TECHNIQUES OF PREDICTION AS APPLIED TO THE PRODUCTION OF OIL AND GAS

M. King Hubbert*

Abstract

Techniques of prediction of future events range from the completely irrational to the semi-rational to the highly rational. Rational techniques of predicting the behavior of a system require first an understanding of its mechanism and of the constraints under which it operates and evolves. This permits the development of appropriate theoretical relations, which, when applied to the data of the system, permit solutions of its future evolution with varying degrees of exactitude. Such a theoretical analysis provides an essential criterion for what data are significant and necessary for the solution.

In the case of the production of petroleum in a given region, present geologic knowledge indicates that oil and gas accumulations occur in limited volumes of underground space, in porous rocks normally filled with water, within or immediately adjacent to basins of sedimentary rocks. These accumulations have resulted from plant and animal material accumulated in the sediments during the last 600 million years, at rates so slow that no significant additions can occur during the period of oil and gas exploitation. Therefore the exploitation of the oil and gas accumulations in a given region represents the continuous reduction of the amounts originally present and is a unidirectional and irreversible process, characterized by a definite cycle of events. Oil production in a given region begins with the first discovery at time t_0 and ends finally at a later time t_k. Hence the cumulative production Q will have the value 0 at $t = t_0$, and a definite finite value Q_k at $t = t_k$. Thus, during the complete cycle of production, Q increases

---

*Consultant, 5208 Westwood Drive, Bethesda, MD 20816.

monotonically during the time interval t_o to t_k, being 0 for $t < t_k$ amd the constant Q_∞ for $t > t_k$. Also, during this cycle the production rate, dQ/dt, will vary as follows:

For

$$\left.\begin{array}{r} t < t_o,\ dQ/dt = 0; \\ t_o < t < t_k,\ dQ/dt > 0; \\ t_k < t,\ dQ/dt = 0. \end{array}\right\} \tag{1}$$

Therefore

$$Q_k = \int_{t_o}^{t_k} (dQ/dt)dt = Q_\infty, \tag{2}$$

where

$$Q_\infty = \int_{-\infty}^{+\infty} (dQ/dt)dt. \tag{3}$$

Because of the definiteness of the limits of Q as compared with those of time, it is more useful to consider the production rate, dQ/dt, as a function of Q rather than of t. Then

$$dQ/dt = f(Q), \tag{4}$$

and the integration of this equation gives the corresponding functions of time, $Q(t)$ and $(dQ/dt)\ (t)$. Comparable relations pertain to other variables of this system such as the rates of discovery and cumulative discoveries, proved reserves, and the rate of exploratory drilling and cumulative drilling. By developing the appropriate equations among these variables, and supplying the data from petroleum-industry statistics, it becomes possible, after production in the region has passed through about the first third of its cycle, to determine with reasonable accuracy the principal constants of the equations: Q_∞, the ultimate cumulative production; various exponential growth constants; various critical dates of the cycle, such as that of the maximum rate of production.

Methods based upon this type of analysis, which have been developed and used by the present author during the last 25 years, have consistently given predictions of the future courses of oil and gas production in the United States which have agreed within narrow limits with what has subsequently occurred.

* * * * * * * * * * * * * * * *

Techniques of Prediction

Soothsaying, while probably not the world's oldest profession, can certainly offer a strong claim for being its second oldest. The techniques of soothsaying may be divided roughly into those that have some rational basis, and those that do not have; and the result of soothsaying, namely the prediction of some event, may be expressed in language whose meaning ranges anywhere from completely indefinite to precise and unambiguous.

The nonrational techniques of prediction are well exemplified by the activities of many of the priestcrafts of the ancient world - notably those of the famous Oracle at Delphi in Greece - and by those of the great variety of fortune tellers of today with their tea leaves, crystal balls, or astrological interpretations of the human consequences of various planetary configurations. These have usually been characterized by a combination of astute guesswork and ambiguous statement, conducted behind a facade of mystical rites.

From examples of the soothsayer's art which have been handed down from antiquity, it appears to have been learned very early in human history that a soothsayer's life expectancy could be considerably enhanced if his professional opinions, while appearing to convey useful information, were actually couched in language of such ambiguity as to cover all likely contingencies. For example, King Croesus of Lydia, in the sixth century B.C., was considering conducting a war of conquest against a neighboring state, but was doubtful as to the outcome. He decided that it would be prudent to consult the Delphian Oracle. The advice that he received was, "If you embark upon this campaign, a great empire will be destroyed." Thinking that to be a good omen, Croesus did embark upon the campaign and a great empire was destroyed, his own.

The technique of prediction by ambiguous statement is still by no means obsolete and it is not unknown to the petroleum industry. During the 1950-decade, one of the most widely quoted dicta released by the propaganda branch of the U.S. petroleum industry was: "The United States has all the oil it will need for the foreseeable future."

Rational techniques of prediction fall into two fairly distinct classes:

1. Techniques based upon empirical extrapolation of real data, with little or no theoretical guidance.
2. Techniques based upon the analysis of data with the theoretical guidance provided by a prior understanding of the mechanism of the phenomena investigated.

Prediction by trends and cycles.— Of the empirical methods, one of those used most commonly at present is based upon the extrapolation into the future of some variable which, during the recent past, has displayed an approximately linear variation with time. By extending this linear trend with a straight-edge, when the data are plotted graphically, a prediction of its future can be made. This is the simplest semi-rational technique to apply, and it probably is the one in widest use at present. But how reliable is it, especially when applied to the exploitation of an exhaustible mineral resource? As an example of the technique, the annual production of pig iron in the United States, plotted on semi-logarithmic paper, for the 60-year period, 1850-1910, is shown in Figure 1.

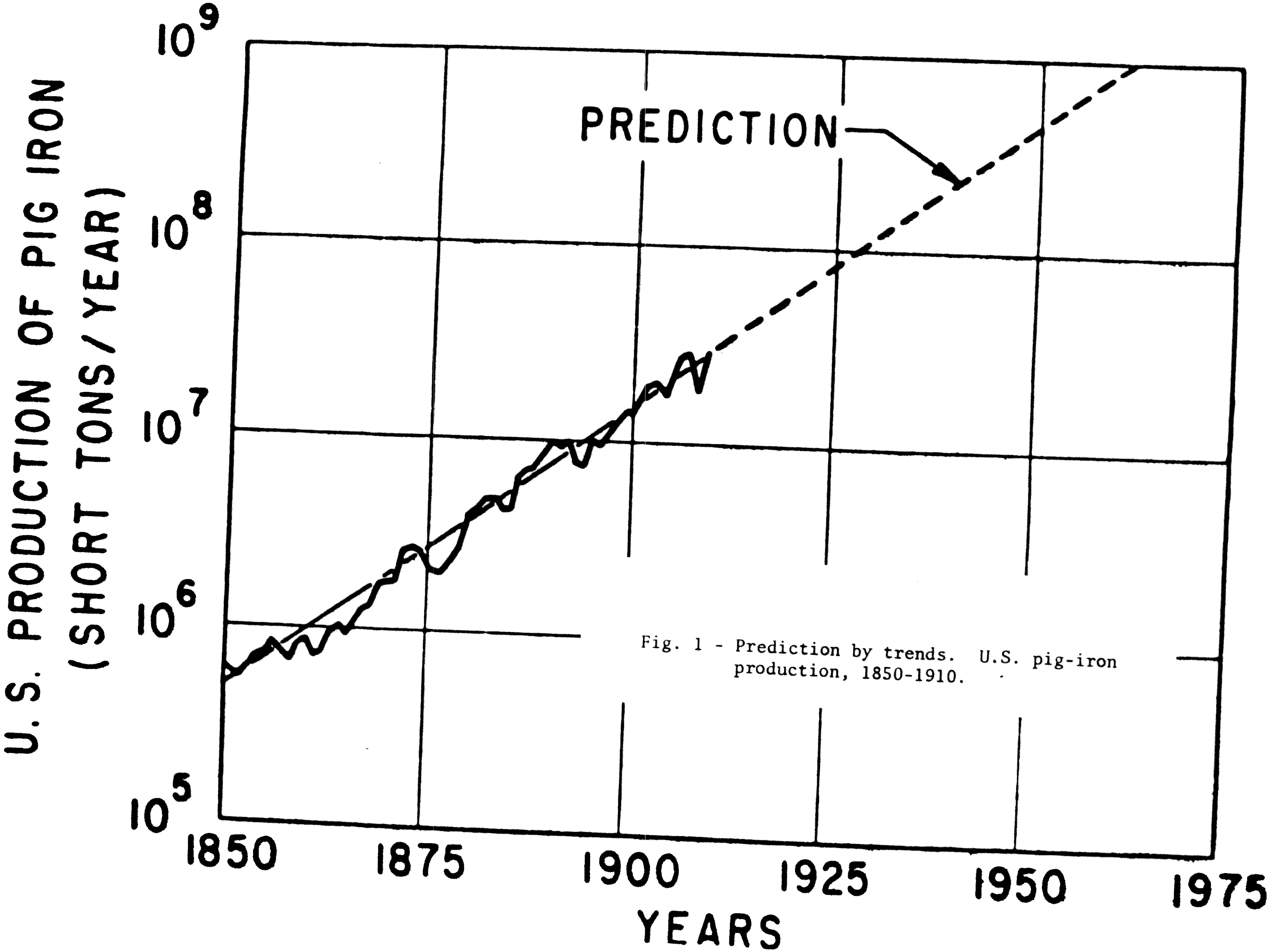

Fig. 1 - Prediction by trends. U.S. pig-iron production, 1850-1910.

The curve plots as an excellent straight line, corresponding to a uniform exponential growth at an annual rate of 6.9 percent. The dashed line shows the prediction of the growth of pig-iron production in the future. The steel industry has been reported to have based its long-range plans on this projection until 1925. Figure 2 shows what actually happened. Within a three-year period from about 1907 to 1910, this curve changed abruptly from a growth rate of 6.9 percent per year to another straight-line segment with a growth rate of only 1.8 percent.

In other instances the quantity of interest may vary in a more or less cyclical manner. In that case, the prediction of its future behavior consists in an extrapolation of this type of behavior into the future. In some cases - astronomical events within the solar system, for example - this can be done with great precision and solar or lunar eclipses can be predicted centuries in advance. Other cyclical or pseudo-cyclical events, such as the "business cycle," can be predicted by this means with much less assurance.

Necessity for knowledge of mechanism.— The fundamental flaw in the use of blind empirical methods lies in the fact that such methods take a minimal account of the inherent constraints in the behavior of a system imposed by its mechanism and physical properties. Thus, when dealing with cyclical events of the solar system, the Newtonian laws of motion and of gravitation, in conjunction with the infinitesimal rate of energy dissipation in the motions of the planets, provides assurance that the cyclical nature of these motions will change extremely slowly over periods of many millions of years. On the other hand, a weight-driven pendulum clock can maintain a precise cyclical operation of its pendulum and its hands, but the fall of the weight is noncyclical, and when the weight reaches its lowest point, the clock stops.

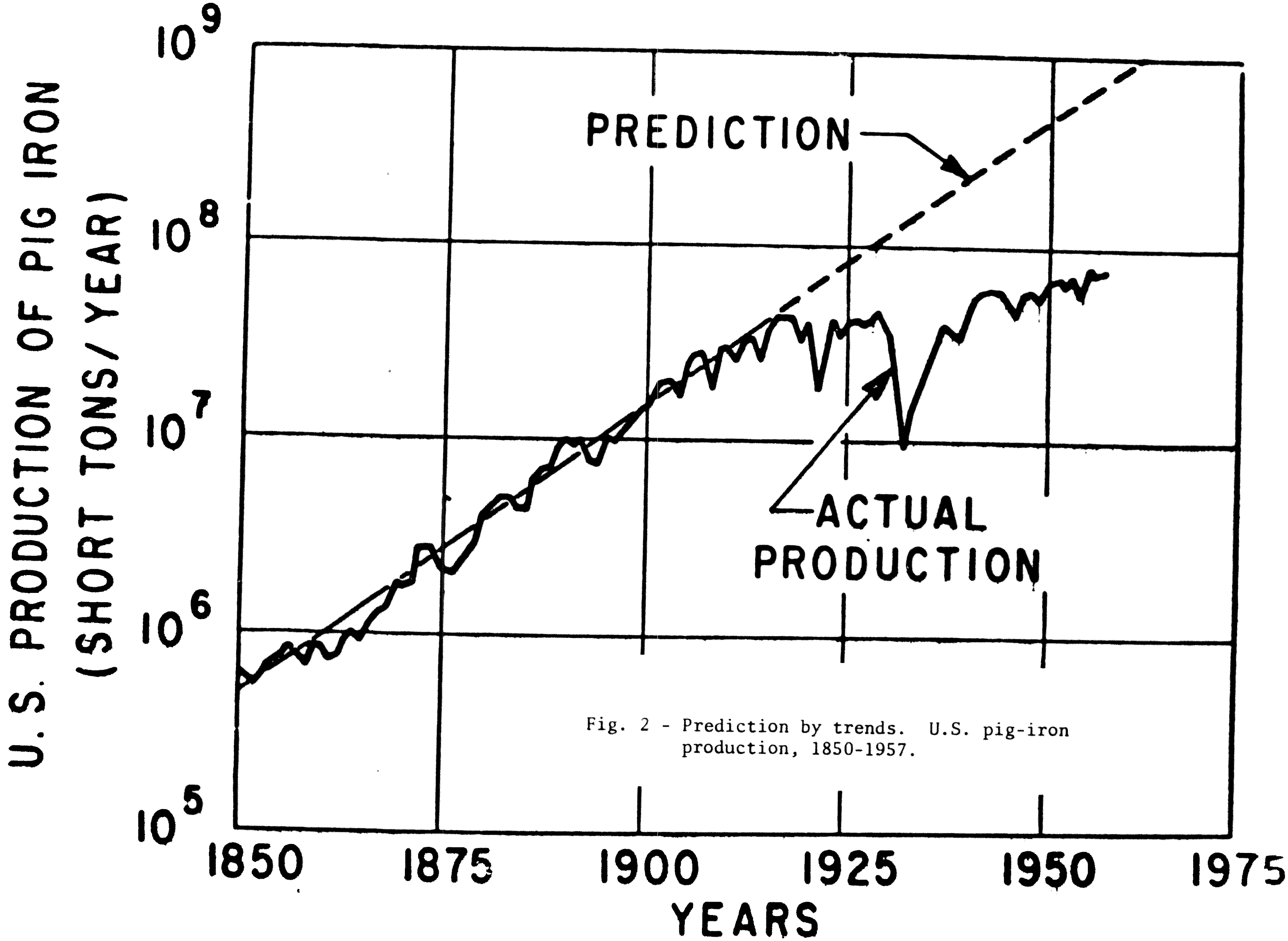

Fig. 2 - Prediction by trends. U.S. pig-iron production, 1850-1957.

A similar situation prevails in the rate of consumption of natural gas in a northern city. This curve consists of two parts, a base rate of consumption, which is noncyclical, upon which is superposed a roughly sinusoidal variation with a period of 1 year. This sinusoidal component is obviously related to the annual rise and fall of climatic temperatures, and hence upon the annual cycle of the earth's revolution. One cannot, however, predict the future of the gas consumption by its sinusoidal component for, say, another century, because by that time there may no longer be any natural gas.

The use of linear trends can be very reliable in cases where the mechanism is understood and does produce a sustained linear variation with time. One of the best known examples of this is the decay of a radioactive isotope. If N_o is the number of atoms of this isotope initially present in a closed system, then, it has been found, the number N of atoms remaining after a time t is given by

$$N = N_o e^{-at}, \tag{1}$$

where e, equal to 2.7183, is the base of natural logarithms, and a is a constant for the given isotope. Taking the natural logarithms of both sides gives the linear equation

$$\ln N = \ln N_o - at, \tag{2}$$

whereby $\ln N$ as a function of t plots as a straight line which may be extrapolated with confidence so long as N is large as compared with its statistical perturbations. This is the basic principle used in the radioactive dating of past geologic events.

Mechanism of Petroleum Occurrence

Similarly, any rational prediction of the future of petroleum production must depend upon a prior knowledge of the manner of origin of oil and gas, the geological situations where these fluids now occur, the procedures of petroleum exploration and production, and the respective time scales. As to their modes of occurrence, oil and gas at present are found as concentrations occupying the pore volumes of coarse-textured or fractured porous rocks in limited regions of space within or adjacent to basins of sedimentary rocks.

In petroleum geology, if a well is drilled at any point, it commonly will penetrate various thicknesses of porous sedimentary rocks - sandstones, shales or mudstones, and limestones. Eventually, if drilled deep enough, the well will reach the bottom of the sediments and will encounter the underlying older non-porous crystalline rocks which are commonly known as the *basement complex*, or the "basement." The upper surface of the basement complex is a continuous, world-wide surface. In certain large regions - a large part of Africa, most of the eastern half of Canada, and much of Scandinavia - the sedimentary cover is either very thin or else absent altogether. In these regions the surface of the basement rocks coincides with the surface of the ground. In other regions, the top of the basement is depressed and these depressions are filled with younger sedimentary rocks, having thicknesses ranging from zero at the edges, to a maximum in their interiors. Commonly, this maximum thickness may not exceed 2 or 3 kilometers, but in a few cases it has been found to be as great as 10 to 15 kilometers.

Unmetamorphosed sedimentary rocks are porous, with the pore volume ranging usually between about 10 and 30 percent of the total volume. Beneath shallow depths of a few tens of meters from the earth's surface, the pore space of these sedimentary rocks is filled with water which extends to depths below which the porosity becomes zero. Oil and natural gas, and also coal, are found in these sedimentary rocks. These are derived from the remains of plants and animals which lived at the times the sediments were being deposited. They became buried in the accumulating sedimentary sands and muds in an oxygen-deficient environment, and have subsequently become transformed chemically into the present fossil fuels.

Oil and gas thus occur as minority fluids immersed in an underground rock-water environment. Being fluids, oil and natural gas are mobile. Occurring originally in a dispersed state, the separate particles have been driven by mechanical forces or else transported in solution, into limited regions of space, corresponding to minimum levels of potential energy with respect to the local environment and lithologic barriers to further migration. The exploration for petroleum consists in determining as accurately as possible from less costly geological and geophysical data the most probable positions in three-dimensional space at which these accumulations may occur, and then drilling exploratory wells at the sites so selected. Most such exploratory wells are failures, but the probability of success is considerably higher than that of random drilling.

As regards time scales, the oldest industrial-sized gas field so far discovered appears to be one in Australia which occurs in rocks of late Precambrian geologic age - some 600 to 700 million years old. Oil and gas fields have been found in the United States and other parts of the world in rocks ranging in age from the Cambrian, 500 to 570 million years old, to sediments of Pleistocene age deposited in the Mississippi Delta of coastal Louisiana during the last million years. The exploitation of petroleum, on the other hand, did not begin on an industrial scale until as recently as the 1850-decade. Statistics of oil production in Romania are reported since 1857, and the initial oil discovery in the United States, by a well drilled for that purpose, was made at Titusville, Pennsylvania, in August 1859. The time required to essentially deplete the initial oil and gas resources of the world can hardly be longer than two or three centuries.

The disparity of these time scales, that for the accumulation of the world's supply of oil and gas, and that for its depletion, is very significant. Although the same geological processes are still operative by which the original oil and gas were accumulated, and at about the same average rate, any additional oil or gas that could accrue within the next few centuries would be infinitesimal as compared with that of the last 600 million years. Therefore, during the period of oil exploitation, the production of oil and gas must consist solely of the continuous withdrawal from a stockpile of an initially fixed and finite magnitude, to which no additions will be made. Therefore, the amount of oil or gas remaining can only decline monotonically as a function of time.

The So-called Geologic Methods of Petroleum Estimation

Petroleum geology and geophysics, which are fundamental to petroleum exploration, comprise the entire complex of existing knowledge regarding the origin, migration, and entrapment of oil and gas, and their present modes of occurrence. This involves of necessity the most detailed knowledge that can be obtained regarding the rocks filling various sedimentary basins, their spatial distribution, and their fluid contents, water, oil, and gas. This information is acquired jointly by surface geological and geophysical mapping, but eventually in most detail from the subsurface geological information provided by wells drilled into the sediments. It is a truism of the petroleum industry that the only tool that actually discovers oil is the drill. Hence it is the record of exploratory and production drilling in a given region that provides the most reliable information available regarding the occurrence of oil and gas, and the probable quantities of these fluids that a given basin may eventually be expected to yield.

However, during the last twenty years, a great deal of confusion has been introduced into the estimates of future petroleum production by the argument that "geological" methods of estimation must somehow be more reliable than so-called statistical methods based upon the cumulative information provided by drilling. By the advocates of this view, the scope of "geology" is seldom defined, but it apparently excludes or minimizes the importance of the information provided by drilling. The estimates obtained by these so-called geological methods during the 1960-decade for the ultimate amounts of crude oil and natural gas to be produced in the Lower-48 states and adjacent continental shelves of the United States were commonly about 600 to 650 billion barrels for crude oil and 2,500 trillion cubic feet for natural gas.

One of the more important uses of geologic methods is in a qualitative evaluation of the petroleum potential of an undrilled area. This is done principally by geological analogy. Suppose, for example, that two contrasting regions, *A* and *B*, have been explored and adequately drilled. Region *A* has been found to be a petroleum-rich region and Region *B* has been found to be barren. Two undrilled regions, *C* and *D*, are under consideration for future exploration. From preliminary geological and geophysical mapping, Region *C* is found to be geologically similar to barren Region *B*, and Region *D* to productive Region *A*. On the basis of this comparison, it would be inferred that Region *D* would merit further development, and that Region *C* should be given a lower priority.

In this connection, the United States is the most intensively explored major petroleum-producing region in the world. Consequently, it often has been used as a standard in the estimation of other less-developed or undeveloped potential petroleum-producing areas of the rest of the world. It is accordingly not surprising that estimates of the ultimate oil potential for the world have a strong correlation with estimates made by the same authors for the United States.

But how are the estimates for the primary areas to be obtained? It is easy to show that no geological information exists, other than that provided by drilling, that will permit an estimate to be made of the recoverable oil obtainable from a primary area that has a range of uncertainty of less than several orders of magnitude. To show this, consider the composite potential of all oil-bearing sediments of a primary region.

Let A be the surface area of the potential oil-bearing region, D be the average thickness of the sediments, and V the total volume of the sediments. Then

$$V = AD \tag{3}$$

Of the total sediments, oil and gas can be produced only from coarse-textured reservoir rocks, which are principally sandstones, and a lesser amount of porous limestones. Let V_{res} be the volume of reservoir rocks and λ the ratio between V_{res} and V. Let V_p be the pore volume of the reservoir rocks and ϕ their average porosity. Finally, let S_o be the average oil saturation of the reservoir rocks and V_o the volume of the oil. Let V_r be the recoverable oil and F the fraction of oil-in-place that can be recovered. Finally, let R be the richness of the region, defined as the ratio of the volume of recoverable oil to the total sedimentary volume. Then

$$V_{res} = \lambda V \tag{4}$$

is the volume of the reservoir rocks;

$$V_p = \phi V_{res} = \phi\lambda V \tag{5}$$

is the pore volume of the reservoir rocks;

$$V_o = S_o V_p = S_o \phi\lambda V \tag{6}$$

is the volume of oil in the reservoir rocks; and finally,

$$V_r = (FS_o \phi\lambda) V \tag{7}$$

is the volume of recoverable oil. Also, the richness R of the region is defined to be

$$R = V_r/V = (F\phi\lambda)S_o. \tag{8}$$

In equation (8), the factors F, ϕ, and λ are known within narrow limits. The recovery factor F may be taken to be about 0.4, the porosity ϕ of reservoir sands has an average value of about 0.15. According to F. W. Clarke, in his classical monograph, *The Data of Geochemistry* (Clarke, 1924, p. 34), shales comprise about 80 percent of the total volume of sediments, sandstones about 15 percent, and limestones and evaporites about 5 percent. Accordingly, we may take about 0.15 as the average value of λ. Hence the combined factor

$$\begin{aligned}(F\phi\lambda) &\simeq 0.4 \times 0.15 \times 0.15 \\ &\simeq 0.009 \\ &\simeq 10^{-2} \\ R &\simeq 10^{-2}S_o. \end{aligned} \tag{9}$$

Then the total amount of recoverable oil that a given region will produce would be

$$V_r \simeq (10^{-2}S_o)V. \tag{10}$$

Therefore the accuracy of a geological estimate of the oil that a given region will produce depends almost entirely upon that of the oil saturation factor S_o. The factor S_o must lie between the limits 0 and 1.0 but how, except by the data of prior drilling can it be determined whether 10^{-1}, 10^{-4}, or 10^{-6} is the better value for this factor?

The cumulative results of exploratory and production drilling, on the other hand, in petroleum-producing regions in advanced states of development, do provide very good information as to the actual magnitudes of the richness, R, and of S_o, within a range of uncertainty of 2 or less for various regions. For example, consider the three following oil-rich regions, each of which is in an advanced state of exploratory and production development:

The Los Angeles basin, California

The Lower-48 states and adjacent continental shelves of the U.S.

The Arabian Gulf basin of the Middle East.

For the Los Angeles basin, about 97 percent of the oil discovered is found in upper Miocene and lower Pliocene sediments, having a volume of 6.67×10^{12} m^3. The cumulative discoveries amount to 7.65×10^9 bbl, or to 1.21×10^9 m^3. (Kilkenny, 1971, p. 170-173; Gardett, 1971, p. 278). From these data, the minimum value of the richness of this basin is

$$R = \frac{1.21 \times 10^9 \text{ m}^3}{6.67 \times 10^{12} \text{ m}^3} = 1.8 \times 10^{-4},$$

and

$$S_o \simeq 100\,R = 1.8 \times 10^{-2}.$$

In other words, the Los Angeles basin has about 180,000 m^3 of recoverable oil per km^3, or 4.7 million barrels per cubic mile, and an average oil saturation in the reservoir rocks of nearly 2 percent.

For the Lower-48 states and continental shelves, the area of potentially oil-producing sediments is 2.0×10^6 square miles, or 5.2×10^{12} m^2 (Cram, 1971, v. 1, p. 5). Then, with an average thickness of 2,500 meters, the total volume would be

$$5.2 \times 10^{12} \times 2.5 \times 10^3 = 13 \times 10^{15} \text{ m}^3.$$

The cumulative oil discoveries for the Lower-48 states amount to about 155×10^9 bbl, or to 24.6×10^9 m^3. From these data,

$$R = \frac{24.6 \times 10^9 \text{ m}^3}{13 \times 10^{15} \text{ m}^3} = 1.9 \times 10^{-6},$$

and

$$S_o = 100\ R = 1.9 \times 10^{-4}.$$

Thus the average oil richness for the potential oil-producing sediments of the Lower-48 states and continental shelves is only about 1,900 cubic meters per cubic kilometer, or 50,000 bbl per cubic mile, and the average oil saturation is only about 2 parts in 10,000 of the total pore volume of the reservoir rocks.

For the Arabian Gulf basin, about 460×10^9 bbl or $73 \times 10^9\ m^3$ of oil has been discovered in a sedimentary volume of $2.5 \times 10^{15}\ m^3$ (Morris, 1978, Preface; Law, 1957, p. 60). This gives for the richness

$$R = \frac{73 \times 10^9\ m^3}{2.5 \times 10^{15}\ m^3} = 2.9 \times 10^{-5}.$$

and

$$S_o = 100\ R = 2.9 \times 10^{-3}.$$

Thus the Arabian Gulf basin has a richness of about 29,000 cubic meters of recoverable oil per cubic kilometer, or about 760,000 bbl of oil per cubic mile, and an oil saturation of the reservoir rocks of about 0.3 percent.

From these comparisons, the Los Angeles basin, the richest known basin in the world, has a richness that is 95 times that of the entire Lower-48 states, and 6 times that of the Arabian Gulf basin. The latter has a richness 15 times that of the Lower-48 states.

From the foregoing discussion it should be clear that arguments over the relative superiority of the so-called geological estimates and those arrived at by other methods serve little useful purpose but can produce a great deal of confusion, especially when the geological estimates are several-fold larger than other estimates. Actually, a petroleum geologist or engineer, when studying a given region, makes use implicitly or explicitly of every kind of pertinent information that may be available. A large and significant part of this information has to be the cumulative knowledge provided by prior exploratory and production drilling.

Complete Production Cycle in Given Region

In any particular oil-bearing region, the production history of the region, or the complete production cycle, has the following essential characteristics. At some initial time t_0 the first discovery well is drilled and oil production in the region begins. The first discovery is of a single field. Additional wells are drilled to develop the field and the rate of production increases. Further exploration in the region is thus stimulated, and, on the basis of geological and geophysical studies, further potential oil-bearing structures are drilled and new fields are discovered. However, since there are only a fixed number of fields in the region, as more and more fields are discovered, progressively fewer fields remain to be discovered, and these are the more obscure fields and commonly the smaller ones. As the undiscovered fields become scarcer, the amount of exploratory effort, including exploratory drilling, per unit quantity of oil discovered increases until eventually exploratory drilling becomes prohibitively costly and ceases.

The rate of oil production, dQ/dt, in the region begins at a near-zero rate at time t_0 and thereafter commonly increases exponentially for a few decades. Eventually, as the rate of discovery slows down, the rate of production follows. It reaches one or more principal maxima, and finally goes into a slow negative-exponential decline. Then, at some definite time t_k, production ceases altogether. This sequence can be stated mathematically as follows:

For

$$\left.\begin{array}{rl} t < t_o, & dQ/dt = 0; \\ t_o < t < t_k, & dQ/dt > 0; \\ t > t_k, & dQ/dt = 0. \end{array}\right\} \tag{11}$$

The period from t_o when production first begins until t_k when it ends comprises the complete cycle of oil production in the region. During that time, the cumulative production Q_t from time t_o to a later time t is given by

$$Q_t = \int_{t_o}^{t} (dQ/dt)\,dt$$

$$= \int_{t_o}^{t} P\,dt, \tag{12}$$

where $P = dQ/dt$ is the rate of production. Then for the complete cycle

$$Q_k = \int_{t_o}^{t_k} P\,dt. \tag{13}$$

Or, since for times earlier than t and later than t_k, $P = 0$, then

$$Q_k = \int_{-\infty}^{+\infty} P\,dt = Q_\infty. \tag{14}$$

Thus Q_∞ may be taken to be the ultimate amount of oil that will ever be produced in the region during an unlimited period of time.

This complete-cycle curve has only the following essential properties: The production rate begins at zero, increases exponentially during the early period of development, and then slows down, passes one or more principal maxima, and finally declines negative-exponentially to zero. There is no necessity that the curve, P as a function of t, have a single maximum or that it be symmetrical. In fact, the smaller the region, the more irregular in shape is the curve likely to be. The crude-oil production curve for the State of Illinois, for example, is shown in Figure 3. Here, oil was first discovered about 1900. Until 1905, the production rate increased very slowly. It rose sharply to a peak rate of about 35 million barrels per year during 1908 to 1910. This was followed by a long negative-exponential decline to about 1 million barrels per year by 1936. Then followed a new cycle of exploration and discovery based upon the use of the reflection seismograph capable of mapping geologic structures beneath the cover of glacial drift. With the new discoveries, the production rate increased sharply to about 145 million barrels per year by 1940, and then declined to 55 million barrels per year by 1953. Next followed a ten-year period of a third production peak of about 80 million barrels per year due to water flooding, and finally a decline to about 35 million barrels per year by 1971.

On the other hand, for large areas, such as the entire United States or the world, the annual production curve results from the superposition of the production from thousands of separate fields. In such cases, the irregularities of small areas tend to cancel one another and the composite curve becomes a smooth curve with only a single principal maximum. However, there is no theoretical necessity that this curve be symmetrical. Whether it is or not will have to be determined by the data themselves.

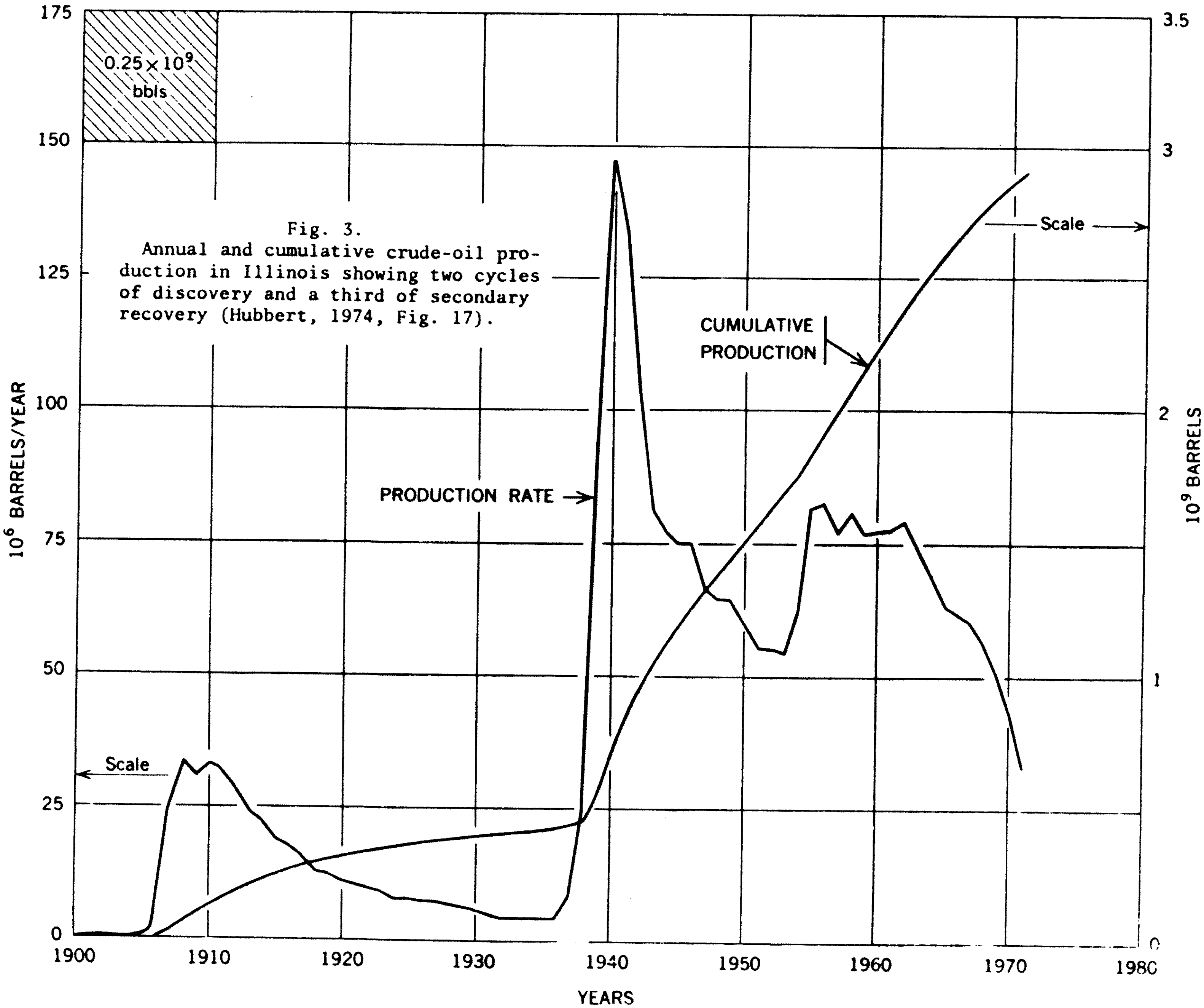

Fig. 3.
Annual and cumulative crude-oil production in Illinois showing two cycles of discovery and a third of secondary recovery (Hubbert, 1974, Fig. 17).

The Complete-Cycle Curve as a Function of Q_∞

The foregoing properties of the complete-cycle curve of production afford a simple but powerful means of estimating the future course of the production-rate curve as a function of Q_∞. After the completion of the cycle, as shown by equation (14),

$$Q_\infty = \int_{-\infty}^{+\infty} P dt.$$

When the curve P versus t is plotted graphically with arithmetic scales, as in Figure 4, the area between the curve and the t-axis to any given time t is a graphical measure of cumulative production. Then, for the complete cycle, the total area beneath the curve is a measure of Q_∞. Also, in plotting such a curve, scales must be chosen arbitrarily for the ordinate P, with a graphical interval ΔP, and for the abscissa t, with a graphical interval Δt. The grid-rectangle, $\Delta P \times \Delta t$, affords a graphical scale for cumulative production. At a constant rate ΔP, the quantity of oil produced during the time Δt would be

$$\Delta Q = \Delta P \times \Delta t. \tag{15}$$

Hence each grid-rectangle beneath the curve represents ΔQ of cumulative production. Therefore, if Q_∞ is known, then the number of grid-rectangles beneath the complete-cycle curve must be

$$n = Q_\infty / \Delta Q. \tag{16}$$

This is the inverse of the usual problem of the integral calculus, where one is given $y = f(x)$, and the problem is to find

$$A = \int y dx.$$

Here, we are given A and the problem is to find the curve $y = f(x)$. There are obviously an infinite number of curves that will satisfy this condition. However,

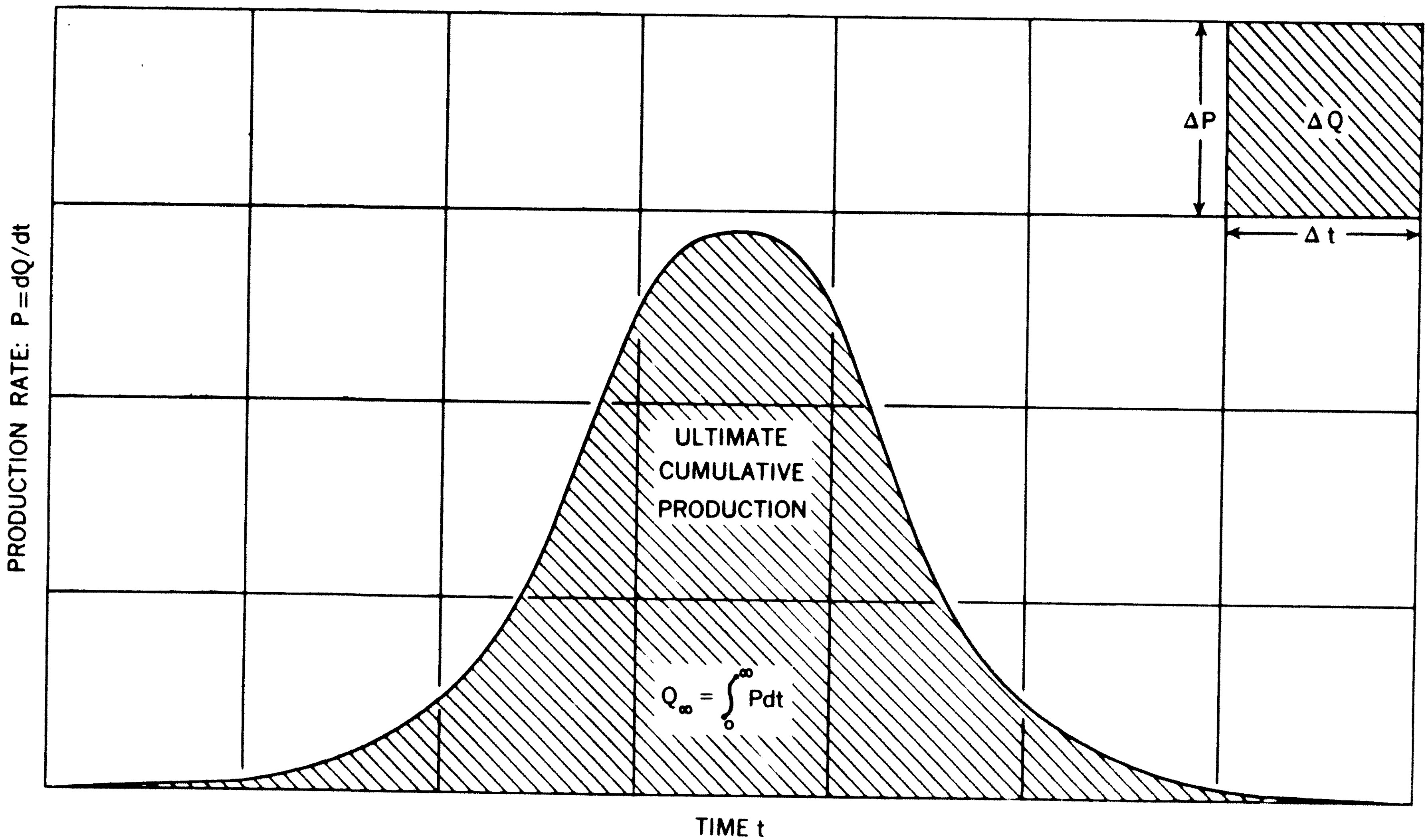

Fig. 4 - Mathematical relations involved in the complete cycle of production of any exhaustible resource (modified from Hubbert, 1956, Fig. 11).

when, in the oil-production case, the further constraints imposed by the technology of oil discovery and production are taken into account, the differences between separate solutions tend to become minor. Accordingly, although there may be an infinite number of complete-cycle curves corresponding to a single value for Q_∞, all of these curves have a strong family resemblance because the area subtended by each is the same, namely n rectangles.

When the petroleum discovery and production in a given region reaches a moderately mature state of development, the history of this development begins to provide a basis for reasonably good estimates of the approximate magnitude of Q_∞ for the region. Suppose that by the time t_1, cumulative production has already reached Q_1. Then the oil remaining to be produced would be

$$Q_2 = Q_\infty - Q_1. \qquad (17)$$

Accordingly the area beneath the curve from t_o to t_1 would be

$$Q_1/\Delta Q = n_1$$

rectangles, and the remaining area beneath the curve from t_1 to ∞ would be

$$(Q_\infty - Q_1)\Delta Q = n - n_1$$
$$= n_2$$

rectangles. Hence the future part of the curve must be consistent with the part that has developed already, and must also be drawn subject to the constraint that it can subtend only n_2 grid-rectangles. In case the peak production rate has not yet been reached by time t_1, then the curve for the future must rise to a maximum and then decline negative-exponentially. If the maximum production rate has already occurred by time t_1, then its future course will be principally the negative-exponential decline.

It must be borne in mind that at time t_1 any figure for Q_∞ is only an estimate, yet the foregoing technique provides a means of determining the figures for estimates of Q_∞, in conjunction with the technique described above, gave the results shown in Figure 6. There one $\Delta P \Delta t$ grid-rectangle has the dimensions

$$\Delta Q = 10^9 \text{ bbl/yr} \times 25 \text{ years}$$
$$= 25 \times 10^9 \text{ bbl}.$$

For the lower figure for Q_∞ of 150 × 10^9 bbl, the total area beneath the curve for the complete production cycle would be 6 grid-rectangles. Of these, 2.1 corresponding to cumulative production of 52.4 × 10^9 bbl had already been developed, leaving 3.9 rectangles for future cumulative production of 97.6 × 10^9 bbl. The lower dashed-line curve of Figure 6 is drawn accordingly. To satisfy this condition in conjunction with a negative-exponential decline, it became impossible to draw this curve very differently from the way it is shown in Figure 6. From 1956, the curve would reach its peak rate of about 2.7 × 10^9 bbl/yr about 10 years hence, and then decline negative-exponentially back to zero.

Assuming that Q_∞ could be as large as the higher figure of 200 × 10^9 bbl would add another 50 × 10^9 bbl to the area beneath the curve of future production, or 2 more rectangles. This curve would rise a little higher than the first, and would reach its peak rate a little later, but the 2 extra rectangles would be the area between the two curves, principally during their decline.

By this analysis, if Q_∞ should be as small as 150 billion barrels, the peak in the rate of production should occur in about 10 years, or about 1966; for the higher figure of 200 billion barrels, the date of peak production would be delayed by about another 5 years, or to about 1971.

The curves drawn in Figure 6 were not based upon any empirical equations or any assumptions regarding whether they should be symmetrical or asymmetrical; they were simply drawn in accordance with the areal constraints imposed by the

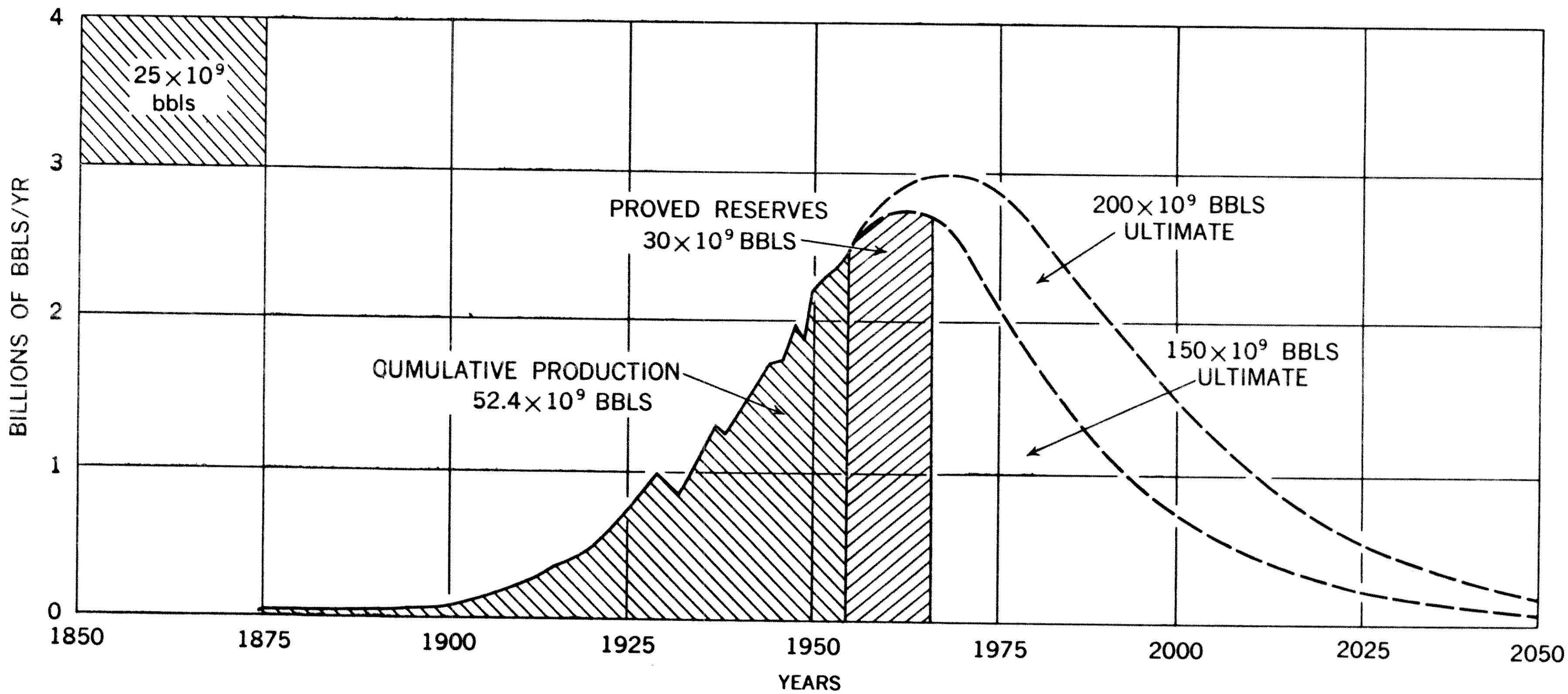

Fig. 6 - 1956 Estimates of two complete cycles of U.S. crude-oil production from Lower-48 states based upon ultimate recoveries of 150 and 200 billion barrels (Hubbert, 1956, Fig. 21).

estimates, and the necessity that the decline be gradual and asymptotic to zero. The strength of this procedure lies in the insensitivity of its most important deduction, namely the date of the peak-production rate, to errors in the estimate of Q_∞. As Figure 6 shows, an increase in the lower estimate of 150 billion barrels by one-third delays the date of peak production by only about 5 years, or to about 1971. If the lower figure were doubled to 300 billion barrels, the date of the peak-production rate would still be delayed only to about 1978.

Without this kind of analysis, the tendency is to extrapolate the production curve prior to 1956 into the future by the method of linear trends which would provide no information whatever regarding the imminence of the date of peak production. In fact, in 1956, the estimates that two or three times as much oil remained to be produced in the future as had been produced during the preceding century led to an attitude of complacency on the part of petroleum geologists, engineers, and oil-company officials alike, that no oil shortages were likely to occur in the United States before the year 2000.

The weakness of this analysis arose from the lack of an objective method of estimating the magnitude of Q_∞ from primary petroleum-industry data. The estimates extant in 1956 were largely intuitive judgments of people with wide knowledge and experience, and they were reasonably unbiased because of the comfortable prospects for the future they were thought to imply. When it was shown, however, that if Q_∞ for crude oil should fall within the range of 150-200 billion barrels the date of peak-production rate would have to occur within about the next 10 to 15 years, this complacency was shattered. It soon became evident that the only way this unpleasant conclusion could be voided would be to increase the estimates of Q_∞, not by fractions but by multiples. Consequently, with insignificant new information, within a year published estimates began to be rapidly increased, and during the next 5 years, successively larger estimates of 250, 300, 400, and eventually 590 billion barrels were published.

This lack of an objective means of estimating Q_∞ directly, and the 4-fold range of such estimates, made it imperative that better methods of analysis, based directly upon the primary objective and publicly available data of the petroleum industry, should be derived. Such methods, which encompass the cumulative records of discovery and production, drilling, and associated knowledge of petroleum geology, will now be developed.

Derivation of the Complete Cycle of Oil and Gas Exploitation from Primary Data

As we have noted heretofore, the complete production cycle of oil or gas, or of any other exhaustible resource, has the following general characteristics: At some initial time t_o production begins. Subsequently, the production rate increases with time, passes one or more principal maxima, and finally goes into a negative-exponential decline until at some time t_k it ceases altogether. During this complete cycle, the cumulative production,

$$Q = \int_{t_o}^{t} P dt, \tag{18}$$

increases monotonically from 0 to a final value Q_∞. The curve of cumulative production is a generally S-shaped curve, being asymptotic to zero initially and to the limit Q_∞ as t increases without limit. If the curve of P versus t has only a single maximum then the cumulative curve Q versus t will have but a single inflection point, coinciding in time with the peak in the production rate. Earlier than that, the cumulative curve will be concave upward; subsequently, concave downward.

A difficulty in analyzing either P or Q as a function of time arises from the asymptotic approaches of these quantities to their respective limits as time increases without limit. On the other hand, Q itself has the definite finite limits 0 and Q_∞. It is convenient, therefore, to consider the production rate dQ/dt as a function of Q, rather than of time. In this system of coordinates, dQ/dt is zero when $Q = 0$, and when $Q = Q_\infty$. Between these limits $dQ/dt > 0$, and outside these limits, equal to zero. While it is possible that during the production cycle dQ/dt could become zero during some interval of time, for any large region this never happens. Hence we shall assume that for

$$0 < Q < Q_\infty, \quad dQ/dt > 0. \tag{19}$$

The curve of dQ/dt versus Q between the limits 0 and Q_∞ can be represented by the Maclaurin series,

$$dQ/dt = c_0 + c_1 Q + c_2 Q^2 + c_3 Q^3 + \text{---}. \tag{20}$$

Since, when $Q = 0$, $dQ/dt = 0$, it follows that $c_0 = 0$.

Then

$$dQ/dt = c_1 Q + c_2 Q^2 + c_3 Q^3 + \text{---}, \tag{21}$$

and, since the curve must return to zero when $Q = Q_\infty$, the minimum number of terms that will permit this, and the simplest form of the equation, becomes the second-degree equation,

$$dQ/dt = c_1 Q + c_2 Q^2. \tag{22}$$

By letting $a = c_1$ and $-b = c_2$, this can be rewritten as

$$dQ/dt = aQ - bQ^2. \tag{23}$$

Then, since when $Q = Q_\infty$, $dQ/dt = 0$,

$$aQ_\infty - bQ_\infty^2 = 0,$$

or

$$b = a/Q_\infty,$$

and

$$dQ/dt = a(Q - Q^2/Q_\infty). \tag{24}$$

This is the equation of the parabola shown in Figure 7 whose slope is

$$d(dQ/dt)/dQ = a - (2a/Q_\infty)Q, \tag{25}$$

which, when $Q = 0$, is $+a$, and when $Q = Q_\infty$, is $-a$. Also, the maximum value of dQ/dt occurs when the slope is zero, or when

$$a - (2a/Q_\infty)Q = 0,$$

or

$$Q = Q_\infty/2. \tag{26}$$

It is to be emphasized that the curve of dQ/dt versus Q does not have to be a parabola, but that a parabola is the simplest mathematical form that this curve can assume. We may accordingly regard the parabolic form as a sort of idealization for all such actual data curves, just as the Gaussian error curve is an idealization of actual probability distributions.

One further important property of equation (24) becomes apparent when we divide it by Q. We then obtain

$$(dQ/dt)/Q = a - (a/Q_\infty)Q. \tag{27}$$

This is the equation of a straight line with a slope of $-a/Q_\infty$ which intersects the vertical axis at $(dQ/dt)/Q = a$ and the horizontal axis at $Q = Q_\infty$. If the data, dQ/dt versus Q, satisfy this equation, then the plotting of this straight line gives the values for its constants Q_∞ and a.

Our problem now is to integrate equation (24) in order to determine how dQ/dt and Q each varies as a function of time. This can be simplified by substituting

$$u = Q/Q_\infty,$$

or

$$Q = Q_\infty u. \tag{28}$$

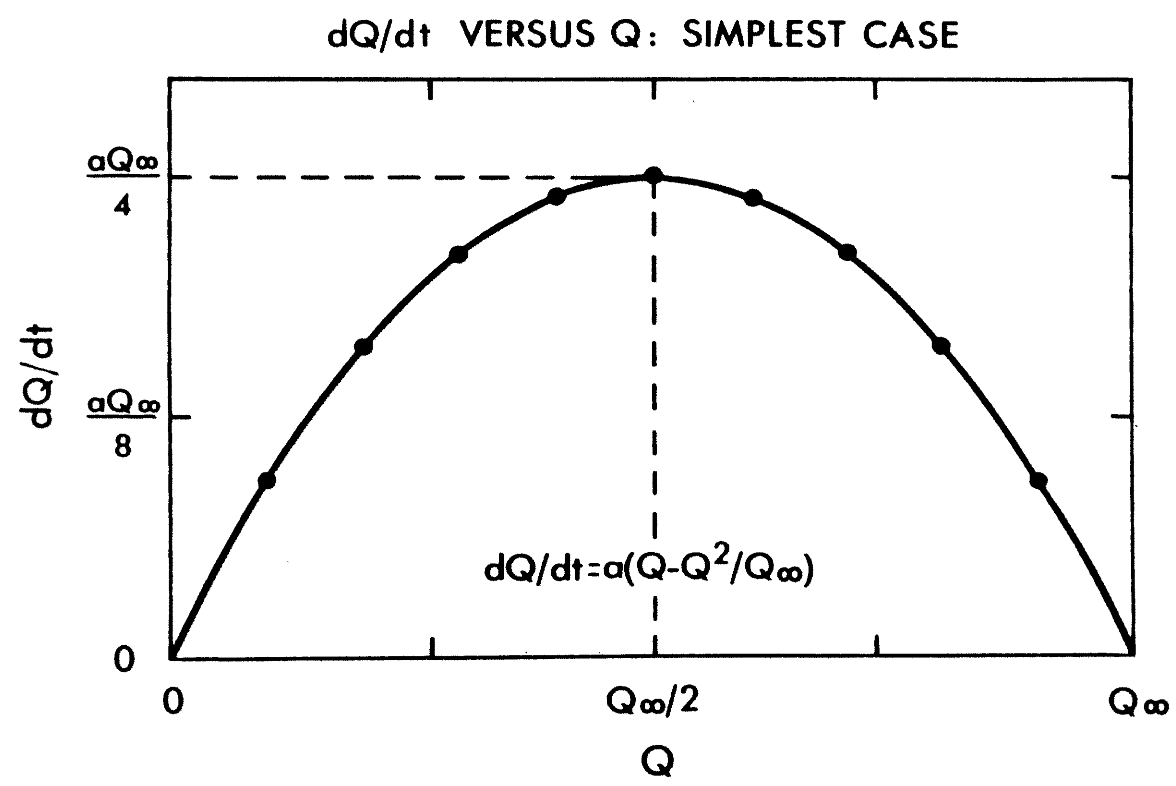

Fig. 7 - Production rate as a function of cumulative production for simplest mathematical case.

Then

$$dQ/dt = Q_\infty (du/dt)$$
$$= Q_\infty a(u - u^2),$$

or

$$du/dt = a(u - u^2),$$

and

$$du/(u - u^2) = adt. \tag{29}$$

This can be simplified by noting that

$$1/(u - u^2) = 1/u + 1/(1 - u).$$

Then the integral of equation (29) becomes

$$\int \frac{du}{u} + \int \frac{du}{1-u} = a \int dt + \text{const},$$

or

$$\ln [(1 - u)/u] = \ln c - at,$$

and

$$(1 - u)/u = ce^{-at}, \tag{30}$$

where c is the constant of integration.

Substituting Q/Q_∞ for u in the left-hand term of equation (30) then gives

$$(1 - u)/u = (Q_\infty - Q)/Q = ce^{-at} \tag{31}$$

Then, when $t = 0$, $Q = Q_0$, and

$$c = (Q_\infty - Q_0)/Q_0, \tag{32}$$

equation (31) becomes

$$(Q_\infty - Q)/Q = [(Q_\infty - Q_0)/Q_0]e^{-at}. \tag{33}$$

Then, letting

$$\left.\begin{aligned} N &= (Q_\infty - Q)/Q, \\ N_0 &= (Q_\infty - Q_0)/Q_0, \end{aligned}\right\} \tag{34}$$

equation (33) simplifies to

$$N = N_o e^{-at}. \tag{35}$$

Taking the natural logarithms of both sides of equation (35) gives

$$\ln N = \ln N_o - at, \tag{36}$$

which is a linear equation between $\ln N$ and t. This plots graphically as a straight line which has the magnitude $\ln N_o$ when $t = 0$, and a slope of $-a$. Or, if common logarithms are used,

$$\log N = \log N_o - (a \log e)t, \tag{37}$$

which has a slope of $-a \log e$.

Solving equation (33) for Q gives

$$Q = Q_\infty/(1 + N_o e^{-at}). \tag{38}$$

This is known as the logistic equation, which was derived originally by the Belgian demographer, P.-F. Verhulst (1838; 1845; 1847) in his classical studies of the growth of human populations.

In equation (38) the choice of the date for $t = 0$ is arbitrary so long as it is within the range of the production cycle so that N_o will have a determinate finite value. It is seen by inspection that as

$$t \to -\infty,\ Q \to 0,$$

$$t \to +\infty,\ Q \to Q_\infty.$$

Also the curve of Q versus t is asymptotic to zero as $t \to -\infty$, and is asymptotic to Q_∞ as $t \to +\infty$. Likewise, the curve of dQ/dt versus t is asymptotic to zero as $Q \to 0$ and $t \to -\infty$, and again as $Q \to Q_\infty$ and $t \to +\infty$.

The maximum value of dQ/dt, from equations (24), (25), and (26), is

$$dQ/dt = (aQ_\infty)/4. \tag{39}$$

This coincides in time with the inflection point of the Qt-curve and occurs when $Q = Q_\infty/2$.

In equation (24), dQ/dt is given as a function of Q. If desired, this can be obtained as an explicit function of t by differentiating the logistic equation (38),

$$Q = Q_\infty/(1 + N_o e^{-at}),$$

with respect to time. This gives the result,

$$dQ/dt = Q_\infty \frac{aN_o e^{-at}}{(1 + N_o e^{-at})^2} . \tag{40}$$

If we then note from equation (35) that

$$N_o e^{-at} = N = (Q_\infty - Q)/Q,$$

and, from equation (38),

$$1/(1 + N_o e^{-at})^2 = (Q/Q_\infty)^2,$$

and substitute these into equation (40), we obtain, as we should, our original differential equation (24),

$$dQ/dt = a(Q - Q^2/Q_\infty).$$

Graphs of the logistic equation (38), and of its time-derivative, equation (40), in terms of the dependent variable, $u = Q/Q_\infty$, as functions of time are shown in Figure 8.

Determination of the Constants Q_∞ and a

Two properties of the logistic equation are of fundamental importance. These are represented by the linear differential equation (27),

$$(dQ/dt)/Q = a - (a/Q_\infty)Q,$$

and the linear equation (36) between ln N and t,

$$\ln N = \ln N_o - at.$$

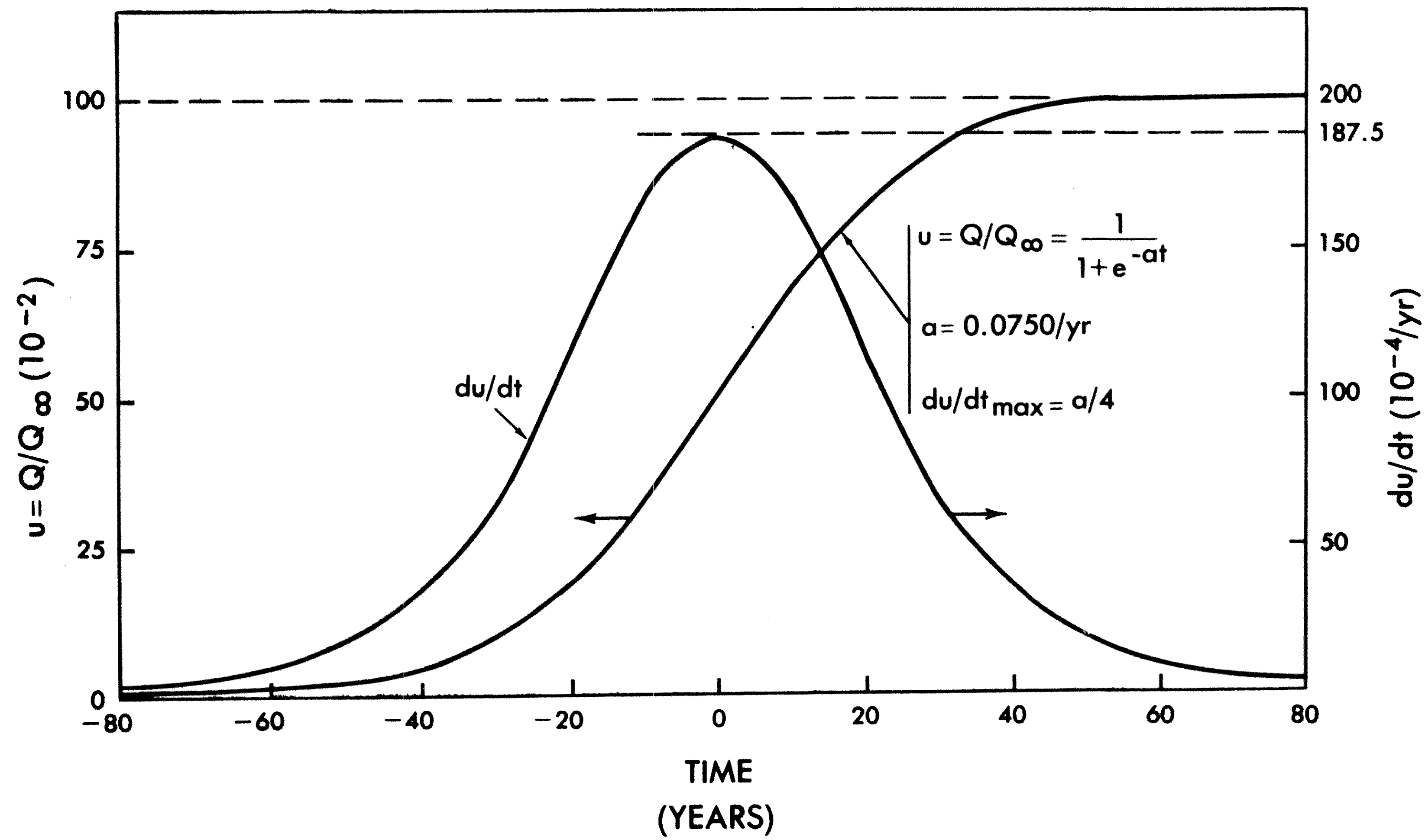

Fig. 8 - Graphical expressions of logistic equation and its time derivative as functions of time.

The virtue of the first of these two equations lies in the fact that it depends only upon the plotting of primary data, $(dQ/dt)/Q$, versus Q, with no a priori assumptions whatever. Using actual data for Q and dQ/dt, it is to be expected that there will be a considerable scatter of the plotted points as $Q \rightarrow 0$, because in that case both Q and dQ/dt are small quantities and even small irregularities of either quantity can produce a large variation in their ratio. For larger values of both quantities, as the production cycle evolves, these perturbations become progressively smaller and a comparatively smooth curve is produced. If the data satisfy the linear equation, then a determinate straight line results whose extrapolation to the vertical axis as $Q \rightarrow 0$ gives the constant a, and whose extrapolated intercept with the Q-axis gives Q_∞. However, even if the data do not satisfy a linear equation, they will nevertheless produce a definite curve whose intercept with the Q-axis will still be at $Q = Q_\infty$.

The use of the second linear equation,

$$\ln N = \ln N_0 - at,$$

is somewhat less direct than the first, because in this case

$$N = (Q_\infty - Q)/Q = Q_\infty/Q - 1.$$

Hence, before the linear graph can be plotted, Q_∞ must be known as a means of determining N and N_0. If the data satisfy the logistic equation, and if the correct value of Q_∞ is used for computing N and N_0, then the resulting graph of $\ln N$ as a function of t will continue as a straight line. If, on the other hand, an incorrect value,

$$Q_b = bQ_\infty, \tag{41}$$

is assumed for Q_∞, then

$$N_b = bQ_\infty/Q - 1, \tag{42}$$

and as

$$Q \to Q_\infty, \quad N_b \to (b - 1).$$

If the assumed value, Q_b, is greater than the correct value of Q_∞, then as t increases, the curve $\ln N_b$ versus t will approach the constant value, $(b - 1) > 0$, and will deflect to the horizontal. If the assumed value, Q_b, is less than Q_∞, then as

$$Q \to Q_b, \quad N_b \to 0,$$

and

$$\ln N_b \to -\infty.$$

In this case, as t increases, the curve deflects vertically downward. This is illustrated in Figure 9, where graphs of log N_b versus t are plotted for the three cases, $Q_b = 1.25\ Q_\infty$, $Q_b = Q_\infty$, and $Q_b = 0.75\ Q_\infty$.

These properties provide an iteration procedure for determining the correct value of Q_∞ and its associated constant a, provided the production cycle is far enough advanced for the deflections of the linear graphs to be perceptible. From the cumulative curve, Q versus t, a rough visual estimate of the value of Q_∞ can be made. Then a series of assumed values Q_{b_1} Q_{b_2} Q_{b_3}, etc. can be used for plotting a family of curves, $\ln N_b$ versus t. For the too large values of Q_b the curve will deflect toward the horizontal, and for the too small values, toward the vertical. The value of Q_b for which the curve continues as a straight line will be the correct value for Q_∞. The slope of that line will be $-a$.

A less cumbersome and more precise variation of this procedure consists in choosing three fixed times, t_1, t_2, and t_3. Let Q_b, regarded as a variable, be an assumed value for Q_∞. Then, at times t_1, t_2, and t_3,

$$\left.\begin{aligned} N_1 &= Q_b/Q_1 - 1, \\ N_2 &= Q_b/Q_2 - 1, \\ N_3 &= Q_b/Q_3 - 1, \end{aligned}\right\} \qquad (43)$$

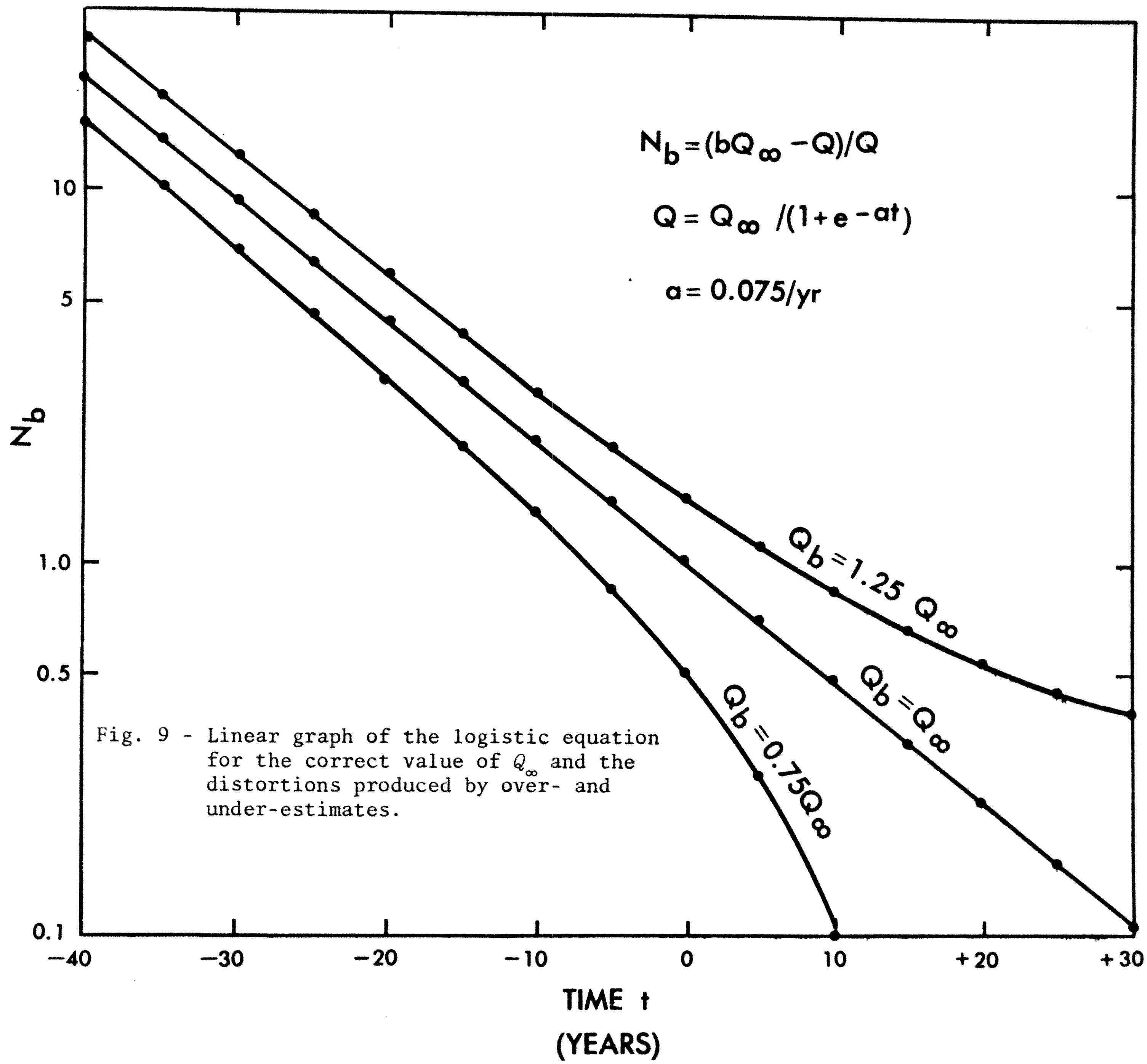

Fig. 9 - Linear graph of the logistic equation for the correct value of Q_∞ and the distortions produced by over- and under-estimates.

and the curve, ln N_b versus t, will consist of two line segments, one from t_1 to t_2, and the other from t_2 to t_3. From these respective line segments the corresponding negative slopes will be

$$\begin{aligned} -S_{12} &= \ln (N_1/N_2)/(t_2 - t_1), \\ -S_{23} &= \ln (N_2/N_3)/(t_3 - t_2). \end{aligned} \tag{44}$$

Hence $-S_{12}$ and $-S_{23}$ will each be a separate function of Q_b. When these two quantities are each plotted graphically as a function of Q_b, the point at which the two curves intersect one another will correspond to

$$S_{12} = S_{23},$$

for which the curve ln N_b versus t will be a straight line. The coordinates of that point will be

$$\begin{aligned} -S_{12} = -S_{23} &= a, \\ Q_b &= Q_\infty, \end{aligned} \tag{45}$$

which are the desired constants of the equation.

This is illustrated in Figure 10, based upon the following data for t_1, t_2, and t_3, and Q_1, Q_2, and Q_3,

Date (t)	1905	1945	1965
Q (10^9 bbl)	3.98	55.5	117.5

As shown in Figure 10, the values for the logistic constants corresponding to these data are

$$Q_\infty = 173 \times 10^9 \text{ bbl},$$
$$a = 0.0750.$$

An independent procedure for determining the constants Q_∞ and a is that based upon equation (27),

$$(dQ/dt)Q = a - (a/Q_\infty)Q.$$

If successive values of dQ/dt are known for successive values of Q, and if these data correspond to the logistic equation, then the curve of equation (27) will be a straight line intersecting the vertical axis as $Q \to 0$, at

$$(dQ/dt)/Q = a,$$

and the horizontal axis at

$$Q = Q_\infty.$$

This is illustrated in Figure 11, for the same constants as those for Figure 10.

After the best value of the constant Q_∞ has been determined, then the linear graph of equation (36),

$$\ln N = \ln N_o - at,$$

can be constructed. From this, the time at which the inflection point on the Qt-curve, or the maximum rate of production dQ/dt, will occur can either be read from the graph or else computed from equation (36). According to equation (26), the peak production rate will occur when

$$Q = Q_\infty/2,$$

or when

$$N = Q_\infty/(Q_\infty/2) - 1 = 1$$

and

$$\ln N = 0.$$

Solving equation (36) for t when $N = 1$ then gives

$$t = [\ln (N_o/N)]/a$$

$$= (\ln N_o)/a. \qquad (46)$$

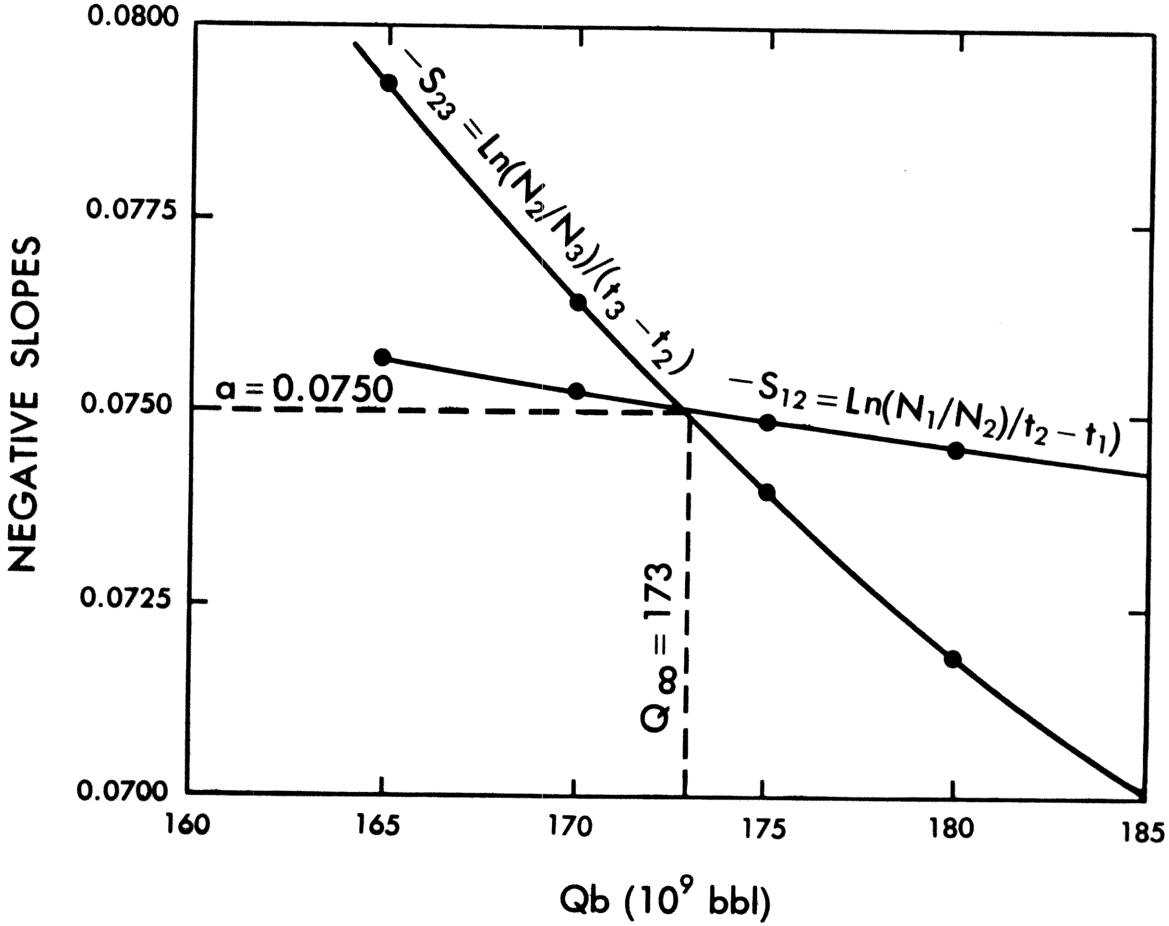

Fig. 10 - Graphical technique for determining the constants a and Q_∞ of the logistic equation.

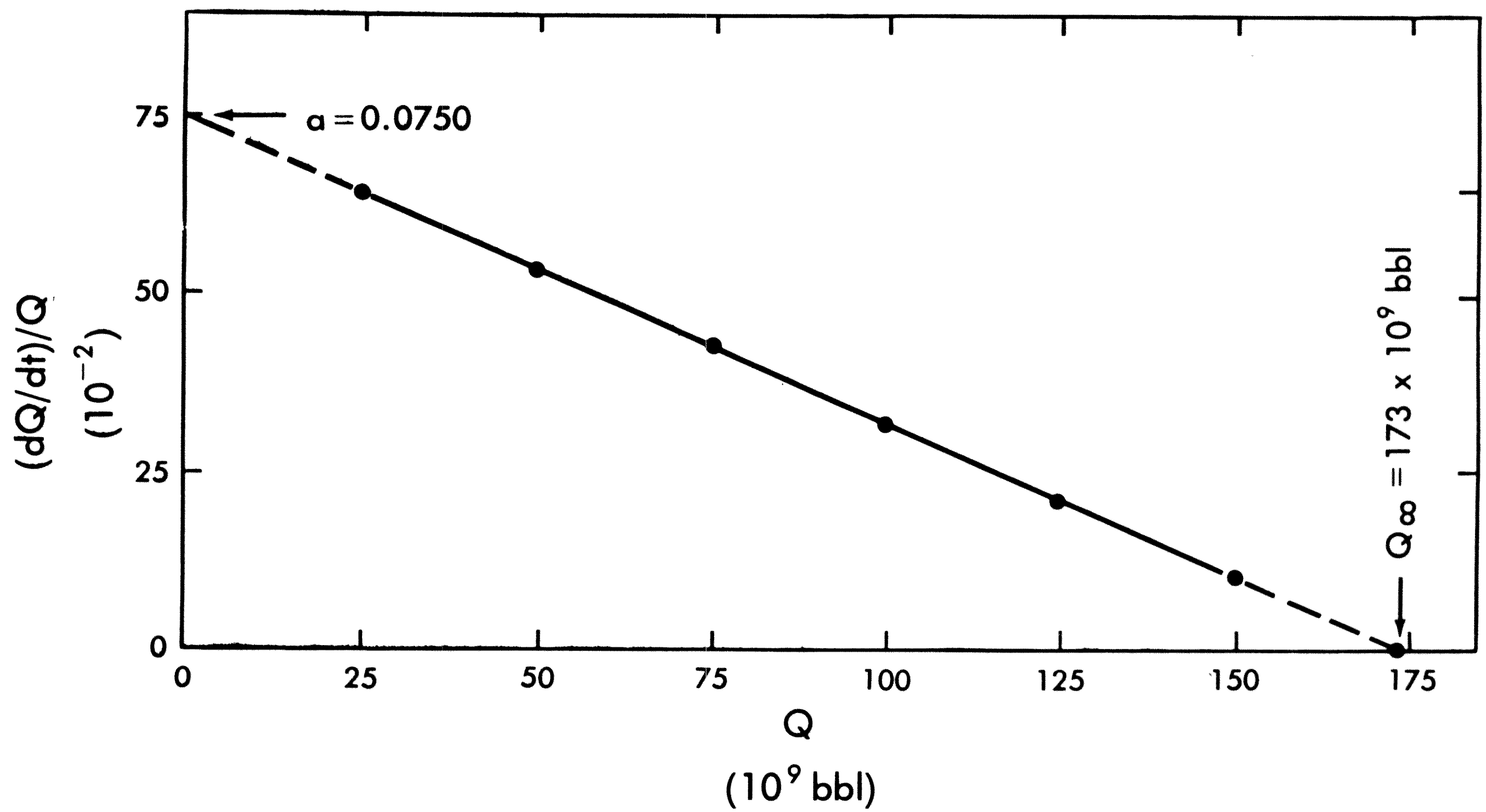

Fig. 11 - Determination of logistic constants by means of the linear graph of $(dQ/dt)/Q$ as a function of cumulative production Q.

Thus, if T be the date in years and T_0 the arbitrarily chosen date for which $t = 0$, then the date of peak production rate would be

$$T_{peak} = T_0 + (\ln N_0)/a. \tag{47}$$

Cumulative Production, Proved Reserves, and Cumulative Discoveries

The foregoing analysis pertains to a single quantity, such as cumulative production, and its variation with time during the complete production cycle. Actually, there is another important variable, based upon additional information, namely, proved reserves. Proved reserves, as defined by the Committee on Proved Reserves of the American Petroleum Institute, represents, essentially, oil in existing fields that has been proved by development drilling and is recoverable by existing installed equipment and technology. Estimates of proved reserves at the end of each year have been made annually for the United States since 1936 by the Proved Reserves Committee, and approximate figures, based upon various earlier estimates, are available back to 1900. Because additions to proved reserves are added annually only as new discoveries are made and older fields are developed, the figure for proved reserves is a conservative figure and is not intended to represent the ultimate amount of oil that the known fields will produce. Over- and under-estimates made in previous years are adjusted as new information becomes available by annual revisions. Proved-reserves estimates are therefore internally consistent and probably have a reliability within a range of a few percent.

A third significant quantity is that of cumulative proved discoveries. This does not represent independent data but is a derived quantity, defined in terms of the primary quantities, cumulative production and proved reserves.

If we let Q_p represent cumulative production, Q_r proved reserves, and Q_d cumulative proved discoveries, then Q_d is defined by the equation

$$Q_d = Q_p + Q_r. \tag{48}$$

In other words, all the oil that can be proved to have been discovered by a given time is the sum of the oil already produced plus the proved reserves.

The manner of variation of these three quantities during the complete production cycle, for a large region such as the United States, is indicated in Figure 12. The cumulative production curve, Q_p, will be a generally S-shaped, logistic-type curve, asymptotic to zero initially and to Q_∞ finally. The curve of proved reserves, Q_r, will be asymptotic to zero initially, and again at the end of the exploitation cycle, and will reach a maximum in the midrange of the cycle. The curve of cumulative discoveries, Q_d, will also be a logistic-type curve similar to that of cumulative production except that it will precede the latter in the midrange by some time interval Δt. This curve also will be asymptotic to zero initially and to the same value of Q_∞ finally as for the curve of cumulative production. This must be so because, as $t \to \infty$, $Q_r \to 0$, and equation (48) simplifies to

$$Q_d = Q_p = Q_\infty.$$

Taking the time derivative of equation (48) gives

$$dQ_d/dt = dQ_p/dt + dQ_r/dt, \tag{49}$$

the terms of which are the rates of proved discovery, or production, and of increase of proved reserves. When proved reserves reach their maximum value,

$$dQ_r/dt = 0,$$

and at that time equation (48) reduces to

$$dQ_d/dt = dQ_p/dt.$$

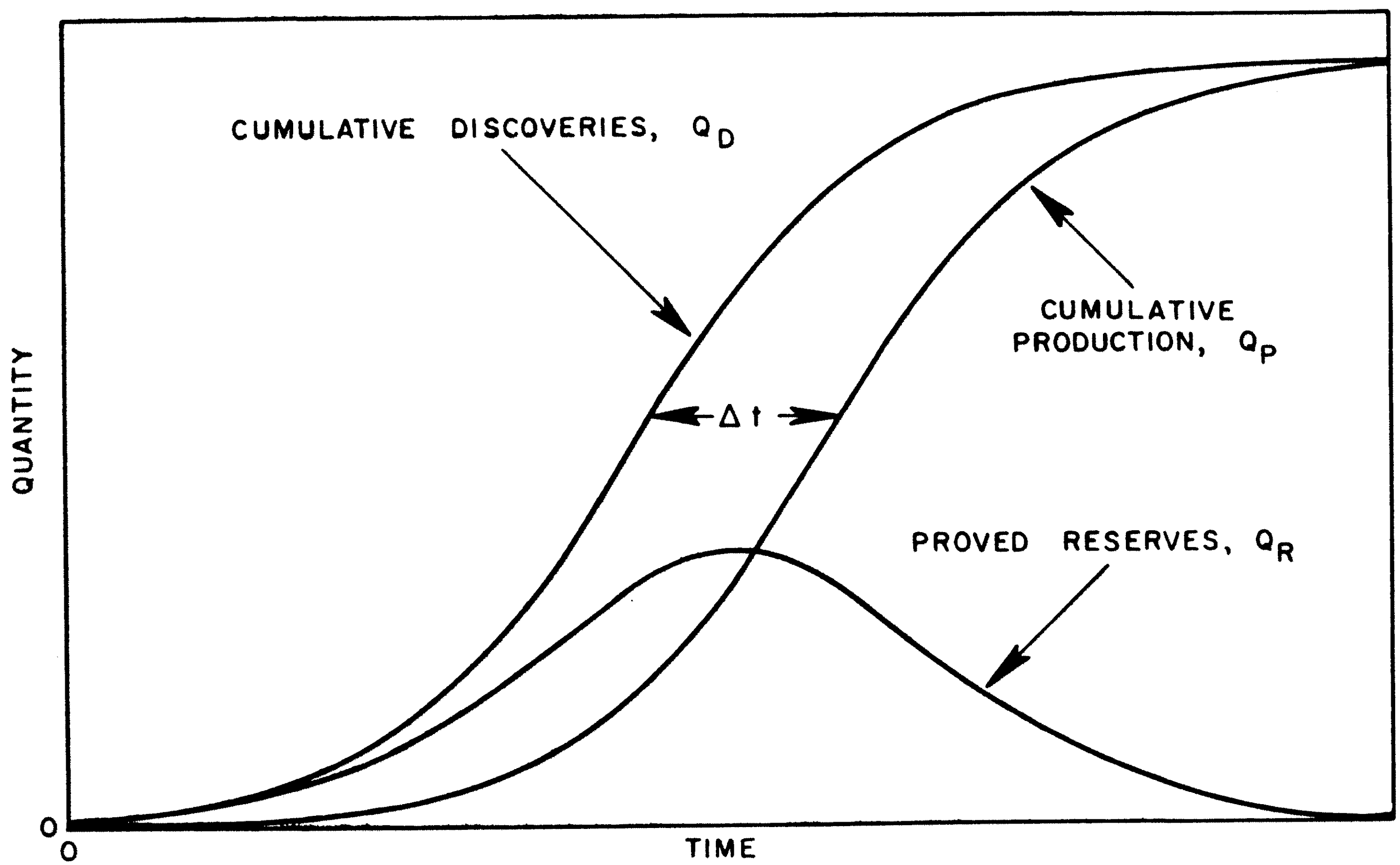

Fig. 12 - Variation of cumulative production Q_p, proved reserves Q_r, and cumulative discoveries Q_d during complete production cycle (Hubbert, 1962, Fig. 22).

This is the time at which the curve of the rate of production, which is still ascending, crosses that of the rate of discovery, already on its descent.

The curves Q_p, Q_r, and Q_d and their time derivatives, shown as functions of t in Figures 12 and 13 provide diagnostic evidence of the approximate stage of evolution in its complete cycle at any given time of the petroleum industry in a large area such as the United States. Because of the geometrical similarity between the curves of cumulative proved discoveries and of cumulative production, and the time lag Δt of the production curve with respect to that of discoveries, it follows that at any given time the discovery curve amounts to an approximate Δt-preview of the curve of production. Thus, if the curve of cumulative proved discoveries passes its inflection point, corresponding to the maximum rate of discovery, at a time t_1, then the curve of cumulative production will reach its inflection point and maximum production rate at a later time of approximately $t_1 + \Delta t$. The curve of proved reserves will reach its maximum at a time about halfway between, or at about $t_1 + \Delta t/2$.

Application to U.S. Petroleum Data as of 1962.---On March 4, 1961, President John F. Kennedy addressed a letter to the President of the National Academy of Sciences asking the Academy to advise him with respect to natural-resources policy. In response, the Academy appointed a Committee on Natural Resources to make the requested study and prepare reports for President Kennedy. I was a member of that Committee and directed the study and wrote the Committee's report on Energy Resources (National Academy of Sciences-National Research Council Publication 1000-D, 1962).

For this study, the technique used in 1956 was no longer appropriate because, during the intervening five years, the petroleum estimates of 1956 of 150 to 200 billion barrels had been progressively increased by various authors to 250, 300,

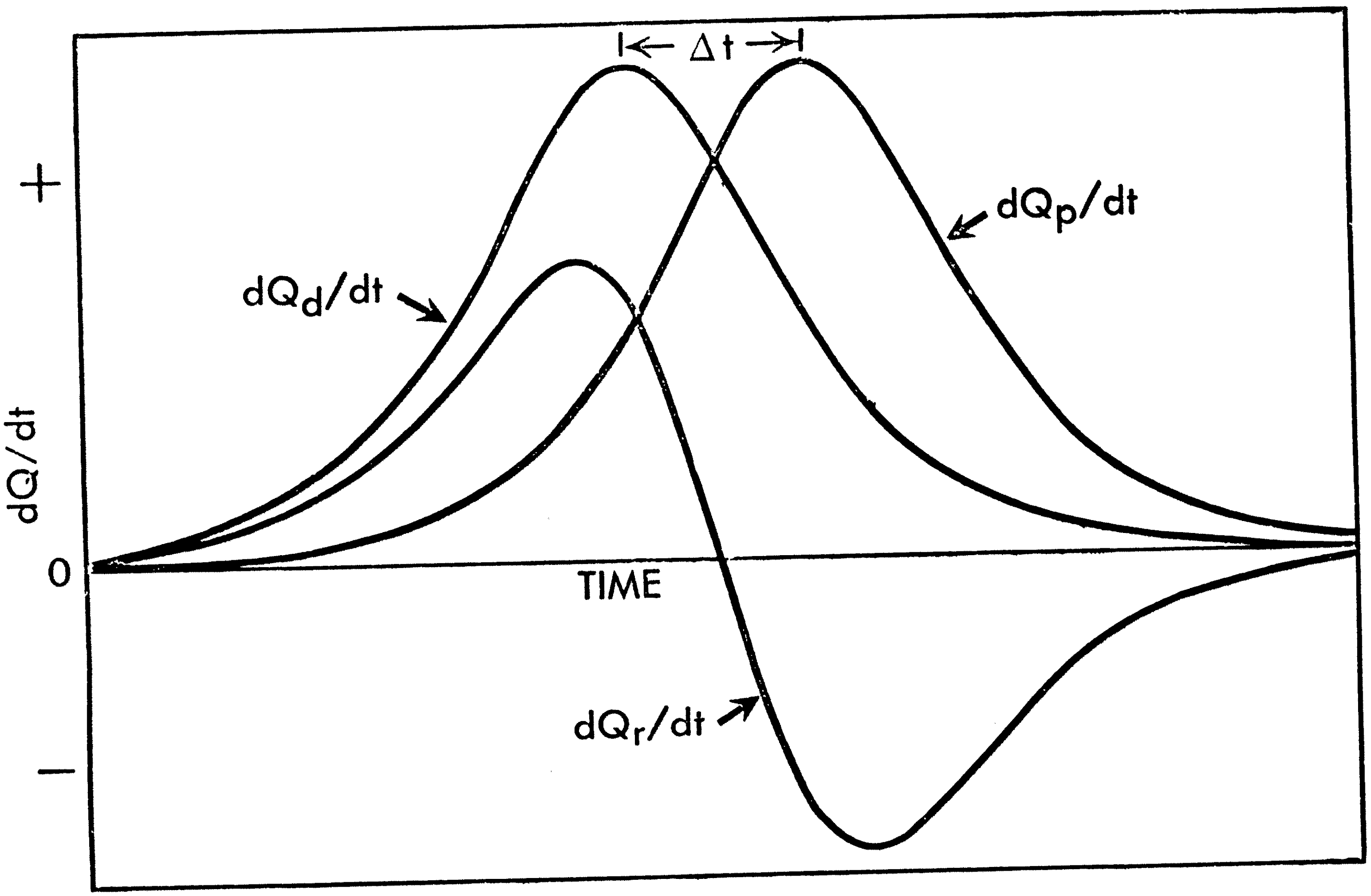

Fig. 13 - Rates of production and of proved discoveries, and of increase of proved reserves during complete production cycle (Hubbert, 1962, Fig. 24).

400, and eventually 590 billion barrels for the ultimate amount of crude oil to be produced in the Lower-48 states and adjacent continental shelves. We thus were confronted with approximately a four-fold range in the magnitudes of these estimates. For the lowest, the United States would reach its maximum rate of oil production at about 1965; for the highest, this would be delayed almost to the year 2000.

Accordingly, it became necessary to disregard the various a priori estimates of Q_∞, and instead let the historical data on discovery and production determine the approximate stage that the petroleum industry had reached in its evolutionary cycle. Of primary interest was the determination of such critical dates as those of the maximum rates of discovery and production, and of the maximum of proved reserves. From these data, as a secondary objective, an estimate of the magnitude of the ultimate production Q_∞ could be derived.

The theoretical basis for this analysis was that shown graphically in Figures 12 and 13. The actual data for the curves of cumulative production, proved reserves, and cumulative proved discoveries for U.S. crude oil from 1901 to 1962 are shown graphically in Figure 14. By visual inspection, the curve of cumulative proved discoveries had passed its inflection point at about 1957; proved reserves appeared to be at about their maximum in 1962; and the time delay Δt between the curve of production and that of discoveries was approximately 10.5 years, and had been so since 1925. Accordingly, the production rate should reach its maximum 10-12 years after 1957, the peak in the rate of discovery, or 5 to 6 years after the proved-reserves maximum in 1962.

For more precise calculations, the three curves of Figure 15 needed to be fitted by analytical equations so that analytical derivatives could be obtained with which to compare the actual annual rates of production, of discovery, and

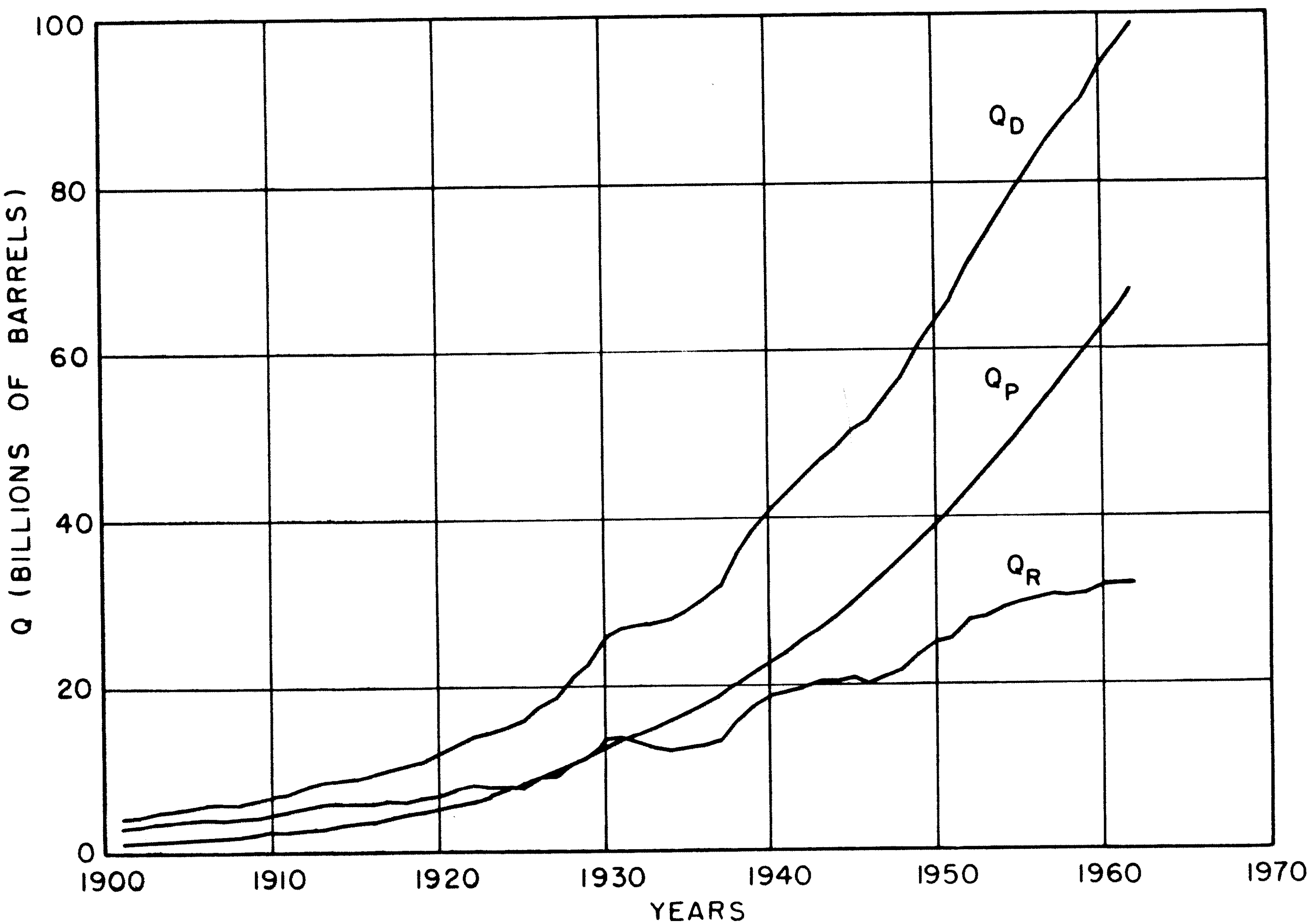

Fig. 14 - U.S. cumulative crude-oil production, proved reserves, and cumulative proved discoveries, 1901-1962 (Hubbert, 1962, Fig. 25).

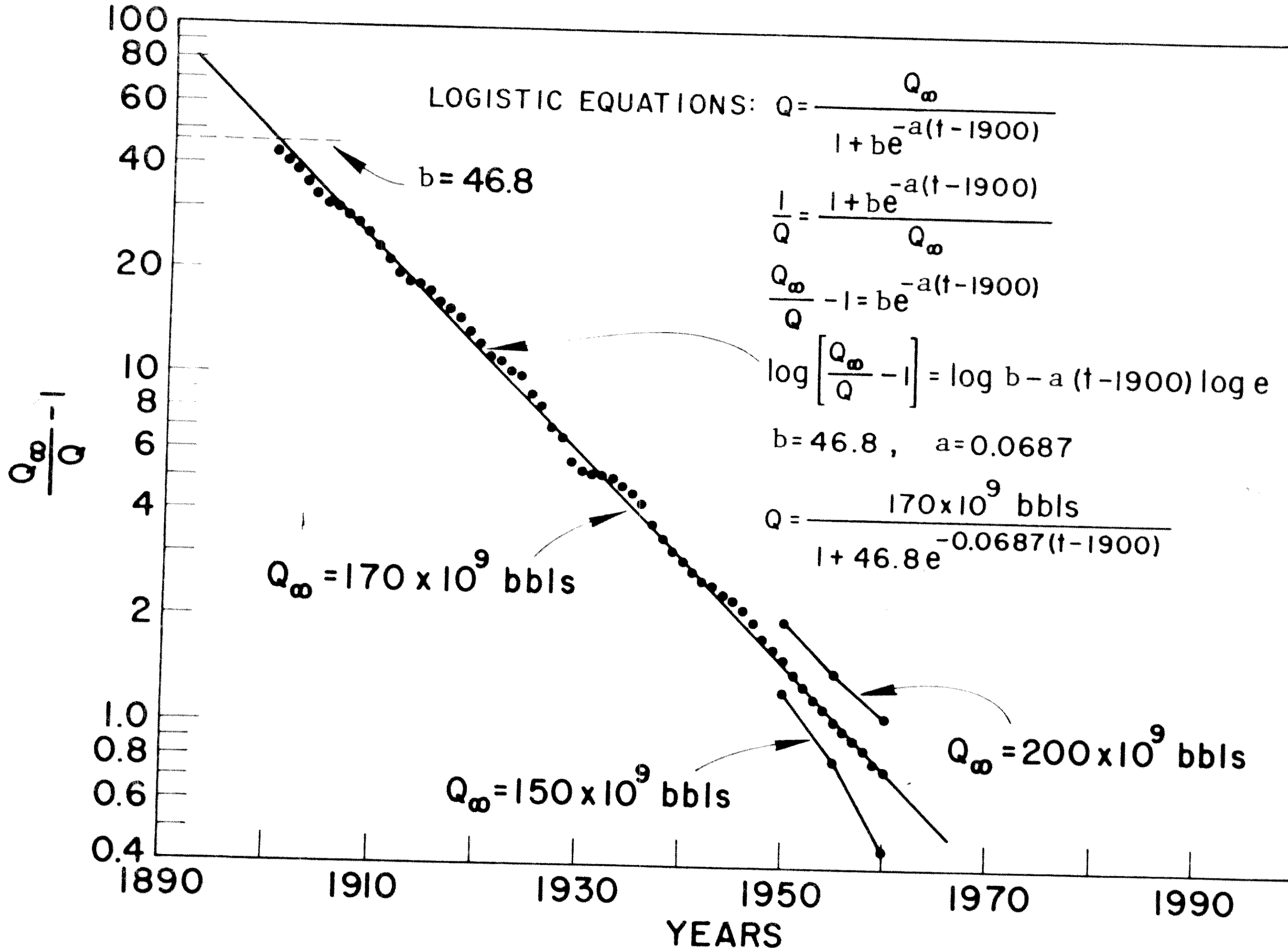

Fig. 15 - Graphical method used in 1962 for determining the constants of the logistic equation for U.S. cumulative crude-oil discoveries.

of increase of proved reserves. For this purpose, various forms of empirical equations were tested, but none gave satisfactory agreement with the data until finally the logistic equation was tried and found to fit the data with remarkable fidelity.

The curve of cumulative proved discoveries, having Δt more years of data than that of cumulative production, was fitted first. This was done by the iterative procedure of equations (41) and (42), as illustrated in Figure 9. The results, shown graphically in Figure 15, were

$$Q_\infty = 170 \times 10^9 \text{ bbl},$$

$$a = 0.0687/\text{yr}.$$

Then the year 1901 was taken for t_o, and

$$N_o = (Q_\infty - Q_o)/Q_o = 46.8.$$

With these parameters, the equations for Q_d, Q_p, and Q_r were

$$\left.\begin{aligned} Q_d^* &= (170 \times 10^9)/[1 + 46.8e^{-a(t-1901)}], \\ Q_p &= (170 \times 10^9)/[1 + 46.8e^{-a(t-1911.5)}], \\ Q_r &\simeq Q_d - Q_p. \end{aligned}\right\} \qquad (49)$$

The graphs of the actual data for Q_p, Q_r, and Q_d superposed upon the theoretical curves are shown in Figure 16. Superposition of the annual increments of proved reserves upon the theoretical derivative curve is shown in Figure 17, and the corresponding superposition of the rates of discovery and of production upon their respective derivative curves are shown in Figure 18.

---

*In the Academy report, t_o was given as 1900, but the data were as of the end of each year. Hence the end of 1900 is actually 1901.0

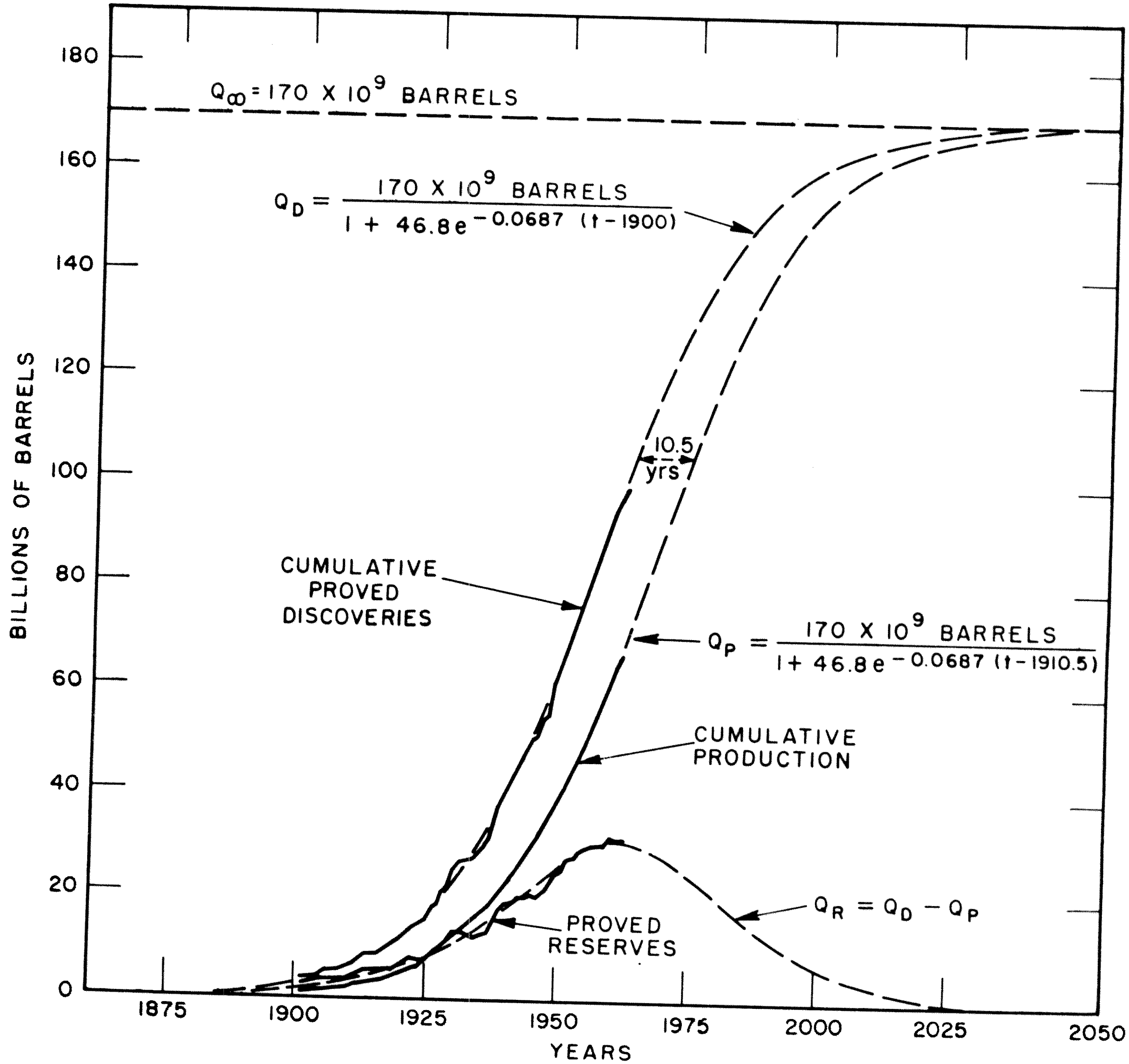

Fig. 16 - Cumulative proved discoveries, production, and proved reserves of U.S. crude oil to 1962 with graphs of the corresponding logistic equations (Hubbert, 1962, Fig. 27).

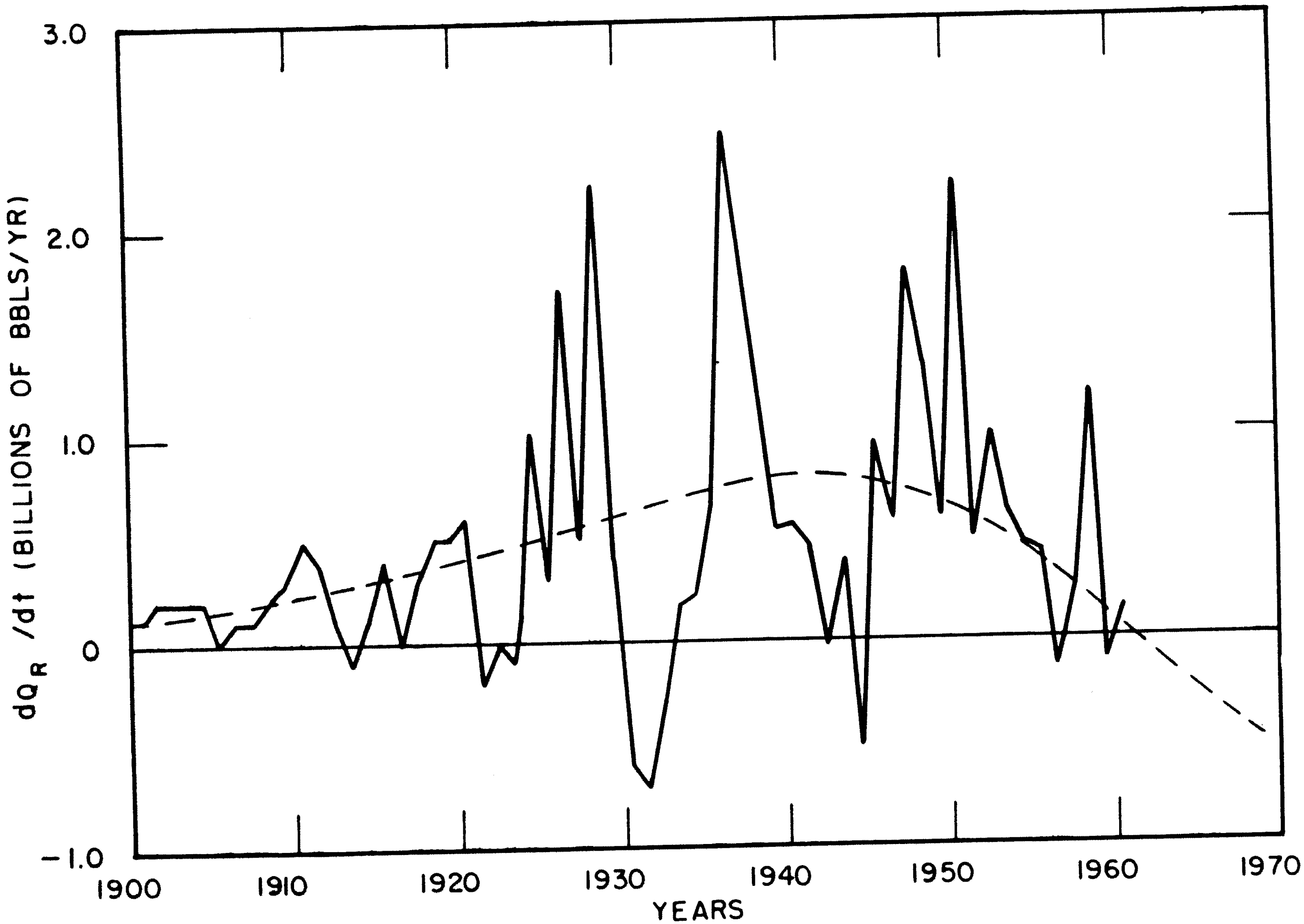

Fig. 17 - Annual increments of proved reserves of U.S. crude oil, 1900-1961, superposed upon the theoretical curve from the logistic equation (Hubbert, 1962, Fig. 29).

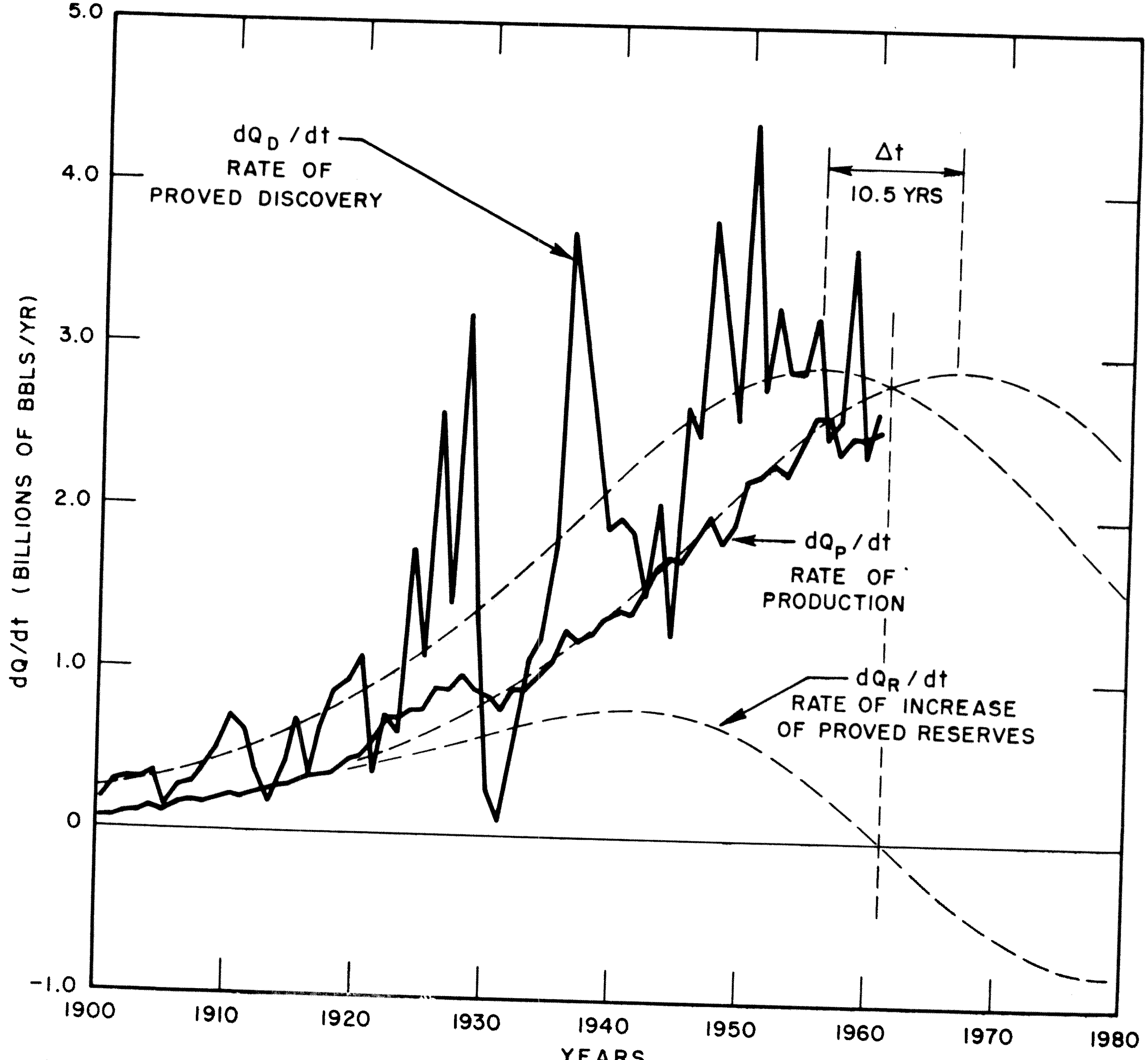

Fig. 18 - Annual rates of production and of proved discovery, 1900-1961, superposed upon derivative curves of the logistic equations (Hubbert, 1962, Fig. 28).

In Figure 17 it is seen clearly that the curve of increase of proved reserves had already gone through its positive loop corresponding to increasing reserves, and was crossing the zero line into its negative loop at just about 1962. This marks the date of the peak of proved reserves.

In Figure 18, the annual discoveries fluctuate widely, but their oscillations still follow the derivative curve and indicate that the peak region had been passed at about 1957. The peak in the rate of production can accordingly be expected to occur near the end of the 1960-decade.

These were the interpretations made graphically in the Academy report of 1962. Although this was not done at the time, more precise figures can be obtained from the equations (49). As shown in equation (46), the peak rate of discoveries occurs at the time

$$t \simeq t_o + \ln N_o/a.$$

Hence, from equation (49), the peak discovery rate should occur at the time

$$t = 1901 + \ln 46.8/0.0687$$
$$= 1957.0.$$

The corresponding date for the peak in the production rate would be 10.5 years later, or 1967.5, and the peak of proved reserves would occur at 1962.25. All of these figures are consistent with those obtained from the graphical interpretation of the data.

An even more informative procedure consists in plotting the linear graphs of the logistic equations of cumulative proved discoveries and cumulative production from 1901 to 1962 on semi-logarithmic graph paper. The corresponding equations for discoveries and production are

$$\log N_d = \log N_o - (a \log e)(t - t_o)$$

and

$$\log N_p = \log N_o - (a \log e)[t - (t_o + \Delta t)],$$

where t_o = 1901, and Δt = 10.5 years.

In these equations, the peak discovery and production rates each occur when $Q = Q_\infty/2$, or when

$$N = [Q_\infty/(Q_\infty/2)] - 1 = 1.$$

Hence the dates at which the respective linear graphs cross the line $N = 1$, or $\log N = 0$, are the dates of peak rates of discovery and of production. These graphs for the cumulative discovery and production data at 5-year intervals from 1901 to 1962 are shown in Figure 19. Note that by 1962 the cumulative-discoveries curve had already crossed the line $N = 1$ at 1957.0, and that the linear graph for cumulative production had been parallel to the discoveries curve since 1925 with the time lag of 10.5 years. Hence only a modest 6-year extrapolation of this curve was required to reach the line $N = 1$, corresponding to the peak in the rate of production at the year 1967.5.

Thus, the combined data on production, proved reserves, and proved discoveries of crude oil in the Lower-48 states were by 1962 sufficient to establish that the peak in the rate of production would have to occur at about the end of the 1960-decade with a range of uncertainty of not more than about 3 years. The corresponding maximum production rate from equation (39) would be about

$$dQ_p/dt \approx aQ_\infty/4,$$

or

$$(0.0687 \times 170 \times 10^9)/4 = 2.92 \times 10^9 \text{ bbl/yr}.$$

The Decade of 1962-1972.— Although the reports of the National Academy Committee were released to the public by President Kennedy in January 1963, the influence on public policy of the Academy report on energy resources was essentially nil. The estimate of 590 billion barrels for the ultimate crude-oil production

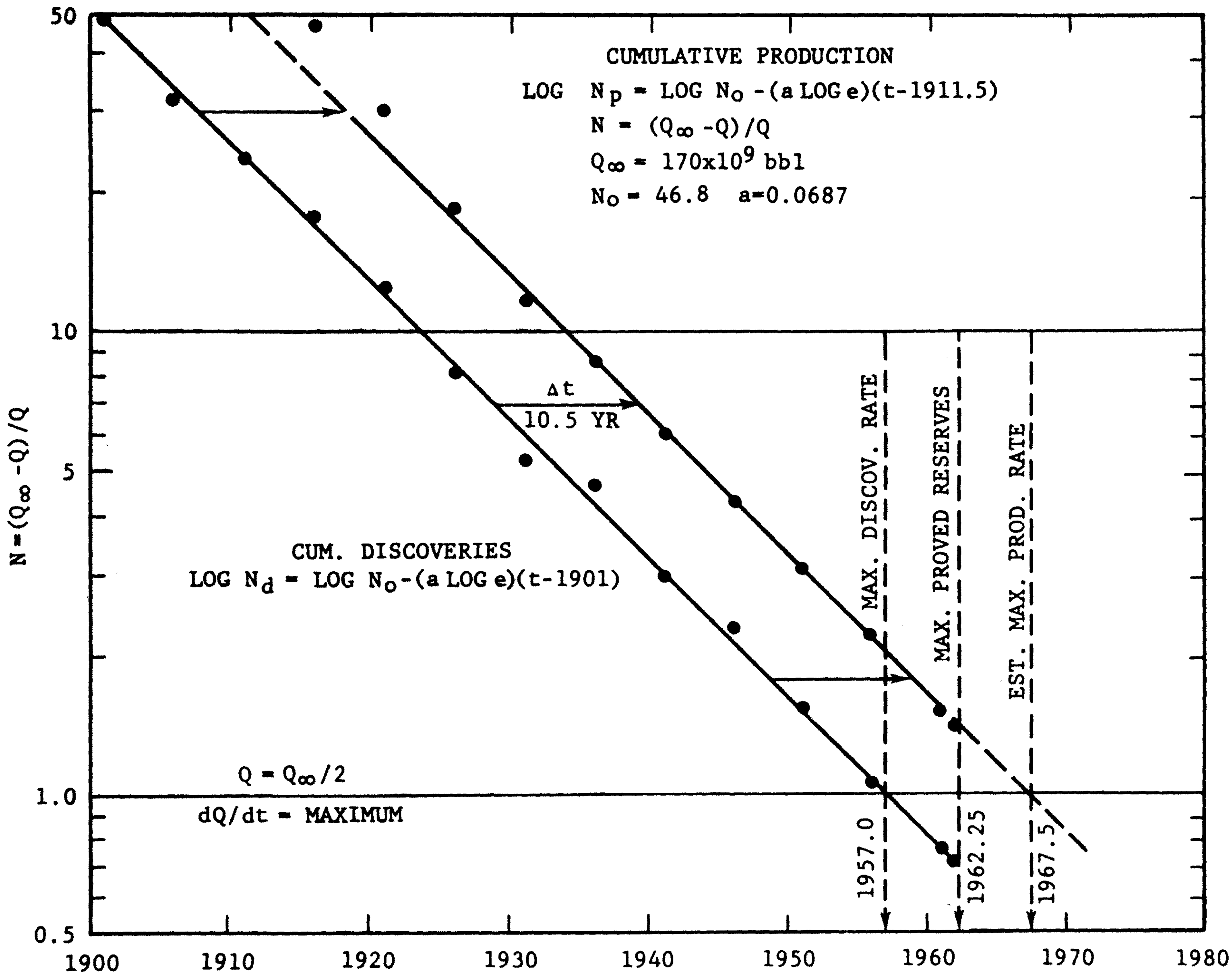

Fig. 19 - Linear graphs of the logistic equations of U.S. cumulative crude-oil discoveries and production, 1900-1962.

from the Lower-48 states had been given to the Academy Committee by V. E. McKelvey, Assistant Chief Geologist of the U.S. Geological Survey, as the official estimate of the USGS. It was cited in the Academy report but, because it could not remotely be reconciled with the petroleum-industry data, it had to be rejected. During the next 5 years, substantially the same figures, modified slightly from year to year, of about 600 billion barrels of crude oil and 2,500 trillion cubic feet of natural gas, were published repeatedly in U.S. government and other publications by McKelvey, either alone or in collaboration with D. C. Duncan (Duncan and McKelvey, 1963; McKelvey and Duncan, 1965; McKelvey, 1967).

In the meantime, the curve of annual crude-oil production from the Lower-48 states continued on the linear upward trend that had prevailed since 1932, except for a distortion from 1957 to 1970 associated with the first Suez crisis and successive Middle East disturbances and the Vietnam War. After 1957, the production rate fell below the linear trend, but by 1970 it rose to 3.24 billion barrels per year, which was just about on the trend. During 1971 it declined slightly to 3.18 billion barrels per year. Consequently, during the decade 1962-1972, there was no perceptible evidence from the curve of annual production alone of the imminence of an impending peak and subsequent decline of the annual rate of U.S. crude-oil production.

One of the first alarms over impending trouble came in the spring of 1969 when the annual report of the Committee on Natural-Gas Reserves of the American Gas Association was released, giving the proved reserves of natural gas as of the end of 1968. The report showed that by the end of 1968 natural-gas proved reserves for the Lower-48 states, which had been increasing steadily since 1947, had dropped 7.2 trillion cubic feet from 289.3 trillion cubic feet at the end of 1967 to 282.1 by the end of 1968. During the following year, the proved reserves dropped by 12.2 trillion cubic feet, as compared with the annual production rate of 20.7 trillion cubic feet.

By 1971, the evidences of impending declines in the rates of oil and gas production were sufficiently clear that the U.S. Senate Committee on Interior and Insular Affairs, under the Chairmanship of Senator Henry M. Jackson, began a new series of hearings on National Fuels and Energy policy. As of July 23, 1971, Senator Jackson addressed a letter to the Secretary of the Interior, Rogers C. B. Morton, requesting my assistance on statistical work for the Committee. What I was asked to do was to bring my earlier studies on energy resources up to date for the use of the Committee. The result was the report, *U.S. Energy Resources, A Review as of 1972*, which was released as a Committee Print in June 1974.

In Figure 20 the curves of cumulative proved discoveries, cumulative production, and proved reserves from 1900 to 1972 are shown as solid-line curves superposed upon the respective mathematical curves, shown as dashed lines. The respective logistic equations are also shown in the figure.

With 10 more years of data after the Academy report of 1962, the best value for Q_∞ was still 170×10^9 bbl and the growth constant a was still 0.0687/yr. The time interval Δt had been increased from 10.5 to 11.0 years, and 1930 was taken for t_0. The corresponding value of N_0 became 6.17. By 1972, proved reserves are plainly seen to have passed their peak about 1962, and the curves of cumulative discoveries and of cumulative production are accurately following their respective logistic curves. From the logistic equations the dates for the maximum rates of discovery and of production, and of the maximum of proved reserves are found to be:

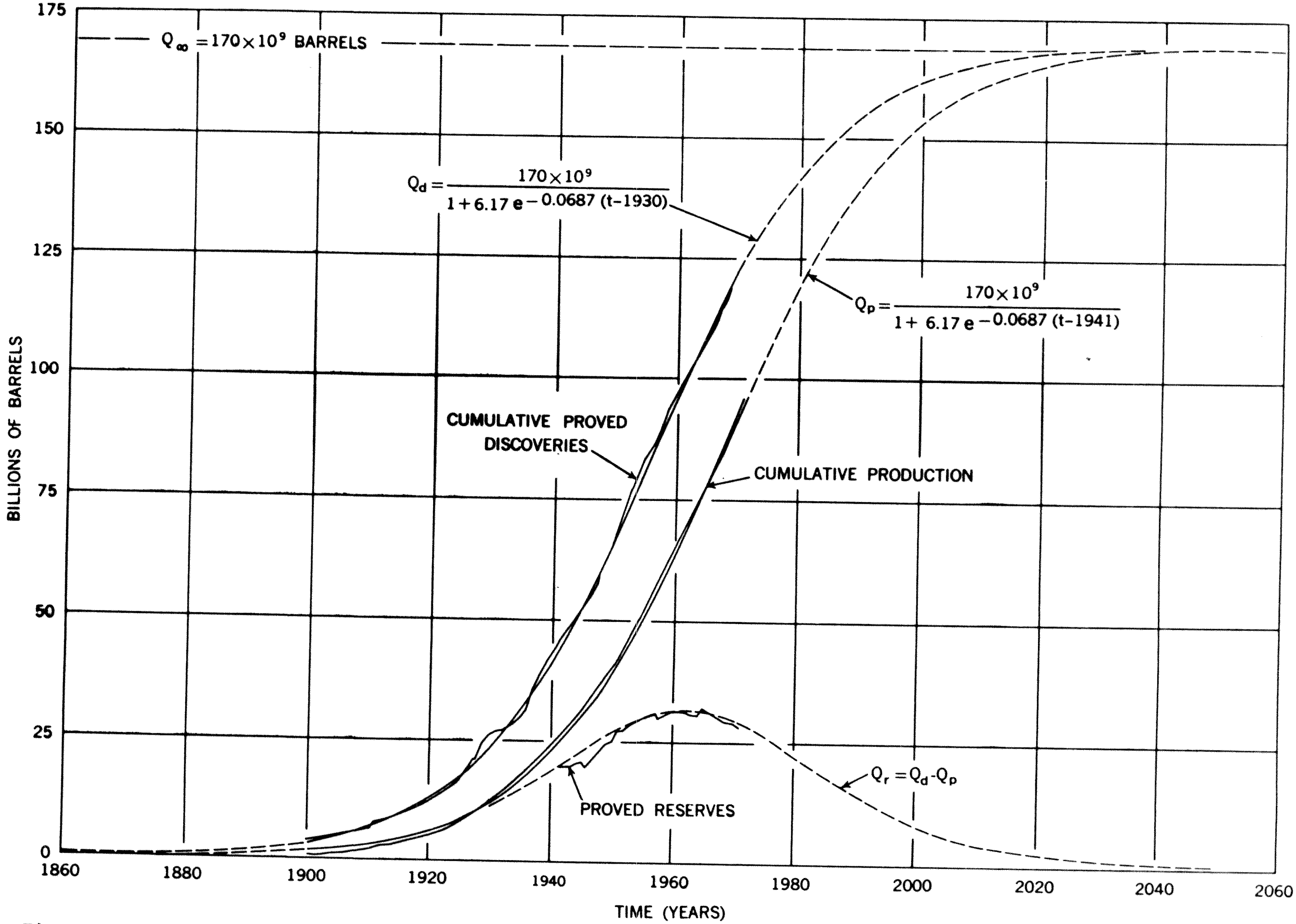

Fig. 20 - Cumulative production and proved discoveries, and proved reserves of crude oil from the U.S. Lower-48 states, 1900-1972, with the corresponding logistic curves (Hubbert, 1974, Fig. 36).

Discoveries	1956.5,
Production	1967.5,
Proved reserves	1962.0.

The time derivatives of these three curves, with the corresponding annual data, are shown in Figures 21 to 24, and the estimated complete cycle for crude-oil production, in Figure 25. Figure 21 shows the rate of increase of proved reserves superposed upon the mathematical derivative. The curve completed its positive loop and crossed the zero line in 1962, and by 1972 was approaching the low point of its negative loop. The rates of discovery and of production are both shown in Figure 22, and separately in Figures 23 and 24. In Figure 23, it is unmistakable that the discovery rate passed its peak before 1960 and is well advanced in its declining phase.

The curve of the rate of production in Figure 24 still shows no definite evidence that its peak has been reached. Instead of reaching a maximum about 1968, the curve fell below the mathematical curve after 1957 and then rose steeply from 1960 to 1970. Whether the slight reversal in 1971 represents the beginning of the decline is an open question. However, the composite evidence of all the data indicate that the reversal of the production-rate curve, if it has not already occurred in 1971, must happen in the very near future.

The estimate of the complete production cycle as of 1972 is shown in Figure 25. Cumulative production by 1972 amounted to 96×10^9 bbl and proved reserves plus additional oil in fields already discovered amounted to another 47×10^9 bbl, giving a total of 143×10^9 bbl for the ultimate amount of oil to be produced from fields already discovered. Then with 170×10^9 bbl for Q_∞, only 27×10^9 bbl are left for future discoveries. Another informative aspect of Figure 25 is the time span involved. The time required to produce the first 10 percent of Q_∞, or 17 billion barrels, was from 1860 to 1932. That required

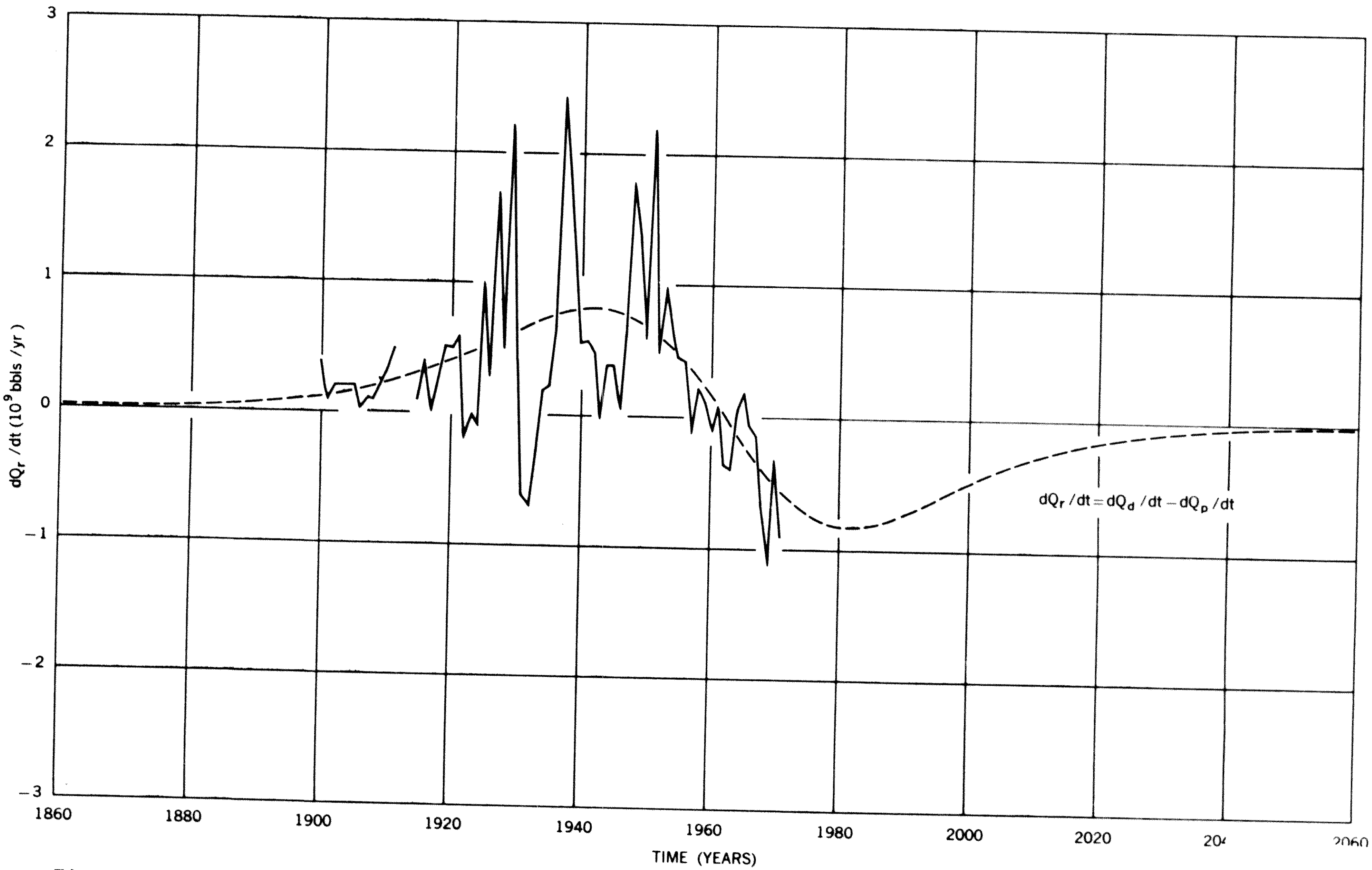

Fig. 21 - Annual increments of U.S. proved reserves of crude oil, 1900-1971, superposed upon the theoretical curve from the logistic equation (Hubbert, 1974, Fig. 40).

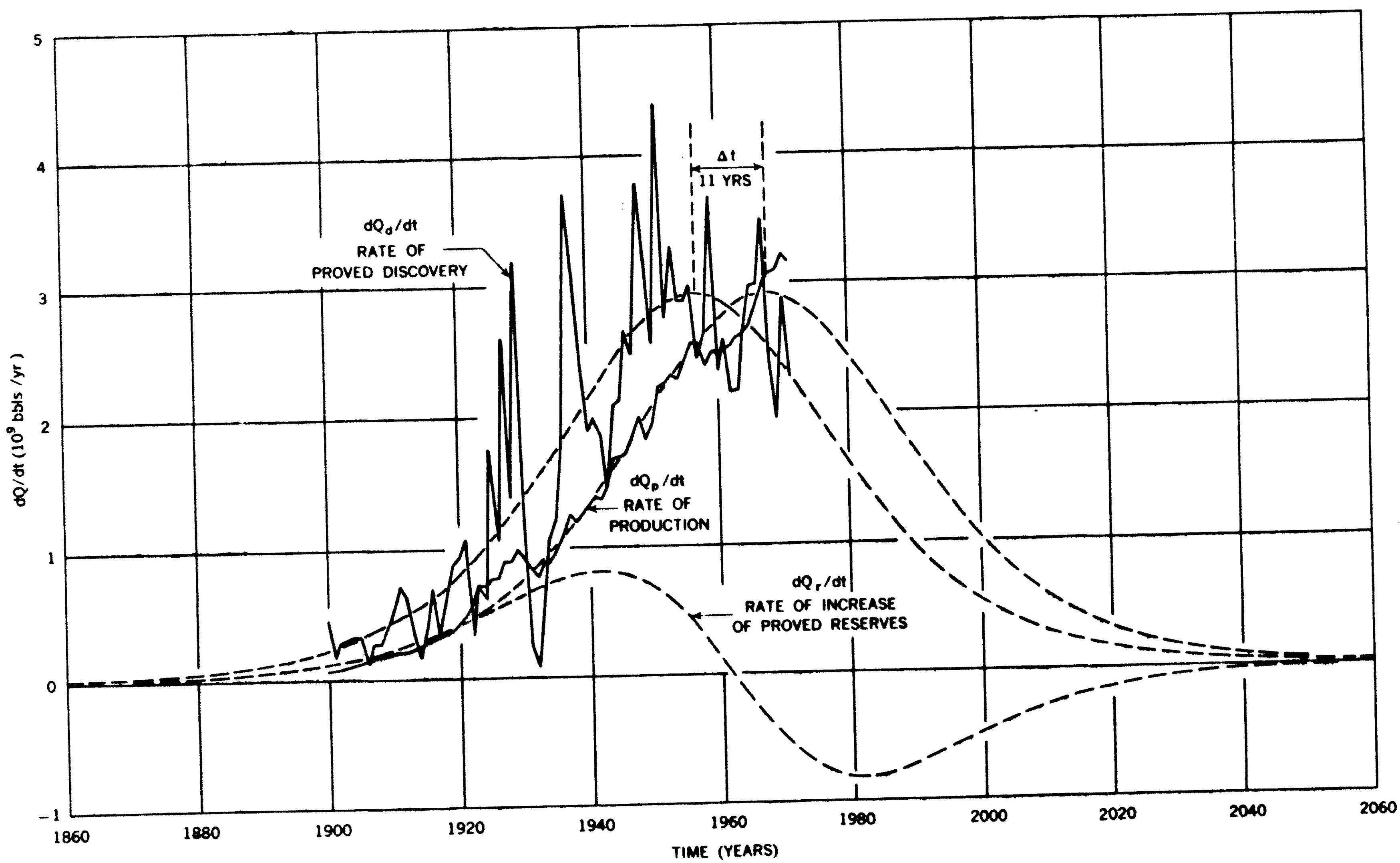

Fig. 22 - Annual production and proved discoveries of U.S. crude oil, 1900-1971, superposed upon curves of derivatives of logistic equations (Hubbert, 1974, Fig. 37).

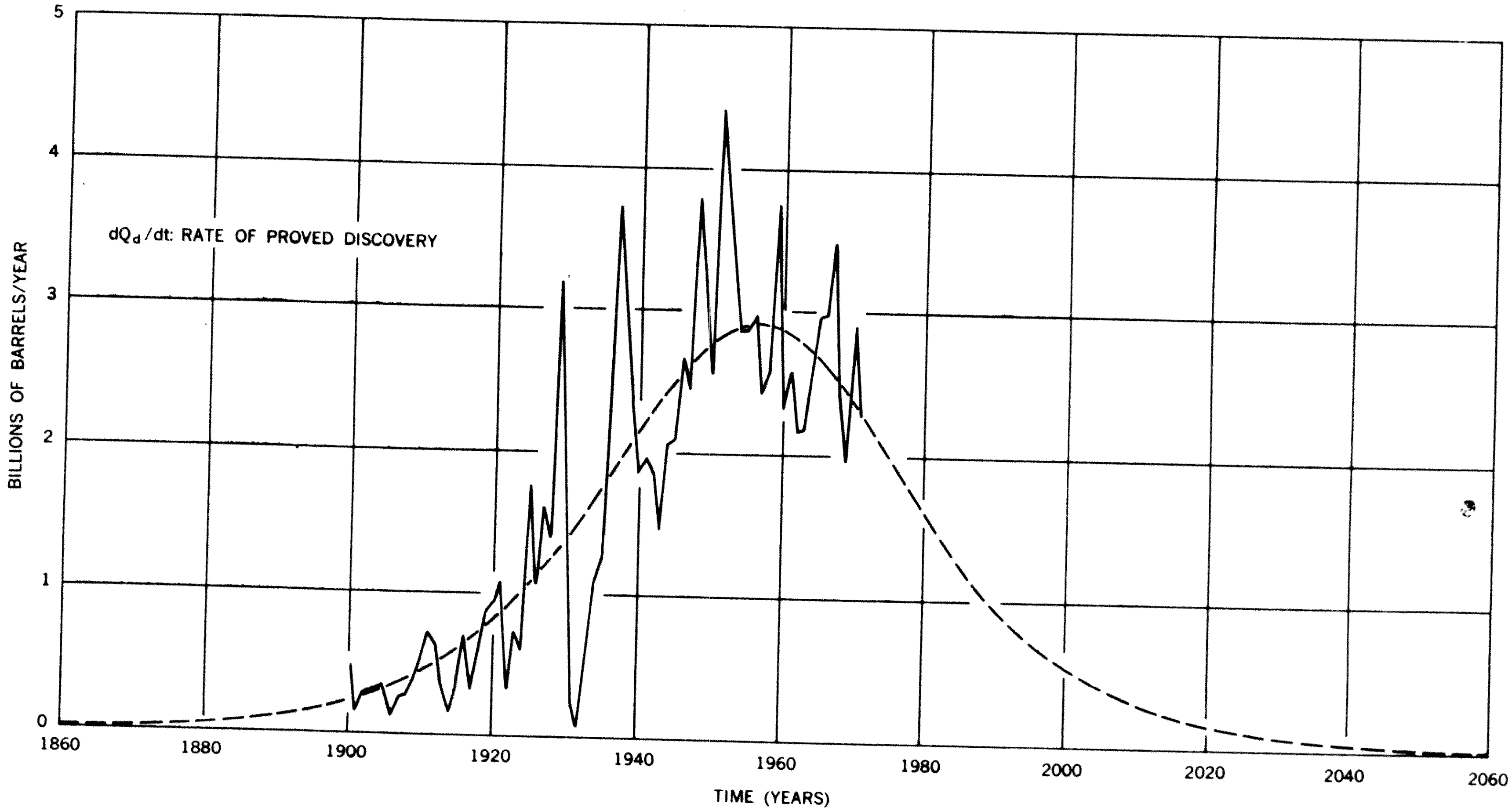

Fig. 23 - Annual proved discoveries of U.S. crude oil, 1900-1971, superposed upon curve of the derivative of the logistic equation (Hubbert, 1974, Fig. 38).

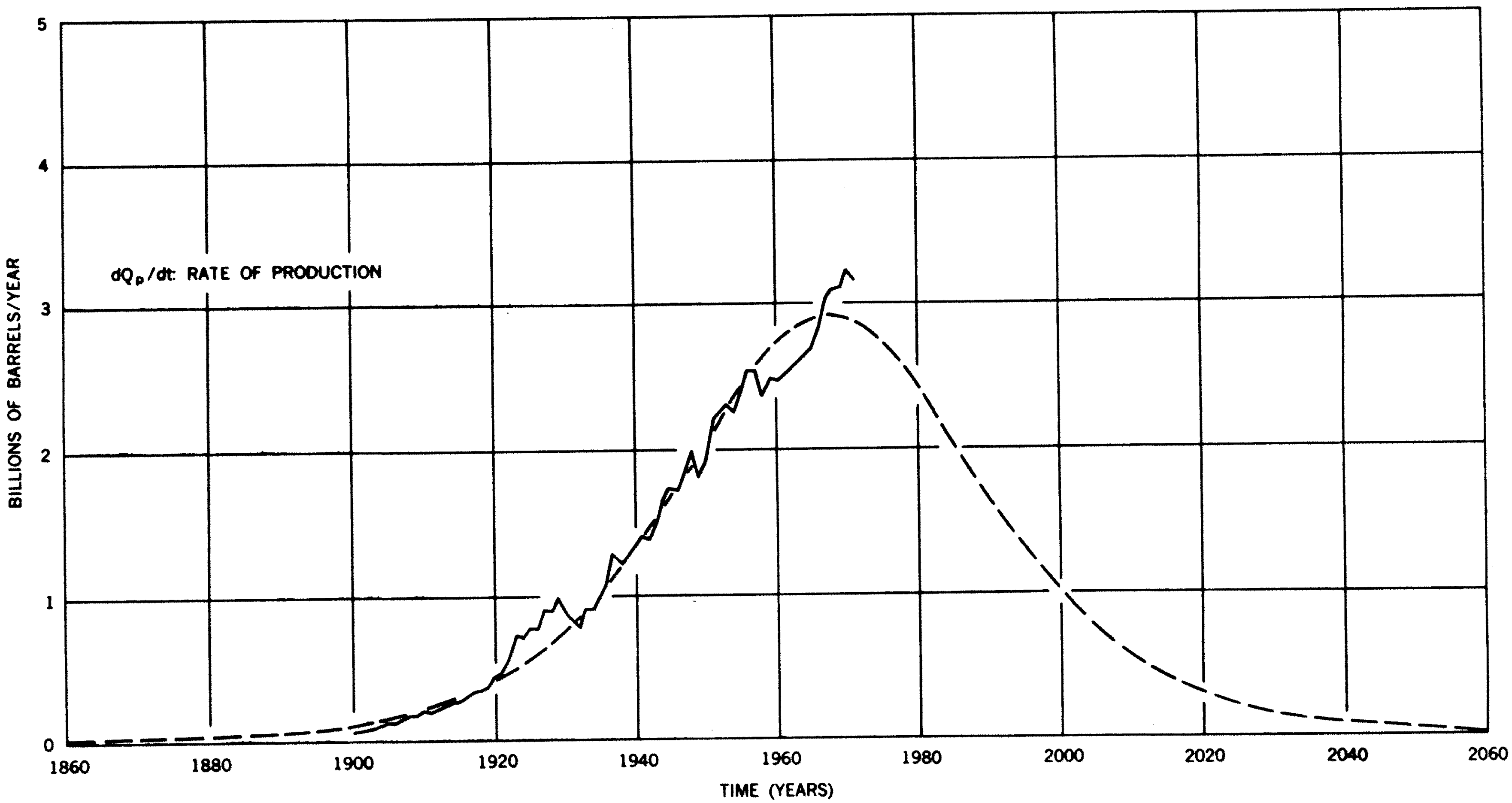

Fig. 24 - Annual production of U.S. crude oil, 1900-1971, superposed upon curve of the derivative of the logistic equation (Hubbert, 1974, Fig. 39).

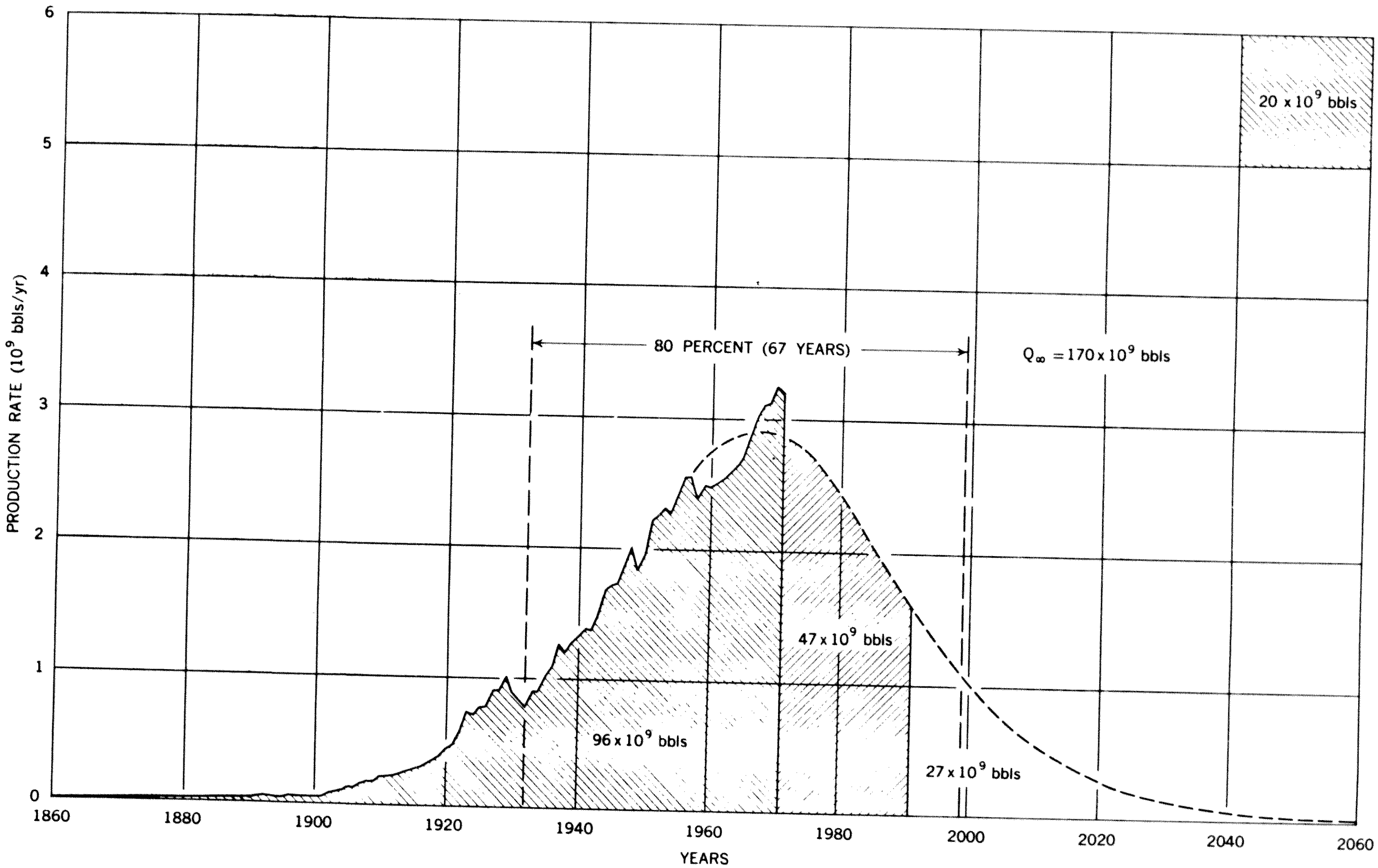

Fig. 25 - Complete cycle of crude-oil production from U.S. Lower-48 states as estimated in 1972 (Hubbert, 1974, Fig. 51).

for the next 80 percent, computed from the logistic equation, would be the 67-year period from 1932 to 1999, and the last 10 percent would occur after 1999.

The Period from 1972 to 1980.— Until 1972, our principal concern was the prediction of the future of the rate of U.S. crude-oil production. The mathematical curve of the rate of production passed its maximum about 1967-1968, but the production rate continued to increase sharply until 1970, yet all the evidence indicated that a decline was inevitable in the very near future. We now have 7 to 8 more years of data by means of which the pre-1972 predictions can be evaluated.

Figure 26 shows the linear graphs of the logistic equations for cumulative discoveries and production of crude oil in the Lower-48 states for the period 1925-1973. In this case a new determination of the logistic constants was made with the results:

$$Q_\infty = 170 \times 10^9 \text{ bbl},$$

$$t_o = 1925,$$

$$\Delta t = 10.7 \text{ yr},$$

$$N_o = 9.05,$$

$$a = 0.0674/\text{yr}.$$

Both curves by 1972 had crossed the line $N = 1$, corresponding to the dates of the respective maximum rates of discovery and production. The date for the discovery maximum rate was the year 1957.7, and that for the maximum production rate, 1968.4.

The curves of cumulative proved discoveries, cumulative production, and of proved reserves have been plotted to 1979 in Figure 27. These are superposed on the logistic curves of 1972 as a means of comparing the more recent developments

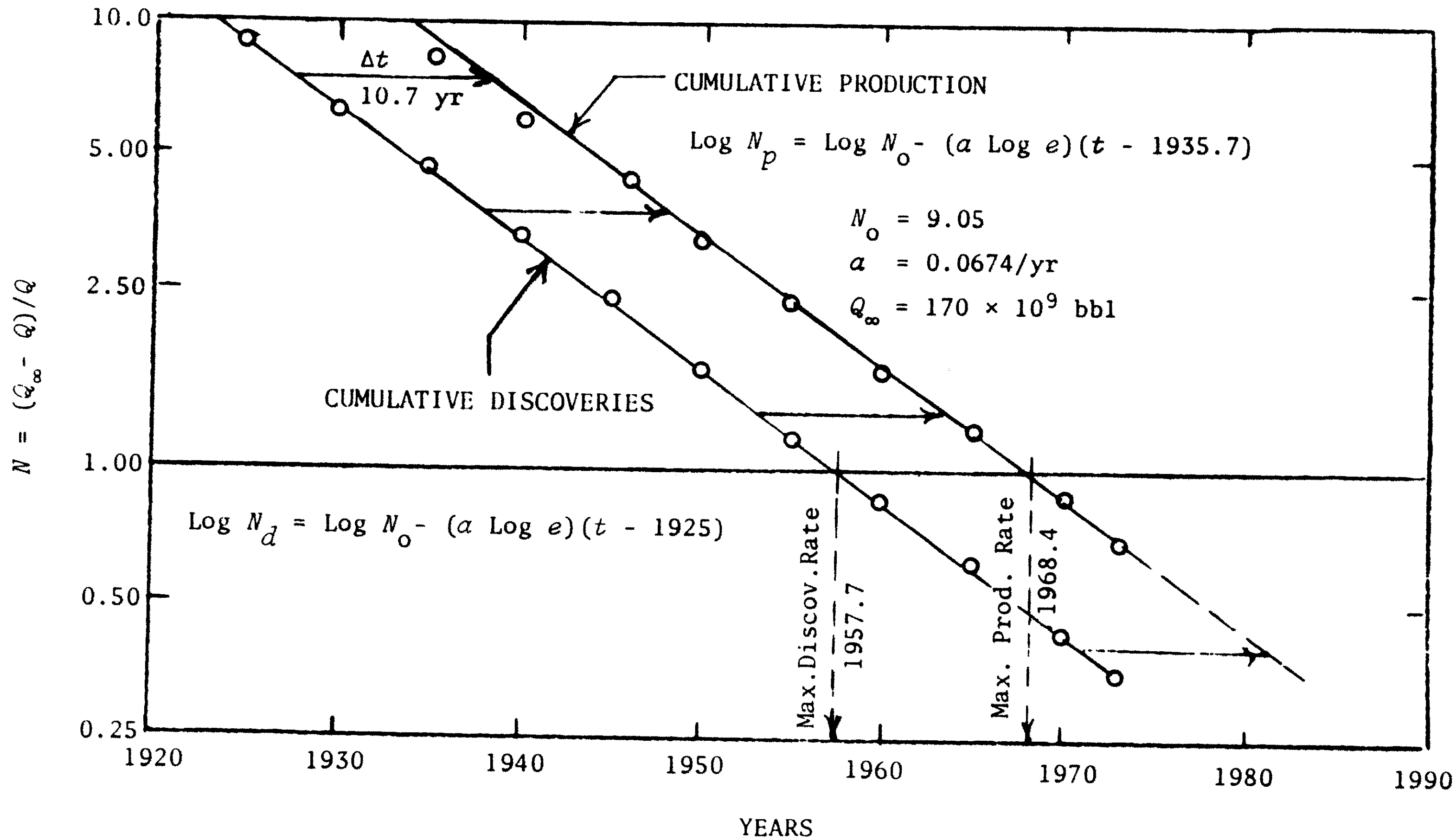

Fig. 26 - Linear graphs of the logistic equations for cumulative proved discoveries and cumulative production of crude oil from U.S. Lower-48 states, 1925-1973.

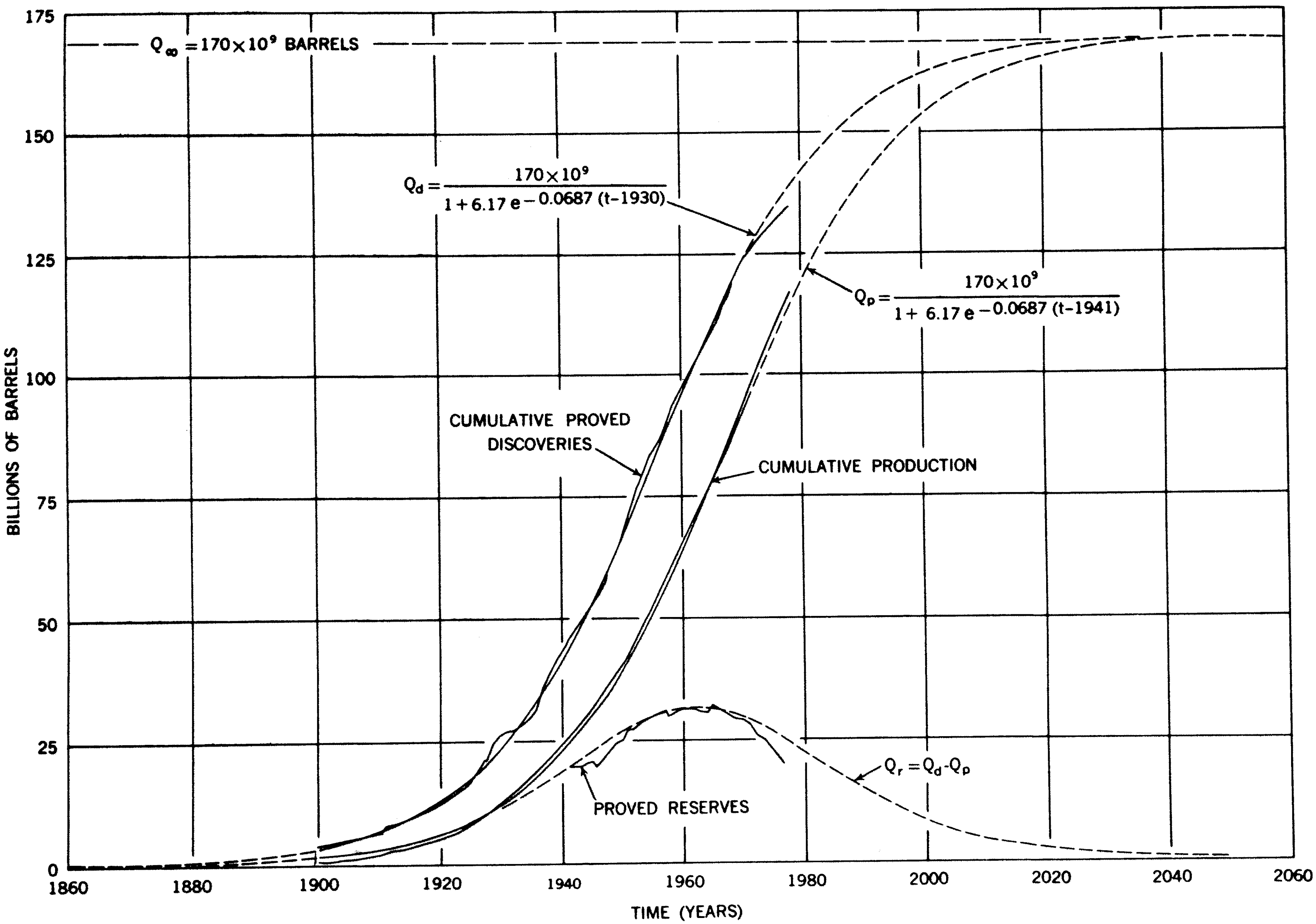

Fig. 27 - Cumulative production, proved reserves, and cumulative proved discoveries of U.S. crude oil, 1900-1978, superposed upon logistic curves of 1972.

with the 1972 predictions. It will be noted that the cumulative production is following the logistic curve rather closely, but both the curves of proved reserves and of cumulative discoveries have fallen significantly below their respective mathematical curves. From this figure, it appears that cumulative discoveries may fall short of the 170-billion-barrel asymptote for Q_∞ by as much as 7 to 10 billion barrels. Should this be so, cumulative production will have to do the same.

This is shown even more clearly in the derivative curves of Figures 28 and 29. Figure 28 shows that the annual rate of increase of proved reserves (which has been negative since 1962) has been well below the mathematical curve during the 1970-decade. The mathematical curve shows a minimum of about 0.8 billion barrels per year occurring in 1980 for the rate of decline of proved reserves. During the period 1973-1979 the average rate of decline was close to 1.5 billion barrels per year, indicating that about one-half of the production during that period was obtained by withdrawal from proved reserves.

Figure 29 shows that from 1972 to 1979 even the high points in the oscillatory curve of annual discoveries were below the mathematical derivative curve of the 1972 logistic equation. Figure 30 shows the corresponding comparison for the rate of production. The year 1970 was indeed the year of peak production, with the production rate falling steeply from its maximum of 3.24 billion barrels in 1970 to 2.45 in 1979.

In view of the departures of the curves of cumulative proved discoveries and of proved reserves during the 1970-decade, as is shown in Figure 27, from the logistic curves of 1972, new calculations have been made with the particular objective of obtaining a new estimate for the magnitude of Q_∞. One procedure has been to make new determinations of the constants of the logistic equation using the data from 1900 to 1980. A second procedure was based upon the linear

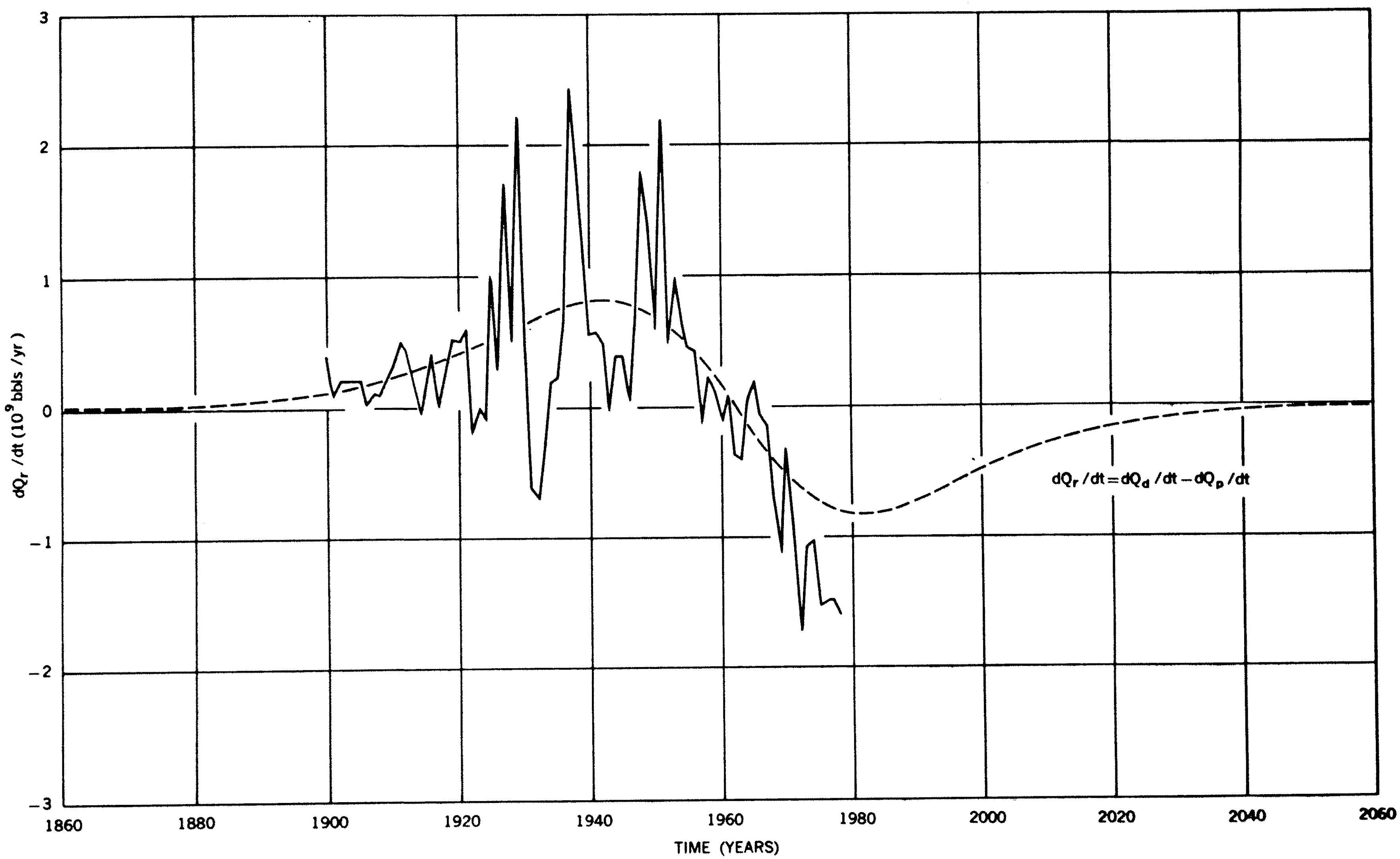

Fig. 28 - Annual increments of U.S. proved reserves of crude oil, 1900-1979, superposed upon theoretical curve of 1972.

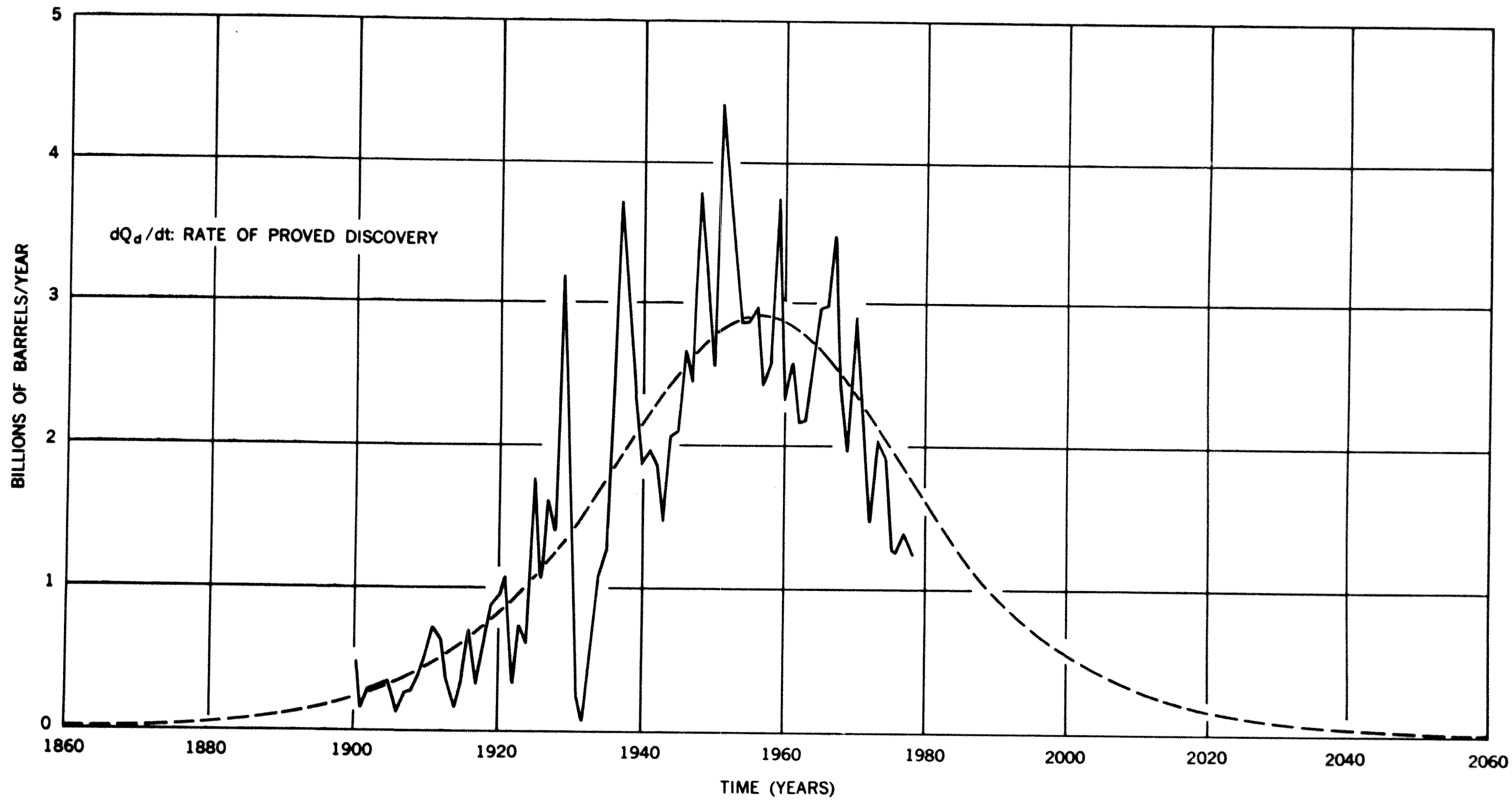

Fig. 29 - Annual proved discoveries of U.S. crude oil superposed upon derivative of logistic equation of 1972.

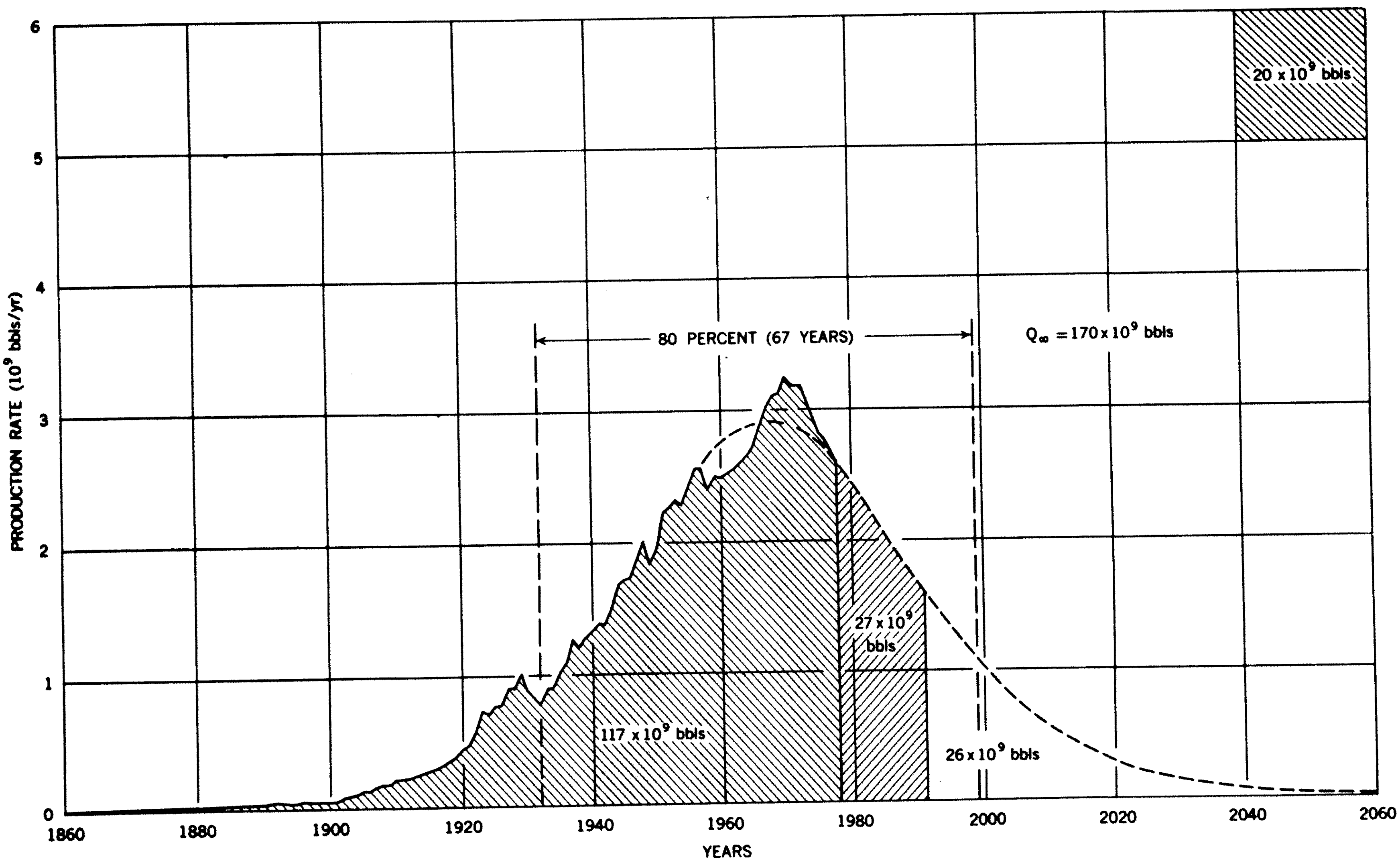

Fig. 30 - Complete cycle of U.S. crude-oil production as estimated in 1980 superposed upon derivative of logistic equation of 1972.

equation (27),

$$(dQ/dt)/Q = a - (a/Q_\infty)Q.$$

Using the method described in equations (43) to (45), the curve of log N_b versus t was constructed at 5-year intervals from 1900 to 1980, using an assumed value of 163.0 billion barrels for Q_b. From this, the following three dates and the corresponding cumulative discoveries were chosen as being points on the smooth almost linear curve:

Dates:	1915	1960	1980
$Q(10^9$ bbl):	8.74	92.61	136.90.

The constants obtained for the logistic equation, as shown in Figure 31, were:

$$t_0 = 1915,$$
$$N_0 = 17.570,$$
$$Q_\infty = 162.3 \times 10^9 \text{ bbl},$$
$$a = 0.0700/\text{year}.$$

A second calculation was made by the same procedure except that the curve of cumulative discoveries, Q_d versus t, was first smoothed by means of an 11-year running average except for the last 5 years. This gave the results:

$$Q_\infty = 161.8 \times 10^9 \text{ bbl},$$
$$a = 0.699/\text{year}.$$

The results of the estimate based upon the linear graph of $(dQ/dt)/Q$ versus Q are shown in Figure 32. In this case, the curve of cumulative discoveries versus time from 1900 to 1980, except for the last 5 years, was smoothed by an 11-year running average, and dQ/dt, at 5-year intervals, was based upon 5-year averages. For the earlier figures, as was expected, there was a wide scattering of the data points, but for the last 25 years from 1955 to 1980, the data points gave a very good straight line. Extrapolation of this linear trend to the Q-axis and to the

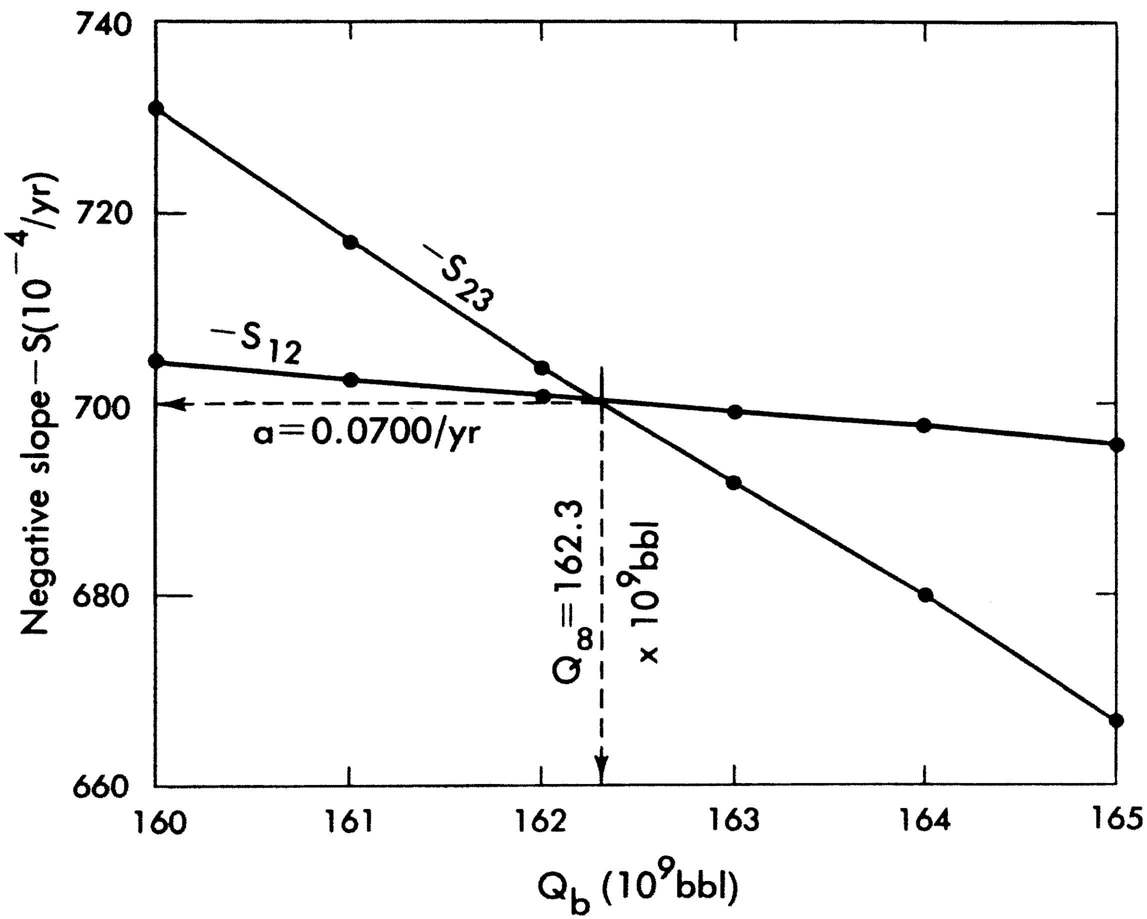

Fig. 31 - Determination of logistic-equation constants for U.S. crude-oil cumulative proved discoveries, 1900-1980.

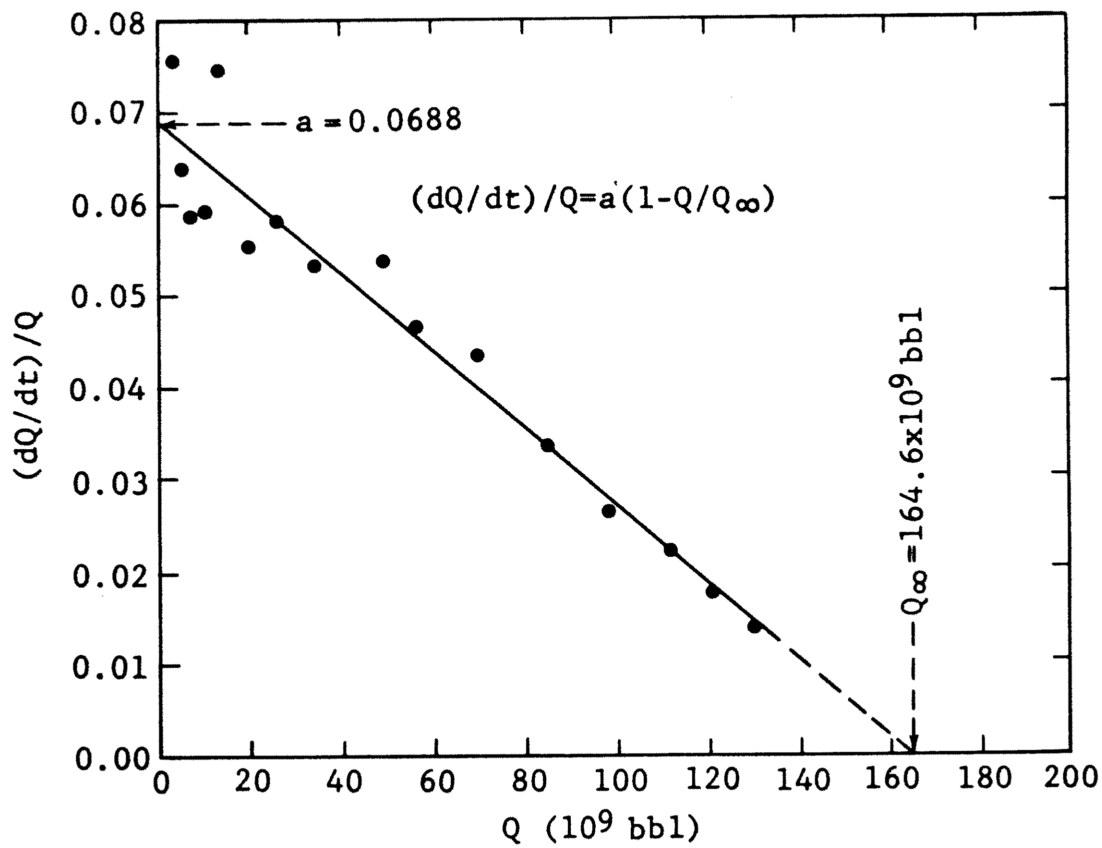

Fig. 32 - Determination of logistic-equation constants for U.S. cumulative proved crude-oil discoveries, 1860-1975, by the technique of $(dQ/dt)/Q$ versus Q.

vertical axis gave the following figures for the logistic constants:

$$Q_\infty = 164.6 \times 10^9 \text{ bbl},$$

$$a = 0.0688/\text{year}.$$

Estimates Based upon Discoveries per Unit Depth of Exploratory Drilling

Development of Theory.— Heretofore we have dealt principally with the variations with time of cumulative proved discoveries and production, proved reserves, and their derivatives with respect to time. One difficulty with variations with respect to time is that such variations are sensitive to economic influences such as fluctuations in prices. A different kind of variation, involving new data not previously used, is represented by discoveries per unit depth of exploratory drilling as a function of cumulative depth of drilling, or of cumulative discoveries. The rate of discovery of oil per unit depth of drilling is determined principally by the geological situation dealt with and by the technology of exploration and production; it is highly insensitive to economic influences.

Let h be the cumulative depth of exploratory drilling in a given region, and Q be the cumulative discoveries. Exploration in the region begins with $h = 0$, and $Q = 0$. Then as h increases without limit, Q tends to a definite finite limit Q_∞. The rate of discovery as a function of h will be dQ/dh.

Because of the indefiniteness of the upper limit of h, but not of Q, it is convenient to consider the variation of dQ/dh as a function of Q, as we have done previously when considering dQ/dt as a function of Q. Thus we consider the variation with Q of dQ/dh within the limits of the complete cycle as Q increases from 0 to Q_∞. During this cycle, when

$$\left.\begin{aligned} Q &= 0,\ dQ/dh > 0; \\ 0 < Q &< Q_\infty,\ dQ/dh > 0; \\ Q &= Q_\infty,\ dQ/dh = 0. \end{aligned}\right\} \tag{50}$$

The variation of dQ/dh with Q, can be expressed by the Maclaurin series,

$$dQ/dh = c_o + c_1Q + c_2Q^2 + \text{---} + c_nQ^n. \tag{51}$$

The lowest degree and the simplest form of this equation that satisfy the conditions of equations (50) is the first degree,

$$dQ/dh = c_o + c_1Q. \tag{52}$$

When

$$Q = Q_\infty,\ dQ/dh = 0,$$

and from equation (52),

$$c_o + c_1Q_\infty = 0,$$

or

$$c_o = -c_1Q_\infty \tag{53}$$

Substituting this into equation (52) then gives

$$dQ/dh = -c_1(Q_\infty - Q),$$

and letting β be substituted for $-c_1$, we obtain

$$dQ/dh = \beta(Q_\infty - Q). \tag{54}$$

Thus dQ/dh varies linearly with respect to Q, as shown in Figure 33, having its intercept on the vertical axis, when $Q = 0$, at βQ_∞; and on the horizontal axis, when $dQ/dh = 0$, at Q_∞. The slope of this line is $-\beta$.

To obtain dQ/dh as a function of h we separate the variables and integrate,

$$\int\frac{dQ}{Q_\infty - Q} = \int\beta dh + c, \tag{55}$$

or

$$\ln(Q_\infty - Q) = -\beta h - c. \tag{56}$$

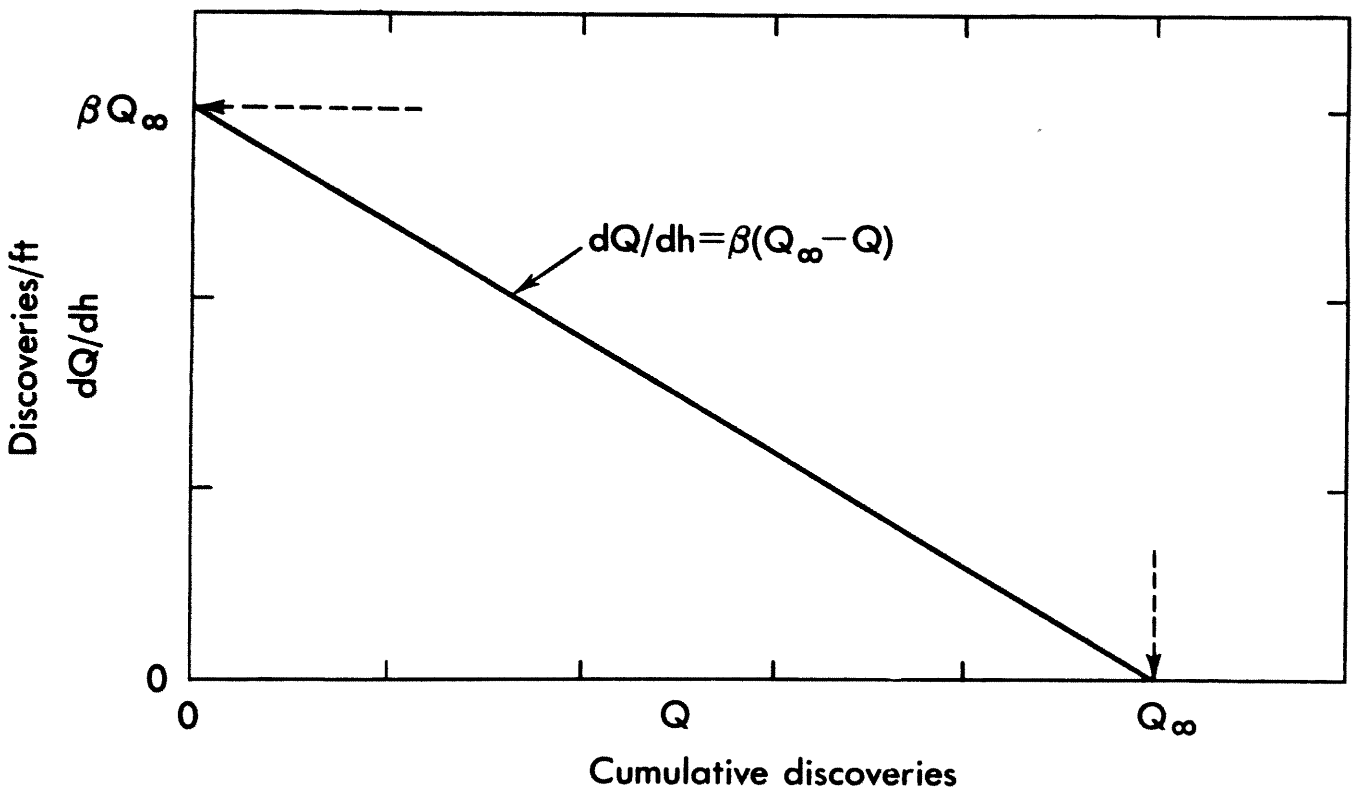

Fig. 33 - Linear relation between oil discoveries per unit depth of exploratory drilling and cumulative discoveries.

Then, when $Q = 0$, $h = 0$, and

$$-c = \ln Q_\infty, \tag{57}$$

Equation (56) then becomes

$$\ln(Q_\infty - Q) = \ln Q_\infty - \beta h \tag{58}$$

or

$$Q = Q_\infty(1 - e^{-\beta h}). \tag{59}$$

Differentiating equation (59) with respect to h then gives

$$dQ/dh = \beta Q_\infty e^{-\beta h}. \tag{60}$$

The manner of variation of both Q and dQ/dh as functions of h are shown graphically in Figure 34.

The foregoing results, as has been shown by Arps and Roberts (1958), Arps, Mortada, and Smith (1971), Menard and Sharman (1975), Root and Drew (1979), and Drew, Schuenemeyer, and Root (1980), are also consistent with the expectations of probability theory. The probability of the discovery of a given amount of oil by a fixed amount of exploratory drilling in a given area is roughly proportional to the undiscovered oil in the region,

$$Q_u = Q_\infty - Q.$$

Consider the idealized case of a region of exploration of total area S which contains fields of uniform oil content ΔQ occurring at the constant depth z. Let A_∞, which is small compared with S, be the total area of such fields, and let A be the area of the fields already discovered by cumulative depth of exploratory drilling h. Then

$$dQ/dh = (\Delta Q/z)[(A_\infty - A)/S]. \tag{61}$$

Then, if

$$\delta = Q/A \tag{62}$$

is the oil per unit area in the fields,

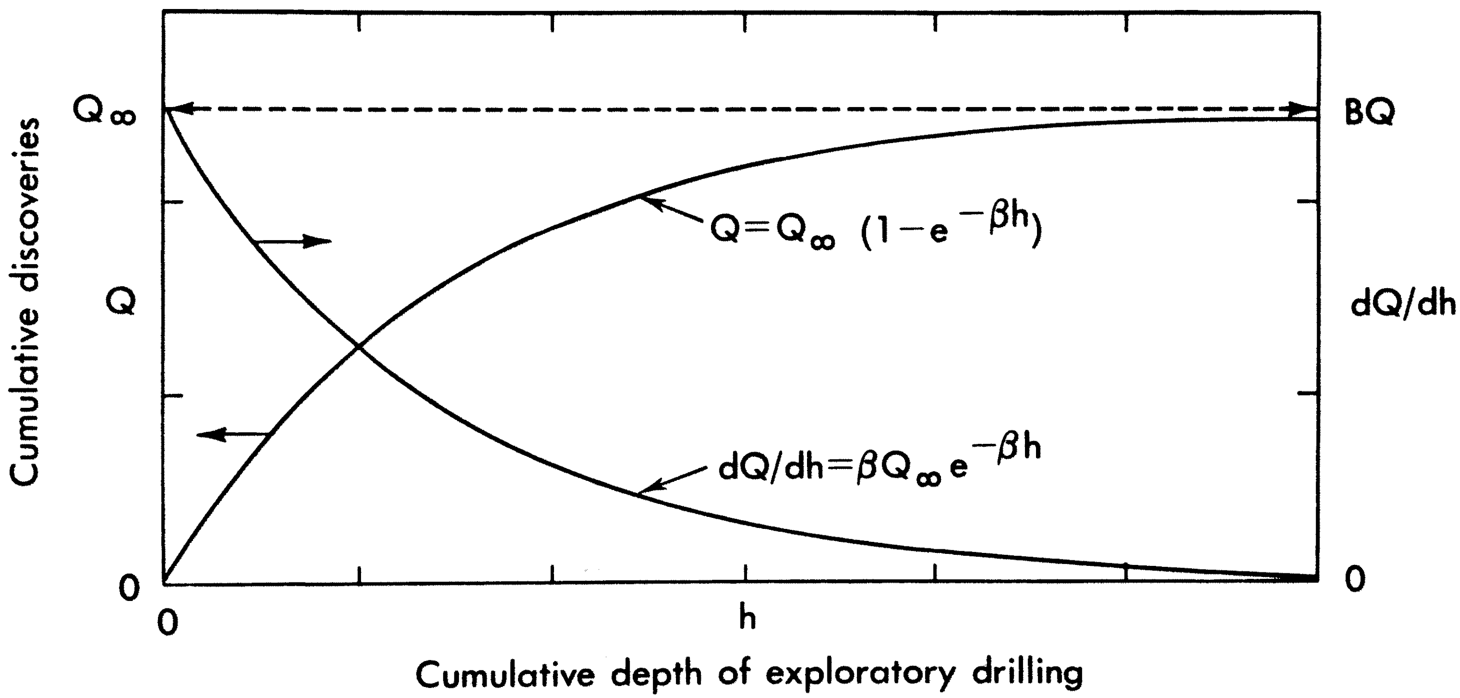

Fig. 34 - Variations of cumulative discoveries and of discoveries per unit depth of exploratory drilling as functions of cumulative exploratory drilling.

$$(A_\infty - A) = (Q_\infty - Q)/\delta, \tag{63}$$

which, when substituted into equation (61), gives

$$dQ/dh = (\Delta Q/zS\delta)(Q_\infty - Q). \tag{64}$$

This is of the same form as equation (54), where $\Delta Q/zS\delta$ in equation (64) corresponds to β in equation (54). Integration of equation (64) then gives for Q versus h an equation of the form of equation (59), and differentiation of that with respect to h gives an exponential decline of dQ/dh versus h of the same form as equation (60).

In an actual oil-bearing region these simplified conditions do not occur. There are commonly a small number of large fields which frequently contain most of the oil, and a large number of small fields. Also the depths of the fields range from a few hundred feet to as much as 20,000 feet. In addition, in a region such as the entire United States, the technologies of discovery and of production undergo progressive improvement during the entire cycle of oil exploitation. This favors the discovery of the larger and shallower fields during the earlier stages of the cycle with a rapid rate of decline of dQ/dh versus h, followed by a slower rate of decline as the sizes of the remaining fields decrease and their depths increase. Offsetting this somewhat is the steady improvement of the techniques of exploration and production which tends to increase the discoveries per foot with respect to the rates expectable by random drilling. The net effect, however, is still roughly an exponential decline of dQ/dh versus h.

In order to correlate the oil discoveries made in a given region with the corresponding exploratory drilling, a different definition of the term "discoveries" from that of "proved discoveries" used previously is required. In this case all of the oil that the fields discovered in a given year will ultimately produce must be credited to the exploratory drilling done during that year.

This involves estimates of "ultimate recovery" of oil from fields discovered in a given year t_i as estimated at a later year t_j. The ultimate recovery from fields discovered during the year t_i as estimated at the later year t_j is defined as the sum of cumulative production from those fields to the year of the estimate plus their proved reserves at the time of the estimate. At the end of the year of discovery the only oil credited to the discovery year is the item "New Field Discoveries," in the American Petroleum Institute annual report on proved reserves. This ordinarily is only a small quantity. During succeeding years, cumulative production from those fields steadily increases, and the sum of cumulative production plus proved reserves gradually approaches asymptotically the quantity δQ_∞, which is the true ultimate amount of oil those fields will produce.

The first study of this kind for the U.S. crude-oil production was that made during World War II by the Petroleum Administration for War (PAW) (Frey and Ide, 1946, Appendix 12, Table 10, p. 442; C.L. Moore, 1962, Table IV, p. 94). In effect, what was done in the PAW study was to combine the cumulative production to January 1, 1945, with the American Petroleum Institute estimate of proved reserves for the same date for all the fields discovered during each successive year from 1860 to 1944. This gave an estimate as of January 1, 1945, of the proved oil discoveries assignable to fields by their year of discovery. Two more such studies were made subsequently by the National Petroleum Council (1961; 1965). The first brought the PAW study up to the date of December 31, 1959, and the second to January 1, 1964. Each of the latter studies gave a lumped increase for all the discoveries made from 1860 to 1920, and then gave separate estimates for the fields discovered during each year from 1920 to the terminal date of the study.

In 1966, the name of the API "Committee on Petroleum Reserves" was changed to "Committee on Reserves and Productive Capacity" and the scope of its activities was expanded. One new item in the committee's annual reports was a table for the year of the report of estimations of ultimate recovery for fields discovered during the pre-1920 period and for individual years from 1920 to the year of the report. These reports are given from 1966 to 1979 in the annual publication, *Reserves of Crude Oil, Natural Gas Liquids, and Natural Gas in the United States and Canada as of December 31,* [given year], issued jointly by the American Petroleum Institute, the American Gas Association, and the Canadian Petroleum Association, for brevity, the API, AGA, CPA "Blue Books."

Using these data, the problem is to estimate the ultimate amount of oil that fields discovered during successive years will eventually produce. Two principal alternative procedures have been developed for this purpose. One, that developed by the present author (Hubbert, 1967; 1974), consists in following the fields discovered in a given year or group of successive years, and plotting their growth with increasing time following their year of discovery. If we let $(\delta Q)_1$ be the initial estimate of the new oil discovered at the end of the year of discovery, and $(\delta Q)_\tau$ be the estimate for the same fields τ years later, then we can plot a curve of

$$y_\tau = (\delta Q)_\tau / (\delta Q)_1 \tag{65}$$

as a function of τ.

Expressed in this manner, y_τ is dimensionless, and is independent of the absolute magnitude of the oil discoveries in any given year, and the time-delay τ is common to the discoveries of all years. Hence the data for all discovery years can be expressed in terms of y versus τ. This will be a curve that rises steeply initially and finally approaches the limit,

$$y \to y_\infty,$$

as τ increases

This was the procedure used in 1967 (Hubbert, 1967, Figs. 4 and 5) using the limited data then available, and again in 1972 (Hubbert, 1974, Figs. 45 and 46) with much more detailed information. In both instances, however, substantially the same growth curves were obtained, which were fitted by empirical equations of the form,

$$y_\tau = y_\infty[1 - e^{-\gamma(\tau+c)}]. \tag{66}$$

The data as of 1972 are shown graphically in Figure 35, which is reproduced from my report of 1974 (Hubbert, 1974, Fig. 45). From these the constants for equation (66) were:

$$y_\infty = 5.8,$$

$$\gamma = 0.076 \text{ per year},$$

$$c = 1.503 \text{ years}.$$

By means of this equation it is possible to estimate how much more the oil discovered in a given year will increase when its magnitude $(\delta Q)_\tau$ is given after τ years of production and development.

This is done by a correction factor α, defined by

$$\alpha = y_\infty/y_\tau, \tag{67}$$

which from equation (66) is

$$\alpha = 1/[1 - e^{-\gamma(\tau+c)}]. \tag{68}$$

Then, since

$$y_\tau = (\delta Q)_\tau/(\delta Q)_1,$$

and

$$y_\infty = (\delta Q)_\infty/(\delta Q)_1,$$

$$y_\infty/y_\tau = (\delta Q)_\infty/(\delta Q)_\tau.$$

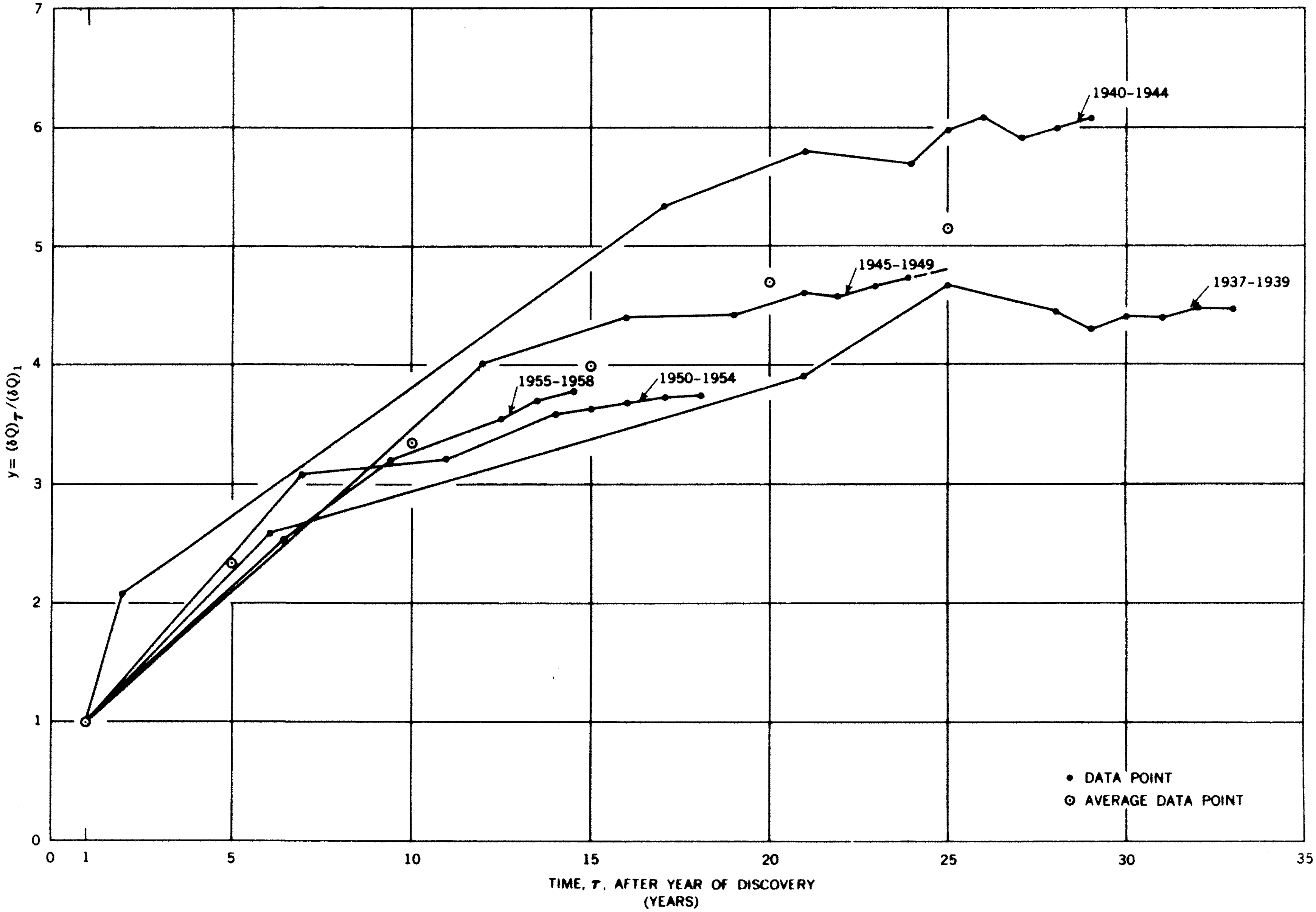

Fig. 35 - Growth in the estimates of the ultimate recovery of crude oil from fields discovered in given years as re-evaluated at successively later years (Hubbert, 1974, Fig. 45).

Hence

$$(\delta Q)_\infty = \alpha(\delta Q)_\tau . \quad (69)$$

This first procedure is based upon the growth during successive years, t_j, of the estimates of oil discovered in a given year, t_i. An alternative procedure is that described originally by Arrington (1960; 1966) and more recently by Marsh (1971). In this case, use is made of the API estimates for the two successive most recent years, t_j and t_{j+1}, for the oil discovered in previous years, t_i. In this case the fields considered are no longer the *same* fields, but instead are the different fields discovered during successive earlier years with an increasing value of the time-delay τ.

In this manner a ratio,

$$r_\tau = (\delta Q)_{\tau+1}/(\delta Q)_\tau , \quad (70)$$

is obtained as a function of τ for the oil discovered during each preceding earlier year. From these successive ratios,

$$r_1, r_2, r_3 \ldots r_\tau ,$$

the growth factor y_τ is obtained by the product

$$y_\tau = (r_1 r_2 r_3 \ldots r_\tau), \quad (71)$$

which tends to y_∞ as τ increases without limit.

As in the earlier procedure, the magnitude of $(\delta Q)_\infty$ for the discoveries made during any given year is given by

$$(\delta Q)_\infty = \alpha(\delta Q)_\tau ,$$

where

$$\alpha = y_\infty / y_\tau .$$

In practice, instead of dealing with the discoveries made during single years, the sum of the discoveries made during a sequence of 5 or so years may be used. In that case, for a common value of τ, the ratios must also be taken during a corresponding sequence of successive pairs of years.

Three Successive Studies of U.S. Crude-Oil Discoveries per Foot of Exploratory Drilling

The results of three successive studies of the U.S. crude-oil discoveries in the Lower-48 states and adjacent continental shelves, as a function of cumulative depth of exploratory drilling, are shown in Figures 36, 37, and 38. The first of these (Hubbert, 1967, Fig. 15) shows the discoveries made by the first 1.5×10^9 ft of drilling, which encompassed the period 1860 to 1966.7. The second (Hubbert, 1974, Fig. 49) gives the results obtained by 1.7×10^9 ft of exploratory drilling during the period from 1860 to just short of 1972. The third, by David H. Root of the U.S. Geological Survey (Root, 1980), gives the discoveries made by the 2.0×10^9 ft of exploratory drilling during the period 1860 to 1977.9. In the first two studies, the method used in estimating the amount of oil ultimately recoverable from fields already discovered was that of the present author, as described earlier. In the third study, Root used his own modification of the method of Arrington and Marsh, which was applied to the API, AGA, CPA "Blue Book" data from 1966 to 1978. For each successive study more and better data were available than for the one preceding.

For the graphical presentation of the data, a convenient unit for Δh is 10^8 ft. Hence, in the three figures, the cumulative drilling amounted to 15, 17, and 20 units respectively. For each unit of drilling, a vertical column is shown representing the quantity of oil discovered by that unit. The lower part of the column represents the proved cumulative discoveries as estimated at the date of the study. The shaded area at the top of each column represents the additional oil those fields are expected to produce as determined by the α-correction. The total amount of recoverable oil discovered by the first 15 units of drilling in the 1967 study was estimated to be 136.04 billion barrels, of which approximately 25 billion barrels were accounted for by the α-correction. In the

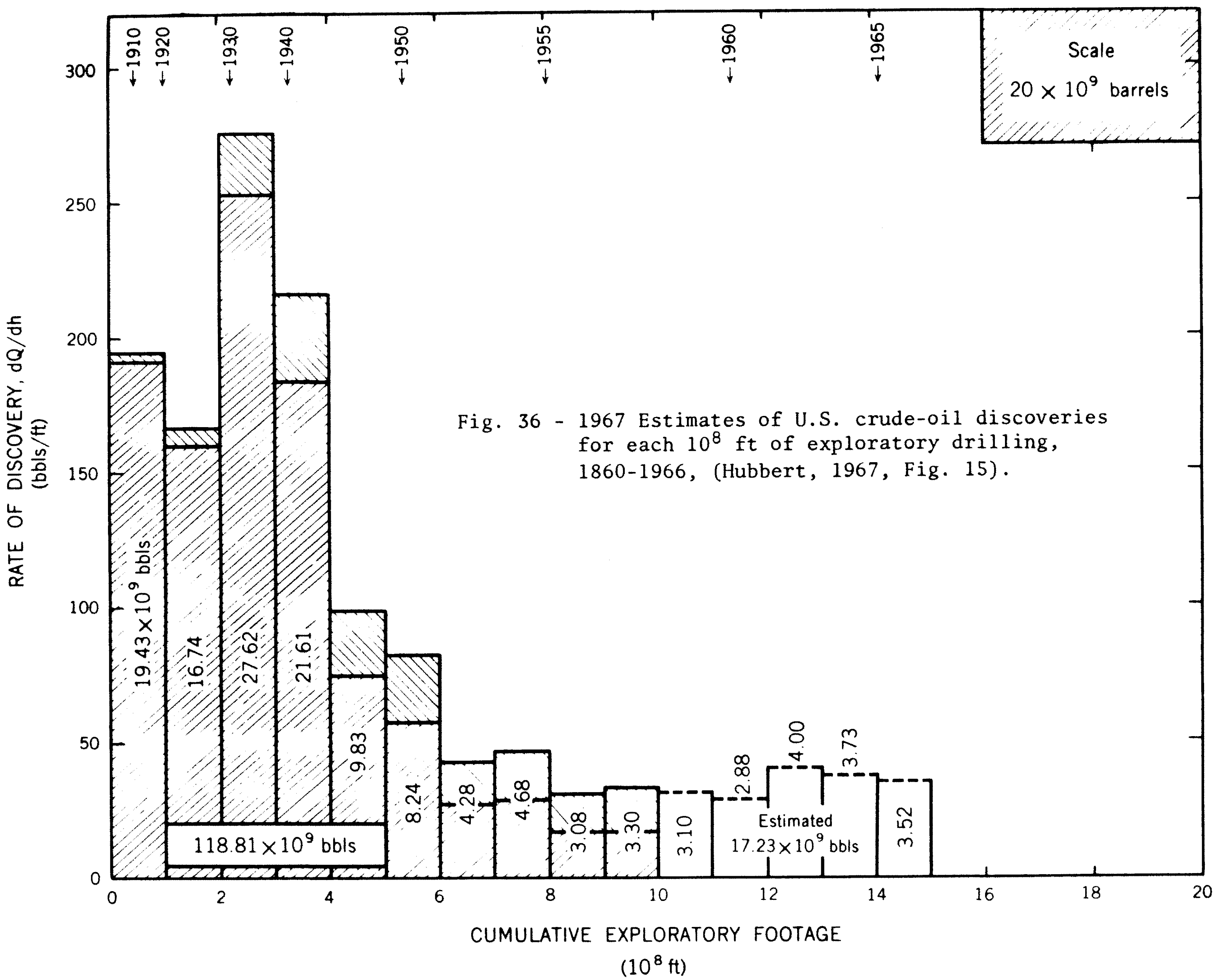

Fig. 36 - 1967 Estimates of U.S. crude-oil discoveries for each 10^8 ft of exploratory drilling, 1860-1966, (Hubbert, 1967, Fig. 15).

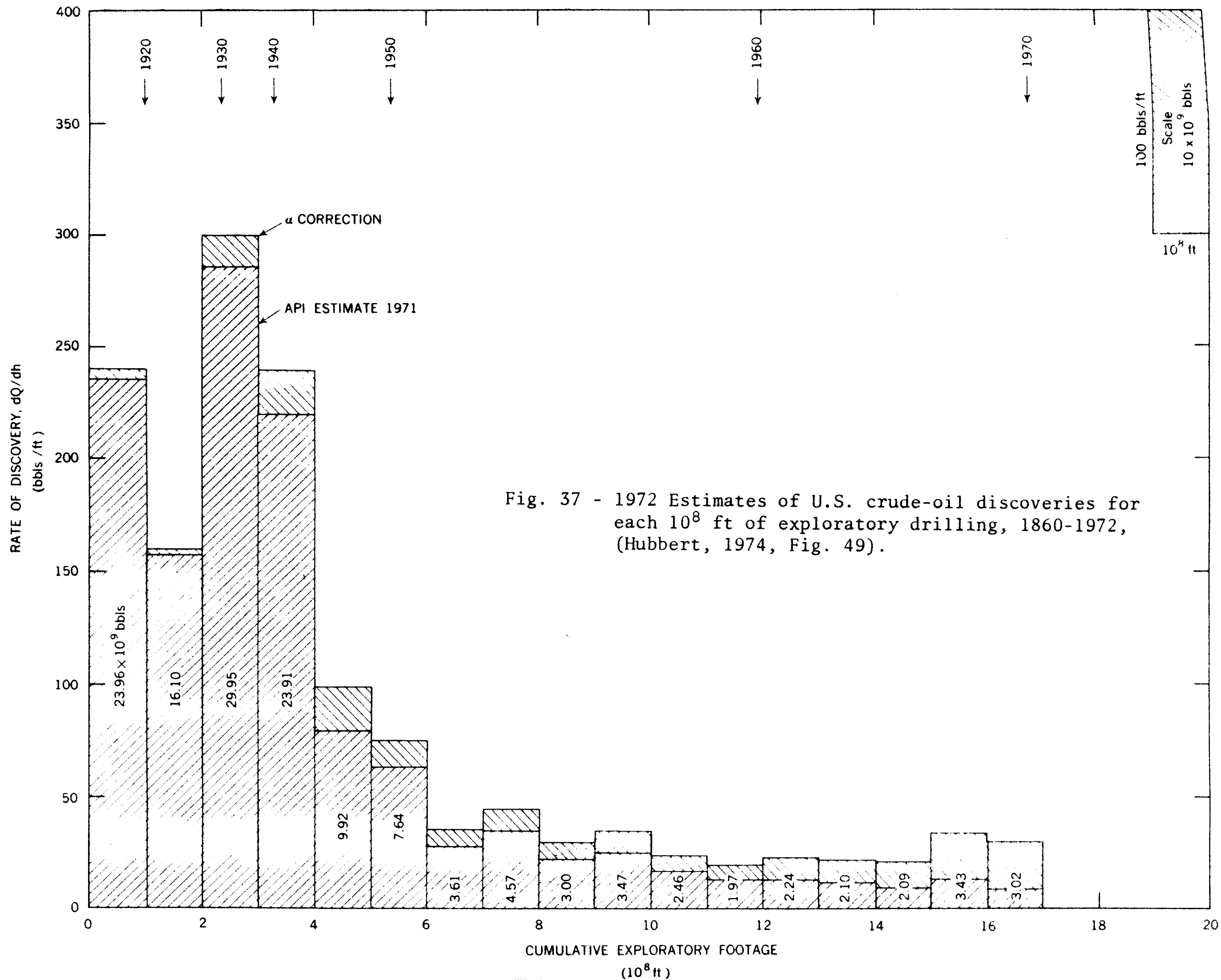

Fig. 37 - 1972 Estimates of U.S. crude-oil discoveries for each 10^8 ft of exploratory drilling, 1860-1972, (Hubbert, 1974, Fig. 49).

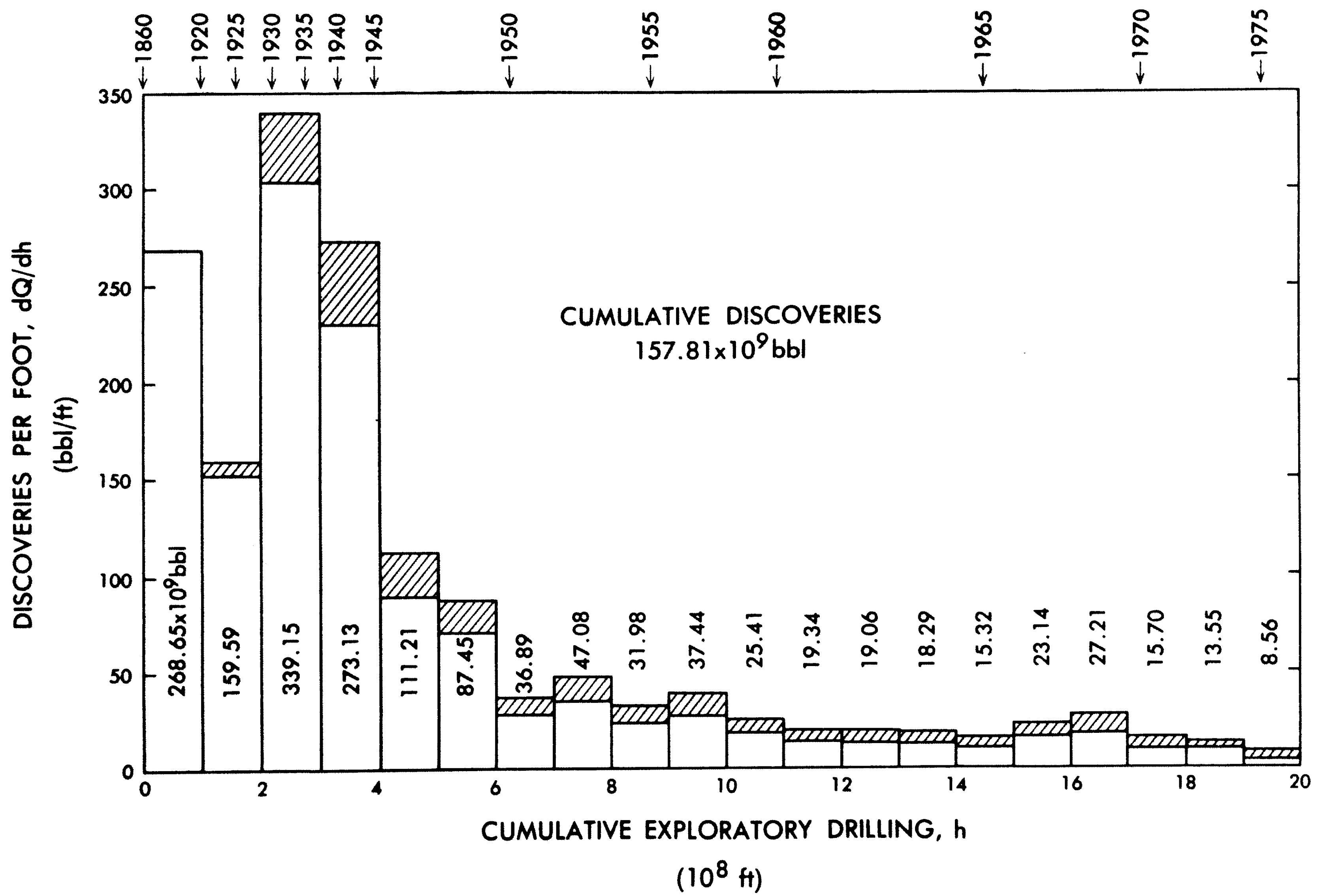

Fig. 38 - 1980 Estimates by David H. Root of U.S. crude-oil discoveries for each 10^8 ft of exploratory drilling, 1860-1977.9, (Root, 1980).

1974 study, 143.44 billion barrels of recoverable oil were estimated to have been discovered by the first 17 units of drilling. Of this, approximately 19 billion barrels were contributed by the α-correction. The total recoverable oil estimated in the third study as having been discovered by the first 20 units of drilling amounted to 157.87 billion barrels, of which 22.92 were due to the α-correction.

Because the results shown in Figures 36, 37, and 38 are all derived from different suites of data, and, in the case of the last study, by a different method of analysis, it is not to be expected that the results obtained would be in close agreement. Nevertheless all of the studies give results of strong similarity. All show high rates of discovery, averaging between 20 and 30 billion barrels per 10^8-ft unit of drilling (200 to 300 bbl/ft) for the first 4 units, followed by a precipitous decline. In Figure 36, the discovery rate had declined to only 3.52 (10^9 bbl/10^8 ft), or 35.2 bbl/ft, for the last or 15th unit; in Figure 37 this decline had reached 30.2 bbl/ft for the 17th unit; and in Figure 38, by the 20th unit, it had declined still further to but 8.56 bbl/ft.

All of these figures show a roughly negative-exponential decline in dQ/dh versus h during the entire cycle. In order to estimate the future, the negative-exponential curve best fitting the data needs to be determined in each instance. To simplify notation, let dQ/dh be represented by R.

Then

$$R = R_0 e^{-\beta h} \tag{72}$$

will be the desired equation of which the two parameters R_0 and β are to be determined from the data. The criterion for best fit will be

$$\int_0^{h_n} R dh = Q_n, \tag{73}$$

where Q_n represents the total discoveries made by the h_n units of drilling, with the curve passing through the last point (R_n, h_n) on the graph. In other words, we wish to determine the negative-exponential curve that equalizes the excesses and defects of the data, and passes through the last point.

Substituting the value of R from equation (72) into equation (73), we obtain

$$Q_n = R_o \int_0^{h_n} e^{-\beta h} dh$$

$$= (R_o - R_n)/\beta,$$

from which

$$\beta = (R_o - R_n)/Q_n. \tag{74}$$

Also, taking the logarithm of equation (72), with $R = R_n$ and $h = h_n$, we obtain

$$\ln\ (R_o/R_n) = \beta h_n$$

or

$$\beta = \ln\ (R_o/R_n)/h_n. \tag{75}$$

Dividing equation (75) by (74) then gives

$$\frac{\ln\ (R_o/R_n)}{R_o - R_n} \cdot \frac{Q_n}{h_n} = 1, \tag{76}$$

of which the only unknown is R_o. This can be solved for R_o by an iteration procedure of substituting for R_o an assumed value R_b whereby the left-hand term of equation (76) becomes $f(R_b)$. When $f(R_b)$ has the value of 1, $R = R_o$. The decline parameter β is then obtained from either of equations (74) or (75).

Once R_o and β are known, the estimate of the ultimate cumulative discoveries, Q_∞, is given by

$$Q_\infty = R_o/\beta, \tag{77}$$

and that of the undiscovered oil, Q_u, by

$$Q_u = R_n/\beta. \tag{78}$$

The negative-exponential curves obtained in this manner from the data of Figures 36, 37, and 38 are shown in Figures 39, 40, and 41. For the study of 1967 (Fig. 36), the values of the parameters are:

$$R_0 = 18.63\ (10^9\ \text{bbl}/10^8\ \text{ft}),$$
$$= 186.3\ \text{bbl/ft},$$
$$\beta = 0.1111\ \text{per}\ 10^8\ \text{ft}.$$

From these parameters and the numerical data of Figure 36,

$$R_{15} = 3.52\ (10^9\ \text{bbl}/10^8\ \text{ft}),$$
$$Q_{15} = 136.04 \times 10^9\ \text{bbl},$$
$$Q_u = 31.7 \times 10^9\ \text{bbl},$$

and

$$Q_\infty = 167.7 \times 10^9\ \text{bbl}.$$

This value of 168 billion barrels for Q_∞ obtained from the 1967 study of discoveries per unit depth of exploratory drilling as a function of cumulative drilling, although based upon different data and a totally different method of analysis, is in very close agreement with the figure of 170 billion barrels obtained in the studies of cumulative production, proved reserves, and cumulative proved discoveries made in 1962 and 1974.

For the negative-exponential curve for the 1974 study of dQ/dh versus h shown in Figure 40, the values of the two parameters are:

$$R_0 = 18.154\ (10^9\ \text{bbl}/10^8\ \text{ft}),$$
$$= 181.54\ \text{bbl/ft},$$
$$\beta = 0.1055\ \text{per}\ 10^8\ \text{ft},$$

and

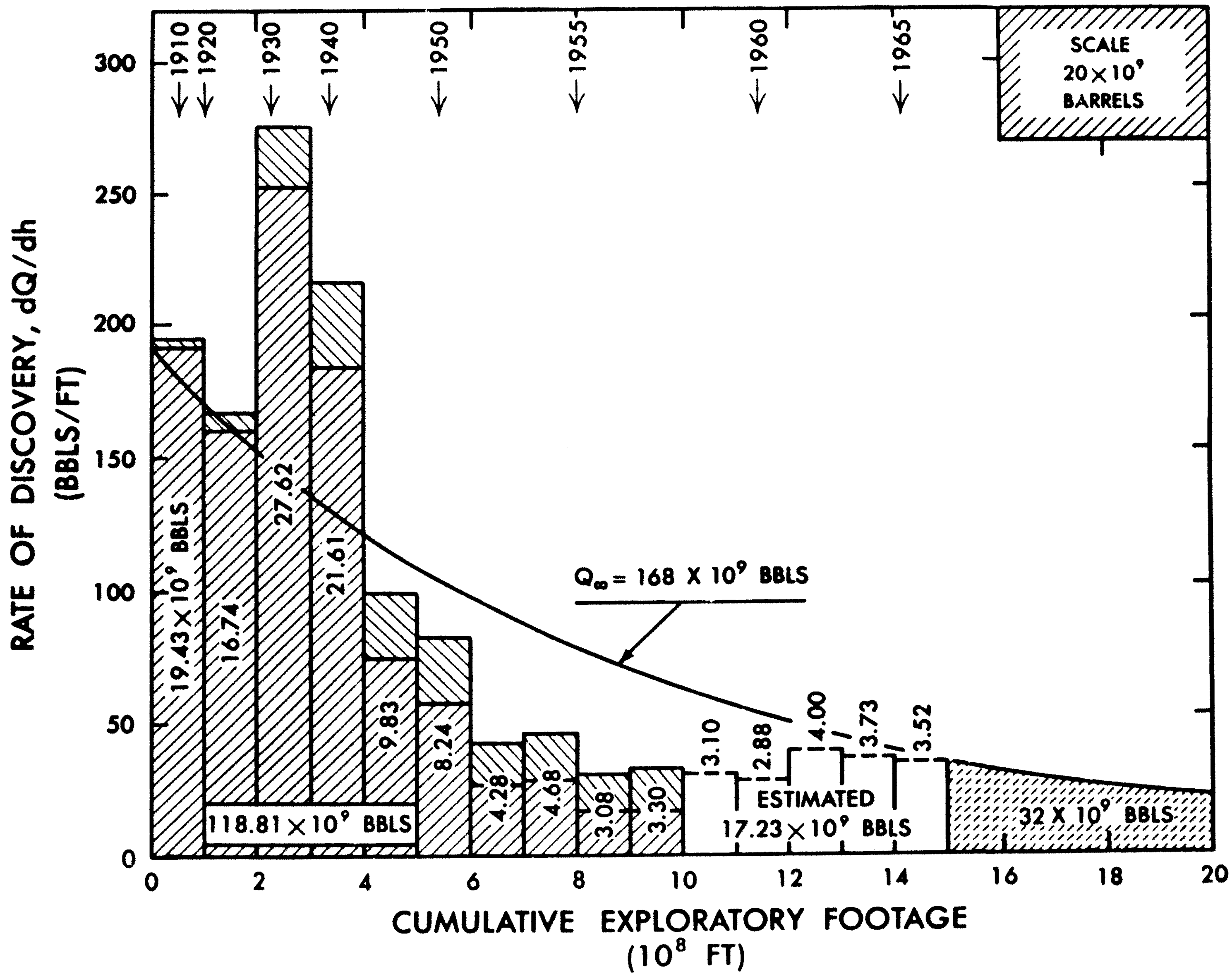

Fig. 39 - Exponential-decline curve superposed upon Fig. 36 giving 1967 estimates of future and ultimate crude-oil discoveries for the U.S. Lower-48 states.

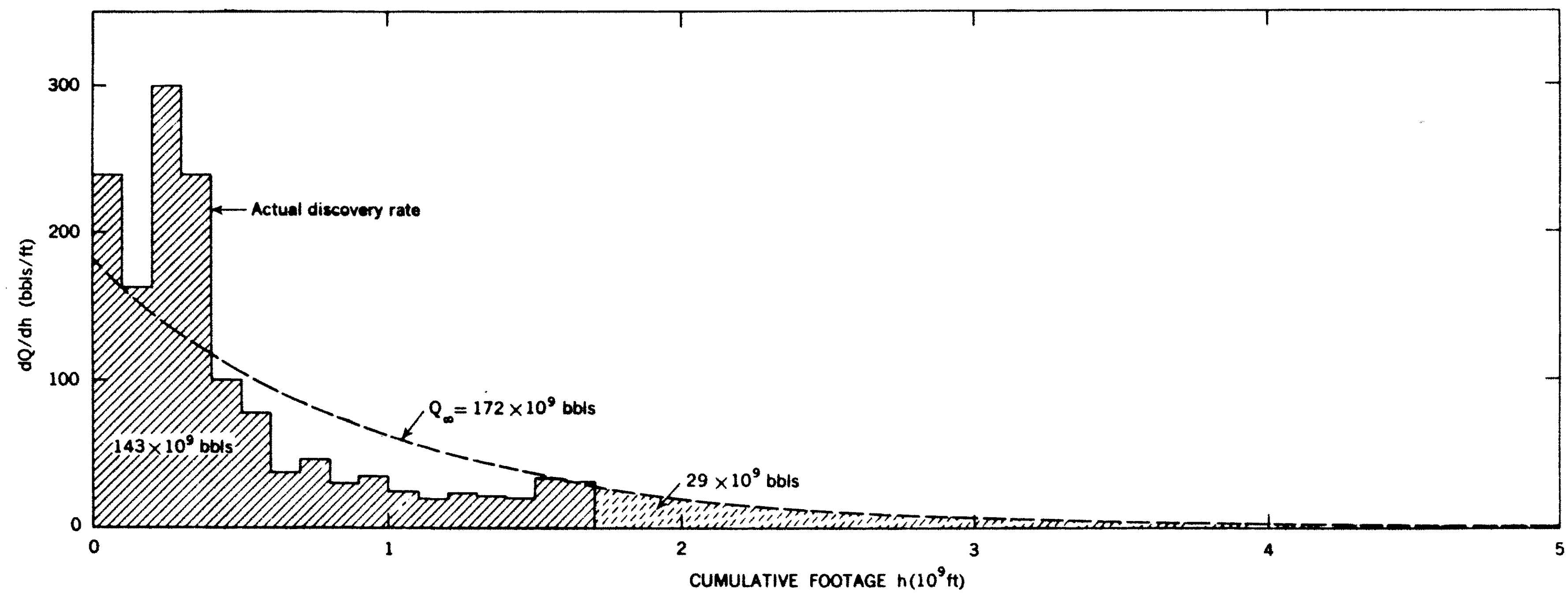

Fig. 40 - Exponential-decline curve superposed upon Fig. 37 giving 1972 estimates of future and ultimate crude-oil discoveries for the U.S. Lower-48 states (Hubbert, 1974, Fig. 50).

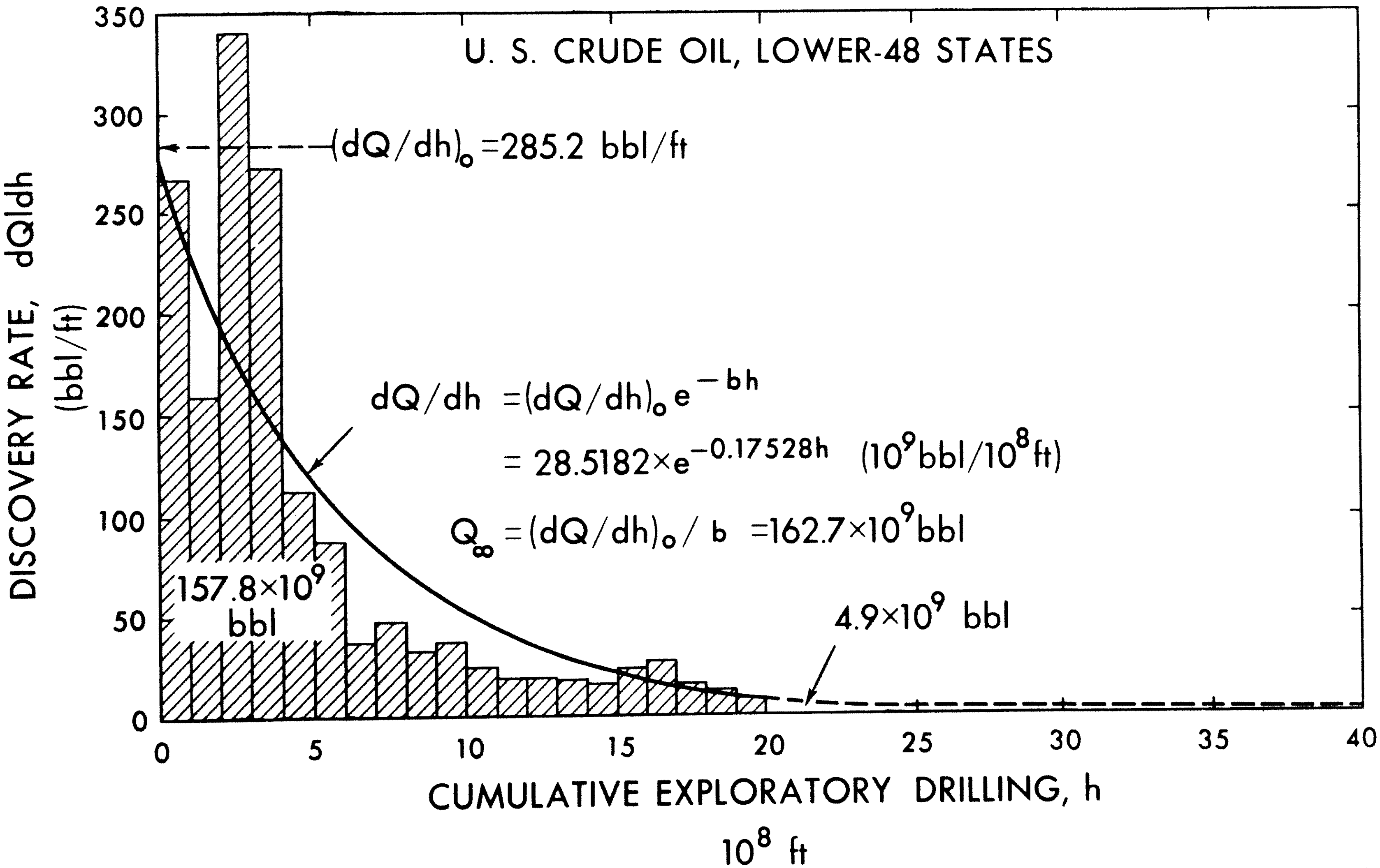

Fig. 41 - Exponential-decline curve superposed upon Fig. 38 giving 1980 estimates of future and ultimate crude-oil discoveries for the U.S. Lower-48 states.

$R_{17} = 3.02\ (10^9\ \text{bbl}/10^8\ \text{ft}),$

$Q_{17} = 143.44 \times 10^9\ \text{bbl},$

$Q_u = 28.62 \times 10^9\ \text{bbl},$

$Q_\infty = 172.06 \times 10^9\ \text{bbl}.$

Again, this is in very close agreement with the estimate for Q_∞ as of 1972 of 170×10^9 bbl obtained by the analysis of cumulative production, proved reserves, and cumulative proved discoveries.

The negative-exponential curve corresponding to the data of Figure 38 is shown in Figure 41. This is of especial interest because data extending to 1978 are included, and a still different method of analysis was used. The significant data of Figure 38 are:

$R_o = 28.518\ (10^9\ \text{bbl}/10^8\ \text{ft}),$

$= 285.18\ \text{bbl/ft},$

$\beta = 0.1753\ \text{per}\ 10^8\ \text{ft},$

$R_{20} = 0.8564\ (10^9\ \text{bbl}/10^8\ \text{ft}),$

$Q_{20} = 157.81 \times 10^9\ \text{bbl},$

$Q_u = 4.89 \times 10^9\ \text{bbl},$

$Q_\infty = 162.70 \times 10^9\ \text{bbl}.$

Again, this value of 162.7 billion barrels for Q_∞, as determined from data extending to 1978, is in very close agreement with that of 162.3 billion barrels shown in Figure 31, obtained from the logistic constants of the curve of cumulative proved discoveries to 1978. It also differs by only 1.9 billion barrels from the figure of 164.6 shown in Figure 32, based upon the linear-decline curve of $(dQ/dt)/Q$ versus Q. The average of these three figures is 163.2 billion barrels, with a range of uncertainty of only about plus or minus two billion barrels.

Estimation of Natural Gas

Estimations of the ultimate amount of natural gas to be produced in the Lower-48 states, and of the future rates of production, are more difficult than the corresponding estimates for crude oil. This is because the statistics of natural gas are less complete than those for oil until after World War II. Prior to that time a large amount of gas was burned (or "flared") in the fields as gas was produced as a by-product of oil and in excess of the pipeline capacity for collection and distribution.

Since World War II, this situation has greatly improved. Large pipelines were constructed for the transmission of gas from the producing areas to the industrial regions of the northeast and north-central United States and to the Pacific coast. Also, in the mid-1940s, the American Gas Association established its Committee on Natural Gas Reserves. From 1946 to 1979, this committee has issued annual reports on natural-gas proved reserves and production, in parallel with those of the corresponding committee on crude oil of the American Petroleum Institute.

Another serious difficulty is that the record of cumulative production, proved reserves, and cumulative discoveries of natural gas is much more irregular than the corresponding record for oil, which makes mathematical analysis of the data more difficult and of a lower level of reliability. Nevertheless, enough information exists to permit reasonably good estimates to be made of the approximate cumulative production and of future production rates.

Earlier Estimates

Estimate of 1956. —— In my 1956 paper, "Nuclear Energy and the Fossil Fuels" (Hubbert, 1956), the same technique was used for estimating the complete cycle of U.S. natural-gas production as was used for crude oil. The best current estimate for the ultimate amount of natural gas to be produced in the Lower-48 states and adjacent offshore areas was about 850 trillion cubic feet. Cumulative production by the end of 1955 amounted to about 150 and proved reserves to 224 trillion cubic feet. This gave the figure of 374 trillion cubic feet for cumulative proved discoveries, leaving 476 trillion cubic feet for future discoveries.

This is shown graphically in Figure 42, which is reproduced from Figure 22 of the 1956 paper. From this it was estimated that the maximum rate of gas production of about 14 trillion cubic feet per year would occur about 1970. As in the case of crude-oil estimates, the published estimates for the ultimate amount of natural gas to be produced began to escalate immediately after 1956 and, by 1961, the highest estimate had reached 2,630 trillion cubic feet, a figure three times that of 1956.

National Academy Report of 1962. —— In view of the lack of agreement as to the approximate magnitude for Q_∞, it became necessary in the National Academy of Sciences report of 1962 to devise a new method of estimation. The statistical data on cumulative production, proved reserves, and cumulative discoveries, which had been available only since 1945, were insufficient for an estimate of Q_∞. To obtain this figure, the parallel study for crude oil was used in conjunction with the ratio of the discoveries of natural gas to those of crude oil during a given period of time. Thus,

$$Q_\infty \text{ gas} = Q_d \text{ gas} + G[(Q_\infty - Q_d)\text{oil}], \qquad (79)$$

where G is the gas/oil-ratio.

At the end of 1961, cumulative proved discoveries of natural gas amounted to 474 trillion cubic feet. Cumulative proved discoveries for crude oil were 99.1 billion barrels, and the estimate for Q_∞ for crude oil was taken as 175 billion barrels, leaving 75.9 billion barrels for the undiscovered crude oil.

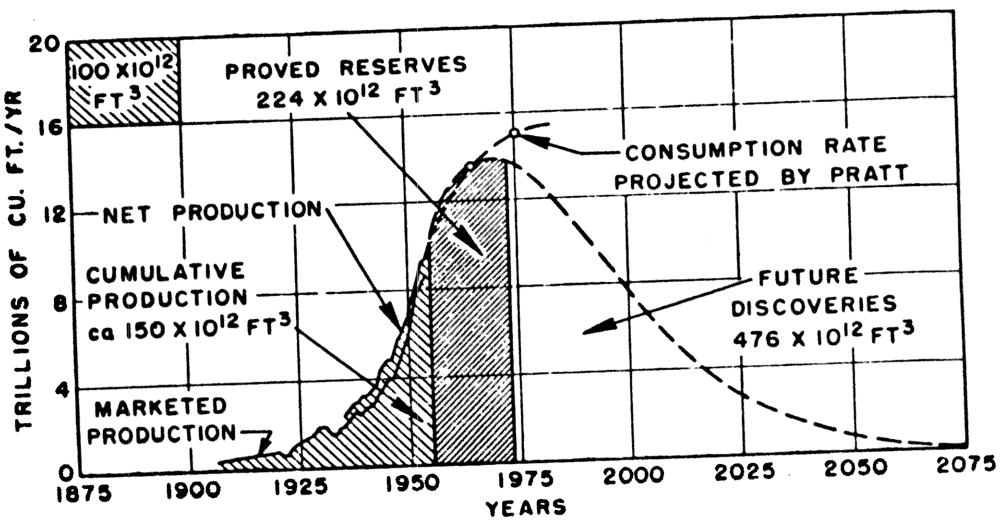

Fig. 42 - Complete cycle of natural-gas production in the U.S. Lower-48 states as estimated in 1956 (Hubbert, 1956, Fig. 22).

For the gas/oil-ratio, two figures were used. The ratio of gas discoveries to crude-oil discoveries during the most recent 20-year period, 1941-1961, was 6,250 ft^3/bbl. However, the possibility was also considered that in response to deeper drilling this ratio might increase in the future to as much as 7,500 ft^3/bbl.

Substituting these figures into equation (79) gave, for the ultimate amount of natural gas to be produced in the Lower-48 states, a low figure of 958 and a high figure of 1,053 trillion cubic feet, or roundly 1,000 trillion cubic feet. Using this figure for Q_∞ in conjunction with the limited data for Q_p, Q_r, and Q_d for natural gas, gave the logistic curves of Figure 43 and their derivatives in Figure 44. The two complete gas-production cycles, based on both low and high estimates of 958 and 1,053 trillion cubic feet, are shown in Figure 45. From these figures, the time delay Δt between cumulative discoveries and cumulative production was estimated to be 16 years. The maximum rate of discovery was estimated to occur at about 1961, the peak in proved reserves in 1969, and the maximum production rate of about 18 to 20 trillion cubic feet per year at about 1977.

Estimate of 1972. —— By 1972 (Hubbert, 1974), despite the fact that 10 more years of data were available, the natural-gas data on cumulative production, proved reserves, and cumulative proved discoveries were still so irregular as to make the use of the logistic equation of doubtful validity. However, by 1972, the proved reserves of natural gas had already reached their maximum in 1967, two years earlier than predicted in 1962. After 1967 they declined steeply. For estimates of Q_∞, two methods were used, that of the gas/oil-ratio in conjunction with the oil estimate, and the gas discoveries as a function of cumulative exploratory drilling. The first method gave an estimate of about 1,000 trillion cubic feet, and the second a higher figure of 1,103 trillion cubic feet.

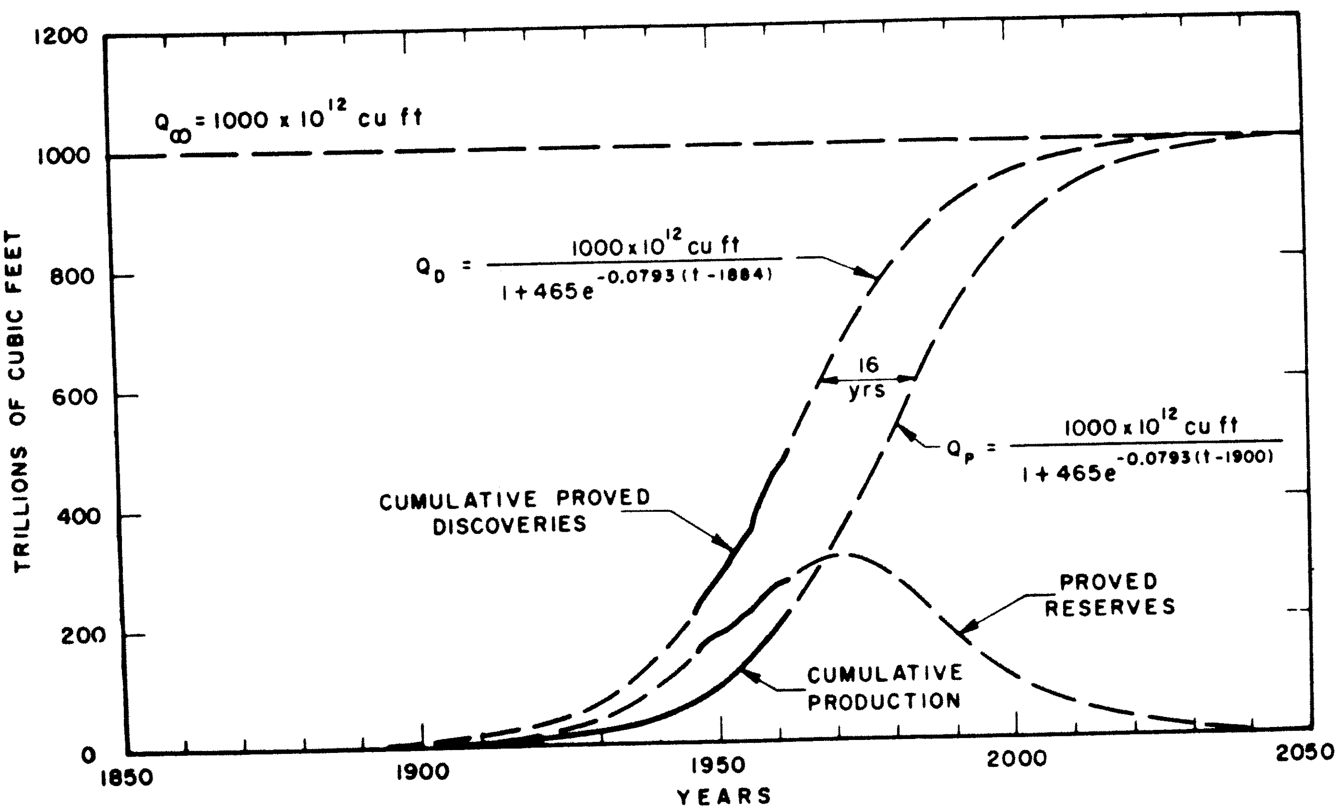

Fig. 43 - Logistic equations and curves for U.S. natural gas in 1962 (Hubbert, 1962, Fig. 45).

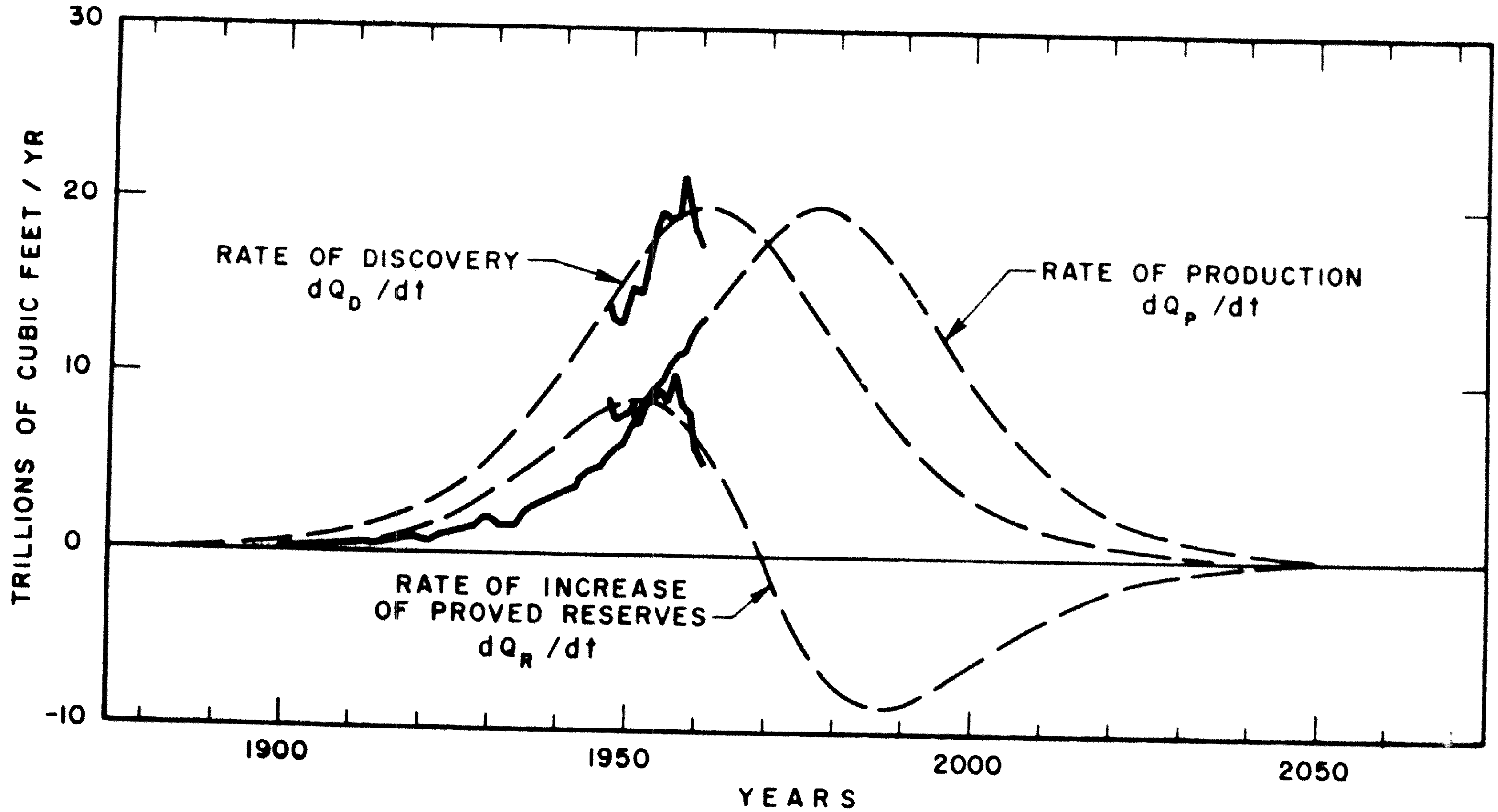

Fig. 44 - Rates of production, discovery, and increase of proved reserves of U.S. natural gas, 1900-1962, superposed upon derivative curves of logistic equations (Hubbert, 1962, Fig. 46).

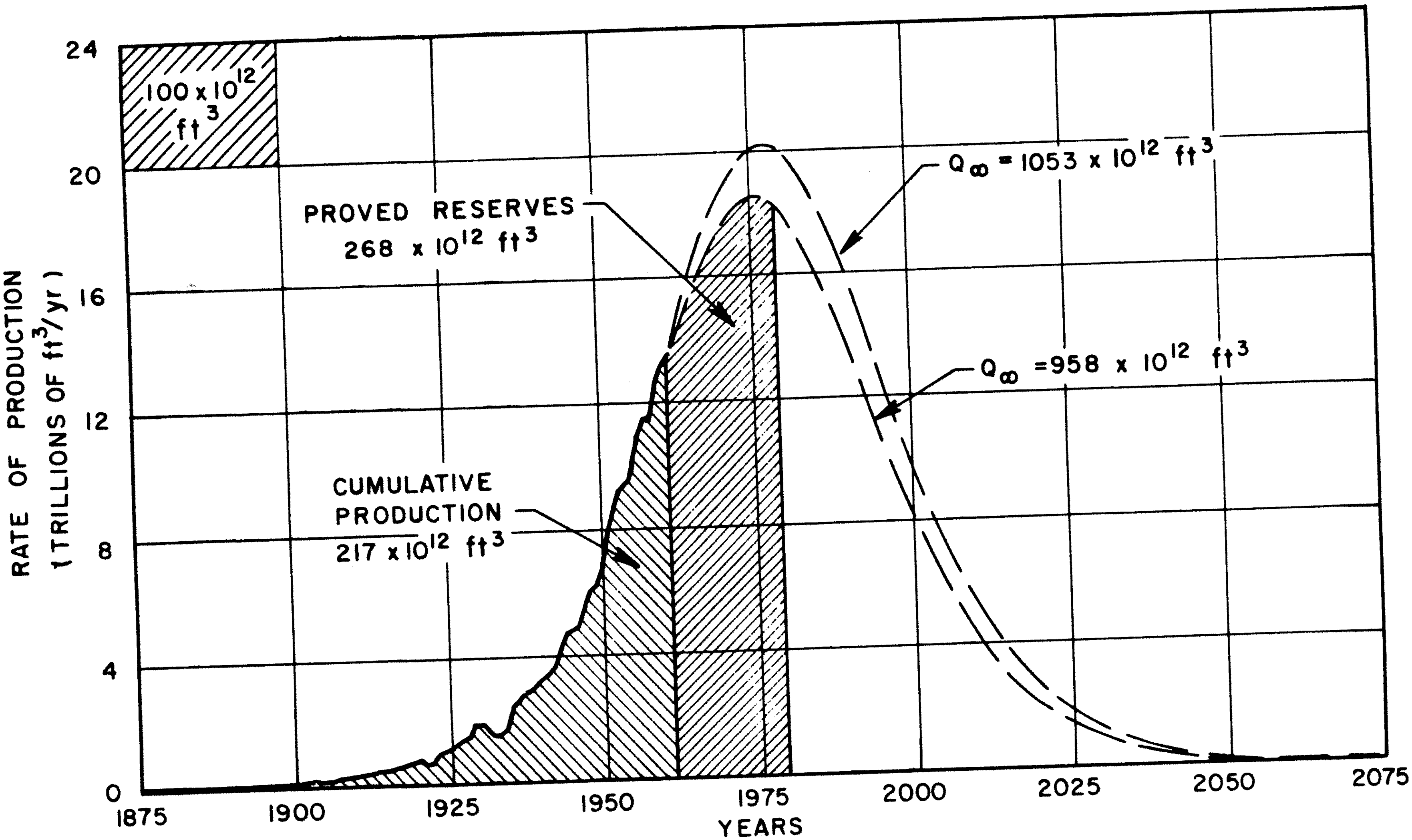

Fig. 45 - Two complete cycles of U.S. natural-gas production based upon high and low estimates of 1962 (Hubbert, 1962, Fig. 47).

The mean of these two figures of 1,050 trillion cubic feet was adopted as the value for Q_∞. This was then used, in conjunction with the statistical data for Q_p, Q_r, and Q_d, to construct Figure 46. Although in 1972 the maximum rate of natural-gas production had not yet been reached, all the evidence indicated that this would have to occur within the next two or three years. This was accordingly estimated to occur about 1975, with the peak production rate of about 24 trillion ft^3/yr. It actually occurred in 1973, with a peak rate of 22.6 trillion ft^3/yr.

The difficulty of trying to fit the cumulative data for natural gas with the logistic equation, using $1,050 \times 10^{12}$ ft^3 for Q_∞, is evident from inspection of Figure 47. The abrupt decline of proved reserves after 1967, and the corresponding downward deflection in the curve of cumulative discoveries combine to suggest that the actual figure for Q_∞ may be considerably less than the value of $1,050 \times 10^{12}$ ft^3.

Estimate of 1980. —— By 1980 the curve of cumulative proved discoveries is far enough advanced beyond its inflection point, which occurred about 1961, to permit estimates of the asymptote to which this curve is tending. For this purpose five different procedures have been used:

1. The linear regression $(dQ/dt)/Q$ versus Q.
2. The negative-exponential approach of the curve of Q_d versus t to Q_∞ as t increases.
3. Estimate of Q_∞ for gas based upon prior estimates for oil in conjunction with the gas/oil-ratio.
4. Estimate of the logistic constants for the curve of cumulative gas discoveries, Q_d, as a function of time.
5. Estimate based upon a new analysis by Root (1980) of gas discoveries per each 10^8 ft of exploratory drilling.

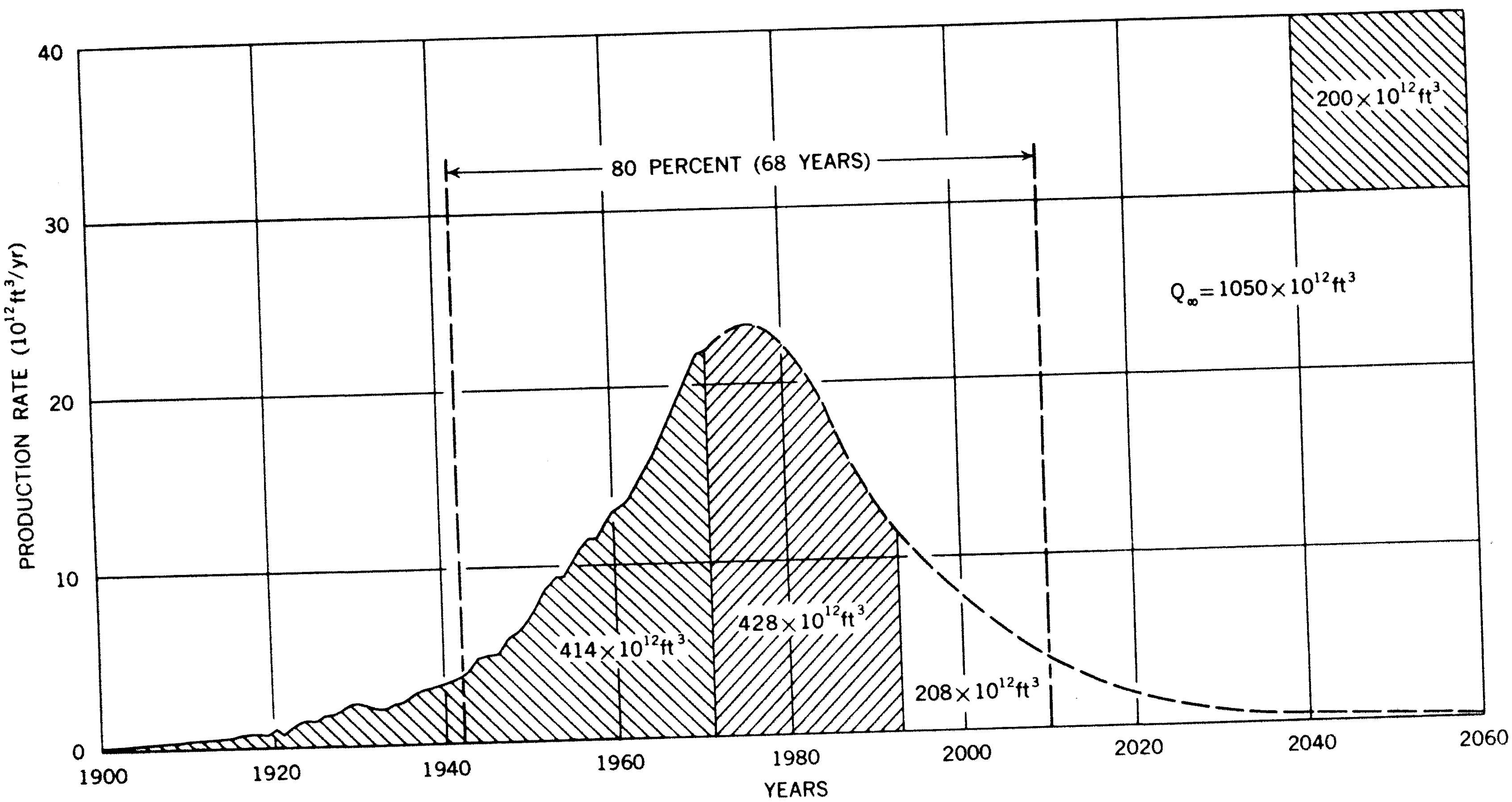

Fig. 46 - Complete cycle of U.S. natural-gas production as estimated in 1972 (Hubbert, 1974, Fig. 64).

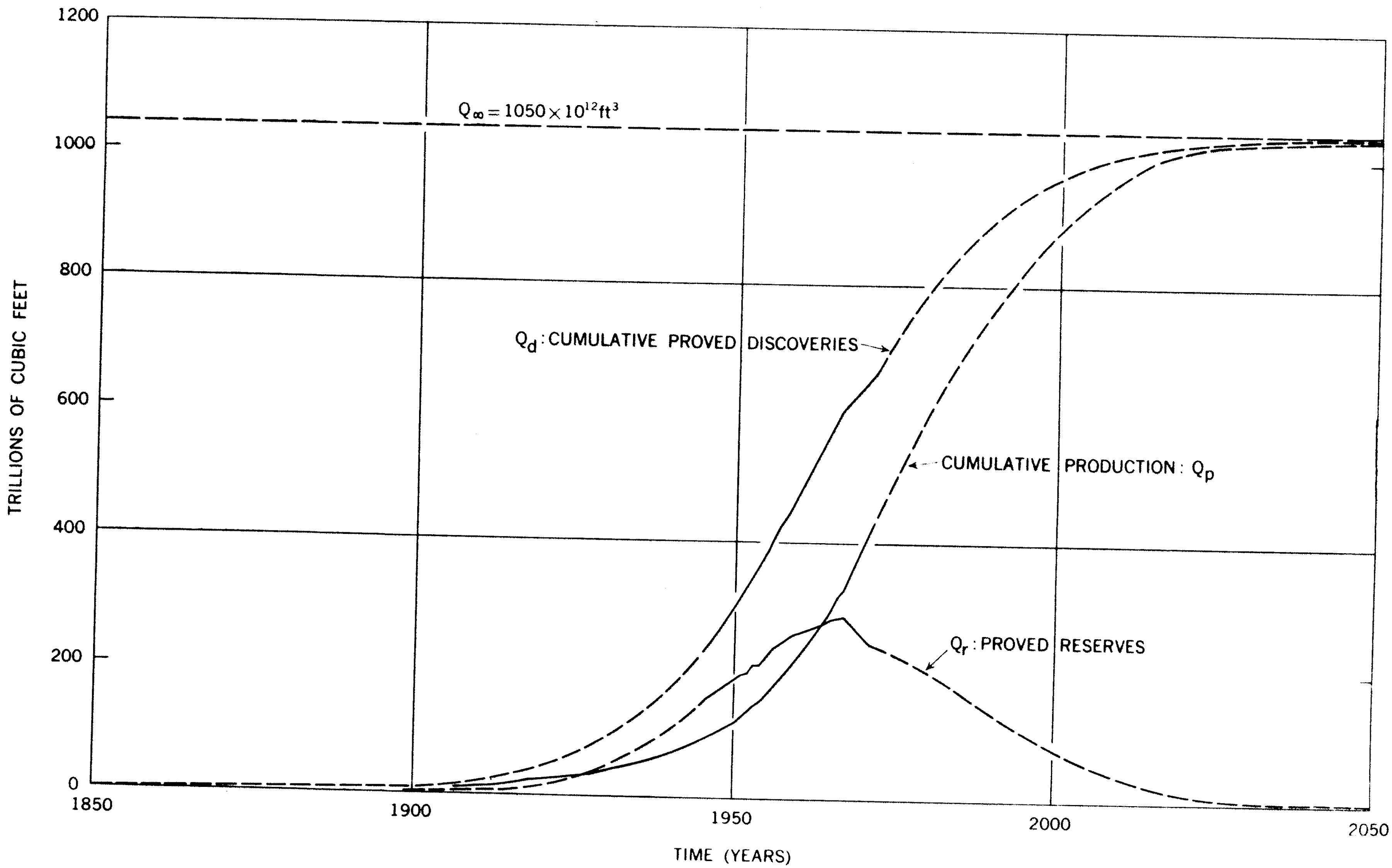

Fig. 47 - Cumulative production, proved reserves, and cumulative proved discoveries of U.S. natural gas, 1900-1972, with graphical estimates of future developments (Hubbert, 1974, Fig. 56).

Estimate based upon $(dQ/dt)/Q$ versus Q. —— The first of the foregoing procedures is based upon equation (27),

$$(dQ/dt)/Q = a(1 - Q/Q_\infty).$$

This is a linear equation between $(dQ/dt)/Q$ and Q, the graph of which intersects the Q-axis at the point $Q = Q_\infty$, and the vertical axis at $(dQ/dt)/Q = a$.

For data, the American Gas Association (May 1980) has recently published a table of cumulative proved discoveries of natural gas for the U.S. Lower-48 states corresponding to the end of each year from 1945 to 1979. Using these data, mean values for dQ/dt were computed for each year from 1950 to 1974 based upon a 10-year running average. Shorter periods of averaging were used for the years 1975 to 1978. For Q, the actual yearly figures were used. Plotting the data for $(dQ/dt)/Q$ versus Q gave a very good linear graph from Q = 480 to 720 trillion cubic feet, corresponding to the period 1960 to 1979. This line, as estimated visually, passed through the point, $(dQ/dt)/Q = 0.0500$; $Q = 400 \times 10^{12}$ ft^3, and intersected the Q-axis at 810×10^{12} ft^3, which is the estimated value for Q_∞. By backward extrapolation the line intersects the vertical axis at the point 0.099, which gives the value of the coefficient a of the corresponding logistic equation.

Although this figure of 810 trillion cubic feet is a surprisingly low figure for Q_∞, the data of the graph are sufficiently linear over the interval stated that very little latitude, possibly $\pm$ 10 trillion cubic feet, is allowable for the uncertainty of the point of intersection.

Estimate by the negative-exponential approach of Q_d versus t to Q_∞. —— The assumption that cumulative discoveries Q approach the asymptotic value Q_∞ in a negative-exponential manner during the later stages of the discovery cycle affords another means of estimation. At a fixed date t_0, let Q_0 be the magnitude of cumulative discoveries. Then using this point as a new origin of coordinates, let $y = Q - Q_0$ be the subsequent increase in Q, and let $\tau = t - t_0$ be the subsequent time coordinate. Also let k be the asymptotic value of y. We then have the equation,

$$k - y = ke^{-b\tau}, \tag{80}$$

or

$$\ln[(k - y)/k] = -b\tau, \tag{81}$$

in which the two parameters, k and b are to be determined. Consider two points of the curve of y versus τ, (τ_1, y_1) and (τ_2, y_2). Introducing these values into equation (81) then gives the two equations,

$$\left.\begin{aligned} \ln[(k - y_1)/k] &= -b\tau_1, \\ \ln[(k - y_2)/k] &= -b\tau_2. \end{aligned}\right\} \tag{82}$$

By taking the ratio of the second to the first, b can be eliminated, and we obtain

$$\frac{\ln[(k - y_2)/k]}{\ln[(k - y_1)/k]} = \frac{\tau_2}{\tau_1}, \tag{83}$$

in which k is the only unknown. This can be solved by iteration if we substitute for k an assumed value k_a. Then the left-hand term of equation (83) becomes $f(k_a)$, and when

$$f(k_a) = \tau_2/\tau_1,$$

$$k_a = k.$$

After k is known, b may be determined from equation (81) by

$$b = \ln[(k - y_1)/k]/\tau_1.$$

Then finally,

$$Q_\infty = Q_0 + k.$$

The data of Q versus t and y versus τ are given in Table 1 at 5-year intervals from 1960 to 1980. Taking τ_1 at 10 years (date 1970), and τ_2 at 20 years (1980), and y_1 and y_2 equal to 169.05 and 267.50 trillion cubic feet, respectively, we find that

Table 1. — Estimates of future natural-gas production of U.S. Lower-48 states by negative-exponential approach to Q_∞.

Date	τ (yr)	Q_d	y	Q_d Computed	Date	τ (yr)	y	Q_d Computed
		(10^{12} ft^3)					(10^{12} ft^3)	
1960	0	466.35	0	466.35	2000	40	358.22	824.57
1965	5	553.52	87.17	562.23	2010	50	377.67	844.02
1970	10	635.40	169.05	635.40	2020	60	388.99	855.34
1975	15	680.17	213.82	691.24	2030	70	395.58	861.93
1980	20	733.85	267.50	733.85	2040	80	399.42	865.77
1985	25	-	(Computed) 300.02	766.37	2050	90	401.66	868.01
1990	30	-	324.83	791.13	∞	∞	404.78	871.13

Source for Q_d: Am. Gas Assoc., May 1980.

$$f(k_a) = \tau_2/\tau_1 = 2$$

when

$$k_a = k = 404.78 \times 10^{12} \text{ ft}^3,$$

and that

$$b = 0.054066.$$

Then

$$Q_\infty = Q_0 + k$$
$$= (466 + 405) \times 10^{12} \text{ ft}^3$$
$$= 871 \times 10^{12} \text{ ft}^3.$$

Estimates based upon the gas/oil-ratio. —— As shown in equation (79), an estimate of Q_∞ for gas is obtained by

$$Q_\infty \text{ gas} = Q_d \text{ gas} + G[(Q_\infty - Q_d)\text{oil}],$$

where the cumulative discoveries for both gas and oil are the most recent figures available, at present those for 1980.0, and G, the gas/oil ratio, is the ratio of gas discoveries to oil discoveries during a recent finite period of time.

By 1980.0

$$Q_d \text{ oil} = 136.9 \times 10^9 \text{ bbl},$$
$$Q_\infty \text{ oil} = 163 \times 10^9 \text{ bbl},$$

and

$$(Q - Q_d) \text{ oil} = 26 \times 10^9 \text{ bbl},$$
$$Q_d \text{ gas} = 734 \times 10^{12} \text{ ft}^3.$$

For gas/oil-ratios during recent decades, cumulative discoveries of both crude oil and natural gas are given for 1960.0, 1970.0, and 1980.0 in Table 2. From these data three separate gas/oil-ratios are obtained for three different periods of time. For the decade 1960 to 1970, the value of G was 6,789 ft^3/bbl; for 1970 to 1980 it had declined to 5,472 ft^3/bbl; and for the 20-year period

Table 2. — Cumulative discoveries of crude oil and natural gas in U.S. Lower-48 states by 1960, 1970, and 1980, and gas/oil-ratios for 1960-1970, 1970-1980, and 1960-1980.

Date (Jan. 1)	a/ Q_d Crude oil (10^9 bbl)	b/ Q_d Natural gas (10^{12} ft^3)	Time interval Δt (yr)	ΔQ_d Crude oil (10^9 bbl)	ΔQ_d Natural gas (10^{12} ft^3)	Gas/oil-ratio G (ft^3/bbl)
1960	94.01	466.35	1960-1970	24.90	169.05	6,789
1970	118.91	635.40	1970-1980	17.99	98.45	5,472
1980	136.90	733.85	1960-1980	42.89	267.50	6,237

Sources:

a/ Am. Petrol. Institute, Am. Gas Assoc., and Canadian Petrol. Assoc., Annual "Blue Books," 1966-1979.

b/ Am. Gas Assoc., May 1980.

1960 to 1980 it had an intermediate value of 6,237 ft^3/bbl.

For use in estimating future gas discoveries the ratio of the last 10 years is evidently the best of the three figures, although the ratio for the last 20 years may also be considered. With the numerical data given above, the solution of equation (79), using the gas/oil-ratio of 5,472 ft^3/bbl, gives an estimate for Q_∞ for natural gas of

$$Q_\infty = 876 \times 10^{12}\ ft^3.$$

Using the ratio of 6,237 ft^3/bbl of the last 20 years gives the slightly higher estimate

$$Q_\infty = 896 \times 10^{12}\ ft^3.$$

Estimate based upon the constants of the logistic equation. — Despite the fact that the curve of Q_d versus t for natural gas cannot be fitted accurately by the logistic equation, still a good approximation can be obtained using the data for Q_d from 1946 to 1980, during which Q_d increased from 233 to 734 trillion cubic feet. Expressing these data in the linear form of equation (36),

$$\ln N = \ln N_0 - at,$$

and then using the technique described in equations (43) to (45), approximate values of the parameters, Q_∞ and a can be determined. For this purpose the following data were used:

Date (t)	1946.0	1965.0	1980.0
Q_d (10^{12} ft^3)	233.18	553.52	733.85

The value of Q_d for which the two line segments, that from 1946.0 to 1965.0, and that from 1965.0 to 1980.0, have the same slope was found to be 840.0×10^{12} ft^3, and $-S = 0.0850$/yr. Hence, the estimates for the logistic constants by this procedure are:

$$Q_\infty = 840 \times 10^{12}\ ft^3,$$

$$a = 0.0850/yr.$$

Estimate based upon gas discoveries per each 10^8 ft of exploratory drilling. —— David H. Root (1980) has just completed a new study of natural gas discoveries in the Lower-48 states, based upon his own modification of the Arrington method of estimating the additional gas that fields discovered each year will ultimately produce. Root has estimated the ultimate amount of gas to be produced by each of the 20 10^8-ft units of exploratory drilling extending in time from 1860 to 1977.9. This is a parallel study for natural gas to Root's crude-oil study, the results of which are shown in Figures 38 and 41.

As in the case of crude oil, the discoveries made by the first 4 units of drilling, which extended from 1860 to 1945.2, were large, averaging slightly more than 100 trillion cubic feet each. However the discoveries per unit for the entire 20 units declined in a roughly negative-exponential manner to a final figure of 13.912 trillion cubic feet for the 20th unit.

Using the method developed in equations (72) to (78), the actual data for dQ/dh versus h can be approximated by a negative-exponential decline curve,

$$R = R_o e^{-\beta h},$$

whose integral from $h = 0$ to $h = 20$ units has the same value as the sum of the actual discoveries, and which passes through the last data point on the curve.

The significant data for this determination are:

$$R_{20} = 13.912\ (10^{12}\ ft^3/10^8\ ft),$$

$$Q_{20} = 844.406 \times 10^{12}\ ft^3.$$

From these,

$$R_o = 95.04118\ (10^{12}\ ft^3/10^8\ ft),$$

$$\beta = 0.09608 \text{ per } 10^8 \text{ ft},$$

$$Q_\infty = R_o/\beta = 989.2 \times 10^{12} \text{ ft}^3,$$

$$Q_u = Q_\infty - Q_{20} = 144.8 \times 10^{12} \text{ ft}^3.$$

Summary of Estimates of Natural Gas

The foregoing estimates for the ultimate quantity of natural gas to be produced in the Lower-48 states and adjacent offshore areas are the following:

Method of estimation	Q_∞ (10^{12} ft^3)
$(dQ/dt)/Q$ vs. Q	810
Q vs. t	871
Gas/oil-ratio	876
	896
Logistic equation	840
dQ/dh vs. h	989
Mean	880

What is most impressive about these separate estimates is the range from the lowest to the highest of 810 to 989 trillion cubic feet, or approximately 900 ± 90, with a mean value of 880 trillion cubic feet. If we omit the lowest and the highest estimates, each of which differs by a large amount from that next above or below, then the remaining four figures fall within the much narrower range of 840 to 896, or 868 ± 28, with a mean value of 871 trillion cubic feet.

In this series, both the lowest figure of 810 trillion cubic feet and the highest of 989 are anomalous, but the latter is especially suspect since it exceeds the average of 871 trillion cubic feet of the middle four estimates by

118 trillion cubic feet. This analysis by Root was a companion study to that of the crude-oil discoveries as a function of cumulative depth of exploratory drilling, the results of which are shown in Figures 38 and 41. In the crude-oil analysis the data used were the API "Blue Book" data on "ultimate recovery" of crude oil by year of discovery. In the natural-gas analysis the corresponding data were the AGA "Blue Book" figures for "ultimate recovery" by year of discovery. However, the natural-gas estimate of 989 trillion cubic feet shows the same inconsistency with the corresponding crude-oil estimate as it does with the other gas estimates given above.

This can be seen by using Root's data for crude-oil discoveries in conjunction with the gas/oil-ratio. In this case,

$$Q_\infty \text{ gas} = Q_{20} \text{ gas} + GQ_u \text{ oil}. \tag{84}$$

Using Root's figures of

$$Q_{20} \text{ gas} = 844.4 \times 10^{12} \text{ ft}^3,$$

$$Q_u \text{ oil} = 4.9 \times 10^9 \text{ bbl},$$

and the two values of the gas/oil-ratio from Table 1,

$$G = 5{,}472 \text{ and } 6{,}237 \text{ ft}^3/\text{bbl},$$

gives the following two estimates for Q_∞ for natural gas:

$$Q_\infty = 871 \times 10^{12} \text{ ft}^3,$$

$$Q_\infty = 875 \times 10^{12} \text{ ft}^3.$$

These figures are consistent with those ranging from 840 to 896 trillion cubic feet obtained by other methods. Combining the mean of the above two figures, 873 trillion cubic feet, with the previous estimates (omitting the low figure of 810), gives as our present best estimate for Q_∞ for natural gas,

$$Q_\infty \text{ gas} = (870 \pm 30) \times 10^{12} \text{ ft}^3.$$

Conclusion

The principal thesis of the present paper has been that the successful prediction of the future behavior of any matter-energy system must be based upon a prior understanding of the mechanism of the system considered, and upon a rational analysis of the data of the system in accordance with that mechanism. Also, the final arbiter of the reliability of any prediction is the future itself. So long as the predicted event is still in the future, whether or not the prediction is valid must remain to some degree uncertain. But after the time has been reached at which the predicted event was to occur, this doubt no longer remains.

In this paper, the results of the application of this philosophical view to the petroleum industry of the United States during the last 25 years have been reviewed. It is now evident that by the mid-1950s the cumulative data of the U.S. petroleum industry were sufficient to permit reasonably accurate predictions of its future development. With the passage of time, more and better data have permitted a refinement of earlier estimates, and also provided a verification of their degree of accuracy. By now, the peak in the rate of crude-oil production has already been passed in 1970, and that of natural gas in 1973, and the production rates of both are now in decline.

The present cumulative statistical evidence with regard to crude oil leads to a figure of approximately 163 ± 2 billion barrels for the ultimate cumulative production in the Lower-48 states. Less exact evidence for natural gas indicates that the ultimate cumulative production from conventional sources will probably be in the range of 870 ± 30 trillion cubic feet. However there still remain geological uncertainties regarding the occurrences of undiscovered oil and gas fields, yet those are being severely restricted by the extent of exploratory activity. In the case of crude oil, there is also the uncertainty regarding the magnitude of future improvements in extraction technology.

With due regard for these uncertainties, estimates for crude oil that do not exceed that given here by more than 10 percent may still be within the range of geological uncertainties; estimates that do not exceed this by more than 20 percent may be within the combined range of geological and technological uncertainties. Estimates for natural gas that do not exceed the upper limit of the range given above by more than 10 percent may likewise be regarded as possible although improbable. But estimates for either oil or gas, such as those that have been published repeatedly during the last 25 years, which exceed the present estimates by multiples of 2, 3, or more, are so completely irreconcilable with the cumulative data of the petroleum industry as no longer to warrant being accorded the status of scientific respectability.

REFERENCES

American Association of Petroleum Geologists, Committee on Exploratory Drilling, 1946 — , Annual reports on exploratory drilling, initially in the United States but later also in Canada and Mexico: Am. Assoc. Petroleum Geologists Bull. (various numbers, annually).

American Gas Association, Gas Supply Committee, May 1980, Summary of United States gas supply statistics: Gas Energy Review, Supply and Production Supplement, p. 1-4.

American Petroleum Institute, 1959, Petroleum facts and figures, Centennial Edition, 471 p.

American Petroleum Institute, American Gas Association, and Canadian Petroleum Association, annually 1967-1980, Reserves of crude oil, natural gas liquids, and natural gas in the United States and Canada as of December 31 (each year).

Arps, J.J., and Roberts, T.G., 1958, Economics of drilling for Cretaceous oil on east flank of Denver-Julesburg Basin: Am. Assoc. Petroleum Geologists Bull., v. 42, no. 11, p. 2549-2566.

Arps, J.J., Mortada, M., and Smith, A.E., 1971, Relationship between proved reserves and exploratory effort: Journal of Petroleum Technology, June, p. 671-675.

Arrington, J.R., 1960, Size of crude reserves is key to evaluating exploration programs: Oil and Gas Journal, v. 58, no. 9, Feb. 29, 4 p.

——— 1966, Estimation of future reserve revision in current fields: *In* Trans. West Texas Geological Society's Symposium: Economics and the Petroleum Geologist, p. 17-30.

Clarke, F.W., 1924, The data of geochemistry, Fifth edition: U.S. Geological Survey Bull. 770, 841 p.

Cram, Ira H., ed., 1971, Future petroleum provinces of the United States – their geology and potential: Summary: Am. Assoc. Petroleum Geologists, Mem. 15, v. 1, Preface, viii; p. 1-34.

DeGolyer and MacNaughton, 1944 — , Twentieth Century Petroleum Statistics, Dallas, annually.

Drew, Lawrence J., Schuenemeyer, J.H., and Root, David H., 1980, Resource appraisal and discovery rate forecasting in partially explored regions: Part A, An application to the Denver Basin: U.S. Geological Survey Professional Paper 1138, p. A1-A13.

Duncan, D.C., and McKelvey, V.E., May 1963, U.S. resources of fossil fuels, radio-active minerals, and geothermal energy: *In* Federal Council for Science and Technology, Research and development on natural resources, Appendix A, p. 43-45, table 9.

Frey, John W., and Ide, H. Chandler, 1946, A history of the Petroleum Administration for War, 1941-1945: Washington, Petroleum Administration for War.

Gardett, Peter H., 1971, Petroleum potential of the Los Angeles Basin, California: *In* Am. Assoc. Petroleum Geologists, Memoir 15, v. 1, p. 298-308.

Hubbert, M. King, 1956, Nuclear energy and the fossil fuels: Dallas, Am. Petroleum Inst., Drilling and Production Practice (1956), p. 7-25.

——— 1962, Energy Resources, A report to the Committee on Natural Resources: Natl. Acad. Sci.-Natl. Research Council, Pub. 1000-D, 141 p.; Reprinted 1973 by Natl. Tech. Inf. Service, U.S. Dept. of Commerce, no. PB-222401.

——— 1967, Degree of advancement of petroleum exploration in United States: Am. Assoc. Petroleum Geologists Bull., v. 51, no. 11, p. 2207-2227.

——— 1974, U.S. energy resources, a review as of 1972, Pt. 1, *in* A national fuels and energy policy study: U.S. 93rd Congress, 2d Session, Senate Committee on Interior and Insular Affairs, Serial no. 93-40 (92-75), Washington, U.S. Govt. Printing Office, 267 p.

Kilkenny, John E., 1971, Future petroleum potential of region 2, Pacific Coastal States and adjacent continental shelf and slope: *In* Am. Assoc. Petroleum Geologists Memoir 15, v. 1, p. 170-191.

Lahee, Frederick H., 1962, Statistics of exploratory drilling in the United States 1945-1960: Am. Assoc. Petroleum Geologists, 135 p.

Law, J., 1957, Reasons for Persian Gulf oil abundance: Am. Assoc. Petroleum Geologists Bull., v. 41, p. 51-69; Reprinted *in* A.A.P.G. Foreign reprint series no. 2, 1978, p. 43-61.

Marsh, G. Rogge, 1971, How much oil are we really finding?: Oil and Gas Journal v. 69, no. 14, April 5, p. 100-104.

McKelvey, V.E., 1967, Contradictions in energy resource estimates: The Technological Institute, Northwestern Univ., The Seventh Biennial Gas Dynamics Symposium, p. 5-23.

McKelvey, V.E., and Duncan, D.C., 1965, United States and world resources of energy, *in* Symposium on Fuel and Energy Economics, Joint with Division on Chemical, Marketing and Economics: 149th Natl. Meeting, Am. Chem. Soc., Div. Fuel Chemistry, v. 9, no. 2, p. 1-17.

Menard, H.W., and Sharman, George, 1975, Scientific uses of random drilling models: Science, Oct. 24, v. 190, p. 337-343.

Moore, C.L., 1962, Method for evaluating United States crude oil resources and projecting domestic crude oil availability: Washington, Dept. Interior, Off. of Oil and Gas, 112 p.

Morris, Anthony E.L., 1978, Preface and bibliography: Geology and productivity of Arabian Gulf geosyncline: *In* Geology and productivity of Arabian Gulf geosyncline: Am. Assoc. of Petroleum Geologists, Foreign reprint series, no. 2, 8 p.

National Petroleum Council, 1961, Proved discoveries and productive capacity of crude oil, natural gas and natural gas liquids in the United States: Washington, Natl. Petroleum Council, 40 p., 7 tables, 6 exhibits.

——— 1965, Proved discoveries and productive capacity of crude oil, natural gas and natural gas liquids in the United States: Washington, Natl. Petroleum Council, 23 p., 6 tables, 1 exhibit.

Root, David H., 1980, Historical growth of estimates of oil and gas field size: Paper presented on June 19, 1980, before Department of Energy and National Bureau of Standards Symposium on Oil and Gas Supply Modeling, Washington, D.C.

Root, David H., and Drew, Lawrence J., 1979, The pattern of petroleum discovery rates: American Scientist, v. 67, no. 6, p. 648-652.

Verhulst, P. -F., 1838, Notice sur la loi que la population suit dans son accroissement: Corr. Math. et Phys., T. X, p. 113-121.

——— 1845, Recherches mathématiques sur la loi d'accroissement de la population: Mémoires de L'Académie Royale, Belgique, T. XVIII, p. 1-38.

——— 1847, Deuxième Mémoire sur la loi d'accroissement de la population: Mémoires de L'Académie Royale, Belgique, T. XX, p. 1-32.

DISCUSSION

DR. HUBBERT (in reply to question by Samuel Kao - Brookhaven): Your statement that all of my curves are symmetric is not entirely correct. I have stated explicitly that the complete-cycle curve of production of an exhaustible resource in a given region has the following essential properties: The rate of production as a function of time begins at zero. It then increases exponentially during a period of development and later exploration and discovery. Eventually the curve reaches one or more maxima, and finally, as the resource is depleted, the curve goes into a negative-exponential decline back to zero. There is no requirement that such a curve be symmetrical or that it have only a single maximum. In small regions such a curve can be very irregular, but in a large area such as the United States or the world these irregularities tend to smooth out and a curve with only a single principal maximum results. If such curves are also approximately symmetrical it is only because their data make them so.

In my figure of 1956, showing two complete cycles for U.S. crude-oil production, these curves were not derived from any mathematical equation. They were simply tailored by hand subject to the constraints of a negative-exponential decline and a subtended area defined by the prior estimates for the ultimate production. Subject to these constraints, with the same data, I suggest that anyone interested should draw the curves himself. They cannot be very different from those I have shown.

DR. HUBBERT: As I have stated before, there is no theoretical necessity for the complete-cycle curve to be symmetrical. When such curves are symmetrical it is only because the data require that they be so. A critical test of whether such a curve is symmetrical or not is the linear equation,

$$\ln N = \ln N_o - a(t - t_o),$$

where

$N = (Q_\infty - Q)/Q$,

Q = cumulative discoveries or production,

Q_∞ = the ultimate value of Q,

a = the growth constant.

This is the linear form of the symmetric logistic equation. If the quantity $\ln N$ plots as a straight line as a function of time, this is evidence that the cumulative data increase in accordance with the symmetric logistic equation.

For cumulative discoveries and production of crude oil in the U.S. Lower-48 states, during the period 1900-1973, the data plot as excellent straight lines in accordance with the above equation. For discoveries, the maximum rate occurred at about 1957. However, from 1973 to 1980, the discovery rate has been declining faster than the equation would predict.

DR. HUBBERT (in response to remarks by David Nissen - DOE): Your kind remarks with regard to my previous studies of the evolution of the U.S. petroleum industry are greatly appreciated. However, you suggest that my estimates of the ultimate amount of oil to be recovered is questionable for reasons of classification and because I have not taken into account the effect of the price of oil on ultimate recovery. You mention oil shale, coal, and the Orinoco heavy oils of Venezuela.

With regard to classification, if unintelligibility is to be avoided, it is essential that one define his terms and then adhere rigorously to those definitions. In the present study I have been concerned with the techniques of estimation as applied to conventional crude oil and natural gas in the U.S. Lower-48 states. This excludes consideration of shale oil, coal, Orinoco heavy oils, natural gas from unconventional sources, and also oil and gas from Alaska.

My analyses are based upon the simple, fundamental geologic fact that initially there was only a fixed and finite amount of oil in the ground, and that, as exploitation proceeds, the amount of oil remaining diminishes monotonically. We do not know how much oil was present originally or what fraction of this will ultimately be recovered. These are among the quantities that we are trying to estimate.

Your statement that the fraction of the original oil-in-place that will be recovered is a function of the price of oil is correct, but the effect may easily be exaggerated. For example, we know now how to get oil out of a reservoir sand, but at what cost? If oil had the price of pharmaceuticals and could be sold in unlimited quantity, we probably would get it all out except the smell. However there is a different and more fundamental cost that is independent of the monetary price. That is the energy cost of exploration and production. So long as oil is used as a source of energy, when the energy cost of recovering a barrel of oil becomes greater than the energy content of the oil, production will cease no matter what the monetary price may be. During the last decade we have had very large increases in the monetary price of oil. This has stimulated an accelerated program of exploratory drilling and a slightly increased rate of discovery, but the discoveries per foot of exploratory drilling have continuously declined from an initial rate of about 200 barrels per foot to a present rate of only 8 barrels per foot.

There is the further question of what fraction of the original oil-in-place is now being recovered. The conventional figure most frequently quoted is about one-third. However, a critical review of this question in a book entitled "Determination of Residual Oil Saturation" by a panel of nationally prominent petroleum engineers has just been published by the Interstate Oil Compact Commission (June 1978). In this study the average value of the residual oil saturation in the depleted reservoir sands of a hundred or so fields was found to be only 28 percent, as compared with previous estimates of 38 percent. According to this study, the recovery factor at present is evidently much higher than has been conventionally assumed, and the remaining oil correspondingly smaller.