiane intersecano la parabola.

Per tali parametri nel modello della caratteristica capacitiva rappresentata a destra in fig.2.8 sono stati utilizzati i seguenti valori:

$\Phi = 0,7V$; $C_0 = 0,0959pF$; $C_1 = 0,12435pF$; $V_0 = 0,36V$; $\Delta V = 0,08V$.

Quello che il lettore ha avuto modo di leggere sopra è il “metodo classico” col quale definire il modello di un semiconduttore, ossia: osservare la curva del comportamento fisico di un dispositivo e interpretarla analiticamente con una funzione che sia in grado di simulare in grandi linee tale comportamento adattando e poi affinando i parametri fino ad ottenere il risultato migliore. Oggi questo lavoro può essere svolto da un calcolatore sul quale viene installato un apposito programma “modellatore” col quale è possibile ottenere anche risultati migliori. Farò un esempio utilizzando un programma che si chiama IC-CAP della Keysight che però è stato progettato per modellare semiconduttori generici di ogni tipo, ma non degli RTD per i quali è indispensabile fornire anche un “aiuto umano” onde ottenere dei risultati accettabili. L'esempio in questione è stato condiviso da Chien M. Ta (SMA), Fujiang Lin (IME) e Subhash R. Chander (IME) nella loro presentazione del giugno 2004 al Simposio Internazionale dell'Elettronica.



Senza entrare troppo nel dettaglio al programma è stata fornita l'indicazione di estrarre le informazioni di tensione di corrente del diodo attraverso degli strumenti di misura prodotti dalla Keysight e di formulare con essi una possibile funzione analitica compatibile col dispositivo sottoposto al test, in questo caso si trattava di un diodo RTD. Il programma ha ricreato la curva $i(v)$ del diodo utilizzando n.6 variabili, quattro per descrivere la parte dove è presente la resistenza differenziale negativa e due per il tratto di curva finale in cui la corrente riprende a salire con una funzione esponenziale. La funzione che ne descrive il comportamento è la seguente:

La funzione che ne descrive il comportamento è la seguente:

$$I_D(V_D) = I_{NDR}(V_D) + I_{FWD}(V_D) \quad [11]$$

dove: $I_{NDR}(V_D) = A(\gamma V_D)^{1+\alpha} \exp[(-\gamma V_D)^\beta]$ [12]

e: $I_{FWD}(V_D) = I_S \left[\exp\left(\frac{V_D}{n V_T}\right) - 1 \right]$ [13]

Qui accanto troverete i valori ottimizzati dei parametri coi quali è stata formulata la [11]. V_D rappresenta la variabile della tensione, V_T invece è una costante. L'ottimizzazione è avvenuta in seguito ad alcuni aggiustamenti "manuali" per ridurre al minimo l'errore RMS rispetto alla curva naturale del semiconduttore che in questo caso si attesta a 3,813%, un risultato decisamente più accurato rispetto al precedente modello.

Parametro	Valori ottimizzati
A	652,7x10 ⁻⁹
α	2,719
β	16,92
γ	2,593
I _s	2,2248x10 ⁻¹⁹
n	0,7882

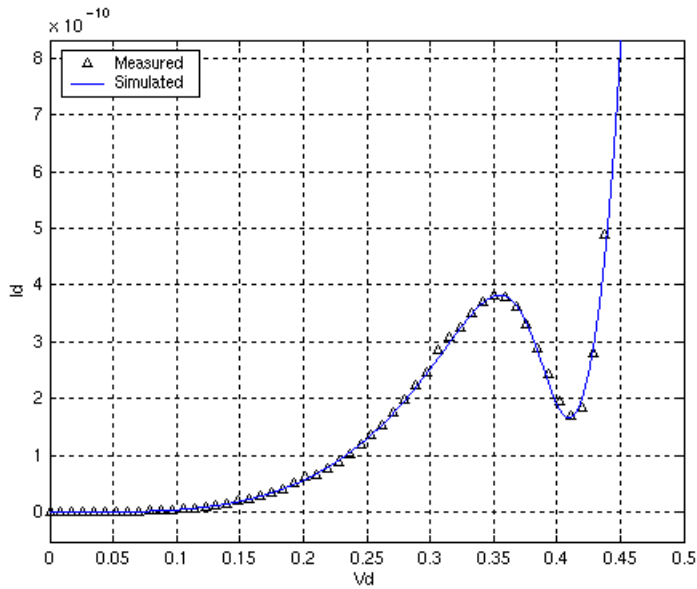


Fig.2.13 Modello parziale di un RTD ottenuto con IC-CAP di Keysight

Si osservi bene la [12] e la [13]: anche in questo caso siamo in presenza di una funzione Gaussiana e una esponenziale, tuttavia stavolta si sta valutando solo la curva nel primo quadrante, con valori positivi della tensione e della corrente. Qualora volessimo estendere simmetricamente la curva anche sul terzo quadrante bisognerà fare qualche modifica alle precedenti equazioni e aggiungerne una, ecco come si fa. La [11] diventa:

$$I_D(V_D) = I_{NDR}(V_D) + I_{FWD}(V_D) + I_{REV}(V_D) \quad [14]$$

dove: $I_{NDR}(V_D) = A(\gamma V_D) |\gamma V_D|^\alpha \exp\left[(-|\gamma V_D|)^\beta\right]$ [15]

con: $I_{FWD}(V_D) = I_S \left[\exp\left(\frac{V_D}{nV_T}\right) - 1 \right]$ se: $V_D \geq 0$ [16]

e: $I_{REV}(V_D) = -I_S \left[\exp\left(-\frac{V_D}{nV_T}\right) - 1 \right]$ se: $V_D < 0$ [17]

La [16] e la [17] si alternano in base al segno di V_D ; il risultato di queste modifiche determinano la curva riportata in fig.2.14.

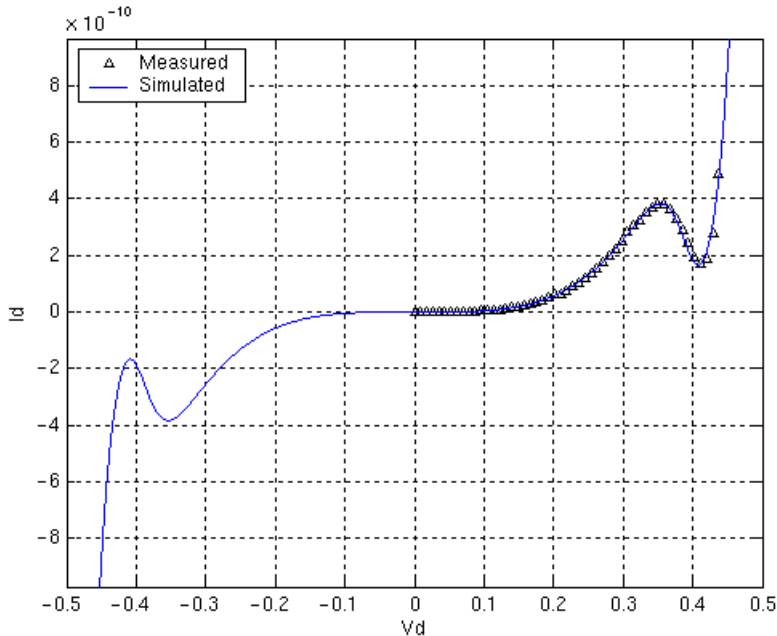


Fig.2.14 Modello completo di un RTD ottenuto aggiungendo un artificio per rendere simmetrica la curva $i(v)$, estendendola anche sul 3° quadrante

Naturalmente la curva così ottenuta peggiora l'errore RMS totale perché è noto che la caratteristica $i(v)$ di un RTD non è poi così simmetrica come vorremmo che fosse, tuttavia questo modello resta valido nel caso di una simulazione, e la conduttanza $g(v)$ rimane la derivata di questa funzione.

Il diodo RTD è diventato ben presto il più valido sostituto del diodo tunnel, oggi è uno dei semiconduttori più studiati in tutto il globo per via delle sue prestazioni eccezionali nel settore delle onde millimetriche e sub-millimetriche. Verso la fine del secolo scorso parlare di semiconduttori in grado di raggiungere frequenze dell'ordine dei 2-300GHz sembrava quasi un'utopia ma in realtà tali dispositivi erano già stati abbondantemente studiati a livello teorico, soltanto la tecnologia che ne avrebbe dovuto misurare le proprietà elettriche non era stata ancora abbastanza affinata per farlo. Il

motivo per cui il diodo tunnel fu abbandonato è proprio grazie all'avvento imponente del diodo RTD che ne ha assorbito tutte le qualità aggiungendo la caratteristica dell'elevato limite di frequenza con cui esso sarebbe stato operativo. Tuttavia l'intera parte teorica dello studio del diodo tunnel può essere trasferita in modo quasi immutato verso il diodo RTD, incluse le sue applicazioni. Per questo motivo esorto il lettore a rivedere quanto già scritto nel paragrafo precedente in merito al diodo di Esaki, trattasi di informazioni tanto necessarie quanto propedeutiche per avere la padronanza scientifica completa anche del diodo RTD. L'unica differenza rilevante fra i due diodi è la loro caratteristica con polarità invertita: nel diodo tunnel la barriera di potenziale viene immediatamente superata a causa dell'eccesso di doping offrendo una resistenza molto bassa al passaggio della corrente; nell'RTD invece gli elettroni in transito devono comunque superare le due barriere e il pozzo quantico, tre sottili strati di semiconduttore completamente prive di doping, quindi la caratteristica V-I inversa ricalca in grandi linee quella che lo stesso diodo offre con la polarità diretta ma con prestazioni inferiori e pertanto non degne di nota in questo contesto.

I concetti di stabilità in funzione della retta di carico di un diodo tunnel, già espressi e sviscerati nella parte 1 e raccolti in fig.1.10, possono essere applicati quindi anche a un diodo RTD; esso infatti può funzionare come dispositivo stabile (come amplificatore), instabile (come oscillatore), oppure come dispositivo bistabile (come cella di memoria). Inizierò questa breve panoramica descrivendo un progetto relativamente recente di un amplificatore bidirezionale in banda C realizzato con un combinatore ibrido a costanti concentrate di tipo "branch-line" collegato a due amplificatori a riflessione che utilizzano altrettanti diodi RTD realizzati per epitassia con tecniche simili a quelle con cui si realizzano i circuiti monolitici. Per questo progetto bisogna ringraziare i due Ingegneri che lo hanno realizzato presso l'Advanced Institute of Science and Technology (KAIST), Daejeon, Republic of Korea, cioè Jongwon Lee e Kyoungsoon Yang. Tale realizzazione non è proprio così "irresistibile" dal punto di vista della complessità, tuttavia sarebbe da prendere come esempio per la diversità di argomenti che esso è in grado di infondere verso uno studente universitario che abbia intenzione di affacciarsi oggi nel mondo delle microonde. La vastità di informazioni tecniche e tecnologiche necessarie per studiare a fondo questo amplificatore è una ot-

tima base di partenza per qualsiasi corso universitario di elettronica applicata alle microonde, un validissimo modo per incentivare gli studenti a studiare tali argomenti con maggiore determinazione.

Il progetto in questione è stato pubblicato nell'IEEE Microwave and Wireless Components Letters (Volume: 27, Issue: 1, gennaio 2017), trattasi di un minuscolo amplificatore a basso consumo (circa 0,42mW) con un guadagno di lineare di 8,6dB a 5,8GHz fino a -25,4dBm e con 11dB di perdita di ritorno. Il circuito elettrico e la realizzazione pratica è riportata in fig.2.15.

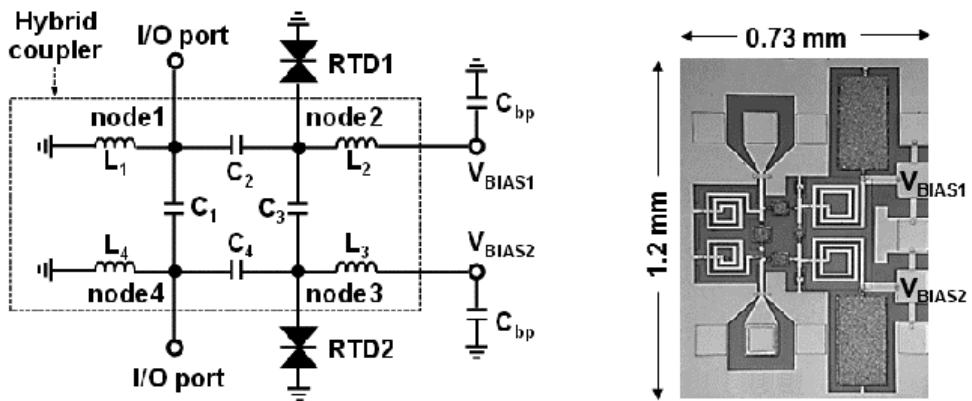


Fig.2.15 Amplificatore bidirezionale in banda C con diodi RTD

In questo caso per trasferire correttamente la polarizzazione verso i due RTD per il combinator è stata scelta la configurazione ad anello capacitivo, in questo caso $L_1 = L_4$, $L_2 = L_3$ e $C_2 = C_4$. Per motivi di compatibilità con il combinator a costanti distribuite di solito la struttura del combinator a costanti concentrate di tipo "branch-line" si realizza facendo il contrario, cioè mettendo le capacità al posto delle induttanze e viceversa (configurazione ad anello induttivo), tuttavia quest'ultima soluzione sarebbe inadatta per polarizzare questo circuito. Sulla foto del prototipo si distinguono chiaramente tutti gli elementi attivi e passivi, inclusi i due grossi condensatori di bypass C_{bp} e le due serie di contatti sui quali collegare le complanari per le misure del caso mediante VNA. Ebbene le guide d'onda complanari (o "complanari" come talvolta qualcuno ancora oggi le nomina erroneamente) rappresentano il metodo più diffuso per collegare uno strumento di misura, di

solito un network analyzer vettoriale o VNA, ad un microcircuito oppure ad un circuito RF miniaturizzato mantenendo una impedenza controllata fino ai contatti delle apposite piazzole già predisposte all'occorrenza su di esso che di solito sono a gruppi di tre, un polo caldo centrale e due poli di ritorno laterali che insieme realizzano una linea "quasi TEM" sbilanciata. Il collegamento fra una linea complanare e una linea coassiale o a una microstriscia avviene attraverso delle apposite transizioni, sempre ad impedenza controllata. In fig.2.16 c'è un esempio di un piccolo circuito sottoposto a misura mediante un apposito attrezzo per VNA che contiene una transizione per microonde, da coassiale a complanare.

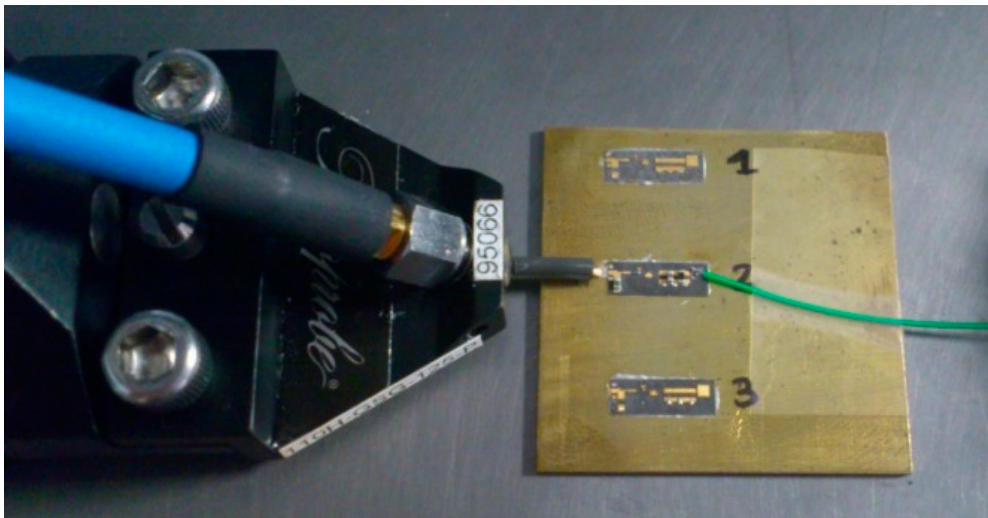


Fig.2.16 Circuiti MW sottoposti a misura mediante complanare

Le complanari in genere si utilizzano fino a 26GHz con transizioni SMA e fino a 40GHz con transizioni per connettori da 2,4mm; oltre a questa frequenza di solito si lavora soltanto in guida d'onda perché i collegamenti coassiali diventano sconvenienti. Spesso per i circuiti miniaturizzati come quello di fig.2.16 si utilizzano anche delle transizioni intermedie, da complanare a microstriscia, ad esempio come quella rappresentata in fig.2.17 che mostra un detector per onde millimetriche in banda K realizzato con un diodo Shottky di tipo ZBD, cioè Zero Bias Diode. In questo caso è stato previsto anche un bonding (cioè un filo termosaldato) per collegare la porta 2 della

transizione alla porta 1 del detector. Esistono tre tipi di linee complanari: quella semplice, quella con uno strato di massa diffuso sulla superficie opposta del substrato e infine c'è anche quella con dei fori passanti che collegano le masse laterali che scorrono superficialmente con quella diffusa sull'altro layer. Questi dettagli solo per offrire al lettore anche un'idea del livello tecnologico necessario per operare con frequenze così elevate.

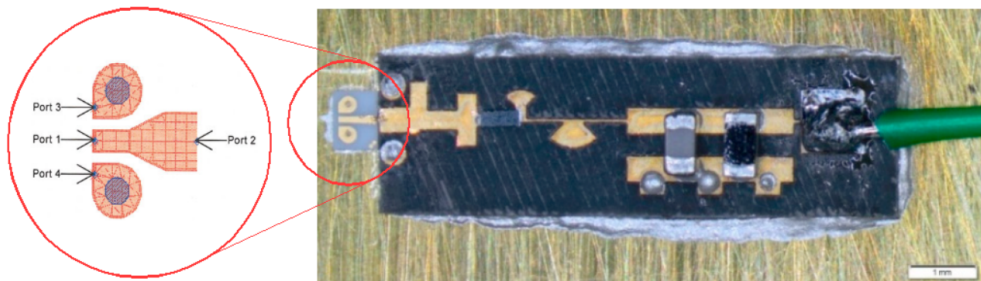


Fig.2.17 Una transizione da complanare a microstriscia per la misura della porta di ingresso di un detector MW

Quasi sempre le complanari sono standardizzate, nel senso che esiste un set di calibrazione per ogni tipo, in grado di azzerare i soliti tre punti lungo l'asse delle impedenze reali sullo strumento (VNA), open, short e load, quest'ultima normalizza anche l'impedenza sul diagramma di Smith.

Cambiamo argomento: passiamo a prendere qualche confidenza con gli oscillatori a diodi RTD, ce ne sono un paio che voglio presentare al lettore. Il primo è un oscillatore classico in cavità in cui le dimensioni del disegno non rendono giustizia alla realtà, purtroppo, perché nella realtà questo oscillatore è veramente un gioiellino della miniaturizzazione meccanica. Siccome il diodo RTD ha una caratteristica tipicamente capacitiva, del resto come lo è un comune diodo tunnel, è necessario montarlo all'interno di una cavità risonante che deve avere una leggera reattanza induttiva per farlo oscillare. In realtà la parte più rilevante di questa induttanza è dovuta al filo di bonding necessario per collegare l'anodo del diodo al contatto collegato al polarizzatore (di solito è l'anodo che prevede il bonding). Anche in questo caso per avere ulteriori informazioni su come calcolare tale reattanza induttiva suggerisco il lettore di riferirsi sempre al paragrafo dove è stata data

un'ampia descrizione del diodo di Esaki, la tecnica con cui si identificano i parametri elettrici di un diodo RTD per ottenere l'oscillazione e calcolare la frequenza di risonanza sono assolutamente identici.

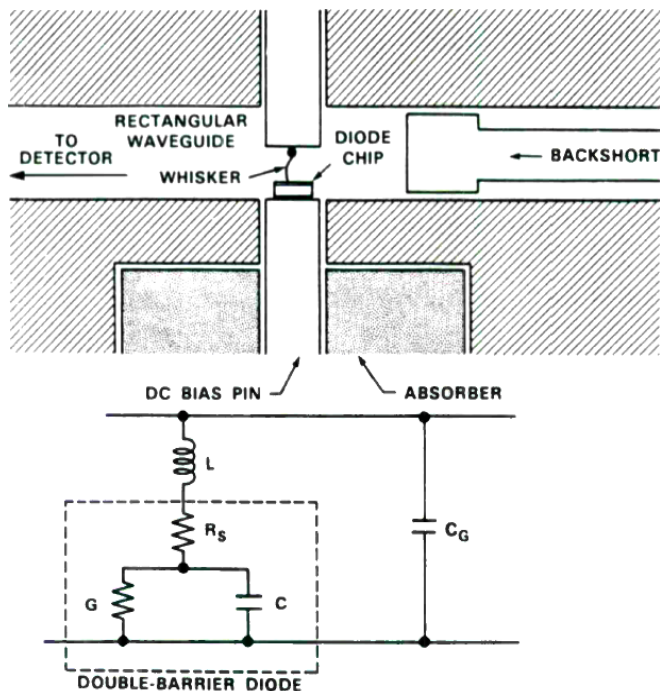


Fig.2.18 Un oscillatore in cavità realizzato con un chip di un diodo RTD

Un oscillatore a RTD per onde millimetriche in genere si presenta come appare in fig.2.18, in basso c'è il suo schema di principio con cui si innesca l'oscillazione. Naturalmente la frequenza di oscillazione dipende dall'induttanza L e dalla posizione del pistone di cortocircuito (backshort) che scorre all'interno della guida, esattamente come accade per un diodo Gunn oppure un diodo tunnel, il principio di innesco dell'oscillazione è identico. Tuttavia quello che ci interessa di più è di conoscere quale è la frequenza massima a cui può oscillare un diodo RTD; essa è la seguente:

$$f_{max} = \frac{1}{2\pi C} \left[\frac{-G_{max}}{R_s} - G_{max}^2 \right]^{1/2} \quad [18]$$

dove:

$$R_S = \rho_{epi} \frac{l}{A} + R_{spread} + R_C \quad \text{e} \quad C = \epsilon \frac{A}{L_{depl}} \quad [19] \text{ e } [20]$$

ρ_{epi} è la resistività degli strati epitassiali, l è la lunghezza e A è l'area della superficie dello strato; R_{spread} è la resistenza dovuta allo spread durante il transito degli elettroni e R_C è la resistenza di carica; ϵ è la costante dielettrica e L_{depl} è l'induttanza che offre la regione di svuotamento.

Faccio qualche esempio: un wafer con una f_{max} di 270GHz con qualche difficoltà si riesce a farlo oscillare a 200GHz utilizzando una guida WR-3 con una potenza di uscita di poco superiore ai 0,1 μ W; con una guida WR-6 lo stesso wafer può oscillare intorno ai 110-150GHz con una potenza di uscita di circa 2 o 3 μ W; invece con una guida WR-15 arriva ad erogare circa 70 μ W a 60-70GHz.

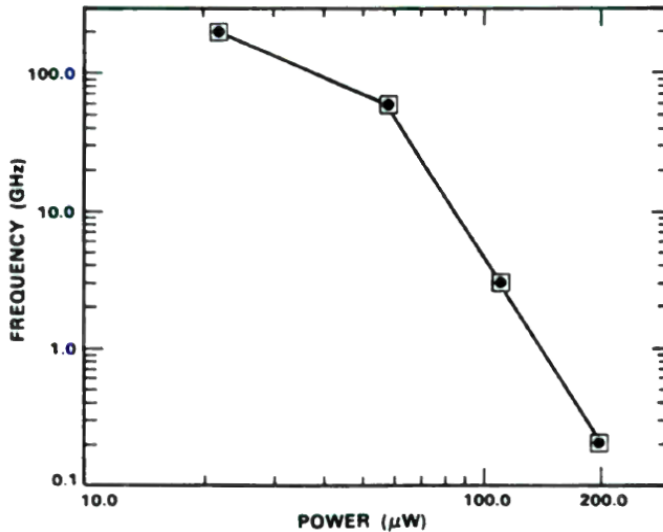


Fig.2.19 Livelli di potenza ottenibili da un chip RTD per 200GHz

Per frequenze inferiori (35-40GHz) si usa una guida WR-22 ma in questo caso è necessario cambiare wafer e usarne uno più adatto per tali frequenze che però erogherà meno potenza, diciamo che essa torna ad attestarsi al

massimo intorno al microwatt, ricordiamoci che l'RTD non è un diodo che può erogare potenze elevate, orientativamente la potenza massima che si riesce ad ottenere da un diodo RTD è pari a:

$$P_{max} = (3/16) \Delta I / \Delta V \quad [21]$$

nel caso del wafer da 200GHz la P_{max} era di $225\mu W$, questo giusto per dare al lettore un'idea sulle grandezze elettriche che si vanno a considerare. Per quanto riguarda invece le grandezze fisiche ricordo che la finestra di apertura di una guida rettangolare WR-3 è pari a: 0,4318mm e 0,8636mm: pertanto alla domanda se si può salire ancora di più in frequenza utilizzando una guida più piccola avete già una risposta.

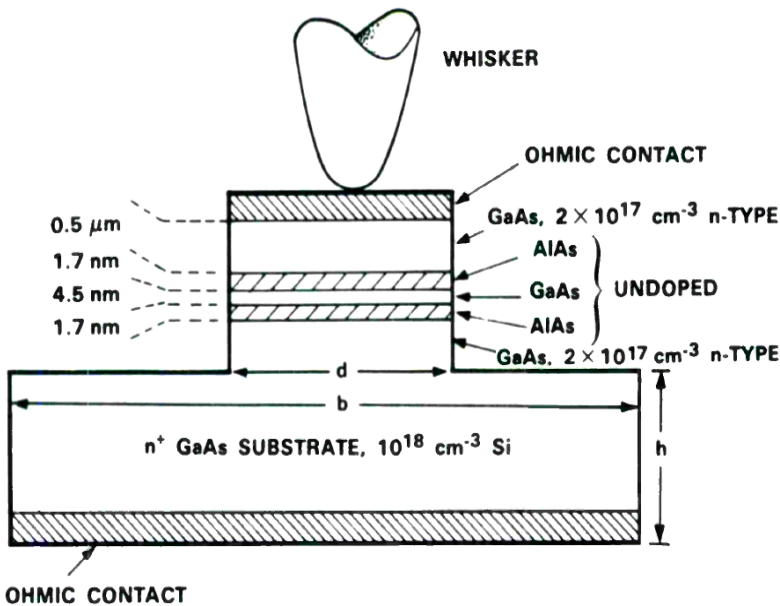


Fig.2.20 Parametri fisici di un diodo Mesa-Etching RTD da 200GHz

A 200GHz è già molto difficile arrivarci con una cavità con le dimensioni di una guida d'onda e oltretutto è anche costosissimo perché per realizzare l'interno di una guida è necessario disporre di una attrezzatura meccanica in grado di garantire un livello di precisione di almeno 10 volte superiore a quello con cui si deve operare; nel caso di una guida WR-3, che richiede un

livello di precisione dell'ordine dei decimi di micron, vuol dire che l'attrezzatura meccanica (la fresa) con cui si intende produrre una guida del genere deve avere un livello di precisione garantito dei centesimi di micron e vi posso garantire che pochissime aziende al mondo dispongono oggi di tali attrezzature meccaniche così sofisticate e con tali livelli di precisione.

Qualora si volesse considerare le dimensioni di un diodo RTD da 200GHz la fig.2.20 può essere utile. La precisione della stratificazione di questo diodo raggiunge i decimi di nanometri ($0,0001\mu\text{m}$), quindi il suo processo di fabbricazione deve garantire almeno un livello di precisione dell'ordine dei centesimi di nanometri, e tale livello può essere ottenuto esclusivamente per crescita epitassiale con tecnologia di asportazione della metallizzazione superflua di tipo "Mesa-Etching", una geniale tecnica giapponese che risale al 1973 e che ancora oggi viene massicciamente utilizzata nella produzione di semiconduttori per microonde e per onde millimetriche. Per far capire il livello tecnologico di tale struttura le dimensioni reali del whisker rispetto a quelle del diodo, ovvero della punta che collega il contatto superiore del diodo con il mondo esterno, sono esattamente quelle che si possono percepire dal disegno tenendo presente che la distanza "d" rappresenta il diametro della superficie circolare Mesa che in genere si attesta fra 1 e $5\mu\text{m}$.

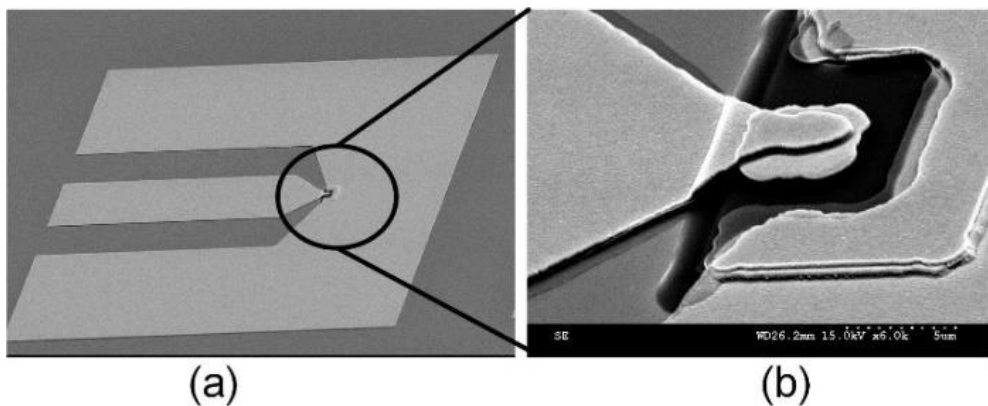


Fig.2.21 Dimensioni reali di un diodo RTD stratificato su substrato (b) e collegato a una linea complanare per la misura delle caratteristiche (a)

I diodi RTD per il montaggio in cavità vengono stratificati direttamente sul supporto metallico opposto a quello dove di collega il whisker e assume

una forma cilindrica; viceversa i diodi RTD destinati al funzionamento su un substrato hanno una forma più squadrata come quella rappresentata in fig.2.21 e si stratificano direttamente su di esso, inclusa la metallizzazione che realizza il collegamento verso il contatto superiore, gli adattamenti si realizzano mediante una linea rastremata (taper). La tabella seguente offre una panoramica di tre tipi diversi di wafer RTD per oscillatori in cavità.

Parametri del materiale	Unità	Wafer 1	Wafer 2	Wafer 3
Materiale della barriera	--	AlAs	Ga _{0,7} Al _{0,3} As	AlAs
Spessore della barriera	nm	2,5	3	1,7
Spessore del pozzo	nm	4,5	4,5	4,5
Doping degli strati esterni	cm ⁻³	1 X 10 ¹⁸	2 X 10 ¹⁷	2 X 10 ¹⁷
Parametri elettrici				
Rapporto picco/valle @300°K	--	1,7/1	1,3/1	3,5/1
Densità della corrente di picco	X10 ⁴ A/cm ²	0,8	1,2	4
Regione di svuotamento +bias	nm	15	30	70
Capacità	fF	100	50	20
Conduttanza negativa max	mS	5	8	13
Resistenza serie	Ω	10	15	15
Caratteristiche di oscillazione				
Bias, IB ; VB	mA ; V	0,7 ; 0,4	2,7 ; 0,32	3,0 ; 0,95
Frequenza di oscillazione	GHz	20,7	43,7	201
Parametri teorici				
Frequenza max di oscillazione	GHz	35	70	270
Frequenza max di svuotamento	GHz	1000	500	230
Frequenza ΔE	GHz	40	1000	400

Tabella 2.1 Parametri e caratteristiche di tre diversi chip RTD per altrettanti oscillatori in cavità; il wafer 3 è quello di fig.2.20

Per quanto riguarda la frequenza massima di svuotamento essa dipende dalla velocità di deriva per la quale si assume un valore di 10⁷cm/s. Invece, la frequenza ΔE si trae dal calcolo dell'ampiezza di trasmissione dell'energia attraverso la struttura a doppia barriera. Per tutti e tre i wafer il diametro della superficie circolare è di 4μm.

Vediamo ora il progetto di un oscillatore modulato in banda W realizzato con due diodi RTD per ottenere una maggiore potenza in uscita, entrambi stratificati su un substrato evitando così di aggiungere il costo della cavità.

L'obiettivo di questo progetto è quello di testare una connettività ad alto bit rate (max 15Gbit/s) sulle brevi distanze (max 50cm) con onde millimetriche al posto dei comuni transponder RF-ID.

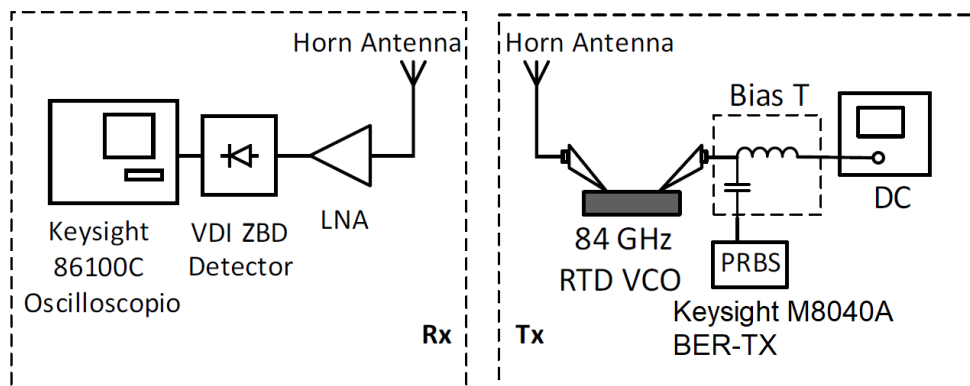


Fig.2.22 Schema di un collegamento wireless in banda W fino a 15Gbit/s

L'idea di questo esperimento, finanziato dalla Comunità Europea, è venuta a Jue Wang (et al.) dell'High Frequency Electronics Group, School of Engineering, dell'Università di Glasgow (UK), in collaborazione con il James Watt Nanofabrication Centre (JWNC), sempre di Glasgow, che ha aiutato gli autori nella costruzione del prototipo.

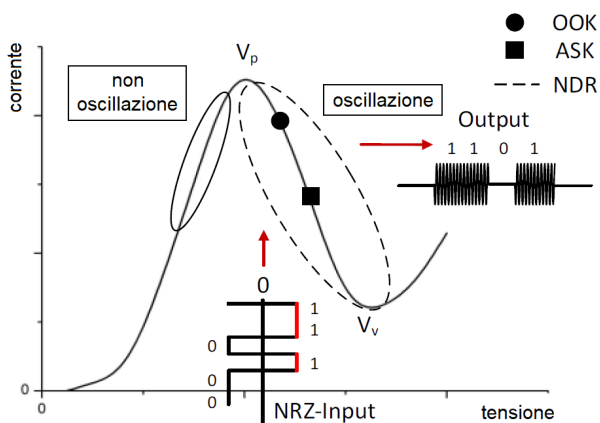


Fig.2.23 Principio di funzionamento del modulatore della parte Tx

Il principio di funzionamento di questo sistema si basa su un oscillatore in banda W a RTD a doppia barriera che viene stimolato direttamente dal segnale modulante attraverso il circuito di polarizzazione, vedi fig.2.23. L'oscillazione si stabilisce soltanto in presenza di una conduttanza negativa (vedi zona tratteggiata), in tutte le altre situazioni l'oscillatore si spegne; pertanto regolando opportunamente l'ampiezza del segnale modulante e il suo offset rispetto al centro della regione in cui la conduttanza diventa negativa è possibile modulare tale oscillatore in due modi diversi: in OOK, vale a dire "On Off Keying" ovvero codifica on-off in cui l'oscillatore si accende in presenza di un "1" logico e si spegne in presenza di uno "0", oppure ASK, cioè "Amplitude Shifting Keying", una sorta di modulazione di ampiezza che distingue il simbolo "0" dal simbolo "1". È chiaro che questa variazione della tensione di polarizzazione determina anche una variazione della frequenza dell'oscillatore oltre alla sua modulazione di ampiezza, tuttavia tale variazione risulta essere abbastanza contenuta come dimostrato dalla curva caratteristica di fig.2.24.

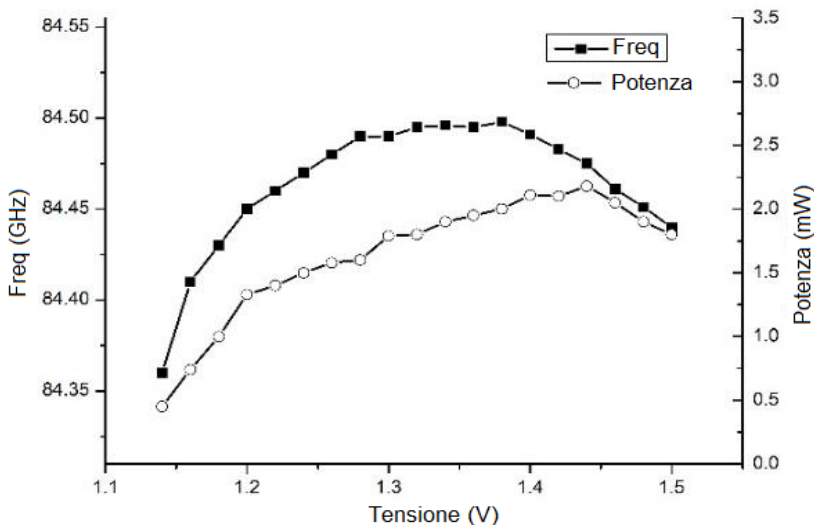


Fig.2.24 Variazione della potenza e della frequenza dell'oscillatore in funzione della sua tensione di polarizzazione in condizioni statiche

Lo schema elettrico e quello di principio che determina il funzionamento dell'oscillatore in questione sono rappresentati in fig.2.25; l'elemento che

determina la frequenza di oscillazione è l'induttanza L che è realizzata direttamente sul substrato e che appare come una microstriscia, ma che in realtà trattasi si una complanare il cui valore è legato all'impedenza della linea, ovvero:

$$L = \frac{Z_0 \tan(\tilde{\nu} l)}{2\pi f_0} \quad [22]$$

dove $\tilde{\nu}$ è il numero d'onda; in questo caso $Z_0 = 50 \text{ ohm}$ e $l = 42 \mu\text{m}$

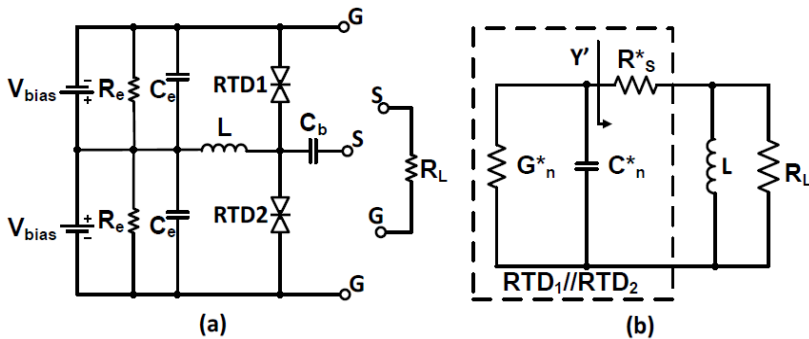


Fig.2.25 Schema elettrico e di principio dell'oscillatore in banda W

La lunghezza dell'induttanza L è determinata dalla posizione delle due capacità C_e rispetto ai due diodi RTD1 e 2; essi sono fabbricati in tecnologia MIM (Metal In Metal) con il suo strato dielettrico SiNx depositato mediante la tecnica ICP, cioè del plasma accoppiato induttivamente in un ambito CVD (Chemical Vapor Deposition); il loro valore è di 2pF ciascuno. Il condensatore di disaccoppiamento della DC in uscita C_b ha invece un valore di 1,5pF. La condizione di oscillazione si ha quando:

$$\text{imag}[Y'] + \omega_0 C_n^* = 0 \quad [23]$$

e la frequenza di oscillazione è calcolata come segue:

$$f_0 = \frac{1}{2\pi L(G_n^* R_s^* + 1)} \sqrt{\frac{L}{C_n^*} - R_s^{*2}} \quad [24]$$

dove Y' è l'ammettenza indicata in fig.2.25(b); essendo un circuito composto da due diodi RTD, ciascuno con la propria capacità C_n , la propria conduttanza G_n e la propria resistenza serie R_s i valori asteriscati di tali parametri da inserire nella [24] sono i seguenti: $C_n^* = 2C_n$ $G_n^* = -2G_n$ e infine $R_s^* = R_s/2$.

La regione della polarizzazione dove si manifesta la NDR nei due diodi, ossia la resistenza negativa o la conduttanza negativa che dir si voglia, si manifesta fra 0,9V e 1,7V, in questo range la conduttanza varia sensibilmente, tuttavia la G_n minima si attesta intorno ai -42mS. Questo significa che il punto di lavoro ottimale per una modulazione ASK è pari a 1,3V; a questa tensione il circuito completo assorbe 88mA con una R_L di 50 ohm. In fig.2.26 c'è una immagine del layout dell'intero oscillatore; in essa si possono distinguere i due diodi RTD, le due capacità C_e che determinano la lunghezza dell'induttanza L , le due resistenze di fuga R_e e infine anche la capacità di blocco della DC C_b . Il circuito appena descritto è veramente molto semplice, anche se il processo per realizzarlo non lo è affatto, oltretutto lavorare un substrato per 84GHz non è proprio un giochino da ragazzi.

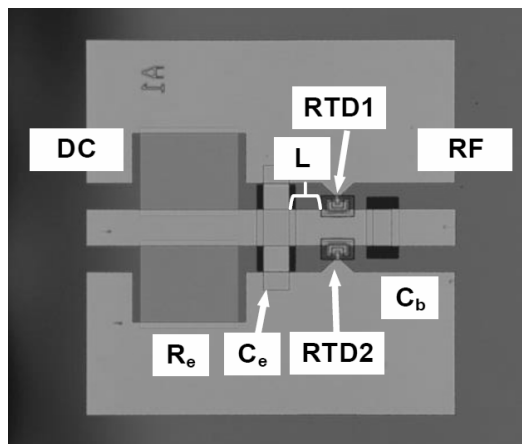


Fig.2.26 Layout completo dell'oscillatore in banda W

A tali frequenze nascono anche numerosi problemi per effettuare correttamente le misure, ad esempio per misurare la frequenza e il livello di uscita è necessario inventarsi diversi trucchetti e sotterfugi.

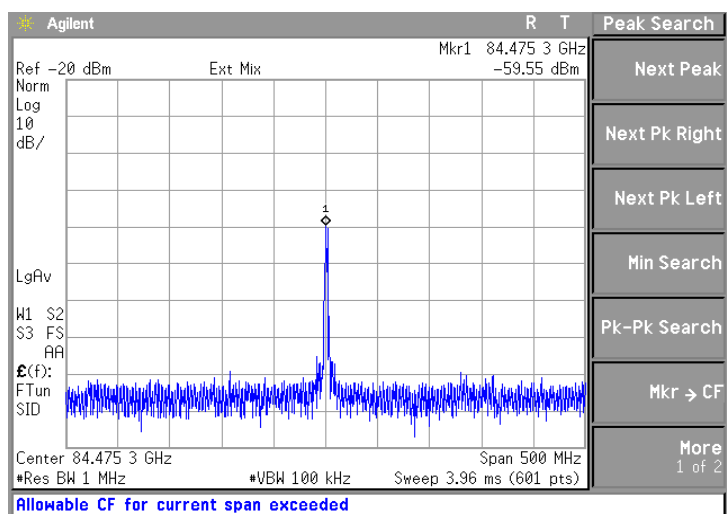


Fig.2.27 Frequenza dell'oscillatore non modulato e polarizzato a 1,3V misurata con un analizzatore di spettro corredato di mixer down-converter

Ad esempio per la banda W ancora oggi è quasi impossibile trovare analizzatori di spettro in grado di offrire una sufficiente precisione nelle misure. In questo caso è stato scelto di lavorare con uno strumento Keysight E4448A che però è limitato a 50GHz. Fra l'oscillatore e lo strumento è stato interposto un mixer down-converter professionale studiato appositamente per la banda W, il Keysight 11970W; questo mixer armonico è in grado di convertire tutte le frequenze fra 75 e 110GHz con una perdita di inserzione massima di 46dB. L'ingresso del mixer però è realizzato in guida WR-10, quindi serve anche un'apposita transizione guida-complanare equipaggiato con una flangia UG-387/U. In fig.2.27 c'è una immagine ricavata dall'analizzatore di spettro in questione che mostra la portante non modulata dell'oscillatore al centro della scala delle frequenze.

Per quanto riguarda invece la parte ricevente di questo sistema lo schema a blocchi del circuito realizzato è sempre riportato in fig.2.22 sul lato sinistro, esso è composto da un Low Noise Amplifier (LNA) seguito da un detector a diodo Schottky per la banda W realizzato dalla famosa Virginia Diodes di Charlottesville in Virginia (USA), il WR10ZBD; infine l'uscita del detector è collegato a un oscilloscopio, per tenere sotto controllo l'apertura del-

l'occhio, e alla sezione ricevente di una coppia NRZ-BERT (Bit Error Ratio Test-set), il Keysight M8040A. Con tutti questi strumenti, e con una coppia di antenne horn disposte a 50cm l'una dall'altra, è stato possibile ottenere dal sistema i seguenti risultati con una modulazione di tipo ASK: a 5Gbps il tasso di errore era di circa 1×10^{-6} , con 10Gbps il tasso era salito a $3,6 \times 10^{-4}$ e infine con 15Gbps si è arrivati a $4,1 \times 10^{-3}$. Quest'ultimo valore non è proprio un gran che, tuttavia si dimostra che con una semplice architettura composta da un LNA e da soli tre diodi, due per la trasmissione e uno per la ricezione, è possibile comunicare a corto raggio con una velocità che era considerata inimmaginabile fino a poco tempo fa.

Concludo questa carrellata di applicazioni per i diodi RTD con quella forse più promettente dal punto di vista della loro diffusione: le celle di memoria ad alta densità. Il comportamento bistabile di un diodo tunnel è stato già sviscerato nel capitolo 2 ma il diodo RTD non è molto diverso e pertanto in grandi linee ricalca le stesse proprietà con qualche vantaggio in più. Uno di questi vantaggi è la sua dimensione fisica, che è molto più piccola di un diodo tunnel e di qualsiasi altro diodo che è stato analizzato fin'ora; oltretutto è estremamente veloce perché ha una capacità intrinseca irrisoria, dell'ordine di qualche decina di femto-Farad, vedi tab.2.1. Pertanto qualcuno ha pensato di realizzare delle memorie SRAM ad alta densità con delle coppie di diodi RTD sfruttando la semplicità di un'architettura bistabile basata su tali diodi ma anche sulla loro affidabilità nel tempo. Il problema che nasce quando in uno spazio estremamente ridotto vengono concentrati numerosi circuiti elettrici indipendenti è la dissipazione termica di tali circuiti.

Il substrato di un diodo RTD è costituito da un semiconduttore o da un cristallo di un elemento chimico che in genere non favorisce una perfetta conducibilità termica verso l'esterno; immaginate un array di coppie di diodi RTD posti uno vicino all'altro, ciascuno collegato a un paio di transistor che funzionano da semplici interruttori per cambiare lo stato logico di queste coppie in modo da scrivere o leggere in tempi rapidi l'informazione contenuta in ciascuna di esse: ebbene se si trattasse di una cella CMOS il calore che si svilupperebbe in quel piccolo spazio sarebbe un calore enorme che difficilmente potrebbe essere smaltito da un dissipatore normale posto a contatto con la superficie opposta del substrato. Ecco ora le belle notizie:

una cella di memoria T-SRAM, cioè le Tunnelling-based Static Random Access Memory, in condizioni statiche potrebbe consumare sulla carta una potenza anche di 7 ordini di grandezza inferiore rispetto a una singola cella di memoria C-MOS; quindi in teoria con la potenza assorbita da una sola cella CMOS si potrebbero alimentare circa 10 milioni di celle di memoria realizzate con diodi RTD. Il lettore non avrà difficoltà a capire che si tratta di una cifra enorme; ma proprio per questo motivo è facile intuire che, malgrado l'alta densità di celle per unità di superficie, tali celle di memoria non hanno alcun bisogno di un sistema di raffreddamento particolare per dissipare il calore che esse stesse producono, basta un sistema non molto diverso da uno di quelli standard già ampiamente utilizzati oggi per raffreddare le memorie o i microprocessori, vale a dire con scambiatori di calore di tipo aria-aria o al massimo con quelli aria-liquido. Al momento si stanno sperimentando queste celle di memoria in condizioni di temperatura estreme perché il comportamento dei diodi RTD è molto sensibile alle variazioni termiche. Naturalmente questi casi limite servono per capire quali potrebbero essere i materiali semiconduttori più indicati per soddisfare le varie applicazioni, commerciali, industriali, difesa, spazio e quant'altro; ma soprattutto le sperimentazioni servono per far emergere tutte le potenziali criticità di queste nuove tecnologie, anche nei processi produttivi.

Tali sperimentazioni per il momento avvengono "in ghiacciaia", nel senso che per i test dinamici di queste memorie è previsto un particolare sistema criogenico per cui esse vengono fatte funzionare quando sono completamente immerse nell'azoto liquido a 77°K (-196°C). C'è un motivo per scendere a queste temperature così estreme: esso è legato alla compatibilità dei materiali utilizzati per integrare su un unico substrato sia i diodi RTD che i MOSFET che fungono da interruttori

per scrivere e leggere i dati messi in memoria. A temperatura ambiente i diodi RTD prodotti con materiali a basso costo e compatibili con le tecnologie MOSFET, in genere silicio e il germanio, hanno un PVCR molto basso ma

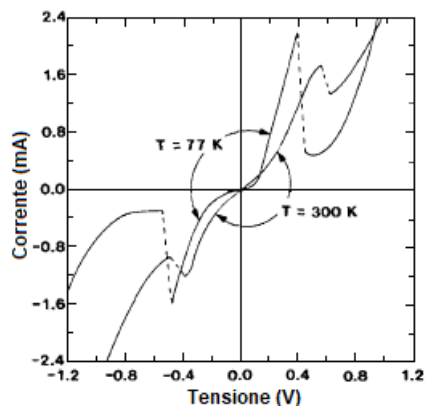


Fig.2.28 RTD a 300°K e a 77°K

per fortuna alle basse temperature questi diodi si comportano molto meglio, assorbimenti a parte, vedi la fig.2.28. Il PVCR aumenta sensibilmente e questo rende più sicura e stabile la collocazione della retta di carico che intercetta la curva al centro della regione dove il diodo manifesta la sua NDR per ottenere la desiderata “bistabilità” del dispositivo. Si rimanda il lettore alla trattazione dell'argomento specifico osservando la fig.1.10 e leggendo il seguito.

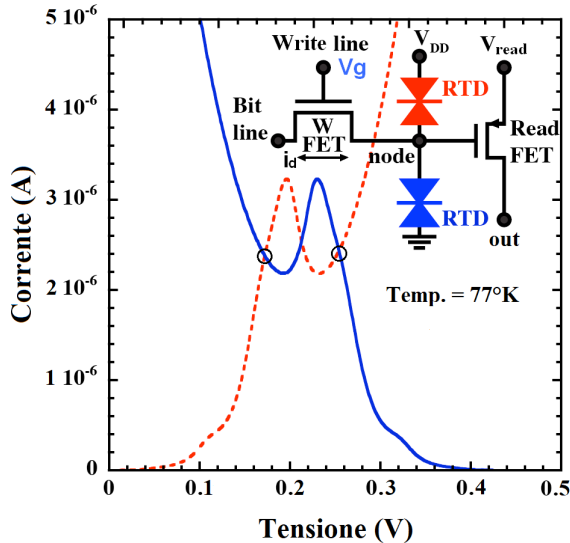


Fig.2.29 Schema di principio di una cella di memoria T-SRAM e dell'intersezione delle curve V-I di due diodi RTD posti in serie

Lo schema di principio di una cella di memoria T-SRAM è rappresentato in fig.2.29. Il nodo centrale che congiunge i due diodi è nominato “storage node”, trattasi del nodo dove viene effettivamente immagazzinato il dato. La curva in rosso rappresenta la caratteristica V-I del diodo collegato alla V_{DD} mentre quella opposta appartiene al diodo evidenziato in blu, essa è speculare perché il diodo blu rappresenta il carico del diodo rosso e viceversa. Il FET di sinistra in realtà è una struttura bidirezionale a MOSFET che è deputata a “scrivere” il dato nella cella di memoria composta dai due RTD; il FET di destra, invece, è quello deputato a “leggere” il dato scritto in memoria e a renderlo disponibile sul nodo “out”. Le intersezioni fra le due curve indicate con un piccolo cerchio nero sono i due punti di stabilità della struttura e

costituiscono i due livelli logici scritti in memoria. In fig.2.30 è rappresentato invece il ciclo di scrittura del dato presente in ingresso, indicato con la linea tratteggiata, sul nodo di memoria, il livello indicato con la linea rossa; il comando di scrittura è invece rappresentato dalla linea tratteggiata blu, e nel diagramma in basso sono invece evidenziate le correnti.

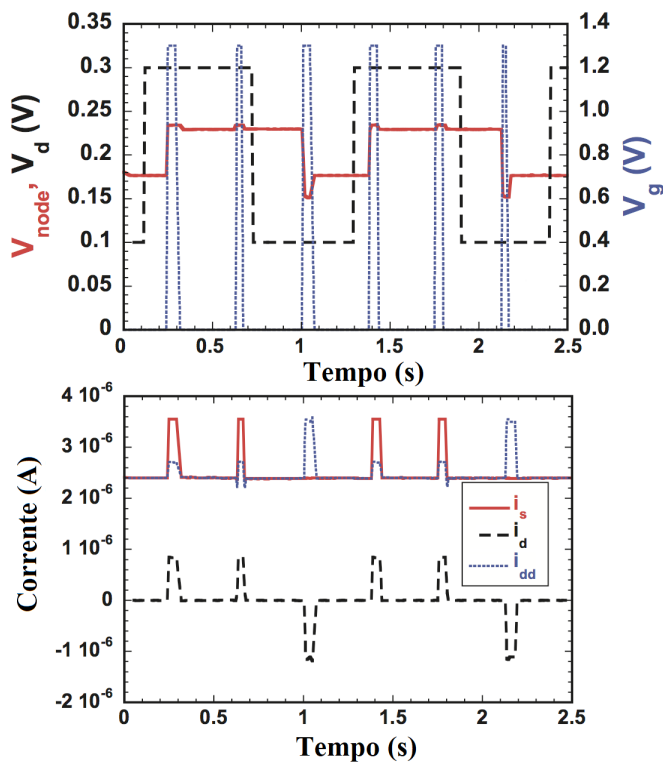


Fig.2.30 Ciclo di scrittura di una cella di memoria T-SRAM

L'obiettivo finale delle attuali indagini condotte su questo nuovo tipo di memorie è quello di individuare un buon compromesso fra le notevoli prestazioni che si ottengono con queste tecnologie ibride con i diodi RTD e la loro applicabilità in campo pratico, cercando di mantenere contenuti i costi di produzione, inclusi quelli legati all'approvvigionamento dei materiali e alle loro lavorazioni. Le memorie T-SRAM saranno disponibili sul mercato non appena sarà raggiunto questo obiettivo.

2.1 Diodi tunnel risonanti a barriera tripla (TBRTD o DWRTD)

Visti i progressi del diodo tunnel risonante a doppia barriera rispetto a quello originale a barriera singola ci sarebbe da domandarsi se vale la pena oppure no aumentare il numero di queste barriere per ottenere un diodo tunnel “perfetto”; ebbene a questa domanda non vi sono ancora delle risposte precise perché in effetti qualcuno ha già realizzato dei diodi a barriera tripla per scopi sperimentali, tuttavia aumentare il numero delle barriere porta inevitabilmente ad alcuni vantaggi e ad altrettanti svantaggi. Facciamo ora il punto, e vediamo meglio come organizzare questo argomento; innanzi tutto cerchiamo di capire cosa succede se aggiungessimo una terza barriera al diodo tunnel risonante RTD, potremo avere una situazione di questo tipo, vedi l'esempio di fig. 2.31.

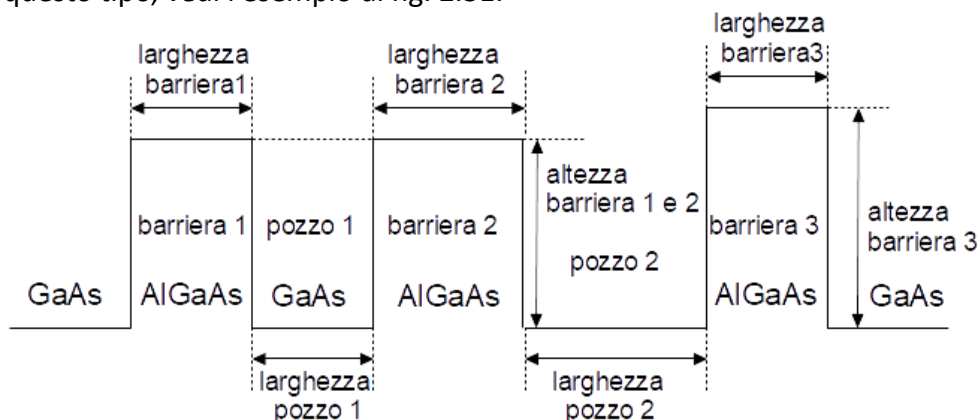


Fig.2.31 Una possibile situazione in un diodo a barriera tripla

In questo contesto potremo trovarci di fronte a tre barriere con caratteristiche fisiche diverse e intervallate da due pozzi quantici, eventualmente anch'essi di dimensioni diverse; l'altezza delle barriere di potenziale (band-gap) può essere regolata dalla percentuale di doping nel reticolo cristallino (superlattice) fra un minimo di 1,42eV (GaAs) e un massimo di 2,16eV (AlAs), vedi fig.2.6; la larghezza delle barriere e dei pozzi quantici, invece, può essere regolata dalla dimensione degli strati depositati sul substrato di base su cui si fa crescere il diodo. In base a numerosi risultati sperimentali è possibile affermare che un diodo a tripla barriera offre i seguenti vantaggi:

- **Rivelazione a polarizzazione zero:** i TBRTD (Triple Barrier RTD) presentano spesso caratteristiche corrente-tensione asimmetriche, che consentono anche il funzionamento a polarizzazione zero. Ciò significa che possono fungere da rivelatori altamente sensibili per segnali ad alta frequenza (onde millimetriche e sub-millimetriche) senza la necessità di avere una polarizzazione in corrente continua.
- **Maggiore flessibilità nella progettazione del dispositivo:** l'aggiunta di una terza barriera crea due pozzi quantici accoppiati anziché uno. Questo parametro di progettazione aggiuntivo consente di ottenere una migliore precisione con cui si determinano i livelli di energia in condizioni di risonanza, la probabilità di trasferimento degli elettroni e le prestazioni complessive del dispositivo semiconduttore.
- **Migliore rapporto corrente picco-valle (PVCR):** introducendo un secondo pozzo quantico, i TBRTD possono sopprimere la corrente di valle (corrente di dispersione) in modo più efficace rispetto ai DBRTD, portando a un PVCR più elevato e, di conseguenza, a migliori prestazioni di commutazione.
- **Prestazioni e potenza alle alte frequenze:** i TBRTD sono in grado di generare e rilevare segnali nella gamma delle onde millimetriche e oltre; offrono una potenza di uscita più elevata ottimizzando il compromesso fra lo spessore dello strato spaziatore nel collettore, i profili di doping e la coerenza degli elettroni nella struttura.
- **Regioni multiple di resistenza differenziale negativa (NDR):** i TBRTD asimmetrici possono presentare regioni NDR multiple e distinte, offrendo così maggiori funzionalità in un singolo dispositivo per applicazioni speciali quali le memorie a tre stati logici e la generazione di segnali criptati.
- **Studio avanzato della coerenza quantistica:** l'accoppiamento tra i due pozzi può essere modulato regolando lo spessore della barriera centrale, rendendoli strumenti eccellenti per lo studio della coerenza quantistica e i fenomeni di trasporto elettronici.

Tuttavia esistono anche degli svantaggi, eccoli elencati di seguito.

- **Ridotta densità di corrente complessiva:** sebbene i diodi a tripla barriera migliorino il PVCR, in genere producono una densità di corrente totale inferiore rispetto ai comuni RTD a doppia barriera.
- **Fabbricazione più complessa:** la creazione di strutture a tripla barriera richiede tecnologie di crescita epitassiale di alta precisione (come MOCVD o MOVPE) per controllare più strati a livello atomico, rendendo la loro fabbricazione decisamente più complessa rispetto ai diodi a doppia barriera.
- **Sensibilità ai parametri di progettazione:** si può dimostrare che le prestazioni del dispositivo a tripla barriera, in particolare la sua NDR, sono altamente sensibili allo spessore della barriera centrale. Una progettazione imprecisa altera sensibilmente le caratteristiche desiderate, esse si possono degradare con estrema facilità, ad esempio, la NDR può scomparire del tutto se la barriera centrale fosse troppo sottile.
- **Difficoltà di misurazione:** poiché questi dispositivi sono altamente instabili nella regione NDR, testare e identificare le loro vere caratteristiche I-V è difficile a causa delle oscillazioni parassite intrinseche.
- **Integrazione limitata:** nonostante le loro potenzialità, l'integrazione pratica e su larga scala dei diodi TBRTD con la tecnologia CMOS convenzionale per realizzare memorie ad alta densità a triplo livello logico (logica ternaria) ancora oggi rimane una sfida significativa.

Quest'ultimo punto è quello principale ma anche quello più critico su cui inizialmente si era concentrato lo sviluppo di questi dispositivi così particolari; la logica ternaria (da non confondere col terzo stato ovvero il "tristate" o "HiZ", lo stato ad alta impedenza tipico dei buffer digitali che scrivono in tempi alterni su un singolo bus di dati) è un sistema logico a tre valori (spesso 0, 1, 2 oppure vero, falso, sconosciuto/incerto/maybe), a differenza della

logica binaria standard (0, 1), si utilizzata spesso in informatica teorica e in alcuni circuiti specializzati, esso promette una maggiore densità di informazione e potenziali vantaggi nei calcoli rispetto al sistema binario.

Ecco i punti chiave della logica ternaria:

1. **Hardware ternario:** Esistono calcolatori ternari che funzionano su base numerica 3 (ternaria), offrendo circuiti teoricamente più efficienti.
2. **Operatori:** Include operatori monadici (come l'inverter ternario negativo, standard o positivo) e diadici, con porte logiche differenti da quelle binarie classiche.
3. **Vantaggi:** Maggiore densità di informazioni, minore complessità di cablaggio e maggiore efficienza in certe operazioni logiche complesse.

Di seguito le applicazioni

- a) **Informatica:** i progetti di ricerca, come il Ternary Spacecraft, esplorano questa logica per futuri scopi spaziali.
- b) **Elettronica:** la logica a tre stati (o ternaria) potrebbe sostituire presto l'attuale logica binaria usata nell'elettronica digitale.
- c) **Filosofia/Teologia/AI (Intelligenza Artificiale):** il concetto di "trinità" viene talvolta analizzato come una struttura logica triadica, che è diversa da quella binaria. La logica triadica è un modello dialettico, celebre soprattutto nella filosofia di Hegel, strutturato su tre momenti dinamici: tesi (affermazione), anti-tesi (negazione) e sintesi (superamento).

In sintesi, la logica ternaria offre un'alternativa più complessa ma potenzialmente più efficiente, capace di gestire l'incertezza (il terzo stato) in modo più nativo rispetto al sistema binario tradizionale. Il "mattoncino" principale di una memoria a tre stati doveva essere proprio il diodo TBRTD.

Riassumendo per grandi linee i diodi tunnel risonanti a tripla barriera offrono indubbiamente qualche vantaggio concreto rispetto a quelli a barriera doppia, nel senso che si migliorano le prestazioni sia come rivelatori che come generatori di segnali alle frequenze più elevate, tuttavia tali vantaggi come sempre non sono del tutto “gratuiti” e comunque non si tratta di qualcosa di rivoluzionario; detto questo il lettore potrà facilmente intuire che tutta la letteratura che oggi è possibile trovare su ipotetici diodi tunnel a barriera multipla con un numero di barriere maggiore di tre è solo frutto di pura teoria accademica, infatti si tratta di esercizi teorici complessi con cui gli studenti di fisica elettronica applicata ai semiconduttori si affacciano per la prima volta su problematiche inerenti nuovi possibili strutture che si possono realizzare con le moderne tecnologie epitassiali di cui disponiamo.

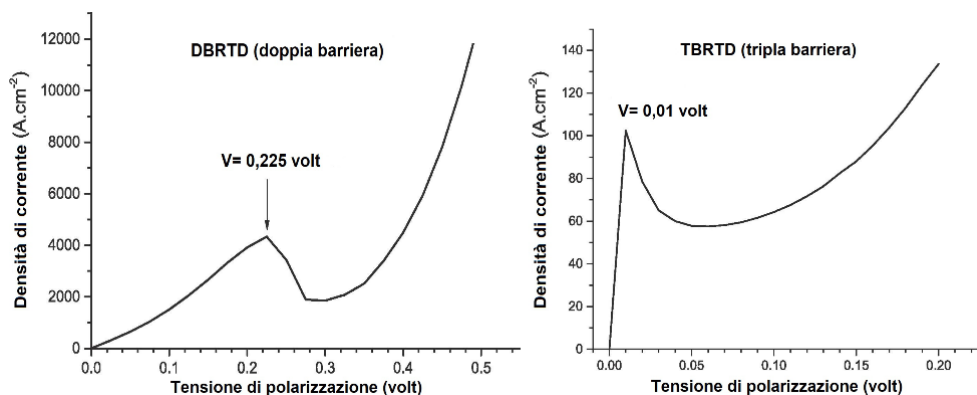


Fig.2.32 Confronto di NDR fra diodi a doppia e tripla barriera

Per valutare le migliori capacità di rivelazione di piccoli segnali (maggiore sensibilità) di un diodo TBRTD rispetto a un DBRTD è sufficiente osservare la fig.2.32; nonostante l'inferiore densità di corrente che attraversa il semiconduttore in condizioni statiche, si può osservare che l'effetto tunnel nel TBRTD nasce immediatamente, con una pendenza molto più pronunciata e senza la necessità di doverlo polarizzare; questo gli consente di avere un primo picco intorno ai 10mV rispetto ai 225mV tipici di un DBRTD. L'assenza di polarizzazione in un diodo rivelatore determina una sensibile riduzione del rumore intrinseco dovuto allo spostamento delle cariche elettriche inutili all'interno del semiconduttore che normalmente inficiano sul rapporto segnale-rumore, un importante elemento qualitativo di un diodo rivelatore,

soprattutto alle frequenze più elevate in cui il rumore che producono le cariche elettriche in transito può essere dominante. Questo primo picco ha un impatto notevole se il diodo venisse utilizzato come rivelatore, ma nel caso in cui si debba utilizzare lo stesso diodo come fonte di un segnale per onde sub-millimetriche si dovrà polarizzare leggermente il diodo in modo che la retta di carico vada a sfruttare il picco immediatamente successivo che in genere determina una NDR molto più estesa e ciò consente di ottenere un PVCR maggiore e quindi una maggior potenza in uscita.

Tornando alla struttura del diodo e osservando la fig.2.2 si può notare che un diodo a barriera multipla, (doppia, tripla, ecc.) non prevede alcun doping nella zona in cui avviene la trasmissione di cariche elettriche per effetto tunnel; in particolare la zona contenuta all'interno delle barriere di potenziale (barriere incluse) in genere non supera mai i 10nm; le uniche variabili che determinano la probabilità di un flusso di elettroni all'interno del semiconduttore sono la concentrazione di alluminio rispetto al gallio nei reticoli cristallini che costituiscono i vari strati che vengono depositati sopra la zona catodica, cioè l'emettitore, nonché lo spessore di tali strati. Poiché in un diodo a barriera tripla gli strati da considerare sono cinque, contro i tre previsti in un diodo a barriera doppia, per una maggiore accuratezza si è deciso di scendere di un ordine di grandezza nel definire lo spessore di tali strati, cosicché nei diodi a barriera tripla invece di parlare di nanometri si parlerà di ångström (Å), una unità di misura che prende il nome dal noto fisico svedese Anders Jonas Ångström, uno dei padri della spettroscopia, sapendo che $1\text{nm}=10\text{Å}$.

Per valutare meglio il TBRTD bisogna innanzi tutto partire dalle sue caratteristiche fisiche; abbiamo già visto in fig.2.31 che tale nanostruttura permette di ottenere anche delle asimmetrie che possono coinvolgere sia le barriere che i pozzi quantici, tuttavia per comodità in questo contesto ci limiteremo ad analizzare solo strutture simmetriche cercando di focalizzarci sull'elemento più sensibile di tutta la struttura, ovvero la barriera centrale. Per prima cosa analizzeremo la variazione dello spessore e della concentrazione di alluminio nel reticolo cristallino che costituisce tale strato di semiconduttore, poi si vedrà cosa accade quando il diodo viene polarizzato con una debole differenza di potenziale elettrico; infine si valuterà quanto il

TBRTD sia sensibile ai campi magnetici, una caratteristica spesso sottovalutata nei semiconduttori che però da decenni ha sempre prodotto numerose applicazioni alternative che un tempo erano definite “impensabili”; basti pensare all'effetto di Hall (da Edwin Hall, 1879), che si basa sulla forza di Lorentz, con cui si dimostra che un campo magnetico trasverso è in grado di deviare i portatori di carica anche in un semiconduttore modificandone così il comportamento.

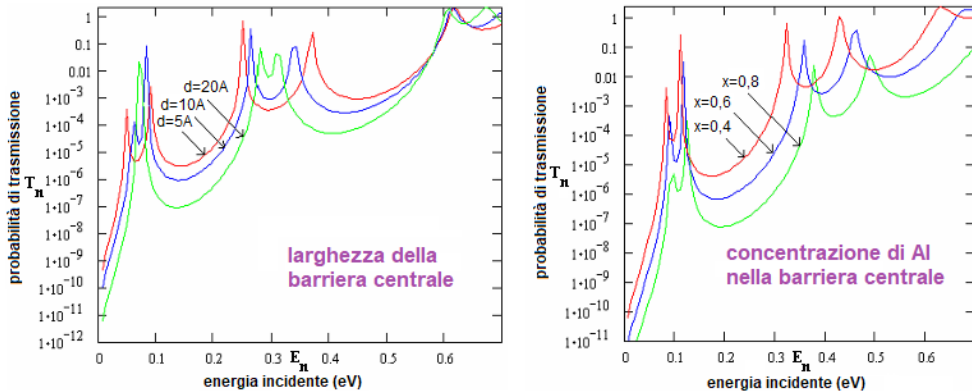


Fig.2.33 Probabilità di trasmissione nel superlattice al variare delle caratteristiche fisiche in un diodo TBRTD

In un TBRTD all'arseniuro di gallio $GaAs/Al_yGa_{1-y}As$ avente una uguale frazione molare di alluminio all'interno delle tre barriere, ovvero $y_{B1}=y_{B2}=y_{B3}=0,67$, uno spessore identico delle due barriere laterali, $L_{B1}=L_{B3}=30 \text{ \AA}$ e lo stesso dicasi per i due pozzi quantici aventi una frazione molare di alluminio pari a zero $L_{QW1}=L_{QW2}=60 \text{ \AA}$, al variare dello spessore della barriera centrale $L_{B2}=d$ otterremo la situazione mostrata in fig.2.33 a sinistra. È immediato notare che con uno spessore di 20 \AA il primo doppietto di picchi di risonanza (livello fondamentale) si fonde in unico picco che determina però il livello più alto di PVCR che si colloca nelle immediate vicinanze del livello più basso di energia incidente; anche il secondo doppietto che costituisce la seconda risonanza (livello eccitato) tende a fondersi ma non riesce ancora a farlo. Pertanto, con spessori più piccoli (5 \AA e 10 \AA) le probabilità di trasmissione degli elettroni per effetto tunnel diventano quattro al di sotto di un valore di energia incidente di $0,6 \text{ eV}$ e questo dipende dalle risonanze che si vengono ad instaurare all'interno delle tre

barriere; i picchi con probabilità di trasmissione più elevata sono il terzo e il quarto (secondo doppietto) e aumentano la loro distanza e il loro livello riducendo lo spessore della barriera centrale.

In fig.2.33 sulla destra invece abbiamo una situazione completamente diversa: in questo caso abbiamo un diodo con una struttura simmetrica che ha una uguale concentrazione di alluminio nelle due barriere laterali, cioè $y_{B1}=y_{B3}=0,6$ ovvero una massa effettiva $m_{B1}^*=m_{B3}^*=0,1168 m_0$ e un livello barriera di $V_0=0,465525 eV$ e con una massa effettiva degli strati di GaAs nel pozzo quantico di $m_{QW1}^*=m_{QW3}^*=0,067 m_0$ e un $V_0=0 eV$; a questo punto variando la frazione molare (x) nella sola barriera centrale si ottengono le tre curve rappresentate in figura. Nella probabilità di trasmissione si osservano i picchi per le energie di risonanza E_{01}, E_{02} e E_{11}, E_{12} .

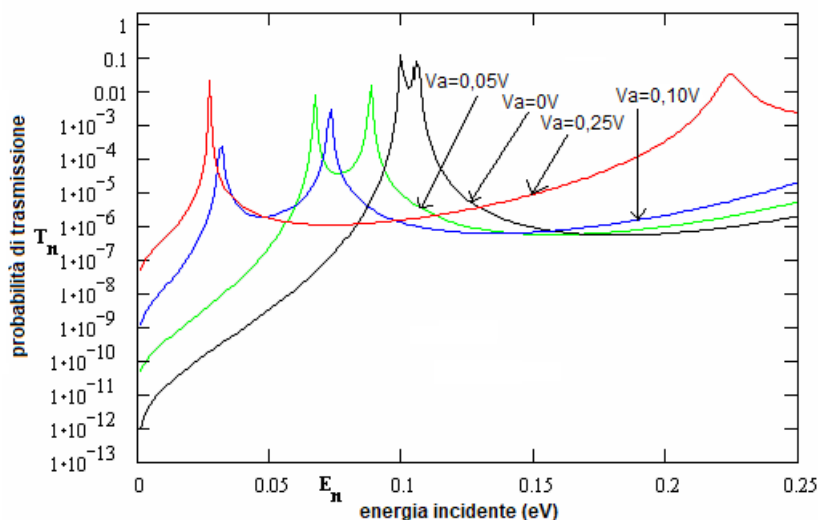


Fig.2.34 Probabilità di trasmissione nel superlattice al variare della polarizzazione in un diodo TBRTD

Quelle rappresentate in figura sono le energie corrispondenti al doppietto quasi-legato fondamentale e al primo doppietto eccitato. Aumentando la concentrazione di alluminio della barriera di accoppiamento, i picchi di trasmissione si spostano verso energie più elevate ma l'energia di accoppiamento per il doppietto quasi-legato fondamentale presenta un incremento

diverso rispetto a quella per il doppietto eccitato; inoltre tale doppietto tende ad avvicinarsi sempre di più a quello fondamentale.

In fig.2.34 è possibile osservare l'effetto di spreading del primo doppietto di risonanza all'aumentare della tensione di polarizzazione del diodo (V_a) che si ottiene applicando una leggera tensione sul suo contatto anodico (collettore); è evidente l'immediata separazione dei primi due picchi e il conseguente spostamento della regione in cui si manifesta la prima NDR.

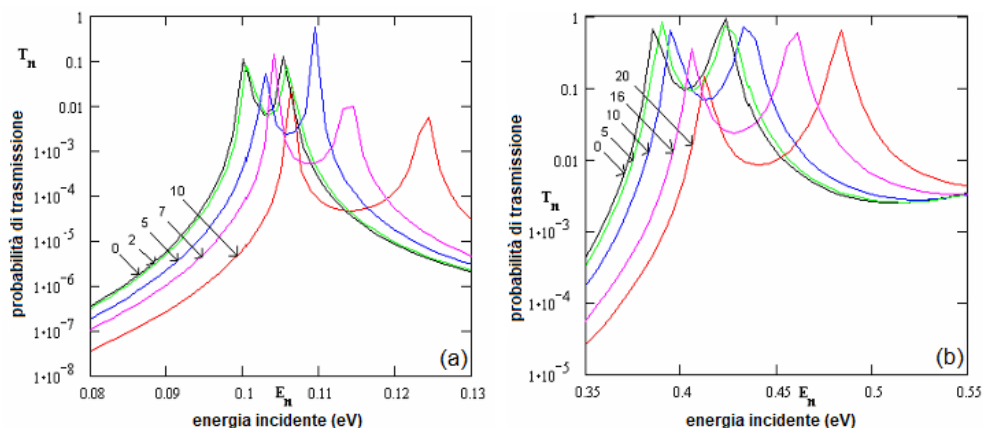


Fig.2.35 Probabilità di trasmissione del primo (a) e del secondo doppietto (b) al variare dei campi magnetici trasversi in un diodo TBRTD

Il fig.2.35, invece, è possibile osservare l'effetto di slittamento del primo e del secondo doppietto di risonanza verso livelli di energia più elevati all'aumentare del campo magnetico trasverso in cui viene immerso un diodo TBRTD, le unità indicate nelle due figure sono espresse in Tesla; anche lo spreading fra le coppie di picchi E_{01} , E_{02} e E_{11} , E_{12} ne risente sensibilmente. Inutile aggiungere che le probabilità di trasmissione e il livelli di energia incidente variano anche al variare della temperatura e di altre condizioni a contorno, ad esempio all'esposizione diretta delle radiazioni elettromagnetiche, inclusa la luce e tutto lo spettro luminoso di cui siamo a conoscenza, tuttavia in questi casi è necessario prevedere contatti elettrici particolari in grado di far transitare liberamente tali radiazioni sugli strati di semiconduttore depositati sul substrato su cui si è fatto crescere il diodo.

RIFERIMENTI E BIBLIOGRAFIA

N.B.: i riferimenti virgolettati sono articoli e pubblicazioni, gli altri sono libri.

ABRIE P. L. D.: *Design of RF Microwave Amplifiers and Oscillators*, Artech House, Boston – London 1999.

BABA R.: “*Resonant tunnelling diodes for THz communications*”, University of Glasgow, 2018.

BUTTON K. J.: *Topics in Millimeter Wave Technology*, Academic Press, USA, 1988.

CHATTERJEE R.: *Elements of Microwave Engineering*, Ellis Horwood Limited, Halsted Press (John Wiley & Sons), Chichester, 1985.

CRISTIAN E. SIMION, CRISTIAN I. CIUCU: “*Triple-barrier resonant tunneling: a transfer matrix approach*”, Romanian Reports in Physics, Vol. 59, No. 3, P. 805–817, 2007.

EID A., HESTER J., TENTZERIS MANOS M.: “*A 5.8 GHz Fully-Tunnel-Diode-Based 20 μ W, 88mV, and 48 dB-Gain Fully-Passive Backscattering RFID Tag*”, IEEE/MTT-S International Microwave Symposium, 2020.

GENERAL ELECTRIC: *Tunnel Diode Manual: Circuits, Applications, Specifications*, Semiconductor Products Dept. General Electric Company, 1961.

GENTILE SYLVESTER P.: *Basic Theory and Application on Tunnel Diodes*, D. Van Nostrand Company Inc., Princeton, New Jersey, 1962.

GHOSH K., SINGISETTI U.: “*Design and Analysis of High Frequency InN Tunnel Transistors*”, Department of Electrical Engineering, University at Buffalo, The State University of New York, Buffalo, NY, 2023.

GUPTA K. M., GUPTA N.: *Advanced Semiconducting Materials and Devices*, Springer, 2016.

GUTIÉRREZ J, ZELJAMI K., PASCUAL J. P., FERNÁNDEZ T., TAZÓN A.: “*Comparison of Microstrip W-Band Detectors Based on Zero Bias Schottky-Diodes*”, MDPI, 2019.

JIT S., PAL B.B.: *"A Simple Analytical Model for the Study of Optical Bistability Using Multiple Quantum Well p-i-n Diode Structure"*, Journal of Semiconductor Technology and Science, 2004.

KESAN V. P., MILLER D. R., REDDY V. K., GULLAPALLI K. K., NEIKIRK D. P.: *"Influence of Space Charge and Quantum Well Negative Resistances on Resonant Tunneling Diodes"*, The University of Texas at Austin, no-date.

KORYU ISHII T.: *Handbook of Microwave Technology*, Vol.1 Components and Devices, Academic Press, USA, 1995.

KORYU ISHII T.: *Practical Microwave Electronic Devices*, Academic Press, USA, 1990.

LIAN W.: *"Resonant Tunneling Diode Mixer and Multiplier"*, Beijing University of Posts and Telecommunications, China 1990.

LIAO SAMUEL Y.: *Microwave Devices and Circuits*, Prentice Hall Inc., Englewood Cliffs, New Jersey, 1980.

LOTFI M., ZOHIR D.: *"PSPICE Implementation of a New Esaki Tunnel Diode Macro-Model"*, International Journal of Smart Home Vol. 10, No. 4 (2016), pp. 153-166.

MILES R.E., X -C ZHANG, EISELE H., KROTKUS A.: *Terahertz Frequency Detection and Identification of Materials and Objects*, Springer + NATO (US Army) + USAF, 2007.

MUNSTERMAN G. T.: *"Tunnel Diode – Microwave Amplifiers"* APL Technical Digest, May-June 1965.

NISHIZAWA JUN-ICHI: *"New Development of Solid State Sub-Millimeter Source"*. Research Institute of Electrical Communication, Tohoku University, Sendai, Japan, 1982.

POZAR D. M.: *Microwave Engineering*, John Wiley & Sons Inc., New York, 2012.

RCA: *RCA Tunnel Diodes for Switching and Microwave Applications*, Somerville New Jersey (USA), 1963.

ROY DILIP K.: *Physics of Semiconductor Devices*, Universities Press, Hyderabad, India, 2013.

SHOCKLEY W.: *"Electrons and Holes in Semiconductors (with applications to transistor electronics)"*, D. Van Nostran Company, Bell Telephones Laboratories Inc. 1959.

SOLLNER T. C. L. G.: *"Submillimeter Quantum Electronics"*, M.I.T. Lincoln Laboratory, 1989.

SORRENTINO R., BIANCHI G.: *Ingegneria delle Microonde e Radiofrequenze*, McGraw-Hill, Milano, 2006.

STEENSON D. P.: *"The High-Frequency Application of Double-Barrier Resonant Tunnelling Diodes"*, Nottingham University, 1993.

STREETMAN B. G., BANERJEE S. K.: *Solid State Electronic Devices*, Pearson, Harlow, London, 2016.

SWANSON D. G., HOEFER W. J. R.: *Microwave Circuit Modeling Using Electromagnetic Field Simulation*, Artech House, Boston London, 2003.

SZE S. M., KWOK K. NG.: *Physics of Semiconductor Devices*, Wiley Interscience, Hoboken, New Jersey, 2007.

TERNENT G., PAUL D. J.: *"Si/SiGe Tunnelling Static Random Access Memories"*, ECS Transactions, The Electrochemical Society, 2013.

TURNER RUFUS P.: *Diode Circuits Handbook*, Howard W. Sams & Co., Inc., The Bobbs-Merrill Company Inc., Indianapolis, Indiana, 1963.

WANG J, AL-KHALIDI A., MORARIU R., OFIARE A., WANG L., WASIGE E.: *"15 Gbps Wireless Link Using W-band Resonant Tunnelling Diode Transmitter"*, University of Glasgow, 2019.

YNGVESSON K. S.: *Microwave Semiconductor Devices*, Springer Science + Business Media, New York, 1991.

ZHAO J., ZHU Z., CUI W., XU K., ZHANG B., YE D., LI C., RAN L.: *"Power Synthesis at 110-GHz Frequency Based on Discrete Sources"*, IEEE Transactions on Microwave Theory and Techniques, 2015.