Metalorganic Chemical Vapor Deposition for Optoelectronic Devices

JAMES J. COLEMAN, FELLOW, IEEE

Invited Paper

The metalorganic chemical vapor deposition (MOCVD) process for electronic and photonic compound semiconductor materials and devices is reviewed. We begin with an introduction to the basic MOCVD chemical reaction process, gas delivery equipment, reaction chambers, and safety. Growth mechanisms, including hydrodynamics, boundary-layer issues, thermal effects, and pyrolysis reactions, are defined, and criteria for growth regimes, growth rate, and alloy composition are described. Material, structural, and dopant considerations, which are particularly important to optoelectronic devices, are presented. Last, a brief description of the selective area epitaxial growth process is presented.

Keywords—Compound semiconductors, epitaxial growth, MOCVD, optoelectronic materials, quantum-well heterostructures.

I. INTRODUCTION

The metalorganic chemical vapor deposition (MOCVD) process for the growth of compound semiconductor materials and devices originated in the pioneering work of H. M. Manasevit [1], at what was then North American Rockwell, in 1968. This process is also called organometallic chemical vapor deposition (OMCVD), metalorganic vapor phase epitaxy (MOVPE), and organometallic vapor phase epitaxy (OMVPE). Manasevit originally named the process MOCVD to emphasize the metal constituent and to avoid confusion with organometallic chemistry research, which generally places more interest in, and emphasis on, higher order organic radicals. His early work included epitaxy on crystalline substrates but mainly involved the more general chemical vapor deposition of amorphous thin films on a number of different solid surfaces, such as sapphire. In the early 1970’s, researchers at Rockwell and other laboratories recognized the limitations of the state-of-the-art (at the time) liquid phase epitaxy (LPE) semiconductor growth process. By applying the lessons of silane chemical vapor deposition for silicon epitaxy to compound semiconductor materials and devices, they determined a need for controllable large-area and thin-layer growth and a manufacturable gas-phase process. The result was an explosion in the research, development, and commercialization of the MOCVD process for virtually all forms of compound semiconductor electronic and optical devices so that, by the early 1980’s, LPE was reduced to a laboratory curiosity.

Since that time, actually a remarkably long period of time when measured on the scale of innovations in crystal growth processes, the MOCVD process, along with the molecular beam epitaxy (MBE) process described in a paper by K.-Y. Cheng in this special issue, has dominated the research, development, and manufacture of compound semiconductor devices. MOCVD is the epitaxial crystal growth technology of choice for an impressive array of commercial devices, including lasers, light emitting diodes (LED’s), photocathodes, heterostructure bipolar transistors, photodetectors, and solar cells. Virtually every alloy compound semiconductor materials system has been grown successfully by MOCVD, including the antimonides for long (> 2 µm) wavelengths, the quaternary InGaAsP for telecommunications wavelengths (1.3–1.55 µm), AlGaAs-GaAs-InGaAs structures for wavelengths in the range of 700–1100 nm, visible InGaP and related compounds, and the column III nitrides for blue and ultraviolet lasers and LED’s. The diode laser market alone, dominated by MOCVD-grown devices, is projected to reach $1.9 billion in sales in 1997 [2]. Even though the blue nitride LED technology is relatively young, Nichia is producing many millions of these devices every month by MOCVD [3]. Hundreds of MOCVD reactor systems have been sold worldwide, and reactors capable of producing 2500 cm² of epitaxial material in a single growth run are commercially available [4].

1 See pp. 1694–1714.
In this paper, we will outline the basic MOCVD chemical reaction process and describe the gas delivery equipment and reaction chambers common to the process. Safe handling of the constituent materials will be addressed. The growth mechanisms, including hydrodynamics, boundary-layer issues, thermal effects, and pyrolysis reactions, will be defined to establish criteria for growth regimes, growth rate, and alloy composition. Material, structural, and dopant considerations that are particularly important to optoelectronic devices are presented. Last, a brief description is included of the selective area growth process, which promises to be a key technology for future photonic integrated circuit devices. In this paper, we present an overview of this remarkable growth process. We encourage the reader interested in learning about the process in more detail to refer to the extensive literature on the MOCVD growth process, starting with some of the books and book chapters on the subject [5]–[14].

II. BASIC PROCESSES AND EQUIPMENT

In this section, we describe the basic chemical reaction that defines MOCVD, the kinds of sources specific to this particular epitaxial growth process, and, in general terms, the hardware used to deliver and process the sources into the resulting thin films.

A. Reaction Equations

We can begin to understand the essence of the MOCVD process for deposition of compound semiconductors and alloys by considering the simplest possible basic chemical-reaction equation, which is in itself very important, and then extending to more complex situations. This simplest case [1] involves a pyrolysis reaction of the vapors of a volatile organometallic compound and a gaseous hydride, given by

\[ R_n A + D H_n \rightarrow A D + n R H \]  

(1)

where \( R \) is an organic radical of some unspecified form but generally of lower order, such as a methyl- or ethyl-radical, and \( A \) and \( D \) are the constituent species for the deposited solid. An important example of this simplest case is given by

\[ (C H_3)_3 G a + A s H_3 \rightarrow G a A s + 3 C H_4 \]  

(2)

A large number of organometallic compounds have been studied as sources for the MOCVD process, with the most important arguably being trimethylgallium (TMGa), trimethylaluminum (TMAI), and trimethylindium (TMIn). Ideal sources should be easily synthesized and easily purified and have reasonable vapor pressures, where “reasonable” is a term relating to the practicalities of using several different sources together. Usually, these are liquids, but there are important exceptions, such as TMIn, which are solids with reasonable vapor pressure. For solid sources, vendors offer a number of different creative solutions to the problem of forcing a relatively constant evaporation surface area. In general, the organometallic constituents are transported to a heated substrate by passing a carrier gas, usually hydrogen or nitrogen or a mixture of the two, over or through the compound contained in a constant-temperature bubbler vessel. There is a simple relationship, discussed below, between the bubbler temperature and the constituent delivery rate, provided the carrier gas flow rate does not exceed the saturation rate of the vapor over the liquid or solid source.

Most MOCVD growth of III–V compound semiconductors and alloys involves the use of hydrides, such as arsine or phosphine, for the column V species. In principle, these are the simplest of column V sources to use because they are already gaseous and supplied from simple cylinder-based delivery systems. In practice, the toxicity of these gases requires that special attention be given to the safe transport, handling, and storage of these gases. Largely for these safety reasons, there also has been considerable attention given to the use of organometallic compounds [15]–[19] for both the column III and V species, in which case the reaction equation becomes either

\[ R_n A + R'_n B \rightarrow AB + n R H + n R H \]  

(3)

or

\[ R_n A + R'_n B \rightarrow AB + n R H' \]  

(4)

The growth of semiconductor alloys by MOCVD is easily accomplished by mixing the vapors of the different alloy constituents in the appropriate vapor phase ratio to form the desired composition. A general equation for a ternary alloy is given by

\[ x R_n A + (1 - x) R_n B + D H_n \rightarrow A_x B_{1-x} D + n R H \]  

(5)

which applies, for example, to ternary InGaAs

\[ x (C H_3)_3 I n + (1 - x) (C H_3)_3 G a + A s H_3 \rightarrow I n_x G a_{1-x} A s + 3 C H_4 \]  

(6)

A general equation for a quaternary alloy is given by

\[ x R_n A + (1 - x) R_n B + y D H_n + (1 - y) E H_n \rightarrow A_x B_{1-x} D_y E_{1-y} + n R H \]  

(7)

Table 1 is a selection of common alkyl and hydride sources used for MOCVD growth. There are other sources and many different combinations that can be used to form a wide variety of III–V and II–VI semiconductor compounds and alloys. More detailed discussion is beyond the scope of this work, and the reader is encouraged to refer to the various books and book chapters listed in the references for further information on specific combinations.

B. Reactor Gas Delivery Systems

MOCVD reactors consist of three major components: the reactor gas delivery system, the reaction chamber, and the reactor safety infrastructure. The reactor delivery system [20]–[25], or gas panel, is a very clean, leak-free network of stainless-steel tubing, automatic valves, and electronic mass flow controllers, as shown in Fig. 1. Each constituent or dopant type (e.g., gaseous hydride, high vapor...
Fig. 1. Schematic diagram of an MOCVD reactor delivery system gas panel, illustrating hydride delivery modules, alkyl delivery modules, and the vent-run configuration.

Table 1  Typical Column III and V Sources

<table>
<thead>
<tr>
<th>Source</th>
<th>Abbreviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trimethylgallium</td>
<td>TMGa</td>
</tr>
<tr>
<td>Triethylgallium</td>
<td>TEGa</td>
</tr>
<tr>
<td>Trimethylaluminum</td>
<td>TMAI</td>
</tr>
<tr>
<td>Triethylaluminum</td>
<td>TELi</td>
</tr>
<tr>
<td>Trimethylindium</td>
<td>TMIn</td>
</tr>
<tr>
<td>Triethylindium</td>
<td>TEl</td>
</tr>
<tr>
<td>Arsine</td>
<td>AsH₃</td>
</tr>
<tr>
<td>Phosphine</td>
<td>PH₃</td>
</tr>
<tr>
<td>Stibine</td>
<td>SbH₃</td>
</tr>
<tr>
<td>Ammonia</td>
<td>NH₃</td>
</tr>
<tr>
<td>Trimethylphosphorus</td>
<td>TMP</td>
</tr>
<tr>
<td>Trimethylarsenic</td>
<td>TMA</td>
</tr>
<tr>
<td>Trimethylantimony</td>
<td>TMSb</td>
</tr>
<tr>
<td>Triethylphosphorus</td>
<td>TEP</td>
</tr>
<tr>
<td>Triethylarsenic</td>
<td>TEl</td>
</tr>
<tr>
<td>Triethylantimony</td>
<td>TESb</td>
</tr>
<tr>
<td>Tertiarybutylarsenic</td>
<td>TBAs</td>
</tr>
<tr>
<td>Tertiarybutylphosphorus</td>
<td>TBP</td>
</tr>
</tbody>
</table>

Table 1  Typical Column III and V Sources

provided for source replacement without contamination of the rest of the delivery system. Source modules for critical alkyls commonly include ultrasonic monitoring of the mass flow, with the possibility for using the output from such monitors for closed-loop control. Additional gas-handling hardware may be required for dilution of either hydride or alkyl modules to increase the effective range of available flows, especially for dopant sources.

An important part of the main gas panel is the supply of carrier gases within a vent-run configuration, as shown in Fig. 1. Constant source delivery is critical for thin quantum-well optoelectronic devices and can have implications in the growth of thicker heterostructures in materials systems that must be lattice matched. Small changes in carrier gas flow can significantly change the source delivery. Therefore, the design of gas delivery systems must avoid transients from switching or dead space. To improve carrier gas consistency further, the vent-run system maintains a relatively large flow rate of carrier gas (typically several liters/min) in the supply line. Often there are separate hydride and alkyl supply lines. Special vent-run valves that have nearly zero dead space couple the individual source modules to the supply line. The source flow rate (usually on the order of tens or hundreds of cm³/min) can then be established and stabilized while the valve is vented to a waste line, shown as a dashed line in Fig. 1, and prior to injection into the run supply line, shown as a solid line in Fig. 1. An additional vent-run supply line is often used to combine the flows from the hydride and alkyl supply lines, either entering or bypassing the reaction chamber. As long as a fixed relationship is maintained among the total pressures.
in all vent and supply lines, transients will be minimized and the actual source flow can be controlled.

C. Reaction Chambers

The reaction chamber is the vessel in which the source gases are mixed, introduced into a heated zone where an appropriate substrate is located, and the basic pyrolysis reactions described above take place. There are two basic reaction-chamber geometries [1], [26] commonly used for the MOCVD growth of optoelectronic materials. The two chamber models, shown in simplified form in Fig. 2, are fundamentally different in operation, but one or the other can be used to describe the operation of most, more sophisticated commercial reaction chambers. There are a number of design features in common. Both designs are cold-wall systems that reflect the basic pyrolysis nature of the process, both contain a relatively small diameter inlet into some form of transition region, and both make use of an indirectly heated (radio-frequency induction heated or infrared radiant heated) silicon carbide-coated graphite susceptor. The chamber itself can be quartz, stainless steel, or quartz-lined stainless steel.

In the vertical reaction chamber [Fig. 2(a)], the process gases enter at the top and are deflected by a baffle before moving downward through a cold transition region and approaching normal to the heated susceptor. The gas flow is forced to the sides by the susceptor with a velocity profile dependent on geometry and affected by the thermal profile in the system. Better uniformity is obtained by rotating the susceptor. Exhaust gases escape through the base of the reaction chamber. Large susceptors can accommodate multiple wafers. In the horizontal reaction chamber [Fig. 2(b)], the process gases enter from the small inlet to the left and expand to an approximately laminar flow across the heated susceptor, which is tilted by a small amount (5–10°) to account for reactant depletion. Multiple wafers can be accommodated along the length of the susceptor, or even side by side if the susceptor width is large enough. Uniformity can be improved by incorporating a rotating disk within the otherwise rectangular susceptor. Barrel reactors, similar to those used extensively throughout the silicon industry for silicon CVD, are occasionally used in MOCVD growth and are similar in operation to the horizontal geometry of Fig. 2(b). Fig. 3 shows a simplified schematic diagram of a commercial vertical-reaction-chamber geometry [27]. This multiple-wafer chamber has a gas-mixing region above a diffuser [28] that forces a uniform normal gas flow. The susceptor sits on a rotating sleeve that isolates the heater from the process gases. A multistage heater block is used to introduce a lateral temperature gradient for improved uniformity.

D. Safety

Whether or not hydride sources are used, safety is of paramount importance in the design and operation of MOCVD growth apparatus [29]–[32]. Although a detailed description of MOCVD safety is beyond the scope of this paper, the main risks and most common ways to handle them can be outlined. Hydrides pose the biggest risk because they are high-pressure toxic gases. The alkyls pose the next highest risk because, although they are toxic and pyrophoric, they are liquids and generally easier to handle. Some of the effluent gases cause risk because they can contain elemental phosphorus and arsenic as well as the oxides of these elements. Ancillary risks include quartz reaction chambers (which are breakable), large volumes of explosive hydrogen gas, high temperatures, and the acids and solvents used for preparing for, and cleaning up after, a growth run.

Handling these risks, in general terms, falls into three categories and applies equally well to all epitaxial growth processes and, indeed, to all semiconductor processing activities. The first is limited access. Access should be limited in a staged or “onion peel” manner, such that each level of greater access to potentially harmful materials requires a higher level of authority and a correspondingly higher level of training. The second category is training. Thorough training in the safe handling of materials must be required, and emergency-response situations should be considered and appropriate plans made. Plans and training
should involve not only those with routine access but also those likely to respond in an emergency, such as a local hazardous-materials response team. The third category is a hardware safety infrastructure. Hazardous gases should be located in a relatively remote area and held in exhausted safety cabinets with automatic cylinder-change hardware and both mechanical and electronic flow-rate sensing and limitation. Process gases should be piped in leak-free coaxial tubing with an inert purge gas or vacuum in the outer tube. Adequate exhaust hoods and cabinets should be used to contain the growth apparatus itself. Waste gases should be processed with filters, combustion discharge, oxidation, wet chemical scrubbing, or a combination of these methods. Automatic shutdown of source gases and a switch to inert purge gases should take place in the event of power failure when inadequate backup power is available. Perhaps most important, adequate toxic- and flammable-gas monitoring should be provided at all points along the process chain, including attics, halls, labs, cabinets and hoods, purge lines, and effluent gas lines.

III. GROWTH MECHANISMS

It is true for MOCVD growth, and for any growth process that takes place in a regime of viscous flow, that the hydrodynamics of the reactor geometry play a key role in the nature of the process and must be understood. This understanding is complicated by the consideration of thermal and chemical effects. While it is not possible to provide an exhaustive treatment of these subjects here, it is useful to outline the basic hydrodynamic, thermal, and chemical processes and then extract important conclusions about such practical concerns as growth rate, composition, and heterostructure interface quality.

A. Hydrodynamics and Boundary Layers

Even in the absence of a heated susceptor or reactive chemical species, practical reactor geometries exhibit fairly complex flow patterns [33]. Understanding the detailed hydrodynamics of such real systems is usually best accomplished by relying on modern numerical analysis methods. On the other hand, remarkable physical insight into the MOCVD process can be gained by considering the hydrodynamic behavior of simplified systems. In the regime of viscous flow, the boundary condition that the gas velocity must be zero at all solid surfaces is the critical issue. Consider the flow of some inert carrier gas without any reactive species present in a cold reaction chamber. Fig 4(a) shows a view of the evolution of the gas velocity profile in a simplified reactor geometry that approximates the horizontal-reaction-chamber geometry of Fig. 2(b). If a gas flow of uniform velocity is incident on the entrance to a tube (or infinite parallel plate) geometry, the zero-velocity boundary condition quickly begins to affect the velocity profile, as shown on the left in Fig. 4(a). At some distance along the path—which may be large and depends on the gas, its total velocity, and the dimensions of the chamber—the velocity profile stabilizes, as shown in the center of Fig. 4(a). In practical reactors, a large mismatch often exists between the inlet tube, which typically is a standard tube size of less than 10 mm diameter, and the characteristic dimension of the reaction chamber. Thus, even in a relatively simple horizontal reaction chamber, the gas may have to travel well into the chamber before a simple parabolic profile stabilizes.

The region nearest any surface, where the velocity profile drops rapidly and nearly linearly to zero, is referred to as the boundary layer or, inaccurately, the stagnant layer. The gas everywhere in the boundary-layer region has a finite velocity, although it may be quite small compared to the velocity elsewhere in the chamber. Normal MOCVD growth processes generally involve large carrier gas flows combined with relatively small source gas flows. The basic MOCVD reactions are pyrolysis reactions; so, with no region of the reaction chamber being heated to a temperature sufficient to drive the reaction, the overall velocity profile in the chamber is largely defined by the carrier gas flow rate and geometry, and the constituent gas-phase concentrations are uniform.

B. Thermal Effects

The addition of a heated susceptor, shown as the shaded region in Fig. 4(a), introduces several important complications, some essential and others undesirable. A heated zone is necessary to drive the pyrolysis reaction and provide for the desired materials deposition. The temperature gradient between the susceptor and the chamber ambient can be very large, often several hundreds of degrees Celsius. For example, the optimum MOCVD growth temperature for many III–V compounds and alloys falls in the range of 600–800°C (the column III nitrides, such as GaN, may be grown at much higher temperatures), while air or water cooling is often used to maintain the chamber walls at temperatures close to room temperature. A large temperature gradient distorts the velocity profile from a simple parabola,
as shown on the right side of Fig. 4(a), which results in a somewhat steeper velocity profile in the boundary-layer region. More important, a large temperature gradient is likely to cause detrimental buoyancy effects such as large convection or recirculation cells. These can increase the effective residence time for the reactive constituent or dopant species and contribute to, for example, less abrupt heterostructure interfaces. In terms of growth mechanisms, the presence of a heated surface results in a sink for reactive constituents and leads to concentration gradients that define the growth rate for the deposited material. However, when the temperature is high enough, relative to the carrier gas velocity and thermal conductivity, the gas far above the surface becomes hot enough to cause detrimental, parasitic, gas-phase pyrolysis reactions.

C. Pyrolysis and Reaction Mechanisms

The basic mechanism driving reaction (1) is a pyrolysis reaction in which the individual constituents decompose to form the desired nonvolatile deposited layer and volatile reaction products. Of course, the process is not that simple [34]–[39] and is complicated by a number of considerations. These include the possibility of intermediate reactions as the source molecules fragment step wise, e.g.,

\[(\text{CH}_3)_3\text{In} \rightarrow (\text{CH}_3)_2\text{In} \rightarrow (\text{CH}_3)\text{In} \rightarrow \text{In}\]  \hspace{1cm} (8)

the possibility of catalysis at the different surfaces present (quartz, semiconductor, silicon carbide-coated graphite), the relative nonvolatility of some reaction products, especially carbon and carbon-containing compounds, and possible interactions with intentional or unintentional impurities. Fig. 5 shows the pyrolysis efficiency for some of the key alkyl and hydride sources for MOCVD growth [14], [40]–[47]. The pyrolysis efficiencies of the column V hydrides, shown in Fig. 5(a), are strongly affected by the presence of a catalyzing surface. The data for the curves in Fig. 5(a) for AsH\textsubscript{3} (solid squares) and PH\textsubscript{3} (solid circles) were measured in the presence of the logical corresponding binary substrates GaAs and InP, respectively. Shown for reference as a dashed line (open circles) in Fig. 5(a) is the pyrolysis efficiency of PH\textsubscript{3} measured in a quartz tube without an InP substrate present. The effect of the P-containing surface is enormous, reducing the effective pyrolysis temperature by more than 300°C. Similar results have been observed for AsH\textsubscript{3} with and without a GaAs catalyzing surface present.

Fig. 5(b) shows the pyrolysis efficiencies for two column III metal alkyl sources, TMGa and TMIn. The behavior of these is typical of all of the column III alkyls. Note that the effective temperatures for complete pyrolysis are somewhat lower for the alkyl sources and that the slopes are somewhat steeper, defining a fairly narrow temperature range for the transition from zero to complete pyrolysis. In addition, no significant change is observed in the pyrolysis efficiency of these alkyl sources in the absence or presence of a catalyzing surface.

Enough of the critical components have been described to allow us to outline in some detail how the deposition process works. A velocity profile and boundary layer have been established, largely defined by the geometry of the chamber and the carrier gas hydrodynamics. The high-temperature heated susceptor has introduced a temperature gradient and promotes the pyrolysis reaction and provides a sink for constituent source material, leading to a concentration gradient above the susceptor. The velocity, temperature, and concentration gradients define a basic gas-phase diffusion process. Since the pyrolysis natures of the column III alkyls and column V hydrides are somewhat different, because of surface catalysis, we can speculate on the deposition mechanisms. As the column III alkyl molecules diffuse toward the surface, they encounter higher temperatures and lower gas velocities (longer residence time), with the result that fragmentation likely begins above the surface and proceeds step wise as described in (8). A similar process occurs for the column V hydrides, except that the strong surface catalysis effects described above argue that these molecules remain relatively intact until they adsorb on the surface. This process is shown schematically in the enlargement of the susceptor surface shown in Fig. 4(b).

D. Growth Regimes

The MOCVD growth process, whether involving column V sublattice alloys or not, generally is performed under conditions of nonstoichiometric, excess, and usually uninterrupted column V species gas flows. This compensates
Fig. 6. Growth rate as a function of temperature showing the three distinct growth-rate regimes.

for the somewhat lower pyrolysis efficiencies and greater temperature dependence of pyrolysis efficiency at typical growth temperatures [see Fig. 5(a)]. More important, it compensates for the much greater volatility of the column V atoms in the desired solid films. Thus, the growth rate of MOCVD epitaxial layers is defined by the column III alkyl constituents. We can define three fairly clear growth regimes [48], as shown schematically in Fig. 6. The absolute growth rate depends on too many chamber- and process-specific parameters to quantify, so an arbitrary relative scale is used in Fig. 6. The first of these three regimes, at relatively low temperatures, is a region of reaction-rate limited growth. The alkyl pyrolysis efficiency shown in Fig. 5(b) usually is steep, implying that the reaction-rate limited growth regime should only occupy a narrow temperature range.

Once unity pyrolysis efficiency is reached, generally for all temperatures greater than 500°C or so, a regime of mass transport limited growth is established [37], [49]–[51]. This region is usually several hundred degrees Celsius. In strict terms, diffusion plays a role in this region, but the gas-phase diffusion coefficient of the alkyls is only weakly dependent on temperature and is moderated by temperature changes in the boundary layer. In any case, the input source rate defined by the user has a much greater dynamic range. At some still higher temperature, the temperature of the gas becomes high enough far enough above the surface that gas-phase pyrolysis of the hydrides becomes important and solid particulate can form without depositing on the substrate. This parasitic spontaneous nucleation [52], along with increased desorption of the reactant species, takes place at the expense of the desired deposition and leads to a reduced growth rate, as shown in Fig. 6.

E. Growth Rate and Composition of Alloys

If we limit the discussion to the region of mass transport limited growth, it is possible to use the conservation of matter and the gas law to establish functionalities for both the growth rate and composition of III–V semiconductor compounds and alloys by MOCVD. Under conditions of excess column V constituent and unity alkyl pyrolysis efficiency, all of the alkyl molecules supplied to the system will yield solid material in some effective surface area. The effective surface area is chamber specific and should be on the order of the area of the heated susceptor. In practice, the effective area typically is approximately three times the actual heated susceptor area. As long as there are no large changes in carrier gas flow rate or growth temperature, this area can be considered a constant. Thus, the growth rate $g$ for a binary compound (1) is given by

$$g = \frac{MW_A}{kTd_Ae} p_A F_A$$

where $p_A$ is the vapor pressure of the column III alkyl and $F_A$ is its carrier gas flow rate. $MW_A$ is the molecular weight of the compound, $k$ is Boltzmann’s constant, $T$ is the temperature, $d_A$ is the solid density, and $e_A$ is the effective area. The same conditions hold for a semiconductor column III sublattice ternary alloy (5), so the growth rate $g$ for the alloy is given simply by the sum of the binary growth rates

$$g = \frac{MW_A}{kTd_Ae} p_A F_A + \frac{MW_B}{kTd_Be} p_B F_B.$$  

Development of a functional form for the composition of ternary column III sublattice alloys is even simpler. The unity pyrolysis efficiencies of the alkyl constituents imply that the solid phase composition will be identical to the ratio of the growth rate of constituent $B$ to the total growth rate [51]. Thus, the solid composition $x$ is given by

$$x = \frac{1}{1 + p_B F_B MW_B d_B p_A F_A MW_A d_A}.$$  

There may be other integer prefactors for the vapor-pressure flow-rate product terms in (10) and (11) if any of the alkyl sources is other than a monomer, as is the case for TMAl, which is a dimer.

An example of alloy compositions and growth rates for ternary Al$_x$Ga$_{1-x}$As is given in Fig. 7. The horizontal axis in Fig. 7 is a metal alkyl flow ratio

$$F_A/[F_A + F_B]$$

in which only the routinely adjusted flow-rate terms are retained. The source vapor pressures are usually held constant to avoid modifying refrigerated bath temperatures, and the other terms are material-dependent constants. Fig. 7(a) shows the growth rate normalized to total metal alkyl flow rate, and Fig. 7(b) shows the solid composition $x$ of the resulting alloy layer. The solid lines in Fig. 7(a) and (b) correspond to (10) and (11), respectively, while selected experimental data points are shown as solid squares. Complete characterization of an MOCVD apparatus generally involves verifying these curves for specific desired compositions and specific growth temperatures. Source or chamber
IV. CONSIDERATIONS FOR OPTOELECTRONIC DEVICES

It would be a most difficult task to provide detail on the large number of different optoelectronic compounds and alloys that have been grown by the MOCVD process with different source combinations, process parameters, and reactor configurations. The best studied compounds and alloys—GaAs, InP, InGaAs, AlGaAs, InGaAsP, and InGaP—have become widely used commercial materials for such applications as compact disc (CD) player lasers, CD read-only memory, laser printer sources, 1.3- and 1.55-μm telecommunications lasers, 980-nm lasers for Er-doped optical fiber amplifiers, and red LED’s. These products, made inexpensively and in large numbers and often utilizing the MOCVD process for the growth of the basic structures, clearly prove the viability of the process. So in this section, rather than elaborate on a few examples, we will generalize the necessary range of device structure parameters and address how well the MOCVD process fares in meeting the requirements.

A. Materials, Structures, and Doping

The quantum-well heterostructure (QWH) laser can serve as a useful tool in outlining the range of materials parameters necessary for optoelectronic devices [13]. Thicknesses in QWH lasers range from as little as 50 Å for the quantum wells and barriers to one micrometer or more for optical waveguide confining layers. Doping in these different layers can vary from unintentionally lightly doped quantum wells—for example, in which the background doping ideally is very low (≤10^15 cm⁻³)—to moderately doped confining layers—for example, in which the background doping ideally is quite high (2 × 10^18 cm⁻³)—to highly doped, such as contact layers, in which the doping is as high as possible. The composition of any of the layers in a QWH laser can vary widely, although generally there are only a few different compositions in any given laser structure. The well size and composition together define the emission wavelength of the laser, while the compositions of the other layers in the structure play a role in defining an optical waveguide. A less clearly delineated characteristic of a laser structure is optical loss, which arises unavoidably from intentional doping as well as from the basic quality (dislocations, background impurities) of the epitaxial material.

A wide range of growth rates is easily available from the MOCVD process. The limits on growth rate arise from the alkyl saturation limits and the range, controllability, and reproducibility of the electronic mass flow controllers used to supply the alkyl constituents. With no special modifications, an MOCVD reactor can reliably provide growth rates in the range of a few angstroms per second to several tens of angstroms per second, all within the growth of a single structure. Thus, the different steps in the process can be optimized separately, with the quantum wells and barriers being grown at a very low growth rate while the thicker confining layers are grown more rapidly. There are a number of suitable dopant sources for MOCVD growth: hydrides...
such as silane or alkys such as diethylzinc. These dopants have the same process requirements as described above for constituent sources, with the additional requirement of well-controlled dilution. Compositional control generally is simple to obtain, with the exception of very low or very high composition alloys when the system has been designed for midrange compositions. For these points near the limits, there are very large gas-phase constituent ratios, and within the limits of the mass flow controllers, controllability or reproducibility may be more difficult to obtain.

**B. Heterostructure Interfaces**

A critical issue for the growth of quantum-well heterostructures and superlattices, by any growth method, is the quality of the heterostructure interfaces. Abrupt interfaces become more and more important as the thickness of the layer becomes smaller. Even for thicker heterostructures of materials, where there is tight tolerance on the lattice-mismatch composition, graded interfaces are to be avoided. This concern arises for gas-phase growth processes such as MOCVD for two different fundamental reasons. The first reason is the large amount of gas moving in the reaction chamber at relatively slow speeds. Thus, even abrupt pulses of constituent gases inevitably are going to broaden by dispersion. This problem can be exacerbated by recirculation cells inside the reaction chamber. The second reason is dead space in either valves or tubing, which can introduce transients if the amount of dead space is significant in comparison with the total amount of injected constituent material.

There are good solutions to these problems. Properly designed reactor gas-panel architectures and high-quality modern zero dead-space valves minimize, though do not completely eliminate, the dead-space problem. Reaction chambers need to be carefully designed, and operated in an appropriate hydrodynamic range, to minimize or eliminate recirculation. If these problems have been resolved, the last problem of gas-phase dispersion can be minimized by using an interrupted growth timing sequence. Only the column III species are interrupted, so that the out-diffusion of the more volatile column V elements from the solid substrate or grown layer is avoided. Interrupted growth has been demonstrated to be successful within a limited range. If the pause time is too short, then dispersion effects continue to be manifested. If the pause time is too long, the possibility for contamination of the surface from unwanted impurities in the column V source gases increases. Typical pause times are 5–20 s. When these growth-process solutions are implemented, very abrupt interfaces can be grown by the MOCVD process [23], [53]–[62]. Fig. 8 shows a high-resolution transmission electron micrograph [57] with corresponding lattice image of a GaAs-AlAs superlattice grown by MOCVD having 17-Å layer thicknesses. The interfaces are regular and abrupt, without undulation.

Ideally, a fully automatic MOCVD reactor system would provide some mechanism for acquiring real-time in situ data on the instantaneous thickness and composition of the growing layer. These data would be processed and used in a closed-loop feedback control configuration directly and precisely to control the thickness and composition. The user would simply define a composition and thickness rather than specific growth times and constituent supply rates, as is presently done. This is especially important for such applications as multiple quantum wells and stacked dielectric mirrors for vertical-cavity surface emitting lasers.
(VCSEL’s). The key to a closed-loop system is data acquisition, but the environment in an MOCVD reaction chamber precludes the use of most standard surface-analysis techniques. A number of promising optical approaches [63], [64] have been studied, however, all having the advantage of using light sources that can be located outside the reaction chamber. These are, for example, reflection difference spectroscopy, which captures differences in the amplitude of light reflected from the growing surface, and surface photoabsorption, which is based on the reflectivity of p-polarized light incident at the Brewster angle and is sensitive to chemical coverage on the surface. Future closed-loop MOCVD systems will likely be designed on the basis of these kinds of remote optical measurements.

V. Selective Area Epitaxy

The MOCVD growth process described in the preceding sections has become a mainstay in the commercial manufacturing of discrete semiconductor lasers. The process has continued to evolve in the last decade, with more sophisticated commercial reactors and better, purer sources. It can be argued that the most important advances in MOCVD growth research and development in the past decade have been in the application of this process to new materials, such as the wide-gap group III nitrides, described by S. DenBaars in this special issue.2 Another important advance in the growth process that seems likely to have enormous impact on future optoelectronic devices or, more accurately, integrated photonic device, is selective area epitaxy (SAE) or selective area growth (SAG). The biggest limitation to MOCVD, or any conventional epitaxial growth process, is that no matter how fine the resolution or how high the quality, the growth is uniform over the entire wafer. This is inconsistent with different optimum design structures for the optical componentry, such as lasers, amplifiers, modulators, detectors, and routing structures, that might make up a photonic integrated circuit. One solution is to design a compromise layer structure. Another is to grow different stacked layered structures and then use etching techniques to access them individually. A third is to grow one form of structure, etch part of it away, grow another kind of structure, etch part of it away, and so on until the desired functionality is obtained. It would be most desirable to have a single growth process that yields the desired structure at any area on the wafer. SAE [65]–[85] most closely approximates this ideal.

A. SAG and Modeling

The optical properties of most photonic devices are defined by the transition energies in the device, which in turn are generally defined by the quantum-well thicknesses and compositions. Thus, if we can control these thicknesses and compositions anywhere on the wafer, we can define different optimum devices where we want them. SAE makes use of an insulating mask, such as silicon dioxide, to inhibit the deposition of material wherever the mask is present. As a result, the deposition elsewhere must be enhanced simply because of the conservation of mass. The enhancement profile will reflect both the mask pattern on the wafer as well as gas-phase and surface-diffusion effects in the reaction chamber. At first glance, it appears that the obvious pattern of choice is to mask the entire wafer and open windows of different sizes and grow separate regions designed for the device intended to occupy the region. This is not the best solution, however, for several reasons. For example, if the size of the masked area is too large, then the growth rates in the unmasked areas are too large and perhaps uncontrollable. Also, optical interconnection of the separate devices is not straightforward.

A much better, more practical solution is to use the mask pattern shown schematically in Fig. 9(a). This dual-stripe mask pattern has a number of useful features. The opening between the stripes ultimately will be the active region of the device, and the width of this opening can be used to define a lateral waveguide because there will be a step in the index of refraction at the edges of the stripe opening. As we discuss below, typical stripe openings are on the order of a few micrometers, which is much smaller than the effective diffusion lengths in the growth process. This means that the growth inside the stripe opening will be essentially uniform, as shown in Fig. 9(b). The growth rate in the stripe openings will be determined by the oxide mask widths, with wider stripes yielding thicker layers (and thicker quantum wells) for the same growth time. The stripe opening is held constant for different devices so that creating an optical path is trivial. Last, the relatively large areas outside the

---

2 See pp. 1740–1749.
Fig. 10. Outline of numerical models for selective area epitaxy. (a) Basic unit-cell cross section for the dual stripe pattern. (b) Example of a computational mesh used for numerical simulations. (c) Isoconcentration profiles calculated using the profiles and meshes shown.

dual stripe region allow significant deposition to occur, with the enhancement decaying to zero far away from the oxide stripes, as shown in Fig. 9(b). This has the effect of softening the amount of enhancement between the stripes to controllable levels of less than a factor of two or three.

Extensive numerical modeling of the selective area epitaxy process has been reported. The kinds of information that can be obtained are outlined in Figs. 10 and 11 [74], [75], [82]. Fig. 10(a) shows the basic unit-cell cross section for the dual stripe pattern of Fig. 9. At some point well above the substrate at the edge of the boundary layer, the concentration of the constituent species is constant and equal to the supply rate. The unit cell must be chosen to be wide enough, in terms of effective diffusion lengths, that the slope of the concentration everywhere along the sides is zero. Since no deposition takes place on the oxide mask, the slope of the concentration must be zero at the mask surfaces. The slope everywhere else along the surface is the growth rate and is necessarily a function of position. An example of the computational mesh that is used for numerical simulations is shown in Fig. 10(b). To optimize computation time and accuracy in the regions of greatest interest, the density of the mesh is designed to be greatest nearest the surface and between the oxide mask stripes. The results of such a simulation are shown in Fig. 10(c). These are isoconcentration profiles, which show pronounced lateral concentration gradients away from the center of the oxide mask stripes and toward the unmasked regions.

More immediately relevant information can be obtained from Fig. 11, which is the selective growth-rate profile of GaAs, normalized to the background (unmasked) growth rate, for two samples having a 4-μm opening between stripes and oxide mask stripe widths of 6 μm (solid line) and 25 μm (dashed line). These data are taken from the slope of the concentration data at the surface of Fig. 10(c). The wings of Fig. 11 show that the effective diffusion length is tens of micrometers and much larger than the stripe opening width. The model indicates less than a 1% thickness variation in the stripe opening. Experimentally, none is observed. The range of growth-rate enhancement is from one to three times. Consider a base quantum-well thickness of 50 Å. The SAE process would then give a controllable quantum-well thickness in the range of up to roughly 150 Å, which should be adequate for most photonic device designs. Note that the oxide mask stripe widths are well within the range of conventional optical lithography, and no special precision is required.
B. Optoelectronic Devices by SAE

The basic element for SAE integrated photonic devices is the buried heterostructure laser. A scanning electron micrograph of the cross section of such a buried heterostructure laser is shown in Fig. 12. In spite of the midgrowth processing steps and regrowth interfaces, these lasers are of remarkably high quality. The lateral waveguide is formed by the stripe opening, and the emission wavelength is defined by the base quantum-well thickness and composition along with the oxide mask stripe width. Any laser emission wavelength, or multiple wavelengths, can be chosen within a single growth run on the same wafer by simply adjusting the oxide mask stripe width. The kinds of ranges available from this process are shown in Fig. 13 [72], [75], which shows the emission wavelength versus oxide mask stripe width for both long-wavelength [1.55 μm, Fig. 13(a)] and shorter wavelength [980 nm, Fig. 13(b)] SAE lasers. By simply defining the oxide mask stripe width appropriately anywhere on the wafer and including appropriate biases and electrical isolation, the wavelength can be adjusted to form red-shifted detectors, slightly blue-shifted electroabsorption (EA) modulators, or transparent routing structures.

One example of an integrated device formed in this manner, a laser-modulator element [85], is shown schematically in Fig. 14. The longitudinal cross section is shown in Fig. 14(a), and the corresponding mask pattern is shown in Fig. 14(b). The laser section is formed where the oxide mask stripe width is largest, and a slightly narrower mask width is used to form a blue-shifted EA modulator, which will be transparent at zero bias. The laser resonator is formed by a cleaved and coated facet and a distributed Bragg reflector grating. Several of these elements can be connected with transparent routing structures and Φjunction couplers to form a monolithic integrated multiwavelength single-output transmitter.

VI. SUMMARY

We have attempted to capture the essence of the MOCVD process for deposition of compound semiconductors and alloys by considering the basic chemical reaction equation and then extending to more complex situations, such as all-alkyl processes and alloys. The hardware comprising MOCVD reactors has been described, including the reactor gas delivery system, the reaction chamber, and the reactor safety infrastructure. We have outlined the basic hydrodynamic, thermal, and chemical processes and then extracted...
important conclusions about such practical concerns as growth rate, composition, and heterostructure interfaces. We have generalized the necessary range of photonic device structure parameters and addressed how well the MOCVD process fares in meeting the requirements. An important advance in the growth process that seems likely to have enormous impact on future integrated photonic devices is SAE. We have presented an introduction to this process, which most closely approximates the ideal of a single growth process that yields any desired structure at any area on the wafer.

It is worth adding a few words to address the competition between MOCVD and MBE for designation as the technology of choice for the growth of compound semiconductor heterostructures. These two growth processes, which are fundamentally very different, appeared and matured at nearly the same time. Both have proponents that can articulate the relative advantages and disadvantages of the respective processes. Two strong conclusions can be made concerning this competition. The first is that, in the hands of experts, both processes can produce similar results, and the limitations are generally those fundamental to the materials chosen rather than the process utilized. The second conclusion is that the development of each process has been accelerated by advances in the other, to the point where both processes, as well as the technical community in general, are the better for the competition.

ACKNOWLEDGMENT

The author wishes to express his gratitude to S. G. Bishop, P. D. Dapkus, A. M. Jones, R. M. Lammert, L. C. H. Laurin, A. Melas, M. K. Pratt, and S. D. Roh for many technical discussions and for their contributions to this manuscript.

REFERENCES


James J. Coleman (Fellow, IEEE) received the B.S., M.S., and Ph.D. degrees in electrical engineering from the University of Illinois, Urbana, in 1972, 1973, and 1975, respectively. From 1976 to 1978, he was with Bell Laboratories, Murray Hill, NJ, where he studied the materials properties of InGaAsP grown by liquid phase epitaxy (LPE) and helped develop proton-isolated room-temperature continuous-wave (CW) lasers operating at wavelengths near 1.3 μm. From 1978 to 1982, he was with Rockwell International, Anaheim, CA, where he contributed to the development of MOCDV-grown CW room-temperature, low-threshold AlGaAs-GaAs laser devices. Since 1982, he has been a Professor with the Microelectronics Laboratory, Department of Electrical and Computer Engineering, University of Illinois. He and his students are involved in the study of the materials growth and processing of various structures such as quantum-well heterostructures, superlattices, strained-layer InGaAs-GaAs lasers, laser arrays, and integrated photonic devices by selective area epitaxy. He is an Associate Editor of Photonics Technology Letters.

Prof. Coleman is a Fellow of the Optical Society of America and the American Association for the Advancement of Science. He has been Guest Editor for two special issues of the IEEE JOURNAL OF QUANTUM ELECTRONICS. He is a LEOS Distinguished Lecturer for 1997–1998.