Types of Novae*

F. ZWICKY
Norman Bridge Laboratory, California Institute of Technology, Pasadena, California

A general classification of novae in terms of integral parameters is proposed. It is shown in particular that the distribution of novae in dependence on the intrinsic luminosity at maximum brightness reveals the existence of two distinct classes, designated as *common novae* and *supernovae*. In Section (A) the usefulness of the study of temporary stars in comparison with the study of permanent stars is discussed. Sections (B) and (C) give a tentative classification of novae in terms of integral parameters as well as a brief tabulation of some of the observational data on these parameters. In the course of this discussion a number of interesting problems suggest themselves. Among them the problems of the primary and secondary images of novae and of supernovae and the investigation of the final states of temporary stars are of particular interest. In (D) some approximate relations among the integral parameters are described. In (E) the frequency of occurrence $n(M)$ of novae in dependence on their absolute magnitude M is explicitly established and graphically represented. The fundamental significance of the fact is emphasized that the function $n(M)$ has two pronounced maxima. In (F) the importance of the fact is analyzed that supernovae preferably appear in those parts of nebulae whose surface brightness is *low*. Two outstanding cases of supernovae which are located in points of excessively low surface brightness are reproduced in Fig. 3. In Section (G) it is shown that the average frequency of occurrence of supernovae in different types of nebulae does not depend essentially on the type. In Section (H) a number of reasons are advanced which appear to be sufficient to rule out the assumption of collisions between luminous stars as a cause for supernovae. Finally in Section (I) a review is given of the decisive observations and considerations which have resulted in the recognition of supernovae as a separate class of temporary stars.

(A) Significance of the Study of Novae

In recent years much time and energy have been devoted to the search for novae and for supernovae. In addition valuable telescopic equipment has been tied up for relatively long periods of time in the photometric and the spectroscopic investigation of all newly discovered objects. Although the results achieved so far are gratifying, it has nevertheless become clear that new fundamental and ever more difficult problems have presented themselves in the course of these investigations, problems the solution of which can be achieved only through the engagement of an increased number of observers and of telescopes. The question naturally arises if the results to be obtained represent a fair return of the efforts and the investments required. In order to arrive at least at a partial answer to this question, it is perhaps useful to compare the relative significance of the study of ordinary "permanent" stars with that of "temporary" stars or novae.

Although there are many aspects under which stars may be investigated, we confine ourselves here to a few questions which are more or less directly related to the problem of the generation of energy in stars and the evolution of stars and stellar systems.

In order to appraise the relative value of observations on permanent stars and on temporary stars, respectively, we sketch briefly what kind of observations may be made on both types of stars with present day telescopic equipment.

(1) Observations on *permanent stars* yield first of all their *apparent luminosity* for light of different wave-lengths which is not subject to complete absorption in the earth's atmosphere. In the case of a limited number of stars such as the sun and some of its nearest neighbors the *distance* and consequently the *absolute luminosity* can also be directly determined. For the majority of the stars the distances and absolute luminosities are subject to statistical considerations which yield results of more or less limited accuracy. Observations on the *spectra* of stars furnish us with data on the chemical composition of stellar atmospheres and their *physical conditions* such as pressure, temperature and the stages of ionization which determine the processes of emission

* This paper which was originally sent to the *Physical Review* has been transferred to the *Reviews of Modern Physics* at the suggestion of the editors.

and absorption of light by the various elements present. In many favorable cases of binary stars the *masses*, the *radii* and sometimes even the *heights of the atmospheres* can be quite accurately determined by direct observations. For some nearby stars the radii may be measured interferometrically. For many other stars the quantities mentioned are subject to an indirect statistical treatment. The photometric and the spectroscopic analysis of the light from stars reveals the degree of *stability* and *permanence* or of *variability*, regular or irregular, as the case may be, of the physical processes which govern the internal constitution of these stars. Additional properties such as the *rotation*, the surface *evaporation*, the *magnetic* and *electric fields* leave their imprints in the spectra of stars.

Theories concerned with the internal constitution of stars are confronted with the difficulty that all of the data available are essentially derived from observations of superficial physical conditions only. It must not be overlooked, however, that theories about the internal constitution of stars and the generation of energy must not account only for the properties mentioned in the preceding, but that they also must be in accord with the facts of the permanence and the stability or the variability of stars. The *permanence* and the *stability* of the majority of stars impose on any kind of theoretical models restrictions which are of a very severe character inasmuch as they refer to a great number of degrees of freedom which characterize the various possible observable changes. Much attention has been given in recent years to the problem of the generation of energy in stars. But it is just as important to account for the stability of stars. In this connection the possibility of the existence of partly or completely degenerate cores, such as neutron cores and neutron stars may be of importance.

With respect to the question of the *evolution of stars* and stellar systems, *counts* of the numbers of different types of stars, the stellar content of different types of nebulae, the *spatial distribution* and the *velocity distribution* of stars are relevant. A satisfactory interpretation of data of this kind is entirely lacking at the present.

(2) Observations on novae and on *supernovae* involve most of the properties which are of interest in the study of permanent stars. The important distinction is that here the theory, instead of being confronted with the problem of the stability of certain stars, must rather account for the interesting fact of obvious *instabilities*. The observations on novae include the determination of the light curves in various spectral ranges, the spectral analysis of the light emitted by the ejected gases and by the stellar remnants, the velocities and the masses of the ejected gases, the initial and the final stages of a nova, the electric and the magnetic fields present and the frequency of occurrence of novae and of supernovae in different stellar systems.

B. Tentative Classification of Novae

The greatest interest of the study of novae and of supernovae lies perhaps in the fact that the observations indicate that the instability in certain stars causes the rate of generation of energy to be multiplied temporarily by factors 10^4 to 10^9 and that the evolution from one stellar type into another proceeds at a pace which far surpasses that of ordinary permanent stars. Here therefore we deal with happenings which may furnish us with direct clues regarding the generation of energy and the evolution of stars and of stellar systems, problems about whose fundamental character we obtain only indirect information from the study of permanent stars. On the other hand, it is not *a priori* certain that the processes which determine the exceptional case of a nova are of essential importance for the generation of energy in permanent stars and for their evolution. Nevertheless, the possibility that novae perhaps form universal and, in many details, observable links in the evolution of stars would seem to justify amply the efforts necessary to study as many novae and supernovae as possible.

In the present investigation we shall be particularly concerned with the problem of the classification of different types of novae. In attempting such a classification we must first of all decide which characteristic properties of novae should be regarded as of prime importance. A superficial scrutiny of the great mass of observations made on novae reveals that no two novae are alike in all of their observable details. An initial classification of novae therefore must restrict itself to less detailed properties. It seems reasonable to confine oneself to what might be called *integral properties* P_i, such as the luminosity L_{max} at maximum brightness, the total energy E_{vis} emitted in the form of visible radiation, the Mass M_G and the average velocity of expansion $\bar{v}$ of the ejected gases, a suitably defined lifetime τ and the frequency ν of various types of novae. The integral properties of the initial and final stars of a novae, such as their mass, radius, luminosity and spectral type presumably must also be included as determining classifying parameters.

If we restrict ourselves to integral properties P_i of the type mentioned in the preceding the problem of the classification of novae requires the answer to questions of the following type:

(1) With what degree of completeness may novae be characterized through the consideration of integral parameters P_i?

(2) Do there exist any relations, exact or approximate, among the parameters P_i of a nova?

(3) If the maximum number of independent parameters P_i is z, how does the frequency of occurrence ν of novae in a given stellar system depend on these z parameters?

(4) How does the frequency ν depend on the type of the stellar system and what is the spatial distribution of novae in this system?

The *data* available at the present time are far too scanty to make possible an exhaustive answer to any of the questions just stated. The *theory* of novae finds itself in a still worse predicament. Except perhaps with respect to the second question no comfort whatever can be derived from any of the theories which have been propounded so far. Under these circumstances it seems useful to try to extract from the available observational material at least some partial answers to our four questions and to outline new programs which give promise of producing results which will fill in the most obvious gaps in our knowledge of novae.

C. The Integral Characteristics of Novae and of Supernovae

Both in the case of novae and of supernovae our knowledge of their integral properties is very incomplete, although for different reasons. Let us consider some of these properties and discuss a few of the conclusions to be drawn from them.

(1) The *light curves* embody the most obvious properties which distinguish novae and supernovae from ordinary stars. Most of these light curves exhibit a sharp rise from an initial luminosity L_i to a maximum luminosity L_{max} followed by a more or less gradual decline in brightness. The major part of the initial rise in brightness often takes place during a few days. The decline may occupy a more or less large number of years. Whereas the light curve of a typical ordinary nova shows a narrow peak or *flash* near maximum, the maximum of supernovae is represented by a more extended and rounded peak. The light curves of ordinary novae often differ considerably from one another and a classification is therefore difficult. Supernovae on the other hand show a far more pronounced *uniformity* inasmuch as their light curves differ only inconsiderably. This will obviously be important for the interpretation of the processes which take place in supernovae. Also, as Baade has pointed out, the similarity of the light curves of supernovae makes possible an accurate determination of the date and the luminosity at maximum when only fragmentary data on the magnitudes are available.

Because of the great distance of supernovae we cannot always be quite certain that all of the light observed comes *directly* from the supernova itself and from the gases which it has expelled. Just as has been observed in the case of ordinary novae, a supernova may be imbedded in clouds of gases and of dust which have been present before the outburst. These clouds, through scattering and fluorescence, may indirectly transmit to the observer some of the light from the supernova which otherwise could not be observed. Suppose for instance that a supernova is located at a distance of Y million light years. Gases and dust clouds surrounding the supernova to a distance of y light years may be illuminated and produce an image of the apparent angular diameter

$$d = 2y/10^6 Y. \quad (1)$$

With the 100-inch telescope such an image can be distinguished from a stellar image only if

$$d > 1 \text{ second of arc} = 1/206265 \text{ radian}; \quad (2)$$

that is, only if

$$y/Y>2.5. \quad (3)$$

Since for all of the supernovae observed we have $Y>1$, the indirect image becomes distinguishable from the direct image at the earliest a few years ($y>2.5$) after the maximum. It will be of great interest to attempt the observation of such indirect images* in order to arrive at independent estimates of the distance.

The superposition of an indirect image on the direct image of a supernova may also make itself felt in other ways. For instance we now know that the light curves of bright supernovae are very similar and that the decline in brightness for several years after the maximum is very slow and does not exhibit any marked fluctuations. Hence, if the image of a supernova a long time after maximum should suddenly and irregularly brighten up, a spectroscopic investigation might reveal that this increase in brightness is due to scattered light of the same spectral distribution as that which was previously observed in the light of the supernova near maximum.

As a result of the impossibility to distinguish at great distances between the direct and the indirect images, Baade and I were, during our early studies of supernovae, confronted with the somewhat disconcerting possibility that supernovae might really be common novae in disguise, since it was imaginable that, because of cumulative effects (reflection, scattering, excitation of fluorescence by the ultraviolet light and by the fast particles ejected from a nova) in the gas and dust clouds originally surrounding a nova, the total apparent luminosity of the combined direct and indirect image might be greatly enhanced. The decision in favor of a *separate class* of supernovae was made possible through our investigations on the uniformity of the light curves of supernovae and especially because of the fact that the function $n(M)$ which expresses the frequency of occurrence of novae in dependence of their absolute photographic magnitude M possesses at least two maxima which are separated by a deep gap. As soon as we had been able to establish the *similarity of the light curves* of different supernovae through our recent investigations, it of course became clear that, at best, only a small fraction of the total luminosity of the supernovae observed so far could be due to light of the indirect images. In addition, the fact that $n(M)$ has a deep minimum near the absolute magnitude $M=-10.5$ suggests strongly that common novae and supernovae are quite different objects as far as the fundamental physical processes which cause the respective outbursts are concerned. The suggestion by Baade and myself of the existence of a separate class of supernovae has found definitive proof in the recently established fact[1] that the spectra of different supernovae are very similar among themselves but that they are entirely different from the spectra of common novae.

In this connection it is suggestive to think of the relatively narrow emission lines λ6299 and λ6359 observed by R. Minkowski in some supernovae and ascribed by him to O_I. These lines which contain a very small fraction of the light from a supernova might be due to the excitation of oxygen atoms by the ultraviolet light from the supernova or rather the photoelectrons ejected by this light in the stationary gases which originally surround the supernova.

In following up the light curves of bright supernovae it will be well to keep in mind the possibilities just discussed. Also it is advisable to investigate carefully all of the faint variable stars which appear superposed on the images of extragalactic nebulae in order to ascertain whether these stars belong to our own galaxy or if they are indirect images of variable brightness of past supernovae in extragalactic systems. In my search for supernovae I have so far found about a dozen variables which are within the image of nebulae or at least as close to these images as some of the supernovae were. So far most of these objects have turned out to be ordinary long period variables. If, however, any stars of very irregular variability should be found superposed on nebulae it is proposed to investigate their spectra.

(2) The *spectra* of novae and of supernovae

* As Professor J. H. Oort has pointed out to us, the illumination of surrounding clouds might become a spectacular phenomenon in the case of galactic supernovae. The search for the illumination of galactic clouds by past suspected supernovae such as Tycho's star and the stars which gave rise to the Crab nebula and perhaps other nebulosities presents an interesting problem.

[1] R. Minkowski, Astrophys. J. **89**, 156 (1939).

are of fundamental importance inasmuch as they enable us clearly to distinguish these temporary stars from all other types of stars. Again, as in the case of the light curves, the spectra of different ordinary novae show a great diversity in their detailed structure whereas the spectra of supernovae are much more similar. The spectra of common novae have already found a satisfactory explanation in most of their fundamental aspects, but the spectra of supernovae, in spite of their uniformity, so far present us with a complete puzzle. We must therefore refrain for the present from comparing the spectra of these two types of novae, except for a few remarks.

A marked distinction between the spectra of common novae and of supernovae lies in the following. Before maximum and near maximum a common nova shows essentially an absorption spectrum. The spectrum at a later time disintegrates into a complicated system of emission bands and absorption lines superposed on a continuous background. The background now is thought to originate in the stellar remnant. On the other hand, the features of the spectrum of a supernova are essentially the same immediately (a few days) before maximum brightness and up to several hundred days after maximum. It will be essential to discover supernovae in stages still earlier than hitherto observed and to follow their spectra to still later stages than has been so far possible. For this purpose it is intended to organize a vigorous search for bright supernovae in nearby extragalactic systems.

(3) The problem of the *original stars* in which nova outbursts take place is one of great difficulties. As the systematic photographic and spectroscopic surveys of stars in our Milky Way proceed the chances increase that we shall get important information regarding the types of stars which give rise to the phenomenon of a common nova. At the present time, however, very little is known about the original stars of novae. For supernovae the problem is still more intricate. Obviously, because of the great distance, there is little hope that we shall be able to identify the characteristics of any individual original star associated with a supernova. So far we can estimate only upper limits for the original luminosity of some supernovae. For instance the extragalactic nebula IC 4182 had been photographed by Baade with the 100-inch telescope a year before the appearance of the bright supernova which was discovered in this system in August, 1937. No star brighter than the apparent magnitude $m_i = 21$ has been present in the place where in 1937 the supernova flared up to $m_{max} = 8.2$. In Table I we give a few of the data available for common novae and for supernovae. The estimated photographic magnitudes m_i of the original stars are given in column 2, whereas the values of the magnitudes m_{max} at maximum brightness are listed in column 3.

TABLE I. *Data on novae. The values of m_i and m_{max} for the common novae are mostly taken from E. Schneller,* Katalog und Ephemeriden veränderlicher Sterne fuer 1939, *whereas for m_f the values given by M. L. Humason* are listed. Numbers in ordinary straight type are visual magnitudes and numbers in italics are photographic magnitudes.*

	YEAR	m_i	m_{max}	Δm	m_f	$\Delta' m$	M_{max}	$\bar{v}$ IN KM/SEC.
Nova Persei	1901	14.0	0.0	14.0	11.8–14	11.8–14	*−8.9*	1570
Geminorum	1912	14.5	3.7	10.8	14.7	11	—	820
Aquilae	1918	10.8	*−1.4*	12.2	10.8	12.2	*−8.9*	1700
Monocerotis	1918	*15.1*	*5.4*	*9.7*	16.5	11.1	—	1570
Lyrae	1919	14.9	6.5	8.4	15.2	8.7	—	950
Cygni	1920	>17.0	1.5	>15.5	15.5	14	—	610–1800
Pictoris	1925	*12.7*	*1.0*	*11.7*	—	—	—	330
Ophiuchi	1933	11.8	4.3	7.5	11.8	7.5	—	70
Herculis	1934	15.0	1.3	13.7	—	—	—	400–1000
Lacertae	1936	>16.0	*1.9*	>14.0	—	—	—	3800
Aquilae I	1936	>16.5	*7.0*	>9.5	—	—	—	1100
Aquilae II	1936	>15.0	*5.9*	>9.1	—	—	—	1900
Sagitarii	1936	*15.0*	*4.5*	*10.5*	—	—	—	3600
Supernova IC 4182	1937	*>21.0*	*8.2*	*>12.8*	*>21.0*	*>12.8*	*−16.6*	>15000
Tycho's star	1572	—	−4.5	—	*>17.0*	*>21.5*	—	—
Nova 1054 Stellar remnant		—	*−2.2*	—	*≥16.5*	*>18.7*	*−13.1*	—
Nova 1054 Crab nebula		—	—	—	*∼8.0*	—	—	1300

* M. L. Humason, Astrophys. J. **88**, 47 (1938).

In column 4 are given the values $\Delta m = m_i - m_{max}$. It is seen that the greatest rise in brightness for ordinary novae amounts to $\Delta m \geqslant 15.5$ (Nova Cygni 1920). No definite statement regarding the possible highest values of Δm for supernovae can as yet be made.

(4) The determination of the brightness and the spectral characteristics of the *final* stages of novae and of supernovae may, in principle, always be carried out if powerful enough instruments are available. Actually a considerable amount of information about the final stages of common novae is already available. For some of the most recent data we refer to the investigations by M. L. Humason[2] who has shown, for instance, that 16 former novae now exhibit very early B- and O-type spectra. Column six contains the values of $\Delta' m = m_f - m_{max}$. It is of importance to notice that $\Delta' m$ for ordinary novae does not exceed the value 14, even if a much larger number of novae for which data are available is considered.

On the other hand, no end product of any supernova has been identified as yet. There is, however, a good chance that soon some knowledge will be gained about the final stages of supernovae both in our own galaxy as well as in extragalactic stellar systems. The investigation of the end products of supernovae promises to throw much needed light on the cause for supernova outbursts. Also, if our suggestion[3] regarding the formation of exceedingly dense cores in supernovae is correct, we shall be able to learn something about matter in bulk whose average density may be of the order of the density usually encountered only inside of atomic nuclei. For these reasons it will be of interest to discuss here, although necessarily in a very tentative way, some possibilities regarding the end products of past supernovae.

As far as objects in our own Milky Way are concerned we have previously pointed out that the Crab nebula in conjunction with the nova observed by the Japanese and the Chinese in 1054, Tycho Brahe's star of 1572 and the so-called Cygnus loop NGC 6960/6992/6995 arouse our suspicion.

[2] M. L. Humason, Astrophys. J. **88**, 228 (1938).
[3] W. Baade and F. Zwicky, Proc. Nat. Acad. Sci. **20**, 254 (1934).

Reasons for the assumption that the Crab nebula is due to a former supernova have been discussed by Baade.[4] The figures given in Table I are taken from his paper. It is not certain which of the stars imbedded in the Crab nebula should be regarded as the stellar remnant of the nova of 1054 A.D. There are two stars, both of about photographic apparent magnitude $m_s = 16.6$ near the center of the nebula one of which may be the remnant. In any case, the stellar remnant is almost certainly fainter than $m = 16$. The integrated photographic magnitude of the nebula is of the order $m = 8$. Much of the light emitted by the Crab nebula is concentrated in emission lines of hydrogen, nitrogen, oxygen, etc., the atoms of which elements must be excited by the ultraviolet light (or by particles) emitted from some central star. From the unusually large difference

$$m_{star} - m_{nebula} \geqslant 8 \quad (4)$$

we conclude that the central star must be exceedingly hot. On the assumption that its ultraviolet light causes the fluorescence in the gases of the Crab nebula it follows from (4) that the total radiation L_{total} from the central star is

$$L_{total} > 1000 L_{vis}, \quad (5)$$

where L_{vis} is this star's radiation in the visible part of the spectrum. To arrive at a *lowest* estimate of the temperature T of the central star we assume that this star emits a blackbody radiation. For frequencies ν in the visual range (and high enough temperatures) the intensity of the radiation is then given by the Rayleigh-Jeans law

$$I_\nu = 2\pi k T \nu^2 / c, \quad (6)$$

where k is Boltzmann's constant and c the velocity of light. Therefore

$$L_{vis} \cong S\beta \int_0^{\nu_v} I_\nu d\nu = 1.34\beta \times 10^8 ST, \quad (7)$$

where β is a numerical constant somewhat smaller than unity ($\beta \sim \frac{7}{8}$), and S is the effective surface of the central star and $\nu_v = 7.5 \times 10^{14}$ sec.$^{-1}$. On the other hand, we have

$$L_{total} = acST^4/4, \quad (8)$$

where $a = 7.63 \times 10^{-15}$ erg/cm^2 sec. deg.^{4} is the

[4] W. Baade, Astrophys. J. **88**, 285 (1938); F. Zwicky, *ibid.* **88**, 529 (1938).

Stefan-Boltzmann constant. Combining (5), (7) and (8) we arrive at the result

$$T > 133{,}000 \text{ degrees absolute.} \quad (9)$$

A more accurate estimate of the lowest possible value of T can be obtained from Zanstra's theory[5] of the excitation of emission lines in nebulosities surrounding very hot stars. For a difference $m_{\text{star}} - m_{\text{nebula}} = 5.7$, Zanstra estimates a $T_{\text{min}} = 200{,}000$ degrees which means that for a difference of the order of 8 or more temperatures of the central star are required which do not fall far short of half a million degrees.

From a series of quite different considerations I had previously concluded[6] that the surface temperatures of the stellar remnants (stellar cores) of supernovae near and after maximum brightness are probably considerably greater than 200,000 degrees, a result which looks very suggestive when compared with the estimate just given of the stellar remnant of the nova in 1054 A.D. which, after more than 800 years, still excites the intense fluorescence in the Crab nebula.

The present velocity of expansion of the Crab nebula of only 1300 km/sec. would seem to be exceedingly low for a supernova. We refer, however, for possible explanations to the discussion of the velocities in Section C 5. The question, whether or not the Crab nebula had its origin in a supernova outburst, may possibly find an indirect answer through the continued investigation of extragalactic supernovae. The absolute magnitude of the Crab nebula is about $M = -3$. This means that if a supernova in one of our neighboring extragalactic nebulae tends towards a final state similar to that of the Crab nebula, the 100-inch telescope and we hope especially the 200-inch telescope, will furnish us with important information concerning the physical properties of this final state. Among the eleven supernovae which I have found since 1936, the supernova in IC 4182 (August, 1937) is the closest, its distance being about 3×10^6 light years. This supernova had a maximum absolute brightness $M = -16.6$. At the present time it is still in the reach of the 100-inch telescope, and has an absolute brightness $M = -3.5$ which is only slightly superior to that of the Crab nebula. The supernova during the last few months has been declining in brightness exceedingly slowly. It is, however, at the present impossible to say whether or not this decline will result in a final object still bright enough to be seen by the 100-inch or the 200-inch telescope. As Minkowski has remarked, the spectroscopic investigation of the end products of extragalactic supernovae, if they are similar to the Crab nebula, will be enormously facilitated because most of the light will be concentrated in a few emission lines.

In this connection attention should be drawn to the remarkable fact that, as the investigations by N. U. Mayall[7] have shown, the absolute brightness of a great number of extragalactic nebulae in the single line λ3727 (OII) is very considerable. It will be worth while to examine the possibility that many of the gaseous masses which produce such emission lines have been ejected in past supernova outbursts and are at the present still illuminated by the very hot and dense stellar remnants of these supernovae.

About the Cygnus loop the information available at the present time is much more scanty. Worst of all, there exist no records of any nova in historical times, which might be associated with the Cygnus loop. On the other hand this nubulosity presumably lies considerably closer to us than the Crab nebula.

According to Hubble (Yearbook of the Mt. Wilson Observatory 1936/37) the diameter of the Cygnus loop at the present expands at an average rate of about 0.06 second of arc per year. Since the diameter is about 4.7 degrees of arc, the date of the hypothetical nova outburst which caused the ejection of the gases in the Cygnus loop would be some 160,000 years ago. However, it is not off-hand certain that the nebula over this long period of time has expanded at a constant rate. The gaseous shells which were ejected from the nova at an initially much greater speed than the present speed may have been much slowed down by sharing a good part of their kinetic energy with the interstellar gases which originally surrounded the nova. It is quite possible that the mass of these gases considerably surpasses the mass of the matter ejected from the nova itself.

[5] H. Zanstra, Astrophys. J. **65**, 50 (1927).
[6] F. Zwicky, Proc. Nat. Acad. Sci. **22**, 557 (1936).
[7] N. U. Mayall and L. H. Allen, Publ. Astr. Soc. Pac. **51**, 112 (1939).

The actual diameter of the Cygnus loop, to give a very rough tentative estimate, may be of the general order of $y=10$ light years, and its volume about 5×10^{56} cm^3. With a possible density of $\rho=10^{-25}$ g/cm^3 the mass of the interstellar gases included in this volume would have been of the order of $\mu=5\times10^{31}$ g which presumably is greater than the mass ejected from novae, with the exception, perhaps, of very bright supernovae. Passing through a distance $y/2$ of interstellar gases of the given density ρ, every atom ejected from the nova would have suffered about ten collisions resulting in a considerable loss of its initial kinetic energy. Nevertheless, unless the density of the original interstellar gases was much greater than 10^{-25} g/cm^3 the hypothetical nova of the Cygnus loop presumably flared up more than 10,000 years ago. A program is under way to learn as much as possible about this interesting case.

Attempts to locate the end product of Tycho's nova of 1572 have been described by Baade.[4] A blue star of apparent magnitude 17 found by Baade may possibly be the remnant of the former nova. In contradistinction to the Crab nebula there is no obvious nebulosity which at the present might reasonably be associated with the location in which according to Tycho Brahe the nova of 1572 made its appearance.

The preceding remarks serve to illustrate that much work remains to be done to determine the properties of the end products of supernovae. This problem, although difficult, is by no means hopeless.

(5) Some values of the *velocity* of *expansion* of the gaseous shells ejected from common novae are listed in Table I. These velocities range from about 100 km/sec. to almost 5000 km/sec. Usually in common novae several shells or jets may be distinguished moving with different velocities. Also, during the outburst acceleration or deceleration of some of the shells may be observed. R. Minkowski,[8] F. J. M. Stratton[9] and others have pointed out that the interpretation of these effects presents us with some puzzling problems. In particular, as Minkowski has shown, it is not possible to understand the observed decelerations as being caused by the gravitational field of the central star as H. N. Russell[10] has suggested. Minkowski calls attention to the possibility that *magnetic fields* are involved and that the radial decelerations observed are the result of the bending of the orbits of charged particles in a magnetic field. There exists, however, still another interesting possibility. As I have shown recently[11] the fission of a star because of a nova outburst results in the separation of considerable electric charges and the building up of enormous differences in electric potential. Charged gaseous masses which escape from an oppositely charged stellar remnant are subject to deceleration effects during their radial expansion. Observations on these decelerating effects may therefore ultimately enable us to determine the differences in electric potentials which are built up during the violent fission of stars. These differences in electric potential in supernovae are large enough to generate *cosmic rays*.[11] As Professor Stratton has suggested, it will be particularly important to observe novae (and supernovae) in as early stages as possible, since the deceleration seems to be particularly pronounced in these stages.

The spectra of supernovae so far have defied all attempts at a definitive interpretation. It is consequently not possible to say what values the velocity of expansion of the ejected gases assumes in the case of supernovae. From the widths of some of the characteristic features in the spectra of supernovae one might be tempted to conclude that the velocity of expansion of the ejected gases in supernovae is not superior to that of bright common novae. Such a conclusion nevertheless is premature until a satisfactory explanation of the progressive red shift in the spectra of supernovae, discovered recently by Minkowski,[1] is available. Although at the present, the theory is confronted with the choice of several interpretations (decreasing Doppler shift, gravitational redshift), these interpretations all lead necessarily to the result[11, 12] that the *initial* velocity $\bar{v}$ of ejection of gases in supernovae lies in the range 15,000 km/sec. $<\bar{v}<100{,}000$ km/sec. A final decision of these important questions could prob-

[8] R. Minkowski, Astrophys. J. **85**, 18 (1937).
[9] F. J. M. Stratton, Observatory **59**, 325 (1936).
[10] H. N. Russell, Publ. Astr. Soc. Pac. **48**, 13 (1936).
[11] F. Zwicky, Phys. Rev. **55**, 986 (1939); Proc. Nat. Acad. Sci. **25**, 338 (1939).
[12] F. Zwicky, Phys. Rev. **55**, 726 (1939).

ably be reached through the spectral investigation of "nearby" supernovae in sufficiently early pre-maximum stages.

(6) The *lifetime* of novae and of supernovae is subject to various definitions, two of which may be found useful. In the first place we are interested in the interval $\tau_{\delta m}$ during which the nova exists as an object brighter than $m_{\max}+\delta m$. For practical purposes it is advisable to choose $\delta m=2$ or 3. Secondly the *total lifetime* τ of a nova outburst can be considered. This lifetime comprises the interval from the outburst to the time when the end product of the nova has settled down in a stationary physical state.

D. Relations Among the Integral Parameters of Novae[13]

As already mentioned little comfort can be derived from the present observational material with respect to the problem of how many independent characteristic integral parameters P_i must be ascribed to novae and how the remaining not independent parameters P_k' may be expressed in terms of the independent P_i's. Some hints about the solution of this problem can be derived from theoretical considerations about nova outbursts.[13] In the absence of any detailed knowledge one must for the present be content with the investigation of models of a very rudimentary type. For instance, one may, as a first approximation, picture a nova outburst in the following manner.

Assume that through nuclear reactions or through excessive gravitational contraction the total energy $\mathcal{E}_t$ is *instantaneously* liberated throughout some stationary star of radius R_0. The sudden increase in temperature will cause an increase of the total thermal pressure (gas pressure plus radiation pressure) which tends to blow a shell of mass M_s radially away from the star. If $\mathcal{E}_t$ is sufficiently large compared with the initial gravitational energy of M_s relative to the star, the mass M_s will be blown away for good. The radiation from the central star in passing through the ejected shell will be transformed into a quasi-black radiation as long as the shell has an optical thickness greater than the critical thickness of μ g/cm² which is just sufficient to make the shell opaque. The maximum luminosity will be reached when the shell has expanded to a stage when its thickness is about μ g/cm². On the basis of this simple model one may derive the following relations:[13]

(*a*) The expansion-luminosity relation is

$$L_{\max}=3\pi R_0 c\mu\bar{v}^2/2, \tag{10}$$

which expresses the energy $L_{\max}$ radiated by the shell (at maximum brilliance) per unit time in terms of previously defined integral parameters of a nova. Introducing absolute magnitudes M instead of luminosities, we have

$$M=-2.5\log_{10}L+M_0. \tag{11}$$

If L is expressed in ergs per second, the constant M_0 is approximately equal to $M_0=88.8$, since for the sun we have $M\cong 4.8$ and $L=3.8\times10^{33}$ ergs/sec. Substituting (10) in (11) we obtain

$$M_{\max}=-5\log_{10}\bar{v}+M_1, \tag{12}$$

where

$$M_1=88.8-2.5\log 3\pi R_0\mu c/2. \tag{12a}$$

(*b*) The energy-expansion relation[13] is

$$E_{\text{vis}}=\gamma\bar{v}^3, \tag{13}$$

where γ is independent of the average velocity of expansion $\bar{v}$, but may still depend on some of the other integral parameters. E_{vis} is the total energy radiated away during the outburst in the form of visible light.

(*c*) The life-luminosity relation[13] is

$$M_{\max}=-5\log_{10}\tau_{\delta m}+M_2, \tag{14}$$

where M_2 is independent of $\tau_{\delta m}$.

In Fig. 1 we indicate the range for $\bar{v}$ and $M_{\max}$ to be expected for novae and for supernovae according to the expansion-luminosity relation (12). The triangular point marks the position of Nova Aquilae (1918) in the diagram. With the values $\bar{v}=1.7\times10^8$ cm sec.⁻¹ and $M_{\max}=-8.9$ for this nova we obtain

$$M_1=32.35 \tag{15}$$

and

$$R_0\mu=2.7\times10^{11}\text{ g cm}^{-1}. \tag{16}$$

The straight line through the point determined by Nova Aquilae is analytically given by

$$M_{\max}=-5\log_{10}\bar{v}+32.35. \tag{17}$$

[13] F. Zwicky, Proc. Nat. Acad. Sci. **22**, 457 (1936); Publ. Astr. Soc. Pac. **48**, 191 (1936).

For different novae we must, because of different initial radii R_0 and, to a lesser degree perhaps also because of different chemical composition and degree of ionization admit different values of $R_0\mu$. The range of these values is presumably not greater than given by the inequality

$$2.7\times10^{10}\text{ g cm}^{-1}<R_0\mu<2.7\times10^{12}\text{ g cm}^{-1}. \quad (18)$$

This estimate is based on the fact, that the original stars which were associated with nova and supernova outbursts, in all well-known cases were only of average brightness. The range in the initial radii of such stars is probably adequately described by a factor hundred. The two straight lines (I, II) which correspond to the upper and lower values of $R_0\mu$ in (18) are also drawn in. If (18) describes the correct range for $R_0\mu$, the *initial* velocities of expansion $\bar{v}$ for an average supernova with $M_{\max}=-14.3$ would be determined by points on the vertical shaded line. Numerically we get

$$6600\text{ km/sec.}<\bar{v}<66{,}000\text{ km/sec.}, \quad (19)$$

a range which in the light of the present data[13] seems reasonable.

We shall not here go further into the possible tests of our three relations, since it is hoped that in the near future more accurate data will become available. Attention should be drawn to the fact that already on our much simplified model not only two, but several integral parameters enter. The space in which the whole multitude of novae must be represented is therefore necessarily more than two-dimensional. This is especially true because, for actual novae, we cannot expect the assumption of an instantaneous liberation of the total energy $\mathcal{E}_t$ to be strictly correct. In agreement with these considerations the available data on ordinary novae indicate a very complex state of affairs. In supernovae physical conditions definitely converge towards greater uniformity. The supernovae so far investigated exhibit far less individuality than the ordinary novae inasmuch as their light curves and their spectra show only insignificant differences from one object to the other. This lack of individuality suggests strongly that the original stars of supernovae and the processes causing the outbursts possess very definite physical and chemical characteristics, a fact which, if correct, should prove very helpful to the theory of supernovae.

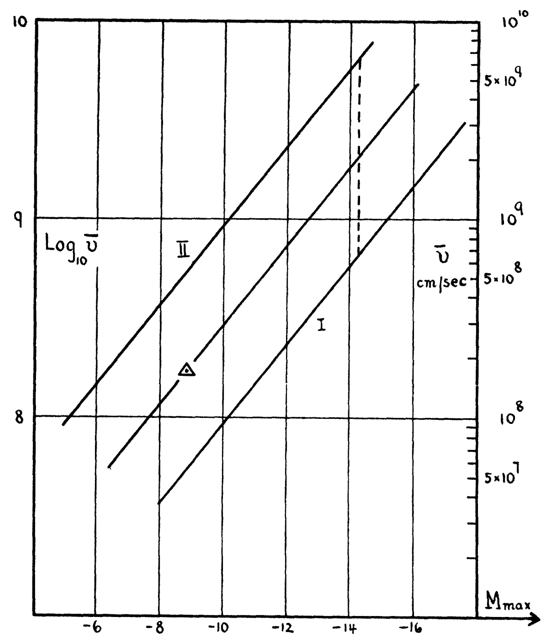

FIG. 1. Expansion-luminosity relation.

The possibility of different values of $R_0\mu$ for different objects suggests the existence of novae of relatively high luminosity and low velocity of expansion, and *vice versa*. Novae corresponding to both extremes, might be called *quiescent novae* (low velocities and high luminosity) and *dynamic novae*, respectively (high kinetic energy of the atomic rays and low luminosity). Nova Pictoris (1925) and Nova Lacertae (1936) are presumably cases of a quiescent nova and a dynamic nova.

E. Dependence of the Frequency of Occurrence on the Absolute Luminosity $M_{\max}$

According to the preceding considerations we must expect that in the theory of ordinary novae a considerable number of independent integral parameters P_i is involved, whereas for supernovae this number is smaller. In fact, it now does not seem impossible that the behavior of supernovae may be described in terms of as few as one or two independent integral parameters.

Concerning the frequency ν of occurrence of novae and of supernovae the important problem at once suggests itself of how this frequency ν

depends on the independent integral parameters $P_1, P_2 \cdots P_z$. Unfortunately the data available are much too scanty to establish the mathematical character of the frequency distribution function $\nu(P_1, P_2 \cdots P_z)$. It is, however, possible to establish a pretty good approximation for the frequency function n expressed in terms of the absolute magnitude $M_{\max}$ alone. Mathematically this function is

$$n(M_{\max}) = \int \cdots \times \int \nu[M_{\max}, P_2, \cdots P_z] dP_2 \cdots dP_z, \quad (19)$$

where the integration has been carried out over the whole range of variability of the parameters P_2 to P_z. Through recent investigations on the frequency of novae and of supernovae in extragalactic systems it has been possible to establish the important fact that the function $n(M_{\max})$ possesses *at least two maxima* which correspond to the mean absolute magnitude $\bar{M}^c{}_{\max}$ and $\bar{M}^s{}_{\max}$ of common novae and of supernovae, respectively. According to the analysis given by Baade[4] the numerical values involved are

$$M_{0c} = \bar{M}^c{}_{\max} \cong -7.0 \quad (20)$$

and

$$M_{0s} = \bar{M}^s{}_{\max} \cong -14.3. \quad (21)$$

Baade has also determined the values of the dispersion in magnitudes. We define this dispersion σ as

$$\sigma = [\langle (M - M_0)^2 \rangle_{\text{Av}}]^{\frac{1}{2}}, \quad (22)$$

which is the square root of the averaged square of the deviation in absolute magnitude. According to Baade[4] the dispersion observed for common novae is

$$\sigma_c \cong 0.53 \text{ magnitude.} \quad (23)$$

For supernovae Baade[4] concluded from the observations which are available so far, that

$$\sigma_s \lessapprox 1.1 \text{ magnitudes.} \quad (24)$$

Concerning the total number N_c of common novae and the total number N_s of supernovae in an average nebula during a given period we may make only a very approximate estimate. The range.

$$1000 < N_c/N_s < 10{,}000, \quad (25)$$

which is based on recent counts[4] of novae and of supernovae in extragalactic nebulae should not be very far from the truth. To arrive at a graphic representation of the function $n(M_{\max})$ we approximate this function by two Gaussian error curves having the proper maxima and the proper dispersions. We therefore write for the common novae

$$n_c(M) = [N_c/\sigma_c(2\pi)^{\frac{1}{2}}] \times \exp[-(M - M_{0c})^2/2\sigma_c^2], \quad (26)$$

where $n_c(M)dM$ is the number of novae which during a given time appear in an average nebula in the interval dM of absolute magnitudes at maximum brightness. It is

$$N_c = \int_{-\infty}^{+\infty} n(M)dM \quad (27)$$

and

$$\sigma_c^2 = \langle (M - M_{0c})^2 \rangle_{\text{Av}} \quad (28)$$

in agreement with the definition (22) of the dispersion σ. Relations (26), (27), (28) also apply to supernovae after we replace the index c by the index s. Since both of the dispersions σ_c and σ_s are small compared with the interval $M_{0c} - M_{0s}$, we may now write for the frequency function $n(M)$ which covers the whole range of magnitudes

$$n(M) = n_c(M) + n_s(M). \quad (29)$$

Since the numerical value of N_c/N_s is very large it is inconvenient in the graphical representation to plot the values of n_c and n_s themselves. We therefore introduce the logarithms and we obtain

$$\log n_c(M) = \log N_c - \log \sigma_c - \tfrac{1}{2} \log 2\pi - 0.217(M - M_{0c})^2/\sigma_c^2, \quad (30)$$

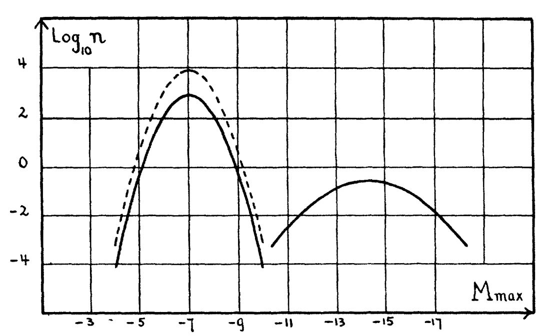

FIG. 2. Frequency $n(M_{\max})$ of novae per average nebula per six hundred years in dependence of their absolute magnitude at maximum brightness.

where the logarithms are to the base ten. A similar relation holds for $\log_{10} n_s$. From my search of supernovae I have obtained the value $N_s=1$ supernova per average nebula per six hundred years.[4] Adopting the interval of *six hundred years* for our counts we therefore have $\log_{10} N_s=0$, and

$$\log_{10} n_s(M)=-0.44-0.18(M-M_{0s})^2. \quad (31)$$

For the two extreme values of N_c given by (25) it is

$$\text{I. } \log_{10} n_c(M)=2.88-0.772(M-M_{0c})^2 \quad (32)$$

and

$$\text{II. } \log_{10} n_c(M)=3.88-0.772(M-M_{0c})^2. \quad (33)$$

In Fig. 2 the expressions (31) and (32) are indicated by the solid lines whereas (33) is drawn as a shaded line. From the figure it is evident that for absolute magnitudes $M_{max}=-10$, -11, the frequency of novae is very low. Actually no nova has ever been observed in the interval $-10>M_{max}>-12$. Such novae, if they had occurred, would have easily been detected in my search for new objects in extragalactic nebulae. The deep minimum at $M=-11$ is therefore real, although we cannot say at the present just how deep it is. The actual shape of the function $n(M)$ in the neighborhood of this minimum can be determined only through the discovery of many more novae than are known at the present. The reality of the minimum, however, must now be regarded as being quite certain and its existence is theoretically of great importance. From the existence of this minimum Baade and I several years ago concluded[3] that supernovae form a class distinct from the common novae. The investigations of the light curves and of the spectra of supernovae have in the meantime furnished additional proof for the correctness of this conclusion. It follows that supernovae and common novae are not due to the same types of physical causes taking place in initially bright stars and faint stars, respectively, as has been suggested by some writers.[14] Also we may conclude that novae and supernovae are not generally caused by stellar collisions as has been proposed by other writers. The question of stellar collisions will be discussed in more detail in Section H.

[14] G. Gamow, Phys. Rev. **54**, 480 (1938).

F. The Spatial Distribution of Supernovae

It is of great interest to investigate the distribution of supernovae in extragalactic space as compared with the distribution of ordinary luminous stars. As we know, these stars are concentrated in stellar systems or nebulae. The surface brightness $I(r)$ of nebulae in points of varying distance r from the center of these nebulae may be taken as a fair indication of the average distribution of luminous stars. We are therefore interested in the frequency of occurrence $f(I)$ of supernovae as a function of the surface brightness associated with extragalactic nebulae.

Let us consider briefly the forms of the function $f(I)$ which one might expect on the basis of some simple assumptions regarding possible types of the original stars in which the outbursts take place.

(*a*) Suppose that the original stars of supernovae are of a type which shows the same spatial distribution as the average luminous star. In this case we should expect

$$f(I)=\alpha I, \quad (34)$$

where $f(I)$ is the total number of supernovae in a given interval of time per unit area in a region of the sky of surface brightness I. This surface brightness I refers of course to extragalactic light sources only. "Local" lights, such as the light of the night sky coming from our own atmosphere, from dust in the planetary system and from the stars in our own Milky Way must be eliminated in the determination of I.

For globular and elliptical nebulae the surface brightness as a function of the distance r from the center of the nebula is approximately represented by

$$I=I_0/(1+r/a)^2. \quad (35)$$

In the ring between r and $r+dr$ we should therefore expect a number of supernovae equal to

$$Fdr=f(I)2\pi rdr=2\pi\alpha I_0 rdr/(1+r/a)^2 \quad (36)$$

or for regions $r>a$ (outside of the main body of the nebula)

$$F\cong 2\pi\alpha I_0 a^2 dr/r. \quad (37)$$

Notice, that if (35) really is correct to very great distances R from the center of nebulae the total number of supernovae in the *intergalactic* space between the nebulae may be considerable. Since, at a certain critical distance r_c from the center of a nebula the night sky light becomes so much stronger than the light from the nebula that the latter is entirely drowned out and unobservable, the discovery of supernovae in "empty" space may furnish the first evidence for the existence of faint extensions of nebulae which reach far into intergalactic space or which perhaps even entirely fill it.

(*b*) If collisions between luminous stars were the cause of supernovae we should expect that approximately

$$f(I)=\beta I^x, \qquad \text{where} \qquad x \geqslant 2, \tag{38}$$

as can be derived from the fact that the total number of collisions per unit volume is proportional to Z^2 where Z is the number of luminous stars per unit volume, and I is proportional to the number of stars integrated along the line of sight over a prism of unit cross section. With increasing r we should get a very rapid decrease of the total number of supernovae, namely

$$F \sim dr/r^{2x-1}; \tag{39}$$

that is, F should decrease more rapidly than r^{-3}. As we shall see, this is conspicuously contradicted, by the facts. Collisions as a general cause for supernovae are consequently ruled out.

(*c*) Supernova outbursts may take place in stars which show a spatial distribution entirely different from the average type of luminous stars. We know indeed that there are types of stars which exhibit a greater tendency towards concentration in the nebulae than does the average luminous star, and also that there are other stars which show a greater independence from the nebulae. Around our own Milky Way these latter stars form Baade's *Globular Substratum*. Baade, in a long series of investigations[15] has proved conclusively that for instance the cluster type variables do not exhibit any marked concentration towards the plane of the Milky Way. In contradistinction to the distribution of most other stellar types the number of cluster type variables per unit solid angle is relatively little dependent on galactic latitude. Also in the direction normal to the Milky Way plane these stars are to be found at very considerable distances. The most distant cluster type variable discovered by Baade[16] is located at a distance of 40,000 parsecs from the sun which is about five times the distance of the center of the Milky Way from the sun.

Interestingly enough, the actual distribution so far observed indicates for supernovae the same lack of tendency to concentrate towards the centers and the main bodies of the extragalactic nebulae as it is exhibited by the stars which constitute Baade's globular substratum in which our own glaxay is imbedded. I am at the present engaged in determining the surface brightness in all of the points where supernovae have been discovered. The function $f(I)$ will then be obtained from the relation

$$N(I)=f(I)A(I), \tag{40}$$

where $N(I)dI$ is the total number of supernovae discovered in regions whose surface brightness lies between I and $I+dI$ and whose total area is $A(I)dI$. I can already state, that the regions of low surface brightness I are vastly favored through the appearance of supernovae as compared with, for instance, the expectation expressed in the relation (37), not to speak about the relation (38) which would demand an exceedingly heavy concentration of supernovae towards the center of nebulae.

Two outstanding examples of supernovae which have appeared in regions of low surface brightness are shown in Fig. 3. In fact the surface brightness in the points where these novae appeared is so low that on the photographic plate it does not produce any contrast against the sky background. For the supernova in NGC 1482 this is also shown by the photometer tracing (Fig. 4) which was run along the straight line L indicated by its end-points in the photograph b, Fig. 3.

In the light of the preceding considerations the organization of the search for supernovae is now confronted with a new problem. So far, I, and no doubt all other observers have concentrated

[15] W. Baade, Mitteilungen d. Hamburger Sternwarte **7**, (No. 36), 28 (1932).

[16] W. Baade, Publ. Astr. Soc. Pac **48**, 277 (1936).

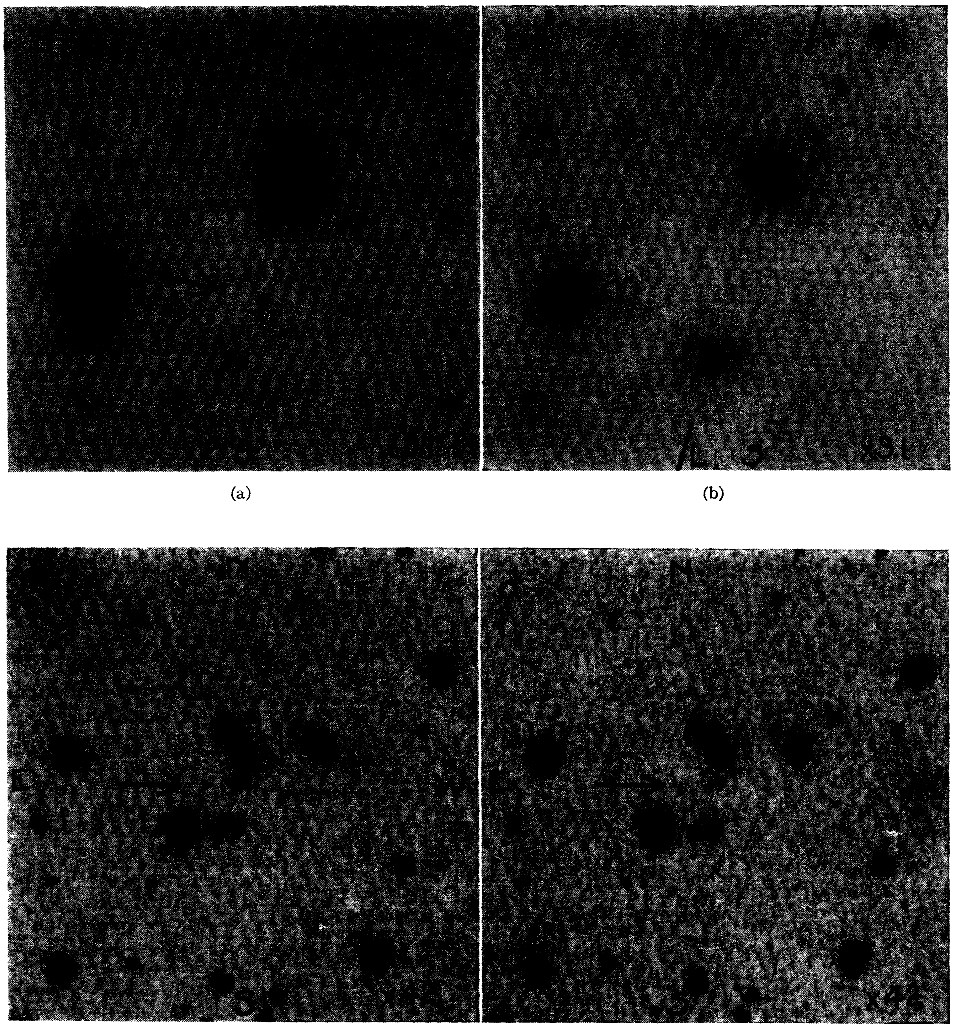

FIG. 3. (a) Shows NGC 1482 [R.A. 3^h 51.9^m, Decl. $-20°41'$ (1938)] with supernova (arrow) about a month after maximum brightness $m_{max} \leqq 14$. The photograph, which was taken by Hubble on February 5, 1938 with the 100-inch telescope on the Mt. Wilson Observatory is of very poor quality due to poor seeing conditions. In addition imperfections of the aluminum coat of the secondary mirror caused the elliptical halos around the bright stars.

(b) Good photograph of NGC 1482, taken by Baade on October 25, 1938. The supernova has become fainter than the apparent photographic magnitude $m=21$ and is entirely invisible on the photograph. The photometer tracing shown in Fig. 4 was made along the line L indicated in (b).

(c) Shows the barred spiral R.A. 2^h 34.0^m, Decl. $+34°$ $10'$ (1938.0) with a supernova (arrow) of the apparent photographic magnitude $m=15.3$. The photograph is a strongly enlarged section of a film obtained with the 18-inch Schmidt telescope on November 17, 1938, a few days after the supernova had reached maximum brightness. The nebula belongs to the Pisces-Perseus cloud of nebulae the distance of which is of the order of 25 million light years. The supernova shown is therefore the *most distant* star ever photographed whose light curve and spectrum were observed and whose distance could be determined. According to a private communication by R. Minkowski the apparent velocity of recession of the above nebula is of the order of 4500 km/sec. which is in agreement with the distance given.

(d) The same field as (c) photographed with the Schmidt telescope on January 17, 1939. The supernova (arrow) has become much fainter, its apparent photographic magnitude being $m=18.2$. Both of the enlargements (*c*) and (*d*) of pictures taken with the 18-inch Schmidt telescope, well illustrate the good optical performance of this instrument. In spite of the fact that the sections shown lie at the edge of the field over four degrees of arc from its center, no systematic distortions of the star images can be detected. The image of the supernova in (d) can clearly be seen on the original film in spite of its faintness ($m=18.2$).

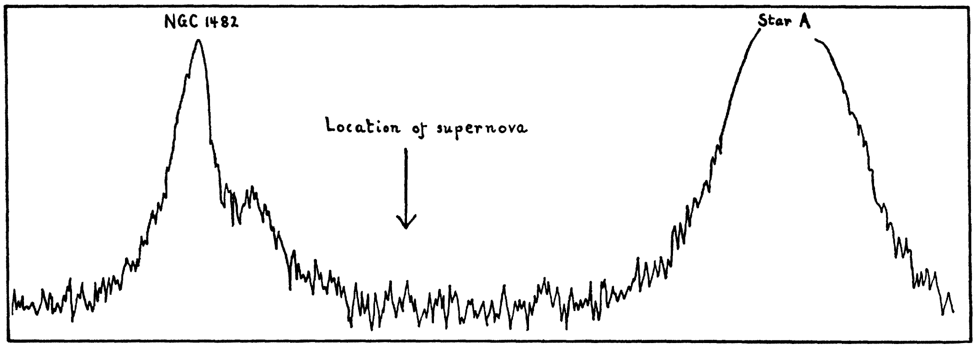

FIG. 4. Photometer tracing along the line L of the 100-inch plate reproduced in (b) of Fig. 3. Distances are twenty times those on the original such that the actual distance between the peak of NGC 1482 and the star A is 170 seconds of arc. The supernova is located in a point where light from the nebula does not, on the photographic plate, produce any blackening distinguishable from the sky background.

our efforts on the examination of the nebulae proper. In my own search, because of the lack of a blink machine, it proved wholly impossible to search on every film the entire field photographed with the Schmidt telescope. Occasionally a new star would be found far out in the "empty" space between the nebulae; but usually no time was available to check up on it by further observations and it was dismissed as an asteroid of small motion. In the future it will be imperative to be more careful and it is intended to check up on all asteroids by taking of each field double photographs separated by an interval of perhaps one or two hours.

The occurrence of supernovae as far away from the main body of nebulae as the objects shown in Fig. 3 suggests that the total number of supernovae which appear in the so-called "empty" intergalactic space may be considerable. Through the discovery of supernovae in interaglactic space we may indirectly get some information about the stellar population of the "empty" parts of space. This is of importance because of the fact that ordinary permanent stars which might thinly populate intergalactic space are not bright enough to be reached by present day telescopes at any great distances. And even at distances which are small compared with the average separation of nebulae it will be difficult to determine the distance of conjectured intergalactic stars with any degree of accuracy. For supernovae, because of their great intrinsic brightness combined with our knowledge of the uniformity of their physical properties the difficulties just mentioned almost entirely vanish. Also, as Baade repeatedly has pointed out in private discussions, the great apparent brightness of not too distant supernovae will eventually enable us to observe their spectra with spectrographs of high dispersion. The analysis of these spectra for interstellar *and intergalactic* absorption lines and for reddening effects promises to result in valuable information about the distribution of gases and of finely divided dust over large parts of interstellar and intergalactic space.

G. Distribution of Supernovae Over Different Types of Nebulae

The eleven supernovae which I have found so far were distributed over various types of nebulae as indicated in Table II. The last column contains the percentages of the different types of nebulae in the general field. The figures given represent an unpublished revision of the frequency of different nebular types for the use of which I am indebted to Dr. E. Hubble of the Mt. Wilson Observatory. The collection of nebulae which I have systematically searched for supernovae probably contains various types of nebulae in proportions very closely given by the percentages listed under the general field. From the data available so far supernovae therefore do not seem to favor to any high degree any particular type of nebula. Caution is still indicated, since the number of supernovae available is very small. Further investigations about the frequency of supernovae in different types of nebulae are desirable in order to correlate the appearance of supernovae with the abundance of different stellar types in different stellar systems.

H. Supernovae Are Not Caused by Collisions Between Luminous Stars

The opinion has often been expressed that novae or supernovae might be caused by either stellar collisions or collisions of stars with dense gas or dust clouds.[17] There is, however, no evidence to be derived from the present data which could elevate such an hypothesis beyond the state of wishful thinking. The following facts make it difficult to postulate collisions as a possible cause for supernovae.

(*a*) The existence of two maxima in the frequency curve $n(M_{\max})$ speaks against collisions. Since encounters between two or more stars become the more frequent the greater the distance of the closest approach is, we should expect that the number of novae due to this cause would increase monotonously with decreasing generation of excess energy in the stars involved, and $n(M_{\max})$ should therefore be a monotonously decreasing function of $M_{\max}$.

(*b*) On the collision theory the number of supernovae should increase rapidly towards the central parts of nebulae, whereas actually the contrary is true.

(*c*) The light curves of various supernovae differ but very little from one another. This is especially true for the light curves of supernovae of the same absolute brightness. Such uniformity of the light curves extending over periods of several hundred days is inconceivable if the outbursts were caused by the collision within nebulae of any two stars chosen at random.

(*d*) Also, since collisions will take place in a random way between stars of all types, it would be impossible to understand the remarkable similarity of the spectra of different supernovae. It must here again be emphasized that this similarity does not refer to one stage of the development only, but that it holds for the whole sequence of events covering at least two hundred days. How accurately the progressive stages of a supernova outburst become manifest in the spectrum may be inferred from the fact that for the supernova in NGC 4621 Minkowski was able from a single spectrum to determine the date of maximum luminosity with surprising accuracy (see Section I 5).

The objections (*b*), (*c*) and (*d*) against the collision theory could be partly met by the assumption that, of the two stars which take part in a collision, at least one must be of a prescribed type in order that a supernova outburst be produced. Such a possibility is of course imaginable, since for a supernova outburst to take place the disturbance caused by a collision, plus an initial lack of great stability of one of the component stars may be required. The supernova outburst might thus be due to a trigger process which is set off by the collision. On this assumption the distribution of supernovae in nebulae would be the same as the distribution of the stars of the special required type, and objection (*b*) might in this way be met. The objection, (*a*), (*c*) and (*d*), however, would only partially have been removed. In addition, on our present knowledge of the total stellar content of nebulae it would be quite impossible to account for the frequency of supernovae if a *selected* type of star were required to take part in the decisive collisions.

Finally it should be remarked that, in itself, the kinetic energy available in a collision is not sufficient to account for the energy of at least 10^{49} ergs which is liberated in a supernova outburst, except in the impossibly rare cases of head-on collisions between very massive stars which, in addition, must move with unusually high relative velocities. A head-on inelastic collision of two stars of the mass of the sun (2×10^{33} g) which move with the extremely high relative velocity of 200 km/sec. would liberate the kinetic energy of only 2×10^{47} ergs. From this it follows that the frequency of collisions which could make available enough kinetic energy to

Table II. *Distribution of supernovae.*

	Number of Supernovae	General Field
Spirals	7	65.4 percent
Barred spirals	2	14.5
Elliptic	2	17.5
Irregular	0	2.6

[17] See for instance F. L. Whipple, Proc. Nat. Acad. Sci. **25**, 118 (1939).

account for the energy liberated in a supernova outburst is entirely insufficient to account for the observed frequency of occurrence of supernovae.[4] The considerations just given are, however, not in themselves decisive in eliminating the collision theory. It is, for instance, possible that the collision serves as a trigger action. An initial generation of excess heat by the impact might start a self-perpetuating chain of nuclear reactions which would generate enough energy to explain the supernova outburst. Although the difficulty of accounting for the liberation of energy in a supernova can in this way perhaps be avoided, we hold that the arguments advanced at the beginning of this section exclude collisions between luminous stars as an origin of supernovae.

I. Review of the Reasons Which Resulted in the Recognition of Supernovae as a Separate Class of Temporary Stars

Some of the brighter novae in extragalactic nebulae which were observed a long time ago, such as S-Andromedae (1885) and Z-Centauri (1895), as well as some of the novae in the Virgo cluster of nebulae are now quite definitely to be classed as supernovae. In view of the long acquaintance of astronomers with at least half a dozen novae of this type it is perhaps somewhat surprising that convincing proof for the existence of a separate and distinct class of supernovae was adduced during the last few years only. It may hence be useful to recall some of the difficulties which stood in the way of an earlier recognition of the true nature of supernovae. The major difficulties involved are related to the following circumstances.

(1) First it had to be shown that the new stars in question are physically related to the nebulae in which they appear. There exists, of course, the possibility that variable stars which belong to our own system may lie in the line of sight of nebulae. This possibility becomes rather acute for extended nebulae, such as the Andromeda nebula, for which the probability of foreground variable stars to be projected on the image of the nebula becomes especially great because of its location in a rich stellar region close to the Milky Way. Z-Centauri, on the other hand, flared up in an irregular nebula (NGC 5253) whose spectrum consisted of bright emission lines and therefore resembled the emission spectrum of a gaseous nebula rather than the absorption spectrum of a system composed of stars. Actually the strongest and perhaps up to the present only acceptable reason for the assumption that NGC 5253 is an extragalactic system lies in the fact that Z-Centauri, because of its own characteristics such as its light curve and its spectrum, could definitely be identified as a supernova. But even this evidence does not quite preclude the possibility that NGC 5253 is a gaseous nebula, galactic or extragalactic, and not a stellar system. Z-Centauri then would have been a supernova in intergalactic "empty" space (see Section F) accidentally lying in the line of sight, in front or behind NGC 5253.

From the preceding discussion it becomes understandable that the physical association of the apparently temporary stars with the nebulae in which they flared up could not be admitted without much more thorough investigations. In fact, even up to now, no satisfactory *quantitative* discussion of the probability has ever been given that variable stars fall on or near the images of a given collection of extragalactic nebulae, such as the collection which is continually examined in my own search for supernovae. In this search I have actually found about as many previously unknown galactic variable stars located in or near the line of sight of extragalactic nebulae as I have found supernovae.

(2) The recognition of supernovae as a distinct class of temporary stars was of course closely tied up with the identification of spiral nebulae as extragalactic stellar systems similar in dimensions and stellar content to our own Milky Way. How intimately the failure to visualize the existence of supernovae is to be linked to the initial difficulties of recognizing the true nature of extragalactic nebulae is best illustrated through a glimpse into the astronomical literature of two decades ago. Actually the contention that stars a hundred million times as bright as the sun could not possibly exist greatly strengthened some astronomers* in their conviction that spiral

* We quote here some passages from H. Shapley, Publ. Astr. Soc. Pac. **31**, 266 (1919). He writes "Moreover, if in real dimensions spiral nebulae were analogous to our galactic system, the absolute magnitude of the novae in spirals would far transcend any luminosity with which we

nebulae could not be "island universes" or stellar systems comparable to our own galaxy.

From these considerations it follows that the problem of the physical nature and of the distances of the spiral nebulae had to be solved first, before the question could sensibly be answered of just how bright the most luminous temporary stars can become at maximum. By 1930, through the extensive work of a number of astronomers, notably Hubble, the problem of the nature and of the distances of spiral nebulae could be considered as solved in its most important aspects.

(3) The final classification of novae into the two distinct classes of common novae and of supernovae required the knowledge of sufficient data regarding the frequency $n(M)$ of novae and of supernovae in dependence of their absolute magnitude at maximum. In order to establish the function $n(M)$ it is necessary to know quite accurately not only the distances of enough individual nebulae in which supernovae have flared up, but one also needs data on the apparent luminosities at maximum brightness of a sufficient number of supernovae. When it is recalled, that even today after an extensive and successful search for supernovae we know the distances of only about fifteen of them and the actual maxima in brightness have been observed in not more than six cases, it becomes clear why the proof for the division into the two magnitude classes of common novae and of supernovae could only be given during the last few years.

In this connection attention should also be drawn to the curious fact that the analysis of the function $n(M)$ even near $M_{\max}=-7$ required the study of *common novae in extragalactic systems.* We do not possess enough data about the distances and the absolute magnitudes of galactic novae to derive from them the mean absolute magnitude and the dispersion in magnitude. The values used in plotting the curve of Fig. 2 were derived by Baade from a study of the common novae in the Andromeda nebula.

(4) With enough evidence at hand to show that the frequency function $n(M)$ possesses two maxima the classification of novae into common novae and supernovae appeared secure enough except for the difficulties related to our inability to distinguish between the direct and the indirect images of supernovae. How these difficulties were finally resolved through the extensive observations of the light curves and the spectra of some of the recently discovered supernovae was already discussed in Section C.

(5) The studies of the light curves and of the spectra of some of the recent supernovae have produced additional decisive evidence for the correctness of the contention that supernovae form a class of temporary stars separate from the common novae. Before the discovery of the supernovae in IC 4182, NGC 1003, 4621, 4636 etc., only some extremely fragmentary light curves were known and the few very poor spectra which

---

are acquainted, and would be at direct variance with present results on intrinsic stellar brightness. For at the distance computed above the absolute magnitude of a nova of the sixteenth apparent magnitude would be −16, nearly two hundred thousand times as bright as the novae of the galactic system for which van Maanen has determined the absolute luminosities. An upper limit to the intrinsic brightness attainable by stars is suggested by recent observational and theoretical work and this limit is much fainter than −16. The study of globular clusters, for example, has yielded sufficient knowledge of the luminosity of more than a million stars to show that not one is within ten magnitudes of this enormous brightness. . . . Hence stellar luminosities of this order [−16] seem out of the question, and accordingly the close comparability of spirals containing such novae to our galaxy appears inadmissable."

In the paper mentioned, Shapley summarizes his conclusions thus: "Observation and discussion of the radial velocities, internal motions, and distribution of spiral nebulae, of the real and apparent brightness of novae, of the maximum luminosity of galactic and cluster stars, and finally the dimensions of our own galactic system, all seem definitely to oppose the 'island universe' hypothesis of the spiral nebulae."

In his detailed discussion of "the scale of the Universe" with H. D. Curtis [Bull. Nat. Research Council, **2**, Part 3, 171 (1921)], Shapley still held the opinion expressed in the preceding quotation from his article of 1919. H. D. Curtis, on the other hand, arrived at a conclusion which in its essentials coincides with the view now taken with regard to the nature of spiral nebulae although he considerably underestimated the dimensions of our own galaxy. His final statement reads "I hold, therefore, to the belief that the galaxy is probably not more than 30,000 light years in diameter, that the spirals are not intragalactic objects but island universes, like our own galaxy, and that the spirals, as external galaxies, indicate to us a greater universe into which we may penetrate to distances of ten million to a hundred million light years." There is also a first inkling of the truth about the possible existence of very bright novae in Curtis' statement, "It seems certain, for instance, that the dispersion of the novae in the spirals, and probably also in our galaxy may reach at least ten absolute magnitudes, as is evidenced by a comparison of S-Andromedae with the faint novae found recently in this spiral. A division into two magnitude classes is not impossible."

For accounts concerning the subsequent investigations on the distances of extragalactic nebulae, their stellar content and some of their additional physical characteristics, we refer to E. Hubble, *The Realm of Nebulae* (Yale University Press, 1936); R. L. Waterfield, *A Hundred Years of Astronomy*, (Macmillan Company, New York, 1938) and other reviews.

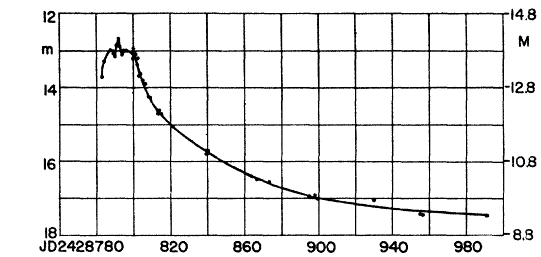

FIG. 5. Photographic light curve of supernova in NGC 1003. m and M are the apparent and the absolute magnitudes, respectively. The dates are given in Julian days. JD 2,428,780 is September 2, 1937. For a more detailed discussion of the data see W. Baade and F. Zwicky, Astrophys. J. **88**, 411 (1938).

had been obtained could not be interpreted. At the present we possess a great number of excellent spectra of about half a dozen supernovae, but still a satisfactory identification of any of their characteristic features is lacking. Nevertheless, these spectra are so similar from one object to the next that a single one of them suffices not only to distinguish a supernova from any other type of temporary, variable or permanent star, but also serves for the determination of the date on which the supernova has reached its maximum brightness. For instance, on May 19, 1939, I discovered a supernova in the elliptical nebula NGC 4621 which is a member of the Virgo cluster of nebulae, seven million light years distant. On the next night, May 20, 1939, Minkowski at the Mount Wilson Observatory obtained a spectrum of this supernova from which he concluded that the supernova at that time was already three weeks past maximum. This prediction that the maximum had occurred around May 1, 1939 was relayed to Professor H. Shapley who reprinted this information on the "Announcement Card 486" of the Harvard College Observatory. On the next card (487) Professor Shapley announced that Harvard patrol plates covered the supernova in question over the period from April 23, 1939 to May 16, 1939. From these plates it was deduced that the supernova had reached maximum brightness around May 1 or May 2, within one or two days of the predicted date. The announcement card commented on this that "The identification of the phase from the characteristics of the spectrum on Minkowski's Mount Wilson plate of May 20 as 'three weeks past maximum' is remarkably exact."

While *one good spectrum* allows us to predict the *date* of maximum brightness with considerable accuracy, *two* observations, ten or more days apart, of the *apparent magnitude* of a supernova enable us to predict *both the date and the apparent magnitude* of the supernova at maximum brightness. The light curves of different supernovae are so similar that, according to Baade, an average light curve may be constructed from which the individual light curves depart but very little. If the difference in apparent magnitude δm of the supernova for a period of say n days is determined through observations on two nights n days apart, we only have to find that region on the average standard light curve of supernovae which for the period of n days possesses the same gradient δm in apparent magnitude. Through this method of superposition, the date of maximum brightness and the corresponding apparent magnitude can immediately be determined. As an example, I mention the supernova in NGC 6946 which I found on a film taken with the Schmidt telescope on Palomar Mountain on July 17, 1939. At that date, the apparent magnitude of the supernova was about $m=14.4$. I also obtained a second picture on August 7, 1939, on which night the apparent magnitude was about $m=16.1$. On the basis of these observations I wired to Harvard the prediction that the supernova reached maximum brightness about July 6, 1939, and that its apparent magnitude at that date must have been about $m=13.4$ (Harvard announcement card 502). Again the Harvard College Observatory had at its disposal a whole series of prediscovery plates of the supernova in NGC 6946. The information obtained from these plates was communicated on the "Announcement Card 506." Actually the supernova had reached maximum on about July 7, one or two days within the predicted date, and the apparent magnitude near the maximum had fluctuated between the limits $13.1<m<13.5$, very close to the predicted $m=13.4$.

In Fig. 5, the light curve of the supernova 1937 in NGC 1003 is reproduced as a typical example. At the present time, we have at our disposal excellent light curves of about half a dozen supernovae covering periods from six months to two years in length. A more complete discussion of these light curves is to be published presently in the *Astrophysical Journal.*

K. Final Remarks

In the discussion given in the preceding it was pointed out that the data at our disposal enabled us to establish the fact that the frequency function $n(M)$ of novae in dependence of the absolute magnitude M at maximum brightness possesses two maxima at $M \cong -7$ and $M \cong -14.3$, from which fact we conclude the existence of two separate classes of novae, designated as common novae and supernovae.* It will be of interest to investigate if the function $n(M)$ can further be differentiated. It is imaginable and even probable that if some integral parameter P' different from the absolute magnitude at maximum $P_1 = M_{max}$ is chosen that the function $n'(P')$ which gives the frequency of novae in dependence on P' will display features which make a further division of novae into different classes possible. In this connection it is suggested to investigate the characteristic features of intrinsically faint novae which may help us to bridge the gap between common novae and ordinary permanent or variable stars. Also much might be learned about the causes of the explosions of stars if novae could be found to fill in the gap between the absolute magnitudes $M_{max} = -9$ and $M_{max} = -12$. On our present views we should expect in this gap objects of equal luminosity at maximum brightness, such that some of these objects display the properties of common novae whereas others behave like intrinsically faint supernovae. This expectation, if correct, would represent the most conclusive proof of the difference in the fundamental physical causes of novae and of supernovae.

I am indebted to Dr. Walter S. Adams, Director of the Mt. Wilson Observatory for the permission to reproduce two of the photographs taken with the 100-inch telescope. I also wish to thank Mr. Edison Hoge of the Mt. Wilson Observatory for the enlargements of the photographs and their arrangement in the plate.

* Baade and I first introduced the term "supernovae" in seminars and in a lecture course on astrophysics at the California Institute of Technology in 1931.

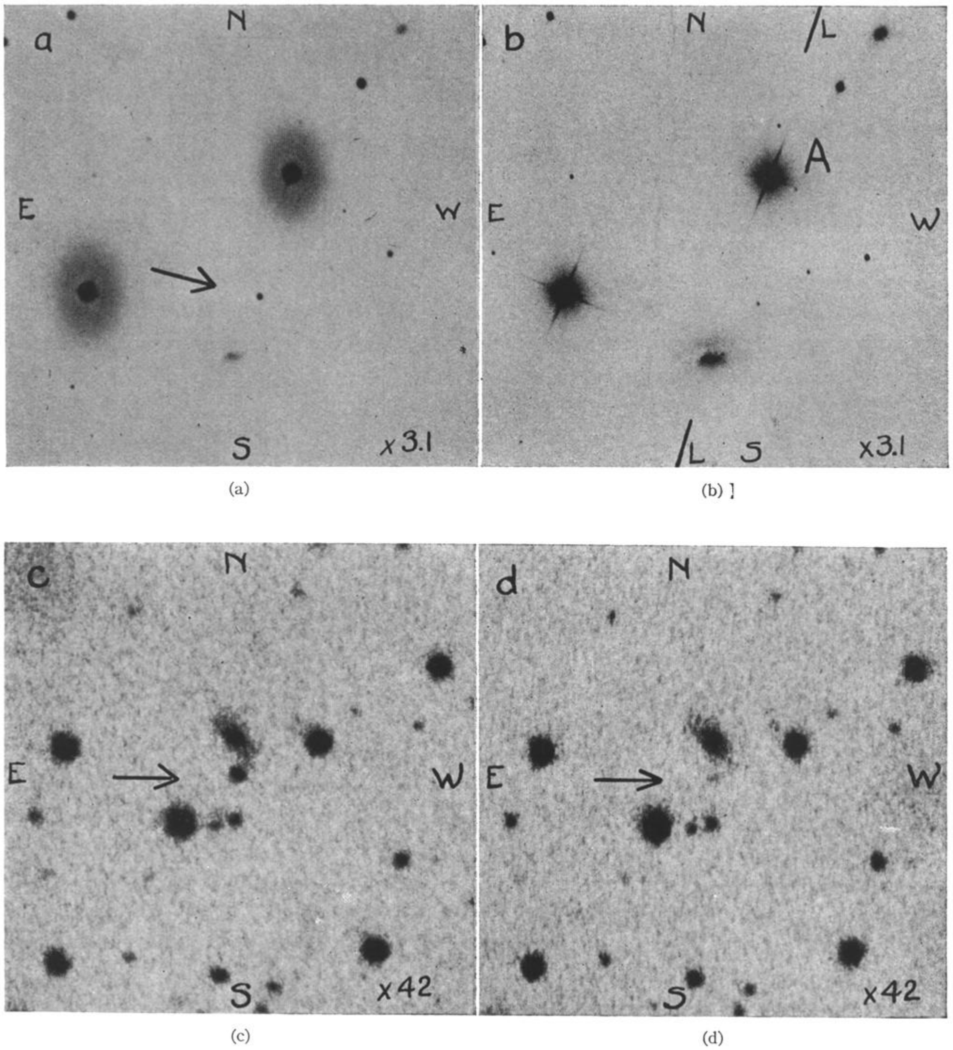

FIG. 3. (a) Shows NGC 1482 [R.A. 3^h 51.9^m, Decl. $-20° 41'$ (1938)] with supernova (arrow) about a month after maximum brightness $m_{max} \cong 14$. The photograph, which was taken by Hubble on February 5, 1938 with the 100-inch telescope on the Mt. Wilson Observatory is of very poor quality due to poor seeing conditions. In addition imperfections of the aluminum coat of the secondary mirror caused the elliptical halos around the bright stars.

(b) Good photograph of NGC 1482, taken by Baade on October 25, 1938. The supernova has become fainter than the apparent photographic magnitude $m = 21$ and is entirely invisible on the photograph. The photometer tracing shown in Fig. 4 was made along the line L indicated in (b).

(c) Shows the barred spiral R.A. 2^h 34.0^m, Decl. $+34° 10'$ (1938.0) with a supernova (arrow) of the apparent photographic magnitude $m = 15.3$. The photograph is a strongly enlarged section of a film obtained with the 18-inch Schmidt telescope on November 17, 1938, a few days after the supernova had reached maximum brightness. The nebula belongs to the Pisces-Perseus cloud of nebulae the distance of which is of the order of 25 million light years. The supernova shown is therefore the *most distant* star ever photographed whose light curve and spectrum were observed and whose distance could be determined. According to a private communication by R. Minkowski the apparent velocity of recession of the above nebula is of the order of 4500 km/sec. which is in agreement with the distance given.

(d) The same field as (c) photographed with the Schmidt telescope on January 17, 1939. The supernova (arrow) has become much fainter, its apparent photographic magnitude being $m = 18.2$. Both of the enlargements (*c*) and (*d*) of pictures taken with the 18-inch Schmidt telescope, well illustrate the good optical performance of this instrument. In spite of the fact that the sections shown lie at the edge of the field over four degrees of arc from its center, no systematic distortions of the star images can be detected. The image of the supernova in (d) can clearly be seen on the original film in spite of its faintness ($m = 18.2$).