Episode Book

EPISODE #5749

From Sand to Trillion-Dollar Chips: Semiconductors in 5 Levels

PUBLISHED	6 October 2026 at 23:56 UTC
DURATION	25:33
LISTEN	myweirdprompts.com/episode/semiconductors-five-levels-explained/
HOSTS	Corn, Herman, Hilbert
SCRIPT MODEL	deepseek-flash
VOICES	chatterbox-regular (chatterbox-local)
PIPELINE	V5.2
BOOK GENERATED	19 October 2026 at 17:39 UTC

SYNOPSIS

Pure silicon is a disappointing material: it barely conducts and barely insulates, and you could hold a crystal in your hand and watch it do nothing. So why is it the foundation of a trillion-dollar industry? This episode climbs five levels of complexity, starting with the basic physics of conductors, insulators, and the narrow band gap in between, then descending into electrons, holes, and the doping process that turns a useless crystal into a working device. From there it's p-n junctions, diodes, and the three-terminal transistor that made amplification and switching possible — and with them, the computer. The top rung lands on the advanced physics governing the chips being fabricated right now, where devices are small enough that quantum tunneling works against the very engineering that created them. Along the way: the numbers behind the industry's record-breaking year, and why the best material for the job is the one that's mediocre at both extremes.

TRANSCRIPT



Corn

Okay. I want to open with a confession, because it's relevant.



Herman

That's never a good sign.



Corn

I have been looking at a computer for most of my life, and until this week I could not have told you what a semiconductor actually does. Like, functionally. I knew the word. I knew it was in there somewhere. I couldn't have explained it.



Herman

And now?



Corn

Now I can, mostly, and I'm annoyed that nobody told me earlier.



Herman

That's a decent place to start, because Daniel's prompt this week is essentially that exact gap. Here's how he put it. He wants us to explain why semiconductors are such important electronic components, and he wants it in five levels of escalating complexity, starting with the most basic explanation and working up to an expert level one.



Corn

Five levels. He's very specific about the ladder. And the questions under that ladder are all real ones. What is a semiconductor versus a conductor or an insulator. How does it actually work at the level of electrons and energy bands. How do you get from a raw crystal of this stuff to an actual diode or transistor. Why does any of that translate into the



Corn (continued)

modern world. And then the top rung, which is the one I'm nervous about, what advanced physics governs the chips being made right now.



Herman

So we climb. Sand that sometimes conducts, all the way up to quantum tunneling in a transistor so small that its own physics is working against it.



Corn

And the funny thing is those two ends are the same object. That's the bit that grabbed me. So let's start at the bottom, because there's an actually good answer at the bottom.



Herman

The plain-language answer. A semiconductor is a material with a filled valence band, an unfilled conduction band, and a relatively small energy gap between them. That's the textbook version and it's correct, but the useful version is simpler than that. In a conductor, electrons move freely, there's no gap to cross. In an insulator, the gap is enormous and the electrons are essentially locked in place. A semiconductor sits in the middle.



Corn

Middle meaning mediocre.



Herman

Middle meaning mediocre, and this is the part people miss. Silicon's band gap is one point one four electron volts. Small enough that ordinary thermal energy at room temperature kicks a few electrons across. Large enough that the material isn't naturally conductive. So it's a bad conductor.



Corn

And a bad insulator.



Herman

Structurally, yes. You could hold a piece of pure silicon in your hand and it would do nothing useful. It wouldn't conduct a meaningful current and it wouldn't insulate particularly well either. By any normal engineering standard it's a disappointing material.



Corn

Which is where I want to stop, because the obvious follow-up is: so why is it the most important material in civilization. And the answer is one word, and that word is control.



Herman

That's the whole thing.



Corn

A conductor is a conductor. You don't get to negotiate with copper. An insulator insulates. Silicon, by contrast, can be pushed either way. Doping pushes it. Heat pushes it. Light pushes it. An applied voltage pushes it. The material isn't valuable because it conducts well. It's valuable because it's *switchable*, and a switch is the only thing a computer actually needs.



Herman

That's the paradox in its purest form. The best material for the job is the one that's mediocre at both extremes. If it were better at conducting, you couldn't turn it off. If it were better at insulating, you couldn't turn it on.



Corn

And just to put a number on how far that's gone, the industry passed one trillion dollars in year-to-date sales this year for the first time ever. One point one four electron volts on one end, a trillion dollars on the other.



Herman

And both of those things happening at the same time as the physics that made the scale possible is hitting hard walls. That's the tension running under this whole episode. Keep it in your back pocket.



Corn

Right. So we've answered level one, which is that a semiconductor is a bad conductor and a bad insulator and that this is the point. But that answer is a description, not an explanation. It tells you what the material is like from the outside. To actually understand it, we have to go down to where the electrons are.



Herman

This is level two, and this is where it gets beautiful, because the mechanism is strange. In a semiconductor, current can be carried by two different things at once.



Corn

Electrons being the first one, obviously.



Herman

Electrons, and holes. When an electron gets enough energy to jump from the valence band up to the conduction band, it leaves behind an empty state. And that empty state behaves, for all practical purposes, like a mobile positive charge. It moves through the crystal the way a bubble moves up through water. Nobody's carrying anything. The vacancy is what propagates.



Corn

So you can describe the same current two ways. Either electrons going one direction, or holes going the other.



Herman

Which is not a bookkeeping trick. It's real. The Hall effect confirms it experimentally. You can measure a positive carrier in a material and find that it behaves exactly like a positive particle should, and that measurement is how engineers pull out carrier density and drift velocity. Holes are a legitimate way of describing what's happening in there.



Corn

Okay. Now give me the number that makes silicon sound as useless as it actually is.



Herman

In pure silicon at room temperature, the thermally excited carrier concentration is about one and a half times ten to the tenth electrons per cubic centimeter. That's the honest figure. For comparison, gallium arsenide comes in around one point one times ten to the sixth. Both of those numbers are orders of magnitude below what you'd get in a metal.



Corn

So a raw silicon crystal is, electrically speaking, a brick.



Herman

A brick with a very tidy crystal lattice. Which is why the next step is the one that turns the brick into a device, and it's the closest thing engineering has to alchemy. Doping.



Corn

Explain doping properly, because I've heard the word for years and never had it explained in a way that stuck.



Herman

You take a silicon crystal, where every atom has four valence electrons and everything is chemically satisfied, and you substitute in an impurity atom with a different number of valence electrons. Drop in arsenic, which has five. Four of those electrons fit into the silicon lattice bonds. The fifth one is spare.



Corn

And a spare electron in a lattice that doesn't want it is a loose electron.



Herman

Very loose. This is the number I want people to remember. The binding energy holding that spare electron to its arsenic atom is roughly zero point zero two electron volts. The band gap of the silicon around it is one point one four. So the energy needed to free that donor electron is something like fifty-seven times smaller than the energy needed to promote an electron across the gap.



Corn

So at room temperature it just falls off.



Herman

Room temperature thermal energy is more than enough to liberate it. Which produces n-type silicon, where you've got mobile electrons wandering around. Now do the mirror image. Substitute boron or aluminum, three valence electrons. Now the lattice has a missing electron, an unsatisfied bond, and that vacancy is exactly the hole we were talking about. That's p-type silicon, and its mobile carrier is the hole.



Corn

And how much of this impurity are we actually adding?



Herman

Parts per million. Sometimes less. Tiny concentrations. But the conductivity of the crystal changes by orders of magnitude, because the donor electrons are so weakly bound that essentially all of them are free at operating temperature.



Corn

Which is why the manufacturing is insane.



Herman

The manufacturing is insane. You cannot do this with dirty material, because if there's uncontrolled contamination in the crystal, you have uncontrolled carriers, and the device is garbage. Every transistor starts as an ultra-pure silicon crystal, precisely doped to parts per million, in a fab that cost more than most countries' infrastructure budgets. The doping is a subtle tweak on paper. It's the entire miracle in practice.



Corn

Alright, we've got a doped crystal with mobile carriers. That's a material, not a device. How do you get from there to something that actually does a job?



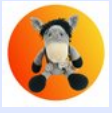
Herman

You put the two types together. Level three. Take a single crystal of silicon, dope one end p-type and the other end n-type, and you've built a p-n junction. And a p-n junction does something remarkable for how simple it is. It only lets current flow one direction.



Corn

One-way valve.



Herman

Almost exactly. At the boundary where the p and n regions meet, carriers diffuse across and recombine, and you get a region at the interface that's been stripped of mobile charge. The depletion region. Apply voltage in the forward direction and you push carriers back into that region, it narrows, and current flows. Apply voltage in the reverse direction and you pull carriers away, the depletion region widens, and the junction blocks.



Corn

And that's a diode.



Herman

That's a diode. The simplest p-n junction device. It conducts one way, blocks the other way. Voltage regulator, rectifier, whichever application you want. That's level three, partly.



Corn

Partly meaning the interesting part is next.



Herman

The interesting part is next. A diode is a valve. A transistor is a valve with a hand on it. You build a device with three terminals. Two of them carry the main current path. The third is a control terminal, and the voltage on that third terminal decides whether current flows between the other two, and how much.



Corn

And that distinction matters more than it sounds like it should.



Herman

It's the whole distinction between an active device and a passive one. A resistor is passive. A capacitor is passive. They do their one thing and that's it. A transistor is active because a small signal at one terminal controls a much larger signal at another. That gives you two things at once. Amplification. And switching.



Corn

And amplification plus switching, at scale, in tiny amounts of power, is a computer.



Herman

A computer is a few tens of billions of transistors wired up so that the output of one becomes the input to the next. Every logic gate, every memory cell, every arithmetic operation, is a particular arrangement of these three-terminal devices flapping back and forth. Everything digital reduces to that.



Corn

So the full arc is: empty crystal, doped crystal, junction, device. Each step actually is an escalation, which is nice, because I was worried this ladder would be five ways of saying the same thing.



Herman

It isn't. And this is the point at which the physics has basically been settled since the nineteen fifties. Diodes and transistors are well understood. Everything after this is not new physics. It's how the economics of the same physics has gone berserk.



Corn

So level four. Why does this translate into a trillion dollars. And I want to lead with the number, because the number is hard to hold in your head.



Herman

August of this year, global semiconductor sales were one hundred fifty-nine point seven billion dollars in a single month. That's up one hundred forty-four percent year over year. Year-to-date sales crossed one trillion dollars for the first time ever, and that's eighteen consecutive months of growth.



Corn

Say that monthly number again.



Herman

One hundred fifty-nine point seven billion. In August. One month.



Corn

And August of last year was sixty-five billion, so the industry roughly two and a half times itself in twelve months.



Herman

Gartner is forecasting one point six trillion dollars in total revenue for the year, up ninety-two percent from eight hundred nine billion last year. And 2027 is projected at one point nine trillion. Omdia's forecast is even hotter, running to about ninety-four percent growth.



Corn

That is not a growth curve. That's a vertical line with a chart attached.



Herman

The driver is entirely AI infrastructure. Five of the major hyperscalers announced something on the order of eight hundred billion dollars of capital expenditure for this year, and about a trillion for next year, overwhelmingly in AI data centers. Nvidia's Data Center segment alone booked seventy-five point two billion in its first quarter of the current fiscal year. That's up ninety-two percent year over year from a business that was already enormous.



Corn

And here's the part I want people to sit with, because it's the strangest fact in this whole episode. High-value AI chips are approaching half of semiconductor revenue. And they're less than zero point two percent of unit shipments.



Herman

That's right.



Corn

So the industry's money and the industry's chips are two different industries now. You've got a small number of enormous, expensive accelerators pulling in half the revenue, and then a vast ocean of cheap microcontrollers and memory and analog parts that make up the actual count.



Herman

It means the semiconductor business is being priced by a handful of design wins at the bleeding edge, not by the volume of chips the world uses. Which is great, as long as that edge keeps paying. And that's the tension I promised you earlier, because we now have to go up one more level and look at what the bleeding edge is actually dealing with physically.



Corn

Level five. The chips being made right now.



Herman

The honest framing is that Moore's Law is not dead, but it has been redefined. For decades, the deal was that every couple of years you shrank the transistor and everything got better automatically. That bargain ended. Dennard scaling, the reason shrinking a transistor also lowered its power, died a long time ago. And below about five nanometers, fab costs double every two to three years.



Corn

So the physics of the shrink stopped paying for the engineering of the shrink.



Herman

Imec's roadmap this year projects zero point three nanometer class nodes by 2038. But contact poly pitch stops scaling at a certain node in 2030. The gains after that come from architecture and packaging, not just from drawing smaller lines.



Corn

Give me the architecture part, because that's the actual physics.



Herman

Take the industry's switch from FinFETs to gate-all-around nanosheet transistors at two nanometers. That switch happened for a specific reason, and it's a good example of physics forcing an engineering decision. A FinFET is a fin of silicon with the gate draped over the top and down two sides. Three-sided gate. It works beautifully down to about a seven nanometer gate length, and then it stops working, because the drain end of the transistor starts to push its own electric field through the channel, and the gate can't hold the barrier up anymore.



Corn

The barrier being the thing that keeps the transistor off when it's supposed to be off.



Herman

The failure is called drain-induced barrier lowering. The transistor starts leaking even when the gate is telling it not to, because the drain's field penetrates too deep into the channel. A three-sided gate simply can't shield against that once the channel gets short enough. So gate-all-around wraps the channel on all four sides. A nanosheet stack with gate material fully surrounding it. Maximum electrostatic coupling, maximum control over the barrier.



Corn

Because the closer you can hold the gate to the whole channel, the less the drain gets to have an opinion.



Herman

And this is where the physics is now openly quantum. Leakage at advanced nodes is dominated by tunneling effects. Band-to-band tunneling. Gate-induced drain leakage. And direct source-to-drain tunneling, where an electron that shouldn't be able to get from one end of the transistor to the other just appears there, because the barrier has gotten thin enough that the wavefunction passes through it. Direct source-to-drain tunneling is widely considered the thing that sets the end of Moore's Law as classically understood.



Corn

Hold on. Say that plainly. The electron is going through a wall.



Herman

There isn't a hole in the wall and it isn't going over the top. At that scale, the barrier is thin enough that quantum mechanics just lets some fraction of the electrons appear on the other side. Which is a fine thing in a purpose-built tunneling device. It's terrible in a switch, because it means the switch is never fully off.



Corn

Which is a beautiful way to say the transistor has become a leaky bucket.



Herman

And the floor on that is hard. There's a quantity called subthreshold swing. It's the number of millivolts you have to add to the gate voltage to increase the current by a factor of ten. And at room temperature, the thermodynamic limit on that is about sixty millivolts per decade. You cannot beat it. It's not an engineering problem, it's a thermodynamics problem. And the best nanosheet gate-all-around devices measured are sitting around sixty-three millivolts per decade.



Corn

Which is basically sitting on the floor.



Herman

Basically sitting on the floor. And everybody knows they can't dig through it.



Corn

So there's a wall on one side, which is quantum tunneling making switches leaky, and there's a wall on the other side, which is the thermodynamic floor on how sharply you can turn one on. And those are two separate walls.



Herman

Two separate walls, closing on the same spot. And notice none of that has anything to do with whether the engineering is clever. You can have the best lithography ever built and still be up against the fact that an electron is a wave and a room is hot.



Corn

So the industry's answer is to stop trying to win on those two axes.



Herman

Increasingly, yes. That's where the wider materials story comes in. GaN, SiC, gallium oxide. Wide-bandgap materials for power electronics, and they behave differently enough that they take some of the pressure off. The interesting bottleneck there is thermal. Diamond has some of the highest bulk thermal conductivity of any material, but if you bond it to GaN, the boundary between the two becomes the dominant thermal bottleneck. You've got this incredible heat spreader attached to the thing it's supposed to be cooling, and the interface between them is where the heat gets stuck.



Corn

The roads are paved with gold and the on-ramp is gravel.



Herman

That is exactly the situation. And one more, because it's worth flagging for the shape of what's coming. China's IMECAS demonstrated experimental stacked nanosheet gate-all-around transistors targeting three nanometers without using EUV lithography at all.



Corn

Which matters because it means the manufacturing race might not stay a single track.



Herman

It means there's a second route being explored, which is a different dynamic than anything in the last thirty years of this industry.



Corn

And IBM showed a sub-one nanometer technology at zero point seven nanometers back in June.



Herman

The roadmap is still stretching. Zero point three nanometer class nodes are projected for 2038. The point isn't that scaling has stopped. The point is that from here on, the engineering has to work around walls rather than through them.



Corn

That's the view from the roadmap. All of it, though, is happening in physical objects, made by people, in rooms. Which is a side of this that no amount of physics papers quite covers.



Hilbert

The field is called the exclusion zone.



Corn

How big.



Hilbert

Big as a car park. You can't go in without the suit, and the suit takes twelve minutes to put on.



Herman

Twelve minutes.



Hilbert

There was a woman on front desk, Rosa. She could do it in nine. I never got under eleven. I was on nights for a while at a place outside Grenoble, late eighties. Handling the wafers between steps.



Corn

Handling meaning moving them.



Hilbert

Moving them. Mechanical stage to mechanical stage, in a sealed carrier, and you don't touch the wafer, obviously. You touch the carrier. The wafer itself is in a little pod and it's got a serial on it and you log every move by hand on a clipboard. The clipboard has to be a special clipboard, because a normal one sheds.



Corn

The clipboard sheds.



Hilbert

Paper sheds particles. Everything sheds. You walk, you shed. The whole room is built to get rid of you the moment you're done being useful. Ceiling pulls air straight down through the floor, so you're standing in a column of air moving about a mile a minute in one direction. That's the reason the suit's a full body thing and not just gloves.



Herman

A mile a minute down. That's the entire ceiling area of the room.



Hilbert

Filtration's the cost. The air handling is most of what the building does. The building is a lung with some rooms attached.



Corn

What actually goes wrong.



Hilbert

Usually nothing. That's the job. Ninety-nine times out of a hundred the wafer comes out fine and you log it and move on. But I had one batch, twenty-five wafers, run of gate oxide, and twenty-four came out clean and one came back from the probe with every die failing in the same pattern. Sent it up to test. They found one particle. One. Somewhere in the oxide.



Herman

Grain of dust.



Hilbert

Finer than that. Looked like nothing had happened. The wafer looks identical to the good ones. Same color, same finish, no mark on it. If you'd handed me the stack and said pick the bad one I couldn't have done it.



Corn

The failure is invisible until it's tested.



Hilbert

The failure is always invisible. That's the whole business. The structures are invisible, the particle is invisible, the defect is invisible. You are making a hundred million objects per wafer that no eye can ever check and the only reason it works is that the process is exactly the same every single time. That was the job. Not cleaning, particularly. Repetition.



Corn

Getting the same thing to happen the same way on a Tuesday night.



Hilbert

On every night. I had a manual for the stage. Two hundred pages for a stage that does one motion. I kept the manual after I left, when they moved the line to a newer stage and the old one got scrapped. I don't have the stage. I have the manual.



Herman

Because it isn't about the machine.



Hilbert

The manual's the part that says what we did.



Corn

Well. That's level five with a name on it.



Herman

There's a line between physics and the people who do it, and Hilbert just walked across it. Every number in this episode lives inside a building like that. Every one of those trillion-dollar months is made out of wafers that a woman on nights in Grenoble moved from one stage to another with a clipboard that had to be particle-safe.



Corn

It retains something that a paper can't. I've been talking all episode about walls and physics and floors on subthreshold swing, and the thing that makes any of it real is a batch of twenty-five wafers where twenty-four were fine and one had a single particle in it and nobody could see which one.



Herman

It's the right note to land on, because it changes the shape of all the numbers. The one-trillion-dollar figure and the sixty millivolt floor are the same object. One of them is a spreadsheet. The other is an electron doing something nobody designed it to do, in a room with the air moving a mile a minute downward, because the room has to keep getting rid of the people standing in it.



Corn

Where does that leave the top of the ladder.



Herman

Two questions I don't have answers to, and I want to be honest about that. The first is the sixty millivolt per decade floor. If that is hard, not just hard for now, then somewhere in the next decade the industry has to buy computing by architecture and packaging rather than by shrinking anything. Chiplets, stacking, specialized silicon. And that means the cost of progress starts rising instead of falling, which is a structural change to the whole economy of computing, not just to a fab.



Corn

The second is the one that bothers me more. Half the revenue of this industry is coming from chips that are less than a fifth of a percent of its shipments. That's not a market, that's a bet. And the rest of the industry, the cheap microcontrollers and analog parts and sensors that everything else runs on, is being funded by a bet it doesn't control.



Herman

Yeah. I don't know how that resolves. But the underlying fact is strange enough on its own. The most important material in modern civilization is one that is a bad conductor and a bad insulator, that does nothing useful until you contaminate it by parts per million, and whose future is now bounded by the fact that electrons are waves and rooms are hot.



Corn

Mediocre at both extremes, and the whole thing rests on it.



Herman

Thanks to Hilbert Flumingtop, our producer.



Corn

This has been My Weird Prompts. If you want to help other people find the show, a review wherever you're listening goes a long way.



Herman

Or email us at show at my weird prompts dot com.



Corn

We'll be back soon.