IN: Proceedings of the Symposium on Silicon Nitride and Silicon Dioxide Thin Insulating Films (Electrochemical Society Proc Vol 89-7), pp. 385-400 (1989)

Si-SiO_2 INTERFACE DEGRADATION MEASUREMENT USING THE FLOATING GATE TECHNIQUE

Albert K. Henning

Thayer School of Engineering
Dartmouth College
Hanover, NH 03755

ABSTRACT

The floating gate technique is used in MOSFETs to characterize and understand gate current, an important component of hot carrier effects and reliability in these devices. The technique itself causes changes in the local interface state density near the device drain. These changes are particularly evident when gate current is measured at low drain bias. The technique then allows measurement of local changes in the interface state density, provided the carrier temperature and current density are known at the point of insulator injection. A 2-D drain-avalanche gate current model is used in simulating these parameters, allowing extraction of the interface state changes.

INTRODUCTION

The ultimate goal in Si MOSFET technology development is to simulate, with physical accuracy, all important device aspects, from fabrication through operation to reliability. Achievement of this goal results in faster overall process development, combined with reduced exposure to the high costs of actual manufacture. To achieve physical accuracy, simulations must be done with 2-D or 3-D simulator tools, on the microscopic level. In terms of MOSFET device design for hot carrier resistance, such simulators seek to predict processing effects on substrate current [1,2], gate current [3,4], and eventually interface degradation [5,6].

No simulator can stand on its own, however; it must rest on accurate measurement. Previous work has explored gate current measurements [7-10] and interface degradation [11-13] as separate phenomena, though an empirical, indirect relationship has been established [14]. This work investigates the direct, microscopic relationship of the two measurements, and thus the direct connection between gate current and interface degradation. In particular, the very measurement of gate current is shown to alter the local interface. The measurements are then combined with a 2-D gate current simulation model, based on the microscopic details of weak-avalanche hot carrier processes. The marriage of measurement and simulation allows calculation of local changes in interface state density caused by charge flux at the local point of emission from the device channel.

Figure One shows the qualitative problem. In a MOSFET (N-channel is described here, but the arguments hold for P-channel as well), channel

current is heated by the high electric field near the device drain. Occasionally, these hot carriers can break Si-Si bonds, thus forming free electron-hole pairs via impact ionization. The separate charges of the pair are assumed to begin with zero momentum [16]. Depending on the gradient of the potential, the charges are swept toward either the source, substrate, or drain; or toward the insulator interface. Charges collected at the source may cause source-drain (bipolar-mode) breakdown. Substrate collection is used to characterize the impact ionization process. Drain collection modifies the channel current slightly, until the thyristor action of strong avalanche occurs. However, charge which reaches the interface can harm device operation. Interface states may be created or filled. Bulk traps can be filled. Mobility degradation and threshold voltage shifts resulting from these phenomena are assumed to create unacceptable reliability problems at the circuit level. Current collected at the gate electrode thus is used to characterize these critical processes.

CHARACTERIZATION

The floating gate technique [8,9] was used previously to measure gate current in NMOS FET's at 300K and 80K [7]. The technique is reviewed here for completeness. The extraction procedure involves the equation:

$$I_G(V_G)=C_G(V_G)\frac{dV_G}{dI_{DS}}\frac{dI_{DS}}{dt} \qquad (1)$$

The capacitance $C_G(V_G)$ includes the MOSFET capacitance, as well as the contribution from the metal pad and contacting probe tip which allow electrical access to the gate. Preferably, the capacitance of the metal pad swamps the gate capacitance of the MOSFET. This allows the metal pad to act as a reservoir of charge which is added to, or subtracted from, during the floating gate measurement. It also allows C_G to be virtually constant over the V_G range of measurement: the long time constants of the measurement mean it uses quasi-static capacitance; and the field capacitor of the metal pad has a threshold voltage beyond the V_G range, so no dips in the quasi-static characteristic occur.

The indirect measurement works as follows. I_{DS} is monitored as a function of time. At some time t<0, the MOSFET is placed under DC bias on all four terminals. Some initial heating may occur in the device, which causes a slight drop in I_{DS}. At time t=0, the probe contacting the metal pad attached to the gate is lifted. Some initial changes in I_{DS} will occur due to redistribution of charge, since the probe capacitance is no longer a part of the circuit. For t>0, I_{DS} will decrease or increase only in relation to changes in charge on the the gate capacitor. These changes are assumed to occur only due to thermionic emission of either holes or electrons from the MOSFET bulk to the floating gate electrode. In an N-channel device, hole emission will increase V_G and I_{DS}; while electron emission will decrease both V_G and I_{DS}.

It should be emphasized that the indirect technique - similar to other techniques - can only measure the *net* charge which reaches the gate

electrode. This increases the need for microscopic simulations which can distinguish between both injection components.

Figure Two shows a schematic of this situation. The local nature of the gate current injection is shown by the current density J_G, which is typically located all along the Si-SiO_2 interface, but peaks sharply in value at or near the maximum of lateral electric field.

Figures Three and Four show the I_G vs. V_G curves resulting from the indirect measurement technique, presented in [7]. Several features are worth noting. First, for V_D=2.5V, the peak electron component of gate current decreases as temperature decreases to 80K from 300K. The component is typically referred to as the channel hot carrier gate current. The decrease occurs because the barrier height to electron emission at the interface (~3.5eV) exceeds the energy available to an electron when it travels from its generation point to the interface. Only those electrons which also acquire energy from optical phonons are able to surmount the barrier. Since the density of these phonons decreases at low temperature, fewer electrons cross the barrier. When V_D=4.5V, however, more than enough energy is available for thermionic emission, despite the reduction of optical phonons at low temperature. This occurs due to the increased mean free path for scattering at low temperature. The model of [14] is sufficient to explain this increase qualitatively, through the lucky-electron probability given by:

$$P_{emission} \sim \exp(\frac{\phi_B}{q\lambda\mathcal{E}}) \qquad (2)$$

where ϕ_B is the interface barrier height, q is the electronic charge, λ is the mean free path for all scattering mechanisms, and $\mathcal{E}$ is the local (in this case, vertical, as discussed below) electric field.

DISCUSSION

Another feature of the gate current curves is the presence of both holes and electrons. This presence indicates both electrons and holes are incident simultaneously on the interface, though the peaks of the two fluxes may be at different locations along the channel. For VD=2.5V, the hole current component disappears at 80K, due to the relative decrease of hole emission vs. electron emission arising from the difference in barrier heights for the two carriers. For VD=4.5V, holes are still noted at 80K; however, they are reduced vs. 300K, both because of the high barrier to hole emission and the increase in threshold voltage at low temperature. Note that the indirect measurement technique can only show the net charge transported to the gate electrode.

Figure Five shows the temperature dependence of the electron gate current emission for a variety of temperatures. The indirect method of [9]

was employed. For the high current levels (> 10^{-14} A), the measurement was generated fairly swiftly (over a few seconds), and the interface was subjected to little total gate charge flux (I_G integrated over time). However, at lower emission levels (see circled area in Figure Five, all of Figure Three, plus $I_G < 10^{-14}$A in Figure Four), the time to sketch out this portion of the curve stretched from seconds to tens of minutes. The result was the steepened curve shown. The slope increase is attributed to an increase in charged interface state density at the local point of injection. Increases in interface state density over broad areas of the channel are ruled out, since the injection is so highly localized; and since the transconductance and threshold voltage, monitored before and after each measurement, showed no perceptible changes. The local increase of interface state density can cause local fluctuations in the surface potential, without necessarily altering the linear transconductance. The lack of slope increase at higher current levels occurs for two reasons. First, the integrated charge passing through the interface plane is less than at lower current levels, because of the short measurement time. Second, the higher current flux can cause screening of the interface charge through the phenomenon of image force barrier lowering [15].

Qualitatively, this understanding is depicted in the next two figures. In the 1-D sense of Figure Six, an electron generated in the device bulk is accelerated toward the interface. This transport process leads to the electron acquiring a distribution of probable energies as discussed in [1], this time using essentially the *vertical* field. At the interface, only that part of the distribution representing electrons with energy in excess of the barrier height contributes to gate current density J_G. The injection is still highly local, however, as shown in Figure Seven. Over time, the passage of charge over the barrier at the injection point leads to the creation and charging of interface states. This charge component, appearing at the interface, causes a local increase in barrier height, at the point of injection. Only when the incident current density J_G exceeds a critical value to cause screening or image force barrier lowering, is this barrier diminished.

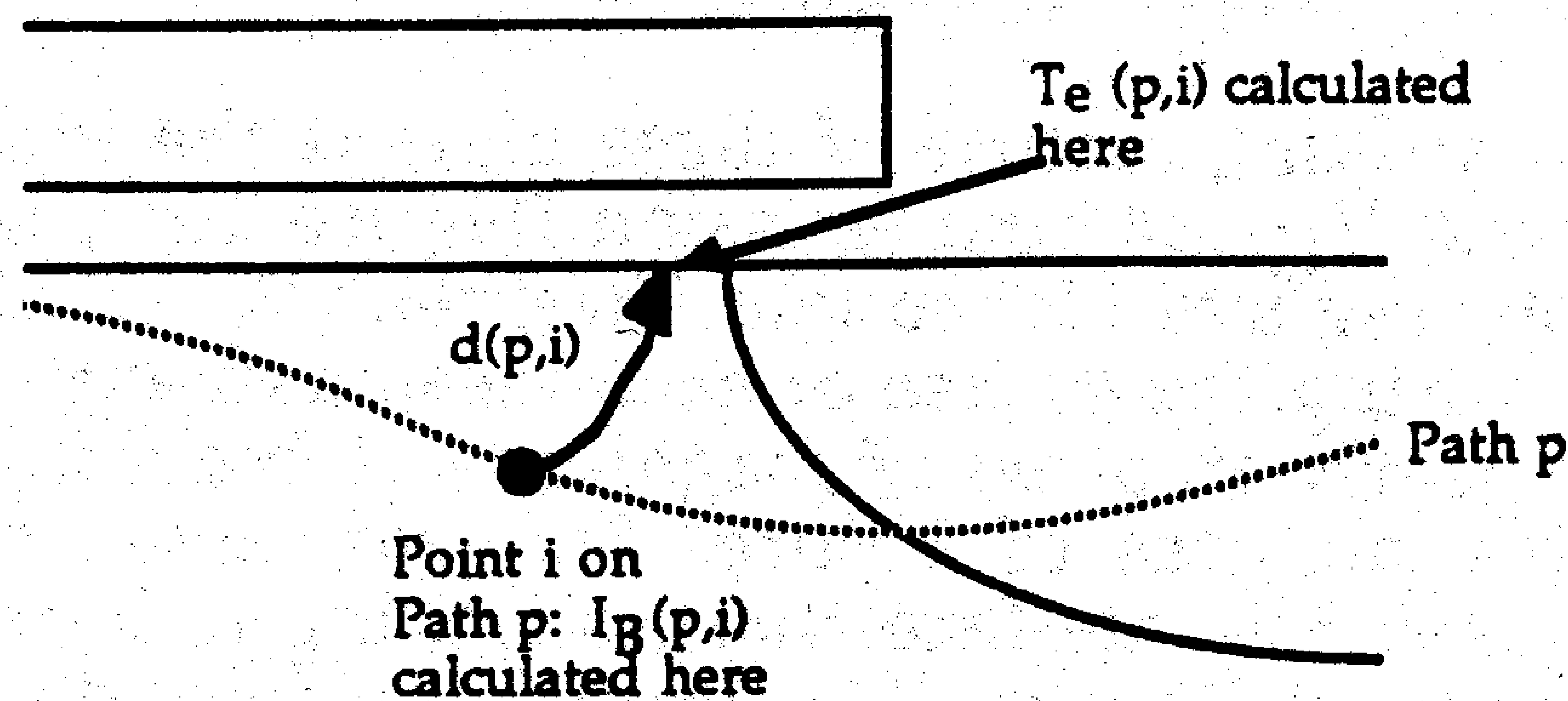
Since exact measurement of J_G is, at present, impossible, a simulator must be used to corroborate macroscopic measurements with microscopic understanding. The gate current transport process is shown in Figures Eight and Nine. Here, the results of simulations based on the 2-D device simulator PISCES II-B [18] are shown. Voltages are applied to the four device terminals, and a solution for the electron carrier concentration and electrostatic potential is obtained. Prior to simulation of gate current, I_{DS} vs. V_{GS} curves were obtained which matched experiment [19]. The simulations were carried out on a vectorized Convex computer at the Kiewit Computation Center of Dartmouth College. Substrate current measurements were also compared to simulations based on a previous model, accurate over wide ranges of bias and temperature [1]. Weak avalanche is assumed, so that the number of secondary carriers is assumed to cause negligible perturbations on potential and carrier density. The simulation environment was a Macintosh running VersaTerm, MacDraw, and CricketGraph under Switcher. Plots were generated on-screen in VersaTerm and transferred to a MacDraw document for inclusion here. Data arrays were also generated for the current flux at the

Si-SiO_2 interface, which resulted later in Figures Thirteen-Fifteen.

The drain-avalanche gate current (DAHC) model used is an extension of the LEED (Lucky-Electron, Energy Distribution) substrate current model used in [1,19]. The calculation is done on the solution file in a post-processing fashion, with the implicit assumption that impact ionization does not perturb the solution significantly. The total gate current is given by:

$$I_G=\sum_{p,i} I_B(p,i)\exp\left(-\frac{d(p,i)}{\lambda}\right)\cdot\int_{\Delta E_C-V(p,i)}^{\infty}\frac{dE}{kT_e(p,i)}\frac{E}{kT_e(p,i)}\exp\left(-\frac{E}{kT_e(p.i)}\right) \qquad (3)$$

I_B is the electron and hole current generated at the point (p,i) due to impact ionization. p and i index the particular channel current contour and specific element along the contour, respectively.



The impact-ionized current generated at this point is then followed along the steepest gradient of potential, taking the proper sign of the driving electric field into account. This assumes, implicitly, that the secondary carriers experience no diffusive force. The amount of charge reaching the interface is reduced by the lucky-electron probability, based on the mean free path for scattering by optical phonons - the main contributor to scattering of high-energy carriers. The carriers which reach the interface have thus arrived ballistically. Those which finally contribute to the gate current are determined as shown in the integral, by finding the number with enough energy to surmount the barrier. ΔE_C is the conduction band discontinuity between Si and SiO_2 at the interface. V(p,i) is the potential energy converted to kinetic energy, gained by transport between the generation and injection points. $T_e(p,i)$ is the carrier temperature at the *injection* point into the oxide. It is not the temperature at the generation point, since secondary carriers are assumed to have zero momentum at (p,i) [15]. Thus, $T_e(p,i)$ depends on the *vertical* electric field at the injection point, not the lateral field in the channel. Similar calculations can be done for holes, making appropriate changes in the mean free path and band discontinuity.

Figure Eight shows one result for $V_G>V_D$. Here, the potential gradient in an NMOS device dictates that electrons are transported toward the gate, and holes go purely toward the substrate. Note the electrons are

injected *inside* the metallurgical junction. In Figure Nine, a similar plot is made, now for $V_G < V_D$. Here, holes are favored to be transported toward the gate; in this instance, they are injected on the *channel* side of the junction. Some holes do not reach the gate, but are transported to the backside substrate contact. Thus, interpolating using Figures Eight and Nine, some terminal voltages may lead to conditions where both charge types are injected, which can alter our understanding of degradation models significantly.

The local injection demonstrated in Figures Eight and Nine is used to explain qualitatively the experimental results shown in Figures Ten through Twelve. They are formed from plots of I_{DS} vs. time obtained during the normal course of the indirect I_G measurement [7]. The transconductance plot is used to relate I_{DS} to V_G, and Eq. (1) allows I_G to be calculated. The results are shown in Figures Ten through Twelve. The "knee" in each curve is due to capacitive charge redistribution in the storage reservoir (metal pad) when the gate probe is lifted at t=0. Scatter occurs due to the extraction procedure used, which relies on a least-squares linear fit to find the local slope dI_{DS}/dt as a function of time. The values of I_G and V_G at the knee are the plotted quantities in Figures Three through Five.

The most striking feature of these plots is the logarithmic decrease in I_G, with increasing V_G for hole emission (Figure Ten), or decreasing V_G for electron emission (Figures Eleven and Twelve). V_G is changing by at most a few tens of mV during the course of each measurement; yet the calculated I_G is decreasing by nearly an order of magnitude. Observations such as these led to the qualitative pictures of Figures Six and Seven to explain them. It should be noted that the interface charges leading to local increases in barrier height, as measured in Figures Ten through Twelve, do not disappear when a new measurement is to be done. Rather, repeated measurements lead to I_G vs. V_G curves which follow the logarithmic trends already begun.

Other possible causes for the observations were examined, such as mobility or capacitance change. First, the channel mobility may be changing due to gate-current-induced changes in the $Si\text{-}SiO_2$ interface. This is ruled out, since monitor of the low-field transconductance and threshold voltage before and after the floating gate current measurement, in both forward and reversed source-drain configurations, shows no perceptible change in either g_m or V_T. Second, surface leakage from the floating gate might cause departures from the expected gate current curve. However, such leakage must be greater than the true, channel-induced gate current in order to be noticed; yet, the observed characteristic is less than the expectation. Also, monitor of the charge leakage off of the floating gate placed an upper bound on surface leakage of 10^{-18}A. The remaining alternative is that thermionic gate current from the device channel alters the interface as it passes. In particular, the thermionic barrier height must be increasing at least linearly with time to cause the observed gate current characteristic.

The DAHC gate current model in PISCES-IIB can be used to explore the local interface charge model quantitatively instead of qualitatively. Results of simulations which lead to, for instance, Figures Eight and Nine,

can be used to calculate J_G along the interface from source to drain. The calculation is done simultaneously for both holes and electrons. Figures Thirteen-Fifteen show a few of the results, paralleling voltage conditions of Figures Ten-Twelve. For hole injection, the peak occurs - as expected, due to the potential gradient from the point (p,i) of peak I_B generation - just inside the p-n substrate-drain junction. The opposite is true for electrons: peak injection occurs just inside the channel.

The final step is to relate the local interface charge measurements in Figures Ten-Twelve to changes in local barrier height. The PISCES simulations show where along the channel the peak charge is injected, and the electric field and carrier temperature at each such injection point. The barrier height and interface state density are then calculated in the following manner, with results reflected in Table One. The injected current density, averaged over the injection area, is:

$$I_G \sim \exp\left[-\frac{q\phi_B}{kT_e}\right] \quad (4)$$

Take I_{G1} and I_{G2} to be the injected gate current at the beginning (knee) and end of a characteristic such as Figure Ten. Note that injection area is assumed to be constant. Then:

$$\frac{I_{G1}}{I_{G2}} = \exp\left[-\frac{q}{kT_e}(\phi_{B1}-\phi_{B2})\right] \quad (5)$$

Re-writing, and expressing ϕ_B as an increase in barrier height at the injection point:

$$\frac{kT_e}{q}\log\left[\frac{I_{G1}}{I_{G2}}\right] = \Delta\phi_B \quad (6)$$

So the simulation yields T_e, while experiment details the rate of decrease in I_G. Interface state densities, averaged over the injection area, may also be calculated. Define:

$$C_{ox} = \frac{\Delta Q}{\Delta V} = \frac{q\Delta N_{it}}{\Delta\phi_B} \quad (7)$$

Or:

$$\Delta N_{it} = C_{ox}\frac{kT_e}{q^2}\log\left[\frac{I_{G1}}{I_{G2}}\right] \quad (8)$$

Table One displays the calculations. X_{cen} is chosen as the injection point, corresponding to the peak of the J_G vs. interface distance curve. The results for $\Delta\phi_B$ are close to 1eV - surprisingly large, until one considers they depend directly on the "vertical", and not lateral, electric field at the injection point. Since T_e is related to the electric field [1,14], the large values of T_e shown are consistent with this relationship, since vertical MOSFET electric

fields are greater than 1MV/cm, while lateral fields are only a tenth as big. The values obtained for ΔN_{it} are quite reasonable, consistent with expectations based on other experiments [12] extrapolated to the low injection levels found in these measurements.

The simulations of gate current left something to be desired in terms of quantitative agreement with experiment, though the qualitative shape of the hole and electron components were correct. It is believed the inclusion of energy transport into the solution may be sufficient to correct the problem, since energy transport alters the current contours quite significantly [20].

It must be noted that these measurements do not preclude creation or charging of interface states caused by hot carriers with energy insufficient to surmount the thermionic barrier. However, based on other experimental work [17], it appears likely that only the gate current emitted thermionically contributes to the change in the interface.

CONCLUSIONS

Gate current measurements at realistic biases in MOSFET's lead to perceptible changes in the Si-SiO_2 interface. These changes are attributed to charged interface states, charged and/or created during the passage of gate current emitted thermionically across the interface. The changes are a direct result of the gate current transported across the interface, and occur even at very low levels of injection. Thus, simulation of the gate current must either subtract out the interface changes, or account for them strictly. Experimental verification of the simulation results must take care to account for the interface changes as well. In particular, because of the build-up of interfacial charge, the sign of the net injected gate current may be reversed.

The 2-D simulator PISCES has been used to examine the microscopic gate current transport process, using a model based on drain-avalanche hot carriers. The results allow calculations of barrier height increase and interface state density increase at the local point of current injection, consistent with present understanding of lucky-electron transport thermionic emission. The simulations set the stage for predicting not just channel, substrate, and gate currents, but changes in the interface which affect transport properties such as mobility and threshold voltage.

Acknowledgements

The author thanks Mr. Daniel Theivanayagam for assistance with programming the 2-D drain-avalanche gate current model into PISCES.

References

[1] A.K. Henning, *et al.*, *Trans. Elec. Dev.* **ED-34**, p. 64 (1987)
[2] W. Hänsch and S. Selberherr, *Trans. Elec. Dev.* **ED-34**, p. 1074 (1987)
[3] A. Phillips, Jr., *et al.*, *IEEE IEDM, Dig. Tech. Papers*, p. 39 (1975)
[4] B. Ricco, *et al.*, *IEEE IEDM, Dig. Tech. Papers*, p. 559 (1986)
[5] S. Baba, *et al.*, *IEEE IEDM, Dig. Tech. Papers*, p. 734 (1986)
[6] S. Bibyk, *et al.*, *Proc. Symp. Low Temp. Electronics and High Temp. Superconductors* (S.I. Raider, *et al.*, eds.), Electrochemical Soc. Proc. Vol. **88-9**, p. 161 (1988)
[7] A.K. Henning and J.D. Plummer, *Proc. Symp. Low Temp. Electronics and High Temp. Superconductors* (S.I. Raider, *et al.*, eds.), Electrochemical Soc. Proc. Vol. **88-9**, p. 130 (1988)
[8] F.H. Gaensslen and J.M. Aitken, *Elec. Dev. Lett.* **EDL-1**, p. 231 (1980)
[9] N. Saks, *et al.*, *Trans. Elec. Dev.* **ED-33**, p. 1529 (1986)
[10] Y. Nissan-Cohen, *Elec. Dev. Lett.* **EDL-7**, p. 561 (1986)
[11] E. Takeda and N. Suzuki, *Elec. Dev. Lett.* **EDL-4**, p. 111 (1983)
[12] F.-C. Hsu, *Elec. Dev. Lett.* **EDL-5**, p. 50 (1984)
[13] P.E. Cottrell, *et al.*, *J. Solid State Circuits* **SC-14**, p. 442 (1979)
[14] C. Hu, *et al.*, *Trans. Elec. Dev.* **ED-32**, p. 375 (1985)
[15] C.N. Berglund and R.J. Powell, *J. Appl. Phys.* **42**, p. 573 (1971)
[16] E.O. Kane, *Phys. Rev.* **159**, p. 159 (1967)
[17] S.J. Wang, *et al.*, *IEEE IEDM, Dig. Tech. Papers*, p. 46 (1987)
[18] M. Pinto, *et al.*, "PISCES II: Poisson and continuity equation solver", Stanford (U.) Electronics Lab Tech. Report (September, 1984).
[19] A.K. Henning, "Hot carrier effects in CMOS field effect transistors at cryogenics temperatures", Ph.D. dissertation, Stanford University (1987). Available from University Microfilms, Ann Arbor, MI.
[20] B. Meinerzhagen, *IEEE IEDM, Dig. Tech. Papers*, p. 504 (1988)

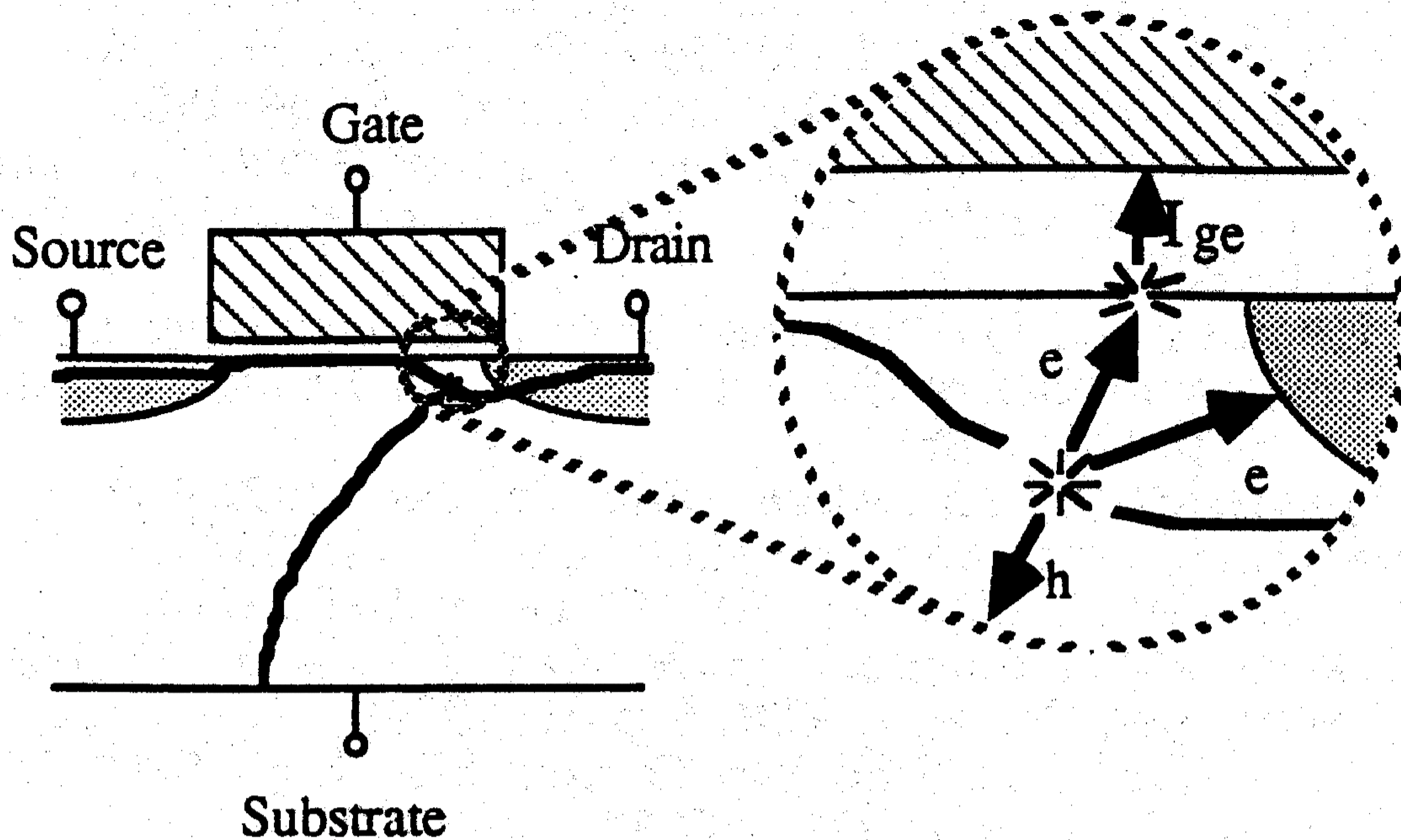


Figure 1. Schematic of impact ionization. Channel current acquires enough energy to break a Si-Si bond. The blow-up shows some of the subsequent processes, including collection of the secondary carriers by the substrate, drain, and gate.

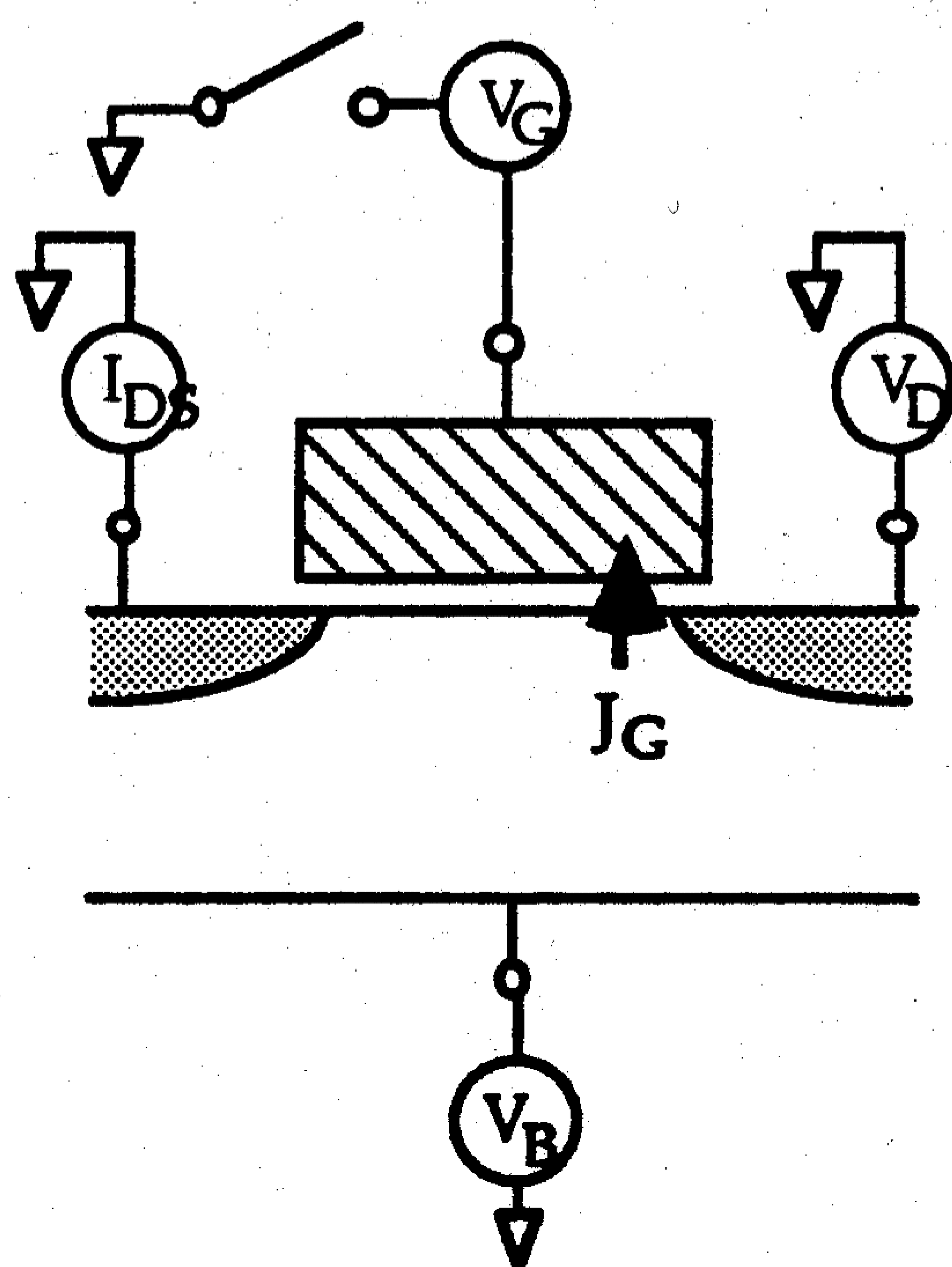


Figure 2. Schematic of the floating gate measurement technique.

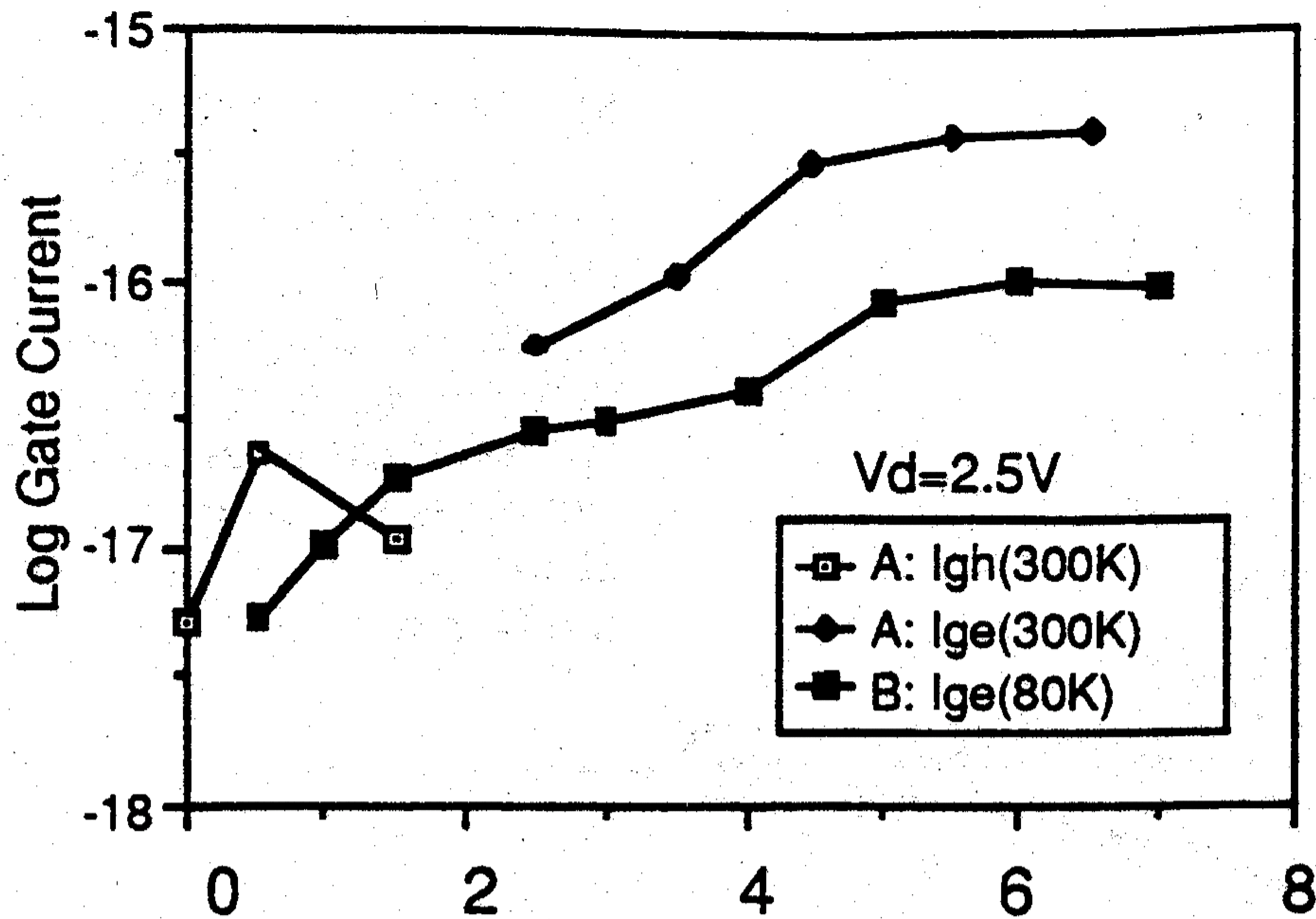


Figure Three: Gate current characteristics for Vd=2.5V, at low and room temperature. Measured on Le=0.8μm, N-channel MOSFETs.

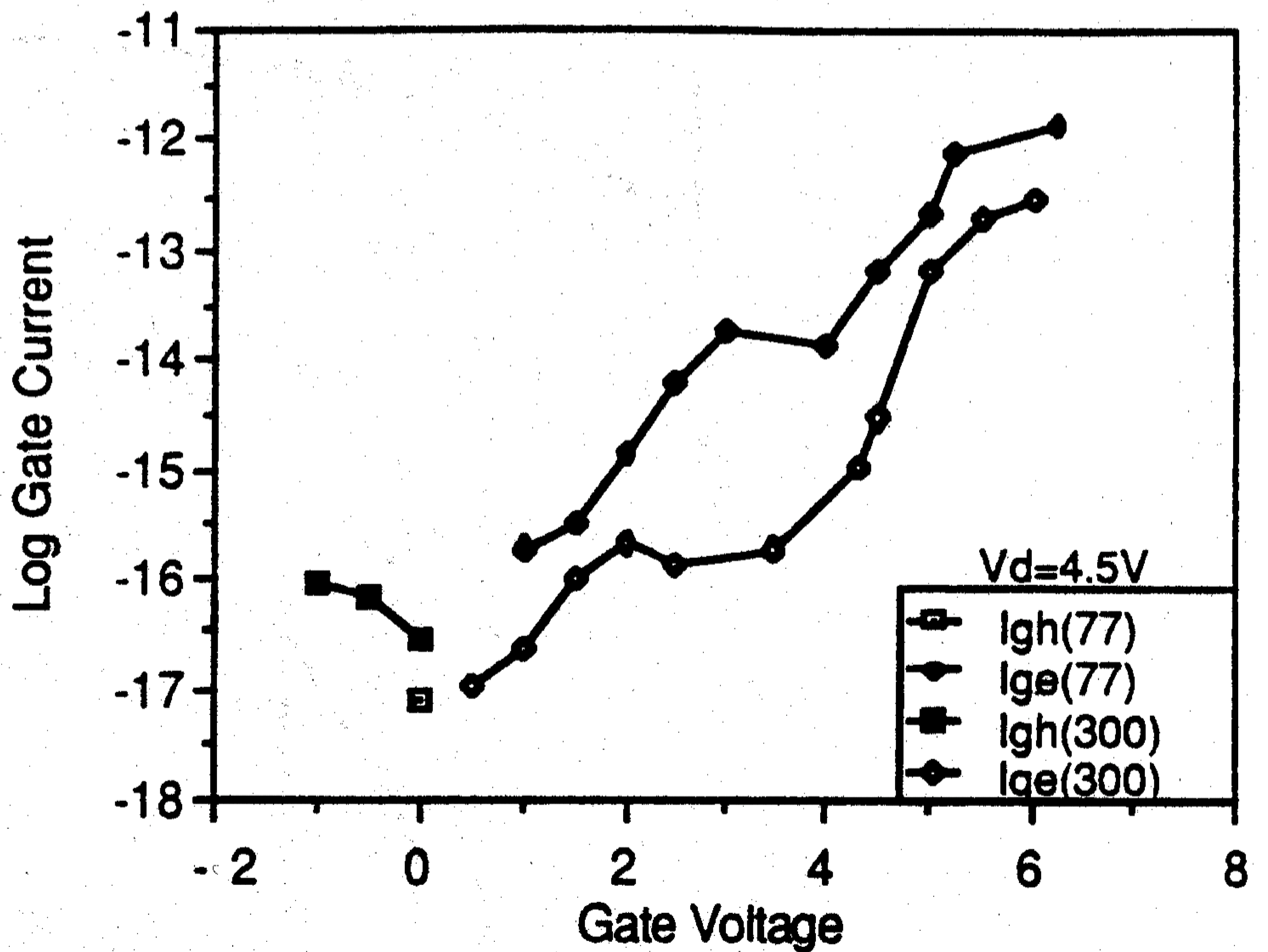


Figure Four: Similar to Fig. Three, but with Vd=4.5V.

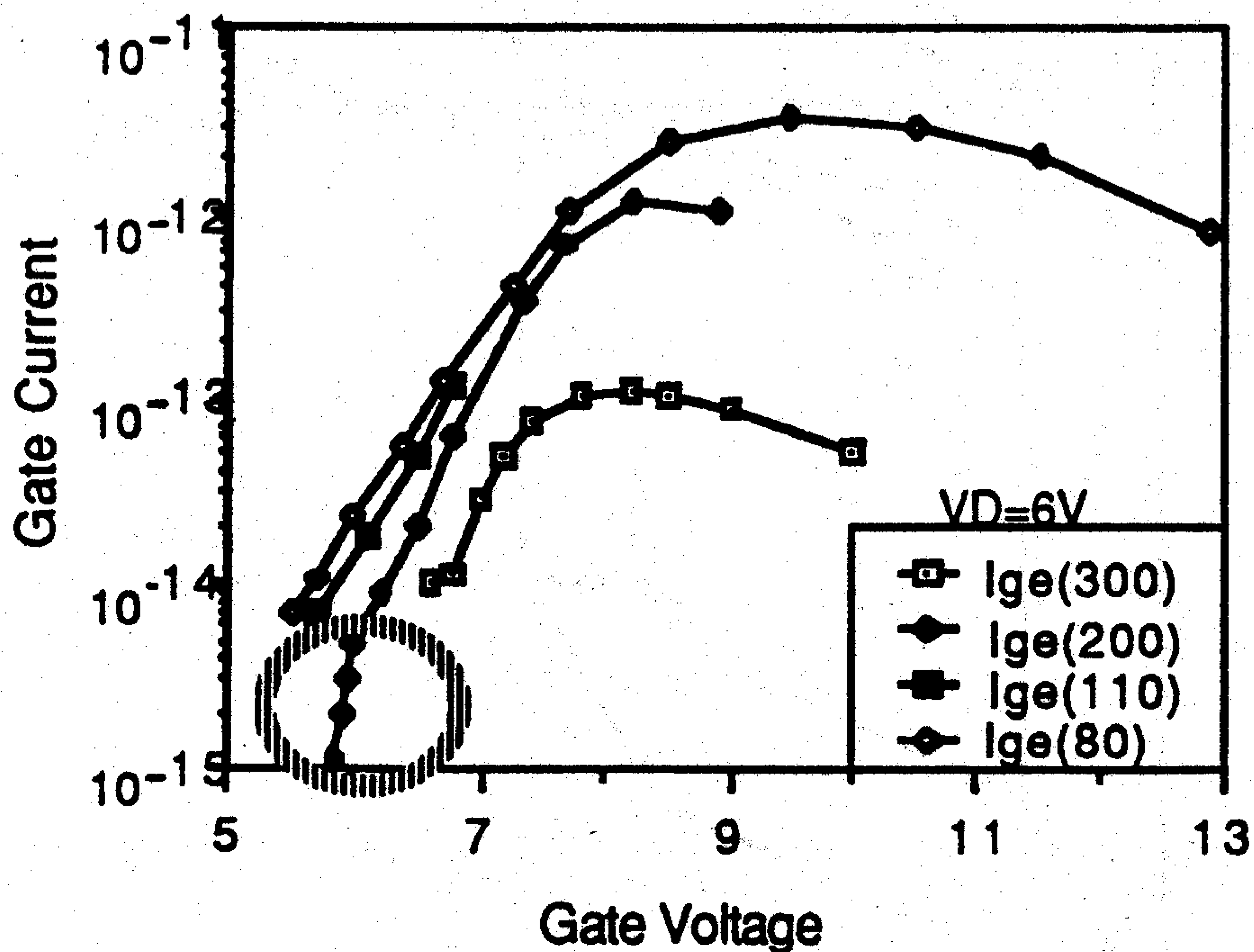


Figure Five: Temperature characteristics of gate current measured with the floating gate technique. The circled area shows signs of interface charge build-up, resulting in a higher slope of (log Ig) vs. (Vg). The time for the floating gate technique to sketch out this low current area is extremely long compared to the time (~ a few seconds) to develop the curves at higher Ig.

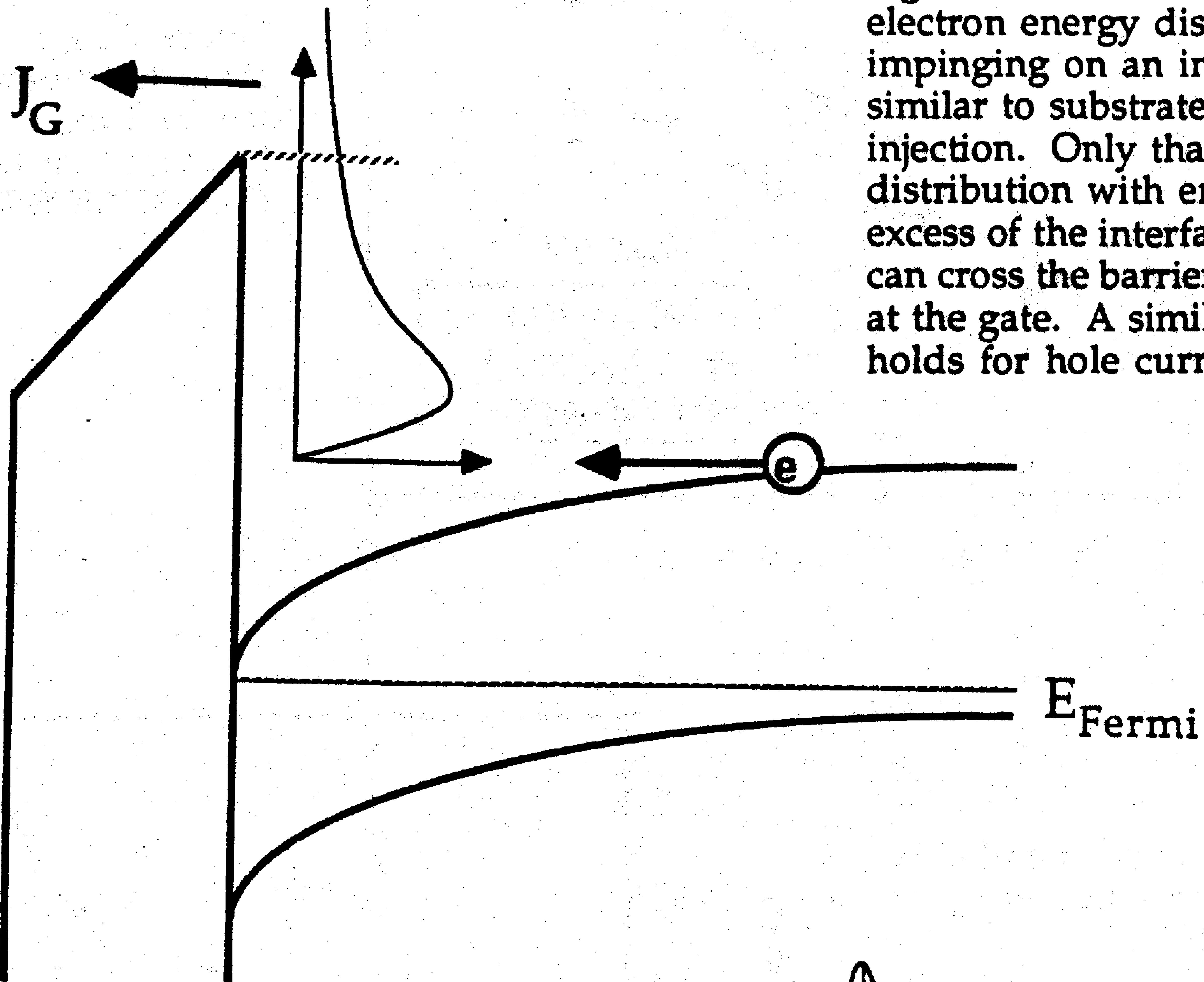


Figure Six: 1-D schematic showing electron energy distribution impinging on an insulator interface, similar to substrate hot electron injection. Only that portion of the distribution with energy in excess of the interface barrier height can cross the barrier and be collected at the gate. A similar configuration holds for hole current injection.

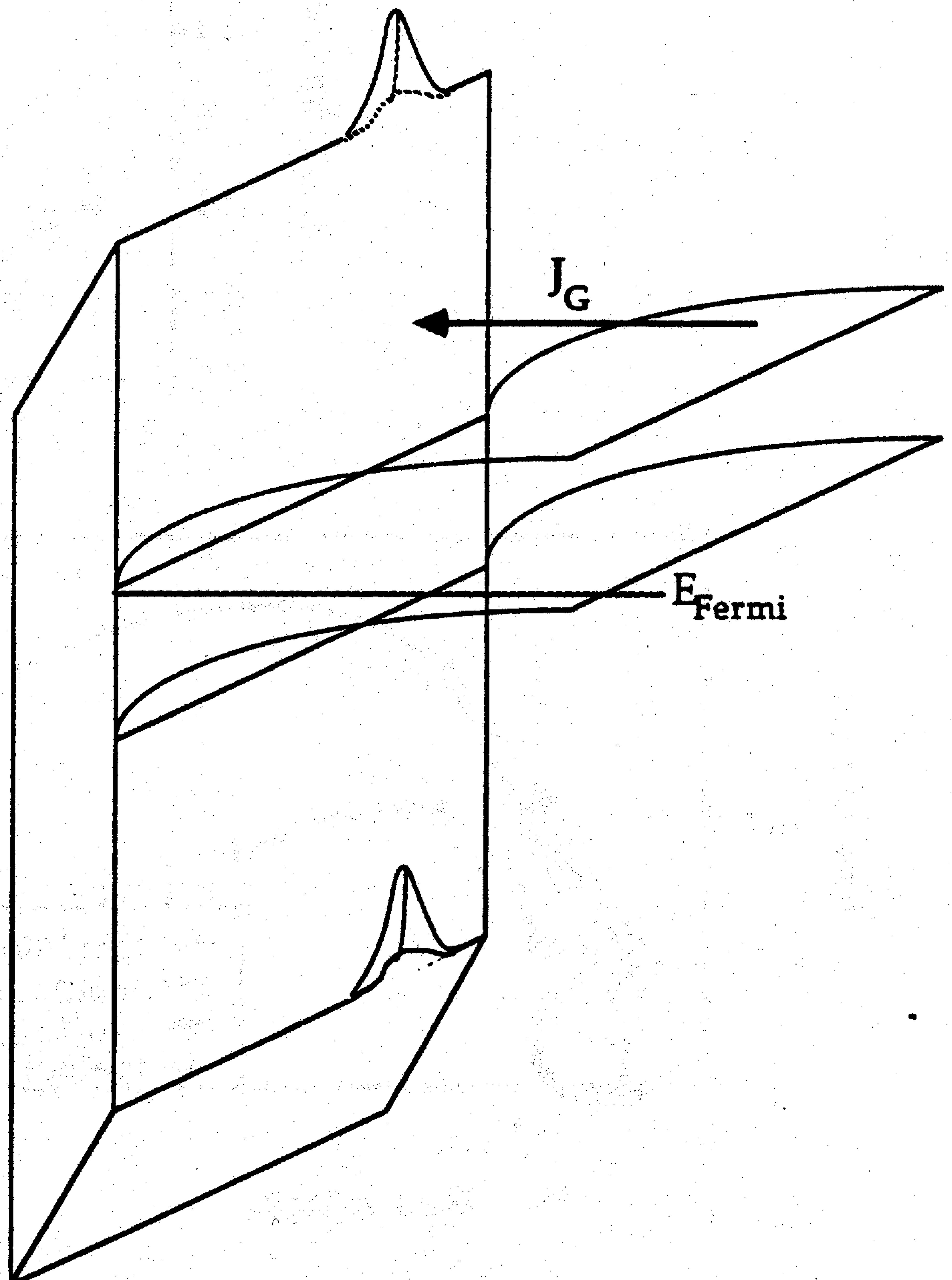


Figure Seven: 2-D schematic showing the result of charge transport across the interface. Generation and/or filling of charged interface states yields an increase in the interfacial barrier at the local point of injection. The barrier for the opposite charge carrier is reduced by the same amount. In the schematic, electron injection has increased the barrier to electron transport, and lowered the barrier to holes.

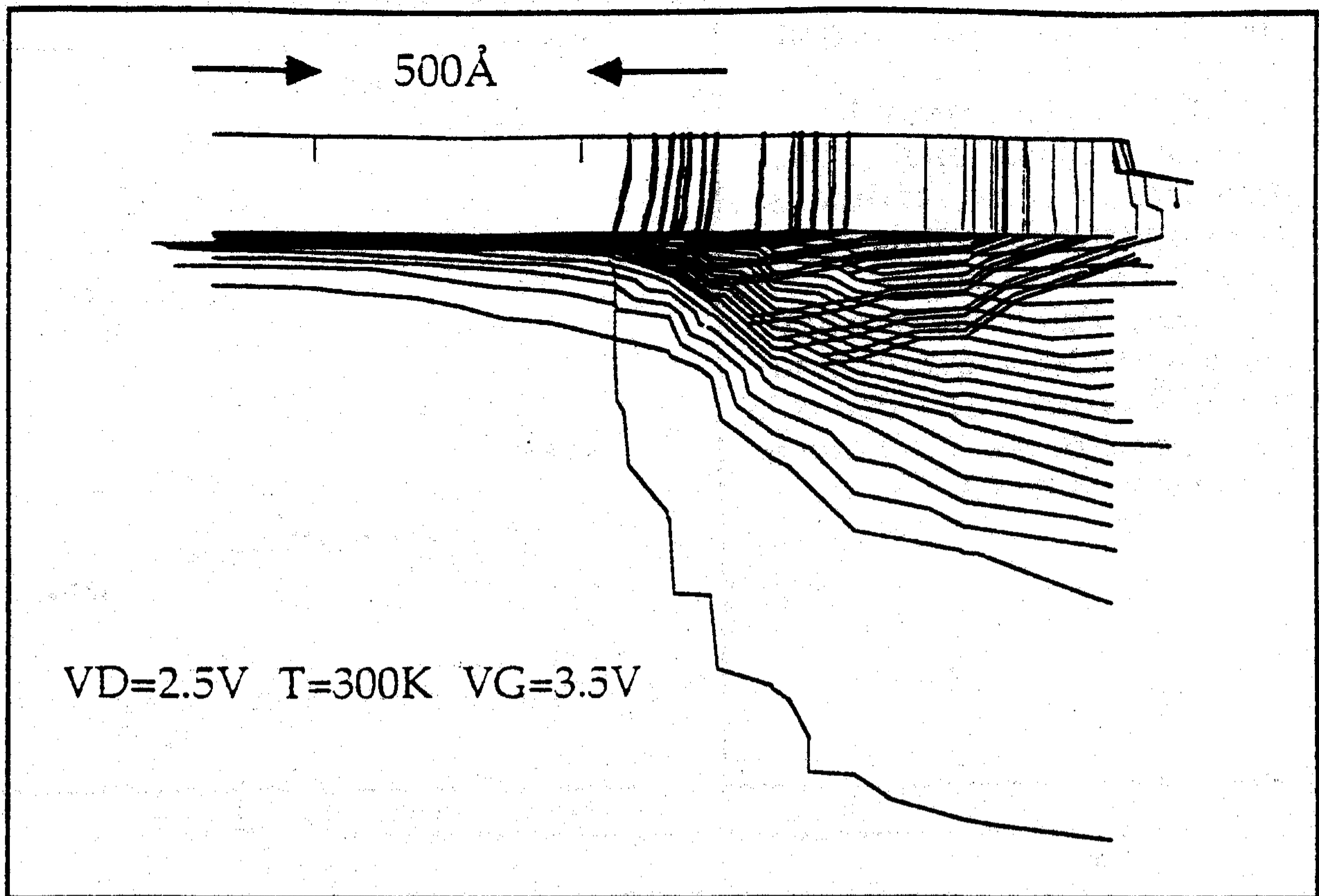


Figure Eight: Electron channel current contours and electron gate current emission contours for the conditions shown (see Figure Three). For reference, the oxide thickness is 385 Å. Only the gate current contributions from two channel current contours are shown, for clarity. Note that VG>VD.

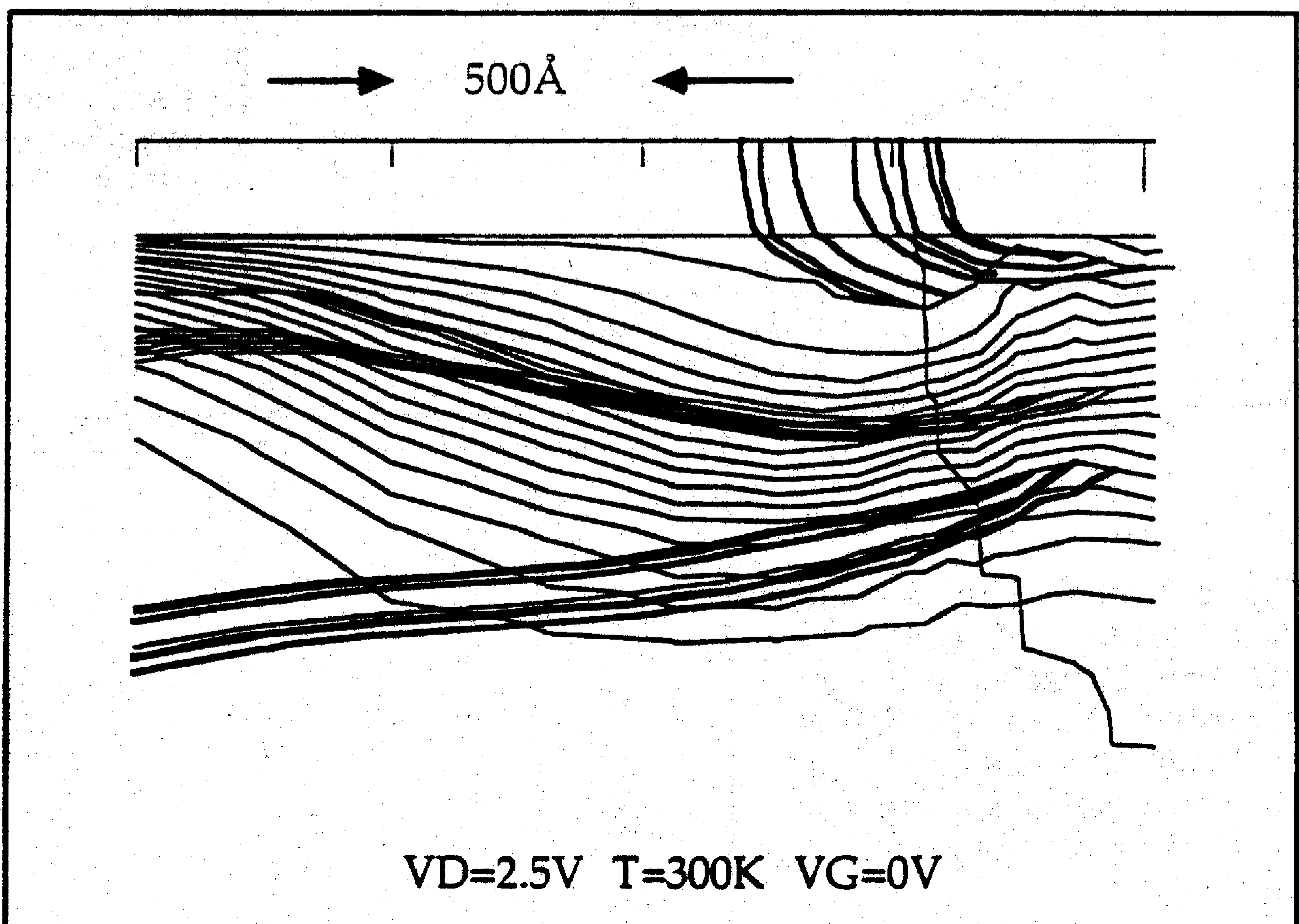
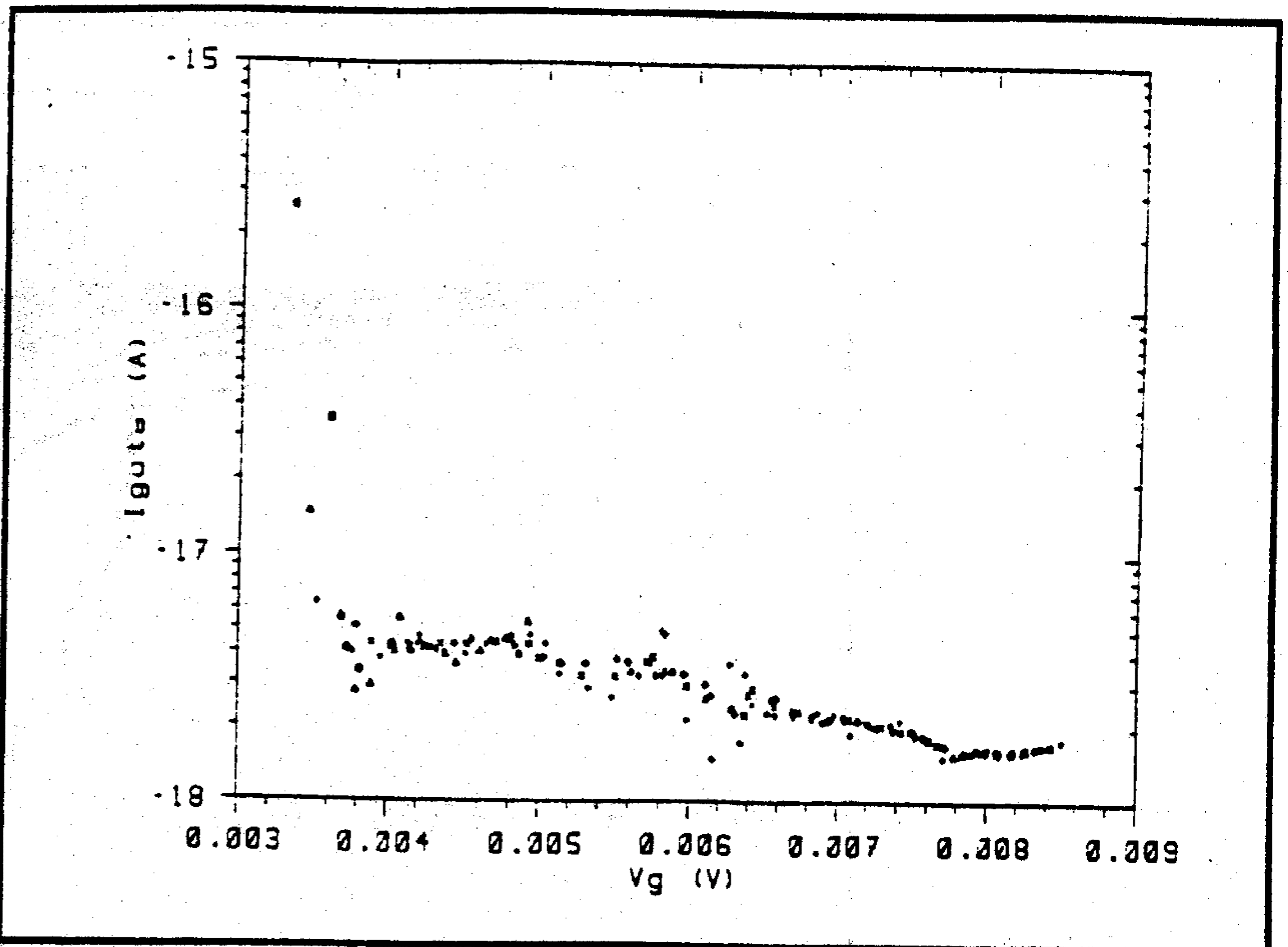


Figure Nine: Electron channel current contours and hole gate current emission contours for the conditions shown (see Figure Three). Only the gate current contributions from three channel current contours are shown, for clarity. Note that VG<VD.

Figure Ten: Full gate current curve for VD=2.5V and VG=0V, extracted from the transconductance, and IDS vs. time data for these conditions. The slope indicates hole emission across the gate insulator. Points are derived by applying least squares fit to five IDS vs. time data points in the vicinity of the VG plotted, in order to obtain ΔIDS/Δt.



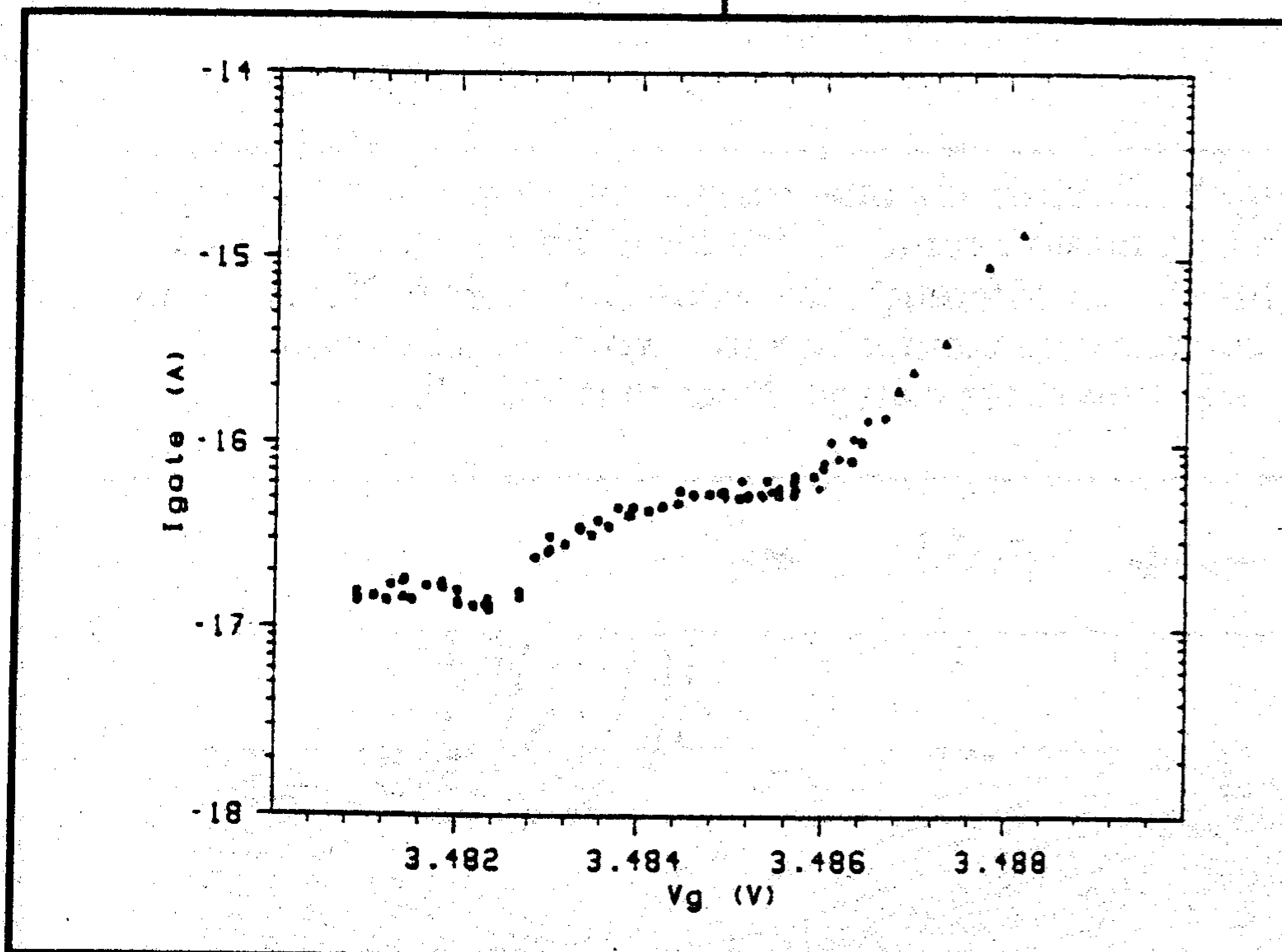


Figure Eleven: Full gate current curve for VD=2.5V and VG=3.5V, extracted from the transconductance, and IDS vs. time data for these conditions. The slope indicates electron emission across the gate insulator.

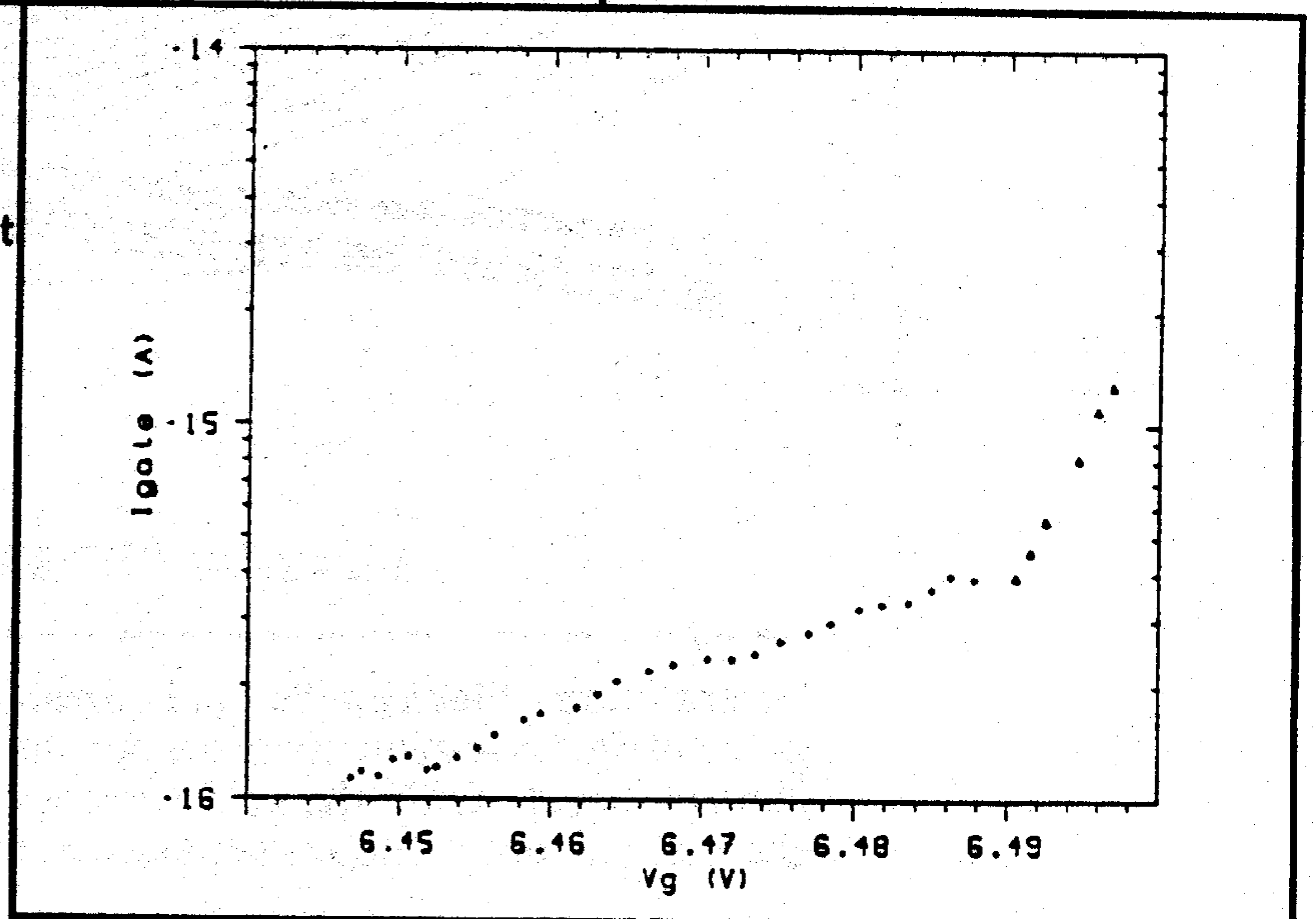
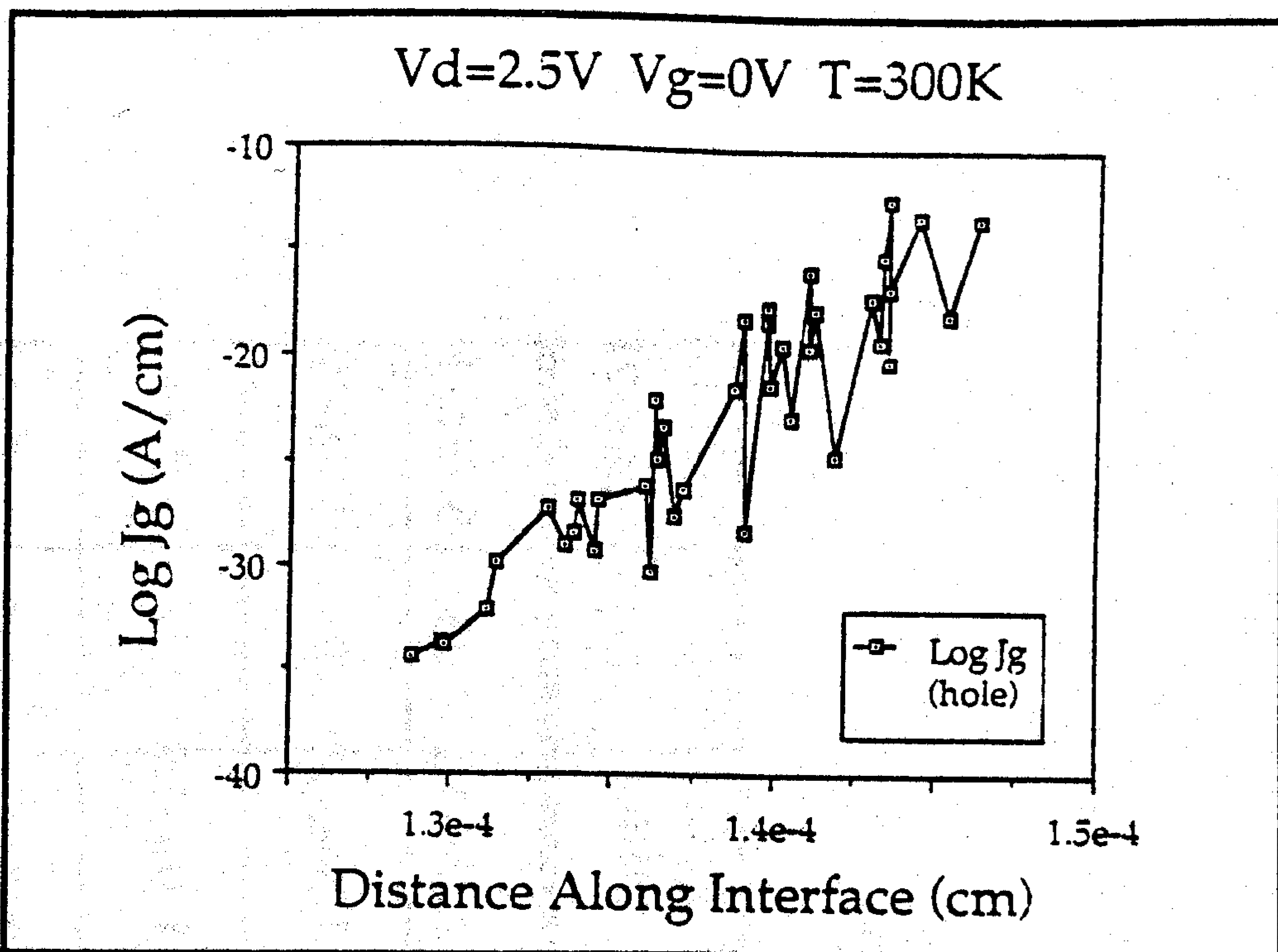


Figure Twelve: Full gate current curve for VD=2.5V and VG=6.5V, extracted from the transconductance, and IDS vs. time data for these conditions. The slope indicates electron emission across the gate insulator.

Figure Thirteen: Hole current density along the Si-SiO2 interface for the conditions in Figure Ten. Scatter is due to the mesh discretization along the interface. The metallurgical junction is located at 1.4μm. Note that the current density has units appropriate for a 2-D simulation.



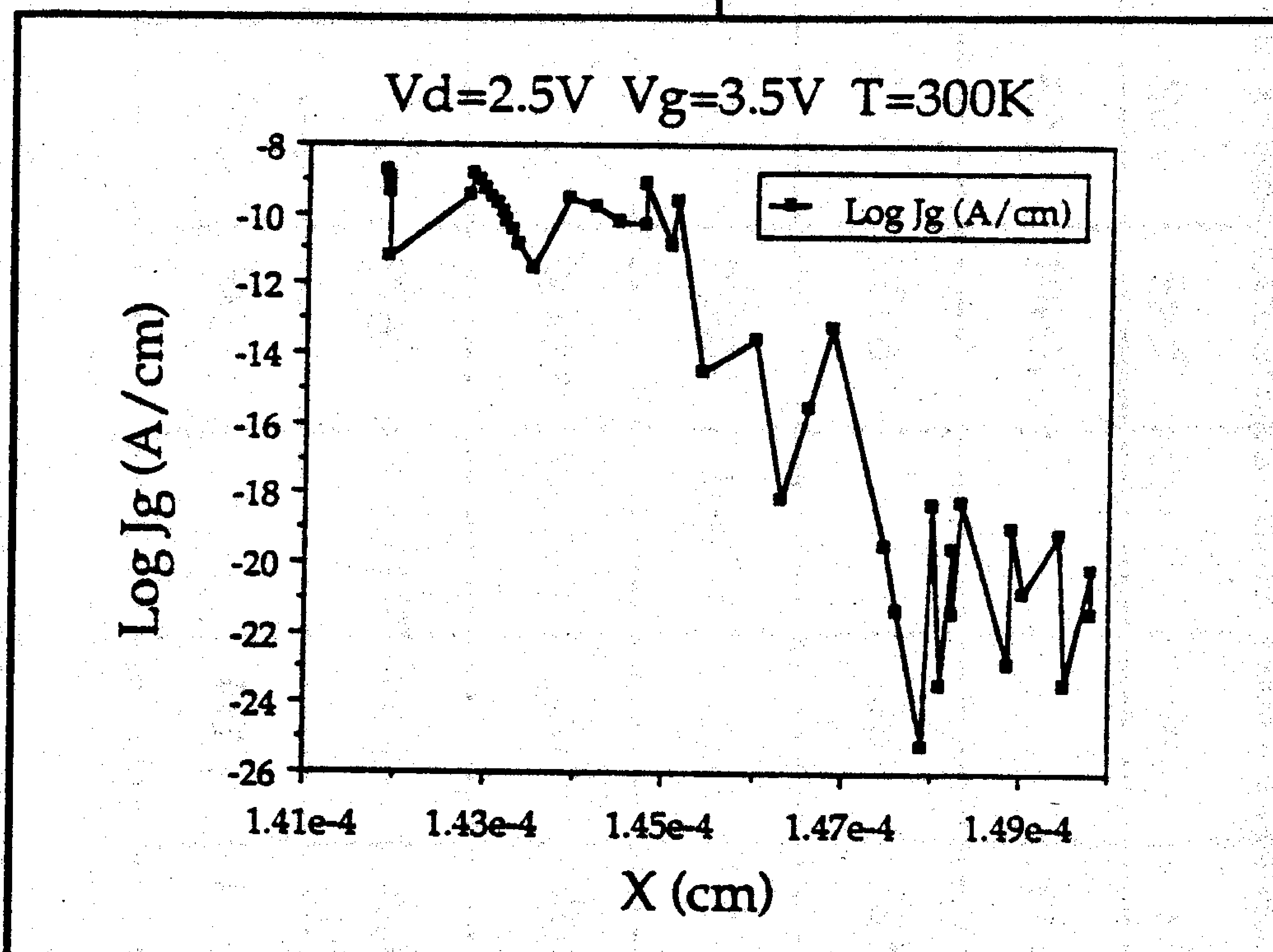


Figure Fourteen: Electron current density along the Si-SiO2 interface for the conditions in Figure Eleven.

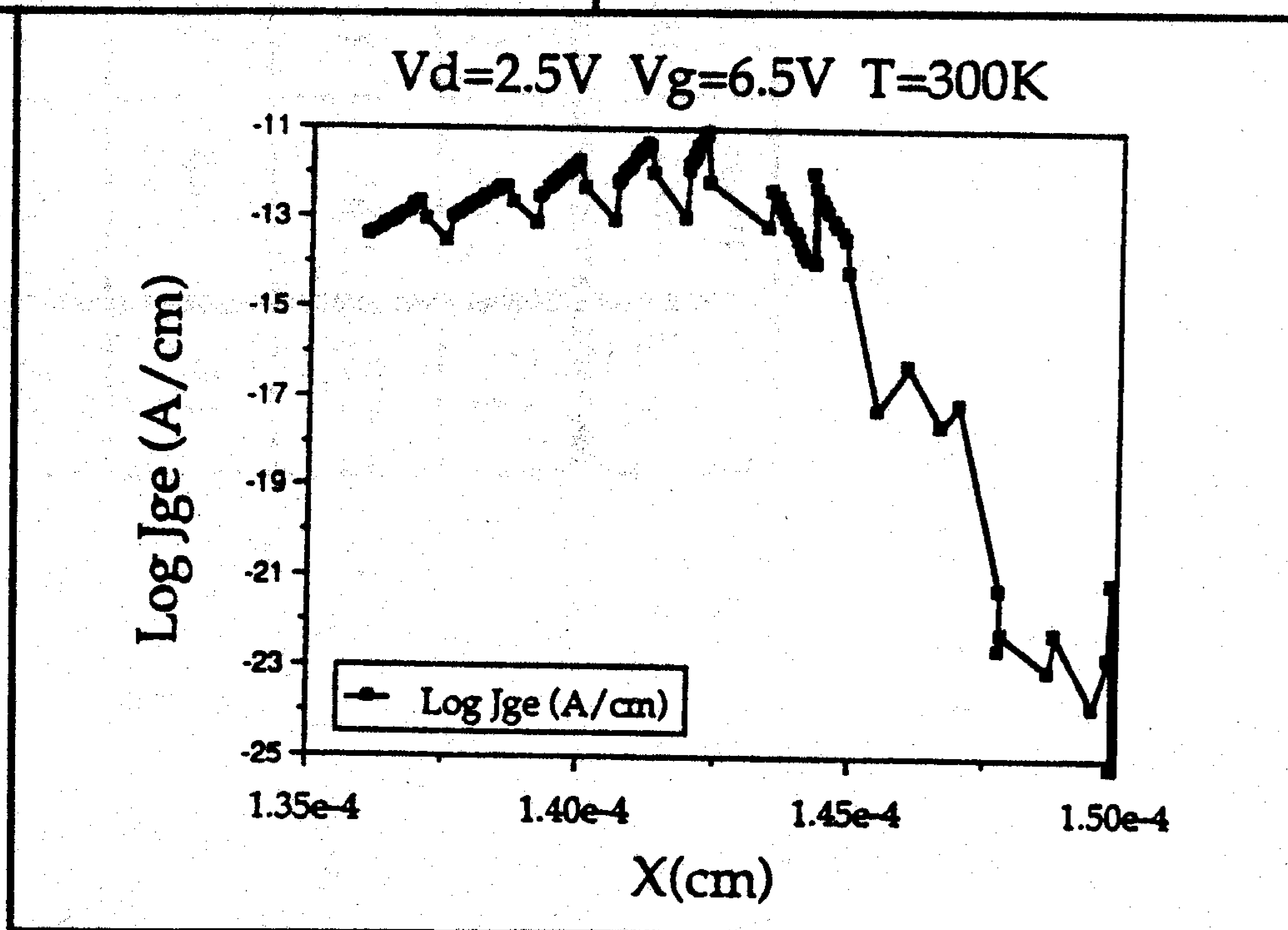


Figure Fifteen: Electron current density along the Si-SiO2 interface for the conditions in Figure Eleven.

Vd=2.5V T=300K

VG	time (sec)	Xcen (μm)	Δx (Å)	Jg (A/cm)	Te (K)	Jg1/ Jg2	Ig (A) [sim]	Δø (eV)	ΔNit (cm-2)
0	1600	1.38	~180	1.0e-12	7845	4.62	2.35e-18	1.034	5.79e11
0.5	1000	1.431			10901	8.33	1.04e-14	1.99	1.12e12
1.5	540	1.421	~50		6537	3.33	1.16e-15	.677	3.79e11
3.5	200	1.424	~120	1.7e-9	14650	4.62	8.09e-16	1.93	1.08e12
4.5	200	1.425	~180	2.67e-9	13307	3.0	1.02e-16	1.26	7.06e11
5.5	100	1.423	~200	7.11e-11	12673	3.75	2.03e-17	1.44	8.09e11
6.5	100	1.423	>200		12544	4.0	5.34e-18	1.50	8.40e11

Table One: Calculated gate current injection parameters for the temperature and biases shown.