THE CONCEPT OF CLIMATE AND THE LIMITS OF MECHANISM

There is a façon de parler among climate scientists that cannot fail to draw the attention of an outsider, especially one approaching the topic from the field of classical philosophy. It is the use of the term 'mechanism' as a shortcut expression for 'adequate explanation.' Of course climatologists are not unique in this way of speaking. Biologists, for example, speak of getting the mechanism for photosynthesis. But there have been certain developments within climate science that suggest that this use of the term 'mechanism' might already be en route to becoming, like so many other words in English, a dead metaphor. I should not venture to propose anything like linguistic reform for climatology, or for any ongoing science. Still, the prevalence of this expression might serve as a point of departure for a consideration of what I take to be the basic 'trajectory' (another dying metaphor) in the concept of climate over the past century, beginning, at the very latest, with Louis Agassiz's posing ice ages as a fact in search of an explanation. Over most of the years during which an explanation has been sought, most of the explanations proposed have indeed been mechanistic. But for some decades now, explanatory models have arisen that are, arguably, non-mechanistic. Which is not to say that mechanistic models have been abandoned—it is quite unthinkable that this shall ever happen. Some limits of mechanistic explanation have nevertheless been uncovered of late and it is this voyage of discovery that I should like to comment upon in the remarks that follow. To this end the old chestnut, 'What is a mechanism?,' will need a brief revisiting. In search of a simpler and more comprehensive account I shall cast a net back to the giant among the founding fathers of science, Aristotle, focusing upon the theory of principles articulated in the opening book of his *Physics* and elsewhere. I do so because no subsequent account of scientific principles has even approached the clarity achieved by Aristotle. In light of this account it should be possible to grasp more precisely just what a mechanism is and why recent developments in climate science point to modes of argument that exceed the limits of mechanism without becoming any less scientific for all that.

A. A Primer on Scientific Principles

At the climax of Hellenic science, 2,360 years ago, most topics of investigation were grasped within the theory of principles then articulated by Aristotle. I would like to invite you to consider whether climate science, especially as it has developed over the past 15 years, can be made more intelligible if grasped in terms of Aristotelian principles. Since modern science has tended to define itself, especially since the 17th century, by its liberation from the authority of Aristotle, my proposal may well seem quixotic. The mechanization of our picture of the world, especially as dramatized in Newton's *Principia*, has, for some centuries, marked, as Alexander Pope famously put it, the advent of light over the murky darkness of the pre-modern, Aristotelian world picture. There is, of course, some truth in this trope. One feature of the Aristotelian world was its geocentricity and its bifurcation of the physical universe into spheres defined by their kind of natural motion: a superlunary realm of quintessentially perfect curvilinear motion and a sublunary realm of essentially incomplete rectilinear motion. As the Copernican revolution unfolded, that bifurcation was swept away and any attempt to resurrect it within the field of cosmology in general would be pointless, despite the fact that, in the more restrictive field of climate science, there has been a certain resurgence of geocentrism in recent years, which I shall touch upon later. At this point I simply want to indicate what I take to be the enduring core of Aristotelian science. And that is his theory of how principles (Greek: *archai*) function in scientific explanations. The essence of his argument is that a science is scientific only when it proceeds in accordance with principles and that there are just three kinds of science: those that operate with a single principle, those that employ two, and those whose principles are three.

He observed that his predecessors in physics, beginning with the Milesians, had sought to explain 'what is' on the basis of a single principle—the water of Thales, the indeterminate (*apeiron*) of Anaximander, the air of Anaximenes, etc. Single-principle explanation may also be called **genetic** because it presumes, as many still do today, that something is explained if its source, origin, or genesis can be identified—that is one reason why evolutionary biology and 'big bang' cosmogony have been so appealing. The early Greek physicists, like their modern counterparts, implicitly assumed that 'matter,' what is potentially other, had to be the principle for explaining change in physical entities. "Ye shall know them by their roots" has long had a plausible ring. It points to a certain *identity* between an origin and its result—the material principle was presumed to be, *qua* that matter, namely, *in* the result. Hence the identity. In genetic explanations the result of the change is presumed to be already contained in the point of departure.

Granted that this explanatory trope will always remain indispensable, the perennial response to single-principle genetic explanations of change (Greek: *metabole*) is that they fail to account for the specific *difference* between before and after. If change is what is to be explained, explanation will be limited if there is no intelligible difference in principle between the change-from-what and the change-to-what. That is one reason why catastrophic meteorological accounts of the end of the age of dinosaurs—and all other appeals to a 'big influence'—tend to be unsatisfying. Of course we shall never abandon the model of genetic explanation, but, when the chips are down, we have again and again wanted, at the very least, to grasp the mechanism and for this there must be two principles.

Adumbrated by the Pythagoreans, the insight that a mechanically conceived process requires two principles was fully articulated by Plato. He saw that a mechanical process would have to involve, in addition to a material principle—something potentially determinable, a source—also a formal principle, a determinate. Hence the emergence of the classical contrast between matter and form which has governed most scientific thought in the meanwhile, making it, in Whitehead's celebrated phrase, a series of 'footnotes to Plato.'

In mechanical explanations the material explanandum counts as explained if it can be fitted into a formal explanans. To take a simple example, diurnal material phenomena are explained when they can be shown to fit into a formally conceived cycle of the earth's non-parallel axial rotations vis-à-vis the sun. The explanation is mechanical because there is a presumed *difference* in principle between the matter to be explained and the formal mechanism in conformity with which the material phenomena are 'saved.' That is why mechanical explanation involves two principles.

Although Aristotle (whose express understanding of diurnal phenomena does not, of course, match this example) recognized the significance of Plato's advance toward a more adequate comprehension of the principles of science, he also observed that Plato's two-principle theory left an important conflation within the dimension of the change-from-what, namely, (a) the material or determinable aspect of the source and (b) the determinate or form aspect of the source. When these, thus conflated, are *qua* matter set in contrast to form as the determinate *telos* of a physical process, the determinate aspect of the source is systematically excluded from consideration. Like his 'Italian' predecessors (to use Aristotle's reference to the Pythagoreans), Plato grasped the form-to-which any process proceeded as the sole determinate aspect to be accounted for in explaining the thing undergoing that process. Thus Plato's form was grasped as the determinate form-through-which a process proceeds, but it left out of account the form-from-which a process proceeds, conflating this with the material principle—a conflation sustained in all of the many 'footnotes to Plato' which have constituted most of science in the meanwhile. Such two-principle explanation may be called **mechanistic** because, as E.J. Dijksterhuis has shown,[1] a 'mechanical' physical system is one in which the formal-mathematical dimension predominates over all determinate aspects of the material phenomena to be explained. This of course entails that the entities which undergo change have to be conceived as indeterminate and characterless. The consummate, and indeed stupendous, realization of this 'Italian' project of explanation was, of course, Newton's *Principia*, in which all acceptable principles of explanation were formal-mathematical and all putative 'entities' were reduced to mathematical points, lines, ratios, and the like. As Pierre Duhem has argued, this was the ultimate result of the quest to realize a way, in the words of Plato's *Timaeus*, "to save the phenomena."

Despite the fact that it has been, by and large, ignored, Aristotle's correction of the defect in the Platonic theory of principles is breathtakingly simple. Any process, to be comprehended, must involve two formal principles as well as a material principle: (1) a determinate form-from-which the process proceeds, (2) a determinable matter which undergoes the process, and (3) a determinate form-to-which the process proceeds. Hence the most distinctive, albeit mostly forgotten, feature of Aristotelian science: that a fully satisfactory explanation must clearly involve three principles, no more and no less. A corollary of this is that, for a process to be theoretically comprehensible, all three principles must be immanent in the process.

Of course Aristotle recognized that not all de facto processes involve three immanent principles. Accidents will happen and, in the world as he conceived it, they happen with considerable regularity. For if, in any process, the form-to-which principle (3) is imposed from a sphere external to the entity undergoing the process, then that process will be, unless its sphere is legitimately enlarged, insusceptible of theoretical comprehension. That is why, for Aristotle, **genetic** and **mechanistic** explanations ultimately fail to satisfy theoretical curiosity, however indispensable they may be for the understanding of certain kinds of physical processes. If we must have a term to describe Aristotle's distinctive three-principle approach to science, let us call it **theoretical**, for its purport is to satisfy theoretical curiosity, regardless of all other considerations, and to facilitate a state of activity called theorizing (Greek: *theorein*), one mark of which is sheer pleasure.

The topic of these remarks is the concept of climate and the limits of mechanism. It is not my point to denigrate mechanistic or even genetic modes of explanation in climate science but rather to

---

[1] E.J. Dijksterhuis, *The Mechanization of the World Picture: Pythagoras to Newton*, C. Dikshoorn, tr., Princeton: PUP, 1986.

show their limits and how they can be, and have been, integrated into a larger, more satisfying, concept of the climate system, one which embraces theoretical explanation in the sense specified.

Climate is a particularly apt sphere of science for the project I propose because it has undergone immanent developments since the 19th century, well known to all climate scientists, that will expedite my argument. And, to be more explicit, these recent developments in climate science have been toward a more **theoretical** model of explanation.

Before I begin allow me to say that climate, as a topic of theoretical inquiry, was first articulated in the three editions of Hegel's *Encyclopedia of the Philosophical Sciences* (1817, 1827, 1830). Although I shall try to refrain from mentioning Hegel in the following remarks, I must acknowledge that I shall be using his thought at virtually every step in my argument. For, as a critical mass, albeit still a distinct minority, of Hegel scholars have come to recognize, Hegel is indeed the 'Aristotle of the modern world' and the thinker who has most clearly recognized the merits as well as the limits of mechanism.

Prescinding, as far as possible, from mentions of Hegel—for my argument must stand on its own—, my aim in what follows will be to trace how, by gradual stages leading up to the present decade, the limits of mechanism have been uncovered.

B. Climate Science—The Mechanistic Phase: From Newton via Agassiz to Milankovitch

Given its astonishing capacity of accounting, as never before, for the mechanism of tides—a stunning revelation to the British maritime nation—, Newton's system of the world won, as scientific innovations go, an exceptionally rapid reception in the 17th and following centuries. For many, including meteorologists, the mathematical elegance of Newton's mechanical picture soon gave way to the more pedestrian notion of a mechanism as 'clocklike.' On this model, the solar system was understood to be, once God the creator had set it in motion, in a steady state. Despite an early-19th-century theoretical challenge by Hegel, this picture of climate first came under serious doubt when, at mid-century, Louis Agassiz was finally able to persuade his scientific peers in Europe that that continent had, in the not too distant past, undergone an ice age. How account for this deviation from the steady state to be expected within the Newtonian mechanism? To be sure, the fact that the physics of orbital mechanics had by the 19th century discovered that the earth, as it tours its heliocentric orbit, spins at an axial tilt from its perpendicular to the sun, and that this tilt undergoes, over a cycle of 22,000 years, a wobble that came to be called a 'precession.' That another jitter in the earth's spin repeats itself every 19,000 years was a fact also known. Finally, it was known that the earth's axis of deviation from perpendicularity to the sun underwent a cycle of deviation, from 21.5 degrees to 24.5 degrees, and that this cycle was traversed in 41,000 years.

Given Agassiz's discovery of ice ages and the known facts of orbital physics, the challenge was to find a way of correlating the data to elicit an explanation of ice ages within the still essentially mechanical conception of the climate system then prevalent. The received Newtonian model had been so successful in accounting for diurnal, tidal, seasonal, and other cycles, that it was natural to search for an analogous cycle to account for the recurrence of ice ages. By general reckoning, this challenge to find a Newtonian mechanism for ice age cycles was most successfully met in the early 20th century by the Serbian mathematician Milutin Milankovitch, whose name still echoes, under the title 'Milankovitch Cycles,' in the contemporary geophysical literature. Given the then-prevailing assumption (later challenged) that all geocentric energy originates with the sun, he was able to calculate, by the use of recondite mathematical models that nevertheless ultimately won wide if not universal acceptance, that there was a correlation between the effects of precession and cycles of variation in the earth's axis tilt vis-à-vis the sun and that this correlation could account for ice age cycles. His critical assumption (analogous to Newton's concentration of the mass of celestial bodies upon geometrical points) concerned the use of temperature data, that they could be focused upon a point, namely, mid-summer measurements, reasoning that this was when winter ice and snow would, or would not, melt. Using the data then available, his model came to persuade a critical mass of fellow scientists that there was a drop in solar radiation reaching the earth by mapping the extremes of the precession cycles upon the earth's axial tilt cycles in a way that seemed to account for the advent and recession of ice ages. Hence the canonical references in the current literature to 'Milankovitch Cycles.'

So the first major challenge to mechanism, Agassiz's ice ages, seemed to have been met within the framework of an only somewhat modified model of Newtonian mechanism. The limits of mechanism had been extended, but not by much.

C. The First Transition from Mechanism to Interaction

As the topic of an on-going interdisciplinary and international research program, climate science only came of age in the years following World War II. In the intervening half century vast strides have been made toward the comprehension of climate as a system that extends well beyond the traditional meteorological focus upon measurable aspects of such cyclic processes as the diurnal, the lunar, the annual, the cometological, and the eleven-year cycle of sun spot phenomena. As long as these processes were thought to be determined by forces external to the sphere of the observed phenomena affected, e.g., atmospheric weather, *the causal relations described were accordingly thought to derive from agencies independent of the processes for which their causal efficacy was invoked as an explanation. Hence these explanatory models were* **genetic** *or, at best,* **mechanical**. The transition from explanatory models based upon external causes, as in weather studies, to the study of climate has been marked by the discovery of intra-systemic or cybernetic feedback processes, facilitating a move toward **theoretical** patterns of explanation. Perhaps the most notable of these has been the detection of carbon dioxide density cycles in the earth's atmosphere which correlate with ice age cycles of 100,000 years (with subordinate cycles) as evidenced by the periodic deposits of carbon in rocks and in polar ice. The basic idea is that greater carbon dioxide concentrations in the earth's atmosphere 'cause' a 'greenhouse effect.'

Once again, as with Agassiz, the initial step toward a new way of thinking about climate emerged with the accumulation of 'natural history' data. This time, around 1976, a new source for the historical record was discovered in the shells of micro-organisms found as a sediment at the base of the oceans. These, called 'forams,' were found to register the cycles of heat and cold correlated with the coming and passing of ice ages. By this means the evidentiary basis for the duration of ice-age cycles was extended to 450,000 years, including reference to subordinate cycles that embraced as few as several thousand years. Thus the 'natural history' data from which climate science had to work became, in a stroke, much more fine grained. The first response to this data was to greet it as empirical confirmation of the mathematically-formulated Milankovitch Cycles, classically, i.e., Newtonianly, understood to be a *formal* hypothesis waiting for empirical, i.e., *material*, confirmation or fulfillment. But it was not to be so.

An ocean chemist, Wallace Broecker, acknowledging that the hydrosphere contains far less carbon than the lithosphere, nevertheless pointed out that the quantity of carbon in the hydrosphere is fifty times that contained in the atmosphere and that these two latter spheres are in interaction. Since the carbon content of the atmosphere is the factor determining the greenhouse effect, he reasoned that the hydrosphere could easily serve as a sink regulating the carbon content of the atmosphere. Hence the notion of the hydrosphere, including especially its organic components, as a regulator of the carbon cycle. The surface waters of the sea are able to absorb carbon dioxide from the atmosphere until the sea reaches a saturation point. This point is governed by the capacity of ocean plant life to consume carbon. With more plant life there would be a greater capacity of the ocean to absorb atmospheric carbon. The result would be less carbon in the atmosphere and a consequent lowering of the greenhouse effect, providing the terrestrial sphere with a plausible cooling, thus accounting for ice age cycles. The argument is simple: more plants, less heat; fewer plants, more heat. The question then turns from the mechanical cycles known to orbital physics, upon which the Milankovitch cycles were constructed, to the exchange of carbon between the atmosphere and hydrosphere, with the mass of plant life as the independent variable.

While it might be argued that the factor controlling the ebb and flow of ocean plant life is, after all, the influx of solar radiation governed by the aspects of orbital physics ingredient in the Milankovitch cycles, that argument becomes moot, i.e., one which must assume the burden of proof, once the key factor controlling solar transmission to planet earth is assumed to be the relative presence or absence of greenhouse gases in the atmosphere. The point at issue, within this model of explanation, has clearly become what governs the growth and diminution of plant life within the hydrosphere and for this the factors of orbital physics remain but candidates for an explanatory role. For the model of explanation has shifted from (a) the consideration of simply mechanical factors governing the transference of solar radiation to the atmosphere to (b) the consideration of organically mediated boundary conditions for such transference.

Since the organism in question is the terrestrial system as a whole, the governing conditions concern *interactions* within the geosphere rather than the simple correlation of terrestrial phenomena, e.g., ice ages, with the formal structures of orbital physics. This model still assumes that the ultimate source of climatic change is heliocentric, but it introduces the non-mechanistic notion that such change is importantly mediated by interactive processes within the earth system.

D. The Second Transition from Mechanism to Interaction

All attempted explanations of ice age cycles considered thus far have retained a top-down model. That is, all have assumed that the ultimate source of heat/energy has been solar radiation, whether mediated by the long-term effects of orbital physics or by the organically mediated carbon cycles governing the transfer of solar heat/energy to the geosphere and its retention there. What none of these models take into account is the possibility that the geosphere itself might be an energetic or thermal source.

That the inner earth is a source of thermal energy has been known as long as volcanoes have been observed, and especially since the advent of *material* equilibrium models of **genetic** explanation among the Greek physicists of the 6th century BC. Early accounts nevertheless remained mythopoeic, tracing the source of volcanism to the iron-forging activities of Hephaestus, who was said (e.g., by Thucydides, 3.88.3) to work in the underworld of the Aeolian archipelago. Operating with a very limited data base, Aristotle marks a transition from mythic explanation when he speculated that volcanoes derive from winds that somehow enter the subterranean chambers and thus generate heat by friction. This view remained dominant for many centuries. In the 18th century, at the outset of the Industrial Revolution, it was natural to suppose that the heat driving volcanism was generated by the intra-terrestrial coal furnaces, on the analogy of the power that drove the factories of contemporary England. At mid-20th-century, awed by Hiroshima, speculation turned to a thermonuclear explanation of volcanism, an explanation still to be found in some reference books at the end of the last century.

The first step toward the explanation that has most recently arisen was made with the discovery, in 1936, that the earth has an inner core. In 1971 it was determined that this inner core is solid and about the size of the earth's moon. Over the past decades shock waves emanating from earthquakes, some of which pass through the inner core, have been monitored at stations around the earth. The resultant data have posed an anomaly. North-pole to south-pole shock waves are transmitted at a more rapid rate than other trans-terrestrial shock waves. This difference, or anisotropy, has suggested that the spin rate of the inner core differs from that of the earth's surface-lithosphere-mantle. That difference, measured at the earth's surface, may amount to less than a km per day, but, given the intensity of the respective electromagnetic fields, it is enough to produce a powerful geodynamo effect.[2] The inner core functions within this dynamo as a coil moving within the magnetic field defined by the mantle and lithosphere. The greater spin rate of the inner core vis-à-vis the earth's surface may derive from an inertially retained angular momentum of the earth's original spin rate (or time-of-day, at ca 5 to 6 hours), meanwhile slowed, over eons, at the surface by lunar and solar tidal drags upon the earth's oceans. Given the liquidity of the outer core, this is not entirely implausible.

At any rate, there are grounds for presuming a differential rotation between the inner core and the earth's mantle/lithosphere-hydrosphere. That would, given the magnetic fields associated with each, account for an earth-endogenous energy source: a geodynamo. But the explanation, as thus far exposed, would only account for a relatively steady evolution of geodynamic output (governed by the very gradual retardation of the earth's relative surface spin rate vis-à-vis the core, as modulated by lunar and solar tides) whereas observed data clearly suggest more radical variations in geothermal energy generation, especially as exhibited by volcanic data that cannot be uniquely accounted for by such an hypothesis.

---

[2] Goa, Tang, Gregori, and Dove, "The Role of the Geosphere in Climate Change," forthcoming.

Data correlations have nevertheless been found that may account for this anomaly. They concern the cycles of sunspot activity. It has long been known that these occur over eleven-year periods. However, the relevant period for geothermal energy generation, and for volcanic activity, turns out to be 22 years.

Now, prima facie, it is not implausible to speculate a correlation between the output of the geodynamo sketched above, given its dependency upon a contrasting electromagnetic field and coil, and sunspot activity, which affects the heliomagnetic field pervading both the intra-terrestrial magnetic field and the rotating magnetic coil that govern the geodynamo. The problem is that natural history data correlating with the heliomagnetic effect seem to be spread out over a period, 22 years, that is twice the length of the observed periodicity of sun spot activity, 11 years. How account for that difference? One plausible solution has been proposed by Antonino Palumbo.[3] It appeals to the fact that the earth's lithosphere/mantle functions as a resistor to heliomagnetic intrusion into the geomagnetic field of the geodynamo. Slowing the absorptive capacity of the earth to heliomagnetic variation, that resistor effect may well reduce the rate of geomagnetic reception of heliomagnetic variations by a half. Hence the 22-year, as opposed to 11-year, cycles.

So far as I have been able to determine, this defines the state of the art in regard to the geodynamic hypothesis. Clearly it is still in its speculative phase, but it holds promise of providing the most theoretically satisfying answer to the question of long-term climate change.

Let us now consider how these developments measure up against the canon of scientific principles laid down by Aristotle. Quite apart from their capacities to 'save the phenomena,' in this case natural history data pertaining to ice age cycles, or to predict phenomena, e.g., volcanic eruptions, it is quite clear that the geodynamic explanation as well as the carbon-cycle model involve consideration of processes for which three, as opposed to two or merely one, principles must be presupposed. In the carbon-cycle model the process in question is the ebb and flow of greenhouse gases regulating the influx of solar radiation. This process, albeit merely governing a genetically conceived exogenous energy-input, is nevertheless thoroughly immanent in the terrestrial system. To consider any phase of this process one must take into account the stuff (material principle) capable of undergoing the process as well as the determinate state (formal principle 1) at the outset and the determinate state (formal principle 2) at the end of the phase of the process under consideration.

If reliance upon a three-principle explanation is evident for the carbon-cycle model, it is even clearer in the case of the geodynamic model. For here the intra-terrestrial process is the source of the energy (and, via Joule's law, the heat) in question and the variations in the heliomagnetic field, based upon the earth's capacity to receive the effects of sun spot activity, are *merely regulatory*.

Much more natural history data will have to be gathered to assess the ultimate significance of the carbon-cycle and the geodynamic models for the science of climate. But as long as such data *are* considered within either of these models, it is clear that three principles will have to be invoked. At each step along the way the science of climate may uncover new limits of the two-principle mechanical model of explanation which has served as a tacit paradigm for centuries.

KENLEY R. DOVE

Venice

---

[3] Antonino Palumbo, "Solar and Volcanic Sources of Climate Change," forthcoming.