

From FinFET to Gate-All-Around Nanosheets

Subthreshold swing, DIBL, the "natural length" argument for wrapping the gate, and effective width.

1. Dennard scaling and why it ended

Scale every dimension and the supply voltage by $1/\kappa$ and doping by κ :

Quantity	Dimensions, V	Electric field	Gate delay	Power / transistor	Power density
Scales as	$1/\kappa$	1	$1/\kappa$	$1/\kappa^2$	1

This relies on threshold voltage scaling too. But off-state current grows exponentially as V_T drops (Section 2), so by the mid-2000s V_T , and therefore V_{DD} , stopped scaling. With dimensions still shrinking, fields and power density rose instead. That is where clock frequency stalled and multicore began.

2. Subthreshold swing and the 60 mV/decade floor

Below threshold, current is diffusion over a barrier controlled by the surface potential, so it is exponential in gate voltage:

$$I_D \propto \exp\left(\frac{q(V_{GS} - V_T)}{n kT}\right), \quad S \equiv \frac{dV_{GS}}{d \log_{10} I_D} = n \frac{kT}{q} \ln 10, \quad n = 1 + \frac{C_{\text{dep}} + C_{\text{it}}}{C_{\text{ox}}}.$$

At 300 K, $kT/q = 25.9$ mV, so $S \geq 59.6$ mV/decade, reached only when $n = 1$, i.e. when the gate fully controls the channel potential. The floor is set by the Boltzmann tail of carrier energies, not by fabrication quality. Steep-slope devices such as tunnel FETs try to beat it by changing the injection mechanism.

Why swing matters: leakage

Extrapolating the subthreshold line from V_T down to $V_{GS} = 0$: $I_{\text{off}} \approx I_T \cdot 10^{-V_T/S}$, where I_T is the current at threshold. Every extra mV/decade of swing costs leakage exponentially at a fixed V_T .

3. Short-channel effects

DIBL (drain-induced barrier lowering): the drain field lowers the source barrier, so threshold falls as V_{DS} rises:

$$\text{DIBL} = - \frac{V_T(V_{DS,\text{high}}) - V_T(V_{DS,\text{low}})}{V_{DS,\text{high}} - V_{DS,\text{low}}} \quad [\text{mV/V}].$$

Well-controlled modern devices sit at a few tens of mV/V. Short planar devices that have lost electrostatic control exceed 100 mV/V. V_T **roll-off** is the related drop of threshold with shrinking gate length. Both raise off-state leakage.

4. The natural length: why more gate sides win

Solving Poisson's equation in a thin body shows that potential disturbances from source and drain decay into the channel as $e^{-x/\lambda}$. The characteristic **natural length** λ shrinks as the gate wraps more of the body. For a body of thickness t_{si} and oxide thickness t_{ox} , a widely used approximation (Colinge) with N effective gates is

$$\lambda_N \approx \sqrt{\frac{\epsilon_{\text{si}}}{N \epsilon_{\text{ox}}}} t_{\text{si}} t_{\text{ox}}, \quad N \approx 1 \text{ (single gate)}, 2 \text{ (double)}, 3 \text{ (tri-gate / fin)}, 4 \text{ (gate-all-} \rightarrow$$

To suppress short-channel effects the gate length should be several natural lengths long, a common rule of thumb being $L_g \gtrsim 5\text{--}10 \lambda$. A smaller λ lets L_g scale further.

Worked comparison

Take $t_{\text{si}} = 6 \text{ nm}$, $t_{\text{ox}} = 1 \text{ nm}$ equivalent, and $\epsilon_{\text{si}}/\epsilon_{\text{ox}} = 11.7/3.9 = 3$:

Gate	N	λ	$L_g \approx 5\lambda$
Single (thin-body planar)	1	4.2 nm	21 nm
Double gate	2	3.0 nm	15 nm
Tri-gate (FinFET)	3	2.4 nm	12 nm
Gate-all-around	4	2.1 nm	11 nm

Illustrative only: real fins and sheets are not ideal slabs, and the "effective N " depends on cross-section shape and corner effects. The trend is the point.

5. Effective width: quantized fins vs. tunable sheets

$$W_{\text{eff}}^{\text{FinFET}} = n_{\text{fin}} (2H_{\text{fin}} + W_{\text{fin}}), \quad W_{\text{eff}}^{\text{nanosheet}} = n_{\text{sheet}} \cdot 2 (W_{\text{sheet}} + T_{\text{sheet}}).$$

Illustration: a fin with $H = 50 \text{ nm}$ and $W = 6 \text{ nm}$ gives about 106 nm per fin, so drive comes in steps of about 106 nm. Three stacked sheets of $T = 5 \text{ nm}$ give $6(W_{\text{sheet}} + 5) \text{ nm}$, so drawing W_{sheet} anywhere from about 15 to 45 nm spans roughly 120–300 nm continuously. Designers regain a sizing knob that fin quantization had taken away.

6. Self-test, worked

Exercise 1: swing vs. leakage

Same $V_T = 0.30$ V and same I_T . With $I_{\text{off}} \approx I_T 10^{-V_T/S}$:

S (mV/dec)	V_T/S (decades)	I_{off}/I_T	Relative leakage
60	5.0	1×10^{-5}	1×
90	3.3	4.6×10^{-4}	46×
100	3.0	1×10^{-3}	100×

Put another way: to keep the 60 mV/dec leakage with a 100 mV/dec device you would need $V_T = 0.50$ V, which leaves little gate overdrive at a 0.7 V supply. Multiply per-transistor leakage by billions of transistors and you have a chip's standby power.

Exercise 2: why gate-all-around wins twice

Electrostatics: each added gate side terminates field lines from the drain that would otherwise reach the source barrier. In natural-length terms, λ shrinks roughly as $1/\sqrt{N}$, so DIBL and swing degradation set in at shorter L_g : GAA beats FinFET, which beats planar. *Design*: fin width is quantized in whole fins of fixed height, while sheet width is drawn per device, so drive strength becomes nearly continuous. One change improves leakage and sizing granularity together.