NanoMEMS SYSTEMS ON CHIP

Héctor J. De Los Santos
NanoMEMS Research, LLC
P.O. Box 18614, Irvine, CA 92623-8614 U.S.A.

ABSTRACT

NanoMEMS exploits the convergence between nanotechnology and microelectromechanical systems (MEMS) brought about by advances in the ability to fabricate nanometer-scale electronic and mechanical device structures. While the “nano” aspect of this field is in its infancy, and is not expected to reach maturity until well into the 21st century, its “MEMS” aspect is a topic of much current and near-term impact in, for instance, RF/Wireless communications. In this context, we discuss the fundamentals of NanoMEMS, in particular, as it relates to its most speculative and futuristic paradigms and applications, and then focus on the RF/Wireless MEMS aspect, specifically in its role as enabler of ubiquitous wireless connectivity.

INTRODUCTION

The field of Nanotechnology, which aims at exploiting advances in the fabrication and controlled manipulation of nanoscale objects, is attracting worldwide attention. This attention is predicated upon the fact that obtaining early supremacy in this field of miniaturization may well be the key to dominating the world economy of the 21st century, and beyond. NanoMEMS exploits the convergence between nanotechnology and microelectromechanical systems (MEMS) brought about by advances in the ability to fabricate nanometer-scale electronic and mechanical device structures. Indeed, the impact of our ability to make and control objects possessing dimensions down to atomic scales, perhaps first considered by the late Richard Feynman in his 1959 talk “There is Plenty of Room at the Bottom” is expected to be astounding (1). In particular, miniaturization, he insinuated, has the potential to fuel radical paradigm shifts encompassing virtually all areas of science and technology, thus giving rise to an unlimited amount of technical applications. Since high technology fuels the prosperity of the world’s most developed nations, it is easy to see why the stakes are so high.

FUNDAMENTALS OF NANOMEMS

Progress in the field of miniaturization benefited from the advent of the semiconductor industry in the 1960’s, and its race to increase profits through the downscaling of circuit dimensions which, consequently, increased the density and the yield of circuits fabricated on a given wafer area. This density, which derived from progress in photolithographic tools to produce the ever smaller two-dimensional patterns (device layouts) of an integrated circuit (IC), has increased since by more than seven orders of magnitude and has come to be captured by Moore’s law: The number of components per chip doubles
every 18 months (2). The culmination of such miniaturization program, it is widely believed, is the demise of Moore’s law, whose manifestation is already becoming apparent due to an increasing predominance of the quantum mechanical nature of electrons in determining the behaviour of devices with critical dimensions (roughly) below 100nm.

This line of development is closely related to the field of quantum devices/nanoelectronics, which was prompted by the conception of a number of atomic-level deposition/manipulation techniques, in particular, molecular beam epitaxy (MBE), originally exploited to construct laboratory devices in which the physics of electrons might be probed and explored, following the discovery of electron tunnelling in heavily-doped pn-junctions (3). Nanoelectronics did produce interesting physics, for instance, the discovery of Coulomb blockade phenomena in single-electron transistors, which manifested the particle nature of electrons, and resonant tunnelling and conductance quantization in resonant tunnelling diodes and quantum point contacts, respectively, which manifested the wave nature of electrons (4-6). These quantum devices, in conjunction with many others based on exploiting quantum phenomena, generated a lot excitement during the late 1980’s and early 1990’s, as they promised to be the genesis for a new digital electronics exhibiting the properties of ultra-high speed and ultra-low power consumption (7-8). While efforts to realize these devices helped develop the skills for fabricating nanoscale devices, and efforts to analyze and model these devices helped to develop and mature the field of mesoscopic quantum transport, the sober reality that cryogenic temperatures would be necessary to enable their operation drastically restricted their commercial importance. A few practical devices, however, did exert commercial impact, although none as much as that exerted by silicon IC technology, in particular, heterojunction bipolar transistors (HBTs), and high-electron mobility transistors (HEMTs), which exploit the conduction band discontinuities germane to heterostructures, and modulation doping to create 2-D electron confinement and quantization, respectively, and render devices superior to their silicon counterparts for GHz-frequency microwave and low-transistor-count digital circuit applications (9-14).

The commercial success of the semiconductor industry, and its downscaling program, motivated emulation efforts in other disciplines, in particular, those of optics, fluidics and mechanics, where it was soon realized that, since ICs were fundamentally two-dimensional entities, techniques had to be developed to shape the third dimension, necessary to create mechanical devices exhibiting motion and produced in a batch planar process (15). These techniques, which included surface micromachining, bulk micromachining, and wafer bonding, became the source of what are now mature devices, such as accelerometers, used in automobile air bags, and pressure sensors, on the one hand, and a number of emerging devices, such as, Gyroscopes, Flow Sensors, Micromotors, Switches, and Resonators, on the other (16). Coinciding, as they do, with the dimensional features germane to ICs, i.e., microns, these mechanical devices whose behavior was controlled by electrical means, exemplified what has come to be known as the field of microelectromechanical systems (MEMS).

Three events might be construed as conspiring to unite nanotechnology and MEMS, namely, the invention of a number of scanning probe microscopies, in particular, scanning tunneling microscopy (STM) and atomic force microscopy (AFM), the discovery of carbon nanotubes (CNTs), and the application of MEMS technology to
enable superior RF/Microwave systems (RF MEMS) (16-18). STM and AFM, by enabling our ability to manipulate and measure individual atoms, became crucial agents in the imaging of CNTs and other 3-D nanoscale objects so we could both “see” what is built and utilize manipulation as a construction technique. CNTs, conceptually, two-dimensional graphite sheets rolled-up into cylinders, are quintessential nanoelectromechanical (NEMS) devices, as their close to 1-nm diameter, makes them intrinsically quantum mechanical 1-D electronic systems while, at the same time, exhibiting superb mechanical properties. MEMS, on the other hand, due to their internal mechanical structure, display motional behavior that may invade the domain of the Casimir effect, a quantum electrodynamical phenomenon elicited by a local change in the distribution of the modes in the zero-point fluctuations of the vacuum field permeating space (16-18). This effect which, in its most fundamental manifestation, appears as an attractive force between neutral metallic surfaces, may both pose a limit on the packing density of NEMS devices, as well as on the performance of RF MEMS devices (19).

NANOMEMS SYSTEM-on-CHIP

NanoMEMS SoCs may be predicated upon a multitude of physical phenomena, e.g., electrical, optical, mechanical, magnetic, fluidic, quantum effects and mixed domain. Therefore, their universe of possible implementations and applications is vast and only limited by our imagination. Possible areas of endeavor already under research include, Nanoelectronics, Nanocomputation, Nanomechanics, Nanoengineering, Nanobiotechnology, Nanomedicine, Nanochemistry, and RF MEMS. In principle, then, there is the potential for conceiving new devices that might spark a revolution as important and wide-ranging as that engendered by the invention of the transistor and ICs. Ultimately, however, the success of the technology may well lie on its ability to deliver improved performance at low cost on technology-blind applications, Fig. 1, as well as in enabling new applications (some of which are right now only limited by our imagination). For the purposes of this paper, we focus on NanoMEMS SoCs in terms of implementation and applications.

NEMS/MEMS SoC Architectures

Regardless of the technology of implementation utilized, a system must perform a definite function and is characterized by how close it comes to meeting certain
technology-blind specifications. Typically, the design process begins with a block diagram of the system in question, which displays an architecture or high-level topological diagram showing how the constituent building blocks are interconnected to transform or process one or more input signals into one or more output signals, Fig. 1. Following this, overall systems analysis assigns or “flows down” the overall system specs to the individual building blocks, which are then designed. In the case of NanoMEMS SoCs this is difficult to do because the field is so premature that, using a circuit analogy, the equivalents of passive components (resistors, inductors, capacitors, diodes) and active components (transistors) is not yet available to the degree of completeness that would allow a complete consistent system implementation. Our course of action, therefore, is to expose a variety of potential NanoMEMS SoC building blocks.

NanoMEMS SoC Building Blocks

Interfaces

The idea behind NanoMEMS is that of creating a system that, in order to accomplish a given function, avails itself of devices and techniques spanning the range from the micro-scale to the nano-scale and beyond. In the most general case, the input signal to a NanoMEMS SoC will be analog, i.e., will exhibit continuous amplitude and will exist at all times, Fig. 1. Processing this signal, therefore, will entail deciding whether it is feasible to act on it as received-detected, or to transform it to a more convenient state. The nature of the interface sensor, in particular, its sensitivity, bandwidth, and dynamic range, will come into play here and will dictate the need for transduction, amplification, digitization, filtering, etc., thus determining the rest of the architecture. In this context, the doubly-anchored Si beam has been considered as a potential mechanical sensing element in future NanoMEMS SoCs, and impressive estimates for its intrinsic force sensitivity $(S_F)$, dynamic range (DR), mass sensitivity (M), and bandwidth (BW) have been obtained by Roukes (20). For instance, a beam of length, width, and thickness $0.1 \times 0.01 \times 0.01$ microns and active mass $10^{-18}$ ag would exhibit $S_F^{1/2} = 3 \times 10^{-17} N / \sqrt{Hz}$, $DR = 35dB$, $M = 1.7 \times 10^{-21} g$, and $BW = 7.7 GHz$, assuming a temperature of 300K and a Q of 10,000. Unfortunately, it is unclear whether the full extent of these parameters will be accessible due to various practical difficulties such as mass variation due to unpredictable adsorbates, and the impossibility of realizing a noiseless read-out. This latter theme is also common to electrostatic- and optically-based sensing interfaces as well. In the former case, which according to Roukes may attain a minimum capacitance of $10^{-18} F$, the parasitic capacitance would preclude resolving it (20). In the latter case, the fact that the spot size of the light delivered by the optical fiber used in AFM displacement-sensing is much greater than nanoscale dimensions, precludes its resolution and, hence, proper detection. In systems with an electronic input signal sensing scheme, however, the sensor may take the form of a quantum superlattice-based analog-to-digital converter, Fig. 2 (21). Here, the pulsating nature of the superlattice’s current-voltage characteristic directly samples/quantizes the voltage axis. The resulting current is used to generate pulses that drive a counter whose output is a digital representation of the input voltage. For highest resolution, the superlattice may be realized with molecular devices.
Interconnects

The communication among the building blocks of SoCs is enabled by interconnect links. These, traditionally consist of transmission links (TLs), where a TL propagates signals, from one point to another, with minimum attenuation and distortion. For NanoMEMS SoCs, these TLs may adopt a number of forms, according to the physical domain being exploited to represent signals, which could be, for example, electrical, optical, or thermal in nature. Thus, the physical realization of the TL would vary accordingly, e.g., an ordinary metal trace transmission line, a suspended mechanical beam, or a dielectric waveguide. At nanoscale dimensions, however, since quantum effects set in, the analysis and design of these TLs must deal with the manifestation of charge discreteness, and discrete phonon and photon transport, respectively.
In the context of miniaturized, narrow metal-trace transmission lines, for instance, new phenomena emerge, in particular, the manifestation of charge discreteness (21). This occurrence is typified by two cases, namely: 1) The so-called L-design (high-impedance) case, in which the governing Schrödinger equation has energy eigenvalues given by,

$$\varepsilon(p, \phi) = \frac{2\hbar}{q_e^2} \sin^2 \left( \frac{q_e}{2\hbar} (p - \phi) \right)$$  \hspace{2cm} (1)$$

where $\phi$ is the magnetic flux threading the TL. Eq. (1) implies that in the discrete charge regime the TL energy becomes a periodic function of $p$ or $\phi$, with maximum amplitude $\frac{2\hbar}{q_e}$ and nulls occurring whenever $p = \phi + n \frac{\hbar}{q_e}$. As a consequence, TL current is given by, $I(\phi) = \frac{\hbar}{q_e L} \sin \left( \frac{q_e}{\hbar} \phi \right)$ even without the application of a signal, as long as there is present a magnetic flux such as might be coupled from adjacent circuits; and 2) The so-called C-design (low-impedance) in which the governing Schrödinger equation has energy eigenvalues given by,

$$\varepsilon = \frac{1}{2C} \left( n q_e - CV \right)^2 - \frac{C}{2} V^2$$  \hspace{2cm} (2)$$

where $n$ is the number of elemental charges describing the TL state. Eq. (2) implies that in the discrete charge regime the TL energy is a quadratic function of the state $n$ of charges, thus the total charge in the ground state is given by,

$$q = \sum_{k=0}^{\infty} \left\{ u \left[ V - \left( k + \frac{1}{2} \right) \frac{q_e}{C} \right] - u \left[ -V - \left( k + \frac{1}{2} \right) \frac{q_e}{C} \right] \right\} q_e$$  \hspace{2cm} (3)$$

where $u(z)$ is the unit step function and, taking the corresponding current is given by,

$$I = \frac{dq}{dt} = \sum_{k=0}^{\infty} q_e \left\{ \delta \left[ V - \left( k + \frac{1}{2} \right) \frac{q_e}{C} \right] + \delta \left[ V + \left( k + \frac{1}{2} \right) \frac{q_e}{C} \right] \right\} \frac{dV}{dt}$$  \hspace{2cm} (4)$$

Eqn (4), in turn, indicates that the current exhibits a series of delta-function impulses with periodicity $q_e/C$, consistent with every time a single electron charge is added, and amplitude proportional to the slope of the voltage source. This leads to the important observation that a low-impedance ideal TL in the discrete charge regime will exhibit current flow dominated by Coulomb blockade.

Clearly, these limiting cases of a charge-transport-based TLs exhibiting persistent currents or Coulomb blockade-type current flow, suggest the TL itself becomes integral part of the NEMS/MEMS SoC, i.e., the usual independence between building block and interconnect becomes blurred, as their properties are not as transparent as those of conventional TLs; they must be absorbed into the building blocks.

**Signal Processing**

While the specific structure of a NanoMEMS SoC is still the subject of much research, a number of potential building blocks for NanoMEMS-based signal processing have been proposed. In what follows, we present a number of these (22).
Fig. 3. This device, theoretically analyzed by Armour and Blencowe, comprises a cantilever resonator operating at radio frequencies disposed over one of the arms of an Aharonov-Bohm (AB) ring containing a quantum dot (QD) (24, 25).

![Figure 3. Sketch of mechanical which-path electron interferometer (21).](image)

The fundamental principle operation of the device relies on the modulation of electron transmission through an AB ring by the electrostatic coupling of a vibrating beam disposed over one of the ring arms. As the beam vibrates, coupling between it and the electrons hopping in/out of a QD in the branch underneath the beam modulates the interference fringes, according to vibration frequency $\omega_0 \tau_d = \hbar/\Delta E_{\text{inc}}$, product, where $\Delta E_{\text{inc}}$ is the electron energy spread. For $\omega_0 \tau_d << 1$, short dwell time, interference fringes are destroyed if $qE\Delta x_{\text{th}} > \Delta E_{\text{inc}}$, where $x_{\text{th}}$ is the thermal position uncertainty of the cantilever and $E$ the electric field. This signals electron dephasing and detection in QD arm. For $\omega_0 \tau_d \sim 1$, the beam-QD behaves as a coherent quantum system, beam vibration and QD exchange virtual energy quanta in resonance, and interference fringes are modulated at beam vibrating frequency. For the largest dwell times, the environment induces lost of coherence.

**Casimir Effect Oscillator.** The experimental demonstration of this device has elicited widespread interest as it represents the first clear demonstration of the impact of the Casimir force in the performance of nanoelectromechanical quantum circuits and systems (NEMX) (21). In essence, the experiment consisted in varying the proximity between a vibrating torsional resonator and a metallic sphere, Fig. 4(a), to measure its behavior in the absence/presence of the Casimir force. In particular, it was found that in the absence of Casimir force influence, i.e., for sphere-oscillator separation distances greater than 3.3 $\mu$m, the oscillator resonance frequency was equal to the drive frequency, 2748Hz, and the angular amplitude frequency response was symmetric and centered around the drive frequency, $\omega_0 = \sqrt{k/I}$, where $k$ is the spring constant and $I$ the moment of inertia, consistent with mass-spring force oscillator behaviour. At smaller separation distances,
however, in particular at 141nm, 116.5nm, and 98nm, the resonance frequency shifted, according to, $\omega_1 = \omega_0 \left[ 1 - b^2 F'(z)/2!\omega_0^2 \right]$, where $F'(z)$ is the first derivative of the external force evaluated at $z$, and the angular amplitude frequency response asymmetric and hysteretic. This behaviour was shown to be consistent with the dynamics of a mass-spring-Casimir force system. The ramifications of this beautiful experiment are enormous, in particular, it may be concluded that the Casimir force will be one of the factors limiting the integration level or density of NanoMEMS SoCs.

RF/WIRELESS MEMS

The field of RF MEMS derives its impetus from exploiting fabrication techniques, in particular, surface micromachining, bulk micromachining, and LIGA to produce superior RF and microwave devices (19).

The high performance of these devices, in turn, is expected to enable manufacturers to meet the challenge for ubiquitous/global (ground, mobile, and space) wireless
connectivity demanded by consumers, which requires ever more powerful and sophisticated future generation wireless appliances running from the same battery. The devices in question, which include, inductors, capacitors, resonators, varactors, and switches, will enable programmable matching network for multi-band/multi-standard transceivers, reconfigurable front-end transceivers and frequency-agile base stations, and high-capacity satellites, Fig. 5 (28). At the heart of RF MEMS-based reconfigurable functions we find electrostatically-actuated varactors and switches, whose operation, however, is limited by the phenomenon of pull-in (19).

**Electrostatic Pull-in Voltage (28)**

Pull-in denotes the voltage at which control of an electrostatically actuated device, e.g., a beam, is lost due to the lost of equilibrium between electrostatic and spring forces. This phenomenon becomes apparent from an examination of the equation of equilibrium between spring and electrostatic forces whose solution gives the displacement of the movable plate of a parallel-plate capacitor, \( x(V) \), Fig. 6(a).

This equation is given by:

\[
\left( kx - \frac{e_0 A V^2}{2(d_0 - x)^2} \right) = 0 \tag{5}
\]

and it implies that, starting at \( V = 0^+ \), \( x \) will adjust itself so that (5) is *eventually* satisfied. For this adjustment in \( x \) to occur, the magnitude of \( V \) must be such that the *initial* value of (5) is greater than zero. In that case (5) has positive real (physical) roots, \( x(V) > 0 \). The voltage that demarcates this regime is that at which the difference between spring and electrostatic forces, embodied by (5) is a minimum. Its value may be found by calculating the derivative of (5) at equating it to zero, i.e.,

\[
\frac{dV}{dx} = \frac{d}{dx} \left[ \frac{2kx(d_0 - x)^2}{\varepsilon_0 A} \right] = 0 \quad \Rightarrow \quad x = \frac{d_0}{3}. \tag{6}
\]

Then, substituting \( x = d_0 / 3 \) into (5) and solving for \( V \) one obtains,
\[ V_{pi} = \sqrt{\frac{8kd_0^3}{27\varepsilon_0 A}} \]  

(7)

This is the pull-in voltage. Fig. 6(b) shows a plot of the normalized displacement versus the applied voltage normalized to the pull-in voltage, indicating that at pull-in, the Displacement becomes imaginary and the slope of the applied voltage no longer controls the displacement. The importance of pull-in cannot be over-emphasized, as it poses, not only performance, but also, important reliability limitations. A number of schemes to avoid pull-in have been advanced. In particular, the use of separate control and signal electrodes, with respective electrode-to-top-plate separations \( d_2 \) and \( d_1 \) obeying the relationship \( d_1 \leq d_2 / 3 \), Fig. 7, has resulted in the tuning range being extended from the theoretical pull-in-limited value of 50% to more than 400% (30).

![Figure 7. Cross-sectional view of MEMS varactor (30).](image)

Conclusions

The field of NanoMEMS SoCs is full of promise. A great deal of work is needed to define architectures and pertinent building blocks that can consistently exploit the plethora of physical phenomena upon which novel applications may be based. SoCs predicated upon integrating micro-, micro-to-nano-, and even atomic-scale devices are foreseeable. The versatility and power of this fabrication technology is well poised to enable high-impact commercial applications and paradigms in the near-term, in particular, ubiquitous RF/Wireless communications.

ACKNOWLEDGMENTS

The author thanks **SBMICRO 2004** Program Chair Prof. Edval J. P. Santos, Universidade Federal de Pernambuco, for encouraging the preparation of this paper.

REFERENCES


Key Words
Nanotechnology, Quantum Mechanics, Nanoelectromechanical, Microelectromechanical, NEMS, MEMS, Casimir force, Interferometer, System on Chip, Quantum device, RF, Wireless, Communications.