Title
Fourth day of creation: The search for a history of star formation

Permalink
https://escholarship.org/uc/item/6cq80309

Journal
STAR FORMATION NEAR AND FAR - SEVENTH ASTROPHYSICS CONFERENCE

ISSN
0094-243X

Author
Trimble, V

Publication Date
1997

License
CC BY 4.0

Peer reviewed
Fourth day of creation: The search for a history of star formation

Virginia Trimble

Citation: AIP Conference Proceedings 393, 15 (1997); doi: 10.1063/1.52780

View online: http://dx.doi.org/10.1063/1.52780

View Table of Contents:
http://scitation.aip.org/content/aip/proceeding/aipcp/393?ver=pdfcov

Published by the AIP Publishing

Articles you may be interested in

Summary: Star formation near and far
AIP Conf. Proc. 393, 617 (1997); 10.1063/1.52830

Cosmic star formation history
AIP Conf. Proc. 393, 481 (1997); 10.1063/1.52821

Current issues in star formation
AIP Conf. Proc. 393, 3 (1997); 10.1063/1.52807

Star formation history of irregular galaxies

The history of the star formation rate in the local disk
AIP Conf. Proc. 393, 437 (1997); 10.1063/1.52755
Fourth Day Of Creation:
The Search For A History Of Star Formation

Virginia Trimble

Abstract. The history of how our present understanding of star formation was achieved appears not to have been previously investigated at any length. We begin here with the ancients; note the difficulty with which the community accepted the existence of widespread interstellar material; digress briefly on the origins of the solar system; explore the dark ages in which star formation was not regarded as a part of serious astronomy, followed by gradual redawning light; and end with a few notes on the first appearances of some ideas now regarded as important.

INTRODUCTION: THE ANCIENTS FROM CHEOPS AND HERSCHEL TO RITTER AND RUSSELL

Vaya'ash Elohim et shney hamorot hag'dolim.... And God made two great lights, the greater light to rule the day and the lesser light to rule the night. V'et hakochabim. And also the stars... Vayechi erev, vayechi boker, yom r'vi-i. And there was evening, and there was morning. A fourth day. There are three things to notice in this scenario for star formation. First, all the stars will be the same age. Second, star formation is not really a separate process, but is more or less incidental to the creation of the universe and the formation of the solar system. We will encounter both of these again. Third, stars did not appear until the fourth day, and since this was a three day conference, we cannot expect to have solved all the problems. When I started to think seriously about this presentation, in July or thereabouts, I began by asking for advice from colleagues involved more directly with history of science. The most succinct and eventually helpful answer came from Owen Gingerich. He said, "There is no history of star formation." I feared briefly that he meant the subject had no intellectual underpinnings. But what he actually meant was that nobody had written about it at length. This is,
therefore, a pioneering effort. Thus I am not sure I have hold of all the right ideas (and additional information and insights would be very welcome). But I am sure I have hold of too many ideas. Thus the oral presentation simply stopped. Quite suddenly. And this one may too. In contrast to the Hebraic view, the Egyptians and Greeks seem to have thought of star formation as an ongoing process, where minor gods and important people (including Pharoah) might take their places among the stars from time to time. Numerous ancient and medieval philosophers speculated on the nature of the stars and possible analogies with the sun. Giordano Bruno, shortly before 1600, was apparently the last person for whom the statement "stars are suns" was a dangerous one to make. Applications of early telescopes by Galileo and others led to the resolution of many fuzzy patches on the sky (nebulae) into stars and to the widespread opinion that all nebulae could be so resolved with sufficient telescopic power. Later, Newton put his considerable scientific weight behind the view that attenuation of starlight on its way to us was always unimportant. William Herschel, in his last three decades, 1784-1814, concluded that the sorts of nebulae that he had been studying would gradually transform themselves into stars. He suggested a sequence from planetary nebulae, to bright nebulae (like Orion), to nebulous stars and clusters (like the Pleiades), with drawings of real objects to illustrate each stage (Hoskin 1963, Berry 1898, Herschel 1784). Herschel’s sequence implied a picture of star formation and evolution in which: “Young stars were the hottest and therefore the bluest [the ones seen around the nebulae]. As they cooled and contracted, they became middle age yellow stars like the sun. Still cooler in old age, they reddened. Gradually their light became redder and feeble until they ceased to shine.” (from Newcomb and Baker, 1932, describing the situation “until recent time.”) Curiously, this picture of stellar evolution lingered in most minds and in the literature long after it had ceased to make sense. Thus we find Astronomer Royal David Gill pondering in 1884: “The nebulae - what are they? Are they, too, condensing into clusters or stars, or will their ghost-like forms remain for ever unchanged among the stars? or do they play some part in the scheme of nature of which we have as yet no conception.” And, in 1885, Newcomb and Holden: “We are thus led to the general conclusions that, so far as our knowledge extends, nearly all the bodies of the universe are hot, and are cooling off by radiating their heat into space.” They go on to speculate about, and reject, regeneration processes, such as might be caused by accretion (for instance of the planets into the sun), and in this were wiser than a later generation. Simon Newcomb is generally said to have been Walt Whitman’s “learned astronomer”. That most of us find the stars more beautiful, not less, for knowing how they shine is why we are astronomers and not poets!

The first half of the nineteenth century saw conservation of energy established as a firm principle, requiring an energy source for the stars. Helmholtz (1853) and Kelvin (1862) are universally credited with the idea of contraction under self-gravity as the main energy source for the sun and stars (though
agreement on the proper reference is less universal!). The corresponding picture of star formation then begins with "meteoric material" that condenses, heats, vaporizes, and, when it can contract no more, radiates away stored heat, reddening and fading as it goes. S.C. Vogel's 1874 classification of stellar spectra implies such a scenario, and it is more explicitly presented by Ritter (1882), and, especially, Norman Lockyer (1887, 1888, perhaps best known as the discoverer of helium in the sun and the founder and first editor of Nature). Lockyer's paper includes a drawing that looks like water spraying up and out of a hose from right to left, but is meant to be viewed left to right, "On one arm of this we have those stages in the various heavenly bodies in which in each case the temperature is increasing, while on the other arm we have that other condition in which we get first vaporous combination, and then ultimately the formation of a crust due to the gradual cooling of the mass, in dark bodies like, say, the companion of Sirius."

The year 1913 appears to have been something of a watershed, and in the same volume of Observatory, we find Eddington (1913) describing B stars as young, M stars as old, and evolution as a cooling process, while Russell (1913) is enunciating what became known as his "giant and dwarf" theory, essentially the equivalent of Lockyer's picture, in which cool nebular material contracts and heats across the top of the HR diagram (giants), and then stars gradually cool off, moving diagonally downward on the main sequence (dwarfs). The phrases "early" and "late" type stars for hot and cool are left from these early pictures of stellar evolution. By about 1923, belief in the giant and dwarf theory was nearly universal. Eddington (1923) noted correctly that the largest mass stars would get hottest, and Shapley (1923) believed (incorrectly) that the largest masses evolved most slowly and that open clusters were globular clusters that had been disrupted by passage through the galactic plane. It is worth emphasizing that, throughout this period, star formation was regarded as an on-going process, not very well understood (though perhaps connected with spiral nebulae among others), but worthy of the attention of a serious astronomer. This is clear, for instance, in the Newcomb and Holden (1885) volume, in Young's (1904) astronomy text, in the first (1926) edition of Russell, Dugan, and Stewart, in Ball's (1902) Royal Institute lecture, and so forth.

THE GRUDGING ACCEPTANCE OF AN INTERSTELLAR MEDIUM

The view espoused by Galileo, Newton, and others that all nebulae could be resolved into stars held sway for a very long time. For instance, Lord Rosse's original discussion of the Crab Nebula regarded it as a fuzzy patch of imperfectly separated star light, and Agnes M. Clerke, the summarizer of the 19th century, opposed any form of diffuse material until she realized that accepting it would spare her having to consider the ridiculous idea of
external galaxies. But in August 1864, William Huggins (1865) put the slit of his spectroscope across “a small but comparatively bright nebula 37 H. iv,” and was most surprised to find three “bright, bluish-green lines, separated by dark intervals”, indicating the presence of “a luminous gas.” He attributed the lines to nitrogen, hydrogen, and an unknown source. (I have not checked whether he was right about N and H, but the unknown line was surely [OII] at 5007 Å, later attributed to nebulium). He was also very firm in saying that this luminous gas (which he later also found in Orion, the dumb-bell nebula, the “annular” nebula in Lyra, and a number of others catalogued by Herschel and Struve) was not Herschel’s “nebulous fluid” from which stars could condense. A firm disciple of Kirchhof and Bunsen, Huggins believed that every line of every element seen in absorption in stellar spectra should appear as an emission line in gas able to make stars. That the emitting gas is mostly hydrogen does not seem to have been fully realized until the work of Struve and Elvey (1938) on diffuse H-alpha emission. They derived a ratio by number of H/Ca=100, which is still much too small by modern standards, but did not seem so far out of line at a time when stars were, perhaps, 10-40% hydrogen (Eddington 1923a, Stromgren 1937). In retrospect, we can say that diffuse absorbing gas was discovered in 1904 when Hartmann (1904) described spectra of δ Ori, saying: “The calcium line at λ 3934 does not share the periodic displacement of the lines caused by the orbital motion of the star.” These stationary lines had also been reported in the hydrogen spectrum of Nova Persei 1901. Hartmann suggested that the gas responsible might have “some relation to nebular masses seen by Barnard,” but also admitted the possibility that the gas might be closely associated with the stars in whose spectra the lines appeared. This remained the majority opinion for a couple of decades, partly because stationary lines were, until quite late, seen only in the light of stars earlier than B3 (this is a matter of crowding and the extent technology).

An important argument against “associated lines” (as the quasar people say) was that the line velocities, corrected for solar motion, were very close to the local standard of rest in Hartmann’s two examples and, soon after, other stars (Plaskett 1924 and references therein). Slipher (1909) opined that stationary lines were probably interstellar, and by 1926, nearly everybody was sure of this (Eddington 1926). An important confirmation was that the strengths of the lines tended to be proportional to our distances from the stars concerned (i.e. inversely proportional to proper motions, Gerasimovich and Struve 1929).

The Hartmann (1904) paper is a translation from a German original prepared by the editors of Astrophysical Journal (a service no longer available, as a quick scan of any recent issue will persuade you). To modern eyes, one of the most interesting aspects of the issue of ApJ in which the “stationary lines” paper appears is the advertisements. Modern scientific journals normally ban ads, or allow only books, telescopes, statistical software, and the like. But ApJ in 1904 featured (a) Pears’ Soap (“beau-begging beauty bubbles from...” indicating some uncertainty about the gender of their readership), (b) Vose
pianos and Weber pianos, (c) Baker’s Breakfast Cocoa (less than one cent a cup), (d) Sapolio hand Kosher soap, recommended to “Gentiles as well as Jews”, and (e) most spectacular of all, Buffalo Lithia water, endorsed by a couple of dozen doctors and professors as of use in Bright’s disease, gout, renal calculi, rheumatism, and so forth. We wonder whether readers then were more or less likely to believe everything they read in ApJ than readers now. Interstellar matter that obscures our view of the stars behind was also seen early and recognized late. Herschel’s first statement (Hoskin 1963) was, “Mein Gott, da ist ein Loch in Himmel” (a hole in the heavens), and the opinion that the coalsack and other dark regions had few or no stars lasted a long time. Real obscuration was suggested repeatedly, by Angelo Secchi in 1853, by A.C. Ranford (1894, not an astronomer), by Max Wolf (1908), and H.D. Curtis (1918) among others. Curtis said that one should study spectra of the stars embedded in the dark nebulosity to learn about it (a very Curtis-like remark). E.E. Barnard was, in some sense, the last to favor true obscuration (because thereafter everybody believed in discrete opaque clouds). He began photographing the Milky Way in 1904 and, 15 years later, said: “I did not at first believe in these dark obscuring masses. The proof was not conclusive. The increase of evidence, however, from my own photographs convinced me later, especially after investigating some of them visually, that many of these markings were not simply due to an actual want of stars, but were really obscuring bodies nearer to us than the distant stars. In this way it has “fallen to my lot to prove this fact.” (Barnard 1919, writing in a style no longer fashionable in the main archival journals). He expressed uncertainty about whether the obscurers were gaseous or something else. Looking ahead for a moment, we find Bart Bok and Edith Faith Reilly pointing to some of the more compact of Barnard’s dark features as protostars (Bok and Reilly 1947). We now call them Bok globules, and many are indeed known to have infrared-emitting young stellar objects at their cores. As late as Lynds’ (1967) article in the Kuiper compendium, it was not certain what the masses of the globules were by more than a factor of 10, because there was no way to be sure whether gas was as over-dense in them as dust. In the same time frame, Minkowski (1949) finally enunciated the idea that absorption and emission nebulae must often be part of a united complex of gas and dust. Why Bok globulae and not Bok-Reilly globules? Perhaps because, apart from some observations of meteors, comets, and variable stars, this was Reilly’s only astronomical paper, and she did not proceed to a PhD, in turn perhaps partly because a physical handicap that interfered both with the clarity of her speech and with walking and manual dexterity (Hoffleit 1996). Meanwhile, Hubble (1922) defined a category of reflection nebulae whose spectra were essentially those of stars in or near them, and Russell (1922) pointed out that these must be made of finely divided dust to produce the observed effect (though he expressed very limited enthusiasm for such quantities of “meteoric material”). The existence of widespread, diffuse interstellar absorption was explicitly denied by pundits.
from Newton (Hoskin 1982) to Shapley (1918) (though advocated by King 1914 and a few others), with unfortunate consequences for the establishment of distance scales within and outside the Milky Way (see Fernie 1969 and Trimble 1996 for various aspects of the story). The official culture hero here is, of course, Robert Trumpler (1930), whose study of the apparent magnitudes vs. angular diameters of open clusters persuaded him and, soon, everybody else, that the clusters were being systematically dimmed by about 1 m/kpc in the galactic plane as well as by their distances. Molecular interstellar gas comes next, beginning with optical observations of unidentified lines in the period 1937-41 and their gradual recognition as belonging to CH and CN (Dunham 1937, Swings and Rosenfeld 1937, McKellar 1940, 1941, Adams 1941, Douglas and Herzberg 1941). Because the CN lines came from two different lower levels, an excitation temperature could be measured. It was just a little less than 3K, a fact which has (so says Gerhard Herzberg 1951 near the end of his classic treatise on the spectra of diatomic molecules) "only a very restricted meaning."

THE ISM THICKENS

Notice that, through the 1920’s, 30’s, and 40’s, though the existence of several components of interstellar material was generally recognized, it was all pretty tenuous, spotty stuff. There were Huggins’s bright and Barnard’s dark nebulae in a few isolated regions. The stationary lines, both of atomic Ca, Na, etc. and of molecular CH and CN came from small, isolated clouds with small velocity dispersions, and a typical line of sight would penetrate only a handful per kiloparsec. Trumpler’s dust might be quite widespread, but it was clearly very tenuous indeed, with a few dozen grains per cubic kiloparsec. There was, in other words, very little evidence of massive, gravitationally bound clouds of gas and dust, waiting eagerly to collapse into stars. The discovery of 21 cm emission by neutral hydrogen changed the situation substantially. The expected wavelength and line intensity had been predicted by van de Hulst (1945) and Shklovski (1949), and the line was seen in quick succession by observers in the US (Ewen and Purcell 1951), The Netherlands (Muller and Oort 1951), and Australia (Pawsey 1951). These three papers appear in that order in a single issue of Nature. In contrast to the spottiness of stationary lines, bright, and dark nebulae, 21 cm emission was, literally, all over the sky. And it quickly became obvious that neutral gas outweighed ionized gas by a factor of 50-100 and comprised at least 5% of the mass of the Milky Way (Kerr and Westerhout 1965). Here at last was a worthy candidate for the raw material of ongoing star formation. The only trouble was that it didn’t seem to be concentrated toward those promising Bok globules (remarked upon by G. de Vaucouleurs in 1957 and undoubtedly many others). The final, missing, phase was, of course molecular hydrogen. Van de Hulst (1948) had said it
could form on grains, but, like all the other phases, it seems to have had to endure a period of, at best, marginal respectability. Though Spitzer (1949) thought it might be quite abundant and came back to this view, at least where Bok globules were concerned in his Compendium review (Spitzer 1968), two of the more serious discussions came from Gold (1961) and Zwicky (1959). They have in mind molecular hydrogen as a significant dark matter contributor in the Milky Way ("Oort limit"), clusters of galaxies, or both. Zwicky indicated that he was tool up to look for the 85µ ortho/para transition, apparently not realizing the extent to which the earth's atmosphere does not want us to do 85µ spectroscopy from the ground! In the same time frame, Bok (1959), also a supporter of H2 as dark matter, suggested that not being able to detect it was the most significant barrier to understanding the interstellar medium. Like any molecule made of two identical atoms, H2 has no strong, low-lying lines that are likely to be conspicuous in emission or absorption if the gas is cool (as it clearly had to be, since the HI was checking in at 100-125K). Thus arose the need for tracer molecules, whose lines would be strong enough to see and whose chemistry might be simple enough to reason back from their abundances to that of molecular hydrogen. First Shklovski (1953) and then Townes (1955) recommended the 18 cm, lambda doubling lines of OH as the best bet. Townes’ accurate frequency predictions led directly to the detection of interstellar OH, first in absorption against Cas A (Weinreb et al. 1963) and then in emission (Weaver et al. 1965). For better or for worse, OH emission, from the very beginning, displayed the peculiar polarization properties and anomalous ratios of line intensities that are the signatures of its being mased. The numbers of photons you see in its lines are, therefore, a very poor indicator of the amount of gas responsible. And so astronomy waited another five years, for the discovery of widespread CO emission (Wilson et al. 1970), before we had the right tracer. The correct conversion factor from CO line strength to total mass of H2 and how it varies with conditions in the gas remain under dispute to this day (Trimble and Leonard 1996), but, at long last, astronomers could be sure that there were massive, fairly dense, bound clouds of gas and dust out there ready and willing to form stars. Other components of the interstellar medium that presumably have some bearing on star formation include cosmic rays and the magnetic field. Just when cosmic rays joined the inventory depends on whether you want one person to understand them, most of the community, or all knowledgeable scientists. By this last criterion, the answer is 1994, when Hannes Alfven died (but, of course, he was also among the first to put them in the galactic inventory!). Alfven (1937) was also an early exponent of a galactic magnetic field, whose function was to confine these same cosmic rays. Spitzer (1946) first predicted a field of 10^{-6-5}G, but, within the same 600 word letter, backed off to 10^{-12}G. Fermi (e.g. 1949) argued that one needs the larger value to accelerate cosmic rays as well as confine them. Davis and Greenstein (1951) proposed the same number for the purpose of aligning interstellar grains to polarize starlight (Hall 1949, Hiltner 1949). The
field briefly climbed to a few \( \times 10^{-5} \)G, sufficient to dominate dynamics of the interstellar medium (Woltjer 1965) before direct measurements (Zeeman broadening of radio lines and rotation and dispersion measures of pulsars) pulled it back down to something \( \times 10^{-6} \)G. Virtually all modern theorists of star formation agree that the magnetic field is important (good vs. bad is another matter). But the cosmic rays seem to have been put there just to make everything more complicated.

A BRIEF DIGRESSION ON FORMATION OF THE SOLAR SYSTEM

Theories of formation of the solar system have a much longer and more distinguished history than those of star formation and have been written about at much greater length (e.g. Brush 1978, 1990, 1996, and many other people and places), frequently under the subject heading “cosmogony”. Very crudely, one can distinguished two streams of ideas. In one, often called the nebular hypothesis, the sun and planets (etc.) form together, more or less at the same time, out of a single rotating cloud of stuff (mostly gas we would now say, but this was not obvious when the sun was believed to be made of the same things the earth is). The official propounders were Swedenborg, Kant, and, especially, Laplace (who met Herschel in Paris in 1802, Lubbock 1933). The modern picture of star and planet formation (e.g. Shu et al. 1987) is a remote descendent of the nebular scenario, though squabbling over details has obscured this from time to time. The primary alternative view of solar system formation was tidal expulsion, first put forward by “the celebrated Buffon” (1707-1788, and somewhere in the writings of G.B. Shaw is a story about confusion between the two of them, the latter as “the celebrated Buffoon”). Buffon’s version was a comet hitting the sun and kicking out the planets. The early 20th century version of tidal expulsion is closely associated with the names of Thomas Crowder Chamberlain (1863-1928) and Forest Ray Moulton (1872-1952), who envisaged a close encounter between two stars. A late variant, due to Raymond Arthur Lyttleton in about 1938, invoked a single star plus a binary system in an effort to overcome various kinematic and thermal difficulties in the basic Chamberlain-Moulton hypothesis. It is hard from our vantage point to realize just how thoroughly dominant tidal expulsion was at one time. Here, for instance, is Struve (1932) reviewing Moulton’s monograph, Astronomy: “The more advanced reader will derive particular satisfaction from chapter xiv, “The Evolution of the Solar System,” which in large measure rests on Dr. Moulton’s own work. It will be recalled in this connection that in 1899 he proved the Laplacian hypothesis to be untenable—a fact which is now universally recognized.” And some discussion among characters in Dorothy Sayers’ novels from the 1930’s illustrates that tidal expulsion was the accepted explanation of the solar system among educated laymen as well.
as among professional astronomers. In a general sort of way, one expects star formation to be an ongoing process in the nebular case, while tidal expulsion fits more comfortable in a picture where all stars formed on Jeans' "long" time scale (of which more shortly), and a few experienced encounters when they were already mature. Both camps made use of spiral nebulae as part of their scenarios at one time or another. Ball (1902) and Russell (1913, 1925) are among those who mention spirals as pre-solar-systems in a nebular context. In contrast, Lowell (as in Lowell Observatory), the polymathic Svante Arrhenius, and the creative spirit T.J.J. See saw spirals as the transition stage between tidal expulsion and condensation of planets. It was in this spirit that Lowell set V.M. Slipher to taking spectra of spirals, and only gradually did Slipher come to believe that he was in the process of demonstrating the existence of external galaxies rather than studying the formation of solar systems. Jeans (1928) at one point described spiral nebulae as new material pouring into the observable universe from elsewhere, an idea which has later echoes in the work of Ambartsumian and others on active galactic nuclei.

DARK AGES AND CONFUSION

The period from about 1935 to 1945 is a very important one in this subject because of a sudden epidemic of disbelief in star formation as a continuing process. It is also, unfortunately, a difficult period to investigate systematically, because the advent of WWII meant that many books and journals did not reach the German astronomers who had been compiling annual, very complete abstract volumes since 1899 (found on your shelves with the word Jahresbericht conspicuous), and one volume, covering 1944-45, seems never to have been distributed. This section is correspondingly incomplete and in need of amplification. At the beginning of the period, the decision had not yet completely been made between the "long" and "short" time scales for cosmic evolution. Sir James Jeans (from about 1904 until very shortly before his death; see any of many editions of his Astronomy and Cosmogony) had long advocated a universe that began in some singular event more than $10^{12}$ years in the past. He had concluded that this much time was needed for an assortment of dynamical processes, including the relaxation of elliptical galaxies, clusters of galaxies, and star clusters, and the modification of binary star orbits by encounters, to bring systems into the conditions we see. (His binaries all began in circular orbits, and the wide ones were gradually eccentricised by passing near other stars, a slow and painful process indeed.) He accepted as a necessary corollary that stars must live on annihilation (presumably of electrons and protons) as their primary energy source, gradually losing mass and moving down the main sequence (a la Russell's giant and dwarf theory) over $10^{12-13}$ years. The competing, "short" time scale came from measurements of the ages of earth rocks (from decay of uranium and thorium to lead)
and carried as a corollary the fueling of stars by some sort of nuclear transformation, which would release only about 1% as much energy as complete annihilation. Eddington's 1926 book, Internal Constitution of the Stars is a standard landmark in establishing the latter view. The balance began to shift decisively in favor of the "short" time scale with Hubble's announcement of a velocity-distance relation for external galaxies (Hubble 1929) that implied a time scale of about 2 Gyr for the age of the universe as a whole. Soon after, new considerations of various dynamical processes suggested that the short time scale was also the right one for spiral galaxies (McVittie 1932), the rotation of the Milky Way (Eddington 1931), and pairs and clusters of stars (Bok 1936, Mineur 1939, Chandrasekhar 1942ff). A 1935 conference, whose proceedings appear in that year's volume of PASP (Epstein 1935 and surrounding papers) showed a clear majority of participants firmly on the side of the short time scale. Unfortunately, it was also a short step from saying that most things seemed to be about 2 billion years old to saying that everything was about two billion years old and that it had all formed from dense, chaotic conditions that no longer exist and hardly deserve study. The quotations that follow are intended to give some feel for the intellectual climate that resulted.

The collection is very incomplete and strongly biased toward the British and American literature. I am aware of, but have not been able to get hold of, relevant papers by Bertil Lindblad and G.C. Armellini (who wrote primarily in Italian and who was the only astronomer at the 1957 Vatican conference on stellar populations whose name is not instantly recognizable by most astronomers today) and of papers by Severny and Pariisky that I could not read if I had them! And there are surely others of which I am not even aware and about which information would be very much appreciated. "The trend of thought in recent years has been favorable to the "short" time scale suggested by the Expansion; but practically everyone is uncomfortable about its brevity. It does not seem sufficiently dignified that the uncompromisingly majestic universe measure its duration as scarcely greater than the age of the oldest rocks on this small planet's surface or the age of life in the crannies of the rocks." (Shapley 1943). "The combined evidence rather suggests that the stars, which constitute the observed clusters did not exist as stars until the clusters came into being. The theory of the Expanding Universe indicates that a "catastrophe" took place $3 \times 10^9$ years ago, and it is tempting to place the origin of stars and stellar systems tentatively at the epoch of this catastrophe." (Bok 1936) "At a certain time, not many billions of years ago, conditions in the universe were so crowded and presumably so different from those which now prevail, that our backward reckoning cannot safely be extended further. It is tempting to look upon this as a time of turbulence, in which various things happened which are very difficult to account for otherwise—such as the origin of the solar system and of double stars" (Russell et al. 1938). "Bright stars rely on a much more generous source [of energy] than synthesis...It is possible to get around it by supposing that heavy stars (and presumably also light ones) are
constantly being made anew from some diffuse material that consists of nearly pure hydrogen; but the idea is not very attractive.” (Atkinson 1936, followed by a detailed explanation of what is unattractive). “Might it not be possible that other Hyades or Pleiades might be formed?... it seems very unlikely...It is rather satisfactory to blame it all on conditions as they were three billion years ago, at a time when, according to the hypothesis of the expanding universe, all matter in the physical universe was packed much closer together than it is at present....But, you may well ask, why should we not admit the possibility that stars are still being formed?....There are, however, several reasons why most astronomers are reluctant to take this view.” (Bok 1945) “But great difficulties remain for the red giants and the highly luminous blue stars. We do not know what keeps these celestial power houses running. We have yet to learn whether they are recent creations or whether they have existed in some prestellar form since the beginning of time, or operate on unknown principles.” (Goldberg and Aller 1943)

**CURIOUS CONNECTIONS**

Three sorts of ideas arose during the same period, each of which now sounds distinctly odd (as well as wrong). The vociferousness of the contemporaneous objections suggest that they were at the time (especially the idea of significant accretion) somehow perceived as threatening. First was an assortment of unified hypotheses for the combined formation of stars and galaxies, some of which survived well past the time when “everything the same age” was still a possible view. Lindblad (1934) put forward such a scenario, as did Stromberg (1934,1935) and von Weizsacker (1937 to 1951). The last exponent of such unified pictures of star formation appears to have been Layzer (1964 and references therein). The second idea is a sort of inverted picture of stellar evolution in which superdense, prestellar stuff comes first and develops into stars which then expel various sorts of nebulae. Part of the idea goes back at least to 1737, when a chap named Jean-Jacques Dortous de Mairan took the solar corona (extensions of which he thought hit the earth’s upper atmosphere, causing aurorae - true, or near enough) as a general model and attributed all nebulous material to streams emanating from stars. A.A. Belopolsky (who attended at least the 11th meeting of the AAS since he appears in the conference photo) seems to have envisaged a sort of balance between attractive forces of gravitation and repulsive forces (like radiation pressure) in the universe. The late Viktor Ambartsumian was apparently much influenced by him (Ozernoi 1996) and this is perhaps part of the origin of Ambartsumian’s (1938, 1953, 1960 and beyond) conviction that star clusters, and, later, spiral arms, clusters of galaxies, and active galaxies form from dense, precursor material and are in a state of expansion. There seems to have been a sort of polarization of opinion in the USSR between viewing all nebulae as
the precursors of stars and all nebulae as the products of stellar expulsion, and one is surprised to find Shklovsky (in his popular treatment, Stars: Their Birth, Life, and Death) expending a good many words to convince the reader that planetary nebulae mark death not birth. Others mixed up two or more of the ideas just mentioned. Waterfield (1938), for instance, sketches out a picture in which star formation is part and parcel of galaxy formation, but all nebulae are, in turn, expelled from stars, and the hottest stars therefore necessarily the most evolved. The idea of dense, prestellar matter had one very attractive feature. It permitted the synthesis of the heaviest elements. Gamow and his co-workers, of course, located the necessary hot, dense stuff in the early universe, but Chandrasekhar and Henrichs (1942) appear to have had prestellar, but post-big-bang material in mind. Hoyle, even before the advent of steady state, was the earliest exponent of late, rather than early, phases of stellar evolution as a solution to the problem. Third, and at least as surprising as the other two, was the proposal that all stars had been born a few billion years ago, with masses small enough that “nuclear transformations” would fuel them down to the present time, but then a few had wandered into nebulae, where accretion of ambient material turned them into short-lived OB stars and supergiants. There is a hint of this in Eddington (1926), who remarks that accretion is most significant in nebulae and at small relative velocity, and you find B stars with these characteristics because that is where they can “grow”. The main advocates of accretion as an alternative to recent star formation were, however, Hoyle and Lyttleton (1939,1940,1941, and beyond). They considered the effects of massive accretion on the orbits of binary stars as well as on the lives of singletons, and supposed that we might be seeing very tenuous accretion on to the sun in the form of its corona. W.H. McCrea considered both star formation and accretion in the same time frame and concluded that “accretion has at least the merit of being known to occur.” You may or may not be surprised to hear that some of the Hoyle and Lyttleton papers are still occasionally cited today. This is not because anybody (or at least anybody I know) still either believes that accretion on main sequence stars is important or feels the need to refute it, but because in the process they derived the formula for accretion by a moving point mass that scales as $v^{-3}$, a formula we all use in a variety of quite different contexts. A last, still unclear, word comes from Greenstein (1951), who considers first accretion and then condensation of new stars and concludes “Therefore only in regions of unusual quiescence can a star condense. Favorable conditions may occur, but, like the accretion hypothesis, the condensation hypothesis now seems suitable only for creation of a few exceptional objects rather than all the stars.” But, rather than telling us next where most of the stars actually come from, he moves on immediately to interstellar polarization.
THE REVIVAL OF STAR FORMATION

Even as the lights of peace were going out all over Europe, astronomical light began to dawn again. Two of the swords that had cut down Jeans' "long" time scale still had some bite to their edges. The first was dynamical time scales for star clusters. Most indeed have held together for the few billion years needed by the "all together boys" hypothesis (Spitzer 1940, e.g.), but some OB associations (starting with the Trapezium in Orion) would surely fall apart in only a few million years (Ambartsumian 1938). Similarly, as work led up to Bethe's (1939) synthesis of our understanding of hydrogen burning in stars, it became clear that a few billion years was a perfectly reasonable age for stars of, say, 0.5 - 2.0 M\(_\odot\), but that very massive ones could live on hydrogen fusion or similar reaction for only millions of years (Gamow 1938). Bethe (1939) said firmly that the B stars must have been born comparatively recently, "by what process we cannot say." Russell (1939) was not prepared to go quite that far, and said of stars like Gamma Cygni that they "have begun to shine rather late in the history of the galaxy" or that "they have still greater stores of energy to draw on." He made life more difficult than necessary for himself by continuing to consider a sun (and so presumably other stars) that was only 1/3 hydrogen by weight. German astronomers were rapidly reaching similar conclusions, with, perhaps, a bit more confidence. Unsold (1944) wrote "dass solche Sterne [OBs] heute noch fortlaufend neu gebildet werden," though the choice of "gebildet" still sounds like a slight weasel. Von Weizsacker (1947) laid down a set of conditions for recognizing new stars that included rapid rotation ("rotierende"), location in spiral arms ("Spiralarme"), location near the galactic plane ("galaktische Ebene"), and a ratio of luminosity to mass exceeding 100 erg/sec/gram ("Hauptreihe"). Continuing reservations persist, however, in the direction of believing that only stars that had to have formed recently did so. For instance: "the origin of a certain class of stars which may have been created recently" in contrast to "the formation of stars in general is still a closed book, since the explosion of the universe a few billion years ago has so far defied any attempts at detailed analysis" (Spitzer 1948). And "We are thus forced to conclude that some stars may be of recent formation and even that the process of stellar birth is still going on." (Menzel 1949). The proposed site was once again the nebulae that Herschel had in mind, and Bok (1948) described the Harvard Centennial Symposium on interstellar matter as having "a central theme that in some cosmic clouds we are now witnessing the operation of the process of star formation." Struve (1949) in a review of the proceedings of the symposium is apparently still not fully convinced and writes of Joy's groups of T Tauri stars, "But perhaps they are not really 'associations,' in the sense of very wide clusters, but represent accidental groupings of stars which have drifted into the dust cloud and show similar spectroscopic features only because they happen to be subjected to the same physical influence - bombardment by interstellar grains." (We would, of course, say

This article is copyrighted as indicated in the article. Reuse of AIP content is subject to the terms at: http://scitation.aip.org/termsconditions. Downloaded to IP: 128.200.104.165 On: Thu, 19 Mar 2015 00:34:15
some similar things, but blame gas more than dust and put the source in a residual disk of material left over from the recent condensation of the stars). And (quoting Struve again), "It is, of course, tempting to search for a connection between the T Tauri stars and Bok's 'globules', but we must admit that at present there is no evidence of any objects that could be considered intermediate between the two groups." (A point that remains nearly true in so far as one has in mind protostars whose energy output is dominated by accretion.) And then in 1951, Schwarzschild, Spitzer, and Wildt (1951) step bravely forward with "The suggestion that all type I stars have been formed from the interstellar clouds may, perhaps, be taken as a working hypothesis." But "Type II stars are all old, having been formed by some different process 3 billion years ago." And the subject got what might be claimed as its first review article (Spitzer 1951). The year 1952 could reasonably be said to mark the admission of current star formation to the inventory of respectable problems in astrophysics. Blaauw (1952) studied the expanding star cluster around Zeta Persei and then the Lacertae group (Blaauw and Morgan 1953), finding them to be only 1.3 and 4.2 million years old (much as Ambartsumian had said before, but now not burdened with the weight of dense "prestellar matter"). Herbig (1952) explained in some detail why Joy's (1945,1946) variable, F5-G5, emission line stars near nebulosity (T Tauris we call them now) had to be young. An important point was their position in the HR diagram, firmly above the zero age main sequence. He returned repeatedly to the topic (e.g. Herbig 1960), gradually widening the pre-main-sequence net to take in what we now call Herbig Ae/Be stars as massive analogues of T Tauri variables and M dwarfs with emission lines as the lowest mass examples. Baade (1952), speaking at the Rome General Assembly of the International Astronomical Union, devoted a major review talk to asking the question, "Are we today in a position to point out two groups of stars, one of which in terms of cosmical time scale is young, the other which is old?" And he answered his own question: "The two groups of stars which I want to discuss are the O- and B-type stars of high luminosity as an example of young stars and the stars in globular clusters as an example of old stars." In this context, "young" meant a few million years (as per Blaauw) and "old" meant $3.4 \times 10^9$ years, as per calculations of stellar evolution by Sandage and Schwarzschild (1952), who assumed CN cycle hydrogen burning, convective cores, and no other source of mixing over the full range of 1-6 $M_\odot$. And, at that same Rome IAU meeting, Hoyle (1952) was still defending accretion as an important source of massive stars. Incidentally, other astronomers referring back to this period sometimes credit Baade with the old/young dichotomy, but attribute it to his 1944 stellar populations paper and indicate that Joy (1945) described his T Tauri stars as young. They must have read different versions of those papers from the ones that exist in our library journal collection! Finally, we close the year 1952 by looking at a non-technical book by Cecilia Payne Gaposchkin (1952, Stars in the Making), one prescient passage of which says: "This, I imagine, is why
the oldest members of the stellar system (as I suppose the globular clusters to be) are so large and populous, for the available material was richer then. As the layer of dust and gas sank toward the galactic plane, stars continued to form. Dust and gas still lie dense in this layer, and stars are still being formed there.” This strikes me as remarkably similar to the ideas we now associate with Eggen, Lynden-Bell, and Sandage (1962).

SOME IMPORTANT (AND UNIMPORTANT) PROCESSES IN STAR FORMATION

We are rapidly approaching the modern era, but it remains to note the first (and sometimes the last) appearance of an assortment of physical processes that have appeared in scenarios of star formation. Important ones include effects that might assemble dense clouds and various kinds of triggers to turn those clouds into stars. There seems to be no rational way to put these into purely chronological order, and I have not tried. It is, however, worth remembering that many of the earlier ones predate the discovery of neutral and molecular hydrogen, so that minds were necessarily being bent to the very difficult problem of trying to turn quite hot, quite tenuous stuff into stars (Spitzer 1941, 1942).

Radiation Pressure. Eddington (1926) was firmly of the opinion that radiation pressure directly regulated the masses of stars. What he meant was that real stars have gas and radiational pressure roughly equal at their centers (which would be true if they were made primarily of elements other than hydrogen). Radiation pressure came to look important again when Whipple (1946), Spitzer (1948a, 1948), Spitzer and Schwarzschild (1951), and Schwarzschild et al. (1951) began to ask how interstellar matter could be concentrated into clouds dense enough to form stars. The idea was that dust grains would shield each other from the general interstellar radiation field and so be pushed together. Later, when a dense enough cloud of grains had assembled, gravitation would kick in and add gas to the mix, resulting in the compositions of stars as they are actually seen. Spitzer and Schwarzschild (1951) and Schwarzschild et al. (1951) introduced the additional refinement that this was relevant only to population I stars, which therefore are metal rich because grain rich. Population II stars had been formed part and parcel with galaxies in the early universe, and so have only their fair share of heavy elements (which, remember, in those days were generally thought to be left over from cosmological or prestellar processes). Only after heavy elements had had time to condense into grains would the current mechanism of star formation kick in. Much more recently, such shielding from radiation pressure (of the cosmic background radiation) has been invoked, somewhat playfully, as a contributor to galaxy formation.
**The Parker Instability.** A very different way of assembling dense interstellar clouds, this is apparently truly the idea of E.N. Parker (1966) and does not favor dust over gas. Instead, part of the support of the gas layer of the galactic disk is attributed to magnetic fields (and perhaps cosmic rays). These are less dense than the gas and will, from time to time, pop loose, leaving clumps of gas behind (perhaps to be further organized and compressed by spiral density waves).

**Contraction.** This was, of course, the main energy source of 19th century stars, as per Helmholtz and Kelvin. Later on, the phase of contraction of a gas cloud to a star (which might then live on some other energy source) came to be seen as interesting in its own right. There are very qualitative treatments in Russell (1913, 1925) and Eddington (1926). More quantitative ones (considering also the process of fragmentation) were given by Ebert (1955), Bonner (1956), and McCrea (1957), partly in the context of condensing material in an expanding universe. That normal clouds would have too much angular moment to contract easily had been one of the reasons for doubting the reality of current star formation, and the problem of contraction of clouds with both angular momentum and magnetic fields (and whether the combination was better or worse than one by itself) was discussed by Spitzer (1948a), Mestel and Spitzer (1956), Burbidge and Burbidge (1958), and especially by Mestel (1965) and many others since. The first "modern" treatment of the contraction of an isolated cloud is generally credited to Larson (1972). A very important piece of information was missing from all early calculations of contraction and of what a pre-main-sequence star ought to look like. This is that cold, molecular gas becomes completely convective, leading to a unique equation state for the gas and unique temporal development of luminosity and effective temperature of a gas cloud of a given mass. The resulting Hayashi (1966) tracks continue to provide a reasonable fit to the locations of known pre-main-sequence stars in the HR diagram today.

**The Jeans' Mass and Fragmentation.** The Jeans' mass (or length) is the smallest one that will contract under its own self gravitation in a homogeneous fluid of given density and temperature. (Expansion of the substrate does not change the critical mass, but greatly slows the contraction.) Credit to Jeans, of course, customarily the 1929 edition of Astronomy and Cosmogony (Jeans 1929). A much cited paper down to the present is Hoyle (1953) on "hierarchical fragmentation," that is, the idea that as a gas mass contracts and radiates, it gets cooler and denser, and the Jeans' mass therefore gets smaller, and what was once the smallest self-gravitating body can then break into additional pieces, the process to continue until the fragments become optically thick or bored. The work of Lifshitz (1946), Ebert, Bonner, and McCrea is also relevant. Chandrasekhar and Fermi (1953) explored the effects of a uniform, oriented magnetic field, and discovered that the Jeans' instability was suppressed for a cylinder of gas representing a spiral arm if B were greater than 7 μG. A early modern treatment of fragmentation is that of Rees (1976),
who concluded that stars ought to contain a number of particles equal to the 3/2 power of the Dirac large number (the ratio, e.g. of electromagnetic to gravitational force).

Coagulation. Layzer (1963) believed that hierarchically fragmenting bits would get in each others’ way so much as to invalidate the entire process. That the fragments are likely at some point to start sticking back together seems plausible (Arny and Weissman 1973) and is part of many very recent calculations of how dense cloud cores become stars.

Triggered star formation. This is one of the ideas that has been in and out of fashion a number of times in the past 50 years (especially in connection with some event that might have initiated the formation of the solar system and simultaneously endowed it with short-lived radioactive nuclides). The traditional triggers are (a) supernovae (Opik 1953, who was concerned with the then-very-real problem of assembling dense gas clouds, and imagined expanding supernova remnants acting like giant brooms), revived in the radionuclide context by Cameron (1962, Cameron and Truran 1977), (b) HII regions expanding around OB associations (Oort 1954), which, like the SN case, is likely to give you a circle or shell of young, secondary stars, (c) cloud-cloud collisions (Kahn, 1953) (d) shocks in general (Dibai 1958), and (e) especially shocks associated with spiral arms in the density wave picture of spiral disks (Fujimoto, 1966, a paper apparently never published in either English or Japanese; Roberts 1969). The idea of stochastic, self-propagating star formation (Seiden and Gerola 1979, 1982) merges several of the possible triggers, so that star formation, once initiated, can go sailing across a good part of a galactic disk, getting twirled into an apparent spiral by differential rotation as it goes.

Bimodal star formation. This phrase has been used to describe a considerable range of ideas in which the stars that form under different conditions or at different times have different mass distributions or other variations. The particular form in which an early generation of stars left many massive (now very faint) white dwarfs behind and not much else apparently belongs to Schwarzschild (1954). He was thinking of dark matter for elliptical galaxies. Very recently, something similar has been suggested to account for MACHO micro-lensing events in the halo of the Milky Way.

L'ENVOI

The Eocene ("dawn of the recent") period in star formation is clearly marked by Spitzer's (1967) article in the Kuiper compendium (which was in press since 1962 or thereabouts). I began this discussion with the words of one anonymous poet and would like to end with the words of another. That he is unknown is not for lack of trying. J.D.G.M. (1943) signed only his initials. I was able to establish that he was not a contributor either to Observatory or to MNRAS at that time nor an IAU member. The current editors of Observatory and the
secretaries of the Royal Astronomical Society and the British Association of Amateur Astronomers have culled their minds and records without providing a firm identification. David Evans who was one of the four 1943 editors has some memory that the poem reached them through his fellow editor, Alan Hunter, and had been sent in by a Chancellor of Melbourne University (where the search is now being continued). The other two editors at the time were George C. McVittie and H.F. Finch (the only one of the four I never met). I quote the poem here for three reasons. It reflects a much higher level of doggerel than we usually see today; it demonstrates a clear belief in star and planet formation as ongoing processes, despite dating from the dark ages when such belief was rare, or at least rarely expressed; and I was born in 1943. It carried the title "III.", being preceded by two shorter meditations on solar and Jovian conditions of life.

Some time ago my late Papa
Acquired a spiral nebula.
He bought it with a guarantee
Of content and stability.
What was his undisguised chagrin
To find his purchase on the spin,
Receding from his call or beck
At several million miles per sec,
And not, according to his friends,
A likely source of dividends.
Justly incensed at such a tort
He hauled the vendor into court.
Taking his stand on Section 3
Of Bailey "Sale of Nebulae."
Contra was cited Volume 4
Of Eggleston's "Galactic Law"
That most instructive little tome
That lies uncut in every home.
"Cease" said the sage "your quarrel base:
Lift up your eyes to Outer Space.
See where the nebulae like buns,
Encurrupted with infant suns,
Shimmer in incandescent spray
Millions of miles and years away.
Think that, provided you will wait,
Your nebula is Real Estate,
Sure to provide you wealth and bliss
Beyond the dreams of avarice.
Watch as the rolling aeons pass
New worlds emerging from the gas:
Watch as the brightness slowly clots
To eligible building lots.
What matters a depleted purse
To owners of a Universe?"
My father lost the case and died:
I watch my nebula with pride
But yearly with decreasing hope
I buy a larger telescope.

ACKNOWLEDGEMENTS

I am most grateful to the following for comments, advice, and other assistance: David DeVorkin, Steven Dick, David S. Evans, Owen Gingerich, Richard Larson, Donald Osterbrock, Leonid Ozernov, Anneila Sargent, and David Stickland. I hope to turn this contribution into a semi-serious discussion of the history of star formation, and input of all kinds would be very much appreciated.

REFERENCES

7. Ambartsumian, V.A. 1960, QJRAS, 1, 152.
44. Fermi, E. 1949, PR, 75, 1169.
47. Gamow, G. 1938, PR, 54, 480.
57. Helmholtz, H. von. 1853, Lecture discussed by Kelvin (1863).
60. Herschel, Wm. 1784-1814, Phil. Trans. Roy. Soc., A74, 437 (1784) and see, Hoskin (1963) and A. Berry, 1898, A Short History of Astronomy (Dover 1961), p. 340.
77. J.D.G.M. 1943. Observatory, 65, 88 (reprinted with permission of the current editors).
105. Oort, J.H. 1954, BAN, 12, 177.
118. Russell, H.N. 1913, Observatory, 36, 290 & 324.
131. Shklovsky, I.S. 1949, A Zh, 26, 10 (in Russian).
138. Spitzer, L. 1946, PR, 70, 777.
140. Spitzer, L. 1948a, Harvard Observatory Monographs, No. 7.
146. Stromberg, B. 1937. quoted by Luyten and Hill 1937.
164. Wolf, M. 1908, Die Milchstrasse und die Kosmische Nebel.