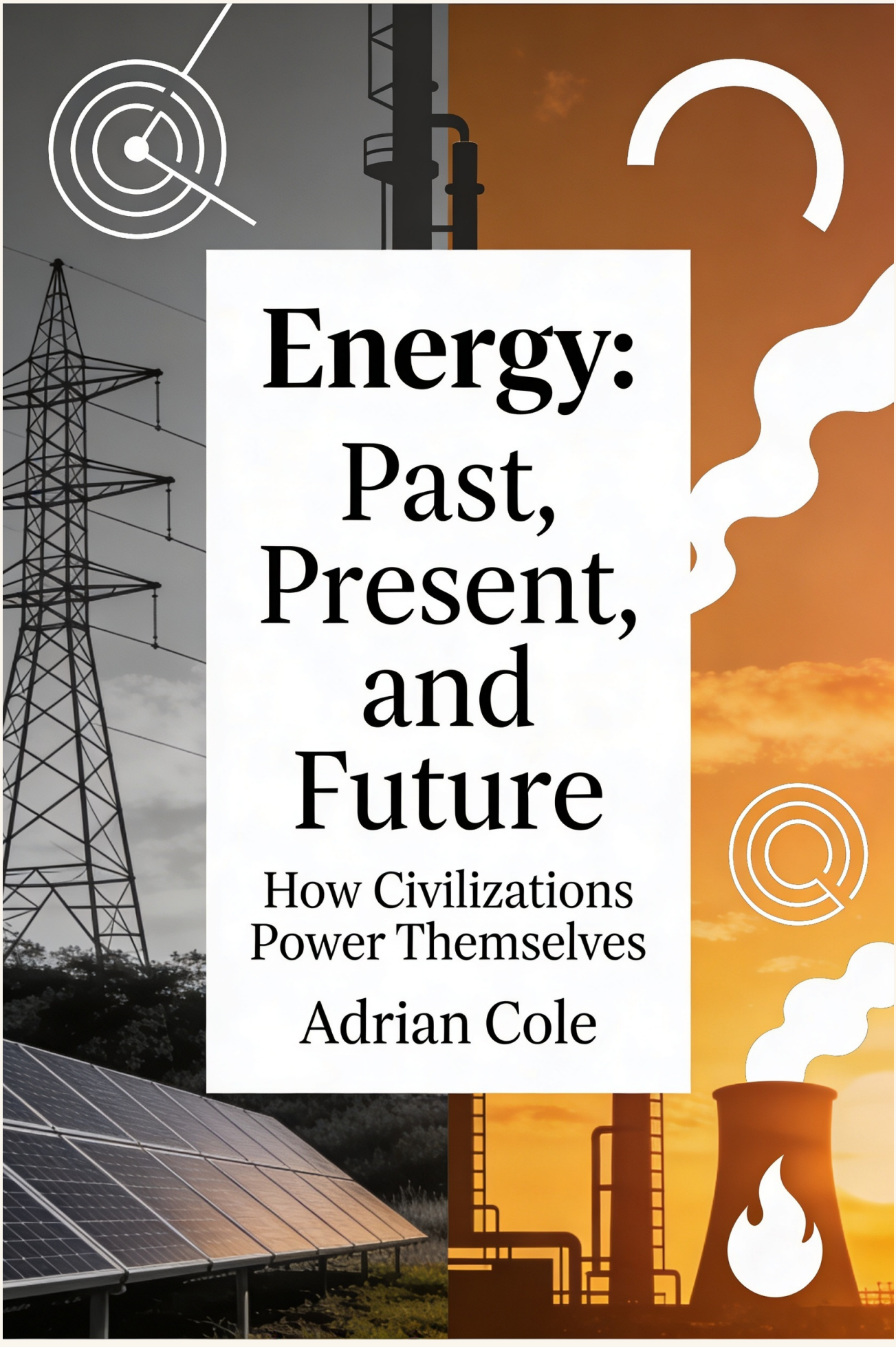


Energy: Past, Present, and Future

How Civilizations Power Themselves

Adrian Cole

Energy: Past, Present, and Future

How Civilizations Power Themselves

by Adrian Cole

CONTENTS

I	What Energy Actually Is
II	Fire and Black Rock: The First Great Transition
III	The Oil Century
IV	Natural Gas: The Bridge to Nowhere, or Somewhere?
V	Nuclear: What the Fear Got Right and What It Got Wrong
VI	Solar: The Cost Curve That Surprised Everyone
VII	Wind: From Windmills to Offshore Giants
VIII	Storage: The Unsexy Problem That Decides Everything
IX	The Grid: Why the Wires Are Harder Than the Generation
X	The Hard Stuff: Steel, Cement, Aviation, and Shipping
XI	The Honest Geography of Energy Poverty
XII	The Next 30 Years: Where the Math Actually Points

CHAPTER I

What Energy Actually Is

In the spring of 2021, a United States senator stood before a microphone and announced that a proposed offshore wind development off the coast of Massachusetts would generate "enough energy to power 800,000 homes." He meant this as a compliment. The project's developers said the same thing in their press release. Journalists repeated it without a second glance. Nobody mentioned that the figure described peak output on a windy afternoon, not the average energy the turbines would actually deliver over a year, and that the "homes powered" metric is a public-relations construction with no agreed definition in engineering or physics.

Nobody was lying. Nobody was even being especially careless by the standards of contemporary energy discourse. They were just speaking a language in which the distinction between power and energy — between a rate and an amount — has been quietly discarded, because it makes the numbers sound better and because almost no one in the audience knows the difference.

This book is going to use the words correctly. Before anything else, that requires a chapter on what energy actually is — not as a metaphor, not as a policy concept, but as a physical quantity, measurable in specific units, governed by laws that cannot be lobbied or subsidized away. The vocabulary built here will carry through every chapter that follows. The scale established here will stop you from being fooled by the impressive-sounding numbers that saturate every energy debate.

The Language Problem

Vaclav Smil, the Czech-Canadian scientist who has spent five decades quantifying how civilization produces and consumes energy, opens his monumental Energy and Civilization (2017) with a blunt observation: most people who write and talk about energy, including many who should know better, do not have a working grasp of basic energy concepts. They confuse energy and power. They use "fuel" and "energy" interchangeably. They describe the economy as "powered by oil" without being able to say what that means in joules. They announce transitions without any sense of how long historical transitions actually take, because they have never looked at the numbers.

Smil is not being elitist. The concepts are genuinely unintuitive. Energy is invisible. You can see a power plant but you cannot see the gigajoules leaving it. You can feel heat but you cannot feel an entropy gradient. The units are abstract — who has an instinctive feel for a terawatt? — and the scale of modern consumption is so vast that the numbers stop meaning anything without deliberate effort to anchor them.

So. From the beginning.

What a Joule Actually Is

Energy is the capacity to do work. That is the textbook definition, and it is correct as far as it goes, but it only moves the question one step back: what is work? In physics, work is done when a force moves an object through a distance. Push a 1-kilogram book across a frictionless table by applying 1 newton of force over 1 meter of distance, and you have done 1 joule of work. The book has acquired 1 joule of kinetic energy.

One joule is a very small amount of energy. A piece of chocolate contains roughly 3 million joules. The human body at rest — lying still, just keeping the heart beating and the neurons firing — burns through about 80 joules per second. A lightning bolt releases approximately 1 billion joules in a fraction of a second, which sounds enormous until you realize it is equivalent to less than 300 kilowatt-hours of electricity, or roughly what a typical American household uses in ten days.

The joule is the SI unit of energy, and it is what physicists use. But energy comes in enough forms — heat, light, motion, chemical bonds, nuclear binding forces — that other units have accumulated over centuries. Calories (food), British Thermal Units (BTUs, beloved by American engineers), barrels of oil equivalent, tonnes of coal equivalent, kilowatt-hours: all of these are just joules wearing different hats.

The kilowatt-hour, which is what your electricity bill uses, equals 3.6 million joules, or 3.6 megajoules. A barrel of crude oil contains roughly 6,100 megajoules. A tonne of hard coal contains around 29,000 megajoules. These conversions matter more than they seem, because any comparison across energy sources requires that everything be expressed in the same currency. When you read that "renewables now provide 13 percent of global energy," you need to know whether that 13 percent is measured in primary energy terms — which includes all the waste heat from thermal power plants — or in final energy terms, which does not. The answers differ by a factor that can make the transition look twice as advanced or half as advanced depending on which accounting convention you choose.

For the scale of civilizations, the useful unit is the exajoule. One exajoule i joules, a 1 followed by 18 zeros. According to the International Energy Agency's 2023 data, global total primary energy supply in 2022 was approximately 620 exajoules. That number will recur throughout this book. Hold it in your mind.

620 exajoules per year. Every year. For 8 billion people.

Power Is Not Energy

Here is where even educated people regularly go wrong.

Power is the rate at which energy is delivered or consumed. Its unit is the watt, named for James Watt, the Scottish engineer who did not invent the steam engine but did make it practical enough to matter. One watt equals one joule per second. Your laptop probably draws between 20 and 60 watts. The incandescent bulb that sat in American living rooms for a century drew 60 watts and converted about 5 percent of that into light; the rest became heat, which is why touching one after it had been on for an hour would burn your fingers. An LED providing the same light today draws about 9 watts.

The distinction between power (watts) and energy (joules, or equivalently watt-hours) is not pedantic. It is the difference between a speed and a distance, between a flow rate and a volume. A fire hose and a garden hose both move water — the question is how fast and how much. A 1,000-megawatt power plant and a 1,000-megawatt wind farm both have the same installed capacity, but they do not produce the same annual energy if one runs 90 percent of the time and the other 30 percent.

That ratio — actual output divided by theoretical maximum output — is called the capacity factor. A nuclear plant in France runs at roughly 75 percent capacity factor. A coal plant in the American Midwest, around 55 percent. A utility-scale solar array in Arizona, perhaps 25 percent. An offshore wind farm in the North Sea, 40 to 50 percent. A rooftop solar panel in cloudy northern Germany, closer to 12 percent.

The senator's 800,000-homes figure almost certainly used peak capacity, not annual energy with capacity factor applied. Apply a realistic offshore capacity factor of 40 percent, and the number of homes powered drops by 60 percent. Apply an honest definition of "homes powered" that accounts for actual average household electricity consumption rather than cherry-picked low-consumption examples, and it drops further still.

None of this makes offshore wind bad. It just makes the rhetoric about it misleading, and the gap between rhetoric and physics is where bad energy policy gets made.

The Laws That Cannot Be Broken

There are two laws of thermodynamics that every energy decision ultimately runs into. They are sometimes stated as jokes: you can't win; you can't break even. That is glib but not wrong.

The first law is conservation of energy: energy cannot be created or destroyed, only converted from one form to another. This sounds reassuring — you never lose anything — until you realize it means you cannot create energy either. You cannot make something from nothing. Every joule that comes out of a power plant went in as coal, gas, uranium, sunlight, or wind. The power plant does not generate energy. It converts it.

Politicians occasionally speak of "generating energy" or "producing energy," and in colloquial terms this is harmless. In physical terms it is wrong in a way that matters. When a solar panel or a wind turbine is called an "energy source," what is really meant is an energy conversion device — one that harvests energy already present in sunlight or moving air and converts it into a more useful form. The sunlight was free; the conversion is not.

The second law is harder. In any spontaneous energy conversion, the total entropy of the system increases. Entropy, loosely, is a measure of disorder — more precisely, of the number of ways energy can be distributed among the particles of a system. High-quality energy (a hot flame, a spinning flywheel, a concentrated beam of sunlight) has low entropy: it is ordered, concentrated, available to do work. Low-quality energy (the diffuse warmth of a room, the faint heat of exhaust gas escaping a smokestack) has high entropy: it is disordered, spread out, difficult to recapture.

Every conversion degrades energy. Some becomes useful output; some becomes waste heat that disperses into the environment and is gone, practically speaking, forever. You cannot run this process in reverse without expending more energy than you recover. The universe tends toward disorder. Physics is indifferent to your plans.

These two laws together mean that every energy system is, at bottom, a system for managing unavoidable loss. The question is never "can we eliminate waste?" — you cannot — but "how much waste is this conversion costing us, and is there a better conversion available?"

Entropy: The Hidden Tax

The efficiency with which a heat engine — a steam turbine, an internal combustion engine, a gas turbine — can convert heat into work is governed by a limit first derived by the French engineer Sadi Carnot in 1824. The Carnot limit depends only on the temperatures involved: the temperature of the hot reservoir (the burning fuel) and the cold reservoir (the environment receiving waste heat).

The formula is simple: maximum efficiency equals 1 minus (cold temperature divided by hot temperature), where temperatures are measured in Kelvin. A modern coal-fired power plant operates at perhaps 600°C, which is 873 Kelvin. The environment is around 25°C, or 298 Kelvin. The Ca ÷ 873), which equals about 65.8 percent. In practice, friction, imperfect insulation, and engineering constraints mean the best real coal plants achieve 40 to 45 percent. The rest — more than half the chemical energy stored in the coal — leaves as waste heat through cooling towers or river water.

This is not a failure of engineering. It is physics. You cannot beat Carnot.

The implications are significant. Of the roughly 620 exajoules of primary energy the world consumed in 2022, only a fraction actually performed useful work — turned motors, lit screens, moved vehicles. The IEA estimates that roughly 65 to 70 percent of global primary energy is lost as waste heat in conversion and transmission. We burn three units of fuel to deliver one unit of useful energy service. The modern energy system is, thermodynamically, enormously wasteful — not because engineers are incompetent but because the second law demands a toll on every transaction.

This has a counterintuitive implication for the energy transition. Electrification of transport and heating, powered by renewables, does not merely swap one fuel for another. It changes the thermodynamic pathway. An electric vehicle uses roughly 85 to 90 percent of the energy in its battery to move the car. A gasoline vehicle uses roughly 20 to 30 percent of the fuel's chemical energy for the same purpose. The rest is heat — Carnot's cut. This is why switching to electric vehicles reduces energy consumption even if the electricity is generated from natural gas, not just from renewables. You are taking a more thermodynamically efficient path.

Entropy is the hidden tax. Understanding it changes how you read every efficiency claim you will encounter.

Power Density: Why Geography Is Destiny

Smil introduced the concept of power density in his 2015 book of the same name, and it is one of the most clarifying ideas in energy analysis. Power density is simply the rate of energy production or consumption per unit of land area, measured in watts per square meter.

It sounds technical. It is actually decisive.

A conventional natural gas power plant produces electricity at a power density of roughly 1,000 watts per square meter of land it occupies. A coal mine with its associated power plant: somewhere between 100 and 1,000 watts per square meter, depending on the richness of the seam and the size of the plant. An oil field in its productive prime can reach 10,000 watts per square meter or more.

Now compare those to the renewable alternatives. A utility-scale solar farm in the American Southwest — one of the world's best solar locations — produces at roughly 5 to 10 watts per square meter. Onshore wind, when you count the full land area between turbines (which farmers can often still use for grazing), averages around 1 to 2 watts per square meter.

The difference is one to three orders of magnitude.

This is not a reason to reject solar and wind — the sun and wind do not send bills, and the economics of renewables have changed dramatically since Smil wrote his power density book. But it does explain a real constraint. The United States consumed roughly 25,000 terawatt-hours of total primary energy in 2022. To supply that entirely from utility-scale solar at 6 watts per square meter would require an area of land on the order of half a million square kilometers. That is larger than California. It would not fit on rooftops. It requires serious thought about where panels go and who decides.

This is not an insurmountable obstacle. The Sahara Desert alone could, theoretically, supply the world's entire electricity demand many times over. But transmission is expensive, deserts are not always near cities, and "theoretically" is doing a great deal of work in that sentence. Power density is why the transition from dense fossil fuels to diffuse renewables involves more land, more infrastructure, and more logistics than most transition advocates acknowledge — and why the advocates who do acknowledge it tend to be taken more seriously in the long run.

There is also an inverse form of the power density question: energy consumption density. New York City consumes energy at a rate of approximately 300 watts per square meter of its area. That is not a city that can be powered by solar panels on its own rooftops. The math does not work. It requires energy imported from elsewhere — whether carried in pipelines, on barges, or through transmission lines. Dense cities have always been energy parasites, in the neutral sense: they concentrate production and consumption, and they require energy supply chains reaching far beyond their borders. That was true when New York burned coal shipped from Pennsylvania; it will be true when New York runs on wind power from offshore farms 100 miles out.

EROI: What It Costs to Get What You Need

Here is a question that sounds simple and turns out to be surprisingly hard: how much energy does it take to get energy?

You cannot pump oil out of the ground for free. The well has to be drilled, the pump has to run, the crude has to be transported and refined. All of that requires energy. The ratio of energy output to energy input is called the energy return on energy invested, or EROI. It was developed and popularized largely by the ecologist Charles Hall at the State University of New York, and Smil uses it extensively as a measure of what an energy source actually delivers to civilization after its own overhead is paid.

The numbers are striking. The first oil wells drilled in Pennsylvania in the 1860s and early 1900s returned somewhere around 100 joules of energy for every joule invested in extraction. EROI of 100:1. You put in a little, you got back a lot. The surplus was enormous. It funded everything — the railroads, the refineries, the infrastructure, and eventually the consumer economy built on cheap fuel.

By the early 21st century, the EROI of conventional oil globally had fallen to something like 15:1 to 20:1. Still very good — a 20:1 EROI means only 5 percent of your output goes to paying the extraction bill — but declining. Tight oil from American shale formations, which drove the production surge of the 2010s, has an EROI estimated by various studies at between 5:1 and 10:1, depending on the field, the methodology, and how you account for the energy costs of water disposal and flaring. Murphy and Hall published an influential 2010 paper in Annals of the New York Academy of Sciences arguing that a society needs a minimum EROI of roughly 3:1 to sustain civilization at all — everything below that ratio, and you are using more energy to get energy than the energy is worth. The exact threshold is contested, but the principle is sound.

Coal, still the world's largest single source of electricity generation, has an EROI estimated at around 30:1 to 50:1 for the better reserves.

Wind turbines, depending on the location and the study, come in at roughly 20:1 to 40:1 over a turbine's operating lifetime — competitive with or superior to most fossil fuels.

Solar photovoltaics are more variable. Manufacturing a panel requires energy — energy to smelt silicon, to run fabrication equipment, to ship the finished product. Early studies in the 1990s found EROI values for solar panels as low as 2:1 to 4:1, which would make solar nearly pointless from an EROI standpoint. More recent analyses, accounting for the dramatic improvements in panel efficiency and manufacturing efficiency since then, typically find EROI values between 10:1 and 30:1 for well-sited utility-scale installations, with the range depending heavily on where the panels are installed and how long they operate. A panel in Arizona harvesting 2,000 kilowatt-hours per kilowatt of capacity per year repays its energy debt much faster than the same panel installed in northern Scotland harvesting 900.

The EROI framing is useful but not the only thing that matters. A low-EROI source is not automatically worthless: if it is the only source available, or if it delivers energy in a particularly useful form, you might accept a lower return. What EROI cannot tell you is whether an energy source is economically viable — that

depends on prices, subsidies, and market structure in ways that pure physics does not capture. And EROI calculations are notoriously sensitive to boundary assumptions: do you count the energy cost of building the road to the wind farm? The energy embedded in the steel in the tower? The energy used to train the engineers who designed it?

Still. As a first filter for taking energy sources seriously, EROI is indispensable. Any energy source with an EROI below about 5:1 is providing very thin margins. Any source below 3:1 is probably doing civilization more harm than good.

The Scale of the Thing

Return to that number: 620 exajoules per year.

Most people cannot feel it. Let me try a few angles.

The average human body, metabolizing food and maintaining life, consumes roughly 80 to 100 watts of power continuously — the heat output of an incandescent lightbulb. Over a year, that is about 2.5 to 3 gigajoules: the energy stored in roughly 60 to 70 kilograms of fat, or about 70 kilograms of coal.

For most of human history, that biological metabolism was essentially the only energy at a person's disposal. A farmer in medieval England commanded perhaps twice their metabolic rate — their own muscles, an ox, some firewood. Call it 200 to 300 watts total, sustained. Their lives were shaped entirely by this limit. Heavy work was possible only in bodies; heat was possible only with firewood; darkness fell at sunset and mostly stayed.

Today, the average person on Earth commands approximately 2,300 watts of continuous power — counting everything: the energy to manufacture their food, to heat and cool their home, to move their goods, to run the factories that made their phone. That is twenty times what a medieval farmer had. Twenty times the leverage on the physical world.

And the average American commands not 2,300 but approximately 9,500 watts — nearly one hundred times the medieval baseline. Smil, in How the World Really Works (2022), puts it this way: in terms of energy services, a typical American is attended by the equivalent of roughly 100 "energy slaves" working continuously on their behalf. This is not a figure meant to produce guilt. It is a figure meant to produce awe at the strangeness of our situation and sober attention to what it would take to change it.

That 620 exajoules is split, roughly, into three kinds of uses. About 40 percent goes to electricity generation. About 35 percent goes directly to transport, primarily in liquid fuels. The remaining 25 percent heats homes and industrial processes. These proportions vary significantly by country — Norway runs its electric grid almost entirely on hydropower while relying heavily on gas for heat; India generates most of its electricity from coal while hundreds of millions of its citizens still cook on biomass. Global averages obscure geography that matters enormously for policy.

Of that 620 exajoules, the IEA's 2023 figures show that fossil fuels — oil, coal, and natural gas — supplied roughly 80 percent. After decades of investment, policy support, and relentlessly falling costs, low-carbon sources combined — nuclear, hydro, wind, solar, and other renewables — still supplied only about 20 percent. This does not mean the energy transition is failing. In 1990, renewables beyond hydro were essentially rounding errors. The speed of solar and wind deployment since 2010 is genuinely extraordinary. But it does mean that anyone who believes the transition is nearly complete, or that fossil fuels are already in their death spiral, is not looking carefully at the numbers.

Wind and solar together produced approximately 33 exajoules of electricity in 2022, according to IEA data. That is roughly 5 percent of global primary energy supply. It was 3 percent in 2019 and nearly zero in 2005. The growth rate is real and sustained. But 5 percent, in a 620-exajoule world, is not a civilization running on clean energy. It is a civilization that has installed the first floor of a building it still needs to build twelve more storeys on.

The Vocabulary Going Forward

The rest of this book will use these concepts without stopping to redefine them. Here is the working vocabulary:

Energy is measured in joules, kilowatt-hours, and at civilizational scale, exajoules. It is a quantity — an amount. It cannot be created or destroyed.

Power is measured in watts and is the rate of energy flow. A 1,000-megawatt power plant has capacity; whether it delivers depends on capacity factor.

Efficiency is the fraction of input energy that becomes useful output. It is always less than 1. Often far less.

Entropy ensures that every conversion extracts a toll in disordered, low-grade heat. The second law is not a problem to be engineered away; it is a property of the universe.

Power density — watts per square meter — describes how concentrated an energy source is. Fossil fuels are dense. Sunlight and wind are diffuse. Both kinds of sources can power civilization, but they require very different amounts of land and infrastructure.

EROI is the ratio of energy out to energy in. Below roughly 5:1, you are working hard for thin returns. Above 20:1, civilization has genuine surplus energy to invest in things beyond energy extraction itself.

Exajoules are the currency of civilizational-scale accounting. Global consumption is approximately 620 per year and rising.

These are not optional definitions that you can swap out for more intuitive ones when the math gets inconvenient. They are the only definitions that correspond to physical reality. Every energy debate you will encounter — every claim about solar reaching grid parity, every argument about nuclear baseload, every projection about electric vehicles reducing oil demand — can be evaluated more clearly through this vocabulary than without it.

Smil's great contribution is not any single finding but a habit of mind: insistence that energy questions be answered in physical units before they are answered in political ones. The joule does not care about ideology. The watt is not impressed by enthusiasm. The entropy gradient will not be moved by a climate pledge, a subsidy, or a well-funded lobby. The laws are what they are, and the scale is what it is.

This does not mean the energy transition is impossible, or even very difficult compared to what civilization has managed before. It means the transition has to be sized correctly in the imagination before it can be built correctly in the world. The chapters that follow will move through the history and technology of specific energy sources — coal, oil, gas, nuclear, solar, wind — and the stubborn sectors that defy easy electrification. At every stage, the questions will come back to the same foundation: how many joules, delivered at what rate, from how much land, at what thermodynamic cost, with what EROI?

Those questions are not the whole story. Economics, politics, human behavior, and institutional inertia all shape energy systems in ways that physics alone cannot predict. But they are the floor. Everything else is built on top of them, and if the floor is not right, nothing else will hold.

One final number to carry with you before the history begins.

In 2022, global energy-related carbon dioxide emissions reached approximately 36.8 billion tonnes, the highest ever recorded. That is the output of a civilization consuming 620 exajoules per year with an 80-percent fossil fuel dependence. It is the number that has made the energy system everyone has agreed is dangerous, and it is the number that justifies the transformation everyone has agreed is necessary.

How that transformation might unfold — and how long it is actually likely to take, given the physics and the history — is the question this book is trying to answer honestly.

Not with optimism. Not with despair. With numbers.

CHAPTER II

Fire and Black Rock: The First Great Transition

Picture a charcoal burner in 14th-century England. He lives in the forest — literally, in a rough hut near his meiler, the circular mound of stacked wood and turf that he tends for weeks at a time, controlling the burn so timber becomes charcoal without combusting entirely. His product is the premium fuel of the medieval world: lighter than wood, nearly smokeless, burning hot enough to smelt iron, forge tools, and fire ceramics. The whole enterprise is skilled, time-consuming, and deeply local. He cannot make more charcoal than the surrounding woodland can grow.

That last sentence is the key sentence in two thousand years of energy history.

The World That Wood Built

For almost all of recorded civilization, wood and its derivatives — charcoal, dried dung, crop residues — provided virtually every joule of energy that human societies ran on. Not most of the energy. All of it. The warmth in a Roman hypocaust came from burning wood. The bronze in a Greek helmet was smelted with charcoal. Medieval breweries, bakeries, glassmakers, salt evaporators, and brickworks consumed wood at industrial rates, long before anyone used the word "industry." Vaclav Smil, whose accounting of energy through human history remains the most thorough ever assembled, estimates that biomass provided somewhere above 95% of global primary energy supply as late as 1800 — a number that should stop you for a moment, given that 1800 is not ancient history. Napoleon was alive. The United States had just elected its third president. Steam engines existed. And the world still ran almost entirely on trees.

This isn't a story about primitiveness. It's a story about limits. The crucial limit was biological: wood is a product of photosynthesis, which captures solar energy at roughly 1% efficiency across a growing season. A hectare of productive European oak forest yields something like 3 to 5 tonnes of dry wood per year on a sustainable cutting cycle. That wood contains about 15 megajoules of energy per kilogram — call it 60 to 75 gigajoules per hectare per year. A modern coal mine extracting bituminous coal delivers the energy equivalent of hundreds of those hectares from a single shaft. The energy density difference is structural, not incidental. Civilizations built on wood hit a ceiling set by biology and geography. Coal breaks through it.

But slowly. Very slowly. And not everywhere at once.

When England Ran Out of Trees

The deforestation of England unfolded across centuries, not decades, and by the time anyone declared it a crisis it had already been ongoing for a long time. By 1600, England retained perhaps 10% of its land area in woodland — down from roughly 75% at the end of the last Ice Age. The iron industry was partly responsible.

The blast furnaces of the Weald, in Sussex and Kent, had been consuming charcoal since the 14th century to smelt iron for cannons, anchors, and agricultural tools. Each tonne of iron required somewhere between 5 and 8 tonnes of charcoal. Each tonne of charcoal required the careful controlled burning of roughly 5 tonnes of wood. The arithmetic was not sustainable, and sixteenth-century commentators knew it.

John Evelyn, the diarist and polymath, published Sylva, or A Discourse of Forest Trees in 1664, written at the request of the Royal Society in response to the Navy's growing alarm about timber shortages. His argument was not aesthetic — it was strategic. England's shipbuilding consumed vast quantities of oak, and Evelyn calculated that the country was consuming its forests faster than they regenerated. He called for a national tree-planting campaign. He was, in a real sense, diagnosing an energy supply problem before the vocabulary to describe it properly existed. His book went through four editions in his lifetime. Trees got planted. The underlying problem did not go away.

London's household fuel had been shifting toward coal since the late sixteenth century, not for ideological reasons but for the simplest possible one: there was no local wood left. Newcastle's coastal collieries began shipping coal south by sea — the famous "sea coal" trade — and London housewives learned, reluctantly, to cook and heat with a fuel that smelled wrong and burned differently from what they'd always used. This is worth noting: the first great energy transition in the modern world began not with an inventor's breakthrough but with a supply shortage. Technology didn't drive the transition. Scarcity did.

The Geology That Changed Everything

Coal is compressed ancient sunlight. This is not a metaphor. The coal seams of Yorkshire, South Wales, Lancashire, and Northumberland are the carbonized remains of vast tropical forests that grew in what is now Britain some 300 million years ago, during the Carboniferous period. Those forests captured solar energy through photosynthesis, died, and were buried under sediment before they could fully decompose — trapping their chemical energy underground rather than releasing it as CO, of decay. thermodynamic sense, a kind of time travel: Carboniferous sunlight releasing itself in a modern furnace.

Britain's particular geological luck was extraordinary. The country sits atop some of the most accessible coalfields in the world, with seams that in many places reached near the surface and extended under navigable rivers, making early extraction and transport feasible without heavy infrastructure. The Tyne and Wear in northeast England, the Severn, the Mersey — all ran through or near coal country. This was not coincidence. It was the geographical precondition for the industrial revolution happening in Britain rather than somewhere else.

The energy density advantage of coal over wood is real but not overwhelming in isolation. Dry wood runs about 15 megajoules per kilogram; bituminous coal runs 24 to 35 megajoules per kilogram — roughly double at the low end, more than double at the high end. That might not sound revolutionary. But consider what it meant for transport: a ship that previously needed to fill its hold with wood to heat a furnace could carry half the weight for the same heat output. More practically, it meant that coal dug out of Northumberland could be

shipped to London and still be economically competitive with local wood in a city that had exhausted its regional forests. The energy density made the trade viable. The geography made it necessary.

Charcoal, the processed form of wood, actually burns hotter than raw coal — but making charcoal requires consuming large amounts of wood for modest output, with energy lost at every step. Coal's real industrial advantage was simpler than that: you dig it out of the ground and it's already a high-quality fuel. No conversion loss. No growing season. No hectares of managed woodland standing between you and your heat.

Iron, Coke, and the Breakthrough That Unlocked Scale

The critical barrier to a full coal-based industrial economy was metallurgical, not domestic. Households could burn coal for heat. But iron smelting — the foundational industrial process — was tied to charcoal, because raw coal contains sulfur and other impurities that contaminate molten iron and make it brittle. The iron industry, the most energy-hungry industry of the pre-industrial world, was trapped in biomass.

Abraham Darby I solved this problem in 1709 at his ironworks in Coalbrookdale, Shropshire. His breakthrough was coke: coal that had been heated in a low-oxygen environment to drive off its volatile impurities, leaving behind nearly pure carbon — exactly as charcoal is wood cooked down to nearly pure carbon. Coke burned hot enough and cleanly enough to smelt pig iron. It worked. Darby kept the process proprietary for decades, and diffusion was slow enough that most British iron production was still charcoal-based in the 1750s. But the door was open.

Once coke smelting became widespread, feedback loops compounded quickly. Cheap iron meant cheaper machines. Cheaper machines meant bigger furnaces and deeper mines. Deeper mines flooded — a problem that demanded its own solution.

Enter the steam engine. Thomas Newcomen's atmospheric engine, first deployed commercially in 1712, was not a general-purpose machine. It was a pump. It pumped water out of flooded coal mines in the northeast of England. Its thermal efficiency was terrible — Smil estimates it converted roughly 1% of the coal's energy into useful mechanical work — but it ran on coal, and coal was essentially free at the pithead. It burned mountains of coal to pump water out of coal mines to produce more coal. This circularity was, perversely, exactly what allowed it to succeed. No other application could have tolerated that level of waste.

James Watt's improvements — the separate condenser patented in 1769, the rotary motion mechanism of 1781 — raised efficiency to roughly 3-4%. Small in absolute terms. Transformative in practice. A Watt engine could power a cotton mill, a forge, a printing press. It began converting the chemical energy of coal into mechanical work at a scale and reliability that no animal or waterwheel could match. By 1800, there were perhaps 500 Watt engines operating across Britain. By 1830, the number was in the thousands, and the railways — themselves powered by improved high-pressure steam — were beginning to stitch the country together.

The Pace of Transition: A Lesson in Humility

Here is where it is easy to get the history wrong. Looking backward, the shift from wood to coal in Britain can look like a rapid, decisive event — a sudden industrial ignition. It was not.

Smil has been making this point with increasing insistence across his career, and it is worth sitting with the numbers. Coal was commercially mined in Britain by the late thirteenth century. Sea-coal was being shipped to London by the sixteenth century. Coke smelting dates to 1709. Watt's rotary engine was operational by the early 1780s. And yet coal did not clearly dominate British primary energy supply — exceeding 50% — until sometime in the early nineteenth century, and it did not reach 90% until after 1870. The transition from "coal is significant" to "coal is dominant" took well over a century in the country most favorably positioned for it.

Globally, the picture is slower still. In 1800, coal provided somewhere around 1-2% of world primary energy supply. Biomass provided the overwhelming remainder. By 1900, coal had risen to roughly 47% of global primary energy. Biomass had fallen to around 50%. The entire century of what we call the Industrial Revolution — which transformed Britain and then Europe and then North America — still left biomass as the world's single largest energy source right up to 1900. Think about what that means: the steam engine existed for nearly two centuries before its underlying fuel actually displaced firewood at the global level.

This is Smil's most important empirical contribution to energy transition debates, and it gets ignored constantly. "Energy transitions are protracted affairs," he writes in Energy and Civilization, and the record shows none has taken less than several decades, and most have taken considerably longer. He plots primary energy shares across the nineteenth and twentieth centuries and the curves are unmistakably gradual — no hockey-stick moments, no overnight flips. Wood didn't vanish when coal arrived. Coal didn't vanish when oil arrived. Each new source layered on top of what came before, as consumption grew fast enough to absorb the old sources alongside the new. The world didn't so much switch from wood to coal as it grew its energy consumption to a level where wood could no longer keep up, and coal filled the gap — while wood kept burning.

What Coal Made Possible

It is hard to overstate how transformatively different a coal-based economy was from a biomass-based one, once the transition reached critical mass. The most dramatic measure is sheer scale. England's coal output was roughly 2.5 million tonnes per year in 1700. By 1850 it was around 50 million tonnes. By 1900 it exceeded 225 million tonnes. At approximately 28 megajoules per kilogram for typical British coal, that 1900 output represents something like 6,300 petajoules of primary energy — from a single island nation of roughly 37 million people. To supply equivalent energy from wood at sustainable forest management rates would require something on the order of 10 billion hectares of productive forest. England has about 13 million hectares of land. Total.

The numbers make the impossibility concrete. No biomass economy could have powered Victorian Britain. The coal economy did not just make the industrial revolution faster or cheaper. It made it possible.

What that energy enabled is the subject of entire libraries: the railways, the steel mills, the cotton looms, the steamships that reorganized global trade, the chemical industries, the gaslit cities. But a few specifics are worth dwelling on because they connect directly to the transition's human consequences. Coal powered the pumps that drained the fens and the mines. It heated the forges that made agricultural equipment cheap enough for ordinary farmers to own. It fired the pottery kilns that made ceramic plates and cups available to households that had previously eaten from wood. These unglamorous improvements in material life were the ground-level product of the energy transition. They were spread unevenly — factory owners and mine owners got rich first, workers in coal towns suffered in specific and documented ways — but in aggregate, per-capita energy consumption rose, and rising energy consumption has tracked rising human welfare with uncomfortable reliability across virtually every civilization in history.

Smil puts a precise number on this worth quoting: in 1800, the average English person commanded, through all energy sources including food, something like 40 to 50 gigajoules of energy per year. By 1900, driven almost entirely by coal, that figure had risen to perhaps 100 to 120 gigajoules. A doubling of usable energy per person in a single century was unprecedented in the entire span of human history before it.

Coal's Costs, Then and Now

The coal transition created enormous wealth. It also created suffering that should not be footnoted away.

Underground mining in the 18th and 19th centuries killed at rates that are difficult to comprehend from a modern vantage point. In Britain alone, estimates of coal mining fatalities across the nineteenth century run to tens of thousands. The Hartley Colliery disaster of 1862 killed 204 men and boys in a single shaft collapse. The Senghenydd colliery explosion in Wales in 1913 killed 439 workers in one morning. Women and children worked underground until the Mines Act of 1842 prohibited it. The Parliamentary reports documenting conditions in British mines in the 1840s make harrowing reading: children as young as five working as trappers, sitting alone in darkness to open and close ventilation doors, twelve hours at a stretch.

Above ground, the air itself became a historical actor. London's famous "pea-souper" fogs were not natural weather phenomena — they were coal smoke chemically interacting with fog, a recurring public health disaster. The catastrophic Great Smog of December 1952 killed an estimated 4,000 to 12,000 people in five days and directly precipitated the Clean Air Act of 1956. The mechanism was the same one that kills people in low-income countries today who cook on open fires: particulate matter and sulfur dioxide attacking respiratory systems already worn down by cold and damp.

The comparison to modern biomass cooking is not rhetorical. The World Health Organization estimates that indoor air pollution from solid fuels — primarily wood, charcoal, dung, and crop residues — kills approximately 3.2 million people per year globally, a figure that dwarfs almost every other energy-related cause of death. Victorian London was not so different in its exposure profile from a modern household in rural sub-Saharan Africa or South Asia cooking over an open hearth. The fuel is different. The pathology is the same.

The Transition That Never Finished

This is the part of the story that gets left out of most energy histories, which are written from the perspective of countries that completed the coal transition and then moved beyond it.

The IEA's most recent World Energy Outlook reports that roughly 2.3 billion people still rely on solid biomass as their primary cooking fuel. That is not a rounding error. It is more than the combined population of Europe and North America, depending daily on the oldest energy technology in human history — the same technology that Homo sapiens mastered at least 400,000 years ago. Hannah Ritchie and the Our World in Data team have mapped this geographically with uncomfortable precision: the biomass cooking belt runs through sub-Saharan Africa, South Asia, and parts of Southeast Asia. In countries like the Democratic Republic of Congo, Ethiopia, Tanzania, and Uganda, traditional biomass accounts for 70 to 90% of total energy supply. Not total cooking energy. Total energy — everything included.

In global average terms, traditional biomass still accounts for roughly 9 to 10% of primary energy supply worldwide, according to IEA data. This number would be higher if you counted all biomass consistently, including modern industrial uses like wood pellets. The point is not a precise accounting but a structural one: two and a half centuries after James Watt built his first rotary engine, roughly a third of humanity's energy still comes from burning plant matter in ways that have not fundamentally changed since the Middle Ages.

The first great energy transition, in other words, is still in progress.

Why? The proximate reasons are economic: coal, natural gas, and grid electricity require infrastructure — mines, pipelines, power lines, distribution networks — that costs money to build and that governments and markets have not seen sufficient return in deploying to the world's poorest communities. The deeper reasons involve the structure of global development and the fact that energy access is a function of income, not just technology availability. When wood is free and electricity costs money, households without money burn wood. When that calculus changes, behavior changes. It changed in Britain because coal was not just available but cheap relative to alternatives — not because someone convinced households to adopt a better technology on principle.

This point matters for how we think about modern transitions. The debate about decarbonization often treats technology availability as the binding constraint. In many low-income contexts, it is not. The solar panels and efficient cookstoves exist. The financing and policy frameworks to deploy them at scale are what's missing. We'll return to this in later chapters. For now, the point is simply that the first transition is an unfinished story, not a completed one.

What the Black Rock Transition Tells Us

Energy historians write about "energy transitions" in the singular — as though one source replaces another in a clean handoff. The actual record is more complicated and more instructive.

Coal did not replace wood. It supplemented wood in rich countries while billions of people continued using wood, and then it grew to dominate industrial economies while wood remained primary in agricultural ones, and then oil and gas began supplementing coal in wealthy countries while coal continued growing in developing ones. Today all four sources — wood, coal, oil, natural gas — coexist in the global energy mix. This is a pattern Smil calls "energy addition" rather than "energy transition," and the distinction matters enormously for how we think about where things are heading.

The specific lag between "transition begins" and "transition completes" is the quantitative baseline this book returns to repeatedly. In Britain — again, the most favorable case — the coal transition took from roughly 1600, when London's household fuel began shifting toward sea-coal, to 1870, when coal dominated 90% of primary energy supply. That is 270 years for a complete transition in the country with the best coal geology, the most developed financial markets, and the most powerful industrial incentives in history. Globally, coal went from 1% to 47% of primary energy over the full nineteenth century. The oil transition from 1% to roughly 40% of global primary energy took from approximately 1890 to 1970 — eighty years, which is the fastest major energy transition in history.

These are not arcane historical data points. They are the empirical anchor for evaluating every claim about how fast a solar-and-wind system can replace fossil fuels. When you hear that the global energy system can be fully decarbonized by 2040, you are hearing something with no historical precedent. When you hear that significant fossil fuel use will persist well past 2060, you are hearing something with extensive historical precedent. Neither observation settles the argument — the circumstances of the current transition are genuinely different in some important ways. But they set the prior, and the prior should be taken seriously.

Fire That Moves

There is one more lesson in the coal transition that tends to get lost in the economic and industrial narrative. It has to do with geography.

Before coal, energy was essentially stationary. You could move charcoal on a cart or a barge, but you couldn't move a forest. The productive woodland that powered an iron furnace had to be within reasonable distance of the ore deposit and the furnace itself — which is why the iron industry was scattered across wooded uplands and river valleys, concentrated in the Weald, the Forest of Dean, the Welsh borders. The energy supply dictated the industrial geography.

Coal untethered this relationship. A coalfield in County Durham could heat a factory in Manchester. A mine in South Wales could power a smelter in Newport and load iron onto ships bound for Brazil. Energy became transportable at scale for the first time in history — a shift so fundamental that it restructured not just industrial geography but the entire logic of where cities could exist and how large they could grow. London grew from roughly 600,000 people in 1700 to 6.7 million by 1900, a growth rate that was thermodynamically impossible in a purely wood-based economy. Cities are concentrations of energy consumption. Coal made concentration at that scale possible.

This mobility of energy also had military and geopolitical implications that shaped the nineteenth and early twentieth centuries in ways still visible today. The Royal Navy's transition from sail to coal-fired steamships in the latter half of the nineteenth century required a global network of coaling stations — strategic depots at Gibraltar, Malta, Aden, Singapore, the Falkland Islands — that were as vital as any fortress. Alfred Thayer Mahan's influential 1890 book The Influence of Sea Power upon History is, among other things, a meditation on the logistics of coal supply for naval power. Control the coaling stations, control the sea lanes. Energy geography was imperial geography.

You can draw a direct conceptual line from this to the oil geopolitics of the twentieth century — to the strategic importance of the Persian Gulf, to the pipeline routes that shaped the Soviet periphery, to Chinese investments in port infrastructure today. The fuel changes. The logic persists. Whoever controls the concentrated, transportable energy source tends to project power. Wood, too diffuse and heavy to move efficiently at scale, never allowed this concentration. Coal did. Oil did it even more completely — which is the subject of the next chapter.

The Inheritance

The coal transition bequeathed the modern world three things that are inseparable from each other.

First: scale. The energy available to human civilization expanded by orders of magnitude, enabling population growth, urbanization, and material welfare improvements no prior civilization had achieved. The average resident of a wealthy country today commands roughly 200 to 300 gigajoules of primary energy per year — five to seven times what an English person commanded at the height of the first coal age. This expansion traces almost entirely back to the transitions coal enabled.

Second: carbon. Every tonne of coal burned released carbon dioxide that had been sequestered for 300 million years. The cumulative effect of 250 years of coal burning — followed by oil and gas layered on top — is the approximately 420 parts per million of atmospheric CO, that existed roughly 280 parts per million in the pre-industrial atmosphere. The warming consequences of this are not a future problem. They are a current one, already measurable in every weather record on the planet.

Third: a template. The coal transition established a structural pattern that every subsequent energy transition has followed: a new source appears, is cheap and abundant in specific locations, attracts vigorous investment, builds out infrastructure that creates economic lock-in, spreads over decades to a fraction of its ultimate reach, and gets declared to have arrived far earlier than it actually has. The pattern holds for oil in the twentieth century. It holds for solar in the twenty-first. The timescales and dominant technologies change. The dynamics — the lag, the unevenness, the incompleteness — persist with stubborn consistency.

That persistence is not cause for despair. The coal transition, for all its real costs, is what made a world of 8 billion people possible. The next transition — whatever combination of solar, wind, nuclear, storage, and efficiency it ultimately involves — is what makes the continuation of that world viable. But making it go well requires understanding how these things actually unfold: at what pace, in what order, for whom.

The charcoal burner in his Wealden hut could not have imagined the collieries of Northumberland. The coal miner in 1850 could not have imagined a solar panel. From inside any energy era, the next one looks like science fiction.

Then, slowly — over decades, not years — it becomes infrastructure.

Then billions of people depend on it.

Then someone starts replacing it.

CHAPTER III

The Oil Century

On the morning of January 10, 1901, a drilling crew working a salt dome on a low hill south of Beaumont, Texas heard a sound nobody had heard before. The drill pipe began shooting up out of the ground. Then mud. Then gas. Then, after a few seconds of eerie quiet, a column of crude oil blasted through the derrick roof and kept going — rising sixty metres into the Gulf Coast sky.

The Lucas gusher, as it came to be known, blew for nine days before the crew figured out how to cap it. In those nine days, it produced approximately 800,000 barrels of crude oil, more than the total output of every other oil well in the United States combined. Spindletop, as the field was called, had just told the world something new: petroleum was not a curiosity for lubricating axles and making lamp oil. It was, latent in the rock, practically unlimited.

That was a lie, of course. Nothing underground is unlimited. But it took the better part of a century to find out exactly how complicated the truth was — and in the interval, humanity built an entire civilisation on the assumption that the oil would keep coming.

Why the Liquid Won

To understand how oil came to dominate the world's energy system, you have to start not with history but with chemistry. Energy analysts spend a lot of time arguing about policy and economics, and those arguments are real, but they rest on a physical foundation that doesn't bend to argument.

Crude oil contains roughly 42 to 45 megajoules of energy per kilogram. That number needs a comparison to land. Dry wood is around 15 megajoules per kilogram. Good bituminous coal sits between 25 and 30. A lithium-ion battery pack in a 2024 electric vehicle stores somewhere between 0.5 and 0.9 megajoules per kilogram — call it 100 times less than the gasoline it replaced. Oil isn't merely better than the alternatives that existed in 1901. It is, on any weight-to-energy ratio that matters for moving things through space, absurdly better than almost anything we have engineered since.

Vaclav Smil, whose career has been largely devoted to quantifying exactly this kind of comparison, puts it plainly in How the World Really Works: modern civilisation was built on fossil fuels not because of greed or conspiracy but because of physics. Energy density is real. The convenience of a liquid that can be pumped, stored in any container, refined into a dozen different products, and transported through pipes across continents is real. No amount of wishful thinking changes the megajoule arithmetic.

But density alone doesn't explain oil's victory. Coal is also very energy-dense. What coal lacks is liquidity — in the literal sense. You cannot pump coal through a pipe. You cannot fill a tank with it and fly across the Atlantic. You cannot, without enormous industrial processing, turn it into the light fractions that make

gasoline or the heavier ones that make lubricating oil or the organic building blocks that become plastics. Oil does all of these things in a single package, and that promiscuity is precisely what made it indispensable.

Before oil, the world ran on wood, coal, and muscle. It was a world with a ceiling. You could build railways and telegraph lines on coal, but your agriculture still depended on draught animals that ate grain that came from fields that required more animal labour to work. The system was intricate and impressive. It was also, in Smil's framing, approaching a hard physical limit. England by 1800 had stripped most of its accessible forests. Urban horses in New York and London produced tens of thousands of tonnes of manure annually, a sanitation crisis with no obvious solution. Energy for moving things was abundant enough; energy for feeding people efficiently was not.

From Curiosity to Substrate

The transition was not Spindletop alone. Edwin Drake had drilled the first commercial oil well in Titusville, Pennsylvania in 1859, at 21 metres depth, producing about 25 barrels a day. John D. Rockefeller built Standard Oil into a monopoly on the back of Pennsylvania crude through the 1870s and 1880s, refining kerosene primarily for lamps. When Edison's electric light began eating into the kerosene lamp market in the 1880s, the oil industry looked briefly vulnerable.

Then the internal combustion engine arrived. And then Henry Ford's assembly line arrived. And then Ford produced the Model T in 1908 at a price that a factory worker could eventually afford — $260 by 1924, down from $850 in 1909. Within two decades, the United States had more registered automobiles than any other country had ever dreamed of owning. The demand for gasoline — a refinery byproduct that had previously been burned off as waste — exploded. Spindletop's heirs in East Texas, California, and the Gulf of Mexico were ready.

By 1920, the United States was producing about 440 million barrels of crude oil per year and consuming the majority of it domestically. By 1950, global production had crossed 4 billion barrels annually. By 1973, when OPEC's Arab members imposed their famous embargo in response to American support for Israel in the Yom Kippur War, the world was consuming around 20 billion barrels per year and had structured its agriculture, its cities, its trade routes, and its foreign policy around the assumption that this flow was permanent.

That agriculture point deserves a full stop. We tend to think of oil as a transport fuel, and transport does account for roughly 60 percent of global oil demand. But the way oil penetrated food production is, if anything, even more profound and far less visible. Modern mechanised farming runs on diesel. Irrigation pumps run on diesel. The trucks and ships that move food from fields to cities run on diesel. And the synthetic nitrogen fertiliser that feeds roughly half the humans alive today is manufactured via the Haber-Bosch process, which runs on natural gas that is chemically almost identical to the methane that travels in oil's shadow. Smil estimates that without synthetic nitrogen fertiliser, the world's soils could support perhaps three to four billion people at current diets. We are at eight billion. The arithmetic of who eats is inseparable from the arithmetic of hydrocarbons.

Plastic, too. About four to five percent of extracted oil never burns at all — it becomes the feedstock for polymers that package our food, insulate our wires, make our medical equipment disposable and therefore safe, and constitute most of the interior of every car, airplane, and building constructed in the last fifty years. The popular conversation about oil imagines a world that replaces gasoline with electricity and is therefore free of petroleum. This is incomplete. Even in a fully electrified transport system, the world would still need oil for chemical feedstocks. The engine disappears; the molecule remains.

The Geopolitics of the Black Substrate

Oil's density and liquidity also made it geopolitically explosive in a way coal never quite managed. Coal is bulky, heavy, and the countries that have a lot of it — the United States, China, Australia, Russia — tend to be large continental powers that can supply themselves. Oil is different. The biggest deposits have no particular obligation to coincide with the biggest economies. The accident of geology placed the majority of conventional reserves under Saudi Arabia, Iraq, Iran, Kuwait, the UAE, Russia, and Venezuela — a map that largely does not overlap with the United States, Europe, or Japan, the three great consuming economies of the twentieth century.

The result was a century of strategic anxiety. Winston Churchill's 1911 decision to convert the Royal Navy from coal to oil — for the speed and range advantages — locked Britain into dependency on Middle Eastern fields it did not control. Franklin Roosevelt's 1945 meeting with King Abdulaziz Ibn Saud aboard the USS Quincy in the Suez Canal, negotiating oil for security guarantees, established an arrangement whose echoes reach to every American foreign policy crisis in the region since. The 1973 embargo cut US GDP growth, triggered gasoline lines stretching around city blocks, and forced the Nixon administration into emergency measures. The 1979 Iranian Revolution and the Iran-Iraq War that followed it caused a second price shock. The 1990 Iraqi invasion of Kuwait triggered the Gulf War. The 2003 invasion of Iraq was certainly about several things, but nobody with a serious grasp of strategic history denies that oil was among them.

OPEC was founded in 1960 in Baghdad, initially by Iran, Iraq, Kuwait, Saudi Arabia, and Venezuela, precisely to counter the pricing power of the "Seven Sisters" — the seven Western oil majors who had controlled Middle Eastern concessions since the 1920s. The cartel's power has waxed and waned over the decades, but the fundamental tension it represents — between states that sit on the resource and states that burn it — has never resolved.

This is not incidental to oil's story. It is central to it. No other energy source has so thoroughly fused physics and politics. The particular combination of high energy density, geographic concentration, and industrial indispensability made oil not just a commodity but a strategic substance in a way that wind and sunlight, distributed across every latitude, can never be.

Hubbert and the Idea That Wouldn't Quite Die

Marion King Hubbert was a Shell Oil geologist — not a marginal figure, not a activist, but a working petroleum scientist who in 1956 presented a paper to the American Petroleum Institute in San Antonio predicting something his employers were unlikely to enjoy hearing. US oil production, he argued, would peak sometime between 1965 and 1970 and then decline in a roughly symmetrical bell curve, following the same mathematical logic that governs the extraction of any finite resource.

His colleagues were sceptical. The industry was booming. New fields were being discovered constantly. The idea that American production had a ceiling seemed, to most geologists of the era, like pessimism dressed up as mathematics.

US production peaked in 1970 at approximately 9.6 million barrels per day.

Hubbert's vindication became one of the most cited results in energy forecasting history. He had been right about the United States, and his method seemed applicable everywhere. Applied globally, the curve suggested a world production peak somewhere around the year 2000. By the late 1990s, with oil prices having collapsed to below $12 a barrel in 1998, a small but vocal community of geologists and analysts — organised loosely around Colin Campbell's Association for the Study of Peak Oil and Gas, founded in 1998 — began arguing that the global peak was imminent.

They were not cranks. Campbell and his colleagues, including the French geologist Jean Laherrère, published careful analyses of reserve data and came to conclusions that were internally consistent and grounded in real geology. The 1998 Scientific American article by Campbell and Laherrère, "The End of Cheap Oil," predicted that global production would peak before 2010. The investment banker Matthew Simmons wrote Twilight in the Desert in 2005, arguing that Saudi Arabian fields were far more depleted than the kingdom's official reserves claimed and that the world was about to confront a geologically imposed supply crisis.

By 2008, with oil prices spiking to $147 a barrel and the global economy convulsing, the peak oil narrative felt like prophecy fulfilled in real time. Mainstream newspapers ran anxious features. The International Energy Agency, historically optimistic about supply, began quietly revising its long-term projections downward.

And then something unexpected happened. Or rather, something that had been quietly happening for thirty years finally arrived at industrial scale.

The Shale Revolution and the Error in Hubbert's Model

George Mitchell was a Texas oilman who had spent much of the 1990s doing something his industry peers considered a waste of money: trying to extract natural gas from shale rock in the Barnett formation near Fort Worth. Shale was known to contain hydrocarbons. The problem was that shale's permeability — its willingness to let fluids flow through it — was, in the technical language, essentially zero. The gas was there. It had nowhere to go.

Mitchell's engineers spent roughly a decade experimenting with hydraulic fracturing — injecting high-pressure fluid into the rock to crack it — combined with horizontal drilling that allowed a single wellbore to snake laterally through thousands of metres of productive formation. By the early 2000s, the combination worked. By the late 2000s, natural gas from shale was flowing at commercial rates across Texas and Pennsylvania and North Dakota. And the same technique, applied to oil-bearing formations like the Bakken in North Dakota and the Permian Basin's tight oil formations in Texas, unlocked petroleum reserves that Hubbert's model — built on conventional geology — simply could not see.

US crude oil production, which Hubbert correctly predicted peaked in 1970 at 9.6 million barrels per day and then declined, hit a new record high of approximately 12.9 million barrels per day in November 2019. The United States became, briefly, the world's largest oil producer — ahead of both Saudi Arabia and Russia. The peak oil movement, which had seemed vindicated in 2008, looked suddenly and badly wrong.

What happened? Two things, really.

First, Hubbert's model correctly described the behaviour of conventional oil — crude that exists in permeable reservoirs under pressure and flows relatively freely to the surface. It was not designed to describe, and cannot describe, oil in tight rock that requires active hydraulic stimulation to extract. The world has vast quantities of unconventional hydrocarbons: shale oil, oil sands in Canada and Venezuela, deepwater deposits. Hubbert's curve was not wrong about the physics of conventional fields. It was incomplete about the total resource base.

Second, and more subtly, the peak oil analysts consistently underestimated the capacity of high prices to drive technological innovation. When oil hit $100 a barrel, it became economic to drill horizontal wells in North Dakota at a cost that would have been absurd at $30 a barrel. The resource boundary is not a geological fact alone. It is a geological fact filtered through economic and technological conditions that change. Higher prices fund research. Research changes what's recoverable. The recoverable amount therefore depends partly on the price trajectory, which depends partly on what governments do, which depends on politics — and Hubbert's model had no term for any of that.

This is not a vindication of the idea that oil is inexhaustible. It isn't. But the peak oil movement made a specific and falsifiable prediction — that supply would be constrained before demand — and that prediction was wrong in the timeframe it specified, for reasons that are now well understood.

What Peak Oil Got Right

Strip away the failed timeline, though, and some of the peak oil diagnosis remains uncomfortably accurate.

Conventional oil production did peak. The IEA's data show that conventional crude — not counting tight oil, oil sands, and other unconventional sources — reached its production zenith somewhere around 2005-2008. Every barrel of shale oil produced since then has been compensating for declining output from the giant fields that were discovered in the 1950s, 1960s, and 1970s in the Middle East, the North Sea, and Alaska. Saudi Arabia's Ghawar field, the largest ever found, was discovered in 1948. The world has not found another

Ghawar.

More importantly, peak oil's underlying message — that treating a finite, unequally distributed, combustion-generating liquid as the permanent foundation of civilisation was a civilisational risk — was correct, even if the mechanism of crisis turned out to be different than the geologists predicted. The peak that is now visible in oil's future is not primarily a geological peak. It is a demand-side peak, driven by electric vehicles, efficiency improvements, and eventually by climate policy. The IEA's Net Zero by 2050 scenario published in 2021 projects that oil demand falls from roughly 100 million barrels per day in 2019 to around 24 million by 2050. That is still 24 million barrels a day of a finite substance being burned. But it is not the infinite-growth trajectory that the twentieth century seemed to promise.

Hubbert was a geologist who worried about the wrong side of the ledger. His heirs are economists who point out that even if supply were unlimited, the atmosphere cannot absorb the combustion products of burning all the oil that exists. Peak oil is real. It just turns out to be a climate-imposed peak rather than a geology-imposed one. Whether that makes us feel better is a matter of taste.

The EROI Trap

There is one more number that the peak oil analysts were right about, even if it doesn't fit neatly into popular narratives about energy abundance.

Energy Return on Energy Invested, or EROI, measures how much usable energy you get back for every unit of energy you spend extracting and processing fuel. Early Saudi and Texas crude had EROIs that geologists now estimate as high as 100:1 — you invested one unit of energy and got a hundred back. Modern conventional oil fields average somewhere between 15:1 and 25:1, depending on the field and the accounting methodology. Shale oil is lower: estimates cluster between 5:1 and 10:1, though the range is contested and depends heavily on whether you count the energy embedded in the steel and water used in fracking operations.

This matters more than it sounds. A society that has access to 100:1 energy sources has enormous slack in its economy — most of the energy budget goes to productive use, and only a tiny fraction goes to energy extraction itself. A society that drops to 5:1 is spending twenty percent of its energy just to get energy. That's not catastrophic, but it changes the economics of everything built on top of the energy system, from steel production to food prices. Smil's work returns to this point repeatedly: the net energy available to civilisation is what drives the economy, not the gross extraction number, and net energy from shale oil is substantially lower than net energy from the giant conventional fields it is replacing.

The shale revolution was real and the production numbers were real. But producing twelve million barrels per day in the United States required burning through capital at a rate that the industry did not sustain profitably for most of its existence. Between 2010 and 2020, the shale sector burned through more capital than it generated in free cash flow — propped up by low interest rates and investor enthusiasm for growth stories. The physics were less generous than the investor decks suggested.

The Infrastructure Problem and Why It Matters for What Comes Next

One figure that doesn't appear in most energy debates, but should: roughly $25 trillion worth of oil infrastructure exists in the world today. Pipelines, refineries, tankers, service stations, drilling equipment, port facilities. This is not a sunk cost that can be written off and forgotten. It is an active system employing millions of people and embedded in the regulatory and financial architecture of dozens of national economies. When Saudi Aramco went public in 2019, it raised $25.6 billion in the world's largest IPO — and was valued at roughly $1.7 trillion, representing the expected future cash flows from extracting and selling oil over the next several decades.

These are not abstract investments. They are commitments. They create constituencies — workers, shareholders, governments whose fiscal revenues depend on royalties — with powerful incentives to extend the life of the system. Every major oil company's capital allocation process contains projections about demand that run to 2040 and beyond. Not because the companies are stupid, but because the physical and financial infrastructure requires decades to build and requires decades of returns to justify itself. The pipeline built today will be operating in 2050.

This is what Smil calls "energy transitions by addition" rather than substitution. In the real history of energy, new sources have generally been added to existing systems rather than replacing them cleanly. Wood did not disappear when coal arrived; it declined gradually over a century and still provides roughly ten percent of global primary energy today. Coal did not disappear when oil arrived. The transitions economists project as clean inflection points are, in actual practice, long and messy processes in which the legacy system continues generating returns while the new one scales up.

For oil, this creates a specific and uncomfortable situation. The physical case for displacement is clear: burning oil produces carbon dioxide that accumulates in the atmosphere and changes the climate in ways that are well-documented and genuinely dangerous. The economic case for electric vehicles is becoming clear too — battery costs have fallen far enough that the total cost of ownership for an EV now rivals a conventional car in many markets, and will undercut it in most markets within this decade. The IEA reported in 2023 that electric vehicle sales had reached a record 14 million globally and were taking a measurable bite out of gasoline demand growth.

But displacing a hundred million barrels per day of crude oil production is not like swapping a light bulb. It requires replacing not just cars but trucks, ships, planes, farm equipment, industrial processes, and chemical feedstock supply chains — most of which have no commercially ready electric or non-petroleum alternative today. The physics of energy density, which made oil's victory so complete, now make its replacement so difficult. A container ship burning heavy fuel oil carries enough energy in its tanks to traverse the Pacific without refuelling. No battery technology on any current development trajectory does the same. A long-haul jet's fuel is forty-four megajoules per kilogram. The best lithium-ion cells available in 2024 hold under one megajoule per kilogram. That gap does not close gradually. It represents a categorical difference in what the two technologies can do.

The Century in Review

Stand back far enough and the sweep of it is genuinely staggering. In 1900, the world used approximately 150 million barrels of oil per year, mostly as kerosene for lamps and lubricant for machines. By 2019, the world used approximately 100 million barrels per day. That is a two-hundred-fold increase in daily consumption in 120 years. No single substance in history has seen demand grow at that rate over that timeframe.

What did it buy? An enormous amount, honestly. Agricultural yields multiplied three- and fourfold through mechanisation and nitrogen fertiliser. Life expectancy in oil-consuming countries rose from around fifty years in 1900 to over seventy-five by 2000. The ability to move people and goods cheaply and quickly across continents created global supply chains that, whatever their fragilities, have distributed economic growth far more broadly than the colonial resource-extraction economies they replaced. Smil, who is constitutionally unsentimental about fossil fuels, is also honest about this: the world's eight billion people are alive and most of them are fed because of what petroleum and its chemical relatives made possible.

The cost is also real. The carbon dioxide already in the atmosphere from a century of combustion is not going away. The geopolitical damage from a century of Western military interventions protecting oil supplies is not easily repaired. The plastic waste accumulating in ocean gyres and Arctic ice cores will outlast any policy response currently being contemplated.

The twentieth century's energy story, told honestly, is not a simple morality play about corporate greed and suppressed alternatives. The alternatives genuinely weren't ready. Oil won because it deserved to win, by the brutal metrics of density, portability, versatility, and cost. It won fair and square by the only standards available at the time. The problem is that winning created a system so large, so deeply embedded, and so physically superior to its available alternatives in certain key applications that the task of unwinding it — even with genuine political will and rapidly falling renewable costs — turns out to be one of the most complex logistical challenges in human history.

Spindletop blew for nine days in 1901 before anyone could stop it. The system it inaugurated has been harder to cap than anyone expected. Understanding why — really understanding the physics and the infrastructure and the lock-in, not just repeating the abstractions — is the precondition for doing anything useful about it.

The chapters that follow are about the alternatives. They are genuinely good news in many places. But every one of them is written in the shadow of this number: as of 2023, oil and other liquid fuels still supply approximately thirty-one percent of global primary energy, a share that has barely moved in a decade despite record renewable deployment. The black rock's liquid cousin is not surrendering the field. It is conceding ground, slowly, one megajoule at a time.

CHAPTER IV

Natural Gas: The Bridge to Nowhere, or Somewhere?

The word "bridge" is doing enormous work in energy policy. It's a reassuring word — it implies you're going somewhere, that both ends of the structure exist, and that you've thought through how long the crossing will take. When natural gas advocates started calling it a "bridge fuel" in the early 2000s, the image was appealing precisely because it seemed to honor the difficulty of the transition while still pointing toward a destination. Coal was the dirty shore you were leaving. Renewables were the clean shore ahead. Gas was the span in between, temporary, pragmatic, mature enough to build on.

Fifteen years later, the question of whether that bridge ever had a far end is one of the most contested in energy policy. Not because the answer is obviously no. Gas has done genuine, measurable climate good in specific places at specific times. But bridges have a way of becoming permanent, and the infrastructure logic of natural gas — the pipelines, the LNG terminals, the combined-cycle plants, the boilers, the industrial furnaces — has a fifty-year half-life built into the steel and concrete. When you build that kind of infrastructure, you're not building a bridge. You're building a neighborhood.

The Chemistry Is Real

Start with what is simply true: burning natural gas to generate electricity produces roughly half the carbon dioxide of burning coal for the same output. Not a rounding error. Half.

The numbers from the IEA are specific. A modern coal plant emits approxima
kilowatt-hour of electricity generated. A combined-cycle gas turbine — the efficient kind that now dominates new gas capacity — emits around 490 grams per kilowatt-hour. Burning gas for heat, as opposed to power, is even cleaner in relative terms because you're converting fuel to warmth directly rather than passing through a turbine with its inevitable thermal losses.

The reason is chemistry, not politics. Methane, the primary component of natural gas, has a molecular formula of CH_4 — one carbon atom and four hydrogen atoms. When you comb
CO_2, and the hydrogen becomes water. Coal is mostly carbon, with a complex
sulfur compounds, and trace metals. Per unit of heat released, coal delivers more carbon to the atmosphere. This is not a contested claim.

There are other advantages that matter in an air quality context rather than a climate context. Natural gas combustion produces almost no sulfur dioxide — the compound responsible for acid rain — compared to substantial SO_2 emissions from coal. It produces less nitrogen oxide. Almost

million people die each year from air pollution, according to WHO estimates, and a significant fraction of those deaths can be traced to coal combustion, both in power stations and in homes. Replacing coal with gas improves that picture, sometimes dramatically.

Vaclav Smil, in How the World Really Works, is characteristically unsentimental about this. He doesn't romanticize gas, but he is precise about what the fuel has accomplished in measurable terms. The global expansion of natural gas consumption over the past four decades has not been irrational. It filled a real need. In many instances, it replaced something worse. To write off that history as a climate delaying tactic requires ignoring the actual lives improved by cleaner air and cheaper heat.

What Happened in America

The most compelling case for natural gas as a genuine climate tool is not theoretical. It happened, in real time, in a real country, and you can see it in the emissions data.

The United States, in 2005, generated roughly 2.4 billion metric tons of CO, 2020, that figure had fallen to approximately 1.45 billion metric tons. A reduction of nearly 40 percent in fifteen years, in a country where electricity consumption stayed roughly flat. This is one of the largest absolute reductions in power-sector emissions any country has achieved, and the story behind it is complicated but real.

Coal tells most of the story. US coal-fired generation fell from about 2,000 terawatt-hours in 2007 to around 800 terawatt-hours by 2020. Some of that displacement was wind and solar — both grew substantially over the same period. But the largest single cause was natural gas, which surged from about 19 trillion cubic feet of annual production in 2005 to over 34 trillion cubic feet by 2020. That surge was entirely the product of one of the most disruptive industrial developments of the early 21st century: hydraulic fracturing.

The shale revolution is a story that deserves its own book, and several exist. The short version is this. Geologists had known for decades that vast quantities of natural gas were locked in shale rock formations across the US — the Marcellus in Pennsylvania, the Haynesville in Louisiana, the Permian Basin in Texas. The gas was there. The problem was getting it out efficiently. Traditional vertical wells barely scratched the shale. What changed was the combination of horizontal drilling — boring sideways along the rock layer — and hydraulic fracturing, injecting high-pressure water and chemicals to crack the shale and release the gas. Neither technique was new. But making them work together at scale, reliably, profitably, took a specific cluster of engineering advances in the 1990s and 2000s.

When it worked, it worked spectacularly. Gas became so cheap in the US that utilities found it cheaper to retire coal plants early than to continue running them. The economics were not subtle. By 2012, US natural gas prices had fallen to around $2.75 per million BTU, their lowest level in a decade. Coal was competing against gas that was essentially free compared to what it had cost five years earlier. The plants that closed were often old, inefficient, high-emitting. The gas that replaced them was burning in combined-cycle turbines operating at thermal efficiencies of 60 percent or more — roughly twice the efficiency of a mid-20th-century

steam plant.

The emissions drop was real. Verified by independent researchers, confirmed by the EPA's own inventory data, visible in the satellite records. You cannot reasonably look at this episode and claim gas did nothing for the climate. In this specific context — the US power grid between 2008 and 2020 — it almost certainly did more to reduce CO_2 emissions than any other single factor.

The case, however, has a serious flaw built into its foundation.

The Ghost in the Pipeline

Methane is not CO_2. Over a hundred-year period, a molecule of methane is ab greenhouse gas than a molecule of CO_2. Over twenty years — a timeframe tha concerned about near-term warming — methane is roughly 80 times more potent. This is the IPCC's assessment, and it is not seriously contested.

Natural gas is almost entirely methane. When you burn it, you get CO_2, whic quantity you're calculating when you compare gas to coal. The problem is that not all the gas gets burned. Some leaks at the wellhead, some leaks from compressor stations, some leaks from pipelines, some leaks at the end-use point — the stove in your kitchen, the furnace in your basement. That leakage, measured as a percentage of total gas produced, is the variable that can flip the entire climate calculation.

In 2018, a paper in Science by Ramón Alvarez and colleagues at the Environmental Defense Fund used airborne measurements to assess methane leakage from the US oil and gas supply chain. Their finding: the US industry was leaking approximately 2.3 percent of gross gas production as unburned methane. This was roughly 60 percent higher than the EPA's official inventory estimate at the time, which was based on facility-level reporting rather than independent atmospheric measurement.

Two-point-three percent sounds small. It isn't. At that leakage rate, the climate benefit of switching from coal to gas — calculated over a 20-year time horizon — shrinks to near zero. The carbon emissions you avoid by burning gas instead of coal are largely offset by the methane you release unburned into the atmosphere along the way. The bridge, on a 20-year accounting basis, may not actually cross anything.

This finding was not universally accepted. The gas industry disputed the methodology. Some researchers argued that the Alvarez study's airborne measurements captured episodic "super-emitter" events that inflate average leakage rates. The EPA revised its estimates upward somewhat but not to the 2.3 percent level. The honest truth is that leakage rates vary dramatically — by company, by basin, by age of infrastructure, by regulatory environment. The Permian Basin in Texas has significantly higher measured leakage than the Marcellus in Pennsylvania. Old, poorly maintained pipelines leak more than new ones. Countries with aggressive leak detection programs leak less.

What you cannot do, honestly, is assume that any particular deployment of natural gas infrastructure achieves the theoretical half-of-coal emissions rate without doing the actual atmospheric measurement. The default

assumption in most policy calculations — that leakage rates are low enough to preserve the climate advantage — has repeatedly failed verification when researchers go out with instrumented aircraft and check.

Europe offers a harder version of this problem. Much of the gas consumed in Germany, Austria, and Eastern Europe has historically come via pipeline from Russia — a pipeline system that was built in the Soviet era, maintained with varying degrees of rigor, and crossed thousands of kilometers of territory where measurement is incomplete. Methane emissions from Russian gas infrastructure are notoriously uncertain. The IEA's Global Methane Tracker has consistently flagged this as a significant unknown in European emissions accounting. You cannot build a clean-energy bridge on a foundation you cannot inspect.

Locked In

The methane leakage problem is solvable in principle. Leak detection technology has improved dramatically. Regulatory frameworks that require operators to find and fix leaks can reduce leakage rates to levels where gas genuinely holds a climate advantage over coal. Some operators have already achieved this. The problem is one of enforcement, monitoring, and incentive structure — difficult, but not physically impossible.

The infrastructure lock-in problem is different in kind. It is not about measuring something better or enforcing existing rules more rigorously. It is about the basic economic logic of capital investment.

A modern combined-cycle gas power plant, built today, has a typical design life of thirty to forty years. An LNG import terminal, like the ones that now line the Gulf Coast of the United States and the ports of Japan, South Korea, and increasingly Europe, costs several billion dollars to build and needs two to three decades of throughput to justify that investment. A gas distribution network for a city — the pipes under the streets, the meters in buildings, the appliances at the end — represents a capital stock that, once deployed, creates powerful institutional and economic incentives to keep using.

The IEA's 2021 Net Zero by 2050 report was unambiguous on this point: in a scenario consistent with limiting warming to 1.5°C, there should be no new investment in oil, gas, or coal supply beyond projects already approved as of 2021. This was not a political statement. It was arithmetic. If the carbon budget is fixed, and if existing fossil fuel infrastructure already commits enough emissions to exhaust most of that budget, then new infrastructure moves the endpoint further away, not closer.

The people who built gas infrastructure between 2010 and 2023 were not wrong by the standards of the information available to them. Some of them were wrong by standards they chose not to apply. But many were operating within a financial and regulatory environment that treated gas as a durable long-term investment. That environment did not arrive by accident. The gas industry lobbied extensively to preserve it, including in policy settings where the explicit justification was "bridge fuel" rhetoric — a framing that sometimes served as genuine transition planning and sometimes served as a delay tactic so refined it barely looked like one.

Here is the concrete version of the lock-in problem. Between 2000 and 2020, the world added roughly 800 gigawatts of new gas-fired generation capacity, according to IEA data. At average capacity factors, that

infrastructure will emit billions of tonnes of CO, over its operational lifetim
debt on a plant that still works, writing off assets, renegotiating contracts — is expensive. Utilities, especially regulated ones with guaranteed returns on invested capital, have every incentive to resist early retirement. The term for an asset that has lost its economic value before the end of its physical life is a stranded asset, and the global fossil fuel sector's exposure to stranded asset risk runs into the trillions of dollars by most estimates. No one with a rational financial interest in that infrastructure is eager to see it strand.

The Crowding-Out Problem

In markets where investment capital is limited — which is every market, always — cheap natural gas creates a different kind of problem for the clean energy transition. It makes the economics of renewable investment harder.

The mechanism is not complicated. When natural gas is cheap and abundant, electricity from gas plants is cheap. When electricity from gas is cheap, the market price for electricity falls. When the market price for electricity falls, the revenue that a new wind farm or solar installation can expect from selling into that market falls with it. When expected revenue falls, some projects that would have been built don't get built.

This is not a hypothesis. The UK offers a well-documented case. The 1990s "dash for gas" — when British utilities rapidly built combined-cycle gas plants after market liberalization — dramatically lowered electricity prices and crowded out investment in renewables that might otherwise have occurred earlier. The UK's renewable transition accelerated only once separate policy mechanisms — the Renewables Obligation, then the Contracts for Difference scheme — effectively insulated wind and solar revenues from the gas-driven spot market.

The United States provides a more ambiguous version of the same dynamic. The same shale-driven gas abundance that displaced coal also suppressed electricity prices during the 2010s, making the economics of new renewable investment harder in markets without specific support mechanisms. States with renewable portfolio standards or production tax credits built wind and solar regardless. States without them largely didn't. The gap between those outcomes tells you something important about what "market forces" mean in an energy system where one technology is receiving enormous, largely invisible subsidies in the form of unpriced methane emissions and unpriced CO,.

Germany is the case most often cited in this debate, usually poorly. The German Energiewende — the energetic and politically contentious effort to transition German electricity generation away from nuclear and fossil fuels — involved a complicated relationship with gas. Germany's decision to phase out nuclear power, accelerated after Fukushima in 2011, created a gap that gas partially filled during the transition period. Whether this delayed the renewable buildout is still debated. The German solar and wind sectors both grew substantially during the 2010s. But Germany's overall electricity-sector emissions took longer to decline than those of countries — like the UK and the US — that made coal-to-gas switches at scale.

The honest framing of the crowding-out problem is this: cheap gas is not automatically a bridge to anywhere. In some markets, with specific policy frameworks in place, it helped displace coal while renewables scaled up. In other markets, it made the financial case for renewables harder and extended the window during which fossil fuel infrastructure was considered a reasonable long-term investment. The outcome depended almost entirely on the policy environment, not the fuel itself.

The Places Where the Bridge Held

It would be intellectually dishonest to write only about the ways gas disappointed the "bridge fuel" narrative. Some bridges stayed standing long enough for something real to cross.

Bangladesh is not a country that appears much in Western energy debates. It should. Over the past three decades, Bangladesh dramatically expanded access to electricity — from below 20 percent of the population in the early 1990s to over 90 percent by the late 2010s — largely by building gas-fired generation at scale. Domestic natural gas reserves made this feasible in a way that coal imports would not have. The health outcomes from replacing kerosene lamps and biomass cooking with grid electricity and LPG stoves were substantial and measurable. Millions of children no longer doing homework by a smoky flame. Maternal mortality declines in part attributable to better hospital capacity. Air quality in homes that shifted from wood and dung to gas.

None of this is an argument for Bangladesh to maximize gas use indefinitely. But it is an argument against the reflexive position — common in some climate advocacy spaces — that any expansion of gas infrastructure in a developing country is simply a bad choice. It is sometimes a less-bad choice than alternatives that are theoretically cleaner but practically unavailable.

The UK's coal phase-out similarly deserves honest credit. Britain in 1990 generated more than two-thirds of its electricity from coal. By 2020, it was below 2 percent. Gas was the primary replacement through the 1990s and 2000s, and in the 2010s wind began to take over as gas itself declined. The transition worked, roughly in sequence — coal to gas to wind — in a way that broadly resembles what "bridge fuel" theorists imagined. UK power-sector emissions fell by more than 70 percent between 1990 and 2020. That is not a failure.

What the UK case also illustrates, however, is the importance of having a plan for the far end of the bridge before you build it. British gas plants built in the 1990s were already running at reduced utilization by the 2010s, as wind and then solar cut into their operating hours. Some became "peakers" — running only when demand spiked — which keeps them economically marginal but physically present. The long-term role of those plants in a near-zero-carbon grid remains unresolved. They are not a bridge. They are a heritage, not entirely wanted, not yet gone.

What the 2022 Crisis Revealed

If you need a stress test for the "bridge fuel" thesis, 2022 provided one that no one had planned.

Russia's invasion of Ukraine in February 2022 triggered the largest energy price shock in Europe since the 1970s oil embargoes. European natural gas prices — which had sat around €20 per megawatt-hour for most of the previous decade — spiked to over €300 per megawatt-hour by August 2022. Electricity prices followed. Industrial energy costs became existential for sectors like aluminum smelting, fertilizer production, and glass manufacturing, many of which curtailed or shut down operations.

The crisis exposed several things simultaneously. First, the depth of European dependence on Russian gas — roughly 40 percent of EU gas imports before the invasion — was a geopolitical vulnerability that energy security analysts had flagged for years without producing adequate political response. Second, gas infrastructure built to take advantage of cheap Russian supply had no Plan B that could be activated quickly. Third, the scramble to build new LNG import terminals, find alternative suppliers, and fast-track renewable permitting all at once revealed how thoroughly the "bridge fuel" years had delayed the necessary hard work.

The European response was, in some ways, impressive. The EU reduced its Russian gas imports by roughly two-thirds in the first eighteen months after the invasion — partly through LNG from the US, Norway, and Qatar, partly through genuine demand reduction, partly through accelerated renewable deployment. The IEA's tracking showed that EU wind and solar generation hit records in 2022 and 2023 precisely because the price crisis made even expensive renewable capacity economically attractive by comparison.

But the crisis also produced perverse short-term outcomes. Germany reactivated mothballed coal plants to cover the gas shortfall. Countries that had been phasing out coal reversed course, briefly. Austria, Italy, and others signed long-term LNG contracts that will constrain their flexibility for a decade or more. The emergency build-out of floating storage and regasification units — cheaper and faster than permanent LNG terminals, but still costly and still steel in the water — created new infrastructure with its own asset logic.

The lesson is uncomfortable. Dependence on any single energy source, in large quantities, from a concentrated geographic source, creates vulnerability that can overwhelm the best-designed transition plan. Gas proponents who argued in the 2000s and 2010s that cheap Russian gas was safe because markets create interdependence and interdependence creates peace were wrong — historically and geopolitically wrong, in ways that turned out to matter. That is not a generic argument against natural gas. It is a specific argument about what "bridge fuel" planning requires: not just a destination, but genuine redundancy, realistic alternative timelines, and an honest accounting of political risk.

The Methane Pledge and What Followed

In November 2021, at COP26 in Glasgow, over a hundred countries signed the Global Methane Pledge, committing to reduce collective methane emissions 30 percent from 2020 levels by 2030. The pledge was, on its face, encouraging. Methane emissions from fossil fuel operations — particularly the oil and gas sector — represent one of the fastest and most cost-effective levers available to reduce near-term warming, precisely because methane is so potent over 20-year timescales and because much of the leakage is technically fixable with existing technology.

The IEA has estimated that nearly three-quarters of methane emissions from oil and gas operations globally can be reduced with existing technology, and that roughly half of that reduction could be achieved at zero net cost — the captured gas is worth more than the cost of capturing it. This is the kind of finding that should produce rapid action. Sometimes it does. In the United States, the EPA's strengthened methane regulations under the Inflation Reduction Act created real enforcement teeth for the first time. Independent monitoring using satellites — including the work from GHGSat, the Environmental Defense Fund's MethaneSAT satellite launched in 2024, and ESA's Sentinel-5P — has made it significantly harder for large emitters to hide routine releases.

But voluntary pledges and satellite surveillance are not the same as regulation with consequences. Outside the US, the EU, and a handful of other jurisdictions, methane emissions from gas operations remain largely unmeasured, underreported, and effectively unpunished. The Permian Basin is monitored more closely than most oil fields in Kazakhstan, Turkmenistan, or parts of the Middle East. The countries that signed the Global Methane Pledge represent a subset of global gas production. The ones with the worst reported leakage rates are not all among the signatories.

This matters for the bridge fuel calculation in a specific way. If you are a country that imports gas — as most European countries do, as Japan and South Korea do — the emissions associated with your gas consumption depend heavily on the production standards of your supplier. Buying LNG from Qatar is not the same climate proposition as buying it from a US operation with continuous third-party monitoring. The product is indistinguishable. The upstream emissions are not.

The Honest Verdict

So is natural gas a bridge to somewhere, or a bridge to nowhere?

The only defensible answer is: it depends where you're standing.

In the US between 2008 and 2018, gas genuinely bridged a coal-heavy grid toward lower emissions, and the emissions reductions were real, documented, and larger than anything renewable energy achieved over the same period in that country. In Bangladesh and similar contexts, gas bridged the gap between energy poverty and functional modern infrastructure, with health and economic consequences that improved real lives. In the UK, gas was one leg of a multi-decade transition that ended with one of the cleanest power mixes among major economies.

In other contexts, gas extended the fossil fuel era rather than shortening it. In markets without strong renewable policy support, cheap gas crowded out investments that might have arrived earlier. In Germany, the gas-heavy transition period coincided with — though was not solely responsible for — higher per-capita emissions than comparable European economies. In the global LNG trade, infrastructure investments made after 2015 committed trillions of dollars to supply chains that will not earn adequate returns unless gas consumption stays high for decades longer than climate targets allow. That is lock-in, not transition.

The hardest version of this problem is the one you can't solve at the level of individual projects or individual countries. It operates at the level of global markets and long-term expectations. When investors, utilities, and governments treat natural gas as a thirty-year solution rather than a ten-year bridge, the infrastructure they build reflects that timeline. Changing those expectations requires not just different rhetoric from politicians — "bridge fuel" rhetoric has been constant for two decades — but different regulatory frameworks, different financing conditions, and different consequences for holding stranded assets.

Vaclav Smil, writing in How the World Really Works, makes the point with characteristic bluntness that energy transitions are not projects with start and end dates. They are processes that unfold over generations, shaped by capital stock, consumer behavior, political economy, and basic physics. The coal transition that began in 18th-century Britain was still unfinished when the 21st century arrived — billions of people still cook over biomass. The shift from coal to gas that appeared to be accelerating in the 2010s stalled, reversed in places, and turned out to depend on geopolitical conditions that no energy planner had modeled.

None of this means the bridge metaphor was wrong. Bridges are useful. Sometimes you build one because there is no other way across, and you build it knowing it won't last forever, and you accept the trade. The problem is not the metaphor. The problem is the way the metaphor was deployed — as a permanent exemption from urgency rather than a tool for managing a finite transition window.

The climate math is unforgiving in one specific way. Methane is potent over twenty years precisely because that twenty-year window is when near-term warming tipping points are most relevant. If gas infrastructure built between 2010 and 2025 operates at high utilization through 2045, it will not have functioned as a bridge at all. It will have been part of the problem it was supposed to solve. Whether that outcome arrives depends less on the fuel and more on the decisions — policy, financial, and political — that determine when the far end of the bridge gets built and whether anyone walks across it before the funding for the structure runs out.

The ambiguity is real. It is not a cop-out. Deciding that natural gas is either a savior or a villain requires ignoring evidence, and ignoring evidence in energy policy is how you end up with cold houses, blackouts, or a warming planet, depending on which evidence you chose to dismiss. The honest position is that gas has been and remains a tool — one with genuine advantages over what it typically replaces, genuine costs that depend heavily on how it is produced and used, and a shelf life in a carbon-constrained world that is almost certainly shorter than most of its advocates have admitted and longer than most of its critics have conceded.

The bridge exists. Whether it leads somewhere depends on whether anyone bothered to build the other shore.

CHAPTER V

Nuclear: What the Fear Got Right and What It Got Wrong

The Safest Dangerous Thing We've Ever Built

Start with a number. Not a slogan, not an emotion — a number.

Coal kills roughly 24 people for every terawatt-hour of electricity it generates. That figure, drawn from Hannah Ritchie's synthesis of global mortality data at Our World in Data, counts mining accidents, air pollution deaths, and the slow respiratory attrition that doesn't make headlines because it arrives in hospital wards rather than reactor control rooms. Nuclear power, over its entire commercial history, kills approximately 0.03 people per terawatt-hour. That's three deaths per hundred terawatt-hours, across every country that has ever run a civilian reactor, including the Soviet Union in 1986 and Japan in 2011.

Those two numbers sit next to each other like a koan. The energy source the public fears most is, by the metric that actually measures harm to human beings, roughly 800 times safer than the one that heats most of the world's homes and powers most of its factories without attracting much fear at all.

This chapter is about why that gap exists, whether it matters, and what it implies for the future — if nuclear has a future, which is genuinely uncertain in ways that have nothing to do with radiation.

Three Mile Island Was Not What You Think It Was

March 28, 1979. A cooling system malfunction at Unit 2 of the Three Mile Island plant in Pennsylvania led to a partial meltdown of the reactor core. It was the worst accident in the history of American commercial nuclear power. For four days, a jittery news cycle followed conflicting statements from the Nuclear Regulatory Commission and the plant's operator, Metropolitan Edison. Pennsylvania Governor Dick Thornburgh recommended that pregnant women and young children within five miles of the plant evacuate as a precaution. About 140,000 people left voluntarily. President Carter toured the site wearing a yellow hard hat.

No one died. The containment structure held. The radiation released to the surrounding population was later estimated at an average of about 8 millirem per person — roughly equivalent to a chest X-ray. The NRC's own post-accident study, completed in 1990, found no discernible increase in cancer rates in the surrounding population.

What Three Mile Island killed was not people. It was confidence. Jane Fonda's film The China Syndrome, a thriller about a nuclear cover-up, had opened twelve days before the accident. The timing was, from a

public-relations standpoint, catastrophic for the industry. Orders for new reactors, which had already been slowing through the late 1970s as construction costs climbed, stopped almost entirely. No new nuclear plant was ordered and completed in the United States for more than three decades. The plant itself, Unit 2, was permanently shut down. Unit 1 kept running until 2019, when it closed not because of safety concerns but because it could no longer compete economically with cheap natural gas — a distinction the industry found difficult to communicate.

The accident was real. The fear was proportionate to the drama of the event. The physical harm was nearly zero. That gap — between the spectacle of nuclear failure and its actual radiological consequences — is the thread that runs through every chapter of nuclear's political history.

Chernobyl: The Real Number and the Imagined One

April 26, 1986. Reactor Number 4 at the Chernobyl Nuclear Power Plant in Ukraine exploded during a botched safety test. The explosion and subsequent graphite fire — the RBMK reactor design used graphite as a moderator, a design choice Soviet engineers had deliberately not built Western-style containment around — sent a plume of radioactive material across much of Europe. It was, and remains, the worst nuclear accident in history.

Here is what it actually killed.

Twenty-eight emergency workers died of acute radiation syndrome in the weeks following the explosion. Two plant workers died in the explosion itself. Two more died from other causes. Those are the confirmed deaths directly attributable to acute radiation exposure. Then comes the much harder question: how many people eventually died of cancer caused by the radiation released?

The World Health Organization, working with a consortium of international experts, published its most thorough assessment in 2006. The conclusion was that approximately 4,000 people among the most highly exposed groups — plant workers, emergency responders, and residents of the most contaminated areas — might eventually die of radiation-induced cancers. "Might" is the operative word. The study also noted that distinguishing radiation-induced cancer from background cancer rates in a population of hundreds of thousands is genuinely difficult, because cancer is common and radiation's contribution, at the doses most people received, is small. For the general population of Europe, the WHO found the statistical cancer risk from Chernobyl exposure to be "very small."

A different estimate, published by Greenpeace and drawing on a wider application of the linear no-threshold model — which assumes that any radiation dose, no matter how small, carries proportional cancer risk — projected nearly a million eventual deaths. That figure has been widely cited, particularly in anti-nuclear literature. It has not been widely accepted by epidemiologists, including many who oppose nuclear power on other grounds, because the LNT model was designed as a conservative regulatory tool, not an epidemiological prediction. Applying it to populations receiving doses well below background radiation levels produces numbers that epidemiology cannot verify and biology does not straightforwardly support.

Vaclav Smil, who is not an apologist for nuclear power, was blunt about this in How the World Really Works: the number of verifiable Chernobyl deaths is far smaller than the public believes, and coal's annual toll from air pollution in Europe alone vastly exceeds even the most expansive Chernobyl death estimates.

What Chernobyl did cause, with certainty, was thyroid cancer in children who consumed contaminated milk in the immediate aftermath of the accident — iodine-131 concentrated in dairy products before authorities in the Soviet Union, operating under the usual reflexes of institutional secrecy, told anyone to stop drinking it. About 6,000 cases of thyroid cancer developed in children and adolescents who were exposed as a result. The survival rate, with treatment, was approximately 98.5 percent. Fifteen people died of thyroid cancer related to Chernobyl exposure. This is not nothing. It is also not what the public narrative suggested.

None of this diminishes what happened to the people evacuated from the 30-kilometer exclusion zone — about 350,000 of them, who lost their homes, their communities, and in many cases their health through the elevated stress, alcoholism, and despair that followed displacement. Those costs are real and human and not fully captured in cancer mortality statistics. But they are the costs of a political and social catastrophe layered on top of a radiological one, not simply the radiological one itself.

Fukushima: When the Evacuation Was Worse Than the Radiation

March 11, 2011. A magnitude 9.0 earthquake off the coast of Japan triggered a tsunami that overwhelmed the seawall protecting the Fukushima Daiichi Nuclear Power Plant. Three reactors lost cooling power and melted down. Hydrogen explosions blew apart the reactor buildings. About 154,000 people were evacuated from the surrounding area. Radiation was released into the ocean and into the air.

The UN Scientific Committee on the Effects of Atomic Radiation, in its most recent assessment, concluded that no discernible increase in cancer rates is expected in the general population of Japan from Fukushima radiation exposure. One plant worker has been officially recognized as having died of cancer attributable to radiation from the accident. One. The radiation doses received by most evacuees were low enough that epidemiological effects are likely undetectable against background cancer rates.

The evacuation itself is a different story. A 2016 study in the Journal of Radiological Protection estimated that approximately 2,000 people died as a result of the disruption and stress of evacuation — elderly patients moved from hospitals and care facilities, people with chronic illnesses who lost access to medical care, suicides. The evacuation was implemented under enormous uncertainty, in the immediate chaos following a disaster that had already killed nearly 16,000 people from the tsunami itself. It may have been the right call given what was known in those first days. But the irony is substantial: the response to the radiological risk killed more people than the radiation did.

Japan shut down all 50 of its operating nuclear reactors in the months following Fukushima. It replaced the lost generation almost entirely with liquefied natural gas and coal, imported emissions from the electricity sector rose sharply. Electricity bills increased by roughly 25 percent. The reactors sat idle, their fixed costs accumulating, while Japan burned more fossil fuels than it had in decades.

By 2023, it had restarted about twelve reactors, with more in the pipeline, but the political and regulatory obstacles remain formidable.

Germany drew the opposite lesson from the same accident. Angela Merkel's government, which had extended nuclear operating licenses in 2010, reversed course within weeks of Fukushima and committed to shutting down all German reactors by 2022. The last three closed in April 2023. Germany made up the shortfall largely with coal and Russian gas — a choice that looked awkward by 2022 when Russia invaded Ukraine and European gas prices spiked violently. German electricity-sector CO, emi declining, plateaued for years through the transition. The costs of the Energiewende energy transition, including grid upgrades and backup capacity, have been substantial.

James Hansen, the NASA climate scientist who is among the most credible scientific voices on global warming, described the German decision as "a catastrophic mistake" — a phrase he doesn't use lightly. He has argued for years that the anti-nuclear movement bears partial responsibility for extended fossil fuel dependence, a position that remains sharply contested but deserves honest engagement rather than dismissal.

What the Fear Got Right

The radiological record is one thing. The institutional and economic record of nuclear power is something quite different, and it is considerably more damning.

The legitimate case against nuclear power — or at least against nuclear power as practiced in most Western democracies over the past fifty years — has three pillars: cost, waste, and weapons proliferation. Each deserves serious treatment, because each involves risks that are not psychological artifacts. They are real.

The cost problem is catastrophic and getting worse. The history of nuclear plant construction in the United States and Western Europe since the 1970s is largely a history of spectacular budget overruns. The Vogtle plant expansion in Georgia — Units 3 and 4, the first new reactors completed in the United States in more than thirty years — was originally budgeted at around $14 billion when construction began in 2009. The final cost exceeded $35 billion. The first unit came online in 2023, seven years late. The second followed months later. Southern Company, the utility that owns the plant, received federal loan guarantees that shielded ratepayers from the full cost, which is another way of saying taxpayers absorbed the overrun.

In the United Kingdom, Hinkley Point C — a new reactor pair being built in Somerset by EDF with Chinese investment — was contracted in 2016 at a guaranteed electricity price of £92.50 per megawatt-hour (in 2012 prices). By 2024, cost estimates for the project had ballooned to £46 billion, with completion now projected for the early 2030s at the earliest. For context, offshore wind in the UK was clearing contracts in the same period at around £40–50 per megawatt-hour, at much shorter construction timelines. The economics require uncomfortable honesty: in much of the West, new nuclear is not currently competitive on pure cost, and pretending otherwise helps no one.

This wasn't always true. The first generation of nuclear plants, built rapidly in the 1960s and early 1970s before regulatory frameworks tightened, were broadly cost-competitive with coal. France built 56 reactors

between 1975 and 1989 — roughly one every three months — using a standardized design and an industrialized construction process, and wound up with an electricity system generating about 70 percent of its power from nuclear at some of the lowest wholesale prices in Europe. The French model is frequently invoked in pro-nuclear arguments, and correctly, but it was achieved under conditions — state ownership, design standardization, industrial momentum, and a political environment largely insulated from public opposition — that no longer exist in France itself, let alone in the Anglo-American regulatory environment.

The waste problem is real and unresolved. The world's operating reactors have generated about 250,000 metric tons of spent nuclear fuel since the 1950s, according to the World Nuclear Association. Almost all of it sits in interim storage — spent fuel pools and dry casks — at reactor sites around the world. A permanent geological repository for high-level nuclear waste does not exist anywhere on Earth.

The United States has been trying to solve this problem since 1982. The Nuclear Waste Policy Act designated Yucca Mountain in Nevada as the candidate site for the nation's permanent repository. Congress approved it in 2002. The Obama administration, bowing to political pressure from Nevada, killed the project in 2010. It has been neither revived nor replaced. American commercial spent fuel sits in temporary storage today, decades after the reactors that generated it stopped running.

Finland is the exception. The Onkalo repository, carved into granite bedrock on the island of Olkiluoto, is the first licensed permanent nuclear waste facility in the world. It began receiving waste in 2024. If you want to understand what the rest of the world's nuclear programs should have been doing for forty years, Onkalo is the answer. That exactly one country has gotten there by 2026 tells you something about the political difficulty of saying "we will bury this here, in this rock, for a hundred thousand years."

Spent nuclear fuel is not the public health catastrophe it's sometimes portrayed as. The volumes are small — the entire American civilian nuclear program's worth of spent fuel would fit in a space roughly the size of a football field stacked to about thirty feet. It is solid, not gaseous, and can be safely stored in dry casks essentially indefinitely. The danger is real but manageable. The political failure to manage it is real and serious, and it has provided legitimate grounds for public distrust.

The proliferation concern is the hardest to quantify and the most serious. The fuel cycle for civilian nuclear power overlaps, at key points, with the technology required to produce fissile material for nuclear weapons. Uranium enrichment, needed to produce reactor fuel, produces weapons-grade uranium if pushed further. Spent fuel contains plutonium, which can be extracted through reprocessing. The international safeguards system administered by the International Atomic Energy Agency is designed to prevent this diversion, and it works reasonably well in countries with functioning governments and honest regulatory institutions. It works less well in others. North Korea developed its weapons program under the cover of a civilian nuclear infrastructure. Iran's enrichment program has been a central concern of nuclear non-proliferation diplomacy for two decades.

This doesn't mean civilian nuclear power is synonymous with weapons proliferation — the countries with the largest civilian programs (France, the United States, Japan, South Korea) are not particularly worrying on

proliferation grounds. But it does mean that the expansion of nuclear power to politically unstable countries, or countries hostile to international inspection, raises genuine security questions that cannot be waved away with reference to deaths-per-terawatt-hour statistics.

The New Reactor Promises, Graded Honestly

Here is where optimism and physics have to have a frank conversation.

The nuclear industry and its advocates have spent the last fifteen years producing a steady stream of announcements about next-generation reactor designs — smaller, simpler, cheaper, safer, and somehow freed from the cost and time pathologies of traditional large-scale nuclear construction. These technologies deserve examination, not as articles of faith or objects of dismissal, but as engineering proposals at specific stages of development.

Small modular reactors — SMRs, in the trade — are reactors with generating capacity below about 300 megawatts, designed to be manufactured in factories and assembled on-site, potentially on shorter construction timelines and with smaller capital outlays than conventional gigawatt-scale plants. The theory is compelling. Factory construction is cheaper and more precise than bespoke on-site work. Smaller units mean smaller upfront commitments and potentially faster regulatory approval. A fleet of standardized small reactors could supply baseload power to regions without the grid infrastructure for large plants.

NuScale Power, based in Oregon, was for years the leading SMR developer in the United States, and the one closest to actual deployment. In 2022, the NRC certified NuScale's 77-megawatt module design — the first small modular reactor design certified in American history. It was a genuine milestone. Twelve months later, NuScale's only contracted customer, a consortium of Utah municipal utilities, cancelled its planned order. The projected electricity cost had risen from about $58 per megawatt-hour to more than $89, driven by supply chain inflation, higher interest rates, and the general tendency of first-of-a-kind nuclear projects to cost more than estimated. NuScale wrote off hundreds of millions in development costs. The project that was supposed to prove SMRs could work economically instead demonstrated, at least for that design, that they couldn't — yet.

This is one data point, not a verdict. Other SMR designs are in various stages of development in the UK, Canada, China, and Russia. China's HTR-PM, a high-temperature gas-cooled pebble-bed reactor, began commercial operation in late 2023 — making it the first fourth-generation reactor design to reach the power grid anywhere in the world. Russia's RITM-200 modules power the world's largest nuclear-powered icebreaker. These are real machines running real power. But they are not yet the scalable, cost-competitive revolution that SMR advocates project. The question of whether factory manufacturing can genuinely drive down nuclear costs, rather than simply relocating the complexity from the building site to the factory floor, will not be answered until multiple units of the same design have been built and operated.

Molten salt reactors are a different class of ambition. Instead of solid fuel rods cooled by pressurized water, a molten salt reactor dissolves fuel in a liquid fluoride or chloride salt mixture, which circulates through the

reactor and acts as both fuel and coolant. The physics are genuinely attractive: the reaction is self-regulating in ways that light-water reactors are not, the fuel cannot melt down because it is already liquid, and some designs can run on thorium — far more abundant than uranium — or on existing spent nuclear fuel.

The Oak Ridge National Laboratory ran a small experimental molten salt reactor for four years in the 1960s. It worked. It was shut down in 1969 for political and institutional reasons more than technical ones. Today, dozens of companies in the US, Canada, China, and Europe are working on molten salt designs. Terrestrial Energy, Moltex, TerraPower's Natrium design (which uses a sodium coolant rather than salt, technically distinct), and others have attracted serious investment and regulatory engagement. None has yet built a grid-connected power plant. The engineering challenges of operating at very high temperatures with chemically aggressive salts — corrosion, materials science, fuel processing — are real and are not solved by the elegance of the underlying physics. This is a technology that could matter in the 2040s. Probably not the 2030s. It is not, despite the excitement it generates in certain corners of the energy world, a near-term solution.

Fusion requires its own paragraph, and that paragraph should begin with an honest statement of where things stand as of 2026.

The ITER project — a massive experimental tokamak being constructed in southern France by a 35-country consortium — has been under construction since 2013, following two decades of design work. Its budget has grown from an initial estimate of roughly €5 billion to something exceeding €20 billion. It will not generate electricity. It is designed to demonstrate that fusion produces more energy than is needed to sustain the plasma — a critical scientific milestone called Q>1 — and first plasma is now expected sometime in the late 2020s, with full deuterium-tritium experiments in the 2030s. The joke that fusion is always thirty years away has been told so many times it has become a cliché even in engineering schools.

Private fusion companies — Commonwealth Fusion Systems, Helion Energy, TAE Technologies, and others — have attracted billions in investment and made credible technical progress, particularly on high-temperature superconducting magnets that could enable smaller, cheaper tokamaks. Commonwealth Fusion's SPARC experiment, designed to demonstrate Q>1 at compact scale, is under construction in Massachusetts and aims for first plasma in the late 2020s. These are serious people doing serious physics. Some of them may succeed.

But "may succeed at demonstrating scientific feasibility" is not the same as "will provide grid electricity." The engineering distance from a plasma that briefly produces more energy than it consumes to a turbine spinning electrons into the grid involves heat exchangers, tritium breeding, neutron-activated materials, and a business model — none of which is trivial. If someone tells you commercial fusion power is arriving in the 2030s, ask them what turbine they plan to use and who's building it. The answers are usually vague.

None of this means fusion research should stop. The potential payoff — a fuel cycle based on hydrogen isotopes, with no long-lived waste and no proliferation pathway — is large enough that continued investment is clearly justified. It means that fusion should not appear in anyone's decarbonization plan for 2050, because

it almost certainly won't be deployable at scale by then.

What France Figured Out (And Then Started to Forget)

The case for nuclear power as a serious tool for decarbonization doesn't rest on unproven future technologies. It rests on what France actually did between 1975 and 1989.

France generates roughly 70 percent of its electricity from nuclear power — a figure that dropped toward 60 percent in 2022 when an unusually large number of reactors required corrosion repairs simultaneously, exposing France's dependence on a single technology more starkly than its planners would have liked. But even in that difficult year, France remained one of the lowest-carbon electricity systems in Europe. Its per-capita electricity-sector emissions are a fraction of Germany's, despite Germany's massive investment in solar and wind. French households pay less for electricity than most of their European neighbors.

The French program worked because it industrialized. Same reactor design, repeated 56 times, with accumulated construction expertise and supply chain depth. That approach is almost certainly not replicable today in the same form — the industrial base has atrophied, the regulatory environment has changed, and the public and political tolerance for nuclear construction has narrowed. But South Korea built six reactors at its Barakah site in the UAE between 2009 and 2024 using a standardized approach and a trained construction workforce, on time and close to budget. The Koreans demonstrated that the French lesson is not entirely lost. It requires political will, industrial continuity, and a regulatory system designed for predictability rather than perpetual challenge.

Whether Western democracies can recreate those conditions is not primarily a technical question. It is a political one.

The Accounting We've Been Avoiding

Let's put the ledger together.

Nuclear power currently provides about 10 percent of global electricity. It generates roughly 25 percent of the world's low-carbon electricity, a share that has been declining as solar and wind grow faster. About 440 reactors operate worldwide. In the countries where those reactors run, they provide reliable, high-density baseload power that solar and wind, without substantial storage, cannot directly replace.

Closing nuclear plants prematurely — before their licensed operating lives expire — reliably increases fossil fuel consumption in the short to medium term. This is not speculation. It happened in Germany. It happened in Japan. When California closed the San Onofre plant in 2013, grid-level emissions increased. When the Diablo Canyon closure was subsequently debated and then reversed in 2022 — the plant received a state-backed life extension — the central argument was that California could not afford to lose its single largest source of carbon-free generation while trying to meet its climate targets.

Extending the operating lives of existing nuclear plants, which involves relatively modest capital expenditure compared to new construction, is almost certainly the most cost-effective nuclear option available today. The IEA has repeatedly emphasized this point in its clean energy scenarios. An existing reactor with its capital cost amortized over decades produces electricity cheaply. The French, who began a long-running debate about phasing out nuclear and then reversed course under President Macron in 2022, eventually acknowledged this arithmetic.

New nuclear construction in the West is expensive and slow — that much is simply true. But the question of whether it should be part of the energy mix is not the same question as whether it should be the only option. The climate math of the 2030s and 2040s is tight enough that foreclosing any low-carbon technology requires confidence in the alternatives that the storage and grid chapters of this book should have made difficult to summon.

The Conversation We're Not Having

Here is the asymmetry that runs through every public debate about nuclear energy.

When a wind turbine catches fire — which happens — no one proposes banning wind power. When a natural gas pipeline explodes and kills workers — which happens regularly — no one holds congressional hearings about whether gas should be phased out. When a coal ash pond fails and poisons a river — which happened at the Kingston Fossil Plant in Tennessee in 2008, releasing a billion gallons of toxic slurry — the story leaves the news cycle within weeks.

Nuclear accidents attract sustained, intense, transnational political response to an extent that no other energy source does. Some of that response reflects the genuine horror of invisible contamination, the loss of control over something that cannot be seen or smelled, and the association with weapons that shaped the technology's public image from birth. Some of it reflects the specific ways that institutional failure — Soviet secrecy at Chernobyl, corporate and regulatory opacity at Fukushima — eroded trust in ways that made rational risk comparison impossible.

The fear of nuclear, in other words, has both legitimate and illegitimate roots. The illegitimate part is the persistent confusion between the radiological risk, which is measurably small, and the psychological and political disruption, which is measurably large. The legitimate part is the set of concerns — cost overruns, unresolved waste, proliferation — that don't disappear just because coal kills more people per terawatt-hour.

A society that wanted to make rational energy decisions would hold the following positions simultaneously: nuclear power has a better safety record than almost any other energy source and deserves credit for that; the cost of new nuclear construction in the West is currently prohibitive and the industry has failed to solve it; existing plants should be operated as long as they are safe; waste storage requires a permanent solution and Finland has shown one is possible; new reactor designs deserve continued research funding but should not appear in near-term decarbonization commitments; and the weapons proliferation risk is real and must be managed through international institutions rather than by abandoning the technology.

That combination of positions is almost impossible to hold in public discourse, where nuclear tends to be either defended with a fervor that dismisses all concerns or opposed with an intensity that inflates all risks. The people best positioned to hold all those views simultaneously — energy engineers, nuclear physicists, climate scientists — are also the people least likely to be asked about it on television.

The reactor that is statistically your best chance of clean, reliable baseload power sits in this uncomfortable space: safer than the public believes, more expensive than the industry admits, indispensable in some grids, already obsolete in others, and promising a next generation of technology that may or may not arrive in time to matter.

The honest answer to whether nuclear power has a future is: it depends on which country, which grid, which decade, and which alternative you're comparing it to. That answer is unsatisfying. It is also, on the available evidence, correct.

CHAPTER VI

Solar: The Cost Curve That Surprised Everyone

The year is 1976. A single watt of solar-generating capacity costs $106. That is not a typo. One watt — enough to power a small LED lightbulb at full brightness, barely. Adjusted for inflation, a single watt of solar electricity would have set you back nearly $550 in 2024 dollars. The oil crisis had just rattled the Western world. Congressional hearings were full of earnest talk about energy independence. And solar power was technically real but practically absurd: it belonged on satellites and in the catalogs of a few dedicated eccentrics willing to pay spacecraft prices for roof tiles.

In 2022, utility-scale solar module prices fell below $0.20 per watt. In some competitive markets, below $0.15. That is a price decline of more than 99.8 percent over forty-six years. The steepest, most sustained cost collapse of any energy technology in recorded history. And the most consequential fact about that collapse — the one worth sitting with before you read anything else in this chapter — is that almost nobody predicted it. Not the oil companies. Not the utilities. Not most governments. Not the IEA. Not, for the most part, even the solar advocates who wanted it to happen and spent careers arguing it could.

Understanding why it happened anyway, who drove it, and what it means for the limits still ahead — that is what this chapter is for.

The Physics You Need and Nothing More

A photovoltaic cell works because photons from the sun knock electrons loose in a semiconductor material, creating a voltage. That is the whole mechanism. Silicon is the dominant semiconductor used in solar panels — purified, sliced into thin wafers, coated with materials that create an electric field to push the liberated electrons in one direction. Thin-film technologies built on cadmium telluride or copper indium gallium selenide use different materials and different manufacturing processes, but the underlying physics is the same.

The theoretical maximum efficiency for a single-junction silicon cell — the Shockley-Queisser limit, derived from thermodynamic principles — sits around 33 percent. Most commercial panels today achieve 20 to 23 percent. Research cells have pushed past 29 percent in laboratory conditions. That gap between commercial and theoretical performance matters for land use, and we'll get there. But efficiency was never the variable that drove the cost story. Manufacturing economics drove it. Specifically, something called Wright's Law.

The Curve That Ate the Forecasts

Theodore Wright was an American aircraft engineer who published a paper in 1936 in the Journal of the Aeronautical Sciences. He had noticed something remarkable about aircraft production costs: every time the cumulative number of planes built doubled, the labor cost per plane fell by roughly 15 percent. Consistently.

Across different manufacturers, different aircraft types, different years. Not by accident. As a result of accumulated experience — workers who had done the same task thousands of times finding faster approaches, engineers who had seen ten thousand units come off the line spotting inefficiencies invisible to someone watching the first hundred, tooling that got better and cheaper as suppliers scaled their own operations.

Wright's Law says: for every doubling of cumulative production, costs fall by a consistent percentage. That percentage varies by industry — somewhere between 10 and 30 percent is typical. For solar photovoltaics, analysis of the full historical data puts the learning rate at roughly 20 to 23 percent. Every time the global installed base of solar panels doubled, panels got about 20 percent cheaper.

Twenty percent per doubling sounds modest. It is not modest. Doubling is exponential, and solar deployment was doubling fast. Between 2010 and 2020, global cumulative solar installations doubled roughly four or five times. Run the arithmetic: four doublings at 20 percent per doubling compound to a 59 percent cost reduction. Five doublings gets you to 67 percent. The IEA's data confirms this directionally: the average auction price for utility-scale solar fell by approximately 90 percent between 2010 and 2020, from around $0.36 per kilowatt-hour in the first year of that period to roughly $0.037 by its end. In the sunniest, most competitive markets — the Atacama Desert in Chile, the Arabian Peninsula, the Rajasthan plateau in India — it fell further still.

Here is the forecasters' problem: Wright's Law works only if deployment keeps growing. In 2005, or 2008, or even 2012, assuming that solar deployment would keep doubling every two to three years required a kind of optimism that looked, to most analysts, indistinguishable from wishful thinking. Energy agencies are not constitutionally built for optimism about unproven technologies. They are built for careful extrapolation from existing trends. And the existing trend, through most of the 2000s, was that solar was an expensive niche product deployed mainly in Europe by policy mandate, at prices that would never be competitive without indefinite subsidy.

Fatih Birol, who became IEA executive director in 2015, acknowledged the failure publicly in 2021: the agency had systematically underestimated solar's growth trajectory in every World Energy Outlook it had published for fifteen years running. In the 2006 WEO, the IEA projected global solar installed capacity would reach 159 gigawatts by 2030. The world crossed that threshold in 2015 — fifteen years early. By the end of 2023, global solar capacity stood at roughly 1,600 gigawatts. The 2006 forecast was wrong by a factor of ten, a decade ahead of schedule. This was not a rounding error. It was a systematic failure of the forecasting methodology — a failure to take Wright's Law seriously, combined with a failure to anticipate the country that would drive deployment to scales where the learning curve would really accelerate.

The Factory That Changed Everything

In 2001, China produced roughly 1 percent of the world's solar panels. By 2010, it produced about 50 percent. By the early 2020s, it produced more than 80 percent of global solar modules, more than 95 percent of the

silicon wafers those modules are built from, and roughly 85 percent of the polysilicon — ultra-purified silicon — that the wafers require. These are not market share statistics in the ordinary sense. They describe a manufacturing transformation with no real precedent in modern energy history.

Chinese solar manufacturers — Jinko Solar, LONGi Green Energy, Trina Solar, JA Solar — scaled in ways that Western manufacturers could not. Lower labor costs were part of the early advantage, but that explanation runs out quickly. Far more important was the systematic industrial policy support: subsidized land, cheap capital from state-backed banks, direct grants, preferential electricity rates for energy-intensive production, and the compounding efficiency of co-locating the entire supply chain within a few provinces. In Xinjiang, Inner Mongolia, and Sichuan, polysilicon could be refined, drawn into ingots, sliced into wafers, processed into cells, and assembled into modules within a few hundred kilometers of end-to-end vertical integration. Western manufacturers, with supply chains stretched across multiple countries and continents, were playing a different game at a structural disadvantage.

LONGi Green Energy, founded in Xi'an in 2000, bet specifically on monocrystalline silicon — higher efficiency, more difficult to produce, but amenable to aggressive manufacturing improvement. For years, analysts questioned whether monocrystalline's higher upfront cost could be beaten down far enough to compete. LONGi's answer was to scale relentlessly and invest in process engineering with the patience that only state-backed capital can sustain through thin margins. By 2022, LONGi had annual production capacity exceeding 85 gigawatts. That number is worth pausing on. It is more than the entire world installed in any single year before 2015.

The consequences for Western manufacturers were severe. Q-Cells, a German company that had been among the world's leading solar cell producers in the early 2000s, filed for insolvency in 2012. Solyndra, the California thin-film manufacturer that received $535 million in U.S. government loan guarantees, went bankrupt in 2011. Solyndra became a political set piece about wasteful government spending, and the politics were real enough. But the engineering story was different: Solyndra had built a genuinely clever cylindrical thin-film design that made sense at high polysilicon prices — it avoided polysilicon almost entirely. When Chinese manufacturers flooded the market with cheap polysilicon and cheap monocrystalline modules, Solyndra's whole business case collapsed. It didn't fail because of government support. It failed because of Chinese manufacturing success.

This manufacturing concentration creates a strategic problem that is not going away quickly. Every country trying to build a solar-powered electricity grid is currently building it on a foundation — panels, inverters, wafers, polysilicon — that largely runs through a single country. The U.S. Inflation Reduction Act of 2022 introduced substantial incentives to re-shore solar manufacturing: domestic content bonuses, advanced manufacturing tax credits, and a newly hostile tariff environment for Chinese imports. The European Union's Net-Zero Industry Act aimed in the same direction. Whether those incentives will work fast enough and at sufficient scale to meaningfully diversify the supply chain is genuinely uncertain. Climbing a learning curve that Chinese manufacturers spent twenty years and enormous state investment scaling is not simply a matter of throwing money at factories. It takes time. The honest answer, as of the mid-2020s, is: partial progress,

slowly, at higher cost per watt than current Chinese production.

There is a harder dimension still. Xinjiang province produces roughly 45 percent of global polysilicon. The evidence of forced labor in Xinjiang — affecting the Uyghur minority population through a state-directed system of coercive employment in industrial facilities — has led to U.S. import restrictions under the Uyghur Forced Labor Prevention Act, signed in 2022. Module imports into the United States fell sharply following enforcement, as panels with opaque supply chains were held at customs. The energy transition discovered it had walked into a human rights catastrophe it didn't create and can't easily extract itself from. This is not an argument against solar power. It is an argument for taking supply chain transparency seriously rather than treating it as a compliance box to be checked.

Germany's Expensive Gift to the World

Long before China dominated manufacturing, Germany decided to deploy solar power at large scale regardless of cost — a political decision rooted in the country's post-Chernobyl ambivalence about nuclear power and its emerging sense of environmental identity. The 2000 Erneuerbare-Energien-Gesetz (Renewable Energy Sources Act, commonly abbreviated EEG) turned out to be one of the most consequential single pieces of energy legislation of the 21st century, though not exactly in the way its architects expected.

The EEG introduced a feed-in tariff: any German household, farmer, or cooperative that installed a solar panel would receive a guaranteed, above-market price for every kilowatt-hour they fed into the grid, locked in for twenty years. Early rates for small rooftop installations reached about €0.51 per kilowatt-hour — against retail electricity prices of roughly €0.15 to €0.20. The incentive was enormous. German rooftops covered with photovoltaic panels became, within a decade, a visible feature of the landscape in Bavaria, Baden-Württemberg, and Saxony. By 2012, Germany had installed about 30 gigawatts of solar capacity — a staggering number for a country whose solar resource is roughly equivalent to Alaska's.

The program was expensive. German electricity consumers paid an annual surcharge — the EEG-Umlage — that by 2013 added approximately €0.063 per kilowatt-hour to household bills. Germany's industrial competitiveness lobby complained bitterly. The total subsidy transferred to renewable generators between 2000 and 2020 has been estimated at over €200 billion. Germany's experience became the cautionary tale that other governments cited when arguing against premium feed-in tariffs.

What that money actually purchased, viewed from a sufficient distance, was something different from what the critics described. It purchased deployment volume during the critical early years when the solar learning curve was still steep. When Chinese manufacturers were making investment decisions in 2006 and 2007 about whether to build the next generation of production capacity, German demand — stable, predictable, guaranteed for two decades — was a decisive signal. The German market validated that volume could exist. Chinese factories scaled to meet it, and in doing so drove costs down for every subsequent buyer. A 2019 analysis from the Potsdam Institute for Climate Impact Research estimated that German solar policy brought forward the global cost trajectory by somewhere between three and seven years. At 20 percent cost reduction

per doubling of deployment, three years of additional early volume is not a trivial gift.

Germans paid premium prices to bootstrap a global manufacturing industry to the scale where Wright's Law could run properly. The rest of the world got cheaper panels. Whether that was wise German policy or German consumers being exploited by Chinese industry through European political naivety is a matter of genuine debate. The outcome — a functional, cost-competitive global solar industry — is not.

The American Version: Tax Credits and Desert Megaprojects

The United States took a structurally different path. Rather than guaranteeing a price for output, American policy primarily supported installation through the Investment Tax Credit, established in its modern form in 2006. The ITC allowed businesses and homeowners to deduct 30 percent of solar installation costs directly from federal tax liability. It was extended multiple times — often at the last possible moment, creating the boom-bust uncertainty that developers consistently identified as the industry's biggest operational problem — before being stabilized by the Inflation Reduction Act of 2022, which locked in 30 percent through 2032 and added domestic content bonuses of up to 10 additional percentage points.

The American solar story is also, distinctively, a utility-scale story. The desert Southwest — California's Mojave, Arizona's Sonoran, Nevada's Basin and Range — has solar irradiance values that Germany's climate cannot approach. By the late 2010s, enormous PV arrays were being built on those lands with a speed and at a cost that genuinely surprised the utilities that had once dismissed them. Projects in Nevada and Arizona were winning competitive auctions at prices below existing natural gas contracts. The California grid, which had been built around large thermal generators supplying power in predictable blocks, was absorbing enough solar by 2017 that it began, on spring days, to curtail — simply throw away — surplus midday electricity.

The IEA's 2023 Electricity Market Report recorded that solar was the single largest source of new electricity generating capacity added globally in 2022. More solar was installed that year than gas, coal, nuclear, wind, and hydro combined. That sentence deserves to be read twice.

What "Cheapest" Actually Means — and Doesn't

By the early 2020s, utility-scale solar photovoltaics had become the cheapest source of new electricity generation by the standard industry metric: levelised cost of electricity, or LCOE. This measure takes the total lifetime cost of building and operating a plant — upfront capital, financing, operations, maintenance, decommissioning — and divides it by the total electricity the plant produces over its life, giving you a cost per kilowatt-hour or per megawatt-hour.

Lazard's 2023 LCOE analysis, one of the most widely cited in the industry, placed utility-scale solar at $24 to $96 per megawatt-hour, with the low end representing large projects in excellent solar resource locations with competitive financing. For comparison: new natural gas combined-cycle plants, $39 to $101. New coal, $65 to $166. New nuclear, $141 to $221 — a range carrying vast uncertainty depending on whether you're building in South Korea, France, or the United States. Solar won. The data is clear.

LCOE is useful, and it is incomplete in a way that becomes more important as solar penetration rises. It measures the cost of the generator. It does not measure the cost of integrating the generator into an operating grid. A solar plant with an LCOE of $30 per megawatt-hour in the Atacama Desert might require $15 to $40 of additional system costs per megawatt-hour — transmission lines to demand centers, backup capacity for cloudy days, storage for evening peaks — depending on the grid architecture around it and what share of total generation solar already represents on that grid.

This is the "value deflation" problem. As more solar enters a grid, solar's market value per megawatt-hour tends to decline, because solar floods the market precisely at the hours when all the other solar generators are also producing. An additional gigawatt of solar on a grid that is already 30 percent solar is worth less to the system than an additional gigawatt on a grid that is 5 percent solar. The LCOE doesn't change. The value does. Vaclav Smil, who has spent decades frustrating both energy optimists and energy pessimists with rigorous arithmetic, has made this point about solar repeatedly: the total system cost of running a solar-heavy grid includes a substantial bill for the backup, storage, and transmission that solar's intermittency requires. That bill is real. It is not prohibitive. It simply needs to be in the accounting.

Land, Rooftops, and the Agrivoltaic Middle Ground

Utility-scale solar requires land, and this fact tends to generate more rhetorical heat than careful analysis. A utility-scale installation typically needs somewhere between 5 and 10 acres per megawatt of capacity, depending on panel efficiency, row spacing to minimize self-shading, and terrain. The Bhadla Solar Park in Rajasthan, India — among the world's largest solar parks at about 2,245 megawatts as of 2023 — covers roughly 57 square kilometers. The land is Thar Desert: hot, arid, receiving extraordinary solar irradiance, and supporting minimal agriculture.

Hannah Ritchie's comparative work at Our World in Data puts solar's land demands in useful context. Solar does require more land per unit of delivered energy than nuclear or natural gas. It requires substantially less than bioenergy grown for combustion — the approach that the European Union has, somewhat absurdly, been subsidizing as "renewable" under its biofuels directives. When you include the full footprint of fossil fuel infrastructure — mines, wellpads, processing facilities, exclusion zones, waste disposal — coal in particular imposes landscape costs that ground-level photographs of solar parks make easy to forget.

Rooftop solar sidesteps most of the land argument. A panel on an existing building roof adds no new land requirement at all. The U.S. Department of Energy's technical analysis estimated that rooftop solar installed on all suitable American buildings — accounting for roof orientation, shading, structural capacity, and existing mechanical equipment — could supply roughly 40 percent of current U.S. electricity demand. That estimate comes with the same grid-integration caveats that apply to utility-scale solar, but the land point stands independently.

Agrivoltaics — placing solar panels above crops in configurations that allow farming to continue beneath — represents a genuine design solution rather than a rhetorical one. Early results from research projects in

Oregon, Germany, and the Japanese island of Hokkaido suggest that shade-tolerant crops including lettuce, spinach, and strawberries perform comparably or better under partial solar shading in hot climates, while panels mounted over irrigated fields run cooler and therefore slightly more efficiently. Full-canopy agrivoltaic systems aren't suitable for grain agriculture at scale, but the approach adds real options where they didn't previously exist.

The Materials That Don't Make Headlines

A standard crystalline silicon solar panel is mostly glass, aluminum frame, and encapsulant. These are abundant, cheap materials. But the functional core — the silicon cells and the electrical contacts and the specialized semiconductor materials in inverters and thin-film panels — involves supply constraints that deserve more attention than they typically receive.

Silver is the most immediate. Conventional silicon solar cells use silver paste to form the electrical contacts on the front surface. Today's panels use roughly 15 to 20 milligrams of silver per watt of capacity — down dramatically from around 400 milligrams per watt in 2008, thanks to consistent process engineering. The trajectory is the right direction, but the absolute demand is growing fast: the Silver Institute estimated that solar accounted for approximately 14 percent of global silver demand in 2023, rising as deployment accelerates. Silver is expensive relative to other industrial metals and is produced largely as a byproduct of lead and zinc mining, which means supply responds sluggishly to solar-driven demand signals. Researchers are working on copper-contact cell designs that would eliminate silver dependency almost entirely; progress is real but not yet at commercial scale.

Cadmium telluride thin-film panels, manufactured predominantly by First Solar and now a significant fraction of American utility-scale deployment, use tellurium — one of the rarest stable elements in Earth's crust, produced as a byproduct of copper refining at roughly 500 tonnes per year globally. If CdTe technology scaled to dominate the market, tellurium would become a hard physical ceiling. It doesn't currently dominate, and supply constraints are part of the reason.

Polysilicon sits at the base of about 95 percent of global solar production. Its geography is the supply chain's most visible geopolitical concentration: the four largest polysilicon producers in the world are all Chinese, and together they account for the overwhelming majority of global output. That's not an immovable fact — polysilicon production can be built elsewhere, and investments in Germany, the United States, and South Korea are underway — but building the chemical engineering infrastructure to produce semiconductor-grade polysilicon at competitive cost takes years and billions of dollars.

None of these material constraints are disqualifying. The history of solar technology is almost entirely a history of material challenges being solved by engineering at scale. Silver intensity per watt has fallen by 97 percent since 2008. Polysilicon purity requirements that seemed exotic in the 1980s are now routine industrial chemistry. But the next decade of growth — getting from 1,600 gigawatts today to the 10 or 15 terawatts that serious decarbonization scenarios require — will stress these supply chains in ways that warrant more

planning attention than the industry typically gives them.

The Subsidy Question, Answered Honestly

Solar critics reliably argue that the cost curve was built on subsidies and would collapse without them. This is partly true and almost entirely beside the point — or rather, beside the wrong point.

Solar has received substantial government support: German feed-in tariffs, U.S. investment tax credits, Chinese industrial policy, direct grants, and guaranteed purchase prices across dozens of markets. The IEA's calculation of global energy subsidies — its broadest measure, including the implicit subsidy of not pricing carbon or local air pollution costs — came to approximately $5.9 trillion in 2020 for fossil fuels alone. That number mixes direct cash transfers with estimated externality costs and is not cleanly comparable to solar's explicit policy support, but the idea that the incumbent energy system operates on unsubsidized market logic is fiction. Nothing in energy ever has.

The more useful question is what the subsidies accomplished. Gregory Nemet's 2019 book How Solar Energy Became Cheap, which traces the causal mechanisms behind the cost curve in careful empirical detail, argues that the policy-to-learning-to-cost-reduction pipeline for photovoltaics worked essentially as intended. Feed-in tariffs drove deployment volume. Volume drove manufacturing scale. Scale drove learning-curve cost reductions. And those cost reductions are now self-sustaining at current deployment levels in high-resource markets: new utility-scale solar projects in India, Chile, the UAE, and Mexico are winning competitive auctions with no direct subsidy, against natural gas and coal on an unsubsidized basis.

The United States still extends the ITC. Offshore wind requires heavier support. Solar in cloudy, high-latitude markets may require policy support indefinitely. These are not uniform facts. They depend on location, grid context, and competing generation costs. But the claim that solar is fundamentally subsidy-dependent — that it reverts to $106 per watt the moment governments stop writing checks — is simply wrong, and has been demonstrably wrong since roughly 2019.

Where the Numbers Stand

Global solar capacity reached approximately 1,600 gigawatts at the end of 2023, generating around 1,800 terawatt-hours of electricity — roughly 5.5 percent of global electricity generation, according to IEA figures. Wind contributes another 7 to 8 percent. Together, wind and solar represent the fastest-growing share of the global electricity mix, by a substantial margin.

Five-and-a-half percent of electricity sounds underwhelming after a chapter about the greatest cost collapse in energy history. It should. Electricity itself is only about 20 percent of total final energy consumption — the rest is heat, transport fuel, and industrial process energy that mostly doesn't run on electrons yet. Solar's 5.5 percent of electricity amounts to roughly 1 percent of total global energy. The distance between 1 percent and the 20 or 30 percent share that serious climate scenarios require is enormous. In absolute terms — the steel, silicon, glass, concrete, and transmission infrastructure that must actually be built — Smil's arithmetic is

sobering: we are talking about deployment rates that have no historical precedent in the energy system's physical scale.

In the countries with the highest solar penetration, the engineering challenges are already visible and instructive. California sourced roughly 27 percent of its electricity from solar in 2023. South Australia has touched 100 percent solar generation on its grid during favorable conditions. Spain crossed 17 percent solar in 2023's annual generation mix. These grids are dealing in real time with the problems that high solar penetration creates: curtailment of surplus midday generation, the steep evening demand ramp when rooftop panels go dark, and the declining revenue per megawatt-hour that value deflation produces for new solar projects in already-saturated morning hours. The California "duck curve" — the shape of net electricity demand when you subtract solar output from total load, with its deep midday dip and steep evening peak — has become the canonical diagram of this challenge.

Managing that curve requires storage, flexible backup generation, demand response, or transmission interconnection to geographies whose solar generation peaks at different times. All of those solutions are technically available. None of them are built at the scale required. That is what storage and grid architecture chapters are for.

The Honest Scorecard

Solar photovoltaics have produced one of the most remarkable industrial achievements of the 21st century. The people who predicted the cost trajectory correctly — a small number of researchers who took Wright's Law seriously and applied it to historical solar data, including researchers at Rocky Mountain Institute and Ramez Naam, who wrote about the trajectory when it seemed wildly optimistic — were right when the IEA, the EIA, and most of the energy establishment were not. The technology works. The costs are real and verified by competitive auction prices in dozens of countries. The deployment is accelerating faster than any previous energy technology at comparable scale of ambition.

The limits are equally real and deserve to be stated without embarrassment. Solar generates useful power for perhaps four to six hours per day in good locations, less in cloudy climates and northern winters, and nothing at night. Its supply chain runs through a single dominant manufacturing nation to a degree that creates strategic vulnerability for every other country trying to build a solar-powered grid. Its material demands require continued engineering progress and supply diversification. Its value to a grid decreases as its penetration rises, making the last 20 percent of solar generation much harder to justify economically than the first 20 percent.

None of that makes solar the wrong answer. It makes solar one part of an answer that requires storage, transmission, flexible demand, and continued decarbonisation of the industrial and transport sectors where electrons can't simply substitute for liquid fuels or high-temperature chemical reactions. The cost curve surprised everyone and delivered something genuinely transformative. What comes next — the grid architecture, the storage chemistry, the industrial decarbonisation — will not be solved by the same

mechanism. Learning rates and manufacturing scale are powerful tools. They are not the only tools.

The sun showed up. The factories came through. The next problem is harder, less photogenic, and already waiting.

CHAPTER VII

Wind: From Windmills to Offshore Giants

There is a formula so simple it fits on a coffee mug, yet it explains almost everything about the modern wind industry. Power equals one-half times air density times swept rotor area times wind speed cubed. That cubed term is the one that matters. It means that if wind speed doubles — from five metres per second to ten — the power available from that wind does not double. It increases eightfold. Build a turbine ten metres taller, capturing air moving just a little faster, and the reward is not linear. It is geometric.

That single relationship — power proportional to the cube of wind speed — has driven four decades of engineering ambition. It explains why the spindly 50-kilowatt machines that dotted California's Altamont Pass in 1982 gave way to the 15-megawatt colossi now being installed off the British coast, machines with rotor diameters wider than the wingspan of an Airbus A380 is long. Every extra metre of hub height, every additional metre of blade radius, bends the physics in the engineer's favour. The industry has not simply built bigger turbines because bigger is better as some kind of aesthetic preference. It has built bigger turbines because the cube law makes each incremental gain in height and sweep area disproportionately rewarding.

This is an engineering story. It is also a policy story, a geography story, and in places a fairly undignified political story involving wealthy waterfront homeowners and strategic opposition dressed up as environmental concern. But the engineering comes first, because without understanding why wind turbines behave the way they do, the decisions — good and bad — that surround them make no sense.

California and the Tax-Credit Farms

The modern wind industry was born partly from crisis and partly from bad incentives, a combination that produces interesting history if not always good technology.

The 1973 oil embargo forced the United States government to take renewable energy seriously for roughly a decade, a seriousness that faded almost entirely when oil prices fell in the mid-1980s. The Public Utility Regulatory Policies Act of 1978 required utilities to buy power from independent producers at their "avoided cost," which opened a door for wind developers. California layered its own tax credits on top of federal ones. By the early 1980s, the Altamont Pass east of San Francisco, Tehachapi in the south, and the San Gorgonio Pass near Palm Springs were filling with turbines.

Most of them were terrible. Not in a subtle way. Danish wind engineers who visited Altamont in the early 1980s came away with a mixture of bemusement and professional horror. American developers, chasing tax credits with a deadline rather than kilowatt-hours with a margin, had rushed machines into service that were mechanically unreliable, sized far too small, and often badly sited relative to the local wind resource. Altamont's turbines had a collective capacity of about 500 megawatts by the late 1980s. Their actual annual

output was a fraction of what that capacity number implied, because capacity factor — the ratio of actual energy produced to theoretical maximum — hovered around 15 percent for many of those early machines.

The Danes, meanwhile, were building something more durable. Danish turbine manufacturers — Vestas, Bonus (later Siemens Wind), Nordtank — developed a culture of mechanical conservatism and incremental improvement that looks, in retrospect, like exactly the right approach. Danish turbines in the 1980s were not glamorous. They were three-bladed, upwind, horizontal-axis machines, a configuration that had been largely settled on by 1980 and has not changed since. They prioritised reliability over raw output. They worked. And because they worked, Danish manufacturers accumulated learning curves while their American competitors were going bankrupt when the tax credits expired.

By 1990, Denmark had roughly 300 megawatts of wind capacity. That number seems trivial against a global backdrop today. It was not trivial as a foundation for an export industry that would come to dominate global turbine manufacturing for the next three decades.

The Physics, Precisely

Before going further, it is worth sitting with the cube law for a moment longer, because it has a second implication that the first glance sometimes obscures.

The theoretical maximum fraction of wind's kinetic energy that any turbine can extract is 59.3 percent. This limit — named after German physicist Albert Betz, who derived it in 1919 — is not a technology constraint. It is a physical one, derived from the fact that a turbine extracting all the kinetic energy from the wind would bring that air to a complete stop, at which point no more air could flow through the rotor. Modern utility-scale turbines operate at efficiencies of around 45 to 50 percent, which means they are already capturing the vast majority of what the physics allows. There is not a factor-of-two improvement waiting in blade aerodynamics.

What there is, instead, is height. Wind speed increases with altitude above the ground because of reduced surface friction — meteorologists call this the wind shear profile. At 30 metres, a typical hub height for a Danish turbine in 1985, a site might average 6 metres per second. At 100 metres, the same site might average 7 metres per second. Cube that difference: $6^3 = 216$, $7^3 = 343$. That 17 percent increase in average wind speed translates to a 59 percent increase in available power. No improvement in blade materials, gearbox efficiency, or power electronics gets you anything close to that gain. You go taller.

The blades have grown too, because swept area appears as a multiplier in that same equation. A rotor diameter of 15 metres sweeps about 177 square metres. The Vestas V236-15.0 MW turbine, introduced in 2022, has a rotor diameter of 236 metres and sweeps 43,767 square metres — roughly 247 times more. Area is cheap, relative to the energy it buys.

This arithmetic explains the machines. A 15-megawatt offshore turbine with a 150-metre hub height and a 236-metre rotor diameter is not overengineering for its own sake. It is the logical endpoint of applying the cube law as aggressively as materials science and marine installation logistics permit.

From Kilowatts to Megawatts: The Scale Progression

Track the growth in turbine size and the trajectory is almost laughably consistent. The Altamont machines of 1982 were typically 50 to 100 kilowatts. By 1990, leading commercial turbines were around 300 kilowatts. By 2000, 1 to 2 megawatts was standard. By 2010, onshore turbines of 2.5 to 3 megawatts were routine. By 2020, onshore machines of 4 to 5 megawatts were available, and the offshore frontier had pushed past 10 megawatts. The GE Haliade-X, which began commercial deployment around 2021, is rated at 13 to 14 megawatts. Vestas' V236 and Siemens Gamesa's SG 14-222 DD push to 14 to 15 megawatts. These are not prototype machines. They are in active installation programs off the coasts of Britain, the Netherlands, and Denmark.

Each doubling in turbine size has not simply doubled output. Because larger turbines can reach higher wind speeds, can spread their fixed installation and maintenance costs across more kilowatt-hours, and can take advantage of better blade aerodynamics at scale, the cost per unit of energy has fallen faster than a naive size-scaling model would predict. This is Wright's Law in action — the empirical observation, first quantified by Theodore Wright in 1936 for aircraft manufacturing, that costs fall by a consistent percentage for every doubling of cumulative production. For wind turbines, that learning rate has been roughly 9 percent per doubling. Every time the industry doubles its cumulative installed capacity, costs fall about 9 percent.

The numbers are striking. The levelised cost of electricity from new onshore wind in good locations has fallen from roughly $135 per megawatt-hour in 2010 to under $33 per megawatt-hour by 2021, according to IEA data. Some auctions in Brazil, India, and the United States have cleared even lower, below $25 per megawatt-hour. To put that in context: in 2010, natural gas combined-cycle plants — then considered the cheapest form of new generation — cost around $65 to $70 per megawatt-hour. Onshore wind in the right locations is now cheaper than almost anything else you can build.

Capacity factors have improved in parallel. An average onshore wind farm in the United States in the early 2000s achieved a capacity factor around 25 percent. By 2020, newly installed projects were averaging around 35 percent, with some sites in the Great Plains exceeding 45 percent. The improvement reflects better siting, taller towers capturing stronger winds, longer blades sweeping more area, and more sophisticated control systems that optimise rotor orientation and blade pitch in real time.

The Land Question

At this point, the sceptic in the room raises an obvious objection. Wind turbines are large. A modern onshore turbine stands perhaps 150 to 200 metres from base to blade tip. You need a lot of them to run a modern grid. What do they actually do to the land?

This is where Hannah Ritchie's comparative analysis, published through Our World in Data, is clarifying. The key distinction is between the land a wind farm occupies in total — the spacing between turbines, the access roads, the substation — and the land a wind farm actually uses. Turbines are typically spaced five to ten rotor

diameters apart to avoid wake interference, which means a wind farm with 100 megawatts of capacity might span several square kilometres on a map. But the turbines themselves, their foundations, and their access roads occupy perhaps 1 to 3 percent of that footprint. The rest can be — and overwhelmingly is — farmland, pasture, or natural habitat continuing to function exactly as it did before the turbines went up.

Ritchie's comparison makes this concrete. A coal power plant and its associated mining operations use roughly 18 times more land per unit of electricity than wind, when you account for the full fuel supply chain. Nuclear power is extremely land-efficient. Solar is more land-intensive per unit of power than wind, though still far less than coal. The narrative that wind farms "industrialise the countryside" in a way that fundamentally differs from conventional energy infrastructure does not survive contact with the data.

This does not mean land use is a non-issue. It means the question needs to be asked carefully. A solar farm on prime agricultural land is a genuine trade-off. A wind farm on the same land usually isn't, because the turbines occupy so little of it. Iowa, which generates more than half its electricity from wind, remains the second-largest corn-producing state in the United States. Those facts coexist without contradiction, because the turbines rise above the crops and the fields between them keep producing.

The genuine land-use concerns for wind tend to be about aesthetics and specific ecosystem impacts rather than agricultural displacement. The Altamont Pass wind farms were sited, inadvisably, in a migration corridor for golden eagles and other raptors. The resulting bird mortality — roughly 2,000 to 4,000 raptors per year at peak operation, according to estimates from the California Energy Commission — was a legitimate ecological problem, and a cautionary tale about siting. Modern wind developers use bird and bat mortality studies as standard practice. Radar-linked shutdown systems, which briefly feather blades when raptors approach, have been retrofitted at Altamont and deployed elsewhere. The problem is not solved — turbines kill birds — but it is manageable, and the numbers are small relative to the annual mortality inflicted by cats, buildings, and vehicles.

The NIMBY Problem

The more politically charged land objection is not ecological. It is aesthetic. Or rather, it is presented as aesthetic but is often something else.

Cape Wind is the most cited American example. The project proposed 130 turbines in Nantucket Sound off the Massachusetts coast. It was first proposed in 2001. It was finally cancelled in 2017 after sixteen years of permitting battles. The opposition was led not by environmentalists but by wealthy property owners — most famously including the Kennedy family, whose compound at Hyannis Port would have had a distant view of the turbines — backed by Alliance to Protect Nantucket Sound, a group that raised tens of millions of dollars to fund legal challenges. The environmental arguments marshalled against Cape Wind were largely pretextual; the same people opposing it rarely showed comparable enthusiasm for other environmental causes. It was, at bottom, a wealthy enclave deploying legal process to protect property values and views.

Cape Wind is not the only story. Opposition to wind projects has been substantial and genuine — communities near turbines do experience audible noise, and some people report sleep disruption from the low-frequency thumping of blades at close range. The UK government imposed a de facto moratorium on new onshore wind development in England between 2015 and 2022 by requiring unanimous local consent — a standard applied to no other form of infrastructure. Denmark, by contrast, has managed community opposition through mandatory profit-sharing arrangements, in which residents within four kilometres of a wind project are offered ownership stakes. The Danish approach has not eliminated opposition. It has reduced it substantially, because it converts potential opponents into co-owners with financial interests in the project's success.

The visual intrusion argument is philosophically awkward territory. Transmission pylons cross the English countryside by the thousand and are largely invisible to cultural consciousness. Gas compressor stations and oil refineries are not commonly described as aesthetic outrages. Wind turbines are large and move, which makes them impossible to ignore. Whether that constitutes a harm sufficient to override their public benefits is a genuine values question that different societies have answered differently. Scotland, which has extensive onshore wind and is approaching 100 percent renewable electricity generation on an annual basis, has largely decided the answer is no. England, until recently, had decided the opposite.

Offshore: A Different Proposition

Offshore wind sidesteps most of the NIMBY problem at a stroke. You cannot hear a turbine ten kilometres out to sea. You can barely see one at twenty kilometres. The visual protests from shoreline homeowners become politically harder to sustain — though not, as Cape Wind demonstrated, impossible — and the wildlife concerns shift from raptors to marine mammals, where the evidence of harm from operational turbines (as distinct from construction noise) is considerably weaker.

The offshore resource is also better. Wind speeds over open water are consistently higher and less turbulent than over land, because there is no surface roughness — no hills, buildings, or treelines disrupting the flow. The North Sea, which has become the proving ground for the offshore industry, combines shallow water suitable for fixed-bottom turbines with some of the most consistent and powerful wind resources in the northern hemisphere. A modern offshore wind farm in the North Sea achieves capacity factors of 45 to 55 percent, compared to 35 to 45 percent for well-sited onshore projects.

The first commercial offshore wind farm was Vindeby, off the Danish coast, which opened in 1991 with eleven turbines of 450 kilowatts each — a total capacity of 4.95 megawatts. It was a demonstration project, almost a curiosity. Horns Rev 1, opened in 2002 off the Danish west coast with 80 turbines of 2 megawatts each and a total capacity of 160 megawatts, was the first genuinely large offshore project and the one that established the template. By the time Hornsea One off the Yorkshire coast was commissioned in 2019 — 174 turbines, 7 megawatts each, total capacity 1,218 megawatts, at the time the world's largest offshore wind farm — the industry had developed purpose-built installation vessels, subsea cable systems, and marine logistics supply chains that simply did not exist when Horns Rev was being built.

Hornsea Two, adjacent to Hornsea One, came online in 2022 at 1,386 megawatts. Hornsea Three is permitted and in development. Dogger Bank, further north in the North Sea, is being built in three phases that will together produce around 3,600 megawatts — larger than most nuclear power stations, though producing power at a lower capacity factor. These are not prototypes. They are large industrial infrastructure projects.

Global offshore wind capacity reached about 64 gigawatts by the end of 2022, according to IEA data, with the United Kingdom holding roughly 14 gigawatts of that total and China growing rapidly from a standing start to become the world's largest offshore wind market. The IEA's Stated Policies Scenario projects offshore wind reaching nearly 380 gigawatts by 2030. Even the more conservative analysts concede that the offshore industry has moved from demonstration phase to mature industry faster than almost anyone expected.

Why Offshore Is Hard (and Expensive)

The problems offshore wind has not solved are real ones, and they deserve honest treatment.

The first is cost. Offshore wind is more expensive than onshore — materially so. In 2021, the levelised cost of offshore wind globally was roughly $75 to $90 per megawatt-hour, versus under $33 for onshore. Installing a turbine in 30 metres of water, 30 kilometres from shore, requires purpose-built jack-up vessels that cost $100,000 to $200,000 per day to operate. The monopile foundations — steel cylinders driven into the seabed — for a 15-megawatt turbine weigh over 1,000 tonnes and require pile-driving equipment that sends acoustic shockwaves through the water column that must be managed to protect marine mammals. The inter-array cables connecting turbines, the export cables running back to shore, the offshore substations transforming voltage — all of this infrastructure sits in a saltwater environment that corrodes, and must be maintained by vessels that can only operate in a narrow window of sea states. The operations and maintenance costs for offshore are roughly twice those for onshore, per unit of energy produced.

The second problem is grid connection distance. Siting turbines far offshore, where wind resources are best and visual opposition is weakest, means running high-voltage cables across substantial lengths of seabed. Beyond about 50 to 80 kilometres, alternating current (AC) cable systems lose too much power to resistive losses to be practical. Projects at greater distances require high-voltage direct current (HVDC) transmission, which adds cost and requires converter stations at both ends. Dogger Bank sits roughly 130 kilometres offshore; its grid connection is HVDC. Floating wind projects off the Norwegian coast, in water too deep for fixed foundations, might sit 150 to 200 kilometres from shore. The cable infrastructure alone can represent 20 to 30 percent of total project cost.

The third problem is hurricanes, which matters specifically for the United States. New England and the Mid-Atlantic states have the best combination of shallow water and political ambition for offshore wind in the country. They are not, in normal circumstances, in the hurricane belt. But major storms track northward along the Atlantic coast with some regularity, and turbines designed for North Sea conditions — mean wave heights, maximum wind speeds, typical storm durations — may not be rated for the worst-case storm events that the Atlantic coast can produce. The 15-megawatt machines currently being installed are designed to

survive 70-metre-per-second wind speeds, which covers most scenarios. Direct hurricane strikes remain a tail risk that is priced into insurance rather than engineered out.

Floating Wind: The Frontier

Fixed-bottom offshore wind — monopile or jacket foundations driven or pinned to the seabed — works in water depths up to roughly 60 metres. Beyond that, the foundations become uneconomic. This matters because the best wind resources in parts of the world, including the Pacific coast of the United States, most of Japan's coastline, Norway, and parts of the Mediterranean, are in deeper water. Scotland's west coast has an extraordinary wind resource over the Atlantic in water depths of 100 to 300 metres.

Floating wind turbines — platforms moored to the seabed by chains or cables, carrying a turbine mast above the waterline — are not a new concept. The physics is well understood. The engineering is solved in principle. What has not yet been demonstrated at commercial scale is whether floating wind can achieve costs competitive with fixed-bottom offshore, let alone onshore wind.

Hywind Tampen, a floating wind project off Norway that came online in 2022, powers five offshore oil platforms and represents the largest operational floating wind installation in the world at 88 megawatts. It is genuinely impressive engineering. Its cost was also roughly three times that of comparable fixed-bottom offshore projects. The learning curve for floating wind is where fixed-bottom offshore was in the mid-2000s: the technology works, but the supply chains, installation methods, and manufacturing scale required to bring costs down do not yet exist. Some analysts project floating wind costs converging with fixed-bottom offshore by the 2030s. That projection requires assumptions about manufacturing scale and supply chain development that are plausible but not guaranteed.

What Wind Cannot Do Alone

By the early 2020s, wind was generating about 7 percent of global electricity, according to IEA data. That share will grow. Denmark generates more than half its electricity from wind on an annual basis — a figure that would have seemed fantastical to a Danish engineer in 1990. Iowa, Texas, and Kansas generate 50 to 60 percent of their electricity from wind, demonstrating that even within the United States' fragmented and politically contentious grid, high wind penetrations are operationally achievable.

But wind is variable in ways that matter, and the variability has a specific, uncomfortable character. It does not vary randomly, like noise. It varies with weather patterns that persist for days. During the anticyclonic, high-pressure "dunkelflaute" events — still air and cloud cover that can sit over northern Europe for a week or more in winter — both wind and solar output drop together. Germany and Denmark, with substantial installed capacity of both, experienced such events in the winters of 2021 and 2022, relying heavily on gas and imports to maintain grid stability. The problem is not that wind produces no electricity during a dunkelflaute. It is that it produces very little electricity at the same time solar produces very little electricity, at times of year when heating demand is highest.

This is the storage and grid problem, which the next chapter addresses in detail. But it is worth flagging here: the intermittency of wind is not a flaw that better turbines will solve. It is a feature of the atmosphere. The wind does not blow steadily because the climate does not warm and cool steadily. No amount of engineering improvement to turbine design changes that.

What wind offers, then, is this: very cheap electricity when the wind is blowing, zero fuel cost, falling capital costs, and an energy source with a land footprint modest enough that it can coexist with agriculture. What it requires is either storage (expensive), complementary generation from sources with different profiles (nuclear, hydro, or some fraction of gas), demand flexibility, or long-distance transmission links to average out regional weather variability. Wind in Texas is not correlated with wind in the Carolinas. Wind in the North Sea is not perfectly correlated with wind in the Baltic. Larger grids with better interconnection naturally smooth the variability problem, which is one reason European ambitions for expanded HVDC transmission are not simply bureaucratic enthusiasm but physical necessity.

The Onshore Arithmetic

Let's be specific about what onshore wind can realistically contribute. The United States had about 145 gigawatts of installed onshore wind capacity at the end of 2022, generating roughly 10 percent of national electricity. At average capacity factors of around 34 percent, that produced about 434 terawatt-hours of electricity that year.

US electricity consumption is around 4,000 terawatt-hours annually. Getting to, say, 40 percent wind penetration would require something in the range of 500 to 600 gigawatts of installed capacity — roughly a tripling of what currently exists. That is physically achievable; the US wind resource is enormous, concentrated in the Great Plains states with sufficient land for many times that capacity. The constraints are not physical. They are transmission (the Plains states are not where the demand centres are), permitting timelines (federal permitting for onshore wind averages years, not months), and the intermittency management problem already noted.

Europe's onshore wind resource is smaller in absolute terms and more constrained by land-use policy, proximity to populated areas, and ecosystem sensitivities. The UK's decision to effectively ban onshore wind in England for seven years was not costless; it meant offshore projects, at roughly double the levelised cost, had to carry a larger share of the renewable buildout than would have been optimal. That moratorium has now ended, and the planning system has been partially reformed. The efficiency loss it caused, measured in billions of pounds of excess consumer energy costs, is real even if it is invisible.

The Bigger Picture

Vaclav Smil, who applies rigorous scepticism to almost every energy technology claim, is not dismissive of wind. He is calibrating. In How the World Really Works, he notes that wind turbines are the largest machines humanity has ever mass-produced, and that manufacturing the steel, fibreglass, and rare-earth magnets they

require imposes its own material and energy demands. A wind turbine does not emit carbon dioxide during operation. It emits a considerable amount during manufacture, transport, and installation. The carbon payback period — the time required to offset the emissions embedded in a turbine's construction — is typically two to three years for a modern onshore machine with an operational life of twenty-five years or more. That is a favourable ratio. It is not zero.

The rare-earth issue is worth a sentence. Direct-drive offshore turbines — which avoid gearboxes by using permanent magnet generators — use significant quantities of neodymium and dysprosium, elements whose global production is currently dominated by China. This is a supply chain vulnerability, not a fundamental resource constraint; deposits exist on multiple continents. But it is a reminder that decarbonising the power sector trades fossil fuel dependencies for materials dependencies, and the geopolitical implications of that trade are not yet fully absorbed by the countries building out their offshore wind programs most aggressively.

A Reckoning

Wind energy's story from 1982 to 2024 is, by any honest accounting, one of the most successful technological learning-curve stories in energy history. A technology that produced expensive, unreliable electricity from small machines has become the cheapest or second-cheapest source of new electricity generation on the planet, depending on location. The engineering trajectory — governed by the relentless logic of the cube law — has produced machines that would have seemed fictional to an engineer standing at Altamont Pass in 1985. Turbines taller than the Washington Monument. Blades longer than a football field. Machines installed in 50-metre swells by vessels that barely existed a decade ago.

The politics has been messier. Opposition that is sometimes principled, often self-interested, occasionally both, has slowed buildout in England, tied up projects in American courts for years, and in the Cape Wind case prevented a project that would have powered most of Cape Cod and the Islands from being built at all. Those delays have costs — in emissions, in money, in years of continued dependence on natural gas that the projects would have displaced. They also reflect genuine democratic processes, which means they cannot simply be dismissed as obstacles to be engineered around.

What remains true is that wind — offshore in particular — is no longer a fringe technology or a subsidy-dependent curiosity. It is a major industrial sector producing genuinely cheap electricity at scale. The 15-megawatt turbines going into the North Sea in the mid-2020s will last twenty-five to thirty years, generating power at costs below most alternatives, without emitting a gram of carbon dioxide during operation.

That is remarkable. It also does not solve the storage problem, the grid problem, or the industrial decarbonisation problem. Wind makes electrons cheap when the wind blows. The rest of the energy system then has to figure out what to do with that fact. Those are the subjects of the chapters ahead.

CHAPTER VIII

Storage: The Unsexy Problem That Decides Everything

The Duck and the Arithmetic

On a mild spring afternoon in California, the sun is generous and the demand for electricity is not. People are at work. Air conditioners are off. Industrial load is moderate. And yet the solar panels covering rooftops and utility farms across the state are pushing power onto the grid at full tilt, flooding a system that doesn't need it. By late afternoon, the sun begins its retreat, industrial shifts end, commuters get home, ovens switch on, and the grid suddenly needs enormous amounts of power in a very short time — power the solar panels can no longer supply.

Grid operators at the California Independent System Operator drew a graph of this pattern in 2013. The net load curve — total demand minus solar generation — dips through the middle of the day, then swoops back up steeply in the evening like the curved neck and rounded belly of a waterfowl. They called it the duck curve. A decade later, the duck was no longer a projection. It was Tuesday.

This is the problem. Not the cost of solar panels, which has collapsed astonishingly (more on that in the previous chapter). Not the efficiency of wind turbines or the speed of their installation. The problem is time. Energy does not wait. Electrons generated at noon cannot be used at seven in the evening unless something, somewhere, holds them in the interim. The physics doesn't care about your renewable energy targets.

Every serious analyst of the energy transition eventually arrives at the same conclusion. Generation is the easy part. Storage is where the physics fights back.

What "Storage at Scale" Actually Means

Before surveying the technologies, it's worth pausing on the numbers. The scale of the storage challenge is routinely underestimated, and that underestimation has real policy consequences.

Global electricity consumption in 2022 was approximately 25,500 terawatt-hours, according to IEA data. That's per year. Divided by 8,760 hours, the average instantaneous global draw on the electrical grid is roughly 2.9 terawatts — and the peaks are considerably higher. A commonly cited estimate for a deeply decarbonised, largely renewable global grid is that you'd need somewhere between one and four terawatt-hours of economically accessible storage just to manage day-to-day and week-to-week mismatches between generation and demand. The upper end of that range — call it four terawatt-hours — is roughly twenty times more grid-connected electrical storage than the world had installed by the end of 2023.

Vaclav Smil, the Czech-Canadian energy scholar whose patient quantitative work underpins much of this book, has a phrase for moments like this: he talks about the "scale and complexity of the challenge." It sounds bureaucratic until you look at the actual numbers. In How the World Really Works, published in 2022, Smil notes that the entire American economy runs on energy flows equivalent to burning roughly 15 billion tonnes of coal per year. The energy stored in all the grid-scale batteries in the world, as of that writing, could power the United States for a matter of hours. Not days. Hours.

This is not an argument against batteries. It's an argument for clarity about what they are — and what else we're going to need.

The Giant No One Talks About

Here is a technology that stores electricity at enormous scale, has round-trip efficiency of roughly 70 to 85 percent, carries no toxic materials concerns, and has been operating reliably for over a century. It is not lithium-ion. It is not hydrogen. It is not the subject of TED talks or breathless technology journalism.

It is water flowing uphill.

Pumped hydroelectric storage works on a principle a child could understand: use surplus electricity to pump water from a lower reservoir to a higher one, then release it through turbines when electricity is needed. The energy is stored as gravitational potential. The world's first pumped storage plant opened in Zurich in 1909. Today, pumped hydro accounts for roughly 90 percent of all grid-scale electricity storage capacity worldwide — approximately 9,000 gigawatt-hours of stored energy spread across some 160 gigawatts of installed power capacity, according to IEA estimates.

Read that again. Ninety percent.

The Bath County Pumped Storage Station in Virginia — tucked into the Allegheny Mountains, largely invisible to most Americans — has 3,003 megawatts of generating capacity and stores about 24,000 megawatt-hours. It was completed in 1985. It remains, as of this writing, the largest pumped storage facility in the United States, and the third largest in the world. The Fengning Pumped Storage Power Station in Hebei, China, which reached full operation in 2023, tops out at 3,600 megawatts.

Pumped hydro has obvious constraints. It requires the right geography — two reservoirs separated by significant elevation, connected by tunnels, located near a transmission grid. Those requirements exclude most of the places that most need storage. There are no realistic sites for pumped hydro in the Netherlands, in Bangladesh, or in the flat agricultural midlands of the United States. Building new facilities takes a decade or more and faces enormous permitting hurdles. And the energy density is low in absolute terms: a cubic metre of water lifted 100 metres stores only about 0.27 kilowatt-hours. You need a lot of water, a lot of height, or both.

But for the places where it exists, pumped hydro provides something batteries currently cannot: storage duration measured in hours or even days, at costs that tend to be far lower per kilowatt-hour than any battery

chemistry. When people talk about the need for "long-duration storage" — a phrase that has become something of a mantra in energy circles — they are often talking about building new pumped hydro, or at least wishing they could.

Australia has been exploring something called "closed-loop" pumped hydro, which doesn't require a natural river or existing reservoir but instead builds two artificial reservoirs. The Australian National University's Andrew Blakers has mapped over 600,000 potential sites globally. Whether enough of them are politically and ecologically acceptable is a different question — one the maps don't answer.

Lithium-Ion: The Darling Has Real Limits

In December 2017, a Tesla Powerpack installation in Hornsdale, South Australia, achieved something no grid battery had done before at commercial scale: it responded to a sudden generation outage in 140 milliseconds — faster, by a factor of roughly 70, than a gas-fired "peaker" plant could spin up. The system, rated at 100 megawatts and 129 megawatt-hours, became an instant icon of the battery-storage era. Elon Musk had bet that it could be installed in 100 days or he'd give it away for free. It was installed in 63.

The story spread because it was genuinely remarkable. Lithium-ion batteries can respond to grid frequency deviations in fractions of a second. They can be installed in months on almost any flat piece of land. Their costs have dropped by roughly 90 percent since 2010, from around $1,200 per kilowatt-hour of capacity to somewhere below $140 by 2023, according to BloombergNEF's annual battery price survey. Round-trip efficiency — the fraction of electricity you put in that you get back out — runs between 85 and 95 percent for modern lithium-ion systems, which is better than every competing grid-storage technology except some novel thermal approaches.

For short-duration applications — two to four hours of storage, the window needed to shift solar generation from the afternoon into the early evening peak — lithium-ion is already cost-competitive with gas peaker plants in many markets. The IEA's 2023 data shows over 45 gigawatts of new battery storage connected to grids globally in that year alone, nearly triple the additions of 2021. The growth curve looks almost as dramatic as the solar cost curve.

So. Problem solved?

No. And this is where the honest accounting gets uncomfortable.

The Hornsdale system stored 129 megawatt-hours. South Australia's average electricity consumption is around 1,800 megawatt-hours per hour. The battery could power the state for about four minutes. Its value wasn't bulk energy storage — it was grid services. Frequency regulation. Rapid response. Ancillary services that keep the lights on moment to moment. It did that job brilliantly. But confusing "grid stabilisation at short timescales" with "solving the storage problem" is like confusing a pressure valve with a reservoir.

The deeper issue is duration. A lithium-ion battery storing electricity for two hours costs, at current prices, roughly $280 per kilowatt-hour of capacity. For four hours, the cost doubles. At 100 hours — the kind of

storage needed to carry a renewable-heavy grid through a cloudy, calm winter week — the economics fall apart entirely. The chemistry doesn't have some inherent long-duration capability that engineers just haven't unlocked. You simply need more cells, and more cells cost more money, and at 100 hours of duration you'd be spending more on batteries than the electricity is worth.

There's also the question of materials. Lithium-ion batteries require lithium, of course, but also cobalt, nickel, and manganese. Global lithium reserves are not trivially small — Australia, Chile, and Argentina together hold the vast majority — but scaling production to meet demand for both electric vehicles and grid storage simultaneously is a genuine supply-chain challenge. A 2020 paper in Nature by Olivetti et al. modelled several scenarios and found that cobalt supply constraints could become binding by the early 2030s under aggressive EV deployment. The industry has moved aggressively toward lithium-iron-phosphate (LFP) chemistries that eliminate cobalt, and China has led that shift, but LFP has lower energy density and performs worse at cold temperatures — trade-offs that matter for grid applications in northern latitudes.

None of this is fatal. It's a set of constraints that shapes where and how lithium-ion fits into a storage portfolio. The mistake is treating it as the entire portfolio.

The Runners-Up: Compressed Air, Flow Batteries, Thermal

Compressed air energy storage works by using off-peak electricity to compress air into underground caverns, then releasing it through turbines when electricity is needed. The first commercial plant opened in Huntorf, Germany, in 1978. There is another at McIntosh, Alabama, operating since 1991. That's it. Two plants, in over four decades.

The reason the technology hasn't scaled is partly geological — you need the right kind of salt cavern or hard-rock formation — and partly thermodynamic. When you compress air, it heats up. When you store that compressed air and later release it, the air cools and loses pressure. The Huntorf plant burns natural gas to reheat the air before expansion, which means it's not purely an electricity-storage device. "Adiabatic" CAES designs that capture and reuse the compression heat look better on paper, with projected round-trip efficiencies of 60 to 70 percent, but remain at the demonstration stage. Hydrostor, a Canadian company, is building what would be the first utility-scale adiabatic CAES system in California. The target commissioning date has been pushed back several times. That's not unusual for novel infrastructure, but it's worth noting.

Flow batteries represent a different approach: instead of storing energy in solid electrode material (as lithium-ion does), they store it in liquid electrolytes held in external tanks. The larger the tanks, the more energy you store. This decoupling of power capacity from energy capacity is genuinely useful — it means you can scale up duration cheaply by just adding more electrolyte, rather than adding more of the expensive electrochemical stack. Vanadium redox flow batteries, the most mature commercial variant, have been deployed at multi-megawatt scale in China and Japan. Round-trip efficiency is roughly 65 to 80 percent. The problem is cost per kilowatt-hour of energy capacity, which remains significantly higher than lithium-ion for durations below about six hours, and the fact that vanadium itself is a somewhat obscure commodity with an

insecure supply chain. Researchers at Harvard have been exploring organic flow batteries using quinones derived from plants, which would eliminate the vanadium dependence entirely — but this work is still in the laboratory.

Thermal storage deserves more attention than it usually gets. The idea is to store energy as heat rather than as electricity. In concentrated solar power plants — the kind that use mirrors to focus sunlight on a central receiver rather than photovoltaic cells — molten salt has been used as a storage medium for over a decade. The Gemasolar plant in Andalusia, Spain, uses a mixture of sodium and potassium nitrates heated to around 565 degrees Celsius, allowing the plant to generate electricity for up to 15 hours after sunset. Crescent Dunes in Nevada used a similar system with 10 hours of storage. These aren't exotic experiments: they're operating commercial plants.

For industrial heat — which accounts for roughly half of global final energy demand and which electricity struggles to supply cost-effectively — thermal storage in materials like molten salt, hot rocks, or phase-change materials may ultimately be more important than electrochemical storage. A startup called Rondo Energy has built commercial units that heat bricks to around 1,500 degrees Celsius using excess renewable electricity and release that heat for industrial processes. This is less glamorous than a Tesla battery wall. It may also be more important for decarbonising a cement kiln or a steel reheat furnace.

Hydrogen: The Great Electrochemical Hope With a Thermodynamic Problem

Hydrogen is the element around which the most ambitious visions of future energy storage are built. The basic idea: use surplus renewable electricity to split water via electrolysis, producing hydrogen; store the hydrogen; burn it or use it in a fuel cell when electricity is needed. The appeal is obvious. Hydrogen has high energy density by mass. It can be stored in tanks, in pipelines, in underground caverns. It can serve as both an electricity storage medium and a feedstock for industrial processes that need a carbon-free reducing agent — steelmaking and ammonia synthesis chief among them.

The problem is efficiency. Every step in the hydrogen cycle wastes energy, and the steps compound. Modern electrolysers operating at green-hydrogen scale convert electrical energy to hydrogen with roughly 65 to 75 percent efficiency. Storing and transporting hydrogen imposes additional losses. Converting it back to electricity via fuel cell or combined-cycle turbine adds another round of losses. End-to-end round-trip efficiency for power-to-hydrogen-to-power is somewhere between 25 and 40 percent, depending on the system. Compare that to lithium-ion at 85 to 95 percent, or pumped hydro at 70 to 85 percent.

That number matters enormously. If you start with 100 units of renewable electricity, you get about 30 units back when you use hydrogen as a storage medium. Lithium-ion gives you back 90. Pumped hydro gives you back 75. Hydrogen is, thermodynamically, a deeply inefficient storage medium for electricity.

This does not mean hydrogen is useless. It means hydrogen makes sense only in specific contexts: where duration of storage is very long (weeks or seasonal), where the alternative to storage is curtailment of large

renewable surpluses, where the end use is not electricity reconversion but direct fuel use (aviation, shipping, steelmaking), or where pipeline infrastructure makes hydrogen transport cheaper than electricity transmission. These are real use cases. They're just narrower than the hydrogen hype cycle has often suggested.

The IEA's 2023 Global Hydrogen Review noted that global clean hydrogen production — meaning hydrogen made with near-zero emissions — amounted to less than 0.1 million tonnes in 2022, against a target of around 380 million tonnes by 2030 in net-zero scenarios. The gap between aspiration and current trajectory is severe. Green hydrogen costs roughly $4 to $9 per kilogram today, against a threshold of around $1 to $2 per kilogram needed for broad industrial competitiveness. The electrolysers are improving, and costs are coming down, but "coming down" is not the same as "arrived."

Iron-Air and the Long-Duration Wild Cards

In a quiet industrial park in Weirton, West Virginia — a town that has seen better decades — a company called Form Energy began construction in 2023 on a factory intended to produce iron-air batteries. The technology is disarmingly simple: iron oxidises (rusts) in air, releasing electrons; running current through the system reverses the reaction, reducing the iron oxide back to metal. The cycle can continue indefinitely. Iron is cheap, abundant, and non-toxic. The round-trip efficiency of about 50 percent is well below lithium-ion, but for long-duration storage the economics may still favour iron-air because the materials cost is so low.

Form Energy claims a target of around $20 per kilowatt-hour of storage capacity, compared to $130 to $150 for lithium-ion as of 2023. At that price point, 100-hour storage becomes economically interesting. You can afford to waste half the electricity if the capital cost of holding it for four days is low enough. Georgia Power signed a purchase agreement with Form Energy in 2022 for a 15 megawatt / 1,500 megawatt-hour system — 100 hours of storage from a single 15-megawatt block. That's the kind of storage that starts to matter for multi-day weather events.

The caveats are real. Iron-air batteries have never operated at grid scale before. The West Virginia factory hasn't yet demonstrated sustained commercial production. "Promising laboratory technology" has a long and humbling history of failing to scale. Gravity storage, using heavy blocks lifted by renewable electricity and lowered through generators when power is needed, has similar appeal on paper — the British company Gravitricity has done promising small-scale tests in an old mine shaft — but no one has yet built a large-scale commercial system.

The honest assessment of long-duration storage alternatives: compelling physics, early-stage engineering, unproven economics. Worth watching carefully. Not yet worth building policy around.

The Curtailment Signal

One way to understand the urgency of the storage problem is to look at what happens when storage is absent. Grid operators deal with surplus generation by curtailment — shutting down renewable plants, or in the case of wind paying operators to stop producing. In 2016, China curtailed approximately 49.7 terawatt-hours of

wind generation, according to China's National Energy Administration. Forty-nine point seven terawatt-hours. To put that in perspective: that's more electricity than Sweden consumes in a year.

California's duck curve stopped being a projection and became a reality by 2017. In April of that year, the grid operator curtailed over 80,000 megawatt-hours of renewable generation in a single month — electricity generated but not used. The number has grown since. In 2022, California curtailed roughly 2.4 million megawatt-hours of solar and wind combined. That's both a waste and a signal: a signal that the grid has more generation than storage and transmission infrastructure can handle.

Germany's Energiewende has produced similar dynamics. Germany has invested enormously in wind and solar — renewable electricity accounted for over 46 percent of German power generation in 2022 — but the country has also struggled with "dark doldrums" (Dunkelflaute in German), periods of cold, still, cloudy weather that can persist for two weeks when renewable output collapses. During those periods, Germany has relied on coal and imports. Without long-duration storage, every renewable-heavy grid needs either backup generation or interconnection to somewhere that has wind when you don't.

Building a Portfolio

So what does a credible grid-storage portfolio actually look like? Not a single technology. Not an ideological bet on the winner. A portfolio, because the grid's problems span different timescales and each technology is better suited to a different slice of the duration spectrum.

For seconds-to-minutes response — frequency regulation, responding to sudden outages, smoothing intermittent generation — lithium-ion batteries are already the best tool and their costs are falling fast enough that they'll displace most gas peakers within a decade in markets with decent solar resources. This problem is, in rough terms, solved.

For hours of storage — shifting solar generation from midday into the evening peak, bridging wind lulls that last through the night — lithium-ion at two to eight hours of duration is competitive and becoming more so. The market is growing explosively. By 2030, the IEA projects something between 500 and 680 gigawatts of battery storage on grids globally in its announced pledges scenario. That's real and significant.

For days of storage — carrying a high-renewable grid through a two-day calm or a three-day overcast period — pumped hydro remains dominant where geography allows, and long-duration technologies like iron-air and adiabatic compressed air need to reach commercial scale. This is where the portfolio has the largest gap today.

For weeks of storage — seasonal storage, bridging the gap between summer solar abundance and winter heating demand — hydrogen or ammonia in underground caverns is the most credible option on the table, but at current efficiency and cost levels it is expensive enough that most analyses suggest a different strategy: overbuilding renewable generation to the point where even a poor week produces adequate power, combined with demand flexibility and cross-regional transmission. Storage at seasonal scale is not nothing. But transmission may often be cheaper.

Hannah Ritchie, whose data work at Our World in Data has done much to clarify the energy transition quantitatively, makes a point that bears on portfolio thinking: storage and transmission are substitutes. A solar surplus in Spain can power Germany if the transmission infrastructure exists. A wind surplus in Morocco can power Egypt if a cable crosses the Mediterranean. The 2022 SuNSET study published in Nature Energy modelled global renewable potential with expanded HVDC transmission and found that continental-scale grids could handle much higher renewable penetrations before requiring significant storage — essentially trading storage capital cost for transmission capital cost. Which is cheaper depends on the specific geography and the cost trajectories of each.

Neither is an excuse for avoiding the other. Both need to happen.

The Round-Trip Tax

There is a way to think about storage that cuts through the complexity of competing technologies. Every time you store electricity and recover it, you pay a round-trip efficiency tax. The tax varies by technology: roughly 10 percent for the best lithium-ion systems, 20 to 30 percent for pumped hydro, 30 to 35 percent for flow batteries, 50 percent for compressed air, and 60 to 75 percent for hydrogen. Every unit of electricity generated by a solar panel or wind turbine that passes through storage is not a full unit when it comes back.

In a world where electricity is cheap — as it is increasingly becoming, in sun-rich and wind-rich regions — this tax is affordable. If you're curtailing electricity anyway, storing it at 50 percent round-trip efficiency is infinitely better than wasting it at 0 percent. But the efficiency tax shapes which technologies are worth deploying where, and it means that any scenario involving large amounts of hydrogen storage is implicitly a scenario that requires substantially more renewable generation than scenarios relying on pumped hydro or lithium-ion. More generation costs money. That cost has to appear somewhere in the accounting.

Smil is characteristically blunt about this kind of arithmetic. In Energy and Civilization, he traces how every major energy transition in history required not just new sources but new infrastructure to move, store, and convert energy — and that building that infrastructure was slow and expensive in ways that always surprised optimists. The coal age required the steam engine, the railway, and a national distribution network for fuel. The oil age required refineries, pipelines, and the internal combustion engine. The transition to high-renewable electricity requires storage and transmission at a scale the world has not yet begun to build.

This is not a counsel of despair. It is a counsel of urgency about the right things.

Scale, Honestly

It is worth ending this survey with a number. The world's total installed grid-scale electrical storage capacity — pumped hydro included — was approximately 9,000 gigawatt-hours as of 2023. Global electricity consumption is roughly 25,500 terawatt-hours per year. If you wanted to store just twelve hours of average global electricity demand, you would need approximately 14,400 gigawatt-hours of storage — more than 1.5 times everything the world has built in the entire history of electricity.

Twelve hours. Not weeks, not seasons. Twelve hours of average load.

This arithmetic is not an argument for slowing down on renewables. Solar and wind are already the cheapest sources of new electricity generation in history, and their deployment is accelerating. But every new solar panel and every new wind turbine installed without a commensurate investment in storage, transmission, or demand flexibility is a panel or turbine that will sometimes be curtailed or will require a gas peaker to back it up. The duck gets flatter — or it gets hungrier.

The storage problem does not have a single solution. It has a portfolio: pumped hydro where the mountains allow, lithium-ion for short-duration and ancillary services, long-duration chemistry like iron-air if it achieves commercial scale, thermal storage for industrial heat, hydrogen for specific long-duration and feedstock applications, and transmission to spread geographic diversity across all of it.

That's a lot of things to build simultaneously. It's expensive. It takes time. And it is, without question, the less photogenic half of the energy transition — no sleek product launches, no dramatic cost curves, no viral battery installations. Just concrete and steel and tunnels and wires and decades of permitting battles.

Pumped hydro. The least glamorous technology in energy. Still doing 90 percent of the world's grid-storage work, four decades after the Bath County plant was carved out of the Virginia mountains. There is a lesson in that. The energy transition needs more of the unsexy thing and less of the assumption that the sexy thing will scale fast enough to substitute for it.

It won't. Not on its own.

CHAPTER IX

The Grid: Why the Wires Are Harder Than the Generation

The moment you flip a light switch, you're relying on a machine that has been running, without pause, for the better part of a century. Not one machine exactly — several thousand interlocking pieces of equipment spanning continents, owned by hundreds of different companies and government agencies, operating under regulatory frameworks that were drafted when Eisenhower was president. The wires that carry electricity across American and European landscapes are, in aggregate, the most complicated engineered system human beings have ever assembled. And in most wealthy countries, they are now being asked to do something they were never designed for.

That gap — between what the grid was built to do and what a low-carbon energy future requires of it — is the binding constraint in decarbonisation that almost no politician acknowledges and almost no energy journalist explains clearly. The past decade has produced a genuine miracle on the generation side: solar and wind have become the cheapest new sources of electricity in history. That is not hype. That is numbers. But cheap generation sitting in a field in West Texas or on a hillside in Northumberland is worth nothing if the electrons it produces cannot reach the people who need them. And increasingly, they cannot.

The Physics of Sixty Hertz

Start with what the grid actually has to accomplish every second of every day.

Alternating current — AC — oscillates. In North America, it completes 60 full cycles per second, which is to say it runs at 60 hertz. In Europe and most of the rest of the world, 50 hertz. That number isn't arbitrary. The motors, transformers, and industrial equipment that civilization depends on are calibrated to it. Let the frequency drift more than about half a hertz in either direction and equipment begins to malfunction. Let it drop below roughly 59.3 hertz in the U.S. grid and automatic relays begin tripping — disconnecting generators to protect themselves, which removes supply and drops the frequency further, which trips more relays. This is a cascade. It happens fast. On August 14, 2003, a software bug at a utility in Ohio allowed a cascade to knock out power to 55 million people across the northeastern United States and Canada in about nine seconds.

Nine seconds. From normal operation to the largest blackout in North American history.

The engineers who built the twentieth-century grid solved the frequency problem with a beautiful, largely invisible mechanism: rotational inertia. Large coal plants, gas plants, and nuclear reactors all turn massive turbines — objects weighing hundreds of tonnes, spinning at 3,000 or 3,600 revolutions per minute. When

demand suddenly spikes or a generator trips offline, the grid tries to slow down (frequency drops). But all those spinning masses resist that slowdown. They keep pushing electrons into the wires for the critical seconds it takes for grid operators to bring backup generation online. The inertia is the buffer. It buys time.

A solar panel has no moving parts. Neither does a wind turbine's power electronics. Neither does a battery. These devices connect to the grid through inverters — boxes of clever semiconductors that convert their output to AC — and inverters, in their traditional form, contribute precisely zero inertia to the system. They are, from the grid's perspective, massless. As coal plants retire and renewable generation grows, the system loses the rotational mass that has been stabilizing frequency for a hundred years. That is not a hypothetical future problem. It is measurable today.

Voltage is the second dimension of the problem. Frequency is a grid-wide phenomenon — it is essentially the same everywhere in a synchronous AC system. Voltage is local. It varies along transmission lines depending on what's being generated and consumed at each node. Managing it requires devices that inject or absorb reactive power — VArs, in grid-engineering shorthand — at the right places. Large rotating generators do this naturally. Inverters, again in their basic form, do not. Running a grid where distributed solar generation is sprinkled across thousands of rooftops, at unpredictable levels, at thousands of nodes the original designers never thought of as generation points, complicates reactive power management in ways that are technically solvable but expensive and not yet solved at scale.

The third stability concern is fault current. When something goes wrong in the grid — a tree branch touches a line, a transformer fails — the resulting short circuit produces an enormous surge of current. Protection systems detect this and isolate the fault. Their ability to detect it depends on the fault current being large enough to trigger the relays. Rotating generators are prolific fault-current sources. Inverters are, by design, current-limited — they protect themselves by not producing the surge. In a grid that has replaced most of its rotating machines with inverter-based resources, protection systems need to be redesigned, or they will fail to clear faults, with potentially serious consequences.

None of this means inverter-dominated grids can't work. They can. But making them work safely requires software, hardware, and operational practices that are still being developed, standardized, and deployed. The physics aren't insurmountable. The timelines are the problem.

What the Old Grid Was Built For

To understand why the grid needs fundamental restructuring, you need to picture what it was originally designed to do.

The model that dominated the twentieth century was simple: large power plants, usually located near fuel sources or water, generated electricity at high voltage, sent it down transmission lines to substations, where it was stepped down and distributed to homes and businesses. Power flowed one way — from big central plants outward to the edges. The customers at the edges consumed. They did not generate. The grid's designers didn't have to worry about bidirectional flows because they never happened.

The system was engineered for predictability. Utilities had decades of data on how demand varied by season, by time of day, by temperature. They knew, roughly, when they needed their peaking plants — the gas turbines that could fire up in minutes — and when their baseload units — coal, nuclear — could run flat out. Grid operators became skilled at this balancing act. The job was complicated, but the variables were manageable.

What a high-renewables system presents to those same grid operators is categorically different. Solar generation traces the sun. In California, on a mild spring day with high solar penetration, output can ramp from near zero at dawn to over 15,000 megawatts by midday, then fall back to near zero as the sun sets — just as evening demand is rising. This creates the "duck curve" that California's grid operator CAISO began warning about in 2013 and has been fighting ever since. On March 29, 2024, California's grid met 100 percent of its instantaneous electricity demand from renewables for several consecutive hours — a genuine milestone — and also had to curtail over 2,000 megawatt-hours of excess solar that couldn't be absorbed or stored. The grid that was designed to dispatch generation on demand now has to manage generation that arrives on its own schedule.

Wind is less predictable than solar on short timescales. A weather front moving through Texas can ramp 10,000 megawatts of wind generation up or down in four hours. No mid-twentieth-century grid designer imagined having to balance those swings. The operators and software managing the ERCOT grid in Texas — which runs largely independently of the rest of the U.S. system — have become world-class at forecasting and balancing wind generation. They've had to.

The distribution grid — the lower-voltage network that runs into neighborhoods and buildings — has the same unidirectional assumption baked into its design. When rooftop solar proliferates, electrons start flowing backward up the distribution lines. Transformers that were designed to step voltage down in one direction can be damaged by sustained reverse flow. Voltage at the ends of distribution feeders can rise unpredictably. Protective relays can operate incorrectly. None of this is catastrophic; it's mostly manageable. But "manageable" in this context means utilities spend money on equipment upgrades in a largely reactive, piecemeal way rather than a systematic one.

The Transmission Gap

The grid's most acute problem, however, is not in the distribution network. It's in transmission — the high-voltage lines that carry bulk power over long distances.

The geography of renewable resources and the geography of electricity demand do not overlap neatly. The best onshore wind in the United States is in a band running from North Dakota south through Texas. The best solar is in the Southwest desert. The bulk of the population and electricity demand is in the coastal cities and the industrial Midwest. Connecting them requires long transmission lines, and building long transmission lines in the United States currently takes, on average, somewhere between ten and fifteen years from the decision to build to the moment the line goes live. The IEA's World Energy Outlook 2023 estimated that the

world needs to add or replace 80 million kilometers of electricity transmission and distribution infrastructure by 2040 — roughly double what exists today — and that current build rates are nowhere near sufficient.

Ten to fifteen years. Think about what that means. A developer decides today to build a wind farm in Wyoming that would deliver clean power to Denver. The turbines could be manufactured and installed in three years. The transmission line to carry that power will not be approved, financed, and built until 2037, if things go well. The generation is not the bottleneck. The wire is the bottleneck.

Why does permitting take so long? The honest answer is: many reasons, stacked on top of each other.

In the United States, building an interstate transmission line requires navigating a patchwork of federal and state authorities that share jurisdiction in ways no single regulator fully controls. The Federal Energy Regulatory Commission sets rules for the bulk power system, but it cannot override state permitting processes. A single 500-mile transmission line might cross a dozen counties, each with their own land-use ordinances, environmental review processes, and legal standing to challenge the project. Native American tribal consultations can add years. Endangered species reviews under the Endangered Species Act can require biological surveys across multiple growing seasons. A project that attracts NIMBY opposition — and most large transmission projects do — can be challenged in federal court at several points along the way.

Europe faces variations of the same problem. Germany's Energiewende — its celebrated energy transition — has been visibly constrained by the inability to build transmission lines from the windy north, where most of the renewable generation sits, to the industrial south, where most of the demand sits. Projects that received approval years ago are still not built. One interconnector planned to carry offshore wind power from Schleswig-Holstein south through Bavaria, called the SuedLink, was initially proposed in 2012. After years of planning disputes, route changes, and legal challenges, it is now expected to be completed in 2028. Sixteen years from proposal to completion for a wire.

The regulatory sclerosis has a compounding effect. Because developers know how long permitting takes, they apply for grid connection early — often before projects are fully designed or financed — to hold their place in line. This produces the "interconnection queue," a list of projects waiting for utilities and grid operators to study how their generation would affect the system and what upgrades would be required. In the PJM interconnection, which covers thirteen mid-Atlantic and Midwestern states and is the world's largest competitive electricity market, the interconnection queue had grown to over 2,600 gigawatts of projects by early 2024. Two thousand six hundred gigawatts. The entire installed generation capacity of the United States is about 1,200 gigawatts. Most of the projects in that queue will never be built — many are speculative, many are duplicative, many will be withdrawn — but the utility has to study all of them, and the process of studying them backs up like rush-hour traffic for years. Actual projects that want to connect have to wait their turn behind the gridlock.

FERC has made several attempts to reform the interconnection process. Order 2023, finalized in 2023, introduced "first-ready, first-served" processing to replace the old first-come, first-served approach that encouraged queue stuffing. Whether it will be sufficient is genuinely unclear. The underlying problem — that

the transmission system itself needs to be upgraded before many of the generation projects in the queue can connect without destabilizing the network — is not a process problem. It's a physical and financial one.

Paying for It

Someone has to pay for the wire. And in the American regulatory system, determining who pays has historically been one of the most vicious fights in the energy business.

The basic issue is this: a new transmission line benefits many parties — the wind developer who can now reach markets, the utility that can import cheap power, the customers in cities who get lower electricity bills, the region as a whole when it can share resources across a wider area. But the costs fall on whoever is closest to the line. A utility in Iowa doesn't want to pay for a transmission upgrade whose primary beneficiary is a data center in Virginia. Under traditional cost-allocation rules, regional transmission organizations tried to assign costs to whoever was deemed to benefit, which led to endless disputes, litigation, and, frequently, projects that were never built because the parties couldn't agree on the allocation.

FERC's Order 1000, issued in 2011, was supposed to fix this by requiring transmission planners to consider long-range scenarios and develop cost-sharing arrangements for regional and interregional projects. It has had modest effects. The interregional piece — lines that cross between different planning regions — has largely not happened. The Eastern and Western Interconnections of the U.S. grid are connected by only four high-voltage DC ties with a combined capacity of about 1,200 megawatts. Moving power between them at meaningful scale is essentially impossible. A wind surplus in the Midwest cannot, at present, flow to a solar deficit in the Southeast at any serious scale. The grid is less integrated than it needs to be, and the political economy of cost allocation is a large reason why.

The IEA estimated in its 2023 Electricity Grids and Secure Energy Transitions report that annual global investment in grids needs to reach $600 billion by 2030, roughly double the level of the late 2010s. In the United States, grid investment has been rising — the Bipartisan Infrastructure Law allocated $65 billion to grid upgrades, the largest federal grid investment in history — but the gap between what's needed and what's planned remains substantial. Money is part of the problem. But money flowing into a system whose permitting processes move at glacial pace doesn't automatically produce wire in the ground.

Hardware the Old Grid Never Had

While the regulatory and financial problems grind slowly toward partial resolution, engineers are working on the technical side, and some of what they're building is genuinely transformative.

High-voltage direct current transmission, HVDC, is not a new idea — the technology has been around since the 1950s, and Sweden and Denmark began operating an HVDC link across the Baltic Sea in 1954. But HVDC has become dramatically more capable and cost-competitive over the past two decades, and for long-distance transmission it now clearly outperforms the AC system that dominates most national grids.

The physics reason is simple. AC transmission lines lose power to resistance and to reactive losses that increase with distance. For transmission routes longer than roughly 800 kilometers on land, or 60 kilometers offshore, HVDC becomes more efficient. It also allows power to flow between grids that aren't synchronized — different frequencies, or AC grids that are electrically isolated from each other — which matters enormously for connecting the U.S. Eastern and Western Interconnections, or for linking the British grid to Iceland's geothermal resources.

The Western HVDC Link in the United Kingdom, completed in 2017, runs 2,200 megawatts of capacity under the Irish Sea and across Scotland, connecting Hunterston in Scotland to Connah's Quay in Wales. It cost roughly £1 billion and carries enough power for about two million homes. Projects of this type are now being planned at much larger scales in Europe — the proposed North Sea Wind Power Hub envisions an artificial island in the central North Sea that would serve as a hub for HVDC cables connecting Danish, German, Dutch, and Belgian offshore wind farms to each other and to demand centers, potentially delivering 10 to 150 gigawatts by mid-century. That is not a done deal. It is an aspiration. But the technology to do it exists and is proven.

The inverter is the other hardware revolution that is quietly reshaping what renewable generation can contribute to grid stability. The first generation of solar and wind inverters were "grid-following" — they locked onto the grid's existing frequency signal and followed it. This made them passive from a stability standpoint. Grid-forming inverters, now entering commercial deployment, can establish and maintain a frequency signal themselves, acting more like a virtual synchronous generator. They can provide synthetic inertia, inject fault current, and manage reactive power — all the things that spinning machines used to do that made inverter-dominated grids problematic. Australia's Hornsdale Power Reserve, a large lithium-ion battery installation in South Australia built by Tesla in 2017 and later expanded, demonstrated that battery storage with grid-forming inverters could respond to frequency disturbances in 140 milliseconds — far faster than any thermal power plant. The Australian Energy Market Operator documented this performance and it has become a reference point for what inverter-based resources can do.

Dynamic line rating is a less glamorous but potentially high-impact technology. Traditional transmission lines are operated at fixed capacity ratings — ratings set conservatively to account for worst-case ambient temperatures and low wind speeds, because hot, still air reduces a line's ability to shed heat, which limits how much current it can carry safely. In practice, most of the time, ambient conditions are better than worst case. Dynamic line rating uses weather data, sensor readings on the line itself, and models to calculate the line's actual real-time capacity rather than its conservative nameplate rating. Studies from National Grid in the UK and from American utilities have found that dynamic line rating can increase effective transmission capacity by 10 to 40 percent on existing infrastructure — no new wires, no new permitting, just better use of what's already there. This is not a silver bullet, but it's the kind of incremental gain that matters when you can't build fast enough.

Flexible AC transmission devices — STATCOMs, synchronous condensers, thyristor-controlled series capacitors — have also expanded in deployment. These are the tools that manage reactive power and voltage

stability in a system where the mix of generation technologies is changing. They are expensive. They require skilled engineers to specify and maintain. Most distribution utilities in the United States do not have deep benches of power electronics expertise, and hiring it is a workforce constraint that gets very little attention in mainstream energy policy discussions. There are only so many power systems engineers. Training more takes time that the transition timeline doesn't comfortably allow.

The Software Layer

Hardware is visible. Software is not, which is perhaps why it gets even less attention than it deserves.

Grid operators run systems that have to balance supply and demand in real time across thousands of nodes, while simultaneously planning for contingencies — the sudden loss of a transmission line or generator — and clearing markets in which generators bid to supply power minutes, hours, or days ahead. These systems were built on computational assumptions appropriate for a world of large, dispatchable generators and slowly varying demand. The software that runs the PJM energy market, for instance, has roots going back decades. Adapting it to handle sub-hourly pricing intervals, distributed generation at thousands of nodes, real-time demand response from smart appliances and batteries, and the stochastic output of weather-dependent generation is not a minor upgrade. It is a near-total architectural rebuild. Several grid operators are in the middle of that rebuild now. It is not going smoothly everywhere.

Advanced distribution management systems — ADMS — are the software equivalent at the distribution level. They give utilities real-time visibility into what's happening on distribution networks that were previously operated largely on estimates and assumptions. They can automatically reconfigure feeders around faults, manage distributed solar and battery storage, and integrate data from smart meters. They are expensive and complex to deploy, and the majority of U.S. distribution utilities are not yet running them. The American Society of Civil Engineers' 2021 Infrastructure Report Card gave U.S. energy infrastructure a C- grade. That is not a rating that suggests readiness for a rapid transformation.

The market design question is ultimately a software and regulatory question as much as a physical one. Electricity markets in most U.S. regions still do not fully compensate the services that a modern grid needs — inertia, fault current capability, fast frequency response, reactive power. They were designed to pay for energy (megawatt-hours) and some capacity (megawatts available when needed), not for grid-stabilizing services. As the mix of generation shifts, the services that were previously free — because spinning turbines provided them automatically — become scarce and valuable and need explicit market mechanisms to procure them. Getting those market rules right is a years-long process of regulatory proceedings, stakeholder comment periods, and technical analysis. It is happening. Slowly.

The Interconnection Queue Is a Symptom

Return for a moment to that PJM queue and its 2,600 gigawatts.

The queue is often presented as a permitting problem or a bureaucratic problem, and it is both of those things. But at a deeper level it is a symptom of the mismatch between where the grid was built and where the best renewable resources are. When a developer proposes to connect a wind farm in a part of the PJM grid that was not designed to carry large power exports, the interconnection study typically finds that significant network upgrades are required — new lines, new substations, new equipment. The cost of those upgrades gets assigned, under current rules, to the developer who triggered them. This can make projects economically unviable even when the generation itself is cheap. The developer withdraws. The next developer in line faces the same study. The queue moves slowly.

There is a conceptual fix that a number of analysts — including Jesse Jenkins at Princeton and various IEA authors — have been arguing for: proactive transmission planning. Rather than waiting for generators to propose connections and then studying them one at a time, grid planners would identify the transmission corridors the energy system will need over the next twenty to thirty years, build them as shared public infrastructure, and let generation connect to a grid that is ready for it. This is how Germany originally envisioned its offshore wind expansion — coordinated offshore grid planning, not a tangle of individual developer cables. The U.S. has no mechanism to do this at the federal level. Regional transmission organizations can do some of it. Getting interstate and interregional projects built requires interstate compacts or federal action of a kind that hasn't materialized.

The problem is not lack of ideas. It is that proactive transmission planning requires someone to spend money before the generation projects that will use the new lines are certain to be built, which requires either a regulatory mandate to spread the costs broadly or a public entity willing to take the risk. Neither is easy in a country where electricity policy is a patchwork of state and federal jurisdiction and where any major federal energy initiative faces intense political opposition.

Texas, February 2021

One week in February 2021 showed what happens when grid infrastructure is not adequate to the conditions it faces.

A severe winter storm descended on Texas. Temperatures fell to -2°F in some places — temperatures the state hadn't seen in decades. Natural gas wellheads and pipelines froze. Coal plants froze. A few wind turbines froze, though the failure of gas generation was quantitatively far larger. Demand for heating surged while supply collapsed. The ERCOT grid came within minutes of a complete, uncontrolled blackout — the kind that can take weeks to restore — before operators implemented rolling outages that lasted, for some customers, four to five days. Somewhere between 250 and 700 people died, depending on how you count, most of them from hypothermia or carbon monoxide poisoning from improvised heating.

The Texas grid failure was not a renewable energy story, though it was reported as one in some quarters. It was a weatherization story — a failure to winterize equipment to match conditions that were foreseeable, just infrequent. It was also an isolation story. ERCOT operates as its own island, with minimal interconnection to

neighboring grids, partly for regulatory reasons (avoiding federal jurisdiction) and partly because the interconnections were simply never built. During the crisis, there was clean power available to the east in the Southeast Electric Reliability Council region and to the west in the Western Interconnection. Texas could not import it at meaningful scale. The interconnections that didn't exist because they were expensive and politically inconvenient were, in February 2021, existentially valuable.

The lesson is not "build more renewables" or "build more gas plants." The lesson is: build a more connected, more resilient grid with enough redundancy to handle rare but predictable extremes. That is a lesson the United States has not fully acted on.

The Binding Constraint

Vaclav Smil has spent decades trying to get energy analysts to take scale seriously — to understand not just that a transition is desirable but what it physically requires in terms of materials, infrastructure, and time. He is generally pessimistic about the pace of transitions. The grid bottleneck is one of the places where his caution feels most warranted.

Because here is where the arithmetic stands. The IEA calculates that reaching net-zero emissions globally by 2050 requires tripling renewable power capacity by 2030. That build rate is achievable — manufacturing capacity for solar panels and wind turbines has grown dramatically. But the IEA's same report notes that grid investments are "lagging critically" behind generation additions. In 2022, about 950 gigawatts of solar and wind capacity was sitting in interconnection queues in the U.S. and Europe, waiting for grid studies to complete and upgrade costs to be negotiated. That is roughly the combined installed capacity of solar and wind that currently exists in both regions. The queue is as large as the existing fleet.

The generation has become, genuinely, the easier part. The physics of silicon and steel, the economics of manufacturing at scale, the engineering of offshore foundations — these problems have been solved or are being solved at impressive speed. A utility-scale solar farm that would have cost $300 per megawatt-hour in 2010 cost roughly $30 per megawatt-hour by 2023 in good solar locations. That 90 percent cost reduction happened in thirteen years. Nothing like it has happened in transmission costs, grid modernization costs, or the pace of permitting reform.

This asymmetry is the central grid problem. You can build cheap generation faster than you can build the infrastructure to connect it, operate it safely, and deliver its output to the people who need it. The generation curve has been heroic. The grid curve has been, in most wealthy countries, nearly flat.

What would it take to bend the grid curve?

Some of it is hardware: more HVDC interconnections, more grid-forming inverters, more flexible AC devices, better sensors and dynamic ratings on existing lines. Some of it is software: overhauled market systems that price all grid services, not just energy; advanced distribution management platforms; better forecasting tools that can handle stochastic renewable output. Some of it is workforce: training tens of thousands of power systems engineers, electricians, and grid operators who don't currently exist.

But the largest constraint is institutional and political. Permitting reform that actually shortens timelines without gutting environmental review. Cost-allocation rules that allow interstate transmission projects to be built and paid for by the regions that benefit. Federal authority, or at least federal coordination, for the kind of long-range interregional planning that no regional transmission organization can do alone. These are not engineering problems. They are governance problems. And governance problems, in the current political environment of most wealthy democracies, are the hardest problems of all.

The generation miracle of the past decade — cheap solar, cheap wind, plummeting battery costs — has been primarily a technology story, shaped by manufacturing economics and policy incentives that worked with the grain of markets. The grid transformation required to make that cheap generation useful is a different kind of story. It is shaped by regulatory inertia, institutional fragmentation, political economy, and the sheer physical difficulty of building infrastructure that crosses property lines, state lines, and national interests.

The wires are harder than the generation. Not because the physics are more difficult — they are not — but because the wires require everything that cheap generation doesn't: long time horizons, shared financing, coordinated planning across competing jurisdictions, and political will sustained over decades. Those are scarce resources in any era. In this one, they are the bottleneck that the energy transition cannot afford but has not yet found a way around.

CHAPTER X

The Hard Stuff: Steel, Cement, Aviation, and Shipping

There is a version of the energy transition story that goes like this: we build enough solar panels, enough wind turbines, connect them to a smart grid, electrify everything, and somewhere around mid-century the carbon problem is essentially solved. It is a tidy story. It is also incomplete in ways that matter enormously.

The grid story — electrons replacing combustion — works beautifully for lighting, heating buildings, and charging the car in your driveway. It gets messier for heavy trucks. And for a significant slice of what the modern economy actually does every day, it runs into walls that are not political or financial but physical. Chemistry. Thermodynamics. The stubborn fact that some industrial processes require heat so intense, or energy sources so concentrated, that the phrase "just electrify it" lands somewhere between premature and naive.

Roughly 30 percent of global carbon dioxide emissions come from this category. The IEA pegged hard-to-abate industrial sectors and heavy transport at around 10 gigatonnes analysis of net-zero pathways. That is not a rounding error. It is approximately equal to the total emissions of the United States and European Union combined. Decarbonising the power sector while leaving this block untouched does not get us to any meaningful climate target.

This chapter is about the hard stuff. Steel. Cement. Aviation. Shipping. Four sectors that are responsible for most of that 30 percent, that underpin nearly every other activity in modern civilization, and that have no clean, cheap, shovel-ready replacements waiting in the wings. The solutions exist — mostly — but they are expensive, they are slow, and they have received a fraction of the policy attention lavished on solar farms and electric cars. Understanding why requires getting into the specific chemistry and logistics of each sector, because the barriers are different in each case, and conflating them produces only confusion.

Steel: When Carbon Is the Point

Vaclav Smil opens How the World Really Works with a blunt inventory of the four materials that constitute modern civilization: cement, steel, plastics, and ammonia. His argument is not rhetorical. In 2022, humanity produced roughly 1.9 billion tonnes of crude steel. That steel is in every building, every ship, every wind turbine, every EV chassis. There is no substitute at scale. And roughly 70 percent of it is made through a process that requires carbon not merely as fuel but as chemistry.

Here is the problem. A conventional blast furnace takes iron ore — mostly ir it to metallic iron by reacting it with coke, a form of nearly pure carbon derived from metallurgical coal. The

carbon does two jobs simultaneously. It burns to generate the roughly 1,600 degrees Celsius needed inside the furnace. And it acts as a reducing agent, stripping the oxygen atoms from the iron oxide molecules. The product of that second reaction is CO,. Not as a byproduct of inefficiency. A the chemistry.

You cannot simply swap in a clean heat source and call it done. Even if you powered the blast furnace entirely with renewable electricity, you still need something to chemically reduce the iron ore. For a century and a half, that something has been carbon, because carbon is cheap, abundant, and extraordinarily good at the job. Eliminating it means finding a different reducing agent.

There are two credible paths. The first is the electric arc furnace, or EAF, which skips the reduction problem entirely by starting with scrap steel rather than ore. You load a furnace with recycled steel, run an electric arc at around 3,000 degrees Celsius, melt everything down, and pour out new steel. No coke required. The electricity, if sourced from renewables, can make genuinely green steel. About 30 percent of global steel production already uses this route — including most of the steel made in the United States, where scrap is abundant.

The problem is the scrap. Steel is extraordinarily durable, and there is only a finite amount of it cycling through the global economy at any given time. The world has been accumulating steel infrastructure since the 19th century, but that stock is growing, not shrinking, because we keep building more. The IEA's analysis suggests that even under optimistic assumptions about collection rates, EAF using scrap can cover at most 45 to 50 percent of global demand by 2050. The rest still has to come from virgin iron ore. And that requires solving the reduction problem.

The second path is hydrogen-based direct reduced iron, or H,-DRI. Instead o into a shaft furnace. Hydrogen, like carbon, strips oxygen from iron oxide — but the byproduct is water vapour, not CO,. The resulting iron, called sponge iron or DRI, can then be f processing. If the hydrogen is made by electrolyzing water using renewable electricity — so-called green hydrogen — the entire chain is essentially carbon-free.

This is not speculative. It exists. In northern Sweden, a consortium called HYBRIT — jointly owned by steelmaker SSAB, mining company LKAB, and utility Vattenfall — has been producing fossil-free steel from a pilot plant since 2020. In August 2021, SSAB delivered the first batch to Volvo for use in prototype vehicles. The engineering works. The question, as always with hard-to-abate sectors, is the cost.

Green hydrogen currently runs between four and eight dollars per kilogram in most markets, compared to roughly one to two dollars for hydrogen made from natural gas. That cost gap flows directly into steel prices. Industry analysts estimate that green steel currently costs 20 to 30 percent more than conventional blast-furnace steel, with some estimates higher depending on local electricity prices. For a car, that premium is modest — perhaps a few hundred dollars on a vehicle that already costs thirty thousand. For a buyer of a million tonnes of structural steel for a megaproject, the arithmetic is less comfortable.

The situation is not hopeless. Green hydrogen costs are falling as electrolyzers scale up, following something resembling the same learning curve that solar panels demonstrated over four decades. The IEA projects that green hydrogen could reach two dollars per kilogram in regions with excellent renewable resources by the mid-2030s, which would narrow the green steel premium substantially. Several major steelmakers — ArcelorMittal, ThyssenKrupp, H2 Green Steel — have announced large-scale DRI-EAF projects for the late 2020s. But "announced" and "operating" are different words in heavy industry. The lead times for these projects are seven to ten years. The capital requirements are in the billions. And the whole chain depends on having cheap, abundant green electricity in the right places.

The political economy is awkward, too. Steel industries are regionally concentrated, heavily unionized, and politically sensitive in ways that solar farms are not. Governments in Europe have responded with carbon border adjustment mechanisms designed to level the playing field between green steel and cheaper imports from countries without carbon pricing. Whether that mechanism survives trade disputes and political cycles is a separate question.

None of this is to say steel decarbonisation is impossible. It is plausible, probably achievable by 2050 for a significant fraction of global production. But it requires sustained policy support, substantial capital investment, and cheap renewable electricity delivered to locations — often remote — where ore and industrial infrastructure already exist. It will not happen by accident.

Cement: The CO, You Cannot Burn Away

If steel's problem is that carbon is the reducing agent, cement's problem is stranger and in some ways more intractable. Part of cement's CO, comes from burning fuel to heat the kiln. T addressable — swap the fuel. But about 60 percent of cement's carbon emissions come from a chemical reaction that happens regardless of what you burn.

The process is calcination. To make Portland cement, you heat limestone — c around 1,450 degrees Celsius. At that temperature, the calcium carbonate dec The CaO, called lime or clinker, is the active ingredient in cement. The CO, atmosphere. It is not waste heat or combustion exhaust. It is a product of the chemistry of making the material.

Global cement production in 2023 ran to approximately 4.1 billion tonnes. That makes cement manufacturing responsible for somewhere between 7 and 8 percent of global CO, emissions shipping combined. China alone produces about 55 percent of the world's cement, mostly to support an infrastructure buildout that has consumed more concrete in some recent decades than the United States used in the entire 20th century. That is not an exaggeration. Historians of Chinese infrastructure point to the scale as genuinely without precedent.

What do you do about the calcination reaction? A few things, none of them fully satisfying.

The most practical near-term option is supplementary cementitious materials — replacing a portion of clinker with industrial byproducts like fly ash (from coal combustion) or blast furnace slag (from steelmaking) that have cementitious properties. Some cement products already replace 30 to 40 percent of their clinker this way. But fly ash supply depends on coal power plants running, and as coal phases out, fly ash supply declines. Slag is similarly constrained by steel production volumes. These are helpful but partial measures.

Geopolymer cements, which use entirely different chemistry — typically activating aluminosilicate materials with alkaline solutions — can avoid calcination entirely. They have been demonstrated in construction projects. The Zeobond E-Crete product was used in a precast element bridge in Brisbane in 2008. But geopolymers have significant limitations around workability, curing conditions, and performance specifications that vary with local raw material chemistry. They have not scaled.

Carbon capture is the third option, and for cement it is arguably more compelling than in most other industries, because the CO_2 released at a cement kiln is highly concentrated dilute flue gases at a power plant. Heidelberg Materials, one of the largest cement producers in the world, is constructing a full carbon capture facility at its Brevik plant in Norway, targeting the removal of 400,000 tonnes of CO_2 per year starting around 2024 or 2025. It is the first such proj captured CO_2 will be liquefied and shipped to subsea geological storage unde Norway's Northern Lights infrastructure project.

Whether this approach can be replicated across thousands of cement plants worldwide is a different matter. It requires proximity to CO_2 transport and storage infrastructure that barely ex and a few other locations. The costs are substantial — estimates for carbon capture at cement plants range from $60 to $120 per tonne of CO_2 avoided, and that adds meaningfully to th traded in tight-margin markets. Building construction is already politically fraught around cost; adding a climate surcharge to concrete requires either regulation or a carbon price that most jurisdictions don't have.

There is also the straightforward option of using less cement. Buildings can be designed to use less concrete — mass timber construction has revived significantly, with cross-laminated timber (CLT) products demonstrating structural performance competitive with concrete for many mid-rise applications. But the world is urbanising fast, mostly in places that will build in concrete regardless, because timber at scale requires managed forests that many of those countries don't have. Efficiency helps at the margins.

The honest assessment of cement is that full decarbonisation by 2050 is possible in technical terms and deeply unlikely in practical ones without a combination of carbon pricing, regulated carbon capture mandates, and substantial public financing for CO_2 storage infrastructure. The IEA's ne emissions roughly halving by 2050 through efficiency, material substitution, and CCUS — but achieving even that half-reduction requires policy frameworks that do not currently exist in most major producing countries.

Aviation: The Tyranny of Energy Density

Let's talk numbers for a moment, because they matter here more than almost anywhere else in the energy transition.

A kilogram of jet fuel — kerosene, essentially, specifically Jet A-1 — contains about 43 megajoules of energy. A kilogram of lithium-ion battery at the current state of the art stores somewhere between 0.7 and 0.9 megajoules usable energy at the cell level, and less than that at the pack level when you account for casings, thermal management, and electronics. The gap is roughly 50 to 1. In a car, that gap is manageable. You carry a 500-kilogram battery pack, accept a range of 300 miles, charge overnight. The car is heavier and the physics are less elegant but the trade works.

In a commercial aircraft, the trade does not work. Not even close.

An Airbus A320 flying from London to New York — a route that doesn't actually exist, that distance requiring something larger, but use it conceptually — would need to carry roughly 20 tonnes of jet fuel. To match that energy content with lithium-ion batteries at current technology, you would need approximately 1,000 tonnes of batteries. The plane cannot carry 1,000 tonnes of anything. It weighs 78 tonnes empty. The whole point of flight is that the energy source has to be light relative to the energy it contains, which is precisely what jet fuel is and what batteries, as of the mid-2020s, are emphatically not.

This is not a pessimistic projection about batteries failing to improve. It is a statement about magnitudes. Even if battery energy density quadrupled — an achievement that would be genuinely revolutionary for the industry — the ratio would still leave batteries approximately 12 to 1 worse than jet fuel on a per-kilogram basis. The physics of electrochemistry place hard limits on how far this number can move. The theoretical maximum energy density of lithium-ion chemistry is somewhere around 3 MJ/kg. Jet fuel at 43 MJ/kg is in a different universe.

For short-haul regional aviation — flights under about 500 kilometres, carrying fewer than 30 passengers — battery-electric aircraft are plausible. The Swedish company Heart Aerospace is developing the ES-30, a 30-passenger regional aircraft, with anticipated entry into service around 2028. Cape Air, the regional carrier, has ordered the Eviation Alice, an all-electric nine-seater. These are real aircraft, with real orders, from real airlines. For island-hopping in Norway, or inter-island routes in Hawaii, or short regional hops where the alternative is a small turboprop anyway, electric aviation may arrive this decade.

But regional aviation handles a tiny fraction of global aviation emissions. The emissions problem is concentrated in long-haul flights. About 80 percent of aviation's CO, comes kilometres. The IEA estimated aviation at roughly 800 million tonnes of CO, of global energy-related emissions — and that number understates aviation's actual climate impact because aircraft contrails and high-altitude NOx emissions warm the climate through though quantifying those effects precisely remains an active area of research.

The leading candidate for decarbonising long-haul aviation is Sustainable Aviation Fuel — SAF. The term covers a range of production pathways, but the basic concept is a liquid hydrocarbon fuel that burns in existing jet engines, works in existing infrastructure, and is produced from low-carbon feedstocks instead of

fossil crude oil. Current SAF is made mostly from waste fats and oils — used cooking oil, tallow from slaughterhouses, municipal solid waste. The carbon accounting works because you are cycling carbon that was recently in the biosphere rather than releasing carbon that was buried underground for 100 million years.

SAF can reduce lifecycle emissions by 50 to 80 percent compared to conventional jet fuel. It is a drop-in fuel — no modification to aircraft or airports required. It is certified for blending up to 50 percent with conventional fuel in existing engines, with work underway toward 100 percent SAF certification.

The catch is supply and price. Global SAF production in 2023 amounted to roughly 300,000 tonnes. Global jet fuel consumption runs to approximately 300 million tonnes per year. SAF is less than one-tenth of one percent of the market. And the feedstocks for first-generation SAF — waste oils and fats — are not remotely sufficient to supply the aviation sector at scale. The world does not generate enough used cooking oil.

Scaling up requires either synthetic SAF, made from green hydrogen and capt power-to-liquid), or advanced biofuels using agricultural residues and dedicated energy crops. Both pathways work technically. Both are currently two to five times more expensive than conventional jet fuel. Power-to-liquid, which can theoretically be produced anywhere with renewab is the longer-term answer — but it is currently three to four times the cost of conventional fuel, and the electrolyzer capacity needed to make a meaningful dent in aviation demand is staggering.

Hydrogen is occasionally promoted as an aviation fuel. Airbus has announced the ZEROe concept, targeting hydrogen aircraft entry into service by 2035. Hydrogen burns cleanly, producing water vapour rather than CO_2. But liquid hydrogen has roughly four times the energy density of batter one-quarter the volumetric density of jet fuel — very bad. Aircraft are designed around volume as much as weight. A hydrogen-fuelled aircraft would need a fundamentally different fuselage design to accommodate cryogenic fuel tanks, stored at minus 253 degrees Celsius. That means not modifying existing aircraft but designing entirely new ones, certifying them, training maintenance personnel, rebuilding airport infrastructure. The timeline to commercial scale, even on Airbus's optimistic framing, is the 2040s at the earliest for meaningful market penetration.

Aviation's most likely 2050 trajectory is a combination: continued electrification of short-haul routes as battery technology improves, growing SAF mandates driving up the blend percentage from the current near-zero toward perhaps 10 to 20 percent by 2035, and a long tail of residual fossil-fuel use in long-haul routes where alternatives remain too expensive. The European Union has mandated SAF blending targets — 6 percent by 2030, rising to 70 percent by 2050. Whether supply chains can actually meet those targets is a question nobody is answering with confidence.

Shipping: Slow and Enormous, in Every Sense

Global shipping carries roughly 80 percent of world trade by volume. That is not a figure that changes much no matter which economist or trade theorist you ask. The modern global economy runs on ships. A single large container ship, a vessel like the class that serves major trans-Pacific routes, burns on the order of 200 to

300 tonnes of heavy fuel oil per day at sea. A voyage from Shanghai to Rotterdam takes about 30 days. The arithmetic of energy consumption at that scale is sobering.

Shipping produces approximately 800 million to one billion tonnes of CO, pe
global emissions. The International Maritime Organization, the UN body that regulates shipping, has revised its decarbonisation targets upward under political pressure — its 2023 revised strategy calls for net-zero emissions by "around 2050" and at least 40 percent reduction in carbon intensity by 2030. Those targets are ambitious relative to current trajectories.

The challenge is different from aviation in its specifics. Ships are not weight-constrained the way aircraft are — a large bulk carrier or container ship can carry enormous quantities of fuel. The problem is not energy density in the same acute sense. Ships travel slowly, over long distances, and the economic logic has always pointed toward the cheapest possible fuel burned in the largest possible engines. Heavy fuel oil, the viscous residue of petroleum refining, has been the answer for a century because it is cheap. It is also extremely dirty — high in sulfur, producing particulate emissions that have genuine local health effects. New IMO sulfur regulations that took effect in 2020 have driven a shift toward very low sulfur fuel oil and LNG, but neither eliminates CO,.

The replacement fuels under serious development for shipping are ammonia, methanol, and hydrogen. Each has a different profile.

Methanol is the nearest-term option. It is liquid at room temperature, which means existing fuel handling infrastructure can be modified rather than rebuilt from scratch. Maersk — the world's second-largest container shipping company — has ordered a series of methanol-capable vessels and took delivery of the first, the Laura Maersk, in 2023. Methanol produces CO, when burned, so it
renewable feedstocks — bio-methanol from agricultural waste, or e-methanol from green hydrogen and captured CO,. Both exist. Both are currently several times more expensive th

Ammonia is a stronger longer-term candidate because burning it produces nitrogen and water rather than CO,. It is already produced and shipped globally in enormous quantities as a
180 million tonnes per year — so there is existing industrial infrastructure. Green ammonia, made by combining green hydrogen with atmospheric nitrogen through the Haber-Bosch process, could in principle provide a genuinely zero-carbon fuel. MAN Energy Solutions, the dominant maker of large marine engines, has developed ammonia-capable two-stroke engines. Several shipowners have placed orders for ammonia-ready vessels.

The complications are real. Ammonia is toxic — it kills at concentrations that would be routine in a fuel leak scenario on a ship. The handling protocols required are substantially stricter than for conventional fuels, and training crews to manage ammonia safely across a global fleet of tens of thousands of vessels is not a trivial undertaking. Ammonia also has a lower energy density than heavy fuel oil, which means larger fuel tanks and reduced cargo capacity. And combustion of ammonia produces nitrous oxide –
gas, if combustion is incomplete. Engine manufacturers are working on this, but it is an unsolved engineering

problem rather than a solved one.

Hydrogen faces the same cryogenic storage problem as in aviation, compressed by the fact that ships often need range of thousands of kilometres. Liquid hydrogen is plausible for ferries and short coastal routes. It is not a near-term solution for a container ship making a 20,000-kilometre voyage.

The economic structure of shipping creates its own barriers. Ships have operational lives of 20 to 25 years. A vessel ordered today will still be in service in 2045. Shipowners making capital decisions now are doing so under genuine uncertainty about which fuel will be cheapest, most available, and regulatory-compliant in 2040. The rational response to that uncertainty, from an individual business perspective, is to hedge — to order "ammonia-ready" or "methanol-ready" vessels that can run on conventional fuel in the interim and be converted later. That hedge is not decarbonisation. It is optionality.

The IMO's carbon intensity index, which requires ships to calculate and reduce their carbon intensity ratings, has pushed some behavioral changes — slower sailing speeds, which reduce fuel consumption by roughly the cube of speed reduction, the same physics that makes wind turbine height valuable. A ship sailing at 18 knots instead of 23 knots uses dramatically less fuel. That is a real near-term lever. It is also a workaround, not a solution.

Port infrastructure matters in ways that rarely make the headlines. A ship running on ammonia needs ammonia bunkering facilities at every port it calls. A ship running on hydrogen needs liquid hydrogen storage. Ports are independently owned and operated, often by port authorities with their own investment cycles and political constraints. The coordination problem of simultaneously upgrading dozens of major global ports to handle new fuels while shipowners simultaneously adopt vessels using those fuels is a classic chicken-and-egg dynamic that markets alone have never resolved.

The Through-Line

Steel, cement, aviation, shipping. Four sectors. Four distinct technical problems. Wildly different scales of difficulty and different timelines to viable solutions.

What they share is this: none of them decarbonise by market forces alone on any timescale compatible with climate targets. Solar panels and wind turbines got cheap through a combination of policy support, learning curves, and falling materials costs, in a market where the product — electricity — is fungible and buyers will simply switch to whatever is cheapest. Green steel, low-carbon cement, sustainable aviation fuel, and zero-carbon ammonia are not yet competitive with their fossil-fuel equivalents. They may become so over time, for some applications, in some locations. But the gap is large enough, and the capital cycles long enough, that waiting for cost competitiveness to arrive organically means locking in decades of additional emissions.

The policy tools that could accelerate these transitions are not exotic. Carbon pricing high enough to close the cost gap. Public procurement — governments buying green steel for bridges and buildings, SAF for military aircraft, low-carbon cement for infrastructure projects — to create demand at scale and drive down costs

through volume. Mandates like the EU's SAF blending requirements, which guarantee a market for producers willing to build supply chains. Investment in shared infrastructure like CO, that no single company will build because they cannot capture enough of the value.

These are not novel ideas. The IEA's net-zero analysis, various national industrial strategies, and a shelf of academic papers have all pointed in roughly the same direction. What has been missing is consistent follow-through. The United States passed the Inflation Reduction Act in 2022, which included meaningful incentives for clean hydrogen production that will help some of these sectors. The European Union has the Green Deal industrial plan. But global coverage is thin, and in most major emitting countries — India, China, much of Southeast Asia — the industrial decarbonisation policy landscape is sparse.

Smil's instinct, honed across decades of studying energy systems, is always to insist on scale. He would tell you — and has, in various formulations — that replacing the physical flows of carbon and fossil energy that underpin modern material civilization requires physical solutions at a scale that societies have not yet fully reckoned with. Green hydrogen for steel and ammonia production alone would require electrolyzer capacity that dwarfs anything currently installed. The biomass and renewable electricity needed to produce SAF at the percentages that regulators are mandating would be substantial fractions of current total renewable energy output.

That is not an argument for inaction. It is an argument for clear-eyed assessment of what the task actually requires, which is consistently larger than the headline targets suggest and consistently larger than current investment levels reflect.

The good news — and there is some — is that the solutions are not waiting on scientific breakthroughs. Hydrogen DRI for steel is a working technology. Carbon capture at cement plants is being demonstrated commercially. SAF is in jet engines today. Ammonia-capable ship engines exist. The question is not whether these things can be done but whether they will be done fast enough and at large enough scale to matter. And that question is answered not in laboratories but in parliaments and procurement offices and financial institutions that have spent most of the past two decades focused on the easier, sexier story of solar panels and electric cars.

The easier story is real, and it is important, and the chapters before this one have told it honestly. But the hard stuff does not disappear because it is less photogenic. Thirty percent of global emissions doesn't vanish because politicians prefer to talk about something else. The next decade's version of the energy transition, if it is serious, will be defined less by how many gigawatts of solar gets installed and more by whether the first commercial-scale green steel plant breaks ground, whether carbon capture at cement kilns moves from pilot to standard practice, whether the price of sustainable aviation fuel starts falling the way solar module prices did in the 2010s.

There are no certainties here. Only trajectories. And the trajectories for the hard stuff are, as of the mid-2020s, pointing slowly in the right direction — and not nearly fast enough.

CHAPTER XI

The Honest Geography of Energy Poverty

Picture a woman named Amara. She is thirty-four years old, lives in a village in northern Nigeria about sixty kilometres from the nearest reliable grid connection, and begins her day at 4:30 in the morning. She starts a fire. She uses three-stone cooking — two rocks and a clay mound supporting a pot — a method that is essentially unchanged from the Neolithic. The smoke fills the interior of her home. She breathes it for two hours while preparing food for her children. She has done this every morning since she was old enough to manage the kindling.

Amara does not appear in most Western discussions about the energy transition. Those discussions are about whether offshore wind should be permitted sixty miles from Cape Cod, whether sodium-ion batteries will displace lithium, whether the grid in Texas can handle another cold snap. They are not unimportant questions. But they are questions about how to decarbonise an energy system that already exists and already powers refrigerators, hospitals, factories, smartphones. They are not questions about the roughly 750 million people who, as of the IEA's most recent tracking, have no electricity access at all — or the 2.3 billion who cook on open fires or primitive biomass stoves every single day.

This chapter is about those people. Their situation is not a footnote to the energy transition. It is a central variable in the transition's arithmetic. Leave it out and the math is simply wrong.

A Map Drawn in Darkness

The IEA's World Energy Outlook 2023 puts the global electricity access deficit at approximately 685 million people, down from around 860 million a decade earlier. Progress has been real. India's Saubhagya scheme, launched in 2017, connected roughly 26 million households to the grid within two years — the largest household electrification effort in human history, executed with a combination of grid extension and small distributed systems. Bangladesh deployed more than 6 million solar home systems between 2003 and 2018, reaching rural households that the conventional grid would not have reached economically for another generation. China, which in 1990 had well over 100 million people without electricity, achieved near-universal access by the mid-2010s, largely through brute-force grid extension financed by its coal-fuelled economic expansion.

Yet the headline number's steady decline can obscure what is actually happening geographically. Nearly all the remaining unelectrified population lives in sub-Saharan Africa — approximately 570 million people, or more than 80 percent of the global deficit. And here the picture is not improving fast enough to track population growth. The region's population is projected to nearly double by 2050. New connections are being added, but the absolute number of people without electricity has barely fallen in Africa because so many new

people are being born into communities that lack it. The IEA estimated in its 2022 Africa Energy Outlook that to achieve universal electricity access by 2030, the region would need to connect roughly 90 million people per year — roughly triple the pace achieved between 2015 and 2022.

The cooking fuel picture is starker and receives far less attention. The 2.3 billion figure comes from the IEA and WHO tracking systems; it represents people using solid biomass (wood, charcoal, agricultural residue, animal dung), kerosene, or coal on open fires or rudimentary stoves with minimal combustion efficiency and no flue to carry the smoke away. The geographic concentration is similar: most of sub-Saharan Africa, large parts of South and Southeast Asia, pockets of Central America and the Pacific Islands. This is not primarily a story about places that recently lost access. It is a story about places that never had it.

Vaclav Smil, whose quantitative rigour on energy questions is without peer among living researchers, frames the issue in terms of primary energy supply per capita. The average American consumes roughly 290 gigajoules of primary energy per year. The average European, benefiting from denser cities and more efficient appliances, manages the modern service level on about 140 gigajoules. The average resident of sub-Saharan Africa (South Africa excluded, since its grid is a significant outlier from the regional norm) consumes somewhere around 20 gigajoules per year — and a large fraction of that is biomass burned at very low efficiency for cooking and heating. The ratio between American and average sub-Saharan African energy consumption is roughly 15 to 1. Not a modest gap. A civilisational chasm.

The Invisible Epidemic

Here is a number that does not receive anything like the media coverage it deserves. According to the World Health Organization, household air pollution from cooking fires and primitive stoves killed approximately 3.2 million people in 2020. Malaria, for all the philanthropic attention it rightly commands, killed around 620,000 people in 2022 — a record high driven by pandemic-era disruptions to prevention programmes, and still a fraction of the cooking fire toll. Indoor air pollution from solid fuel combustion kills more people every year than malaria, tuberculosis, and HIV/AIDS combined.

The mechanism is not complicated. Burning wood, dung, and charcoal in enclosed or poorly ventilated spaces produces carbon monoxide, nitrogen dioxide, particulate matter (PM2.5), and polyaromatic hydrocarbons in concentrations that can exceed outdoor air quality standards by factors of ten to one hundred. Women and young children suffer the most, because they spend the most time near the cooking fire. The health effects are a catalogue of chronic damage: reduced lung function, chronic obstructive pulmonary disease, low birth weight, increased susceptibility to pneumonia in children under five. Pneumonia alone kills about 800,000 children under five each year globally; the WHO estimates that between 45 and 50 percent of those deaths are attributable to indoor air pollution from solid fuels.

Smil notes, with characteristic bluntness, that this situation "should be as unacceptable as any other mass health catastrophe." It rarely features in the energy debates of wealthy nations with the weight it deserves. When it does appear, it tends to be framed as a solvable technical problem — improved cookstoves, biogas

digesters, solar cookers — which it is, in principle. The Global Alliance for Clean Cookstoves (now the Clean Cooking Alliance) spent years promoting improved biomass stoves that burned more efficiently and reduced indoor smoke. The results were useful but modest. A stove that cuts smoke emissions by 50 percent is better than nothing. It still produces dangerous concentrations in a typical unventilated kitchen. What actually solves the problem is LPG, piped natural gas, or electric induction cooking. And that requires infrastructure, supply chains, and purchasing power that much of the affected population currently lacks.

The Leapfrog Story: Sometimes Right, Sometimes Convenient

The dominant narrative about energy access in the global development community runs something like this: sub-Saharan Africa and other energy-poor regions do not need to repeat the West's fossil-fuel-intensive industrialisation. They can skip directly to distributed solar, mobile payments, and modern LED lighting. They can leapfrog the century of coal and oil that preceded electrification in Europe and North America.

This narrative is not entirely wrong. In fact, for one specific application, it is demonstrably correct. Mobile money and pay-as-you-go solar — the combination pioneered by M-KOPA in Kenya beginning around 2011 — has genuinely brought basic electric light and phone charging to millions of households that would have waited decades for the conventional grid. The model works as follows: a household pays a small deposit for a solar home system (typically a 6- to 20-watt panel, a small battery, a few LED lights, and a phone charging port), then makes daily micro-payments via M-PESA, Kenya's mobile money platform. If payments lapse, the system locks remotely. By 2023, M-KOPA reported serving more than 3 million customers across Kenya, Uganda, Ghana, Nigeria, and South Africa. Similar operators — SunKing, Bboxx, d.light — together reached tens of millions of households across East and West Africa during the 2010s.

This is not trivial. Electric light after dark means children can study. It means women working from home can extend their productive hours. Phone charging means connectivity, market price information, mobile banking. The health benefits of eliminating kerosene lamps — which also produce harmful particulates — are real. Hannah Ritchie's analysis of energy access data shows consistent evidence that even this minimal tier of electricity access correlates with measurable improvements in education outcomes and household income over five-year horizons.

But here is where the leapfrog narrative starts to require some interrogation. A 10-watt solar home system and a phone-charging port is not a productive economy. It cannot power a sewing machine, a welding torch, a grain mill, a water pump, a refrigerator for a small health clinic, a lathe in a workshop. The IEA uses a tiered energy access framework — Tier 0 through Tier 5 — where Tier 1 is basic lighting and phone charging and Tier 5 is the kind of supply that powers a modern household appliance suite. Most of the leapfrog solar success stories have delivered Tier 1 or Tier 2 access. The economic transformation that comes with industrialisation requires Tier 4 and Tier 5, which means reliable grid power at sufficient capacity to run productive machinery.

No country has industrialised on distributed rooftop solar. Not one. China industrialised on coal. South Korea on coal and oil. Taiwan on coal and nuclear. Germany on coal. The United States on coal, oil, and hydropower. The industrial foundations of every wealthy nation were built on energy sources that are dense, continuous, and capable of powering factories twenty-four hours a day. You cannot run a cement plant on solar panels and batteries — not today, not at any cost that would permit the cement to be affordable. You cannot smelt steel with intermittent power. The economic activities that create the jobs, tax revenues, and supply chains that enable countries to accumulate wealth and then invest in cleaner technologies require a quality and quantity of energy supply that the leapfrog vision, honestly examined, does not currently provide.

None of this means distributed solar is worthless — it clearly is not — but it does mean that the leapfrog narrative sometimes functions as a way for wealthy countries to argue that low-income countries should develop differently from the way wealthy countries actually did develop. And that asymmetry is worth naming plainly.

The Bill That Was Never Sent

In 2021, at COP26 in Glasgow, John Kerry, then the United States Special Presidential Envoy for Climate, stated bluntly that Africa should not build new coal plants. The continent's energy ministers were not uniformly enthusiastic about this advice. Amani Abou-Zeid, the African Union Commissioner for Infrastructure and Energy, pointed out, not for the first time, that Africa produces roughly 3 to 4 percent of global cumulative CO_2 emissions while containing 17 percent of the world's p with 4 percent of the world's population, has produced roughly 25 percent of emissions since the industrial revolution began. The United Kingdom, which invented the large-scale coal economy in the eighteenth century, has roughly the same cumulative emissions profile per capita as the US.

This is not a rhetorical point. It is a fairness point with quantitative teeth. The carbon that has already warmed the atmosphere was overwhelmingly emitted by wealthy countries during the period when they were using fossil energy to get wealthy. The countries now being told that coal is off-limits are largely the ones that never got the chance to use it for development. The argument that they should now bear the cost of technologies that are cleaner but more expensive — and less reliable, at current penetration levels, without the storage and grid infrastructure that wealthy countries are still struggling to build — is not obviously just.

Smil addresses this directly in How the World Really Works, published in 2022. He argues that the energy transition in wealthy countries is a real and necessary undertaking, but that anyone who thinks global emissions can be decisively addressed without addressing the energy aspirations of sub-Saharan Africa, South Asia, and Southeast Asia is doing arithmetic with half the variables. The populations of these regions are young, growing, and beginning to demand the material conditions — reliable electricity, clean cooking, motorised transport, industrial employment — that wealthy-country populations have considered baseline for generations. The question is not whether their energy consumption will grow. It will. The question is what it will grow on, and whether the international community will make clean alternatives genuinely affordable and accessible rather than simply moral-instructing from a position of accumulated advantage.

What the Numbers Actually Require

Let's do the arithmetic, because the scale is sobering.

Sub-Saharan Africa's population in 2023 was approximately 1.4 billion. The IEA's Africa Energy Outlook projects it reaching roughly 2.5 billion by 2050. If those 2.5 billion people were to reach even half the per-capita electricity consumption of present-day Portugal — a relatively modest European country — they would require approximately 2,800 terawatt-hours of electricity annually. Total electricity generation in sub-Saharan Africa in 2022 was roughly 850 terawatt-hours, most of it concentrated in South Africa, Nigeria, and Ethiopia. The gap between current supply and even a modest development scenario is not tens of percent. It is multiples.

Now consider the global carbon budget. Under the IPCC's pathways for limiting warming to 1.5°C, the remaining global carbon budget from 2023 was approximately 380 billion ton would be consumed in about nine years at then-current global emission rates. Any scenario in which sub-Saharan Africa substantially develops its energy system using fossil fuels — coal plants, gas peakers, diesel generators — will eat into that budget. This is the genuine tension at the heart of the energy poverty problem. It is not resolvable by pretending one side of it does not exist.

What honest analysis suggests is that there are two separate questions that get conflated. The first is what low-income countries are entitled to do — a question of equity and historical responsibility on which reasonable people can disagree but which has a clear moral logic on the side of development. The second is what would actually serve those countries' long-term interests given current technology costs — a question of economics and pragmatics that sometimes, but not always, points toward renewables.

On that second question, the picture has changed substantially since 2015. The IEA's Africa Energy Outlook 2022 found that in most African contexts, new utility-scale solar is now cheaper than new coal on a levelised cost basis — not because of moral persuasion, but because the cost curves described in Chapter 6 have done their work. In Kenya, the Lake Turkana Wind Power project, completed in 2017 at 310 megawatts, produces electricity at costs competitive with the country's existing coal plant at Lamu — which, after years of controversy, was ultimately cancelled. In Morocco, the Noor Ouarzazate concentrated solar complex provides electricity to 1.1 million homes and helped Morocco reach a target of 42 percent renewable electricity by 2021. These are not peripheral examples. They are evidence that for grid-scale power generation in sun- and wind-rich regions, the economics have genuinely shifted.

But grid-scale solar still requires a functioning transmission grid to distribute that power to households and industries. Sub-Saharan Africa's transmission and distribution infrastructure is, by global standards, severely underdeveloped and underfunded. Nigeria — the continent's most populous country with over 220 million people — had total installed generation capacity of roughly 13 gigawatts in 2023, of which typically less than 5 gigawatts was operational at any given time due to gas supply constraints, transmission losses, and grid instability. To put that in perspective: the state of New Jersey, with a population of 9 million, has a peak

demand of around 14 gigawatts. The mismatch between Nigeria's generation capacity and its economic potential is staggering, and it does not resolve itself simply by arguing that solar panels are now cheap.

The Financing Gap That Policy Neglects

Clean energy investment in Africa ran to roughly $40 billion in 2022, according to the IEA. That sounds large. It isn't, relative to need. The IEA estimates that achieving universal electricity access in sub-Saharan Africa by 2030 would require sustained investment of around $190 billion per year across the entire region's energy sector — nearly five times what was actually invested. The gap is not primarily technical. It is financial, regulatory, and political.

Part of the problem is risk perception. Investors from Europe and North America applying standard portfolio logic see sub-Saharan Africa as high-risk: weak regulatory frameworks in some countries, currency volatility, political instability, limited revenue certainty for power purchase agreements. The risk premiums they demand make project financing significantly more expensive than in comparable projects in Europe or the United States. A solar plant in Germany might access capital at 3 to 4 percent interest; an equivalent project in Tanzania might face rates of 10 to 15 percent. The technology is the same. The financing environment produces wildly different outcomes.

Ritchie, in her data-driven writing on energy access, is blunt about this: "The biggest barrier to energy access in low-income countries isn't technology. It's finance." She argues that the multilateral development banks and wealthy-country climate funds — the World Bank's private sector arm, the IFC, the US International Development Finance Corporation, the European Investment Bank — have historically been far more willing to make rhetorical commitments to African clean energy than to actually deploy capital at the scale and risk profile required. The trend is improving, slowly, but years of pledges at climate summits have been followed by years of underdelivery on actual disbursement.

There is a further complication specific to the cooking fuel problem. Distributed solar handles lighting and phone charging reasonably well. It does not handle cooking. Electric cooking on induction stoves is extraordinarily efficient — around 85 percent of input energy becomes useful heat, compared to 10 to 15 percent for three-stone fires — but it requires a supply of reliable grid electricity and an appliance that costs between $30 and $100, which remains out of reach for many of the households in question. LPG is a cleaner alternative that several countries — India, Indonesia, Brazil — have successfully subsidised at scale. India's Ujjwala scheme, launched in 2016, distributed free LPG connections to 80 million households below the poverty line and included subsidised cylinder refills. By 2022 it had reduced the proportion of Indian households cooking on solid biomass from around 58 percent to roughly 35 percent. That is the largest reduction in cooking-fuel poverty achieved in any country in any comparable timeframe. It was not accomplished through leapfrogging to solar cooking. It was accomplished with LPG — a fossil fuel — targeted at a specific population with specific subsidies.

No one designing an emissions-minimising global energy system from scratch would choose this path. But designing from scratch is a privilege only available in textbooks.

Telling the Difference: When Leapfrog Works

The distributed solar success in East Africa is real, as noted, and understanding why it worked there is essential to knowing where it will and won't work elsewhere.

Kenya had several specific features that enabled the M-KOPA model. M-PESA — the mobile money platform launched by Safaricom in 2007 — had already penetrated 80 percent of the adult population by 2013, creating the payment infrastructure on which pay-as-you-go solar depends. Mobile phone ownership was high, producing both the payment mechanism and a clear, immediate use case for solar energy (charging). Kenya's geography means most of the country receives strong solar irradiance year-round. And critically, the alternative — grid extension into dispersed rural communities — was genuinely uneconomic for the near term, which meant distributed solar was not competing against a functional alternative but against nothing.

Where these conditions don't hold, the calculation changes. In a country with an existing, semi-functional grid — say Nigeria or Ghana — every dollar of distributed solar investment is partially competing against money that could upgrade the transmission and distribution infrastructure that would enable reliable grid power. The distributional question of whether the households best served by distributed solar are also the ones most in need of productive-use electricity capacity becomes important. A 20-watt solar home system is transformative if the alternative is a kerosene lamp. It is insufficient if the household's primary energy need is grinding grain, running a small welding shop, or refrigerating vaccines at a rural health clinic.

The IEA's Tier framework again: different tiers of access are appropriate for different needs, and the assumption that distributed solar will eventually supply all tiers is not supported by current technology economics. Tier 4 and 5 access — the kind that powers productive enterprise — currently requires something resembling a functioning grid, not a series of isolated household systems.

Uganda is an instructive case. The country achieved substantial penetration of solar home systems through programs similar to Kenya's, and off-grid solar company activity was high throughout the 2010s. Yet Uganda's economy, particularly its manufacturing and agricultural processing sectors, remained constrained by the absence of reliable three-phase grid power in most rural industrial areas. The leapfrog to solar worked for households. It did not substitute for the grid investment that productive enterprise required.

The Arithmetic of Inclusion

Return now to the global carbon budget. Here is the number that changes everything depending on how you handle energy poverty.

Scenarios produced by the IEA, the IPCC, and institutions like the Rocky Mountain Institute typically model the global energy transition with varying assumptions about what happens in low-income countries. In the scenarios where sub-Saharan Africa and South Asia maintain low per-capita energy consumption — which is to say, scenarios where energy poverty is essentially a permanent condition in the model — the carbon budget arithmetic for the 1.5°C target is tight but arguably achievable if wealthy countries move fast enough. In the scenarios where those regions develop toward even 30 or 40 percent of current European per-capita energy use — which is a modest and morally defensible aspiration — the budget arithmetic becomes extremely challenging even with rapid clean energy deployment.

This is what Smil means when he writes, in Energy and Civilization, that scale is the central variable that most energy commentary fails to reckon with honestly. The discussions about whether Britain should permit a new gas peaker plant, or whether California should extend the life of Diablo Canyon nuclear station, are conducted in a frame that treats global energy demand as approximately fixed. It is not fixed. It is going to grow substantially, the growth is going to occur overwhelmingly in regions that currently have very low energy consumption, and whether that growth happens on clean or fossil energy depends on financing decisions and technology deployment strategies that are being made — or avoided — right now.

The IEA's Sustainable Development Scenario, which attempts to model universal energy access alongside significant emissions reduction, finds that achieving both requires roughly doubling global clean energy investment annually through 2030, with a significant share of that investment directed at low-income countries. As of 2023, global clean energy investment was running at approximately $1.8 trillion per year — a record, and genuinely impressive. But the IEA's own estimates suggest the 2°C pathway requires approximately $4.5 trillion per year by 2030. And the distribution of that investment matters as much as the total: most of it is currently flowing to Europe, China, and North America.

The Right to Develop

There is a version of this argument that gets made in bad faith — where "African countries have the right to develop using fossil fuels" is used as a pretext for continued coal exports, continued financing of fossil infrastructure by development banks under political pressure, or continued delay in wealthy countries' own transitions. That version deserves the scepticism it receives.

But there is also a version made in good faith, and it is harder to dismiss. Rose Mutiso, a Kenyan energy researcher and co-founder of the Energy for Growth Hub at the University of Chicago, has written extensively about what she calls the "energy poverty premium" — the systematic undervaluation, in global energy policy discourse, of the energy services that poor people need to live full lives. She argues that the dominant framing treats energy poverty as a side issue to be addressed with minimal-viable-product solutions (solar lanterns, cookstoves) while the real energy policy action happens in wealthy countries debating the marginal decarbonisation of already-rich grids. The result is a policy architecture that is simultaneously inadequate for development and insufficient for climate, because it refuses to honestly integrate the two goals.

The integration, when you do it honestly, reveals some uncomfortable conclusions. If low-income countries are genuinely going to achieve development-level energy access, and if that energy is going to be clean rather than fossil, then the cost of that clean energy deployment must be substantially borne by countries that have already emitted their historical share of the carbon budget. That means not just pledges and partnerships but actual capital, at concessional rates, for grid infrastructure, transmission, storage, and manufacturing capacity in regions that currently lack it. It means supporting the development of domestic renewable manufacturing industries in Africa and South Asia rather than simply exporting Western panels and turbines. And it means accepting that some countries will continue using natural gas for transition-period power generation, because the alternative — indefinite energy poverty — is neither just nor climatically helpful if it prevents the development of the tax base and governance capacity needed for long-term clean energy investment.

None of this is comfortable for the COP negotiating room. It involves wealthy countries accepting financial obligations larger than they have so far been willing to accept, and it involves letting go of the conceit that energy poverty can be solved cheaply with minimal-technology off-grid systems while the real decarbonisation happens elsewhere.

What Actually Moves the Needle

The evidence from countries that have successfully reduced energy poverty at scale points to a consistent set of factors. Targeted subsidies for clean fuels — India's LPG program being the clearest example. National electrification programs with political will and central coordination — India's Saubhagya, Bangladesh's Rural Electrification Board, which connected 20 million rural households between 1977 and 2015 using a cooperative model. Transmission grid investment in advance of demand, rather than after it. Carbon-finance mechanisms that allow developing countries to attract international capital for clean projects without bearing the full cost premium of frontier-market finance.

And, critically, a willingness to use whatever technology is cheapest for the specific application, rather than ideological commitment to a particular solution. Bangladesh's 6 million solar home systems were the right choice for dispersed rural households far from the grid. The country's subsequent investment in natural gas and — controversially — two coal plants at Rampal and Matarbari represented an attempt to build industrial baseload power for manufacturing and urban growth. Whether Rampal and Matarbari were good decisions depends on what alternatives were actually available and financeable at the time. Retrospective critique from wealthy countries that declined to offer affordable clean alternatives is not a substitute for having offered them.

Amara, the woman from northern Nigeria at the start of this chapter, needs electricity. She also needs her children to be able to run a business, get an education in a lit building, have food refrigerated, be treated in a clinic with reliable power. She does not particularly need the abstraction of "the energy transition." She needs the concrete, material reality of reliable electrons delivered at an affordable price to the place where she lives. The question of where those electrons come from is secondary. Ideally, and increasingly feasibly, they come from the sun and wind. But "ideally" has to mean "in reality, with real financing, deployed at real scale, on a

timeline that matters to a thirty-four-year-old woman breathing wood smoke right now."

The transition is not honest, and it is not complete, if she is not in it.

CHAPTER XII

The Next 30 Years: Where the Math Actually Points

Anyone who tells you they know what the energy system will look like in 2055 is selling something. The honest answer is a range — bounded on one side by physical laws that don't negotiate and on the other by cost curves that keep moving in directions almost no one predicted five years ago. What this chapter offers is not a forecast. It is a map of probabilities, organised by how confident the evidence actually allows you to be.

That distinction matters enormously. The energy conversation has been plagued, for decades, by two equally useless postures: the techno-optimist who tells you the transition is basically done if we'd just get out of the way, and the doomer who insists civilisation cannot function without oil and coal and that solar panels are a rounding error. Both positions feel satisfying. Neither survives contact with data.

Vaclav Smil has spent fifty years quantifying energy systems, and the single most important thing he insists on — more than any specific number — is that scale has a way of humiliating ambition. When he writes in How the World Really Works that we are a fossil-fuel civilisation, he doesn't mean it as a slogan. He means it as arithmetic. Fossil fuels supply roughly 80 percent of the world's primary energy today. Every year, humanity burns about 500 exajoules worth of oil, coal, and gas — a number so large it is essentially meaningless until you translate it into something physical. It is, as Smil notes, roughly equivalent to burning 14 billion tonnes of oil. And global energy demand is still growing, driven by the entirely legitimate desire of billions of people in low-income countries to live longer, cooler, better-fed lives.

Hannah Ritchie, working from Our World in Data, makes a complementary point with more optimism: things that look impossible have a habit of happening anyway, once the economics align. Solar photovoltaics cost 99 percent less than they did in 1976. Wind turbines produce electricity that is, in many locations, cheaper than running an existing coal plant. Electric vehicles sold 14 million units in 2023 alone, according to the IEA, representing roughly 18 percent of all new car sales globally — a number that was essentially zero fifteen years ago. The transition is real. It just isn't finished, and the remaining work is the hard part.

So. Where does the math actually point?

The Nearly Certain Column

Start with electricity generation, because that is where the evidence is most unambiguous. The IEA's 2023 World Energy Outlook stated, in language that would have been inconceivable in a government report a decade ago, that "the era of fossil fuel growth is over." That is probably true for power generation specifically, even under pessimistic policy assumptions.

In 2023, the world added roughly 440 gigawatts of solar capacity. Four hundred and forty. To put that in context: it took the entire nuclear industry about sixty years to accumulate 400 gigawatts of installed capacity. Solar is now adding that in a single year. The learning curve — Wright's Law, which predicts that costs fall by a fixed percentage for every doubling of cumulative production — shows no sign of flattening. Utility-scale solar already achieves levelised costs below $30 per megawatt-hour in high-irradiance locations. Onshore wind is similarly cheap across three continents.

By 2055, in almost any plausible scenario, electricity generation in wealthy countries will be dominated by variable renewables. The IEA's Net Zero by 2050 scenario projects that solar and wind together supply over 60 percent of global electricity by mid-century. Even the more conservative Stated Policies Scenario — which just extrapolates current commitments without additional ambition — has renewables at around 45 percent of global power by 2050. The range of outcomes here is genuinely narrow. The physics is favourable, the economics are settled, and the manufacturing base is enormous.

Light-duty vehicles belong in this column too, though with a slight asterisk. EV adoption is following a classic S-curve, and there are good reasons to think the inflection point is already behind us in the key markets. Battery pack prices fell from roughly $1,200 per kilowatt-hour in 2010 to around $139 per kilowatt-hour by 2023, according to BloombergNEF. The point at which an EV is cheaper to own over its lifetime than a comparable internal combustion vehicle — accounting for fuel and maintenance savings — has already arrived in several segments in China and Europe, and is arriving in the U.S. within this decade. Once purchase price parity lands, the transition tends to accelerate rapidly. Norway, which reached EV purchase parity earlier than most countries because of its tax structure, saw electric vehicles hit 80 percent of new car sales in 2022.

By 2055, a world in which the majority of passenger vehicles on the road are electric is not optimistic — it is the base case. The caveat is that cars last fifteen to twenty years on average, so even a complete halt to ICE vehicle sales tomorrow would leave hundreds of millions of combustion engines in service through the 2030s and into the 2040s. The transition is certain. Its pace is not.

The Plausible but Not Assured Column

This is where honest analysts spend most of their uncertainty budget. Heavy industry, space heating, short-haul aviation, and medium-distance shipping are all sectors where credible decarbonisation pathways exist — but where the technology is more expensive than the incumbent, the policy support is inconsistent, and the timelines are genuinely unclear.

Take space heating. Buildings account for roughly 30 percent of final energy consumption globally, and most of that comes from burning natural gas, oil, or biomass to keep people warm. Heat pumps — which move heat rather than generating it, achieving efficiencies of 200 to 400 percent — are the obvious replacement, and in new construction in mild climates they already make economic sense. The problem is retrofit. The world's building stock turns over slowly; the average building in Europe and North America was built

decades ago, with insulation standards from a different era. Installing a heat pump in a poorly insulated Victorian terrace costs significantly more and delivers less performance than in a modern passive-house structure.

Germany's attempt to mandate heat pump installation via the 2023 Buildings Energy Act — subsequently watered down after significant political backlash — illustrates the gap between what is technically possible and what is socially achievable on any given timeline. By 2055, a majority of new buildings in wealthy countries will almost certainly be heated electrically. What fraction of existing buildings will have been retrofitted is genuinely unclear, and it matters: in a high-electrification scenario, it changes grid demand profiles dramatically.

Short-haul aviation is more tractable than it gets credit for. Flights under 500 kilometres — which represent the majority of departures by number, though not by fuel consumption — are within reach of battery-electric aircraft on current technology trajectories. Startups like Heart Aerospace and Eviation are targeting regional routes of 200 to 400 kilometres with aircraft entering service in the late 2020s. The energy density of current lithium-ion batteries, at roughly 250 watt-hours per kilogram at the pack level, is marginal for this mission but workable. If battery energy density reaches 400 watt-hours per kilogram — plausible but not certain by 2035 — the economics improve substantially.

Hydrogen-powered short-haul flight is also on the table. Airbus has its ZEROe programme targeting hydrogen aircraft by 2035, though the programme has faced repeated internal questioning about timeline realism. Green hydrogen — produced by electrolyzing water with renewable electricity — is the only version of hydrogen that delivers meaningful emissions reductions, and its production cost has fallen faster than most projections suggested but is still, as of 2024, roughly three to five times the cost of fossil hydrogen in most markets.

For medium-distance shipping — coastal freight, ferries, short sea routes — ammonia and methanol are gaining traction as alternative fuels, and several major shipping companies have ordered dual-fuel vessels. Maersk launched its first methanol-capable container ship in 2023. Whether these fuels scale fast enough to transform the sector by 2055 depends largely on whether green ammonia and green methanol production can follow a cost curve analogous to solar — which requires both policy support and massive capital deployment that is not yet materialising at the required rate.

The Column Where Current Trajectories Are Insufficient

Some honesty is required here, and this book has tried to model it throughout. There are sectors where decarbonisation is not only hard but where the current trajectory of technology development and policy support would, if extrapolated, leave the problem substantially unsolved by 2055.

Long-haul aviation is the clearest case. A Boeing 777 carrying 300 passengers from London to Singapore consumes roughly 100 tonnes of jet fuel over about 11,000 kilometres. The energy in that fuel — approximately 1,200 megawatt-hours — has to come from somewhere. Batteries at current energy density

would weigh roughly 4,800 tonnes to store the same energy. The aircraft cannot fly. Sustainable aviation fuel, made from biomass or synthesised from green hydrogen and captured CO,, ca and is genuinely compatible with the current fleet. The problem is supply. Global SAF production in 2023 was approximately 600 million litres — a fraction of a percent of the 300-plus billion litres of jet fuel the aviation sector burns annually. Scaling SAF to the required level by 2050 would require either an enormous amount of agricultural land (if biobased) or an enormous amount of cheap renewable electricity (if synthetic). Both pathways face constraints that don't disappear by wishing.

The IPCC's Sixth Assessment Report and the IEA's aviation analyses both conclude that aviation is one of the most challenging sectors to decarbonise fully, and that even under aggressive policy scenarios, residual emissions may require carbon removal to offset. This is not an argument for despair. It is an argument for honesty about what "net zero" means in practice — it means some sectors will decarbonise fully and others will require permanent offset, and that is a different promise than the slogan implies.

Deep-sea shipping faces similar constraints at enormous scale. Shipping carries roughly 80 to 90 percent of global trade by volume, and the largest container ships consume the equivalent of 200 to 300 tonnes of heavy fuel oil per day. The energy requirements are staggering. Ammonia is the leading candidate fuel for deep-sea shipping decarbonisation: it contains no carbon, it can be synthesised from green hydrogen and atmospheric nitrogen, and it is liquid at manageable conditions. But it is also toxic, energy-intensive to produce, and roughly three times the cost of fossil fuels at current green production prices. The IMO's 2023 revised strategy commits to net-zero by "around 2050" but without binding fuel mandates that would actually force the economics. Without those mandates or a meaningful carbon price on bunker fuel — which has been discussed for decades and not implemented — the commercial incentive to switch is weak.

Blast-furnace steelmaking deserves its own paragraph because it illustrates the chemistry problem that pure electrification cannot solve. Making steel from iron ore requires stripping oxygen from iron oxide. The traditional method uses coke — derived from coking coal — as both fuel and reducing agent. You can replace the heat with renewable electricity. You cannot replace the chemical reduction with electricity directly. The two alternatives are direct reduced iron using green hydrogen (HDRI-EAF), which works but requires enormous quantities of cheap green hydrogen, or electric arc furnace steelmaking, which works brilliantly but primarily requires scrap steel as input. Global scrap availability is finite and growing slowly. A world producing 1.9 billion tonnes of steel per year cannot run on scrap alone. HYBRIT in Sweden and Boston Metal in the U.S. are demonstrating green steelmaking at pilot scale, and several major producers have announced commercial-scale green steel plants. The question is whether these can scale to displace blast furnaces across the 2030s and 2040s at a rate consistent with climate targets. The answer, looking at capital investment commitments versus the scale required, is probably not without policies that make it significantly more expensive to operate a blast furnace. Carbon contracts for difference — which the EU has begun to deploy — are the right instrument. They are not yet deployed at the right scale.

Cement is the other hard case, and the one that attracts less attention than it deserves. Global cement production emits roughly 4 billion tonnes of CO, per year — about 8 percent

roughly 60 percent of that comes not from burning fuel to heat kilns but from the chemical decomposition of limestone into clinker. You heat calcium carbonate and you get calcium oxid of the rock. There is no way to prevent this reaction using renewable energy; you need either carbon capture at the kiln, alternative clinker materials that don't require limestone calcination (geopolymers, belite cements), or a dramatic reduction in cement use. All three pathways are viable in principle. None is deployed at anything approaching the required scale. Carbon capture and storage for cement kilns is technically proven but expensive and requires CO, transport and sequestration infrastructure tha the cement sector will almost certainly be less carbon-intensive than today. Whether it is on a net-zero trajectory is, to be direct, doubtful without a step-change in policy ambition.

The Pace Problem, Revisited

Here is the number that should be in every conversation about energy transitions: fifty years. That is approximately how long it took coal to go from roughly 5 percent of global primary energy to 50 percent. Oil's rise from the same starting point took about forty years. Natural gas, faster still, took about thirty-five years to reach its current share of around 24 percent. Smil has documented this repeatedly: energy transitions are measured in generations, not election cycles.

Solar and wind are currently at around 4 to 5 percent of global primary energy — total primary energy, not just electricity. The electricity share is more impressive, at roughly 30 percent globally in 2023. But electricity is only about 20 percent of final energy consumption worldwide. When you account for the heat used in industry, the fuel burned in vehicles and aircraft and ships, and the direct combustion in buildings, renewables are still a small fraction of the total.

This is not a reason for pessimism. The pace of solar deployment is historically unprecedented. The cost declines are real and durable. The policy momentum — imperfect and uneven as it is — is stronger than at any previous point. But the speed at which a new energy source can displace an existing one is constrained by the lifespan of the infrastructure it needs to replace. A coal plant built in 2010 will run for thirty to forty years without extraordinary political intervention. A gas distribution network installed in 2020 will be expected to earn returns for decades. The people who own those assets have lawyers and lobbyists and, in many countries, implicit or explicit government backing.

None of this means the transition fails. It means it is slower than the optimists' models suggest. The IEA's Net Zero by 2050 scenario is technically feasible by the IEA's own analysis. It requires, among other things, no new fossil fuel development beyond projects already in operation or under construction as of 2021 — a commitment that essentially no major oil-producing country has made. It requires tripling global renewable capacity by 2030, which is the trajectory only if every announced project gets built on time. It requires electric vehicles to reach 60 percent of new car sales globally by 2030, from 18 percent in 2023.

These targets may be achievable. They are not the central case.

What the Numbers Look Like at Mid-Century

A more defensible projection for 2055 looks something like this. It is a range, not a point.

In power generation, the transition will be substantially complete in wealthy countries. The United States, the EU, Japan, South Korea, and China will all have electricity systems dominated by renewables, with gas providing flexibility and declining firm capacity. Coal will be a marginal fuel in wealthy countries — kept alive, where it survives, by politics rather than economics. Nuclear will contribute a stable or modestly growing share, primarily through the extension of existing plants and the addition of new capacity in China, which is currently building more nuclear reactors than the rest of the world combined. Small modular reactors may be commercial by 2035 — the first NuScale plant in the U.S. was cancelled in 2023 due to cost overruns, which is the kind of setback that tends to recur in nuclear — but they will not be a large fraction of global capacity by 2055 under any realistic projection.

In transport, the majority of passenger vehicle kilometres driven in wealthy countries will be electric by 2055. Two-wheelers, which are the most common motor vehicle in the world by units, are already predominantly electric in China and trending that way elsewhere. Aviation and shipping will have lower carbon intensities than today, primarily through efficiency improvements and partial fuel switching, but will not be fully decarbonised. These two sectors together account for roughly 5 to 6 percent them to zero by 2055 is implausible on current technology and policy trajectories.

In industry, steel and cement will be on transition pathways, with meaningful deployment of green hydrogen DRI and alternative cements, but significant residual emissions. The fraction will depend heavily on whether carbon pricing or equivalent policies are implemented at scale. If the EU Carbon Border Adjustment Mechanism works as intended — and if other major economies adopt analogous measures — it could pull investment toward green industrial processes at a rate that current voluntary commitments cannot. If it stalls in political negotiation and trade disputes, the transition in heavy industry will be slower.

In buildings, heat pumps will dominate new construction in temperate climates. Retrofit rates will be substantially lower than targets. Natural gas will still heat a significant fraction of existing residential stock in 2055, especially in countries where insulation standards have not been reformed. This is not a counsel of despair; it is what happens when you need to retrofit hundreds of millions of buildings at a cost of $15,000 to $50,000 each and the subsidy programmes are perpetually underfunded.

In energy-poor countries — the billion-plus people currently without electricity access, the 2.3 billion cooking over open fires — the picture is more genuinely optimistic in some respects. Distributed solar has demonstrated it can reach communities that grid extension cannot reach economically. M-KOPA in East Africa has connected over 3 million homes through solar home systems financed on mobile payment plans. This is real progress. It is also not the same as full electrification at the level that drives industrial development. The right to grow energy consumption, which every wealthy country exercised through coal and oil and gas, remains contested and under-financed in international climate negotiations. By 2055, more

people will have electricity. Whether they will have enough electricity to develop economically, without the freedom to use the cheapest available sources — that question is political, and the answer is not yet written.

The Real Binding Constraints

The grid, as the previous chapter argued, is the honest binding constraint in wealthy-country decarbonisation. A power system that is 80 percent renewables requires either massive overbuilding, continental-scale storage and interconnection, or demand flexibility that the current grid architecture cannot provide. All three require transmission investment, permitting reform, and software upgrades that are proceeding at a fraction of the required rate in the U.S. and Europe.

The U.S. interconnection queue — the backlog of projects seeking connection to the grid — exceeded 2,600 gigawatts in 2023. That is more than double the entire current installed capacity of the American power sector. The projects in that queue represent a genuine willingness to build. The reason they are not being built is not technology or economics. It is permitting timelines, transmission access disputes, and the absence of a planning authority with the mandate and resources to coordinate grid expansion. The average wait time in the U.S. interconnection queue is now over five years. In the UK, some offshore wind projects have waited eight to ten years for grid connection.

Solving this requires regulatory and political change more than it requires new technology. Which makes it, in some ways, harder.

Materials supply chains are the second binding constraint that wasn't visible until recently. Solar panels need silicon and silver. Batteries need lithium, cobalt, nickel, and manganese. Wind turbines need rare earth elements for permanent magnets. None of these materials is scarce in absolute geological terms — the earth contains enough lithium to electrify every vehicle several times over. But mining and refining them at the required rate, while managing the environmental and social impacts of doing so, is a bottleneck that takes decades to ease. The Democratic Republic of Congo currently produces roughly 70 percent of global cobalt. China refines roughly 60 percent of global lithium and processes a dominant share of virtually every other critical mineral in the clean energy supply chain. This is a geopolitical dependency that has not escaped the attention of the IEA, whose 2023 Critical Minerals report documents it in uncomfortable detail.

Battery chemistry is shifting toward less cobalt-intensive formulations — lithium iron phosphate, or LFP, batteries contain no cobalt and now represent the majority of EV batteries sold in China — but other minerals create analogous dependencies. The transition will create new resource geopolitics even as it diminishes the geopolitical weight of oil and gas. Different chokepoints, not no chokepoints.

What Happens If We Get It Mostly Right

Global average temperatures are very likely to exceed 1.5°C above pre-industrial levels. That ship has probably sailed. The IPCC's Sixth Assessment Report gives a greater than 50 percent probability of exceeding 1.5°C in the 2030s, even in scenarios with strong mitigation. The number that matters now is whether we are

heading for 2°C, 2.5°C, or 3°C — and that is a question the next thirty years will largely answer.

The IEA's central scenario under current policies — not net zero, not any new commitments, just what has already been announced and implemented — results in around 2.4°C by the end of the century. That is lower than the 3.5°C-plus trajectory of fifteen years ago. The policy progress is real. It is not enough.

Getting to around 2°C — not safe, but substantially better than 2.4°C — requires roughly the trajectory of the IEA's Announced Pledges Scenario: every country meeting every commitment it has already made, on time, with no backsliding. Whether you find that reassuring or alarming probably depends on your prior assessment of governments' record of delivering on announced commitments.

Getting to 1.5°C almost certainly requires carbon dioxide removal at gigaton
atmosphere using direct air capture, enhanced weathering, bioenergy with carbon capture, or reforestation. Direct air capture currently costs between $400 and $1,000 per tonne of CO,, peer-reviewed estimates. It needs to get to roughly $100 to $150 per tonne to be economically deployable at scale. Whether that cost reduction happens depends on learning curves, policy support, and capital that is not yet flowing at the required magnitude.

The Honest Conclusion

Here is what the evidence actually supports.

The energy transition is not a speculation. It is happening. Solar and wind are the fastest-growing energy sources in human history. Electric vehicles crossed the 10 percent global market share threshold in 2022 and have not looked back. Battery prices continue to fall. Green hydrogen is expensive today and probably won't be by 2040 in regions with abundant cheap renewable electricity. The direction of travel is clear.

The speed is not clear. Smil's historical record of fifty-year transitions is not an iron law, but it is a heavy prior. The sectors where transition is easy are the sectors where it is already happening. The sectors where transition is hard are hard for reasons that won't dissolve under optimism: chemistry, physics, infrastructure lock-in, supply chain bottlenecks, and the entirely rational behaviour of people who made long-term investments in the current system.

The sceptics are wrong that the transition is impossible or economically ruinous. The advocates are wrong that it is essentially complete or imminent. Both postures are emotionally satisfying and analytically useless.

Ritchie's point is the right one to close on: the question has never been whether we can build a low-carbon energy system. The question is whether we can build it fast enough, at large enough scale, while simultaneously ensuring that the people who currently lack energy access get to use more of it, not less. That second condition is not a footnote. In a world where 750 million people lack electricity and 2.3 billion cook over open fires, an energy transition that secures comfort for the already-comfortable while preventing the poor from developing is neither just nor stable.

The math points to a world in 2055 that is less carbon-intensive than today. Almost certainly. That will have involved the fastest large-scale industrial transformation in history. It will have created new dependencies — on critical minerals, on Chinese manufacturing, on software systems that didn't exist twenty years ago — while dissolving old ones. It will have been slower than the Paris Agreement's architecture requires and faster than the fossil-fuel industry's lobbyists predicted. It will have left work undone: in long-haul aviation, in deep-sea shipping, in cement and steel and high-temperature industrial heat, where solutions exist but haven't been deployed at scale.

The useful response to that reality is not cheerleading and not despair. It is the kind of clear-eyed, quantitative, historically grounded attention that Smil has modelled for decades: knowing what a joule is, knowing what an exajoule is, knowing how long it took the last transition to happen and why, and refusing to let the emotional appeal of either the utopian or the catastrophist narrative substitute for arithmetic.

We built this energy system over two centuries. We will not replace it in thirty years. We will replace significant parts of it, faster than anyone thought possible fifteen years ago, and the parts that remain will be the hardest parts — the ones that require not just cheaper panels and better batteries but political will, international cooperation, and the kind of patient investment in infrastructure that democratic governments find structurally difficult to make.

That is the honest account. It is enough to work with.