

Semiconductor Device Physics 2006

Professor D.R. McKenzie
mckenzie@physics.usyd.edu.au

Reference Texts

- C. Kittel Introduction to Solid State Physics J.Wiley New York
- S.M. Sze Physics of Semiconductor Devices J.Wiley New York

Lecture 1-Revision and Scaling of MOSFET

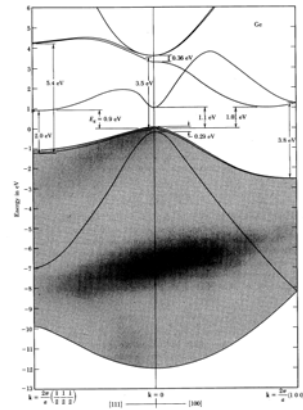
- Schrodinger equation $i\hbar \frac{\partial \psi}{\partial t} = [-\frac{\hbar^2}{2m} \nabla^2 + V]\psi$
- state of system \Rightarrow wavefunction $\psi(\mathbf{r}, t)$
- physical quantities \Rightarrow operators
- Energy eigenstates are states with well defined energy E and in regions of constant potential have the form of a plane wave

Periodic Potential

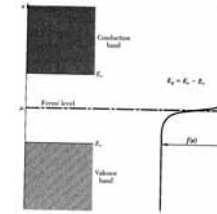
- the energy eigenstates have the form of a Bloch wave
 $\psi_k(\mathbf{r}) = u_k(\mathbf{r}) \exp(i\mathbf{k} \cdot \mathbf{r})$
- wave packets can be formed from the Bloch waves and describe electrons in a semiconductor. These electron states behave in a similar way to those for free electrons but have an *effective* mass different to that of a free electron
- The magnitude of the group velocity of the wavepacket is

$$v_g = \frac{d\omega}{dk} = \frac{1}{\hbar} \frac{dE}{dk}$$

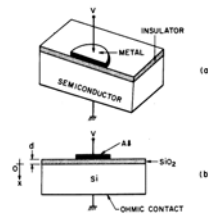
Energy Dispersion Relation-Ge



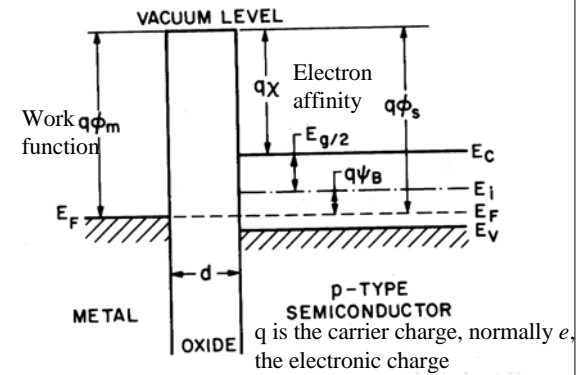
Schematic of Semiconductor Energy Bands Showing Fermi Function



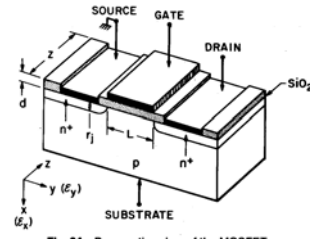
MOS Diode



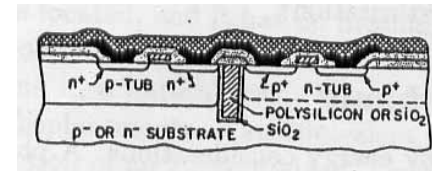
Energy Diagram-MOS Diode



MOSFET -Perspective



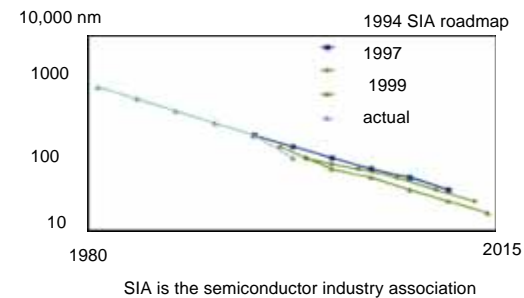
CMOS MOSFET



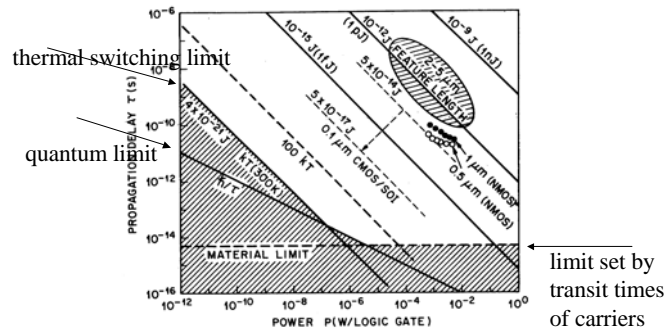
Moore's Law

- A result of 20 years experience in the manufacture of very large scale integration (VLSI) the semiconductor industry has found that progressive miniaturisation has been a successful recipe and is encapsulated in the experience known as "Moore's Law". Moore's law states that the transistor density on integrated circuits doubles every couple of years. This exponential growth and ever-shrinking transistor size result in increased performance and decreased cost .
- What are the limits to this miniaturisation process in conventional VLSI?
- There are two kinds of limits, fundamental limits and practical limits

Minimum feature size of CMOS chips-historical and future trends



limits to miniaturisation



Materials Issues in MOSFET performance

- We need to get the best performance from the MOSFET, that is small size, speed, and low power consumption for digital operation.
- We consider first small size and speed, as these go together

Scaling relationships in the standard MOSFET

- Scaling is the reduction in minimum feature size by a factor k . The results shown are the necessary result of scaling with no change in the *relative* dimensions of each quantity to maintain normal operation

Feature size	$1/k$
voltages	$1/k$
fields in semiconductor	1
speed	k
Dopant concentration	k
Current density	k
Power per circuit	$1/k^2$

Field, Charge and Voltage Scaling

- The same field is required normal to the semiconductor surface in the channel to produce inversion.
- At inversion, the semiconductor surface will carry the same areal charge density, so the total charge on the semiconductor surface and gate is scaled down by k^2
- The same source drain field is required to preserve the relative size of the two fields.
- This means that the lateral voltage for normal operation must scale down as k since lateral field is $E=V/L$
- The gate voltage must also scale down to preserve relativity with lateral voltage (also follows from scaling up of gate areal capacitance by k as gate insulator thickness goes down)

Dopant Concentration Scaling

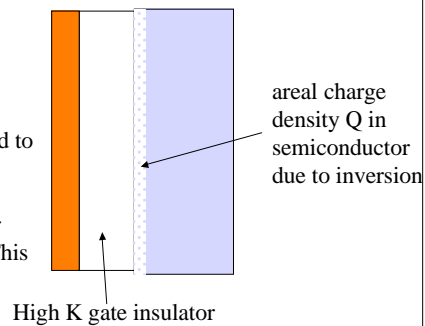
- The gate voltage to turn on the MOSFET is given by $V = \frac{eNd^2}{2\epsilon\epsilon_0}$, where d is the depletion layer depth
- We require d to scale as $1/k$ to preserve relative dimensions and since V must be constant, this means that Nd is constant so that N must scale up by factor k

Changes in materials required because of new effects

- As the gate insulator thickness goes down, the unwanted tunnelling current from the gate goes up exponentially
- Solution: high K oxide

Effect of high K dielectric

At fixed voltage the thickness required to achieve Q can be increased by a factor equal to the insulator dielectric constant. This reduces tunnelling.



Heat production per unit area

- Since the power scales down as k^2 , with the area, the intensity of power production per unit area is basically unaffected. However, there are other reasons for wanting further power reductions (new designs of MOSFET such as FinFET which have smaller area requirements).
- To achieve this, the operation of the transistor requires as high a value of the source-drain conductance as possible to minimise heat production. From the formula for source drain mutual conductance (not derived here), this leads to the requirement that C be made as large as possible. This will also be achieved with high K gate insulators.
- Conductance can be further increased with high mobility semiconductors and new geometries (eg. FinFET)

$$g_m = \frac{W}{2L} \mu C V_D$$

Mobility of Carriers

- The ability of the carriers to carry a current depends on their *drift velocity*

$$j = nev_e + pev_h = \sigma E$$

- Mobility is defined as the *drift velocity per unit electric field*.

$$v = \mu E$$

$$\text{hence } \sigma = j / E = (ne\mu_e + pe\mu_h)$$

- From an elementary derivation,

$$v = e\tau E / m^*$$

- the mobility is therefore strongly dependent on the effective mass

Values of Effective Mass

Table 2 Effective masses of electrons and holes in direct gap semiconductors

Crystal	Electron m_e/m	Heavy hole m_{hh}/m	Light hole m_{lh}/m	Split-off hole m_{so}/m	Spin-orbit Δ , eV
InSb	0.015	0.39	0.021	(0.11)	0.82
InAs	0.029	0.41	0.025	0.08	0.43
InP	0.073	0.4	(0.075)	(0.15)	0.11
GaSb	0.047	0.3	0.06	(0.14)	0.80
GaAs	0.066	0.5	0.082	0.17	0.34
Cu ₂ O	0.99	—	0.58	0.69	0.13

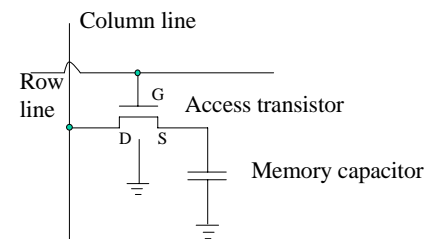
Mobility of carriers-2

Table 3 Carrier mobilities at room temperature, in cm²/V.s

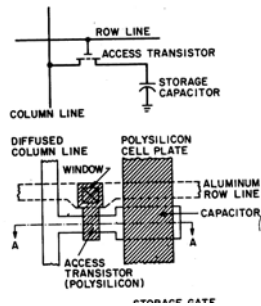
Crystal	Electrons	Holes	Crystal	Electrons	Holes
Diamond	1800	1200	GaAs	8000	300
Si	1350	450	GaSb	3000	1000
Ge	3600	1800	PI-S	550	600
InSb	800	450	PI-Se	1020	930
InAs	30000	420	PI-Te	2500	1000
InP	4500	100	AgCl	50	—
AlAs	280	—	KBr (100 K)	100	—
AlSb	900	400	SiC	100	10-20

The carrier mobility affects the speed of devices and work is under way to increase it

NMOS DRAM Memory Cell



Devices-2 The NMOS DRAM Memory Cell

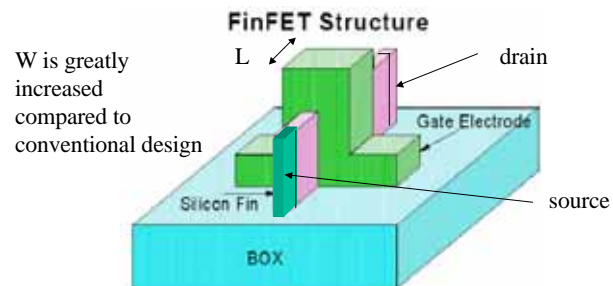


the capacitor must be refreshed every few ms

The FinFET

- The FinFET is a design that increases W/L while reducing chip footprint, giving greater miniaturisation of the “ground floor”

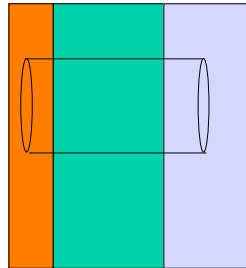
The FinFET Structure



Materials Issues in MOSFET Gate insulator

- Trapped charge in insulator unwanted (shifts characteristics)
- Defect density (trap states) at Si surface relatively low
- Traditionally made from SiO_2 because of ease of growth (thermal oxidation), low trap density, high dielectric strength.
- As miniaturisation proceeds, SiO_2 is no longer the best choice because of the quantum tunnelling effect which is exponentially dependent on thickness
- Need good high K materials ZrO_2 , HfO_2 , Zr SiO_4

Effect of Trapped Charge in Gate Insulator



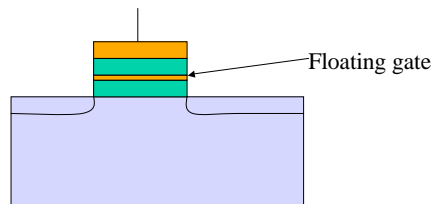
Effect of trapped charge is to keep charge on the semiconductor

Metal oxide semiconductor

The 4 Types of MOSFET available in CMOS

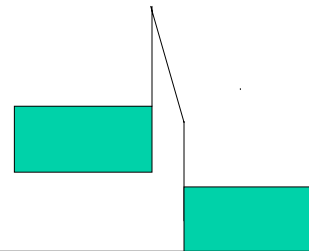
TYPE	CROSS SECTION	OUTPUT CHARACTERISTICS	TRANSFER CHARACTERISTICS
n-CHANNEL ENHANCEMENT (NORMALLY OFF)			
n-CHANNEL DEPLETION (NORMALLY ON)			
p-CHANNEL ENHANCEMENT (NORMALLY OFF)			
p-CHANNEL DEPLETION (NORMALLY ON)			

Flash Memory



Fowler Nordheim Tunnelling

$$J = J_0 e^{-\frac{a\phi^{3/2}}{E}}$$



Power Scaling and Current density scaling in interconnects

- The power is chargevoltage/speed and therefore is scaled down by $k^2 \cdot k/k = k^2$. The maximum current is P/V and so is scaled down by a factor of k , since the voltage is scaled down by k .
- However, the current is carried by interconnects that are scaled down in cross section by k^2 so the current density is scaled up by k

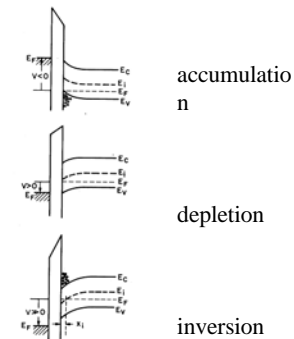
Materials effects in Interconnects

- Originally, the MOSFET only needed 5 elements Si, O, B, P and Al. Al was the interconnect material of choice (easy deposition, compatible with Si, resistant to oxidation)
- Because of the need to reduce heat dissipation as current density scales up as k , Cu is now used. C nanotubes may be used in future as these have higher conductivity
- Also, there is a need to reduce the capacitance associated with interconnects. This problem will get worse as the proximity of interconnects increases. Need *low k* dielectric to replace SiO₂

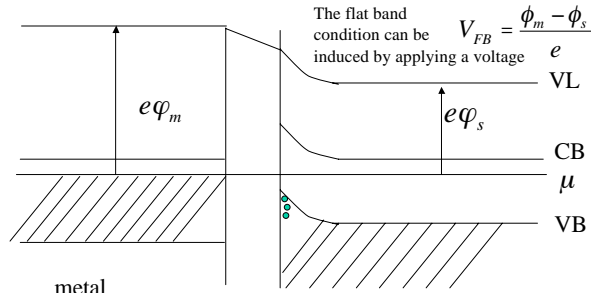
MOSFETs as chemical and biological sensors

- All sensors work by detecting changes at an interface between a solid surface and a gas or liquid.
- Let us first consider the detection of gas in a form of MOSFET termed a ChemFET

Three states of MOS Diode



Real MOS Energy Diagram



The semiconductor is driven into inversion if the metal work function is large enough.

Gas Sensors (Continued)

Using Pd as the metal, we find that the work function is strongly affected by the absorption of hydrogen and hence the extent of inversion is changed. Any change in the work function changes the source drain current at constant V_g . A double layer introduces an additional contribution to the work function

- Other gases can be detected using the same principle: H_2S , NH_3 , CO

Ion Sensitive FET (ISFET) for pH sensing

- The MOSFET is used with the solution as gate electrode
- A potential difference V_G arises between the oxide surface and the silicon substrate of the chip due to an ion sheath
- The potential depends on the pH of the solution

pH Sensing

- $pH = -\log[H^+]$
- Charge neutrality requires the electrode charge to be balanced by charge in the solution $Q_M = Q_S$

$$\frac{d^2\psi}{dx^2} = -\frac{\rho(x)}{\epsilon\epsilon_0}$$
- Gouy Chapman Theory 1910 $n_+(x) = n_0 \exp\left(\frac{-ze\psi(x)}{kT}\right)$
- Poisson's equation $n_-(x) = n_0 \exp\left(\frac{ze\psi(x)}{kT}\right)$
- Apply Boltzmann Statistics
- Poisson-Boltzmann equation $\rho(x) = ze[n_+(x) - n_-(x)]$

$$\frac{d^2\psi}{dx^2} = -\frac{ze n_0}{\epsilon\epsilon_0} \left[\exp\left(-\frac{ze\psi(x)}{kT}\right) - \exp\left(\frac{ze\psi(x)}{kT}\right) \right]$$

Derivation of the pH result

- Take the simple case

$$\frac{ze\psi(x)}{kT} \ll 1$$

- The equation reduces to

$$\frac{d^2\psi}{dx^2} = k^2\psi(x), k = \left(\frac{2(ze)^2 n_0}{\epsilon\epsilon_0 kT} \right)^{1/2}$$

- the Debye length

$$L_D = \frac{1}{k}$$

Introduce Charge Neutrality Condition

- The solution is $\psi(x) = \psi(0)\exp(-kx)$

- The constant prefactor is determined by charge neutrality where the charge on the electrode is Q_M

$$\int_0^{\infty} \rho(x) dx = -Q_M$$

$$\psi(0) = \frac{Q_M}{\epsilon\epsilon_0 k}$$

Surface Chemistry determines the Charge

- Introduce the electrostatic potential difference ψ_L between the bulk of the solution and the potential at the electrode surface, determined by the surface chemistry
- This is related to the voltage required to induce the flat band condition in the semiconductor

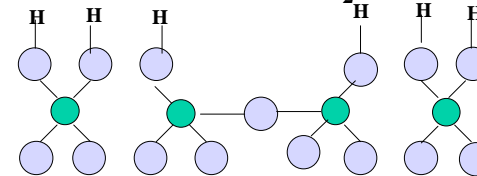
$$V_{FB} = \phi_E - \phi_S - \psi_L$$

- The flat band condition can be deduced from the source drain current and so we can sense the value of

$$\psi_L$$

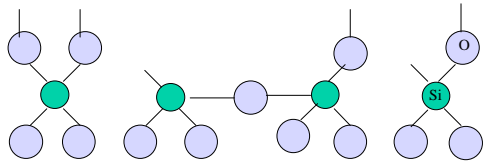
The SiO₂ Surface in Water

Water is in dissociative equilibrium



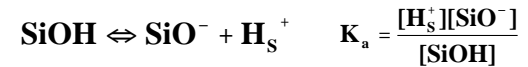
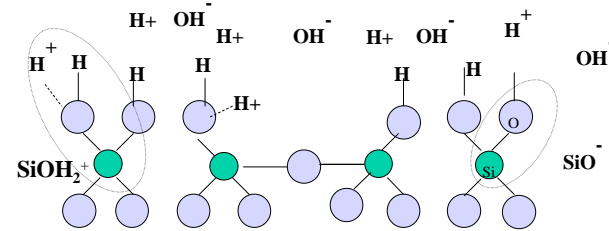
- The dangling bonds are hydrogen passivated by H and OH groups when the surface is in water

The Interface Chemistry-Passivation of Surface



- Dangling bonds on vacuum cleaved SiO₂

Acidic/Base reactions at the surface



$$\frac{K_a}{K_b} = \frac{[\text{H}_s^+]^2[\text{SiO}^-]}{[\text{SiOH}_2^+]}$$

Use Boltzmann Relation to relate H⁺ ion concentration to potential

$$[\text{H}_s^+] = [\text{H}_b^+] \exp\left(\frac{-e\phi_L}{kT}\right)$$

$[\text{H}_b^+]$ The concentration of H⁺ in the bulk solution

$$-\ln[\text{H}_b^+] + \ln\left(\frac{K_a}{K_b}\right)^{1/2} = \frac{-e\psi_L}{kT} + \ln\left(\frac{[\text{SiO}^-]}{[\text{SiOH}_2^+]}\right)^{1/2}$$

$$-\ln[\text{H}_b^+] + \ln\left(\frac{K_a}{K_b}\right)^{1/2} \approx \frac{-e\psi_L}{kT}$$

$$\text{pH} = -\log_{10}[\text{H}_b^+]$$

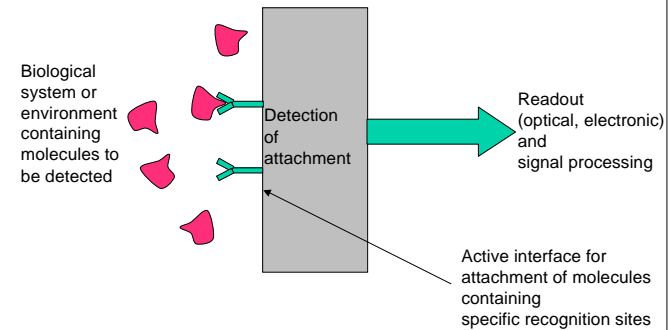
$$\text{pH}_{\text{pzc}} = -\log\left(\frac{K_a}{K_b}\right)^{1/2} = \text{const.}$$

pH Sensing

$$\psi_L = 2.303 \frac{kT}{e} (pH_{pzc} - pH)$$

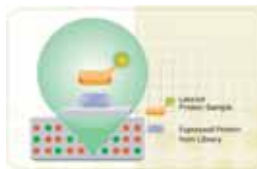
$$\Delta\psi_L = 59mV / pHunit$$

Schematic of a Bio Sensor



Protein Array Technologies

Examples



Array technologies are revolutionising medicine and medical research

Microspot protein array. Readout is optical, using fluorescence from the label.



Microspot antibody array (ELISA)

Existing technologies use nitrocellulose based surfaces- But there are some problems with repeatability, attachment density, attachment strength (washability).

Illustrations from Schleicher and Schull

Confined Electron Systems

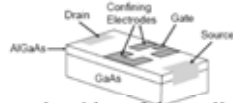
- Confined electron systems containing few electrons can be made by forming a confining potential in 3D
- Method 1-Start with a 2D confined electron gas and apply additional confining potential
- Method 2-Use a small volume of metal-quantum 'dot'

Reference work for this section

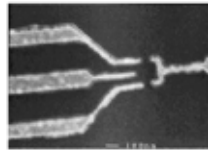
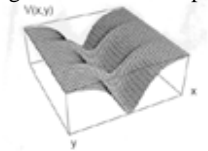
M.A. Kastner Ann. Phys Leipzig 9 (2000) 11-12 885-894

What is a 2DEG?

- A 2DEG (2 D electron gas) is created in a MOS diode during inversion, but we require 3D confinement



Confining electrodes create the sides of the well, the gate controls the depth



SET compared to FET

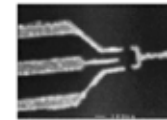
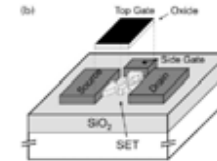
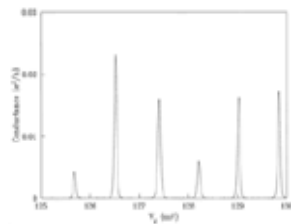


Fig. 8. Electronmicrograph of the top surface of the SET used in the experiment of Goldhaber-Gordon et al. (p. 5), see also Fig. 1.

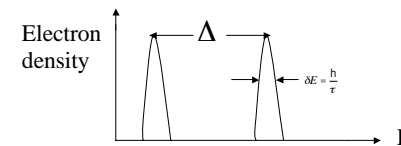
Source drain conductance as a function of gate voltage



The source drain current only exists at specific gate voltages-why?

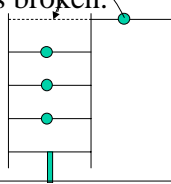
Quantisation of charge in a system- the Coulomb “Blockade”

- Of course charge is quantised, but what are the conditions that a specific number of electrons can be held in a system?
- An atom is an example of a quantum quantised charge system
- In general any confined system of electrons (even classical) has charge quantisation if a) the energy difference $D = E(N) - E(N+1) > kT$ and b) the tunnelling time t for an electron out of the system is long enough such that the energy uncertainty is less than D
- There is a blockade on the addition or subtraction of charge because of the energy gaps D

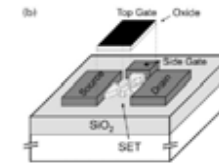


Breaking the “Coulomb Blockade”

- If the gate voltage is adjusted so that the total energy of the system with N electrons in the box is the same as for N-1 electrons in the box, the charge is not well quantised and charge fluctuations occur. The blockade is broken.



Quantum Dot SET



Applications of SETs

- SETs are useful whenever charge transfer needs to be measured accurately (electrometer function)
- Example 1 charge qubit readout in the quantum computer

Assignment due 27 October 5pm

1. Show that the solution for the electrostatic potential near an electrode in an electrolyte is

$$\psi(x) = \psi(0) \exp(-kx)$$

Where

$$\psi(0) = \frac{Q_M}{\epsilon \epsilon_0 k}$$

2. From the formulae in the notes show that

$$-\ln[H^+] = -\ln\left(\frac{K_a}{K_b}\right)^{1/2} - \frac{e\psi_L}{kT} + \ln\left(\frac{[SiO^-]}{[SiOH_2^+]}\right)^{1/2}$$