Design Considerations for Monolithic Microwave Circuits

ROBERT A. PUCEL, FELLOW, IEEE

MTT National Lecture Invited Paper

Abstract—Monolithic microwave integrated circuits based on silicon-on-sapphire (SOS) and gallium arsenide technologies are being considered seriously as viable candidates for satellite communication systems, airborne radar, and other applications. The low-loss properties of sapphire and semi-insulating GaAs substrates, combined with the excellent microwave performance of metal-semiconductor FET's (MESFET's), allows, for the first time, a truly monolithic approach to microwave integrated circuits. By monolithic we mean an approach wherein all passive and active circuit elements and interconnections are formed into the bulk, or onto the surface of the substrate by some deposition scheme, such as epitaxy, ion implantation, sputtering, evaporation, and other methods.

The importance of this development is that microwave applications such as airborne phased-array systems based on a large number of identical circuits and requiring small physical volume and/or light weight, may, finally, become cost effective.

The paper covers in some detail the design considerations that must be applied to monolithic microwave circuits in general, and to gallium arsenide circuits in particular. The important role being played by computer-aided design techniques is stressed. Numerous examples of monolithic circuits and components which illustrate the design principles are described. These provide a cross section of the world-wide effort in this field. A glimpse into the future prospects of monolithic microwave circuits is made.

I. INTRODUCTION

THE LAST two to three years have witnessed an intensive revival in the field of analog monolithic microwave integrated circuits (MMIC's), that is, microwave circuits deposited on a semiconductor substrate, or an insulating substrate with a semiconductor layer over it. In this paper, we shall address the design and technology considerations of monolithic microwave integrated circuits as well as the potential applications of these circuits to microwave systems, such as satellite communications and phased-array radar, as well as instrumentation.

It is important that the reader understand what we mean by the term "monolithic"circuit. By monolithic, we mean an approach wherein all active and passive circuit elements or components and interconnections are formed into the bulk, or onto the surface, of a semi-insulating substrate by some deposition scheme such as epitaxy, ion implantation, sputtering, evaporation, diffusion, or a combination of these processes and others.

It is essential that the full implication of this definition be understood, since it strikes at the very core of why one would want to design and fabricate a microwave monolithic circuit. The reasons are embedded in the following promising attributes of the monolithic approach:

1) low cost;
2) improved reliability and reproducibility;
3) small size and weight;
4) multioctave (broad-band) performance; and
5) circuit design flexibility and multifunction performance on a chip.

The importance of this development is that systems applications based on a large number of identical components, for instance, space-borne phase-array radars planned for the future which require lightweight and reliable, low-cost transmit–receive modules, may finally become cost effective. One might consider this type of application as the microwave system analog of the computer (which spurred the growth of the silicon digital monolithic circuit market), since both require a large number of identical circuits.

Maximum cost effectiveness, as well as improved reliability, derives in part from the fact that wire bonding is eliminated in MMIC's, at least within the chip itself, and is relegated to less critical and fewer locations at the periphery of the chip. Wire bonds have always been a serious factor in reliability and reproducibility. Furthermore, wire bonding, being labor intensive, is not an insignificant factor in the cost of a circuit.

Small size and volume, and their corollary, light weight, are intrinsic properties of the monolithic approach. Small size allows batch processing of hundreds of circuits per wafer of substrate. Since the essence of batch processing is that the cost of fabrication is determined by the cost of processing the entire wafer, it follows that the processing cost per chip is proportional to the area of the chip. Thus, the higher the circuit count per wafer, the lower the circuit cost.

The elimination of wire bonding and the embedding of active components within a printed circuit eliminate many of the undesired parasitics which limit the broad-band performance of circuits employing packaged discrete devices. The monolithic approach will certainly ease the difficulty of attaining multioctave performance. Furthermore, such broad-banding approaches as distributed

Manuscript received January 16, 1981. This work summarizes the lecture given by the National Lecturer throughout the United States, Canada, and Europe, during 1980–1981.

The author is with the Research Division, Raytheon Company, Waltham, MA 02254.

0018-9480/81/0600-0513$00.75 ©1981 IEEE

amplifier stages, heretofore shunned as too wasteful of active elements, will now become feasible, because a cost penalty will not accrue from the prolific use of low-gain stages, and because the unavoidable parasitics associated with the active devices will be incorporated in the propagating circuit and rendered less harmful.

The small circuit size intrinsic to the monolithic approach will enable circuit integration on a chip level, ranging from the lowest degree of complexity such as an oscillator, mixer, or amplifier, to a next higher "functional block" level, for example a receiver front end or a phase shifter. A still higher level of circuit complexity, for example, a transmit–receive module, will be integrated, most likely in multichip form.

So far we have discussed only the virtues of the monolithic approach. Now let us consider some of its disadvantages and problem areas. These are principally the following:

1) unfavorable device/chip area ratio;
2) circuit tuning (tweaking) impractical;
3) trouble-shooting (debugging) difficult;
4) suppression of undesired RF coupling (crosstalk), a possible problem; and
5) difficulty of integrating high power sources (IMPATT's)

The first item refers to the fact that only a small fraction of the chip area is occupied by devices, hence the high processing cost and lower yield associated with active device fabrication is unavoidably applied to the larger area occupied by the circuitry. A corollary of this is that the lower yield processes of device fabrication dominate the overall chip yield. Although these problems diminish as the chip size becomes smaller, that is, for higher frequencies, they are absent in the hybrid approach where the circuit and device technologies are separated.

The second and third items are related and can be considered together. The small chip sizes characteristic of the monolithic approach make it virtually impossible to tune ("tweak") and troubleshoot circuits. Indeed, to want to do so would violate one of the precepts of this approach, namely, to reduce costs by minimizing all labor-intensive steps. What then can be done about these very real problems?

First, it is necessary to minimize the need for tweaking. This can be done by adopting a design philosophy which leads to circuits that are insensitive to manufacturing tolerances in the active devices and physical dimensions of the circuit components. This will be a difficult compromise to accept on the part of the circuit designer, who expects the ultimate in performance from each active device by circuit tuning. However, here computer-aided design (CAD) techniques come to the rescue. Not only will CAD techniques play a major, if not mandatory, role in monolithic circuit design, they will also be used to assess the effect of tolerances on circuit performance during the design phase—and rather easily. CAD program for doing this

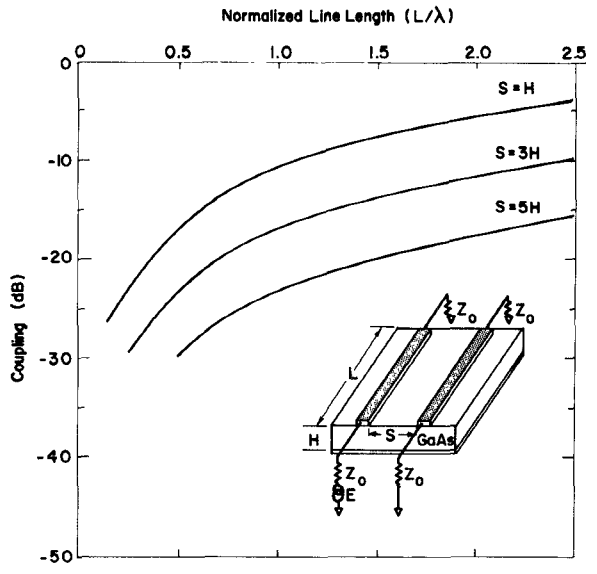


Fig. 1. Calculated coupling between adjacent parallel microstrip lines as a function of spacing and frequency.

reside on many internal computer systems and are also available commercially [4].

The use of CAD also helps alleviate the problem of troubleshooting a working circuit. Until microwave probes suitable for monolithic circuits become practical [19], troubleshooting must be based on terminal RF measurements of the circuit, usually the input and output ports. If a certain component is suspected of being faulty, it is a simple matter of building this defect into the CAD data file and comparing the resultant calculated circuit response with that measured. This can be done for a series of suspected faults, and convergence to the true fault can be achieved rather expeditiously.

The potential problem of undesirable RF coupling within the circuit is real because of the small chip sizes involved. To illustrate this point, Fig. 1 is a theoretical calculation of the coupling between two parallel microstrip lines on a GaAs substrate, one of which is excited by a generator. Both lines are matched at either end. Shown is the fraction of power coupled from the excited line to the adjacent line as a function of line length and line spacing. It is obvious that the coupling can become unacceptably high for long line lengths approaching a wavelength or more. Even for short lines, of the order of a quarter-wavelength or less, a feedback problem may exist if, say, a high-gain amplifier exists in one of the lines. In practice, line spacings of the order of three substrate thicknesses or more ($S>3H$) have been found adequate in most cases. This proximity "rule" plays a major role in determining the chip area and, hence, the chip cost. This restriction on circuit packing density, somewhat unique to MMIC's, can be alleviated measurably if direct-coupled circuitry is used, that is, if no distributed or lumped componentry is involved. We shall see examples of this approach later.

Turning to the fifth item, though both low-noise and power FET circuitry can easily be integrated on the same

chip, where very high powers, more precisely, power densities are involved, the monolithic approach may face some fundamental limitations. These limitations are associated with the need for special means of removing heat from the device. A case in point is the diamond heatsink used with millimeter-wave IMPATT diodes. Though it would be desirable to integrate avalanche diode sources in monolithic circuits for millimeter wave applications, the high-power densities involved cannot be handled by heat transfer through the chip. This is not a problem with power FET's, but of course, FET's cannot deliver the powers available from IMPATT's. Integration of high power sources in monolithic circuits is a problem that, as yet, has not been addressed.

Even for FET power amplifiers, tradeoffs must be made between good thermal performance and good RF design. For example, to minimize the thermal resistance through the substrate, it is desirable to use as thin a wafer as practical. However, a thin wafer increases the circuit skin effect losses, and hence the attenuation. Furthermore, since heat-sinking requires metallization of the chip bottomside, additional parasitic capacitance to ground is introduced and corrections must be made to planar inductors to account for "image" currents in the ground plane.

Despite these limitations on power, it is possible that with on-chip power combining techniques applied to FET's which are thermally isolated from each other [17], power outputs of the order of 10-W CW or so may be realizable from a single chip at the lower microwave frequencies, that is, at *X*-band.

II. MMIC's—A Brief History

The concept of MMIC's is not new. Its origin goes back to 1964 to a government-funded program based on silicon technology, which had as its objective a transmit–receive module for an aircraft phased-array antenna. Unfortunately, the results were disappointing because of the inability of semi-insulating silicon to maintain its semi-insulating properties through the high-temperature diffusion processes. Thus, very lossy substrates resulted, which were unacceptable for microwave circuitry [12]. Because of these and other difficulties the attempt to form a monolithic circuit based on a semiconductor substrate lay dormant till 1968 when Mehal and Wacker [15] revived the approach by using semi-insulating gallium arsenide (GaAs) as the base material and Schottky barrier diodes and Gunn devices to fabricate a 94-GHz receiver front end. However, it was not until Plessey applied this approach to an *X*-band amplifier, based on the Schottky-gate field-effect transistor, or MESFET (MEtal-Semiconductor Field-Effect Transistor), as the key active element that the present intense activity began [16].

What brought on this revival? First, the rapid development of GaAs material technology, namely, epitaxy and ion implantation, and the speedy evolution of the GaAs FET based on the metal Schottky gate during the last decade led to high-frequency semiconductor device performance previously unattained. A few examples are high-efficiency and high-power amplifier performance through *Ku*-band, low-noise amplifiers, variable-gain dual-gate amplifiers, and FET mixers with gain. The dual-gate FET will play a major role in MMIC's because of its versatility as a linear amplifier whose gain can be controlled either digitally or in analog fashion. With dual-gate FET's, multiport electronically switchable RF gain channels are feasible. Second, resolution of many troublesome device reliability problems made FET's more attractive for systems applications. Third, recognition of the excellent microwave properties of semi-insulting GaAs (approaching that of alumina), removed the major objection of silicon. Fourth, hybrid circuits were becoming very complex and labor intensive because of the prolific use of wire bonds, and hence too costly. Fifth, the emergence of clearly defined and discernible systems applications for MMIC's became more apparent. Thus it was the confluence of all of these factors, and others, which stimulated the development of GaAs MMIC's within the last five years.

TABLE I
Some Properties of Semiconductors and Insulators

Property	GaAs	Silicon	Semi-insulating GaAs	Semi-insulating Silicon	Sapphire	Alumina
Dielectric Constant	12.9	11.7	12.9	11.7	11.6 (C-axis)	9.7
Density (gm/cc)	5 32	2.33	5.32	2.33	3 98	3.89
Thermal Cond. (watts/cm-°K)	0.46	1.45	0.46	1.45	0.46	0.37
Resistivity (ohm-cm)	---	---	10^7 - 10^9	10^3 - 10^5	$>10^{14}$	10^{11} - 10^{14}
Elec. Mobility (cm^2/v-sec.)	4300*	700*	---	---	---	---
Sat. Elec. Vel. (cm/sec.)	1.3×10^7	9×10^6	---	---	---	---

* At $10^{17}/cm^3$ doping

III. Silicon or Gallium Arsenide?

It is ironic that this revival based on GaAs technology has, in turn, restimulated the interest in silicon MMIC's—but now based on the silicon-on-sapphire (SOS) approach [13]. There are understandable reasons for this. First, the use of sapphire as a substrate eliminates the losses associated with semi-insulating silicon mentioned earlier. Second, silicon technology is an extremely well developed technology—much more so than GaAs. Third, the availability of the simpler MESFET technology, developed in GaAs, could now be used in place of the more complex bipolar technology, which, however, was still available should it be needed. Nevertheless, gallium arsenide has the "edge" for reasons to be discussed next.

Table I lists some of the pertinent physical and electrical properties of GaAs and silicon (n-type) in their insulating and semiconducting states, as well as that of sapphire and alumina. As is evident from this table, as a high-resistivity substrate, semi-insulating GaAs, sapphire, and alumina are, for all practical purposes, comparable. Also evident is that the carrier mobility of gallium arsenide is over six

times that of silicon. For this reason and others, GaAs MESFET's are operable at higher frequencies and powers than silicon MESFET's of equivalent dimensions. For example, silicon MESFET's based on 1-μm gate technology will be limited in operation to upper *S*-band at best, whereas GaAs MESFET's operate well at *X*-band and higher. Therefore, it is highly likely that the performance of 1-μm gate silicon MESFET's will be matched, and perhaps exceeded by that of 2-μm gate GaAs MESFET's at *S*-band. The near-future availability of much larger GaAs wafers, approaching 3.5 in in diameter [20], obtained by the Czochralski method, will overcome the size limitations imposed by the present 1-in wafers grown by the Bridgman method. The early success of direct-coupled FET analog circuitry [11], [21], which leads to high component density at *S*-band, will also help overcome wafer size limitations in GaAs. Finally, the proven success of gigahertz high-speed GaAs logic circuitry will allow, for the first time, complete integration of logic and analog microwave circuitry. This opens up the feasibility of high-speed signal processing on a chip.

We do not wish to imply that MMIC work based on SOS technology should be diminished; however, we believe its major role will be found in the range below 2 GHz, for example, in IF circuitry and other applications. In light of this conclusion, we shall direct the following discussion to the GaAs approach. However, much of what we shall say, as will be obvious to the reader, will also apply, with minor changes, to the SOS approach or to other approaches which may emerge in the future. Nevertheless, we maintain that before this decade is over, it will be GaAs monolithic integrated circuits that will exert the greatest influence on the way solid-state device circuitry is used in microwave systems.

IV. The Gallium Arsenide Approach

A cornerstone of the monolithic approach will be the availability of a highly reproducible device technology. This in turn is related, in part, to the control of the starting material, especially the active (semiconducting) layer.

Two general techniques are available for forming this layer on GaAs substrates, namely, epitaxy and ion implantation. Of the two approaches, the former at present is more widely used and developed. In this approach a doped single crystal semiconducting layer is deposited on a semi-insulating GaAs substrate, usually with an intervening high-resistivity epitaxial "buffer" layer to screen out diffusion of impurities from the substrate during the active layer growth. With ion implantation, dopant atoms are implanted directly into the surface of a semi-insulating GaAs substrate. This procedure requires a higher state of purity of the substrate—a problem at present.

Expitaxy does not have the control or flexibility associated with implantation. With implantation, more uniform conducting layers are possible over a large area—more uniform in doping level as well as in thickness. Furthermore, with implantation, selective doping is easy, that is, formation of different doping profiles in different parts of the wafer is easy to achieve, whereas with epitaxy it is difficult. The potential device reproducibility achievable with implantation is a definite advantage for it.

It should be added that implantation can also be used in conjunction with epitaxy. One such application is the isolation implant, wherein oxygen is implanted in the unused portions of the epitaxial layer to produce a high-resistivity region within the epitaxial layer onto which microwave circuitry may be situated. Thus a truly planar surface is maintained, since no mesa etching is required to remove the undesired regions of epitaxial layer. This also eliminates yield problems associated with metallization patterns extending over mesa steps.

It is likely that, once substrate purity reaches the necessary level for ion implantation (as it is approaching with unintentionally doped Czochralski-pulled crystals), ion implantation will supplant epitaxy as the preferred method for monolithic circuits.

The processing technology used for FET fabrication is also applicable to the monolithic circuit elements. The high degree of dimensional definition associated with FET photolithography is more than adequate for the circuit elements.

V. General Design Considerations

We turn now to a discussion of the design considerations for MMIC's.

A. Constraints on Chip Size

Present substrate sizes corresponding to that of GaAs boules are approximately 1 in in diameter, though larger boules approaching 3 in in diameter are now being grown by the Czochralski method. Given the expected limits on substrate size, it is instructive to estimate the circuit count/wafer achievable as a function of frequency, since the processing cost per circuit is inversely proportional to this density.

We assume that the maximum linear dimension per circuit will fall between $\lambda_g/10$ and $\lambda_g/4$, where λ_g is the wavelength of the propagation mode (microstrip-coplanar, etc.) in GaAs. The lower limit takes into account the approximate maximum size of lumped elements; the upper limit, the typical maximum size of distributed elements. It seems reasonable to assume that in the vicinity of 10 to 20 GHz some distributed elements of the order of $\lambda_g/4$ (for example, hybrid and branch line couplers) will be used. Therefore, above this frequency range, linear circuit dimensions of the order of $\lambda_g/4$ will be the rule. Let us choose 16 GHz as the demarcation frequency. We then postulate a "linear" admixture of lumped- and distributed-element weighting so that we obtain $\lambda_g/10$ at 1 GHz and $\lambda_g/4$ at 16 GHz as the probable linear dimension of a circuit function "chip."

Fig. 2 is a plot of the approximate density of these circuits as a function of frequency for two sizes of wafer. (The 2-in square wafer corresponds to a 3-in diameter

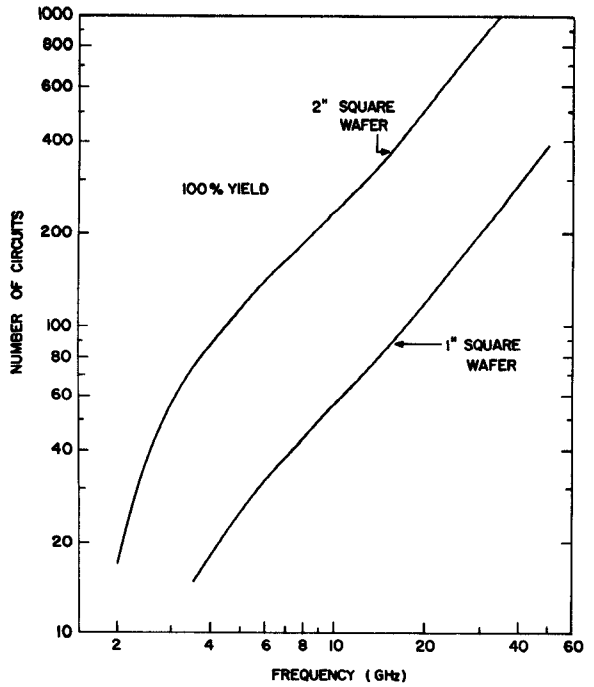


Fig. 2. Estimated number of circuits per wafer taking dicing and edge waste into account.

wafer.) The circuit density estimates take into account edge and cutting waste, but not "proximity effects," wafer yield, and other factors which will reduce these numbers.

A yield factor is associated with each fabrication step, the overall yield being the product of the individual yield factors. Thus, since active devices generally require the most processing steps (about 40 for an FET), the overall yield is determined by the device processing technology. The "proximity effect," that is, the RF coupling problem mentioned earlier, will put stringent limitations on how closely packed the signal lines may be, and hence how much circuitry can be compressed into the chip area, which is fixed by wavelength or lumped-element dimensions as just described.

The circuit count estimate must be modified for very low microwave frequencies (below *C*-band) if active components such as FET's are used to simulate resistors and capacitors and if inductors are dispensed with because tuning is not necessary. In this so-called direct-coupled design, packing densities approaching those normally associated with digital circuitry is possible [21], that is, much higher than that indicated in Fig. 2. However, it must be cautioned that this circuit approach is not suitable for all monolithic applications, for example, high-efficiency power amplifiers or low-noise circuits. The reason is that the use of active (FET) devices as resistive elements in the gate and drain circuits introduces high dc power dissipation and mismatch, as well as additional noise [11].

It is appropriate at this time to point out that the size advantages of GaAs MMIC's will be lost if proper packaging techniques are not used. Perhaps the efficient techniques adopted for low-frequency and digital circuitry can

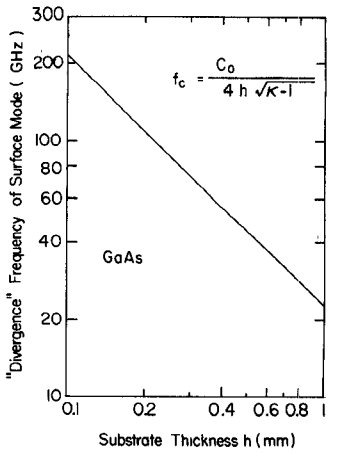


Fig. 3. Frequency of onset of lowest order TE surface wave on a GaAs substrate as a function of substrate thickness.

be suitably modified for microwave applications. Much thought must be devoted to this very important problem.

B. Constraints on Wafer Thickness

So far we have discussed requirements on the substrate area. There also are constraints imposed on the substrate thickness. Some of these are:

1) volume of material used;
2) fragility of wafer;
3) thermal resistance;
4) propagation losses;
5) higher order mode propagation;
6) impedance-linewidth considerations; and
7) thickness tolerance versus impedance tolerance.

Obviously, to keep material costs down one wishes to use as thin a substrate as can be handled without compromising the fragility. Thermal considerations also require the thinnest wafer possible. On the other hand, a thin wafer emphasizes the effect of the ground plane. For example, propagation losses increase inversely with substrate thickness in the case of microstrip. Furthermore, the *Q*-factor and inductance of thin-film inductors decrease with decreasing substrate thickness. In contrast, undesired higher-order surface mode excitation is inhibited for thinner substrates.

Fig. 3 is a graph of the frequency denoting the onset of the lowest order (TE) surface mode as a function of substrate thickness. For example, for a substrate thickness of 0.1 mm (4 mils) the "safe" operating frequency range is below 200 GHz. It appears that, for presently contemplated circuit applications, surface mode propagation is not a limiting factor in the choice of substrate thickness. The linewidth dimensions for a given impedance level of some propagation modes, such as microstrip, are proportional to substrate thickness. Therefore, thicker substrates alleviate the effect of thickness and linewidth tolerances.

The point being made here is that the choice of substrate thickness is a tradeoff of the factors listed above, being strongly dependent on the frequency of operation and the

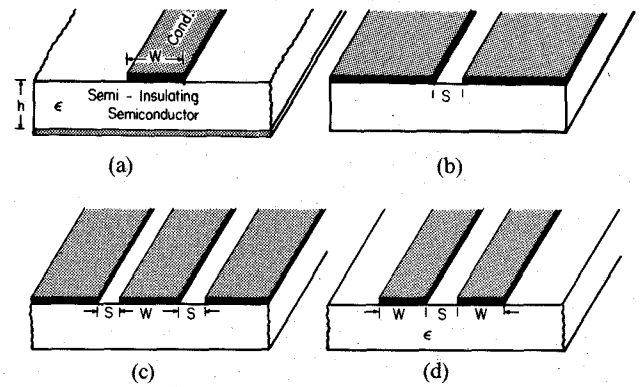


Fig. 4. Four candidate propagation modes for monolithic circuits. (a) Microstrip (MS). (b) Slot line (SL). (c) Coplanar waveguide. (d) Coplanar strips (CS).

power dissipation of the circuit. It is true that perhaps the most important of the factors is the thermal consideration. We believe that in the frequency range up to 30 GHz a substrate thickness of the order of 0.1 mm to 0.15 mm is appropriate for power amplifier circuits, with thicknesses up to 0.6 mm tolerable for low-noise amplifiers and similar circuits, provided a satisfactory means of dicing the thicker wafers can be found.

C. Propagation Modes

At microwave frequencies, the interconnections between elements on a high dielectric constant substrate such as GaAs, where considerable wavelength reduction occurs, must be treated as waveguiding structures. On a planar substrate, four basic modes of propagation are available, as illustrated in Fig. 4. The first mode (Fig. 4(a)) is microstrip (MS), which requires a bottomside ground metallizaton. Its "inverse," slot line (SL), is shown in Fig. 4(b). The third mode is the coplanar waveguide (CPW) shown in Fig. 4(c); it consists of a central "hot" conductor separated by a slot from two adjacent ground planes. Its "inverse," the coplanar stripline (CS), is illustrated in Fig. 4(d); here, one of the two conductors is a ground plane. Both the coplanar waveguide and coplanar strips are generally considered to be on infinitely thick substrates. Of course, this condition cannot be met. We shall see the implication of this later.

Of the four modes, only the slot line is not TEM-like. For this reason, and because it uses valuable "topside" area, we do not expect slot line to be a viable candidate for monolithic circuits, except possibly in special cases.

The principal losses of microstrip and the coplanar modes consist of ohmic losses. Since the coplanar structures are, in essence, "edge-coupled" devices, with high concentration of charge and current near the strip edges, the losses tend to be somewhat higher than for microstrip, as verified by experiment [5].

The lack of a ground plane on the topside surface of the microstrip structure is a considerable disadvantage when shunt element connections to the hot conductor are required. However, as we shall see later, there are ways to overcome this disadvantage. Table II summarizes, in a qualitative way, the features of the four modes of propagation illustrated in Fig. 4.

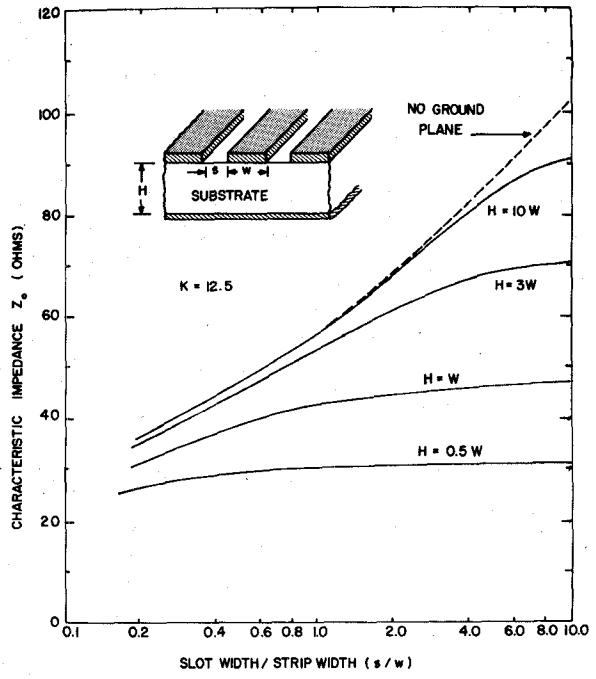


Fig. 5. Effect of ground plane on characteristic impedance of a coplanar waveguide.

TABLE II
QUALITATIVE COMPARISON OF PROPAGATION MODES

	MICROSTRIP	COPLANAR WAVEGUIDE	COPLANAR STRIPS	SLOT LINE
Attenuation Loss	low	medium	medium	high
Dispersion	low	medium	medium	high
Impedance Range (ohms)	10-100	25-125*	40-250*	high
Connect Shunt Elements	diff.	easy	easy	easy
Connect Series Elements	easy	easy	easy	diff.

*Infinitely thick substrate

The impedance range achievable with CPW and CS is somewhat greater than for MS, particularly at the higher end of the impedance scale, provided an infinitely thick substrate is assumed for CPW and CS. This range is reduced considerably when practical substrate thicknesses are used and the bottomside of the chip is metallized. Fig. 5 shows how the high impedance end of the scale is lowered when substrates of the order of 0.1–0.25 mm thick are mounted on a metal base (for heat-sinking purposes). The considerable reduction in Z_0 makes the design of monolithic circuitry with CPW nearly as dependent on substrate thickness as with MS, at least at the high end of the impedance scale.

Microstrip has its own unique restriction on the realizable impedance range. This is dictated by technology considerations. The limitation stems from the fact that for MS

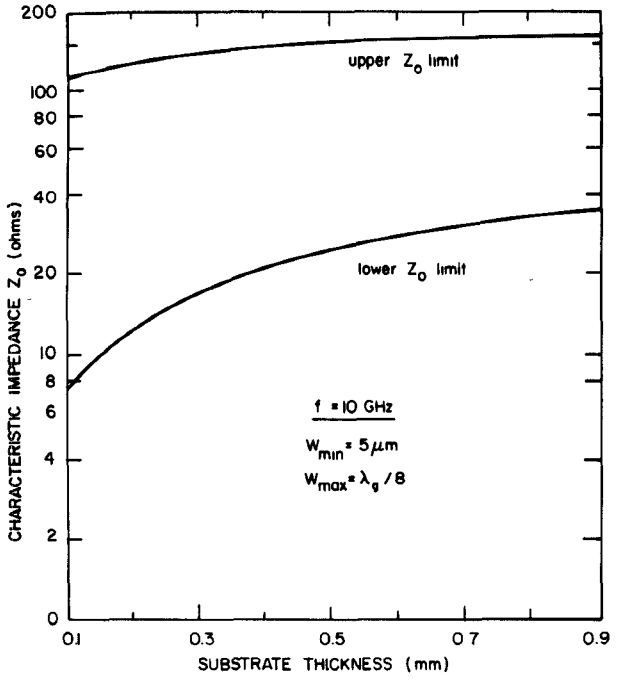


Fig. 6. Range of characteristic impedance of microstrip on GaAs substrate as a function of substrate thickness.

the characteristic impedance Z_0 is a function of the ratio W/H (see Fig. 4). The highest achievable impedance is determined by the smallest linewidth W that can be realized with acceptable integrity over a long length, say, a quarter of a wavelength. Our experience is that a minimum linewidth of 5 μm is reasonable. With this restriction, and an additional limit imposed on the maximum linewidth to be well below a quarter-wavelength, say, one-eighth-wavelength, the realizable range of characteristic impedance as a function of substrate thickness and frequency is constrained within the range indicated by Fig. 6

It is evident that the usable impedance range for a 0.1-mm thick substrate is approximately 10–100 Ω, and somewhat less for thicker substrates and higher frequencies. For higher frequencies, the lower curve moves "up." This limited impedance range is a severe restriction in the design of matching networks, a problem not faced in the hybrid approach.

Weighing all of these factors, we believe that of the four candidate modes, MS and CPW are the most suitable for GaAs monolithic circuits, with preference toward MS. Indeed, there will be instances where both modes may be used on the same chip to achieve some special advantage. The transition from one mode to the other is trivial. Most of the examples of MMIC's to be described later are based on MS.

Fig. 6(a) is a graph of the wavelength of a 50-Ω MS line on GaAS as a function of frequency, with dispersion neglected. The wavelength of CPW is similar. Fig. 7(b) illustrates the attenuation of MS as a function of characteristic impedance and substrate thickness at 10 GHz. The loss in decibels per centimeter increases as the square root

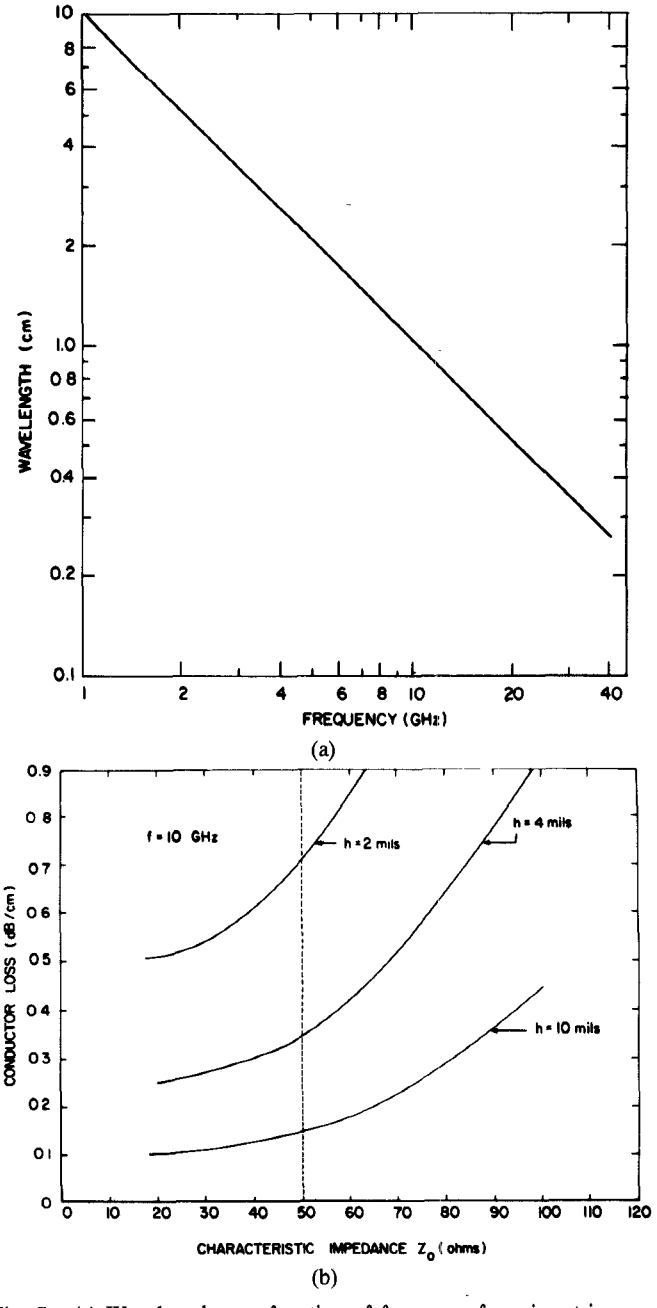


Fig. 7. (a) Wavelength as a function of frequency for microstrip on a GaAs substrate for $h=0.1$ mm. (b) Conductor loss of microstrip on a GaAs substrate as a function of substrate thickness and characteristic impedance for $f=10$ GHz.

of frequency. The loss per wavelength, on the other hand, decreases as the square root of frequency. Note the inverse dependence of loss on substrate thickness.

D. Low Inductance Grounds and Crossovers

Microstrip and coplanar waveguide are adequate for interconnections that do not require conductor crossovers or that are not to contact the bottomside ground metallization. Often, however, such connections are needed. In particular, with MS, which does not have any topside ground planes, some means of achieving a low-inductance ground is essential.

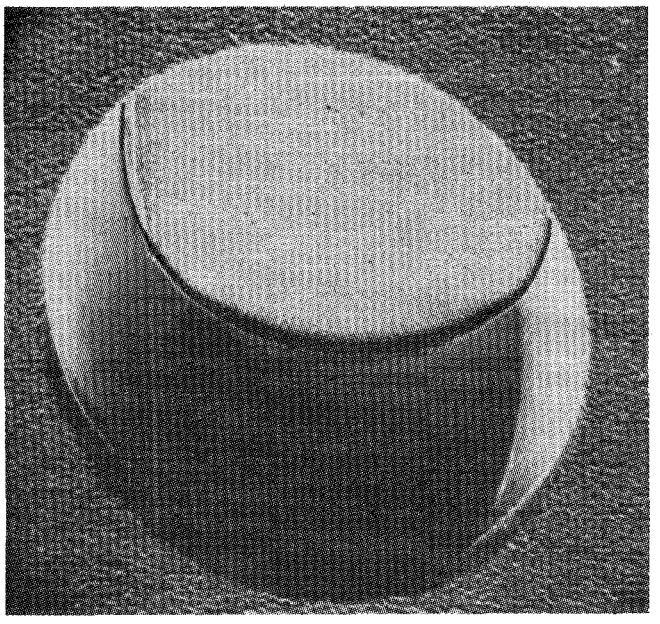


Fig. 8. 50-μm diameter "via" hole etched in a GaAs wafer.

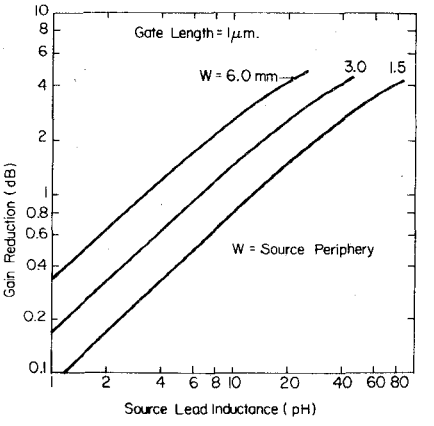


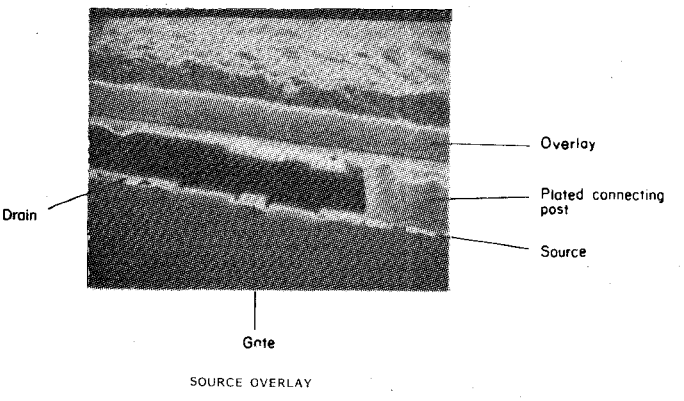
Fig. 9. Calculated gain reduction of a GaAs power FET as a function of source lead inductance.

Two general methods of grounding are available: 1) the "wrap-around" ground; and 2) the "via" hole ground. The former requires a topside metallization pattern near the periphery of the chip which can be connected to the chip ground. The "via" hole technique, on the other hand, allows placement of grounds through the substrate where desired. Holes are chemically milled through the substrate until the top metallization pattern is reached. These holes are subsequently metallized at the same time as the ground plane to provide continuity between this plane and the desired topside pad. A microphotograph of such a "via" hole etched through a test wafer of 50-μm thickness is shown in Fig. 8. The hole diameter, in this case, is only 50 μm, much smaller than those normally used in monolithic circuits. The estimated inductance of a via hole is approximately 40–60 pH/mm of substrate thickness. Examples of circuits using both grounding techniques will be described later.

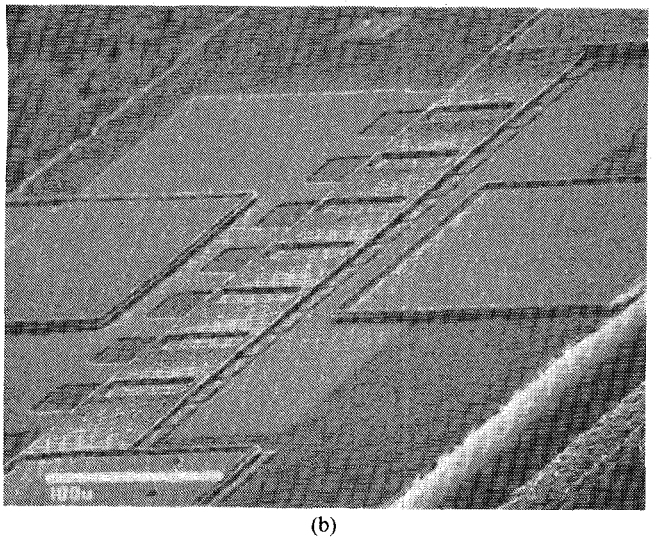
Low inductance grounds are especially important in source leads of power FET's. An inductance in the source lead manifests itself as resistive loss in the gate circuit, and hence a reduction in power gain. To illustrate this, Fig. 9 is a graph of the calculated gain reduction as a function of source lead inductance for an unconditionally stable power FET, corresponding to a power output of 1, 2, and 4 W (W=1.5, 3.0, and 6.0 mm).

The second interconnect problem arises when it is necessary to connect the individual cells of a power FET without resorting to wire bonds. A requirement is that these interconnects also have a low inductance. Here the so-called "air-bridge" crossover is useful. This crossover consists of a deposited strap which crosses over one or more conductors with an air gap in between for low capacitive coupling.

Fig. 10(a) is a cross-sectional view of a source crossover which interconnects two adjacent source pads of a power



(a)



(b)

Fig. 10. (a) SEM microphotograph of a segment of source overlay (airbridge) of a power FET showing gate and drain contacts. (b) Top view of a GaAs power FET showing an air-bridge overlay connecting all source pads.

FET. The air gap is approximately 4 μm. Clearly shown is the 1-μm gate and the larger drain pad underneath the crossover. Fig. 10(b) is a closeup, angular view of a power FET which employs an airbridge (overlay) interconnect bus.

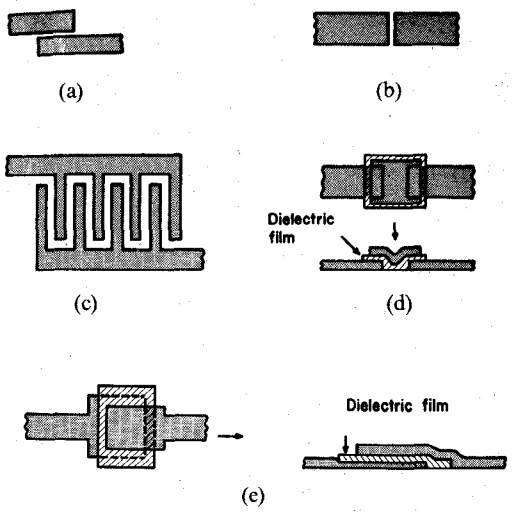


Fig. 11. Some planar capacitor designs. (a) Broadside coupled. (b) End coupled. (c) Interdigitated. (d) End-coupled overlay. (e) Overlay.

It is evident that the airbridge technology allows one to interconnect all cells without recourse to wire-bonding and therefore nicely satisfies the criterion for a monolithic circuit. Airbridge interconnects, of course, are also useful for microstrip and other crossovers. A good example is a planar spiral inductor, which requires a contact to the inner terminal.

E. Thin-Film Components

A flexible monolithic circuit design philosophy must include both lumped elements (dimensions <0.1 wavelength) and distributed elements, that is, elements composed of sections of transmission line. Lumped elements, R's, C's, and L's, are also useful for the RF circuitry, and in some cases mandatory, as for example, in thin-film resistive terminations for couplers. Lumped thin-film capacitors are absolutely essential for bias bypass applications, because of the large capacitance values required. Planar inductors can be extremely useful for matching purposes, especially at the lower end of the microwave band where stub inductors are very large, physically.

The choice of lumped or distributed elements depends on the frequency of operation. Lumped elements are suitable through X-band up to, perhaps, 20 GHz. It is likely, however, that beyond this frequency range, distributed elements will be preferred. It is difficult to realize a truly lumped element, even at the lower frequencies, because of the parasitics to ground associated with thin substrates. In this section we shall review the design principles of planar lumped elements.

1) Planar Capacitors: There are a variety of planar capacitors suitable for monolithic circuits—those achieved with a single metallization scheme, and those using a two-level metallization technology in conjunction with dielectric films. Some possible geometries for planar capacitors are shown in Fig. 11. The first three, which use no dielectric film and depend on electrostatic coupling via the substrate, generally are suitable for applications where low values of capacitance are required (less than 1.0 pF) for instance, high-impedance matching circuits. The last three geometries, the so-called overlay structures which use dielectric films, are suitable for low-impedance (power) circuitry and bypass and blocking applications. Capacitance values as high as 10 to 30 pF are achievable in small areas.

Two sources of loss are prevalent in planar capacitors, conductor losses in the metallization, and dielectric losses of the films, if used. Since the first three schemes illustrated in Fig. 11 are edge-coupled capacitors, high charge and current concentrations near the edges tend to limit the Q-factors. At X-band, typical Q-factors measured to date have been in the range of 50, despite the fact that no dielectric losses are present. The last three geometries distribute the current more uniformly throughout the metal plates because of the intervening film. However, even here, Q-factors only in the range of 50–100 are typical at X-band (10 GHz) because of dielectric film losses. Let us turn now to a more detailed analysis of the overlay structures depicted in Fig. 11, in particular the structure in Fig. 11(e).

First, we review briefly some general requirements of dielectric films for the overlay geometry. Some properties of dielectric films of importance are 1) dielectric constant, 2) capacitance/area, 3) microwave losses, 4) breakdown field, 5) temperature coefficient, 6) film integrity (pinhole density, stability over time), and 7) method and temperature of deposition. This last requirement is obviously important, because the technology used for film deposition must be compatible with the technology used for the active device (FET). Dielectric films which easily satisfy this criterion are SiO_x and Si_3N_4.

Some useful figures of merit for dielectric films are the capacitance-breakdown voltage product

$$F_{cV} = \left(\frac{C}{A}\right) V_b \tag{1a}$$

$$= \kappa \epsilon_0 E_b \tag{1b}$$

$$\cong (8-30) \times 10^3 \text{ pF} \cdot \text{V/mm}^2$$

and the capacitance-dielectric Q-factor product

$$F_{cq} = \left(\frac{C}{A}\right) Q_d \tag{1c}$$

$$= \frac{(C/A)}{\tan \delta_d} \tag{1d}$$

where C/A is the capacitance per unit area, V_b is the breakdown voltage, E_b is the corresponding breakdown field, κ is the dielectric constant, and $\tan \delta_d$ is the dielectric loss tangent. Breakdown fields of the order of 1–2 MV/cm are typical of good dielectric films. Dielectric constants are in the order of 4–20. Loss tangents can range from 10^{-1} to 10^{-3}. It is desirable to have as high figures of merit as possible. Table III is a list of candidate films and their properties.

We return, now, to the overlay structure of Fig. 11(e). A closeup perspective view is shown in Fig. 12. Taking into

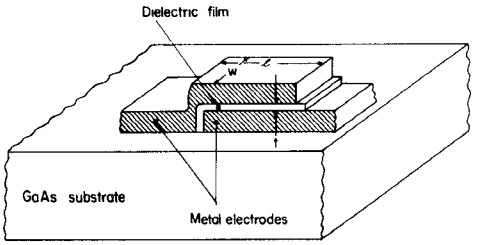


Fig. 12. Perspective of an overlay thin-film capacitor.

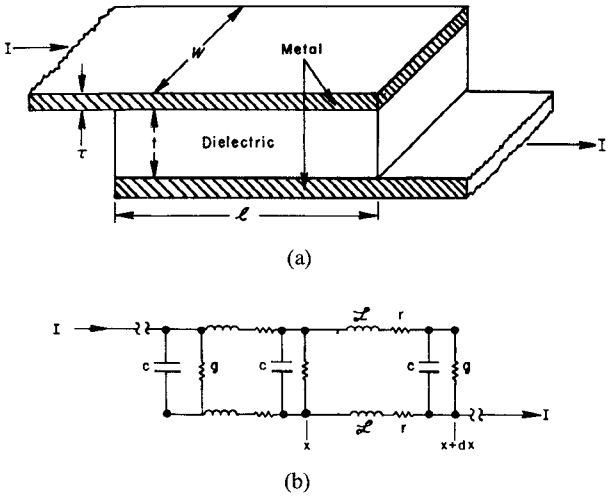


Fig. 13. Diagrams relevant to analysis of impedance of a thin-film capacitor. (a) Thin-film capacitor. (b) Circuit model.

TABLE III
PROPERTIES OF SOME CANDIDATE DIELECTRIC FILMS

DIELECTRIC	κ	TCC (ppm/°C)	C/A* (pF/mm²)	(C/A)·Q_d	(C/A)·V_b	COMMENTS
SiO	4.5-6 8	100-500	300	low	medium	Evaporated
SiO_2	4-5	50	200	medium	medium	Evaporated, CVD, or Sputtered
Si_3N_4	6-7	25-35	300	high	high	Sputtered or CVD
Ta_2O_5	20-25	0-200	1100	medium	high	Sputtered and Anodized
Al_2O_3	6-9	300-500	400	high	high	CVD, anodic oxidation, sputtering
Schottky-Barrier Junction	12 9	--	550	very low	high	Evaporated Metal on GaAs
Polyimide	3-4 5	-500	35	high	--	Spun and Cured Organic Film

*Film thickness assumed = 2000 Å, except for polyimide, 10,000 Å

account the longitudinal current paths in the metal contacts, one may analyze this device as a lossy transmission line as indicated in Fig. 13. In Fig. 13(b), $\mathcal{L}$ and r represent the inductance and resistance per unit length of the metal plates, and c and g denote the capacitance and conductance per unit length of the dielectric film. The relation between g and c is determined by the loss tangent, $g=\omega c \tan\delta_d$. The series resistance in the plates is determined by the skin resistance if the metal thickness exceeds the skin depth, or the bulk metal resistance if the reverse is true. Usually, the bottom metal layer is evaporated only, and hence is about 0.5 μm thick, which may be less than the skin depth. The top metal is normally built up to a thickness of several micrometers or more by plating.

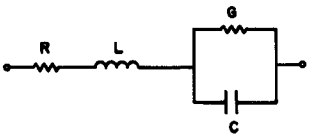


Fig. 14. Equivalent circuit of a thin-film capacitor. $R=2/3\ rl$. $C=cl$. $G=\omega c \tan\delta$. $L=\mathcal{L}l$. r = resistance/length (electrodes). c = capacitance/length. $\mathcal{L}$ = inductance/length (electrodes). $\tan\delta$ = loss tangent of dielectric film.

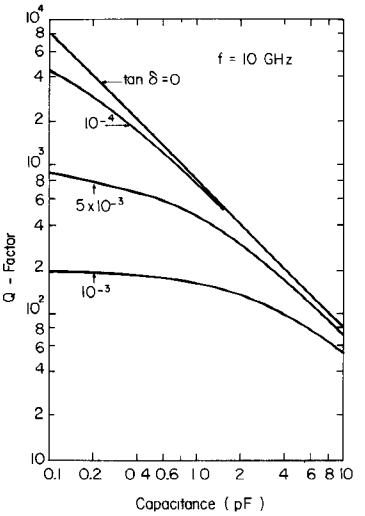


Fig. 15. Quality factor of a square thin-film capacitor as a function of capacitance and dielectric loss tangent for $f=10$ GHz.

For a well-designed capacitor, the longitudinal and transverse dimensions are small compared with a wavelength in the dielectric film. In this case, a good approximation to the capacitor is the equivalent circuit shown in Fig. 14. When the skin loss condition prevails, the Q-factor corresponding to these losses is given by the expression

$$Q_c = \frac{3}{2\omega R_s (C/A) l^2} \tag{2}$$

where R_s is the surface skin resistivity and l is the electrode length (see Fig. 12). Note the strong dependence on electrode length. This arises because of the longitudinal current path in each electrode. Note that if one electrode, say the bottom plate, is very thin, Q_c is decreased.

The dielectric Q-factor is $Q_d = 1/\tan\delta$, and the total Q-factor is given by the relation $Q^{-1} = Q_d^{-1} + Q_c^{-1}$. Fig. 15 is a graph of the calculated Q-factor as a function of capacitance for various loss trangents. Note that for a 1-pF capacitor, and no dielectric losses, the predicted Q-factor is approximately 800! Yet, experimentally, values more like one-tenth of this are obtained, suggesting that dielectric films are extremely lossy—much more so than their bulk counterparts. No satisfactory explanation for this observation has yet been advanced.

2) Planar Inductors: Planar inductors for monolithic circuits can be realized in a number of configurations, all achieved with a single-layer metallization scheme. Fig. 16 illustrates various geometries that can be used for thin-film inductors. Aside from the high-impedance line section, all

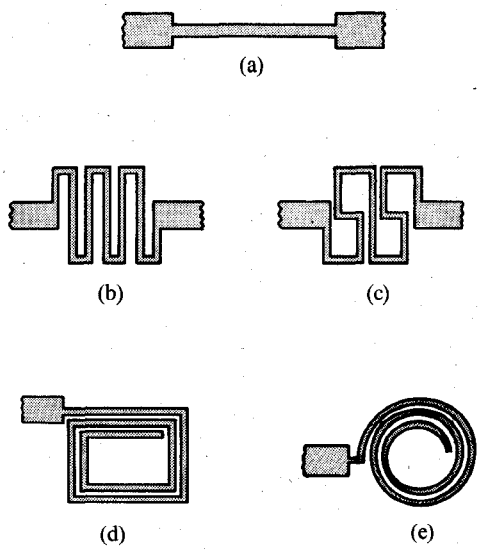
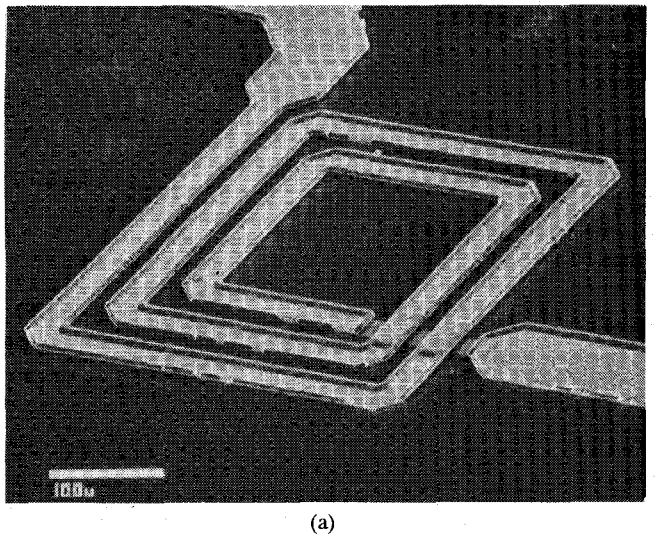
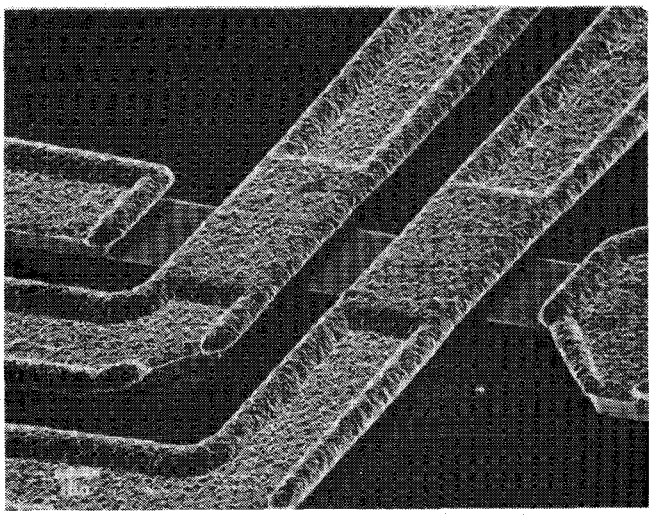


Fig. 16. Some planar inductor configurations. (a) High-impedance line section. (b) Meander line. (c) *S*-line. (d) Square spiral. (e) Circular spiral.

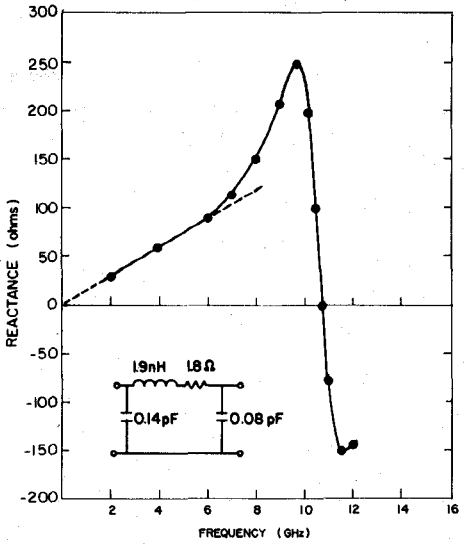


(a)



(b)

Fig. 17. SEM photographs of a thin-film square spiral inductor showing air-bridge crossovers.



MEASURED REACTANCE OF A TEN-SEGMENT SQUARE-SPIRAL GROUNDED INDUCTOR ON A 0.1 MM THICK SI-GAAS SUBSTRATE

Fig. 18. Measured reactance of a ten-segment square spiral inductor on a 0.1-mm thick GaAs substrate (equivalent circuit shown in inset).

of the structures depend on mutual coupling between the various line segments to achieve a high inductance in a small area. In any multisegment design, one must insure that the total line length is a small fraction of a wavelength, otherwise the conductor cannot be treated as "lumped." Unfortunately, this latter condition is not often satisfied. Fig. 17 is a SEM photograph of a multisegment square–spiral inductor. Note the crossover connections.

When thin substrates are used, corrections must be made to the calculated inductance to account for the ground plane. These corrections are always in the direction to reduce the inductance, and are typically in the range of 15 percent, though for large-area inductors, the reduction can be as high as 30 percent.

Typical inductance values for monolithic circuits fall in the range from 0.5 to 10 nH. The higher values are difficult to achieve in strictly lumped form because of intersegment fringing capacitance. A more serious problem is that of shunt capacitance to ground, especially in the case of the microstrip format. This capacitance to ground can become important enough to require its inclusion in determining the performance of the inductor.

An illustration of the serious effect of capacitance to ground is demonstrated by the data of Fig. 18. This is a graph of the measured reactance of a ten-segment square spiral inductor as a function of frequency. The inductor is approximately 0.4 mm square, consisting of segments 1 mil wide, separated by 1 mil (see Fig. 17). The inductance, as designed, was nominally 1.9 nH. Note that above 10 GHz the reactance becomes capacitive! The equivalent circuit, as deduced from two-port *S*-parameters, is shown in the inset. The substrate thickness was 0.1 mm.

Of course, the inductor is usable, provided all of the parasitics indicated in Fig. 18 are taken into account.

Unfortunately, these parasitics are not known in advance. Thus, in a computer-aided approach, corrections to the circuit in which the inductor appears must be made in later iterations. This can become a costly procedure. It is often more sensible to use an inductive transmission line segment whose electrical behavior is known in advance.

Some of the skin losses in the inductor reside in the ground plane (assuming a metallized bottom side) and increase as the ground plane approaches the film inductor (not unlike shielding losses). However, the dependence on substrate thickness is mild, since most of the losses reside in the film turns, because of their small cross section.

In practice, inductor Q-factors of the order of 50 are observed at X-band, with higher values at higher frequencies. There appears to be no way to improve the Q-factor significantly, because of the highly unfavorable ratio of metal surface area to dielectric volume.

Somewhat higher Q-factors are achievable with microstrip resonant stub sections. These are more properly considered as distributed inductors, or more correctly, as distributed resonant elements. Three sources of loss are important here, skin losses, dielectric losses, and radiation losses. For microstrip stubs, the skin losses are those associated with microstrip, as are the dielectric losses. Skin losses vary inversely with the substrate thickness, and increase as the line impedance increases. Assuming negligible dielectric losses, one may show that the conductor Q-factor for a quarter-wave open circuit stub is given by

$$Q_c = \frac{27.3}{(\alpha\lambda g)} \tag{3}$$

where $(\alpha\lambda g)$ is the loss in the line section in decibels per wavelength. Since $(\alpha\lambda g)$ decreases as $f^{-1/2}$, Q_c increases as the square root of frequency, as for thin-film inductors. On the other hand, radiation losses from the open circuit end vary as [8]

$$Q_r = \frac{R}{(fh)^2} \tag{4}$$

where h is the substrate thickness and R is a function of w/h and the dielectric constant of the substrate. (The radiation factor R is considerably larger for a quarter-wave stub grounded at its far end.) Note that the radiation Q decreases as the square of the frequency and the substrate thickness h. Thus any attempt to increase the conductor Q-factor by increasing the frequency and substrate thickness is eventually overcompensated by the decrease in radiation Q. Fig. 19 illustrates this fact for practical substrate thicknesses. Thus, above X-band, open-circuit stub resonators are dominated by radiation losses, unless the substrate is less than 0.25 mm thick. This radiation also can cause coupling to adjacent circuits. A way to overcome both problems is to use a ring resonator.

The choice then as to whether reactive lumped elements or distributed elements should be used must be considered for each individual application. If high-Q narrow-band

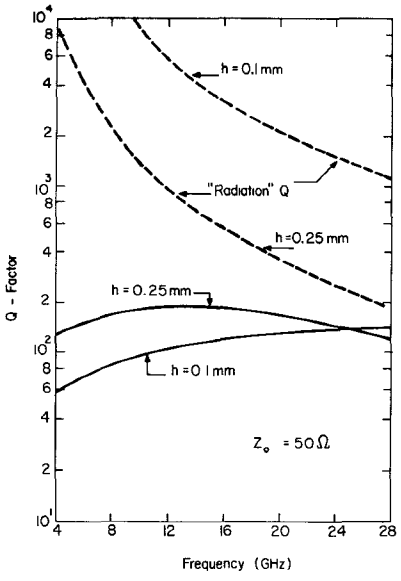


Fig. 19. Quality factor of a quarter-wave microstrip resonator on a GaAs substrate.

circuits are to be realized, distributed elements are recommended, provided space is available. On the other hand, broad-band circuits are probably easier to design with lumped elements, though even here synthesis techniques based on transmission line stubs are now available. Some circuits are more readily designed with distributed elements. Examples are four-port couplers and power combiners/dividers.

3) Planar Loads: Planar loads are essential for terminating such components as hybrid couplers, power combiners and splitters, and the like. Some factors to be considered in the design of such loads are: 1) the sheet resistivity available; 2) thermal stability or temperature coefficient of the resistive material; 3) the thermal resistance of the load; and 4) the frequency bandwidth. Other applications of planar resistors are bias voltage dividers and dropping resistors. However, such applications should be avoided in monolithic circuits, where power conservation is usually an objective.

Planar resistors can be realized in a variety of forms but fall into three categories: 1) semiconductor films; 2) deposited metal films; and 3) cermets. Resistors based on semiconductors can be fabricated by forming an isolated land of conducting epitaxial film on the substrate, for example, by mesa etching or by isolation implant of the surrounding conducting film. Another way is by implanting a high-resistivity region within the semi-insulating substrate. Metal film resistors are formed by evaporating a metal layer over the substrate and forming the desired pattern by photolithography. These techniques are illustrated in Fig. 20. Cermet resistors are formed from films consisting of a mixture of metal and a dielectric. However, because of the dielectric, they are expected to exhibit an RC frequency dependence similar to that of carbon resistors, which may be a problem in the microwave band.

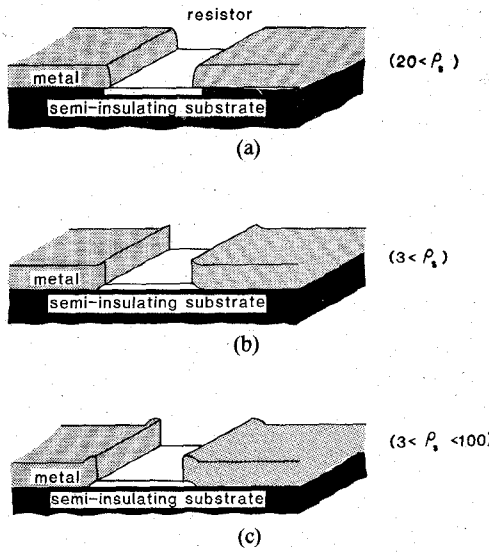


Fig. 20. Examples of planar resistor designs. (a) Implanted resistor. (b) Mesa resistor. (c) Deposited resistor.

TABLE IV
PROPERTIES OF SOME RESISTIVE FILMS

MATERIAL	RESISTIVITY ($\mu\Omega$-cm)	TCR (ppm/°C)	METHOD OF DEPOSITION	STABILITY	COMMENTS
Cr	13 (BULK)	+3000 (BULK)	EVAPORATED SPUTTERED	G-E	EXCELLENT ADHERENCE TO GaAs
Ti	55-135	+2500	EVAPORATED SPUTTERED	G-E	EXCELLENT ADHERENCE TO GaAs
Ta	180-220	-100 TO +500	SPUTTERED	E	CAN BE ANODIZED
Ni Cr	60-600	200	EVAP. (300°C) SPUTTERED	G-E	STABILIZED BY SLOW ANNEAL AT 300°C
TaN	280	-180 TO -300	REACTIVELY SPUTTERED	G	CANNOT BE ANODIZED
Ta_2N	300	-50 TO -110	REACTIVELY SPUTTERED	E	CAN BE ANODIZED
BULK GaAs	3-100 ohms/sq.	+3000	EPITAXY OR IMPLANTATION	E	NONLINEAR AT HIGH CURRENT DENSITIES

Metal films are preferred over semiconducting films because the latter exhibit a nonlinear behavior at high dc current densities and a rather strong temperature dependence—as some metal films do. Not all metal films are suitable for monolithic circuits, since their technology must be compatible with that of GaAs. Table IV lists some candidate metal films along with GaAs.

A problem common to all planar resistors used as microwave loads is the parasitic capacitance attributable to the underlying dielectric region and the distributed inductance of the film, which makes such resistors exhibit a frequency dependence at high frequencies. If the substrate bottomside is metallized, one may determine the frequency dependence by treating the load as a very lossy microstrip line.

For low thermal resistance, one should keep the area of the film as large as possible. To minimize discontinuity effects in width, the width of the resistive film load should not differ markedly from the width of the line feeding it. This means that the resistive element should be as long as possible to minimize thermal resistance. This length is specified by the sheet resistivity of the film and is given by

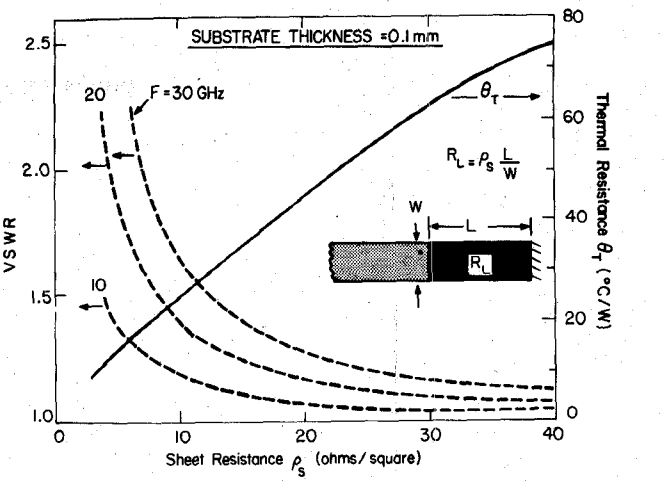


Fig. 21. Thermal resistance and VSWR of a planar resistor as a function of sheet resistance and frequency.

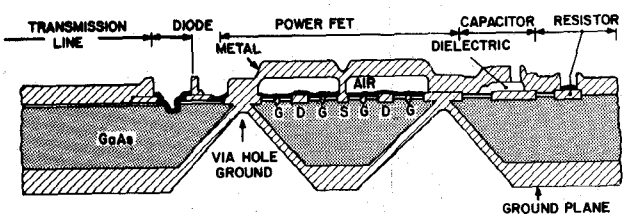


Fig. 22. Composite sketch illustrating technologies used in monolithic circuits.

the formula

$$l = \frac{wR}{\rho_s} \tag{5}$$

where w is the width of the film, R the desired load resistance, and ρ_s the sheet resistance of the film.

If one increases the length of the load (by decreasing the sheet resistivity) to achieve a low thermal resistance, one may get into trouble because the load may begin to exhibit the behavior of a transmission line (albeit a very lossy one) rather than a lumped resistor. Fig. 21 shows how the VSWR increases dramatically at low values of ρ_s because the length of the load becomes too large. Also shown is the thermal resistance. Clearly, a tradeoff is necessary between VSWR and thermal resistance.

All of the technologies we have discussed above are conveniently summarized in the cross-sectional view of a hypothetical monolithic circuit shown in Fig. 22.

4) Transmission Line Junction Effects: The many junctions and bends required of transmission lines in monolithic circuits to achieve close packing introduce unwanted parasitic inductances and capacitances. Fig. 23 illustrates some of the circuit representations of these junctions. Since such discontinuities cannot be avoided, but only minimized, the frequency dependencies must be taken into account, especially when the frequency is above X-band. It is particularly important to include junction effects in any broad-band design, that is, octave bandwidths. Unfortunately, though much work has been done on this topic, the results are not generally in a form useful for the circuit

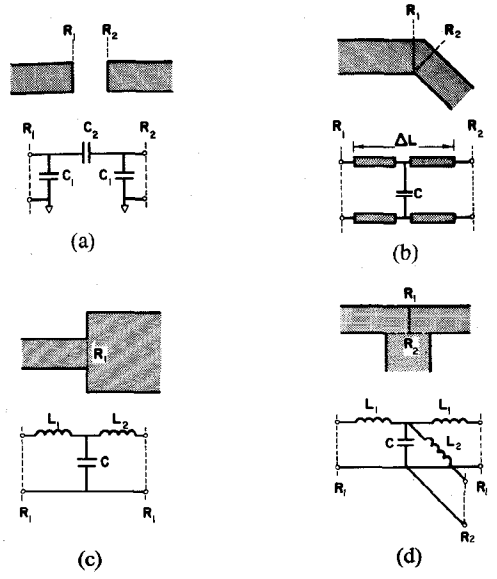


Fig. 23. Some microstrip discontinuities and their equivalent circuits. (a) Gap. (b) Bend. (c) Width discontinuity. (d) Tee junction.

designer. As a consequence, computer-aided design programs do not incorporate corrections for junctions at present.

VI. Examples of Monolithic Circuits

We shall present examples of some practical monolithic circuits which demonstrate the design principles discussed above. These circuits are representative of the research being conducted at laboratories around the world.

Fig. 24 is a photograph of a GaAs chip containing a single-stage four-FET power combiner designed at Raytheon (Research Division) [17]. This amplifier, an *X*-band microwave circuit, was the first to dispense with wire bonds on the chip by use of "via" holes for grounding the source pads. Built on a chip $4.8\times6.3\times0.1$ mm in size, and using a microstrip format with on-chip matching to a 50-Ω system, the circuit exhibited a 5-dB small-signal gain at 9.5 GHz and a saturated CW power output of 2.1 W at 3.3-dB gain (see Fig. 25). Bias was supplied through bias tees via the RF terminals. Although large by present standards, the chip area could be reduced by 30 percent if the capacitive stubs were replaced by thin-film capacitors, which were not available at the time.

An extension of this technology to a two-stage *X*-band power amplifier also designed at this laboratory [22] is shown in Fig. 26. In this circuit, thin-film capacitors, based on SiO or Si_3N_4 technology, were incorporated on the chip for RF blocking and bias applications. Another innovation, clearly evident in the future, is the use of extended integral (grown) beam leads, an offshoot of the airbridge technology. The beam leads allow off-chip bonding of the RF and dc supply connections to the chip, thus avoiding damage to the chip. The amplifier, built on a $2.5\times3.2\times0.1$-mm chip, exhibited a saturated CW power output of 550 mW and 8.5-dB gain at 9.5 GHz and a small-signal gain of 10 dB.

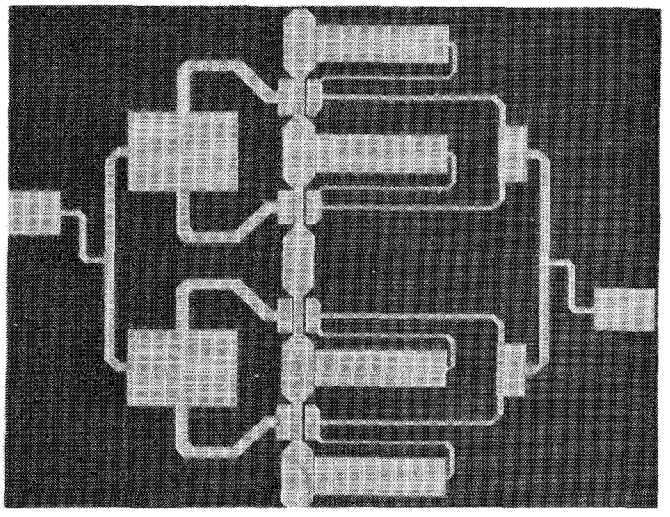


Fig. 24. Monolithic GaAs four-FET *X*-band power combiner. Chip size is $4.8\times6.3\times0.1$ mm. (Raytheon Company.)

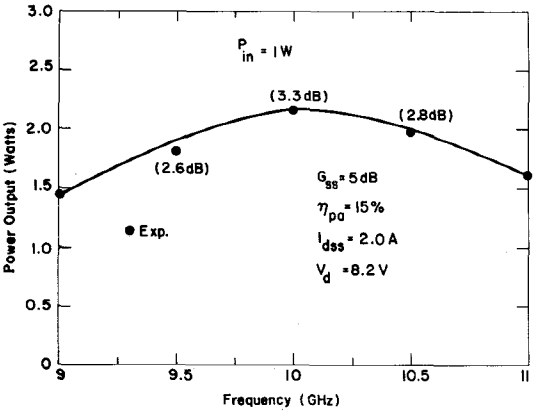


Fig. 25. Power output–frequency response for monolithic GaAs four-FET power combiner.

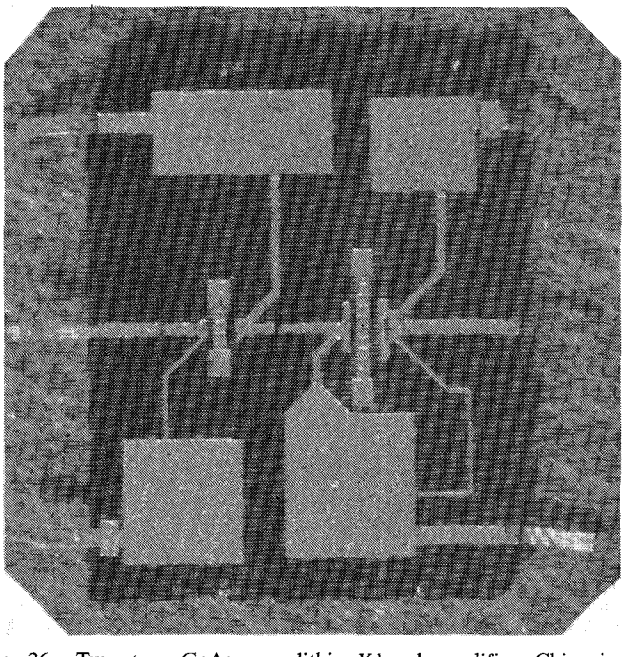


Fig. 26. Two-stage GaAs monolithic *X*-band amplifier. Chip size is $2.5\times3.2\times0.1$ mm. (Raytheon Company.)

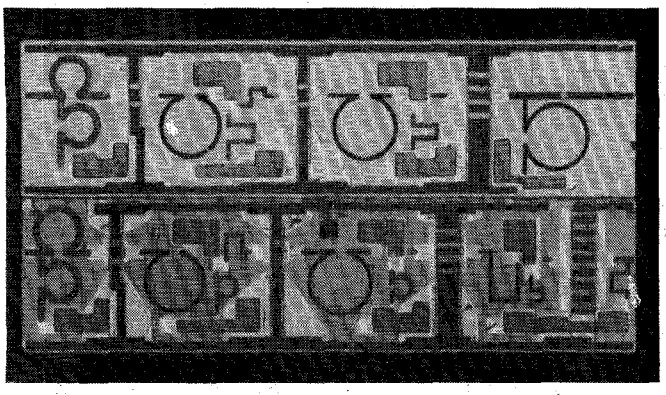


Fig. 27. Three- and four-stage GaAs monolithic *X*-band power amplifiers. Circuit sizes are 1.0×4.0×0.1 mm (Courtesy, W. Wisseman, Texas Instruments, Inc.)

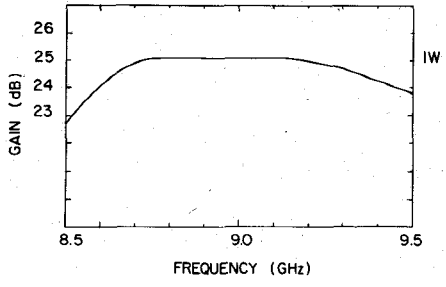


Fig. 28. Measured power gain–frequency response of four-stage amplifier of Fig. 27. (Courtesy, W. Wisseman, Texas Instruments, Inc.)

Turning to results obtained at other laboratories, Fig. 27 represents some of the research at Texas Instruments [18]. Shown is a chip containing side-by-side *X*-band amplifiers: the top, a three-stage FET amplifier; the bottom, a four-stage amplifier. Each chip is 1×4×0.1 mm in size. Both designs are based on a lumped-element approach which uses spiral inductors, clearly evident in the photographs, and thin-film capacitors of the end-coupled variety (Fig. 11(d)). Grounding is achieved by means of a metallized peripheral strip, and bias connections are made by wire-bonds to pads on the chip. The three-stage amplifier delivers 400 mW at 23-dB gain and the four-stage delivers 1 W at 27-dB gain and 15–17-percent power-added efficiency in the 8.8 to 9.2-GHz range (see Fig. 28).

Another circuit reported by this laboratory [18] is the push–pull amplifier shown in Fig. 29. Each channel is a two-stage power amplifier, again based on the lumped-element approach, situated on a 2.0×2.0×0.1-mm chip. Although not monolithic in the strict sense of the word because inductive wire bonds interconnect the two channels, the design is unique in that a "virtual" ground is achieved by connection of the corresponding source pads of the adjacent channels; thus the need for a low inductance ground for the sources is avoided. Over 12-dB gain was obtained at 9.0 GHz with a combined CW power output of 1.4 W. All three amplifiers interface with a 50-Ω system.

An octave bandwidth GaAs amplifier designed at Westinghouse (R. and D. Center) is shown in Fig. 30. This

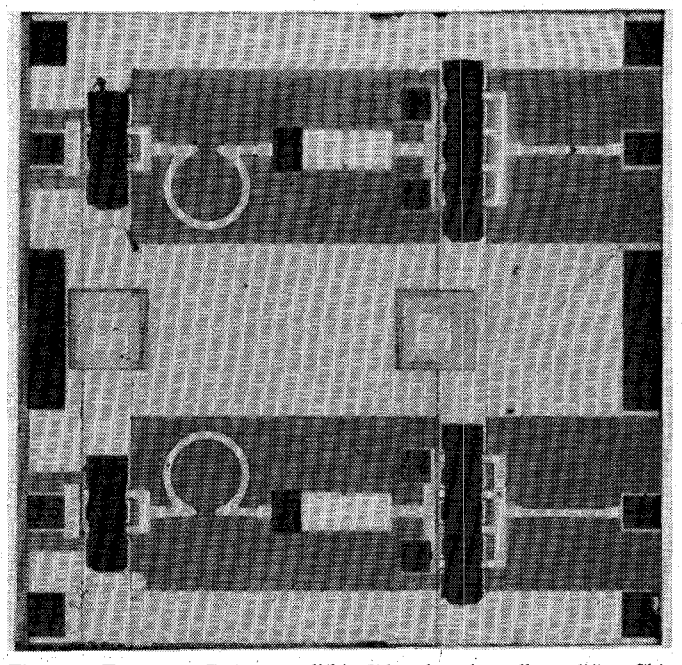


Fig. 29. Two-stage GaAs monolithic *X*-band push–pull amplifier. Chip size is 2.0×2.0×0.1 mm. (Courtesy, W. Wisseman, Texas Instruments, Inc.)

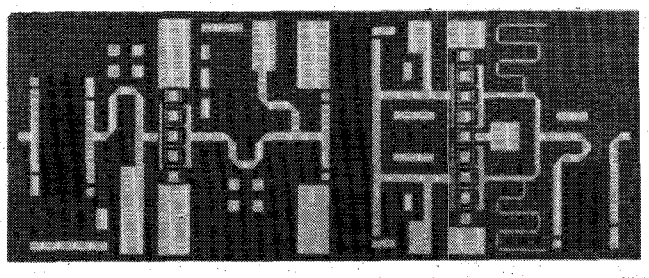


Fig. 30. Two-stage 5.7–11-GHz GaAs monolithic power amplifier. Chip size is 2.0×4.75×0.1 mm. (Courtesy, J. Oakes, Westinghouse.)

circuit, similar to the one reported by Degenford *et al.* [7], consists of 1200-μm and 2400-μm periphery power FET's in cascade formed by selective ion implantation into a semi-insulating substrate. Built on a 2.0×4.75×0.1-mm chip, the circuit is based on a microstrip format with via holes, and makes liberal use of interdigitated capacitors. Source pads are grounded individually with vias. The amplifier produces a power output of 28±0.7 dBm at a gain of 6±0.7 dB across the 5.7 to 11-GHz band.

Another monolithic wideband amplifier is the 4–8-GHz eight-stage GaAs circuit reported by TRW [3] shown in Fig. 31. The design, based on the lumped-element approach (spiral inductors and SiO_2 thin-film capacitors) uses a coplanar feed at the input and output 50-Ω ports, with coplanar ground planes extending the full length of the 2.5×5.0-mm chip.

A departure from the GaAs approach is the SOS three-stage *L*-band amplifier built at Raytheon (Equipment Division) [13] (Fig. 32). This circuit, occupying a chip 7.5×7.5×0.46 mm in size, delivers 200-mW CW output at 20-dB gain at 1.3 GHz. The circuit uses spiral inductors. Dielec-

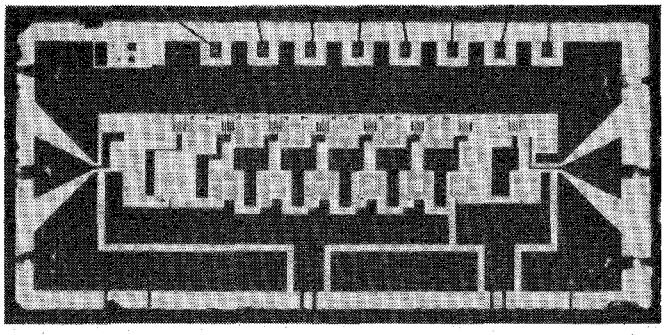


Fig. 31. Eight-stage 4–8-GHz GaAs monolithic amplifier. Chip size is 2.5×5.0 mm. (Courtesy, A. Benavides, T.R.W., Inc.)

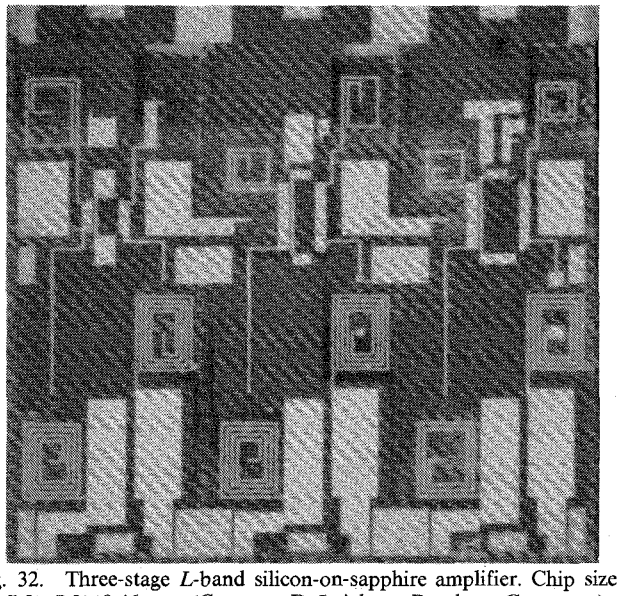


Fig. 32. Three-stage *L*-band silicon-on-sapphire amplifier. Chip size is 7.5×7.5×0.46 mm. (Courtesy, D. Laighton, Raytheon Company.)

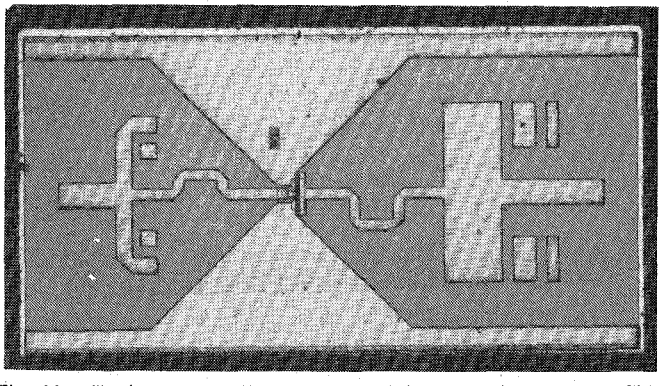


Fig. 33. Single-stage 20-GHz GaAs monolithic low-noise amplifier. Chip size is 2.75×1.95×0.15 mm. (Courtesy, A. Higashisaka, Nippon Electric Company.)

tric (SiO_2) films are used for capacitors and conductor crossovers.

So far we have described power amplifiers only. The first monolithic low-noise amplifier was reported by NEC (Central Research Laboratories) [10] (Fig. 33). This is a one-stage circuit on a 2.75×1.95×0.15-mm GaAs chip. The matching circuits use microstrip lines and stubs to interface with a 50-Ω system through bias tees. Large topside pads are used for the source RF grounds. The circuit, using a 0.5-μm gate, exhibited a noise figure of 6.2 dB and an associated gain of 7.5 dB in the 20.5–22.2-GHz band.

Most of the circuits we have described so far are based on the lumped-element or the microstrip approach or on a

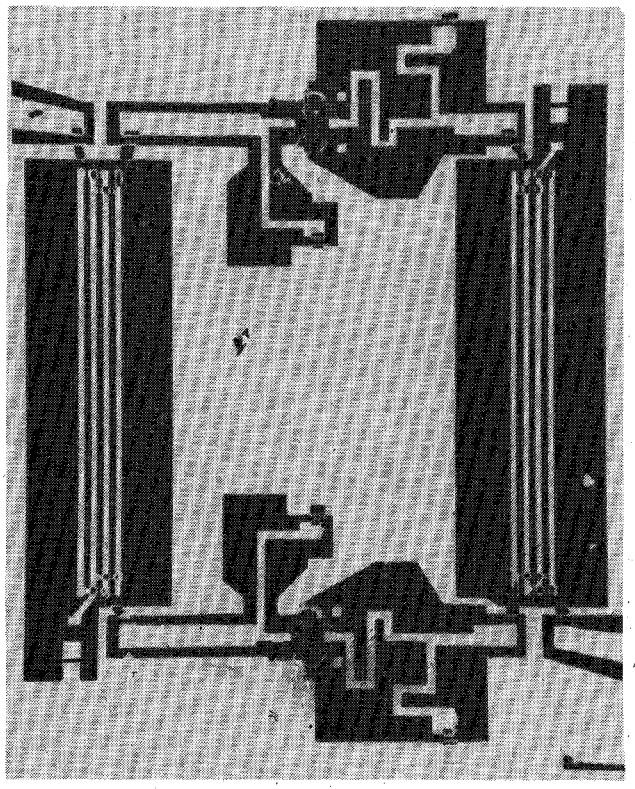


Fig. 34. *X*-band GaAs monolithic balanced amplifier using coplanar coupler. Chip size is 4.0×4.0 mm. (Courtesy, E. M. Bastida, CISE SpA.)

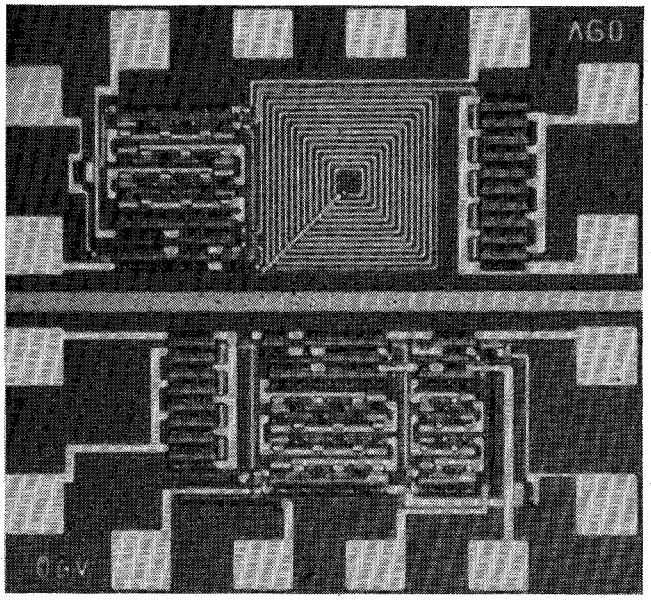


Fig. 35. Multistage direct-coupled GaAs monolithic amplifiers. Circuit sizes are 300×650 μm. (Courtesy, D. Hornbuckle, Hewlett Packard.)

combination of the two. Fig. 34 is a photograph of an *X*-band circuit using coplanar waveguides. This is a balanced amplifier reported by CISE SpA [2] built on a 4.0×4.0-mm GaAs chip, which uses two 90°, 3-dB broadband couplers. The couplers employ CPW rather than MS to obviate the need for micron-line spacings which are necessary with MS couplers. Lumped inductors and thin film (SiO_2) capacitors are used for RF matching and bypass. The circuit utilizes 0.8-μm gate MESFET's and has demonstrated a gain slightly below 10 dB between 8.5 and 11 GHz.

The next circuits, Fig. 35 represent a complete departure from the design philosophy considered so far. Shown is a photograph of two wide-band (0–4.5 GHz) amplifiers de-

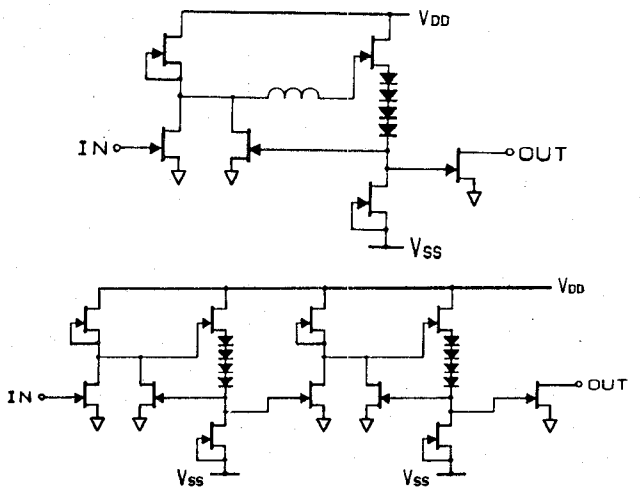


Fig. 36. Circuit schematics for direct-coupled amplifiers shown in Fig. 35. (Courtesy, D. Hornbuckle, Hewlett Packard.)

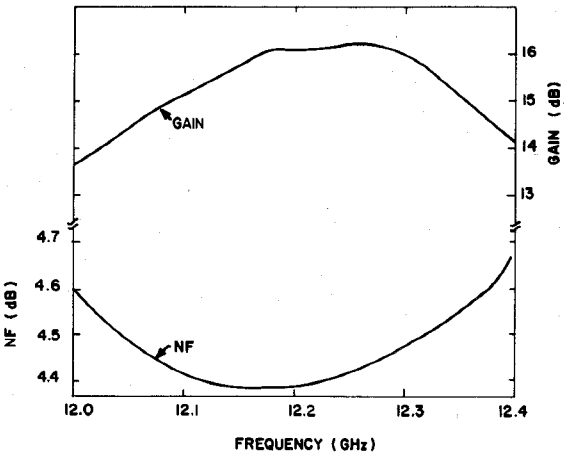


Fig. 38. Performance curves for receiver front end shown in Fig. 37. (Courtesy, P. Harrop, LEP.)

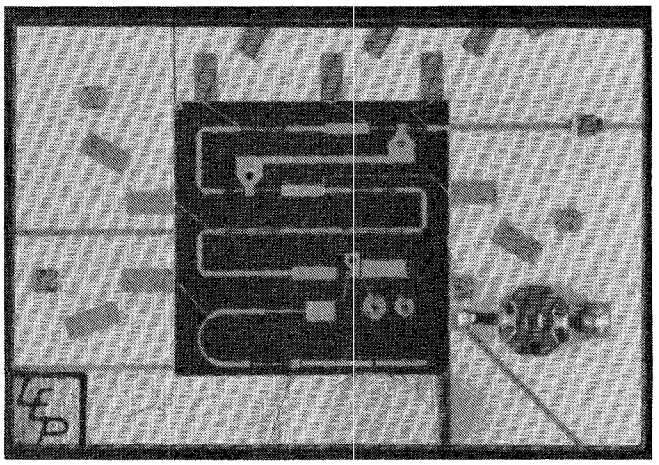


Fig. 37. 12-GHz GaAs monolithic receiver front end. Chip size is 1.0×1.0 cm. (Courtesy, P. Harrop, LEP.)

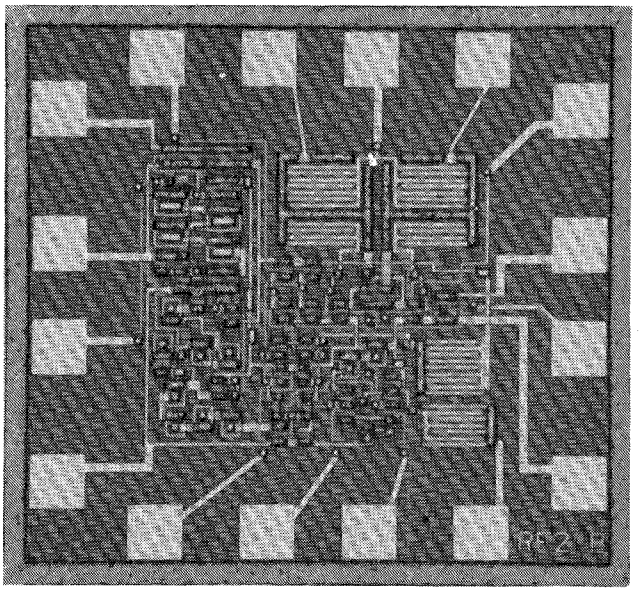


Fig. 39. Direct-coupled GaAs monolithic FET RF signal generation chip. Chip size is 600×650 μm. (Courtesy, R. Van Tuyl, Hewlett Packard.)

signed at Hewlett Packard [11]. What is unique about these circuits is the fact that, except for the spiral inductor, MESFET's are used throughout as active devices and as replacements for resistors and capacitors. The elimination of lumped elements, in conjunction with a direct-coupled circuit approach, allows a very high circuit packing density. Fig. 36 illustrates the circuit complexity achieved in each 0.3×0.65-mm area. Both amplifiers exhibited a gain in excess of 10 dB over the band.

Up until now we have described circuits which earlier we referred to as the lowest level of complexity. The next series of circuits represent integration on a functional block level. The first circuit (Fig. 37) is an integrated receiver front end on a GaAs chip intended for 12-GHz operation. This was reported by LEP [9]. The circuit, deposited on a large 1-cm square chip of GaAs, consists of a two-stage low-noise 12-GHz MESFET amplifier, an 11-GHz FET oscillator, and a dual-gate FET mixer. The matching circuits use microstrip lines and quarter-wave dc blocks. The oscillator is stabilized by an off-chip dielectric resonator. Bias circuits are included on the surrounding alumina substrate. Preliminary results are summarized in Fig. 38. The circuit is intended for a potential consumer market for domestic satellite-to-home TV reception planned for Europe.

Another example of the functional block approach is the monolithic GaAs FET RF signal generation chip (Fig. 39) designed at Hewlett Packard [21]. An extremely high degree of integration was achieved by use of the direct-coupled approach described earlier. Contained within the 0.65×0.65-mm chip is the circuit shown in the schematic of Fig. 40. The local oscillator is resonated by an off-chip inductor which is tuned over the 2.1–2.5-GHz range by an on-chip Schottky barrier junction capacitor. The circuit is intended for an instrument application.

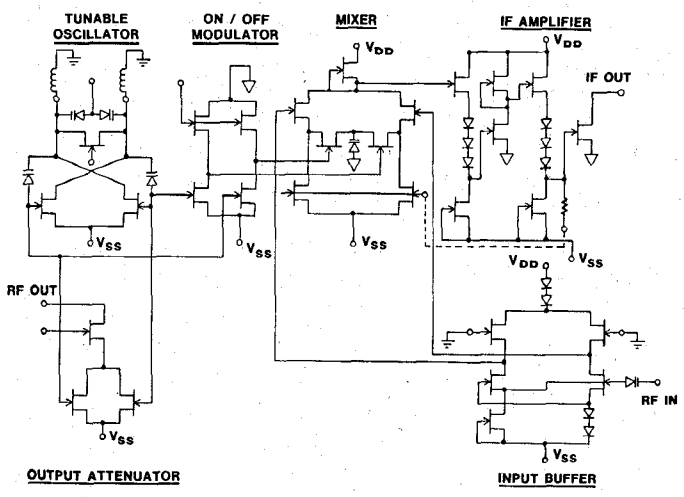


Fig. 40. Schematic for direct-coupled signal generation chip shown in Fig. 39. (Courtesy, R. Van Tuyl, Hewlett Packard.)

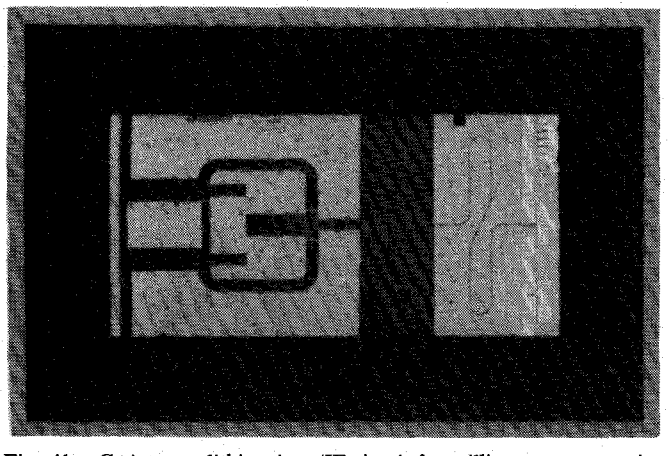


Fig. 41. GaAs monolithic mixer/IF circuit for millimeter-wave receiver applications. Chip size is 2.7×5.3×0.18 mm. (Courtesy, R. Sudbury, Lincoln Laboratories.)

Our final functional block circuit is the monolithic balanced Schottky-barrier diode mixer/IF FET preamplifier chip illustrated in Fig. 41. This MS circuit, reported by Lincoln Laboratories [6], is built on a 2.7×5.3×0.18-mm GaAs chip in MS format. The circuit operates between a 31-GHz signal source and a 2-GHz IF output. An external oscillator signal is injected through one of the coupler ports. The circuit exhibits an overall gain of 4 dB and a single-sideband noise figure of 11.5 dB.

We now turn to some special passive components fabricated in monolithic form. The first is a Wilkinson combiner/divider reported by Raytheon [23] shown in Fig. 42. Built on a 1.5×2.5×0.1-mm chip, the circuit uses a thin-film titanium balancing resistor and was designed to operate at a center frequency of 9.5 GHz. Note the extended beam leads. As an illustration of the extremely good electrical balance that one can achieve with the high-resolution photolithography intrinsic to the monolithic approach, we show in Fig. 43 a graph of the power division and phase balance measured for the two 3-dB ports.

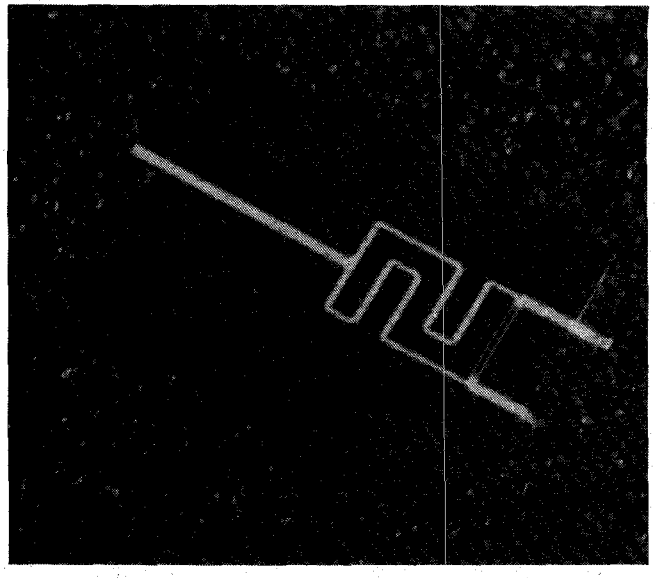
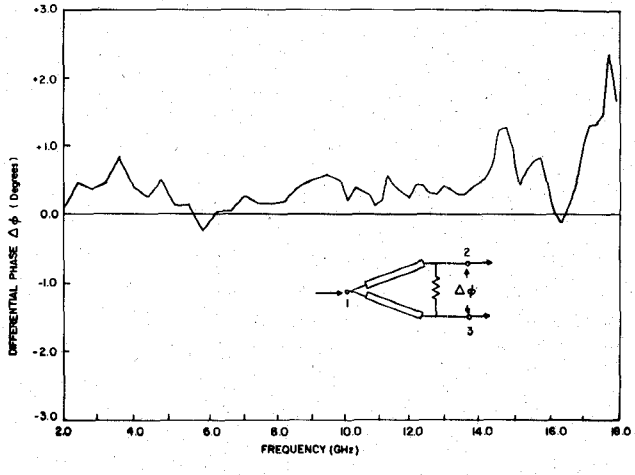
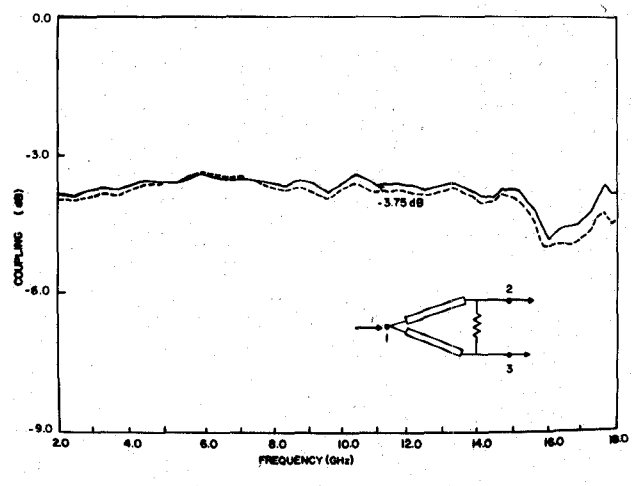


Fig. 42. GaAs monolithic *X*-band Wilkinson combiner/divider. Chip size is 1.5×2.5×0.1 mm. (Raytheon Company.)



(a)



(b)

Fig. 43. Measured phase and power balance of Wilkinson divider shown in Fig. 42.

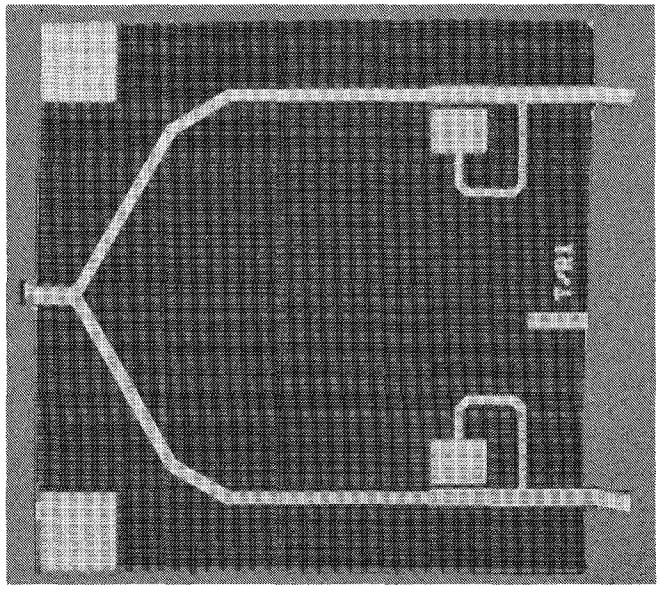


Fig. 44. GaAs monolithic *X*-band transmit/receive switch. Chip size is 3.0×3.0×0.1 mm. (Raytheon Company.)

Another component designed at this laboratory is the all-FET T/R switch shown in Fig. 44 [1]. This switch, intended for phased-array applications at *X*-band, requires no dc hold power in either state. Built on a 3.0×3.0×0.1-mm chip, the switch exhibits an isolation in excess of 33 dB between the transmitter and receiver ports in the 7–13-GHz range, and an insertion loss as low as 0.5 dB within this band. An alternative approach, also using FET's, was reported by McLevige *et al.* [14]. Both approaches utilize the change in source–drain resistance with gate bias.

The examples we have shown, though not exhaustive, are representative of the work reported so far (December 1980) and are intended to give the reader a good perspective of the advances made in the field during the last two years. Needless to say, the next several years will see the emergence of a still higher level of circuit integration in this rapidly developing field.

VII. Future Developments

We have so far concerned ourselves primarily with the technical aspects of monolithic circuits—their technology, design considerations, and microwave performance. Problems have been described and their solutions demonstrated. This is as it should be in the early stage of development of a new technical venture. No major unsolvable technical problems are evident; therefore, on the basis of technical considerations alone, there is no reason why the steady rate of progress already established cannot be maintained, indeed, accelerated.

What then will determine the future course of progress? The answer is simple—cost! Because the development of MMIC's requires a large capital investment and involves time-consuming and expensive processing technology plus a sophisticated testing procedure, the future development of this field will rest squarely on the as yet unproven expectations of reductions in cost and, to a lesser extent, improvements in reliability and reproducibility accruing from the monolithic approach.

The matter of cost reductions, in turn, rests on the answers to two questions.

1) Will the many complex technology steps required of MMIC's lend themselves to a high-yield production process?

2) Will a mass market develop in the microwave system area—a mass market necessary to capitalize on the high-volume low-cost attributes of batch processing?

Both of these requirements were eventually satisfied for silicon technology. Will this happen for gallium arsenide microwave technology? Time will tell. Since the silicon development was helped along by a vast domestic market (radios, TV's, etc., and more important, the commercial computer) and military markets, what are the expected large-volume markets for MMIC's?

Two potentially large markets appear to be developing, one military, the other consumer. In the military area, one such market includes electronically scanned radar systems, especially airborne and space-borne systems being planned for the future. For it is in the phased-array antenna, which may require modules as high as 10^5 in number, that we find a microwave system analog of the computer, which gave impetus to the growth of the silicon IC market. The antenna module requirements have already spurred developments of such module subsystems as transmitters, low-noise receivers, phase shifters, and transmit–receive (T/R) switches, some examples of which were described earlier. Here, along with cost, important design performance criteria will be reliability and small weight and size.

Another military application is in ECM systems, which require low-cost high-gain broad-band amplifiers. The difficult technical problems and projected high manufacturing cost associated with the hybrid integrated approach to this task have in essence mandated the use of monolithic circuits. Finally, the possibility of merging high-speed GaAs digital and microwave circuitry on the same chip may encourage use of such circuitry in signal processing at the RF level.

Turning to the nonmilitary markets, one potentially large outlet may be receiver front ends for the direct satellite-to-home-TV consumer market. Numerous such systems are being planned, for example, in Europe. We have described earlier one circuit intended for this market.

A third potential market, though much smaller in size, is instrumentation. Here cost and possibly circuit packing density are most important. Several examples of circuits earmarked for this application have been described.

We have not said much of the millimeter-wave spectrum. It is perhaps premature to do so, as this field itself is in its early stage of development. Here monolithic applications might develop, more for technical reasons than for economic reasons, because of the important role played by undesired packaging parasitics associated with discrete devices at these high frequencies. It is not unlikely that here too, as at lower frequencies, military applications may spur initial development. Now we turn to the question of costs.

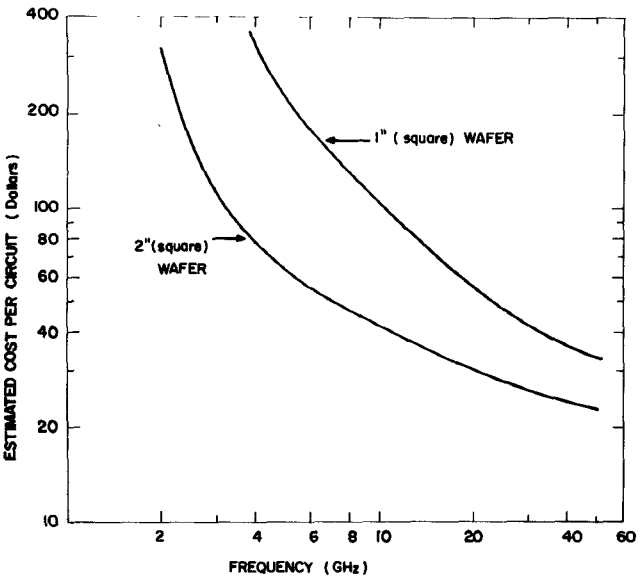


Fig. 45. Zero-order estimate of circuit cost as a function of frequency and wafer size.

Cost being the crucial item that it is, what are the factors contributing to it? They are the following:

1) materials;
2) materials processing;
3) circuit/device technology;
4) circuit assembly and packaging; and
5) testing (dc and RF).

These items, as is evident, do not include the important but nonrecurring costs such as capital investment, engineering, and mask design. In the materials category we include cost of substrate qualification, epitaxial growth and/or ion implantation, and profile evaluation, among others. Testing includes both dc and RF testing at the wafer probe level as well as circuit performance evaluation at the jig or package level.

Can a dollar figure be attached to these costs? At this stage, no! It is hazardous, at best, to attempt an accurate cost analysis based on laboratory experience, for large volume production, because ultimate module costs will be directly dependent on circuit yield in a manufacturing environment.

It is helpful, nevertheless, to attempt at least a "zeroth" order estimate of potential circuit costs, not so much to obtain an absolute level of cost, but to pinpoint the high cost items in the list above. To do this we have estimated the number of available circuits per wafer as a function of circuit operating frequency. This estimate was shown in Fig. 2. In the context of our present discussion, a circuit is equivalent to one submodule, for example, a transmitter stage or two, a phase shifter, etc. Using this estimate, we have determined the cost per circuit as a function of operating frequency. This data is shown in Fig. 45.

The cost estimates were derived by assumption of a 50-percent processing yield, independent of frequency. The base cost includes material cost, fully loaded labor cost, and circuit qualification at the dc and RF level. We feel that the cost estimates shown are useful guidelines but they should not be considered accurate in any absolute sense. For instance, depending on circuit complexity, current laboratory yields at X-band range from near zero to 20 percent. The development of a 50-percent yield fabrication process technology, deemed necessary, requires much additional experience and substantial simplification of monolithic circuit fabrication techniques.

Adjustments may be necessary at either end of the frequency scale. For example, in the range below 3–4 GHz, a drastic cost reduction may ensue, at least for some circuit applications, if the direct-coupled approach can be used. At the other end of the scale, above, say, 10–12 GHz, the cost figures should be elevated. The reason is that, because of the necessity of submicron gate technology, the lower throughput of the ultrahigh resolution electron beam (EB) lithography will increase costs substantially. Here what will be needed is optimization of the processing technology by appropriate merging of the EB lithography for the active devices and the higher throughput photolithography which is more than adequate for the circuit elements. This problem has not yet been addressed.

On the basis of our cost analysis, certain definite conclusions can be reached about the expected relative cost of the several items listed above. First, the two materials factors, under large production lots (>100 K parts) will contribute a negligible amount to the total cost—of the order of 5 percent or less. Second, next to wafer processing, the cost of packaging and microwave testing will be the largest cost factor. Indeed, because these latter costs will be fairly independent of the frequency band, and because of the decreasing processing cost per circuit with increasing frequency and wafer size, it is expected that packaging and testing will be the dominant cost factor at the higher frequencies (perhaps above 10 GHz).

It seems evident that, in light of this conclusion, the reduction of assembly and testing costs will be of paramount importance and must be addressed rather soon. Not only must as many functions as possible be integrated on one chip, consistent with high yield, but RF testing of chips and monolithic circuits and modules must be automated, just as dc tests have been. This will be very difficult because RF probes small enough for chip use are still in the laboratory stage, and their extension to performance tests on an entire circuit are nonexistent.

VIII. Conclusions and Summary

Monolithic microwave circuits based on gallium-arsenide technology have finally become a practical reality. Owing its origin to early experiments based on silicon bipolar technology, the gallium-arsenide approach, except for some scattered results in the sixties, emerged as a serious development only within the last three years.

The factors most responsible for this rapid growth can be traced to: 1) the development of the Schottky-barrier field-effect transistor; 2) the excellent microwave properties of semi-insulating GaAs as a low-loss substrate; 3) the perfection of GaAs epitaxy and ion implantation; 4) the

establishment of GaAs crystal pulling facilities capable of large-diameter crystal growth; and 5) the emergence of potential systems applications for monolithic microwave circuits.

We have attempted to demonstrate in this paper some of the many design considerations and tradeoffs that must be made to optimize the performance of GaAs monolithic microwave circuits. Attention has been focussed, primarily, on the nondevice aspects of monolithic circuit design.

Despite the small physical size of the circuitry, interconnections between components often must be treated as wave-propagating structures because of the high dielectric constant of GaAs, which reduces the wavelength within the substrate. Both coplanar waveguide and microstrip lines, as well as combinations of both, are appropriate for monolithic circuits.

A typical circuit design may use both distributed and lumped-element components. Lumped elements, it was shown, are not truly lumped, because of built-in parasitics arising from the dielectric substrate. These must be taken into account at *X*-band and higher frequencies. A major drawback of thin-film inductors and capacitors is the limited *Q*-factor achieved to date. Much has yet to be learned about loss reduction in thin dielectric films.

We have shown that MMIC's are realized rather easily. Via hole grounding and source airbridge interconnections are eminently suited for them. Computer-aided design techniques are a "must" to reduce the number of iterations necessary.

Many examples of monolithic circuits have been shown which demonstrate the design principles described. These circuits, representing a world-wide cross section of the efforts in this field, have emerged within the last two to three years, and demonstrate the variety of circuit applications amenable to the monolithic approach. The promising attributes of the monolithic technology to cut fabrication costs, improve reliability and reproducibility, and reduce size and weight will overcome many of the shortcomings of the hybrid approach.

We have argued that, based on the cost considerations, the potential markets for MMIC's will be for the most part systems requiring large quantities of circuits of the same type. Because of this, and because of the large capital expenditures required of an organization to become a viable contender for these markets, it is most likely that the major efforts in MMIC's will eventually reside in the systems houses themselves.

ACKNOWLEDGMENT

The progress reported in this paper represents the cumulative effort of many people, too numerous to mention individually. However, the author wishes to express his deep appreciation to his associates at Raytheon, whose work is described here, and to his colleagues at many other laboratories who so graciously gave him permission to use their photographs and latest results to help describe their research.

REFERENCES

[1] Y. Ayasli, R. A. Pucel, J. L. Vorhaus, and W. Fabian, "A monolithic *X*-band single-pole, double-throw bidirectional GaAs FET Switch," in *GaAs IC Symp. Res. Abstracts*, 1980, paper no. 21.

[2] E. M. Bastida, G. P. Donzelli, and N. Fanelli, "An *X*-band monolithic GaAs balanced amplifier," in *GaAs IC Symp. Res. Abstracts*, 1980, paper no. 25.

[3] A. Benavides, D. E. Romeo, T. S. Lin, and K. P. Waller, "GaAs monolithic microwave multistage preamplifier," in *GaAs IC Symp. Res. Abstracts*, 1980, paper no. 26.

[4] L. Besser, "Synthesize amplifiers exactly," *Microwave Syst. News*, pp. 28–40, Oct. 1979.

[5] D. Ch'en and D. R. Decker, "MMIC's the next generation of microwave components," *Microwave J.*, pp. 67–78, May 1980.

[6] A. Chu, W. E. Courtney, L. J. Mahoney, G. A. Lincoln, W. Macropoulos, R. W. Sudbury, and W. T. Lindley, "GaAs monolithic circuit for millimeter-wave receiver application," in *ISSCC Dig. Tech. Pap.*, pp.144–145, 1980.

[7] J. E. Degenford, R. G. Freitas, D. C. Boire, and M. Cohn, "Design considerations for wideband monolithic power amplifiers," in *GaAs IC Symp. Res. Abstracts*, 1980, paper no. 22.

[8] E. Denlinger, "Losses of microstrip lines," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-28, pp. 513–522, June 1980.

[9] P. Harrop, P. Lesarte, and A. Collet, "GaAs integrated all-front-end at 12 GHz," in *GaAs IC Symp. Res. Abstracts*, 1980, paper no. 28.

[10] A. Higashisaka, in *1980 IEEE MTT-S Workshop Monolithic Microwave Analog IC's.*

[11] D. Hornbuckle, "GaAs IC direct-coupled amplifiers," in *1980 IEEE MTT-S Int. Microwave Symp. Dig.*, pp. 387–388.

[12] T. M. Hyltin, "Microstrip transmission on semiconductor substrates," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-13, pp. 777–781, Nov. 1965.

[13] D. Laighton, J. Sasonoff, and J. Selin, "Silicon-on-sapphire (SOS) monolithic transceiver module components for *L*- and *S*-band," in *Government Microcircuit Applications Conf. Dig. Pap.*, vol. 8, 1980, pp. 299–302.

[14] W. V. McLevige and V. Sokolov, "A monolithic microwave switch using parallel-resonated GaAs FET's," in *GaAs IC Symp. Res. Abstracts*, 1980, paper no. 20.

[15] E. Mehal and R. W. Wacker, "GaAs integrated microwave circuits," *IEEE Trans. Microwave Theory Tech.* vol. MTT-16, pp. 451–454, July 1968.

[16] R. S. Pengelly and J. A. Turner, "Monolithic broadband GaAs FET amplifiers," *Electron. Lett.* vol. 12, pp. 251–252, May 13, 1976.

[17] R. A. Pucel, J. L. Vorhaus, P. Ng, and W. Fabian, "A monolithic GaAs *X*-band power amplifier," in *IEDM Tech. Dig.* 1979, pp. 266–268.

[18] V. Sokolov and R. E. Williams, "Development of GaAs monolithic power amplifiers," *IEEE Trans. Electron Devices*, vol. ED-27, pp. 1164–1171, June 1980.

[19] E. Strid and K. Reed, "A microstrip probe for microwave measurements on GaAs FET and IC wafers," in *GaAs IC Symp. Res. Abstracts*, 1980, paper no. 31.

[20] R. N. Thomas, "Advances in bulk silicon and gallium arsenide materials technology," in *IEDM Tech Dig.*, 1980, pp. 13–17.

[21] R. Van Tuyl, "A monolithic GaAs FET RF signal generation chip," in *ISSCC Dig. Tech. Pap.* 1980, pp. 118–119.

[22] J. L. Vorhaus, R. A. Pucel, Y. Tajima, and W. Fabian, "A two-stage all-monolithic *X*-band power amplifier," in *ISSCC Dig. Tech. Pap.*, pp.74–75, 1980.

[23] R. C. Waterman, W. Fabian, R. A. Pucel, Y. Tajima, and J. L. Vorhaus, "GaAs monolithic Lange and Wilkinson couplers," *GaAs IC Symp. Res. Abstracts*, 1980, paper no. 30.