FLUORESCENT ORIGIN OF THE LIGHT FROM A SUPERNOVA OUTBURST*

P. Morrison

Department of Physics and Laboratory for Nuclear Science,
Massachusetts Institute of Technology, Cambridge, Massachusetts

and

L. Sartori

Science Teaching Center, Massachusetts Institute of Technology, Cambridge, Massachusetts
(Received 6 January 1966)

Supernovae of type I (SN I), even though they occur in different regions within a variety of galaxies, form a remarkably homogeneous group. Their properties may be summarized as follows[1,2]:

(1) The photographic light curve (log luminosity in the spectral region 3800-5200 Å) shows a sharp sudden rise, reaching a maximum after a few days; the average maximum luminosity[2] is more than 10^9 times solar. After maximum there occurs a rapid dimming by about three magnitudes in some 20-30 days, followed by a slower decline which is nearly exponential, with a half-life about 50 days, continuing with no marked change for as long as 640 days.[3,4] However, the lifetimes for different supernovae are not the same; they lie in the range 40-80 days.[5]

(2) The color reddens rapidly during the early period of observation. After about 40 days the image turns somewhat more blue,[5] and after about 100 days the color remains approximately constant.

(3) The optical spectrum consists largely of broad bands, about 100-200 Å wide. The most prominent of these has its peak near 4600 Å in the early spectra, and contains about half the total intensity in the photographic region. Appreciable continuum is seen only during the first few days.

(4) Most of the observed SN I have been too faint to permit spectral analysis beyond a short period after maximum. Our knowledge of the time development rests largely on Minkowski's series of spectra,[6] now nearly 30 years old, of the supernovae in IC 4182 (t = 8-339 d) and NGC 1003 (t = 0-115 d). Both sets of spectra retain the same general character throughout the period of observation. The most interesting change is a gradual red shift in the location of four peaks in the blue part of the spectrum: the principal one near $\lambda 4600$ and three less prominent ones near $\lambda 4900$, $\lambda 4800$, and $\lambda 4300$. The over-all displacement in each case is about 100 Å. One peak in the red (near $\lambda 5900$) can be followed for a long time and does not red shift. The other red features behave erratically: Some bands split, peaks disappear and appear; it is not possible to follow any other single red peak unambiguously as a function of time.

Adequate explanation for the optical properties of SN I has thus far been lacking. The near-exponential decline, with a half-life close to that of Cf^{254}, has given rise to the proposal that fission radioactivity may play a dominant role.[4,7] It is of course far from self-evident, even if this hypothesis should be correct, that the radioactive energy input is transformed into visible light with the same time dependence.

We present here a simple explanation of the SN I light output, fundamentally kinematic, which entails that the exponential decline in the light curve has nothing whatever to do with radioactivity. We explain the nature of the visible spectrum, including widths, frequency shifts, and intensity variations, assuming that the bulk of the direct energy release is a sudden burst of radiation which lasts at most a few days. At first, this emission is concentrated in the ultraviolet (or even beyond), as is to be expected in view of the presumed high excitation at the time of explosion. With time, as the supernova expands, the output shifts toward lower frequencies; the visible power therefore increases even while the total emission is decreasing. The "direct" emission contributes only during the days of initial rise and fall.

After maximum, we suppose that the direct emission continues to decrease so rapidly that it soon becomes negligible. The light observed after 10-20 days is fluorescence, i.e., visible lines excited by the initial burst of ultraviolet continuum as it impinges on the interstellar gas which surrounds the supernova.[8] As time passes, one observes the fluorescence emanating from regions of space progressively farther and farther away from the explosion; the observed near-exponential decline in intensity with time simply reflects the essentially exponential attenuation of the exciting ultraviolet pulse as it moves outward in space. We call this phenomenon "optical reverberation." The most important emitter of fluorescence is the ion He II, whose lines account for the majority of the observed spectral features.

In addition to creating the ionization and exciting the fluorescence, the photon outburst also accelerates outward everything in its path. The radial motion of the emitting atoms is responsible for the broadening of the spectral features, and also for their shifts: During the early period most of the observed radiation comes, for kinematic reasons, from the near side of the emitting region, and therefore appears blue shifted. Later, more of the observed radiation must emanate from the far side; the peaks therefore shift toward the red.

Theory of Optical Reverberation.—Suppose that a δ-function pulse, suddenly emitted at some point, travels out and excites secondary radiation. An observer at a distance L sees the direct pulse after a time L/c. At a "local time" t, measured from the observer's first glimpse of the event, he sees secondary radiation from those points for which the total light path is $L+ct$; the locus of such points is an ellipsoid,[9,10] but since the secondary emitters are so much nearer to the source than is the observer, this ellipsoid effectively reduces to the paraboloid

$$r(1-\cos\theta)=ct. \qquad (1)$$

We fix our attention on a single line of the secondary spectrum. Let N denote the number of photons in the primary pulse which fall in the proper frequency band to excite this line. The flux of such photons which arrives at point r at time t' is

$$\Phi(r,t')=(N/4\pi r^2)A(r)\delta(t'-r/c), \qquad (2)$$

where

$$A(r)=\exp[-\int_0^r a(\rho)d\rho]$$

measures the attenuation due to absorption, $a(\rho)$ being the linear absorption coefficient. (Isotropy is assumed.) If the absorption is uniform, we have merely

$$A(r)=e^{-r/\Lambda}, \qquad (4)$$

where Λ is the mean free path. The strength of the secondary emission is

$$\epsilon(r,t')=\frac{\alpha N\bar{\sigma}n}{4\pi r^2}A(r)\delta\left(t'-\frac{r}{c}\right)\frac{\text{photons}}{\text{cm}^3\text{ sec}}, \qquad (5)$$

where n is the concentration of absorbers, $\bar{\sigma}$ is the average cross section over the absorption line, and α is the relative probability that an atom, having absorbed a photon from the primary beam, decays by a mode which includes the optical photon under consideration. From Eq. (5) it follows that the effective luminosity (in photons/sec) which an observer infers from his measurements at a local time t is

$$\mathcal{L}(t)=\frac{\alpha c}{2}\int_{\frac{1}{2}ct}^{\infty}N\bar{\sigma}n(r)A(r)\frac{dr}{r}. \qquad (6)$$

The average cross section can be written as

$$\bar{\sigma}=(\pi e^2/mc)(f/\Delta\nu), \qquad (7)$$

where f is the absorption oscillator strength and $\Delta\nu$ the effective width of the absorption line. Hence Eq. (6) becomes

$$\mathcal{L}(t)=\alpha\frac{\pi e^2}{2m}f\frac{dN}{d\nu}\int_{\frac{1}{2}ct}^{\infty}\frac{n(r)A(r)}{r}dr, \qquad (8)$$

where we have put $N\cong(dN/d\nu)\Delta\nu$. Equation (8) is the fundamental equation of our theory. It shows directly how the time variation of $\mathcal{L}(t)$ reflects the space variation of the product $n(r)A(r)$. If, in particular, the absorbers are uniformly distributed and the width $\Delta\nu$ is independent of r, then $A(r)$ is given by (4) and the integral in (8) is the exponential integral:

$$\mathcal{L}(t)=\alpha\frac{\pi e^2}{2m}\frac{dN}{d\nu}nfE_1\left(\frac{ct}{2\Lambda}\right);$$

$$E_1(x)\sim\frac{e^{-x}}{x}(1-x+\cdots), \qquad (9)$$

where[11]

$$\Lambda=\frac{1}{\pi c r_{\text{cl}}}\left(\frac{\Delta\nu}{nf}\right)=37\text{ sec cm}^{-2}\frac{\Delta\nu}{nf}. \qquad (10)$$

Suppose further that the emitters of secondary radiation are in radial motion with respect to the primary source. Then each photon is observed Doppler-shifted by the fraction $(v/c)_{\text{line of sight}}$. The red-shifted radiation observed at local time t originates from the retion $\frac{1}{2}ct<r<ct$, whereas the blue-shifted component comes from $r>ct$ (Fig. 1). But, because of the attenuation of the exciting flux, most of

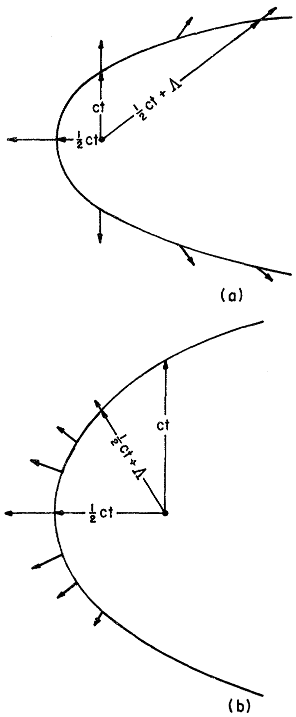

FIG. 1. The region kinematically accessible to observation at local time t. (a) $\frac{1}{2}ct \ll \Lambda$; (b) $\frac{1}{2}ct \gg \Lambda$.

the intensity originates roughly from the region between $\frac{1}{2}ct$ and $\frac{1}{2}ct+\Lambda$. Consequently, for $ct \ll \Lambda$ [Fig. 1(a)] a preponderance of the radiation originates from the near side, and the line is observed as blue shifted. When $ct \gg \Lambda$ [Fig. 1(b)] the opposite is true. The net effect is therefore a gradual reddening, which takes place over a time roughly Λ/c. We emphasize that the prediction of a red shift in no way depends on a uniform-density model.

Figure 1 also provides an estimate of the apparent size of the "indirect image" at a local time t. The diameter is approximately

$$D = 2(2\Lambda ct)^{1/2}. \qquad (11)$$

Notice that for some time the image expands at a rate faster than c.

The Optical Spectrum.—Since the direct outburst lasts a short time we are justified in approximating it by a δ function. The intensity of each fluorescent line is then given by Eq. (8), in which the absorber concentration $n(r)$ and the attenuation factor $A(r)$ denote effective time averages of these quantitites over the interval during which the actual pulse passes the point r. (These quantities in fact change as the pulse both creates new ionization and accelerates the ions.)

We identify the Paschen α line of He II ($n=4 \to n=3$, λ4686) as the origin of the dominant blue band in the SN I spectrum. The initial blue shift of this line, and the subsequent reddening, are just what one would expect from the argument presented above. In addition to Pα, the visible lines of He II which can be excited by absorbtion from the ground state are the Brackett series components np-$4s$ and np-$4d$; all these lines up to $n=13$ (as well as a merged band from the remainder of the series) can be clearly recognized in the spectra from NGC 4496 shortly after maximum (Fig. 2), published by Bloch, Chalonge, and Dufay[12]; each line is blue shifted by some $1\frac{1}{2}\%$, about the same as the principal peak. The He II lines account for almost all the spectral features in the photographic region,[13] including three of the four peaks which Minkowski observed to redden.[6] (The fourth might be He I λ5015.)

The remarkable decay of the light curve with a single lifetime is mainly the consequence of two properties of the helium II level structure. First, the levels from which the Brackett lines originate are so close in energy that the Doppler broadening causes a considerable overlap of the absorption lines which excite the visible spectrum. (All the levels with $n \geq 5$ lie within 2 eV of the ionization edge at 54 V, and a broadening of a few percent can be inferred from the observed width of the emission features.) Consequently, the flux which excites the Brackett lines effectively constitutes a single band, 2-3 eV wide, which attenuates with a single mean free path. Moreover, it happens that the summed oscillator strength for the absorptions which excite the Brackett lines is nearly equal to that for the single absorption which excites Pα.[14] The mean free paths, and the corresponding lifetimes for the decline of the emission lines, are therefore similarly close.

The second circumstance which acts to synchronize the time variation of all the He II lines is the fact that the final state of the Brackett transitions is the initial state for the Pα transition. In each of the Brackett transitions, moreover, the $4s$ component of the final state is highly favored; from the $4s$ state, the relative probability of decay via a 4-3 transition is about 0.4. Consequently, each Brackett photon is followed, on the average, by 0.4 Pα photon; this indirect excitation of the Pα line increases its total intensity by a factor of almost 2 and thus helps to explain its prominence. And even if the direct emission of Pα should decline more rapidly than that of the Brackett lines, the indirect mode ensures that the actual intensities display almost the same time dependence.

In view of these circumstances, and of the virtually total domination of the photographic spectrum by He II lines, we expect the photographic luminosity to follow a simple time dependence, which in the uniform model is a single expression, Eq. (9). Figure 3 shows our best fit to the IC 4182 light curve, with $\Lambda/c = 52$ days; the agreement over the period 20-640 days is better than that provided by a pure exponential.

The spectral features not accounted for by He II are due to minor constituents of the interstellar medium—carbon, nitrogen, oxygen, or possibly even heavier atoms—in various stages of ionization. The much lower concentration of the heavy elements does not necessarily imply that their fluorescent lines are correspondingly weaker. However, it does tend to increase the mean free path for the

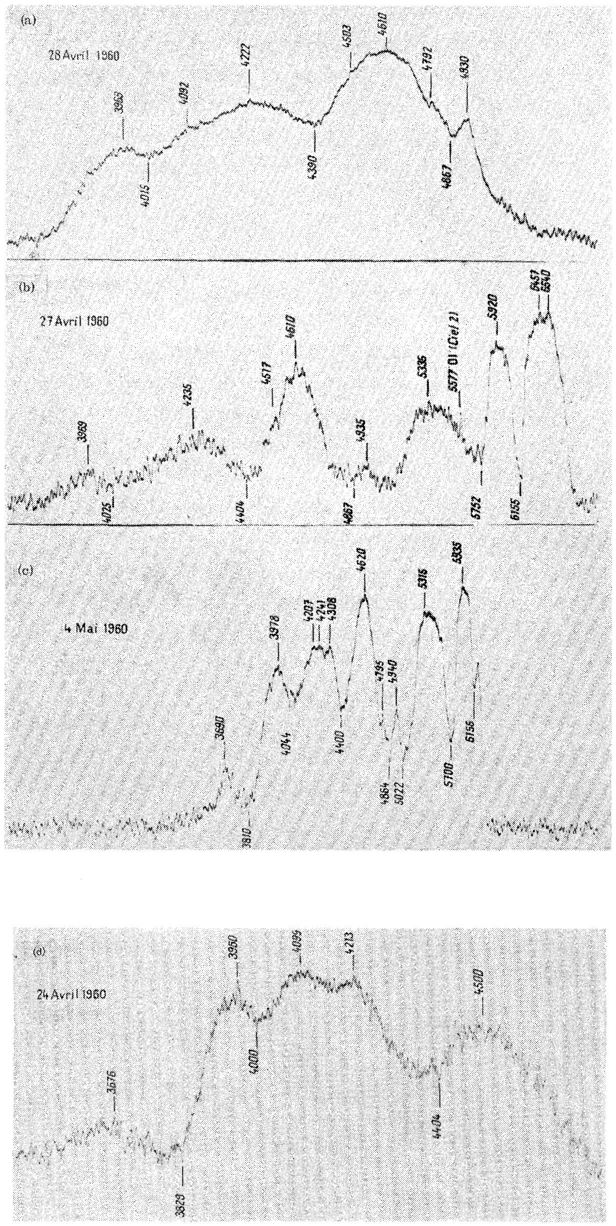

FIG. 2. Spectra of the supernova in NGC 4496, shortly after maximum, from Ref. 12. Spectra (a) and (b) were taken with a flint prism, (c) with a quartz prism, and (d) with a grating spectrograph. In these spectra we identify the following lines of He II, each blue shifted by $1\frac{1}{2}$-2%: 4-3 (λ4686), and n-4 for $n = 6$(λ6562), 7 (λ5412), 8 (λ4861), 9 (λ4543), 10 (4340), 11 (4200), 12 (4100), and 13 (4025), as well as a possible merged band from $n \geq 15$.

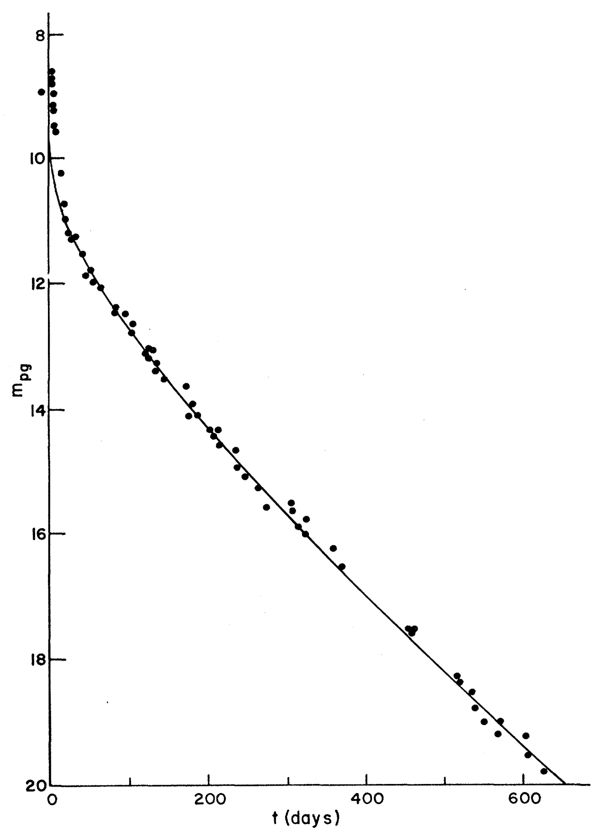

FIG. 3. Photographic magnitude of the supernova in IC 4182, as a function of time (dots) (Refs. 1 and 3). The solid curve is Eq. (9), with Λ set equal to 52 light days, normalized for best visual fit.

exciting radiation [see Eq. (10)]; the high Λ associated with many of the red lines explains their lack of observable red shift[15] and also accounts for the initial reddening of the color. The subsequent sharp decline of some red features, which causes the downturn in the color curve, probably reflects sharp boundaries of the regions in which the corresponding ions are abundant. The late spectra contain mostly helium features, which decline at the same rate. Therefore the color remains roughly constant.

Characteristics of the Fluorescent Medium.— The mean concentration of He II required to fit the light curve, according to Eq. (10), is about 4 ions/cc. This is appreciably higher than one may expect as interstellar gas even in the central plane of spiral galaxies. But the environment is enriched by all the material ejected from the star during its entire presupernova history. A tenth of a solar mass, which represents a plausible total mass loss, can easily supply the required concentration if distributed over a sphere ~1 light year in radius.

From the limit on the available mass, and the total energy emitted as reverberation light, we estimate that the average helium ion emits $\sim 10^4$ fluorescence photons. Inasmuch as the integrated excitation and ionization cross sections have comparable magnitudes, and recombination is negligibly slow at the low densities involved, such a fluorescent yield implies that the ultraviolet flux at frequencies just beyond the ionization edge of He^+ must be drastically attenuated. The required filtering is plausibly supplied by a layer of helium around the central star, sufficiently dense to make the recombination time short, and sufficiently thick to have an optical depth $\gtrsim 10$ for the ionizing radiation beyond 54 eV.

The hypothesis that most of the material in the emitting region originates from the presupernova star explains the remarkable uniformity among SN I lifetimes. It would be hard to believe that the undisturbed interstellar environment is so nearly the same everywhere that SN I are observed, whereas in view of the similarity of SN I in all other respects it is reasonable to suppose that the integrated "solar wind" should be similar, varying over a factor of only two.

The uniform model on which the specific equations (9) and (10) are based is unrealistic for a number of reasons. The concentration of a given ion cannot be strictly uniform with r, because of both the decreasing over-all density and the variation in ionization. The ion velocities, which determine the absorption width, are not purely radial, nor is the effective v_r independent of r; the attenuation of the accelerating flux should cause v_r to decrease rather strongly with r. Finally, the photons which excite the fluorescence are scattered a number of times, and do not attenuate exponentially even in a uniform medium.

The effect of the decrease of $\Delta\nu$ with radius is minimized for the He II lines by the strong overlap of the absorption lines. Even if the width of an individual line is much reduced, the overlap remains appreciable. Notice also that the expected decreases in $n(r)$ and in $\Delta\nu$ at large distance have opposing effects on the absorption rate; the departure from exponential attenuation can therefore remain slight even when $n(r)$ and $\Delta\nu$ are decreasing more markedly. That this is indeed the case for SN I is strongly indicated by the long persistence of the near-exponential decline. We further remark that as long as $A(r)$ is nearly exponential, then so is the integral (8) for large t, even if $n(r)$ varies. For example, a model in which $n(r)$ is uniform out to $\frac{1}{2}$ light year, and then falls as $1/r^2$, is not inconsistent with the observed light curves as long as the width falls off in a manner not too different from the concentration.

We gratefully acknowledge cogent discussions with Professor A. G. W. Cameron and Professor M. Oda. A detailed account of this theory has been submitted to the Astrophysical Journal.

---

*Work supported in part through funds provided by National Aeronautics and Space Administration through Grant No. NsG-386, and also in part by the National Science Foundation.

[1]F. Zwicky, in Stellar Structure, edited by L. H. Aller and D. B. McLaughlin (University of Chicago Press, Chicago, 1965), p. 367.

[2]R. Minkowski, Ann. Rev. Astron. Astrophys. 2, 247 (1964).

[3]W. Baade and F. Zwicky, Astrophys. J. 88, 411 (1938).

[4]W. Baade, E. Burbidge, F. Hoyle, G. Burbidge, R. Christy, and W. Fowler, Publ. Astron. Soc. Pacific 68, 296 (1956).

[5]D. M. Mihalas, Publ. Astron. Soc. Pacific 75, 256

(1963).
[6]R. Minkowski, Astrophys. J. 89, 156 (1939).
[7]E. Burbidge, G. Burbidge, W. Fowler, and F. Hoyle Rev. Mod. Phys. 29, 547 (1957).
[8]That some of the observed light may originate from the region outside the supernova was first pointed out by F. Zwicky, Rev. Mod. Phys. 12, 66 (1940).
[9]P. Couderc, Ann. Astrophys. 2, 271 (1939).
[10]S. van den Bergh, Publ. Astron. Soc. Pacific 77, 269 (1965).
[11]Expression (10) for the mean free path is based on the premise that only the particular line under study absorbs the primary flux. If some competitive process absorbs at the same frequency, Λ of course depends on the total cross section.
[12]M. Bloch, D. Chalonge, and J. Dufay, Compt. Rend. 250, 3952 (1960).
[13]A tentative identification of some of the SN I features as belonging to He II was given already by Minkowski (Ref. 6, Table 4).
[14]See, for example, C. W. Allen, Astrophysical Quantities (The Athlone Press, University of London, London, England, 1963), 2nd ed.
[15]According to our theory the Bβ and Bγ lines (λ6562 and λ5412) must red shift. This is not excluded by the spectra of Ref. 6.

OBSERVATION OF ~500-keV PROTONS IN INTERPLANETARY SPACE WITH MARINER IV

S. M. Krimigis and J. A. Van Allen
Department of Physics and Astronomy, University of Iowa, Iowa City, Iowa
(Received 10 January 1966)

We report herein the observation of protons of kinetic energy $E_p \sim 500$ keV during a 10-month period of interplanetary flight by the Mariner-IV spacecraft.

The impulsive emission of energetic protons ($E_p \sim 1$ BeV) from the sun was shown first by Forbush.[1] Massive studies of the arrival at the earth of solar cosmic rays (protons, alpha particles, and heavier nuclei) in the energy range ≳10 MeV have been made by ionospheric methods,[2] by direct detection with satellite equipment,[3] and by other methods. The frequent occurrence of beams of protons of $E_p \sim 1$ MeV, below the energy and/or intensity threshold of previously used techniques, was inferred by Gregory[4] from 2.3-Mc/sec ionospheric scatter measurements in Antarctica during 1959, 1960, and 1961. The inference of Gregory was confirmed and placed on a firm foundation by an extended series of direct interplanetary observations in late 1962 by Van Allen, Frank, and Venkatesan[5] using a detector having an energy threshold $E_p = 0.5$ MeV on the Mariner-II spacecraft.

The University of Iowa detector complement on the Mariner-IV spacecraft consists of three thin-window Geiger tubes (EON type 6213) and a thin (~35-μ) surface barrier nontotally depleted solid-state detector (Nuclear Diodes, Inc.). A description of the detectors and of other experimental details has been published elsewhere.[6] In this note we are concerned primarily with results from the solid-state detector which has two discrimination levels set to count protons in the energy ranges $0.50 \leq E_p \leq 11$ MeV (channel D_1) and $0.88 \leq E_p \leq 4$ MeV (channel D_2). Both channels are insensitive to galactic cosmic rays and to electrons of any energy, in the intensities found to be present by the Geiger tubes during the Mariner-IV flight. The detector is equipped with a weak $_{95}Am^{241}$ source of alpha particles to provide assurance of its proper operation in flight. The conical collimator of the detector has a half-angle of 30°, and the spacecraft is oriented continuously so as to point the center line of the collimator toward the inner solar system at 70° to the spacecraft-sun line. The absolute value of the unidirectional geometric factor is 0.065 ± 0.003 cm² sr. Simultaneous data from the three Geiger tubes assure that all protons reported herein are in fact entering the solid-state detector through its collimator and not through the protective shield.

Figure 1 shows the daily averages of the counting rates of channels D_1 and D_2 during 10 months of interplanetary flight. Our apparatus performed properly throughout. "Cruise science" was commanded "off" for days 197-214 and days 243-245 of 1965 in order to use the spacecraft for other purposes and was finally commanded "off" on day 274 for an indefinite period, thus terminating the present series of observations.

The daily average of the counting rate of the Deep River neutron monitor (courtesy of H. Carmichael) is also plotted in Fig. 1 as is the daily sum of the three-hour geomagnetic disturbance indices K_p. These auxiliary data provide di-

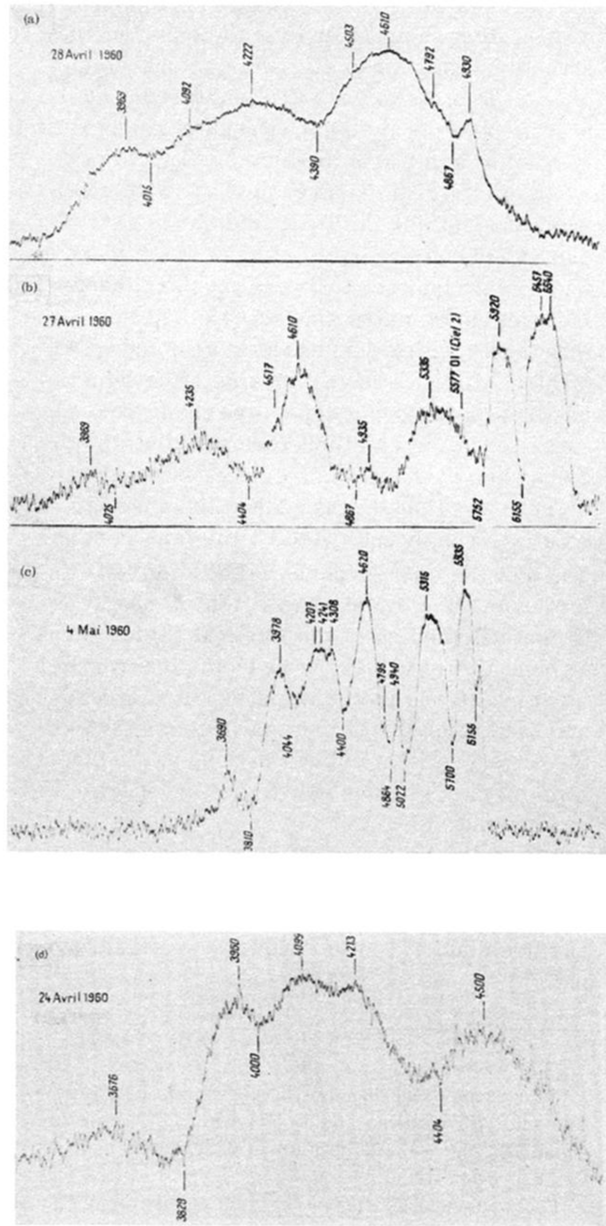

FIG. 2. Spectra of the supernova in NGC 4496, shortly after maximum, from Ref. 12. Spectra (a) and (b) were taken with a flint prism, (c) with a quartz prism, and (d) with a grating spectrograph. In these spectra we identify the following lines of He II, each blue shifted by $1\frac{1}{2}$–2%: 4-3 (λ4686), and n-4 for $n = 6$(λ6562), 7 (λ5412), 8 (λ4861), 9 (λ4543), 10 (4340), 11 (4200), 12 (4100), and 13 (4025), as well as a possible merged band from $n \geqslant 15$.