The High Electron Mobility Transistor at 30: Impressive Accomplishments and Exciting Prospects

Jesús A. del Alamo

Microsystems Technology Laboratories, Massachusetts Institute of Technology, Cambridge, MA 02139
alamo@mit.edu; T. +1-617-253-4764

Keywords: high electron mobility transistor, HEMT, III-V, two-dimensional electron gas, modulation doping

Abstract
2010 marked the 30th anniversary of the High-Electron Mobility Transistor (HEMT). The HEMT represented a triumph for the, at the time, relatively new concept of bandgap engineering and nascent molecular beam epitaxy technology. The HEMT showcased the outstanding electron transport characteristics of two-dimensional electron gas (2DEG) systems in III-V compound semiconductors. In the last 30 years, HEMTs have been demonstrated in several material systems, most notably AlGaAs/GaAs and AlGaN/GaN. Their uniqueness in terms of noise, power and high frequency operation has propelled HEMTs to gain insertion in a variety of systems where they provide critical performance value. 2DEG systems have also been a boon in solid-state physics where new and often bizarre phenomena have been discovered. As we look forward, HEMTs are uniquely positioned to expand the reach of electronics in communications, signal processing, electrical power management and imaging. Some of the most exciting prospects in the near future for HEMT-like devices are those of GaN for high voltage power management and III-V CMOS to give a new lease on life to Moore’s Law. This paper briefly reviews some highlights of HEMT development in the last 30 years in engineering and science. It also speculates about potential future applications.

INTRODUCTION

The HEMT was first demonstrated by Mimura and colleagues at Fujitsu Labs in 1980 [1]. The invention of the HEMT represented the latest triumph of bandgap engineering and molecular beam epitaxy that had earlier brought to the world the heterostructure laser, the heterojunction bipolar transistor, heterostructure avalanche photodiodes and other electronic and photonic devices that would come to revolutionize communication systems.

The HEMT was based on the concept of modulation doping first demonstrated by Dingle and collaborators at Bell Labs in 1978 [2]. A modulation-doped structure creates a two-dimensional electron gas at the interface between two semiconductors of different bandgaps. The spatial separation that is created between dopants and electrons confers these with a mobility that exceeds the bulk value even at relatively high carrier concentrations.

The first demonstrations of modulation doping and the HEMT took place in the AlGaAs/GaAs system. These would not have been possible without the atomic layer precision growth capabilities of molecular beam epitaxy. Through MBE, heterostructures approaching monolayer-level abruptness, nearly perfectly coherent interfaces and precise dopant placement became possible. All these are essential ingredients for realizing modulation-doped structures with improved mobility over bulk values. The initial mobility enhancements that were observed were quite modest at room temperature but much more significant at low temperatures. Over time, the low temperature mobility would dramatically improve leading to fundamental discoveries in solid state physics [3].

This paper traces the most significant steps of the evolution of the HEMT from its inception to its current state in various semiconductor systems. It also outlines some of the new discoveries that the development of high crystalline quality AlGaAs/GaAs heterostructures enabled in solid-state physics. It finishes by sketching a future for new applications for HEMT-like structures in several domains.

In this paper, the term HEMT is used with a slightly broader meaning than the one intended by its original inventors. Over the years, we have learned that not only modulation doping but also polarization doping and the application of a vertical electric field to a suitable heterostructure all result in strongly confined two-dimensional carrier systems with potentially very low scattering and high mobility. The term HEMT will be loosely used here to refer to device structures based on all these approaches.

HEMT ELECTRONICS

Right from its outset, the high mobility and highly confined nature of the two-dimensional electron gas suggested that modulation doping could be exploited to make high-speed field-effect transistors with excellent short-
channel effects and great scaling potential. This spurred efforts to demonstrate high-speed logic circuits. Progress was spectacular. Within one year, Fujitsu had demonstrated the first HEMT integrated circuit featuring an enhancement/depletion-mode logic ring oscillator that showed a switching delay of 17.1 ps, “the lowest of all the semiconductor logic technologies reported thus far” [4]. Soon the advantages of HEMTs over contemporary GaAs MESFETs and Si ECL in terms of speed-power performance become evident and Fujitsu and others launched a race to demonstrate HEMT circuits for high-speed computers. Fujitsu demonstrated 1Kbit SRAMs in 1984, 4 Kb in 1987 and 64 Kbit in 1991 [5]. Several companies demonstrated many kinds of circuits with various levels of complexity. In the end, in spite of interesting demonstrations in supercomputers and other systems, this first commercial direction for HEMT ICs did not pan out. This technology, just as Si ECL, could not compete with the increasing densities, much lower power consumption, and huge economies of scale of Si CMOS.

The first commercial application of a HEMT came from a totally unexpected area. In 1983 Fujitsu demonstrated a four-stage HEMT amplifier operating at 20 GHz [6]. Its gain and noise performance was observed to improve significantly as the operating temperature was reduced. This was attributed to the improvement in mobility at low temperatures. These cryogenically cooled HEMTs significantly outperformed state-of-the-art GaAs MESFETs of the time. This result came to the attention of the Nobeyama Radio Observatory in Nagano (Japan) who ordered several low-noise amplifiers for their 45 m radio telescope. These were installed in 1985 and in 1986, this telescope discovered new interstellar molecules in the Taurus Molecular Cloud, about 400 light years away [5].

The great stability of the HEMT, compared with the then conventional parametric amplifiers, allowed prolonged observations, something that had great value to astronomers. Since then, HEMTs have been widely used in radio telescopes throughout the world [7].

The first mass-market applications of AlGaAs/GaAs HEMTs came in the communications arena after the recognition of their superior high-frequency noise characteristics over MESFETs. This was attributed to the excellent fT and the high aspect ratio of the channel. Direct Broadcasting Satellite receivers including low-noise HEMT amplifiers first went commercial in 1987. They enabled a reduction in antenna size by one half. By 1988, the yearly world-wide production of HEMT receivers was about 20 million [8].

Two key innovations in the mid 80’s would greatly expand the high-frequency capabilities of GaAs-based HEMTs and would open up many new application areas in communications and radar systems. In 1985, Ketterson and colleagues at U. Illinois demonstrated the Pseudomorphic HEMT, or PHEMT [9]. This is an AlGaAs/InGaAs/GaAs quantum-well structure where the enhanced transport properties of electrons in InGaAs coupled with the tight quantum-well confinement and large conduction band discontinuity between channel and barrier resulted in improved device scalability and performance. A second innovation was planar or delta doping of the AlGaAs barrier [10]. This allowed the thinning down of the barrier yielding improvements in transconductance, channel aspect ratio and device scalability. The removal of dopants from directly underneath the gate also improved the breakdown voltage of the device. The combination of planar doping and pseudomorphic InGaAs channel propelled PHEMTs to the 0.1 μm gate length regime and the achievement of record noise and power performance up to very high frequencies.

The extraordinary capabilities of the PHEMT in terms of noise, power, and low-loss switching at very high frequencies have made this device technology a success in the commercial arena. Some of the most notable applications are cell phones, broadband wireless, satellite communications, instrumentation, cable TV, DBS, GPS, space, radar, fiber optic communications, sensing, military, and many others.

PHEMT IC technology is now manufactured around the world on 6” GaAs wafers and is available in foundry mode from several companies. Following a path similar to Si, new PHEMT IC technologies are emerging that monolithically integrate enhancement- and depletion-mode devices, HBTs and high quality passive elements to enable enhanced functionality single-chip systems [11]. These will make possible the development of highly integrated front-end modules for high volume cellular phone applications.

Market estimates for GaAs HEMT MMICs show a rapid rise in the last ten years with an expected volume of $1.2B in 2011 (Fig. 1). There is an additional estimated $100M market for discrete HEMTs in 2011 [12]. The market segmentation for GaAs HEMT MMICs in 2009 is shown in Fig. 2 (total estimated market value $944M) [12]. The biggest sector for GaAs HEMT MMICs is wireless communications followed by military applications. The most significant applications, in order of importance, are cell phone handsets, WLAN, base stations and CATV [12].

In parallel with progress in GaAs HEMT technology, modulation doping and HEMTs were also demonstrated in other material systems, most notably, InAlAs/InGaAs [13,14], AlSb/InAs [15,16], AlGaN/GaN [17,18], and even SiGe/Si [19,20]. Similarly, modulation-doped two-dimensional hole gases and High Hole Mobility Transistors have also been realized in several heterostructures systems such as AlGaAs/GaAs [21,22], AlGaAs/InGaAs [23,24], InAlAs/InGaAs [25,26], Si/SiGe [27,28], AlAsSb/GaSb
This spurred the co-integration of n-channel and p-channel devices which was achieved in the AlGaAs/GaAs [32] and InAlAs/InGaAs systems [33] and the demonstration of complementary circuits. Some of these circuits attained amazing levels of integration, in excess of $10^5$ transistors [34].

InAlAs/InGaAs HEMTs on InP have particularly excelled at the extremes of noise and frequency. Increasing the InAs content of InGaAs rapidly improves electron transport properties and increases conduction band discontinuity which results in enhanced carrier confinement. This unique set of attributes confers ultra-scaled pseudomorphic InAlAs/InGaAs HEMTs on InP with outstanding low noise and high-frequency characteristics. These are in fact the first transistors of any kind in any material system to simultaneously exhibit $f_T$ and $f_{max}$ values in excess of 640 GHz [35]. Particularly remarkable is the fact that this result was obtained at an operating voltage of 0.5 V, showcasing the outstanding low power potential of this technology. Separately, $f_{max}$ values in excess of 1 THz [36,37] and a noise figure as low as 0.71 dB at 95 GHz [38] have also been demonstrated in this material system. Fig. 3 shows the evolution of the current-gain cut-off frequency in GaAs MESFETs, GaAs PHEMTs and InP HEMTs as a function of time.

InGaAs HEMTs on InP are uniquely suited for mm and sub-mm wave applications in radar, radio astronomy, high-resolution imaging arrays, high capacity wireless communications, high data rate optical fiber systems and sub-mm wave spectroscopy. A variety of sub-mm wave amplifiers and other circuits have been demonstrated operating at frequencies that approach 500 GHz [39,40]. A 10 Gb/s wireless link operating in the 120 GHz band was demonstrated at the 2008 Beijing Olympics were live uncompressed high-definition video was transmitted over a distance of 1 km [41,42].

Similar InAlAs/InGaAs HEMTs with high InAs channel content have also been metamorphically grown on GaAs substrates [43,44]. These metamorphic HEMTs, or “MHEMTs”, provide comparable performance with improved manufacturability (e.g. larger wafer size, improved reproducibility, greater ease of wafer handling, GaAs backside process flow), lower cost, and better established packaging technology. While InP HEMT technology is not yet commercial, InAlAs/InGaAs MHEMT processes have been available in foundry mode since 2003 [45]. A wide range of circuits with performance that rivals the best InP HEMT technologies have been demonstrated, including sub-mm wave amplifiers that operate at 460 GHz [46].

One of the most exciting device structures at the present time is the AlGaN/GaN HEMT. Modulation doping in this system was first demonstrated in 1992 [17] and a working transistor was first reported in 1993 [18], in both cases by
Khan and collaborators at APA Optics. The enormous current interest in this system stems from several factors. First, the wide bandgap of GaN yields a high breakdown voltage which enables the high voltage operation of transistors. Also, the saturation velocity of electrons in GaN is over a factor of two higher than Si, which is valuable for high frequency operation. In addition, in a HEMT structure, a high electron concentration is induced at the AlGaN/GaN interface yielding high current. The confluence of all these factors results in devices with unique power amplification attributes at frequencies all the way up to the millimeter-wave regime. In fact, recently a power level of over 933 mW has been reported at 94 GHz [47] which is over a factor of two higher than the previous record established by InAlAs/InGaAs HEMTs [48]. A perspective on the progress of power amplification by HEMTs in the 94-95 GHz range is shown in Fig. 4.

AlGaN/GaN HEMTs have been in volume production since 2008 and are starting to appear in cellular phone and WiMax base stations. Many more applications are ripe for deployment such as radar systems, cable TV amplifiers, satellite communications and a variety of military systems [49].

Towards this goal, a great deal of effort has been devoted to eliminating all forms of scattering and improving the mobility of electrons at the AlGaAs/GaAs interface. A continuous stream of innovations in MBE technology brought the low temperature mobility from an early value of around 1.5x10^{4} cm^{2}/V.s to a current record value of 3.5x10^{7} cm^{2}/V.s [50]. This is the highest mobility on record in any semiconductor structure and attests to the extreme purity and perfection of the materials involved.

As the mobility increased, a veritable treasure trove of new and unexpected physics emerged that was previously hidden by disorder [3]. The fractional quantum Hall effect (FQHE) was discovered on a sample with a mobility of 9x10^{5} cm^{2}/V.s [51]. Zero resistance states in magnetotransport induced by microwaves were discovered in a sample with a mobility of 3x10^{6} cm^{2}/V.s. Interactions between composite fermions (bound states of electrons and quantized vortices introduced to explain the FQHE) were first postulated on a sample with a mobility of 1x10^{7} cm^{2}/V.s. A very fragile even-denominator 5/2 quantum Hall state was discovered on a sample with a mobility of 3.1x10^{7} cm^{2}/V.s. This is speculated to possibly reveal the existence of “non-abelian quasi-particles” that hold potential for future quantum computers.

In parallel with this effort, using gates of various shapes on the surface of AlGaAs/GaAs heterostructures and applying appropriate potentials, the 2DEG can be carved into electron distributions of capricious geometries. In this way, 1DEG (quantum wires), 0DEG (quantum dots) and many other structures have been created where electrons are coherent and show prominent quantum phenomena over very long distances. In quantum wires, for example, quantized conductance has been observed [52]. Quantum dots exhibit a discrete set of energies, much like in an artificial atom. Field-effect induced quantum dots can be made so small that they are able to contain just one electron [53]. Electrons can be induced to tunnel in and out of this artificial atom one at a time in what is called a single-electron transistor [54].

These findings already go beyond physics. The fact that the Hall resistance plateaus in the IQHE are only linked to fundamental constants (Planck’s constant and the charge of the electron) and the extreme accuracy with which the Hall resistance can be measured has led to a new international standard for resistance. The Ohm today is defined through a measurement of the Hall resistance of the first plateau in an array of quantum Hall bars fabricated in AlGaAs/GaAs heterostructures [55]. A precision of a few parts in 10^{9} is achieved.
EXCITING PROSPECTS

The future is bright for structures and devices based on two-dimensional carrier confinement in electronics, communications, physics and other disciplines. GaAs, InP and Ga HEMTs will continue their march towards higher integration, higher frequency, higher power, higher power efficiency, lower noise, and lower cost. Society will derive great benefits from this. GaN in particular, offers the promise of penetrating into the high-power high-frequency territory of vacuum tubes and lead to lighter, more efficient and more reliable radar and communication systems on the ground and in space. Material systems currently under research, such as the antimonides, will soon emerge into the real world and enable a new generation of ultra-low power and high speed systems.

HEMTs will continue to morph themselves into other kinds of FETs that will exploit the unique properties of 2DEG in various materials systems. The GaN system, for example, offers the prospect of a revolution in power electronics which will reach and deeply impact consumer, industrial, transportation, communications and military systems. Here, an enhancement-mode device with very low leakage current is favored and MOS-HEMT or MISFET structures are likely to prevail.

With Si CMOS rapidly advancing towards the 10 nm gate length regime, concerns have emerged about the future of Moore’s law. Managing power dissipation in future generation ultra-dense chips will demand significant operating voltage reductions. It currently seems difficult to accomplish this with Si CMOS while maintaining adequate performance. Due to their outstanding electron velocity, quantum-well MOSFETs based on InGaAs or InAs offer tremendous potential [56]. A future sub-10 nm CMOS technology could well be based on the integration of an n-channel III-V FET with a Ge- or III-V-based p-channel device on a Si platform. The challenges to realize this are enormous but the benefits to society of extending Moore’s law for several more generations are simply gigantic.

If the past serves as any guidance, it is reasonable to expect that continued progress in the quality of AlGaAs/GaAs structures will result in new insights into the often bizarre physics of quantized electrons. Similar improvements in other heterostructure material systems will also bring surprises. The FQHE, perhaps the greatest exponent for impeccable purity and atomic order, has also been observed in ZnO, SiGe and GaN [57] so the field is ripe for surprises.

The area of sensors is also pregnant with possibilities. The ultra-high mobilities that are possible in InAlSb/InAsSb system enable high-sensitivity micro-Hall sensors for many applications including scanning Hall probe microscopy and biorecognition [58]. Three-axis Hall magnetic sensors have been demonstrated in micromachined AlGaAs/GaAs structures [59]. These devices have the potential of measuring the three vector components of a magnetic field and they may find use in future electronic compasses and navigation. 2DEG-based devices can also be used for THz detection, mixing and frequency multiplication [60] as has been demonstrated in the AlGaAs/GaAs system [61]. GaN and related materials are chemically stable semiconductors with strong piezoelectric polarization. “Functionalized” GaN-based HEMT structures can be used as sensitive and robust chemical and biological sensors in gases and liquids [62]. More sophisticated sensors can be realized by combining functionalized GaN-based 2DEG structures and free-standing resonators [63]. These offer the possibility of simultaneous measurements of several properties such as viscosity, pH and temperature.

CONCLUSIONS

The first 30 years of the HEMT, rich and exciting as they have been, represent just the opening act. The best for HEMT electronics is yet to come. Kroemer’s Lemma of New Technology is appropriate here: “The principal applications of any sufficiently new and innovative technology have always been – and will continue to be – applications created by that technology” [64]. We are just now starting to see what 2DEG semiconductor heterostructures are capable of. From the cosmic to the personal, the next 30 years will bring us wonderful surprises and satisfactions.

REFERENCES

ACKNOWLEDGEMENTS

The author is grateful to several colleagues for many valuable suggestions: Ray Ashoori (MIT), Brian Bennett (NRL), Bobby Brar (Teledyne), P. C. Chao (BAE Systems), Takatomo Enoki (NTT), Augusto Gutierrez-Aitken (Northrop Grumman), Eric Higham (Strategy Analytics), Jose Jimenez (Triquint Semiconductor), Marc Kastner (MIT), Richard Lai (Northrop Grumman), Takashi Mimura (Fujitsu), Tomas Palacios (MIT), Loren Pfeiffer (Princeton University), Philip Smith (BAE Systems), Tetsuya Suemitsu (Tohoku University) and Ling Xia (MIT).

ACRONYMS

0DEG: Zero-Dimensional Electron Gas
1DEG: One-Dimensional Electron Gas
2DEG: Two-Dimensional Electron Gas
2DHG: Two-Dimensional Hole Gas
CATV: Cable Television
CMOS: Complementary Metal-Oxide Semiconductor
DBS: Direct Broadcasting Satellite
ECL: Emitter-Coupled Logic
FQHE: Fractional Quantum Hall Effect
GPS: Global Positioning System
HBT: Heterojunction Bipolar Transistor
IC: Integrated Circuit
IQHE: Integral Quantum Hall Effect
HTEM: High Electron Mobility Transistor
LSI: Large Scale Integration
MBE: Molecular Beam Epitaxy
MESFET: Metal-Semiconductor Field-Effect Transistor
MHEMT: Metamorphic High Electron Mobility Transistor
MMIC: Monolithic Microwave Integrated Circuit
PHEMT: Pseudomorphic High Electron Mobility Transistor
WLAN: Wireless Local Area Network