

**THE NATURE AND THEORY
OF THE GENERAL CIRCULATION
OF THE ATMOSPHERE**

BY

EDWARD N. LORENZ

WMO - No. 218. TP. 115

**WORLD METEOROLOGICAL ORGANIZATION
1967**

CHAPTER I

THE PROBLEM

I think the causes of the General Trade-Winds have not been fully explained by any of those who have wrote on that Subject...

George Hadley (1735)

We have chosen the opening words of Hadley's famous paper for the opening words of this monograph because they seem to afford an apt description of the state of the same subject today. We have no desire to imply that tremendous progress has not been made, because, in the light of today's knowledge, Hadley's remark appears to be a considerable understatement. Yet not in any of the thousand or more excellent works which have appeared since that time, nor in any combination of these works, is a full explanation of the distribution of easterly and westerly winds to be found.

It is evident that the validity of this claim depends very much upon what constitutes a full explanation. It is not to be expected that there will ever be complete agreement on this matter. At this point we shall simply express the opinion that the requisites for a complete answer to a qualitative question differ considerably from those for a complete answer to a quantitative question. Before considering this matter in greater detail, we shall present an account of Hadley's paper, which will serve to illustrate some of the points involved.

Prior to Hadley's time there had been sporadic attempts to account for the trade winds, and one of these which pictured the winds as exhalations from the sargassum weed in the subtropical seas nevertheless found its way into a scholarly journal. In sharp contrast was the notable work of the astronomer Edmund Halley (1686), who presented a detailed and methodical account of the trade winds as observed in three separate oceans, and sought a common cause for them. He rejected an earlier notion that the air by reason of its lightness simply could not keep up with the Earth's surface in its diurnal rotation, and ascribed the north-easterly trades on the north side of the Equator and the south-easterly trades on the south side to the tendency of the air to converge toward the most strongly heated region, as this region progressed about the equatorial belt. For reasons which are not clear he assumed that the cumulative effect of the afternoon tendency to move toward the western sun would outweigh that of the morning tendency to move toward the east.

In concordance with Halley, Hadley concluded that the distribution of solar heating would lead to a general rising motion in lower latitudes and a sinking motion in higher latitudes, the circuit being completed by equatorward motion at low levels and poleward motion aloft, but he rejected the idea that motion toward the sun would lead to any average westward or eastward movement. He then noted that in the absolute sense the Earth's surface moves most rapidly eastward at the lowest latitudes, and he maintained that if the air were initially moving equatorward with no relative eastward or westward motion it would, in attempting to converse its absolute velocity, arrive at lower latitudes moving westward relative to the earth. He found, in fact, that air travelling considerable distances would acquire a much greater westward velocity than any ever observed, and assumed that the frictional drag of the

Earth's surface would in the course of a few days reduce the velocities to those actually found — thus the trade winds.

He next noted that the required counter-drag of the air upon the Earth would continually slow down the Earth's rotation unless opposed by an opposite drag in other regions: this he assumed to occur in the belt of prevailing westerlies in middle latitudes. To account for the westerlies he maintained that the air initially moving directly poleward at high levels would soon acquire an eastward relative velocity, and upon reaching higher latitudes and being cooled would sink and become the prevailing westerlies.

Although Hadley's remarkable paper contains scarcely a thousand words, many hundred thousand words have since been written about it, and it is not surprising to find that some of these have criticized it adversely. One fault requires immediate correction: in the absence of eastward or westward forces, air moving equatorward or poleward conserves its absolute angular momentum rather than its absolute velocity. This tendency to conserve angular momentum is identical with what is now designated as the east-west component of the deflective force, whose proper formulation has been credited to the nineteenth-century mathematician Coriolis among others. But Hadley preceded Coriolis by a century, and perhaps he deserves credit for being as nearly correct as he was. Hadley's error caused him to underestimate the Coriolis force by a factor of two, but since the remainder of his argument was entirely qualitative, his error did not influence it.

Far more significant are his positive contributions. Hadley realized what today seems fairly obvious, that, by reason of continuity of mass, general equatorward motion at one level requires general poleward motion at some other level; and, what is less obvious, that, by reason of conservation of total angular momentum, general westward motion dragging upon the Earth's surface at one latitude requires general eastward motion at some other latitude. His ideas embody the concept of a global circulation, no one of whose major branches can be explained independently of the remaining branches.

Hadley stated that he felt it unnecessary to consider the changes in solar heating with the seasons, and he rejected the diurnal variations of heating, which had played a dominant role in Halley's hypothesis, as having any important effect. He did not consider the presence of oceans and continents, whose contrasting thermal capacities could have destroyed the symmetry of the heating, nor the mountains and other obstacles which could have distorted the flow. He did not consider the presence of water vapour, whose thermodynamic properties were in any event not known in his day. Had he been questioned on these omissions, he might have maintained that these influences would alter the flow to some extent, but not so greatly as to render his arguments invalid.

Many theoreticians today would take a different attitude. They would maintain that what they were studying was not the Earth's atmosphere at all, but an idealized atmosphere, consisting of a gas of uniform composition enveloping a planet with a level homogeneous surface, and driven by an external heat source not varying with longitude or time. They would regard the Earth's atmosphere as only one of many conceivable planetary atmospheres, which in turn comprise but one type of many conceivable types of thermally driven rotating fluid systems. Certainly the general theory of planetary atmospheric circulations is as suitable a subject for theoretical study as the specialized theory of the circulation of the Earth's atmosphere. Moreover, although one cannot deny that simplifications are often made solely to facilitate theoretical treatments, it would appear that, within the collection of possible planetary atmospheres, one which is devoid of irregularities occupies a more central and fundamental position than one with any specific arrangement of irregularities.

It is noteworthy that Hadley adopted an approach which has characterized numerous subsequent attempts to account for the atmospheric circulation, not to mention many other natural phenomena.

He attempted to describe how the final steady circulation which he envisioned would have developed from a previous simpler circulation which lacked the specific features whose development he wished to account for. In his case the simpler circulation was the one which he assumed would prevail in the absence of rotation. In many subsequent studies it has been a state of rest.

Hadley has been criticized for disregarding the north-south component of the Coriolis force altogether, and it is unlikely that he was aware of its existence. Consideration of this force would have been useless, in any case, in an argument making no reference to pressure. As a consequence he apparently supposed that the vertical and meridional (north-south) motion would not change during the development of the zonal (east-west) motion, and his task of describing the development was relatively simple. In reality, as soon as zonal motion has been produced by the deflection of the initial meridional motion, additional meridional motion will be produced by the deflection of the zonal motion, whereupon additional zonal motion will be produced by the deflection of the additional meridional motion while additional north-south pressure gradients will concurrently be produced by the convergence and divergence of the additional meridional motion. Both the additional pressure gradients and the deflection of the additional zonal motion will produce further additional meridional motion, etc., and it is reasonable to conclude that Hadley would have had a difficult time in carrying his argument to completion. Indeed, it is difficult to see how any argument of this sort, involving two or more processes whose effects may alternately combine and cancel, and requiring more than two or three steps, can be carried to a successful conclusion unless it is made quantitative, so that the accumulated changes of each quantity can be properly recorded. In this event the argument is converted into a stepwise numerical integration. Recently such integrations have been widely used with excellent results, but they often require hundreds of steps for completion.

A modern theoretician attempting to reproduce Hadley's description of the development of the trade winds in a rigorous quantitative fashion would in fact find that many years would be needed for the circulation to become nearly steady if he represented the effect of friction through a coefficient of molecular viscosity. To achieve a steady circulation within a few days he would be forced to introduce the much larger coefficient of turbulent viscosity. Use of this coefficient can be justified only in combination with a further idealization.

It is utterly impracticable to describe every gust of wind or even every cumulus cloud occurring at a particular time, even if the description is to appear only in the memory of the largest existing digital computer. It is therefore customary in problems of global scale to define the circulation as a smoothed circulation, from which motion systems of thunderstorm size or less have been subtracted. Meanwhile the effects of these systems cannot be disregarded. Ordinarily it is postulated that the statistical properties of the small-scale motions can be described in terms of the smoothed circulation, although really suitable formulae which accomplish this have yet to be established. The simplest way to represent these properties is through the use of coefficients of turbulent viscosity and conductivity, which may exceed the corresponding molecular coefficients by a factor of 10^5 or more. Qualitatively, this idealization treats the atmosphere as a highly viscous, highly thermally conductive fluid.

Evidently Hadley unknowingly used this idealization in his argument, since he assumed that the trade winds would be reduced to their observed velocities within a few days. It is interesting to speculate as to whether, in an atmosphere with very high molecular viscosity and conductivity but otherwise like the Earth's atmosphere, the troublesome small-scale motions would actually fail to develop. If this is the case, the present idealization, like the ones previously described, replaces the Earth's atmosphere by a physically conceivable system.

In any event, in a comprehensive study of what is known about the global atmospheric circulation, it is necessary to recognize both the real and the idealized atmospheres. The idealized atmosphere has

formed the subject of the great majority of theoretical studies. The observations needed to confirm the results of these studies have of necessity been restricted mainly to the real atmosphere. Since the two atmospheres are not the same, certain discrepancies between theory and observation are inevitable.

It is remarkable that a few changes in wording, entailing, however, a considerable change in approach, would have eliminated all the shortcomings of Hadley's work thus far mentioned. Hadley sought a steady-state circulation, independent of longitude. In such a circulation there must be at least one latitude, separating low-level easterlies from low-level westerlies, where the flow is directly toward the Equator. If Hadley had referred to a particular parcel of air crossing this latitude at some initial time, instead of referring to an initial circulation where all the air flowed directly equatorward or poleward, his ensuing sentences would have formed a qualitatively acceptable account of the nature and maintenance of the steady circulation which he envisioned.

Hadley's only fault which cannot be remedied by a slight rewording of his arguments is less obvious, and it lies in his original assumption about the vertical motions. It can be shown that in a thermally forced system the temperature and the upward motion are positively correlated, but the correlation need not be perfect nor even very high. Hadley assumed in essence that all of the air would rise in low latitudes and sink in high latitudes. From this point on, barring further errors in reasoning, he was forced to obtain the picture of the circulation which he did. Observations which were unavailable in the eighteenth century but have since become superabundant reveal that this picture is incorrect. Yet it is within the realm of possibility that there somewhere exists a planet whose circulation conforms by and large to Hadley's picture. Such a circulation, whether real or hypothetical, is now known as a Hadley circulation.

If such a planet exists, Hadley's work, with the indicated changes in wording, is not only a description of the circulation there but also an essentially correct account of the basic reasons why this circulation occurs. Yet it is in no way a demonstration that the envisioned circulation must take place in preference to some other one. It lacks quantitative considerations, and on a qualitatively similar but quantitatively different planet there are alternative possibilities, one of the more obvious being the type of circulation which actually occurs on the Earth. Stated otherwise, Hadley's work lacks mathematical rigour. For this reason, we cannot look upon it as a full explanation.

A demand for mathematical rigour is not a demand for mathematical symbols and formulae. It is perfectly possible for a purely verbal argument to be mathematically rigorous. But, particularly when the argument is very complicated, a non-rigorous qualitative approach offers numerous opportunities for errors in reasoning. One of the best ways to avoid such errors is to formulate the problem in mathematical symbols, and manipulate these symbols according to established procedures.

What, then, constitutes a full or complete explanation? This depends upon whether the question being answered is qualitative or quantitative.

Consider, for example, the problem of explaining why the average surface wind at latitude 20°N is directed from 15°N of E at 5 metres per second (or whatever the exact direction and speed may be). The wind is influenced by the field of pressure, which in turn is influenced by the field of temperature. Certainly then the precise wind velocity depends upon the precise amount of energy received from the sun, and upon the precise values of the physical constants which characterize the Earth and its atmosphere. Water vapour and liquid water ultimately affect the wind velocity by altering the thermodynamic properties of the atmosphere, and perhaps to an even greater extent by influencing the fields of incoming and outgoing radiation. The distribution of water in the atmosphere is in turn affected by the locations of oceans and continents, and of course by the field of motion itself. If all the relevant physical factors

could be properly incorporated into the governing equations, and if the equations could be solved in a rigorous fashion, the proper numerical values would be found. The observed wind velocities would then seem to be completely explained, whereas no simpler procedure could be expected to give the correct result.

A correct answer to the quantitative question of why the wind blows from 15°N of E at 5 metres per second is of necessity an answer to the qualitative question of why the wind blows from a general easterly direction, but it may not be a very satisfactory answer. It may not indicate which of the many physical factors involved are needed to bring about the easterly wind, and which are mere modifying influences. In short, it may fail to answer the more general question as to why planetary atmospheres sharing certain features with the Earth's atmosphere possess easterly surface winds at low latitudes.

This difficulty need not arise if an analytic expression for the wind velocity in terms of the various physical constants has been found, but analytic solutions of meteorological equations are rather rare. If the solution has been obtained by numerical means, it would have to be repeated many times, with different values of the constants, in order to apply to the general case. At best this would be an extremely roundabout way of obtaining a desired answer which is not quantitative at all.

Moreover, even if the irrelevant physical factors are all eliminated, and a rigorous solution of the resulting simplified equations is obtained, the reader who has followed the demonstration from beginning to end may still gain little physical insight as to why easterly winds must exist, particularly if the demonstration is complicated or lengthy, or depends upon mathematical theorems whose proofs he does not recall or understand. Whereas a lack of rigour may lead to incorrect results, rigour alone does not guarantee understanding. An argument of the type presented by Hadley, if correct, may well prove more satisfying. Thus an acceptable answer to a qualitative question may well be more difficult to produce than an acceptable answer to a quantitative one.

Both quantitative and qualitative questions concerning the global circulation frequently arise. The most complete answer to the problem should therefore consist of a rigorous quantitative solution of the governing equations, yielding the observed circulation, together with a qualitative and possibly verbal explanation of the basic reasons why the principal qualitative features occur. In this event the qualitative explanation need not be rigorous, but it should be correct, and it must certainly be consistent in every respect with the quantitative solution which it accompanies.

From what has been said it appears that the motion of the atmosphere cannot be explained without full consideration of the accompanying fields of pressure, temperature, and moisture, and that these fields in turn cannot be explained independently of the field of motion. Such a statement cannot be made for all fluid systems. The future motion of a homogeneous incompressible fluid, for example, is completely determined by the present field of motion together with the external mechanical forces, and the circulation of such a system may be regarded as synonymous with the field of motion. In the case of the atmosphere it is more logical and certainly more convenient to regard the circulation as consisting of the field of motion together with the accompanying fields of the remaining meteorological variables.

The question naturally arises as to why no complete explanation of the global circulation has yet been produced. As already noted, the laws governing the real atmosphere are very complex, and are not perfectly known. We shall attempt to show now why the circulation of even the idealized atmosphere has yet to be fully explained.

The equations governing the idealized atmosphere appear to possess a steady-state solution which is also independent of longitude; this solution describes the Hadley circulation. If all other particular solutions could be shown to converge toward this solution, the problem of determining the circulation

would be simply the problem of finding this solution. The determination of steady-state solutions of various systems of equations is one of the more frequently encountered problems in fluid dynamics.

When the general solution does not approach the Hadley solution asymptotically, the equations are likely to possess periodic solutions. Again, if all particular solutions, excluding those exceptional ones which converge toward the Hadley solution, could be shown to converge toward the periodic solutions, the problem of determining the circulation would reduce to the problem of finding these solutions.

Observations reveal, however, that the behaviour of the real atmosphere is neither steady nor periodic. Theoretical studies imply that the idealized atmosphere is likewise non-periodic; indeed, if the atmosphere has been idealized to the extent that it is forced to behave periodically, it has probably been over-idealized for the present purposes. Except in special instances it is not possible to express a complete non-periodic solution, even approximately, with a finite number of symbols, and the goal of determining the complete life-history of the idealized atmosphere must be abandoned.

This state of affairs is brought about by the non-linearity of the equations. Among the non-linear terms are those representing advection — the displacement of the field of motion, temperature, water vapour, or some other quantity, by means of the field of motion itself. In a sufficiently idealized atmosphere with crudely represented heating and friction, advection is the only non-linear process. Since the motion which brings about the displacement is generally not uniform, different portions of the field of each displaced variable undergo different displacements, and the field as a whole is distorted as well as displaced. Under continual distortion it may soon acquire a shape bearing little resemblance to its earlier configurations, and possessing much fine detail. With such an infinite variety of shapes there is no need for a pattern ever to repeat itself in all its features simultaneously, and the circulation need not vary periodically.

Yet non-linearity does not assure non-periodicity. The number of possible circulation patterns, none of which bears any resemblance to any of the others, is limited, and ultimately a pattern must occur which resembles a previous pattern rather closely, particularly in its coarser features. If the further evolution of the pattern is stable, in the sense that small differences between separate solutions of the equations will not amplify, the previous history will tend to repeat itself and the pattern will continue to recur at regular intervals, at least in an idealized atmosphere where the external conditions are steady. If, instead, the behaviour is unstable, approximate repetitions of previous history will ordinarily be only temporary, and periodicity need not develop.

Since it is not feasible to determine the complete history of the circulation theoretically, we must turn our attention to slightly less ambitious problems. One of these is the problem of explaining each pattern in a long but finite succession of circulation patterns; this in essence is the problem of long-range forecasting. A different problem, and the one with which this monograph is concerned, is that of explaining the characteristic properties, or statistics, of the collection of all circulation patterns which ever occur.

The equations governing the circulation are most readily written in a form expressing the time-derivative of each atmospheric variable — velocity, temperature, water-vapour content, etc. — in terms of the current values of the same set of variables. They do not directly account for any particular circulation pattern, except in terms of some other pattern which has just occurred or is just about to occur. It is as though the laws had been created for the convenience of the weather forecaster.

But the problem of determining long-term statistics is not the problem of weather forecasting, even though the governing equations may be the same. The latter is strictly an initial-value problem; the former does not *a priori* involve any initial values, even though initial-value procedures sometimes offer the only tractable means of solution. Whereas the latter is strictly a problem in differential equations,

the former is a problem in ergodic theory, which is concerned with long-term statistical properties of solutions of equations.

The results of ergodic theory do not assure us even of the existence of long-term statistics, since there are systems of equations for which the average values of particular solutions over long intervals do not converge to any limit as the period of averaging becomes infinite. Assuming that the atmospheric equations are not of this peculiar and possibly exceptional type, each particular solution possesses its own long-term statistics, but there is no assurance that different solutions possess the same statistics. For a large class of systems of equations, however, there is only one set of statistics which a randomly chosen particular solution has a greater-than-zero probability of possessing. Such systems are called *transitive*. A transitive system may possess in addition any number of particular solutions having different sets of statistics, but the probability that a randomly chosen solution possesses one of these sets of statistics is zero (in the same sense that the probability is zero that a number chosen at random from the set of real numbers between zero and one will be a rational fraction). For example, in an atmosphere whose general solution is unsteady, the probability of choosing at random a solution which asymptotically approaches the Hadley solution is zero. If two or more sets of statistics have greater-than-zero probabilities of being chosen at random, the system is called *intransitive*.

Ergodic theory has not yet provided us with a general rule for determining whether a given system is transitive or intransitive. We therefore do not know whether the atmosphere is capable of possessing more than one set of statistics. Lest it appear implausible that the atmosphere could actually behave in an essentially different manner from what is observed, let us note that certain laboratory systems designed to simulate the atmosphere have proven to be intransitive. Unfortunately for our understanding of the atmosphere, but perhaps fortunately for the continuation of the human race, we cannot halt the atmospheric circulation and then see whether it will redevelop in a different manner.

Assuming that the atmosphere is transitive, we must then decide which statistics ought to be determined. There is no hard-and-fast rule, but the long-term time-averaged circulation, or more specifically the limiting form of the time-averaged circulation as the period of averaging approaches infinity, might be regarded as a minimum requirement. Undoubtedly this average circulation has received the most theoretical attention in recent years.

Yet time averages *per se* are not necessarily the statistics of greatest interest. Perhaps the average circulation is of more interest as a first approximation to the particular circulation to be expected at any given time. The trade winds, for example, are so persistent that an explanation of the time-averaged trades might be considered tantamount to an explanation of the time-variable trades. The upper-level westerly flow in middle latitudes, while less persistent, is still far more than a mere statistical residual.

Other regularly occurring features are poorly represented by time averages of the meteorological variables. Nothing indicates the frequency or even the presence of migratory cyclones and anticyclones. The jet stream appears only in attenuated form, and most of its familiar meanders are lacking.

All of these features are indicated by suitably chosen statistics, and hence by the collection of all long-term statistics. This collection includes such quantities as joint probability distributions, and it is of course impossible in practice to explain all of these, simply because an infinite amount of labour would be needed. Conceivably it might be possible to explain any particular statistic. Yet not even the long-term time-averaged circulation has thus far been fully explained.

The underlying difficulty is again the non-linearity. By rendering the general solution of the equations non-periodic, non-linearity makes it impossible to solve the equations by analytic methods and then obtain statistics by integrating with respect to time.

The most feasible method of solving non-linear equations with non-periodic solutions is as an initial-value problem by numerical means. This method yields finite segments of particular solutions. Statistics may be easily evaluated from these.

Such segments possess one of the principal disadvantages which characterize real meteorological data; they are finite samples from a population, and are not necessarily representative. The best method of assuring reasonably representative results is to extend the solution over a long time-interval, but this method may entail a prohibitive amount of computation.

More than any other theoretical procedure, numerical integration is also subject to the criticism that it yields little insight into the problem. The computed numbers are not only processed like data but they look like data, and a study of them may be no more enlightening than a study of real meteorological observations.

An alternative procedure which does not suffer this disadvantage consists of deriving a new system of equations whose unknowns are the statistics themselves. This procedure can be very effective for problems where the original equations are linear, but, in the case of non-linear equations, the new system will inevitably contain more unknowns than equations, and can therefore not be solved, unless additional postulates are introduced.

Moreover, even if the new system of equations could be solved, it would not necessarily yield the desired result. The separate solutions of the new system would include the statistics of all solutions of the original system. The statistics of the Hadley solution could perhaps be recognized as such and eliminated, but there would remain the statistics of an infinity of periodic and otherwise special solutions.

The separate solutions of a system of equations whose unknowns are statistics will therefore show nearly as wide a variety as the statistics evaluated from separate finite segments of solutions of the original equations. For example, there are presumably special periodic solutions representing circulations which are permanently of the "high-index" or "low-index" type, with well developed or poorly developed middle-latitude westerlies; there are presumably a great many more special periodic solutions which oscillate between high-index and low-index régimes, but do not divide their time between the régimes in the same proportion as does the general solution. The statistics of these special solutions are included among the solutions of the new system of equations.

In short, the only presently feasible procedure for determining quantitative statistics consists of evaluating them directly from particular time-dependent solutions of the original equations, and the only known procedures for solving these equations are numerical. Even these procedures are feasible only because of high-speed computing machines. Ultimately with the development of much larger and faster computers it may become possible to estimate the statistics of the general solution with a high degree of precision, even for the real atmosphere, although the proper representation of the effects of small-scale systems may prove to be a stumbling block. At present the procedure is limited to a rather idealized atmosphere. Moreover this procedure, being numerical, is of the type which contributes least to a qualitative understanding of the circulation.

There remains the possibility of rigorous procedures which are not quantitative at all. Any qualitative statement about the circulation may be formulated as a mathematical inequality; for example, the statement that the trade winds blow from a general easterly direction is equivalent to the statement that the eastward wind component in these latitudes is less than zero. There is no difficulty in deriving various incomplete systems — systems with more unknowns than equations — whose unknowns are statistics. Sometimes enough inequalities connecting the statistics may be established to complete the system. In this event it may be possible to solve the system of equations and inequalities for upper and

lower bounds of the statistics, and thereby obtain qualitative descriptions of certain features of the circulation.

Possibly the relevant systems derivable in this manner are intractable. We feel, however, that the current failure to have obtained a qualitative explanation through this procedure must be attributed mainly to failure to have exploited the procedure.

If the causes of the circulation have not been fully explained, what can be the nature of the thousand or more excellent studies previously alluded to? Some of these have dealt principally with observations, thereby providing a better picture of the phenomena to be accounted for. Some have sought to reproduce the circulation or some of its features by means of laboratory models, or with the aid of electronic digital computers, thereby making it possible to perform controlled experiments. Some have aimed to establish relationships between various features of the circulation by analytical means. Some have presented comprehensive assessments of the current state of progress. In the following chapters we shall examine some of these studies, and attempt to identify the contributions which they have made to our present understanding of the problem.