THE ORIGIN AND EARLY HISTORY OF THE SUN AND THE PLANETARY SYSTEM IN THE CONTEXT OF STELLAR EVOLUTION

G. H. Herbig
Lick Observatory, Board of Studies in Astronomy and Astrophysics, University of California, Santa Cruz

ABSTRACT: A plausible scenario for the early history of the sun can be constructed by combining the results of stellar astronomy with lunar and meteoritic chronologies. The meteorites apparently contain material exposed to two nucleosynthetic events, one about 10^8 yr and another a few 10^6 yr before solidification. Following H. Reeves, these are associated with supernovae occurring in star clusters in molecular clouds that formed during passage through successive galactic arm shocks. The Orion Trapezium Cluster may be a modern example; its density is such that encounters between members would have been close enough and frequent enough to have had major effects upon their circumstellar 'solar nebulae,' as would recurrent FU Ori-like eruptions of the stars themselves. The lunar bombardment continued for 7 x 10^8 yr following formation of our sun. If this represented disk cleanup, disks must persist for that long, and hence circumstellar activity may still be in progress around some young stars in the solar vicinity. The observed time decay of axial rotation and surface activity in solar-type stars can be extended backwards, and indicates that the ultraviolet radiation of the young sun would have had major photochemical consequences upon the primitive earth.

About 40 years ago, recognition of the short nuclear lifetimes of OB stars and of the short expansion ages of OB associations forced astronomers to accept the formation of massive stars as an ongoing process that must be taking place today, before our eyes so to speak. No such obvious time scale urgency existed for stars of lower mass,[1] yet the accumulation of observational evidence since that time shows that young, lower-mass stars can indeed be found in large numbers in just the places where they would be expected, namely in molecular clouds and in many of the younger OB associations.

Richard M. West (ed.), Highlights of Astronomy, Vol. 6, 15–28.

https://doi.org/10.1017/S1539299600004962 Published online by Cambridge University Press

A considerable body of information has now accumulated for these T Tauri stars, whose masses are believed to range up to about 2-3 times that of the sun. It is tempting to reflect upon their properties for a larger purpose: what they may be able to tell us about the early history of stars like the sun. A further possibility is that they might also be able to shed some essential light upon the circumstances that led to the formation of the planetary system.

Over the past 20 years, there has grown up another substantial body of evidence that certainly bears upon the early history of the solar system, namely the physical evidence from meteorites, from returned lunar samples, and from planetary astronomy. A third, quite independent field that is currently very active is the numerical modelling of cloud collapse and of the evolution of the resulting disk-shaped nebulae that may exist around young stars. All of these endeavors, unless we are completely mistaken, must illuminate in some way the same central issue. I do not think it is premature to speculate upon how they might all fit together, the goal being to find mutual constraints and insights. In that spirit, I propose to describe a scenario for the origin and early history of the sun and its planetary system, drawn from these diverse fields of investigation, that seems to account for much of the compelling information that is on hand today.

You will recognize that much of this picture is not original with me, and I would simply acknowledge that it owes much of its merit to the ideas of Hubert Reeves, to evidence from ancient radioactivities painstakingly gathered by Reynolds, Wasserburg and many other geochemists, and to the work of planetary scientists like Hartmann, Wetherill and a host of others.

One can begin with an interpretation, put forward originally by Reeves (1978) (see also Schramm 1974), of the origin of two of the extinct radioactivities in meteorites: Fig. 1. A massive interstellar cloud (a Giant Molecular Cloud in today's terminology) is formed by compression of interstellar gas as it flows into a spiral shock. Stars — let us call them the first generation — form in the compressed material, and not long thereafter Type II supernovae (or possibly some other high-energy events) contaminate the gas of the cloud with nuclear debris, among it ^{129}I (half-life 17 x 10^6 yr) and ^{244}Pu (82 x 10^6 yr). Observations of GMCs show that their lifetimes, in the sense of being detectable in CO surveys, are much less than the time required for this gas to reach the next spiral shock; lifetimes of roughly 30 x 10^6 yr are often mentioned (Blitz and Shu 1980; Cohen _et al_. 1980; Boulanger _et al_. 1981). Therefore this gas, contaminated with decaying radioactivities and stars of the first generation, reexpands and streams into the interarm region. There it must be diluted by mixing with material having different histories, but after about 1-2 x 10^8 yr, some is recompressed again in the following arm. Note that this value of the waiting-time is forced upon us by the rates of decay of ^{129}I and ^{244}Pu, not by appeal to galactic structure as it is today.

https://doi.org/10.1017/S1539299600004962 Published online by Cambridge University Press

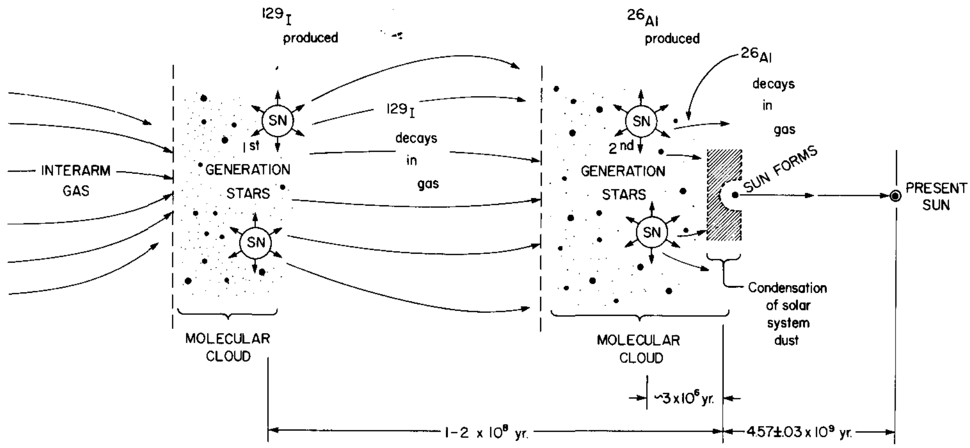

Fig. 1. Schematic diagram of the flow of interarm gas (from the left) into the first spiral shock (vertical dashed line), the formation of the first generation of stars, and the subsequent contamination of the gas of that molecular cloud with long-lived radioactivities (as ^{129}I and ^{244}Pu). The cloud reexpands into the following interarm region, and 1-2 x 10^8 yr later some of the gas is recompressed into a second molecular cloud (second vertical dashed line). Again, a generation of stars forms and their supernovae create ^{26}Al which quickly (after a few million years) is incorporated into dust grains and the solar nebula. Some 4.5 x 10^9 yr later (far right), the daughter products of these two nucleosynthetic events are found in meteorites.

This recompression of some of the original gas into a new GMC results in a second generation of stars, the massive members of which become supernovae, and among the short-lived products of their ejecta is ^{26}Al (half-life 0.7 x 10^6 yr), which after a few million years in the gas, is itself frozen out in certain special dust grains. The 0.1% of the original ^{129}I which remains after 10^8 yr also goes into the dust but perhaps because it pervades the entire cloud, is distributed much more uniformly. It is then, from gas mixed with dust in which the decay products ^{26}Mg and ^{129}Xe tell this story, that a certain condensation in that GMC could have begun its contraction which was to lead to the formation of our sun and planets.

One recognizes here that if this sequence of events is correct, it must be concluded that the sun could not have been born in isolation: the necessity of nearby supernovae from massive stars points to a spectrum of masses, as well as a second-generation star cluster having a significant spread in ages. It seems unlikely that any stars of the first generation would be caught up in the formation of the second GMC. This idea that the sun formed in a rich star cluster has interesting

https://doi.org/10.1017/S1539299600004962 Published online by Cambridge University Press

consequences, to which we shall return. Of course, one can say nothing of the subsequent history of the other cluster members. The overall lifetime of ordinary galactic clusters today is a few 10^8 yr, so the brothers and sisters of the sun have long since been lost in the field.

I must point out that nothing has been said as to how the gas goes so quickly into dust in that GMC, as is demanded by the ^{26}Al decay, but clearly there is no time to manufacture that dust by cycling gas through red giants or planetary nebulae or any other low-mass post-main sequence objects. Another issue that was not faced is why one should assume that the gas and dust that went into that first-generation GMC was pure and pristine, in the sense of no prior nucleosynthetic or chemical inhomogeneities at all. It would be unreasonable to suppose such a simplistic starting-point. It may well be that the complex story told by the stony meteorites reflects such still earlier history; that is, that our "first generation" event began with a slate not wiped at all clean. This view has been vigorously argued especially by D. D. Clayton (1979), and I think deserves respectful attention. It implies, of course, a somewhat different chronology than described here; for a survey of the subject by a detached reviewer, see Begemann (1980).

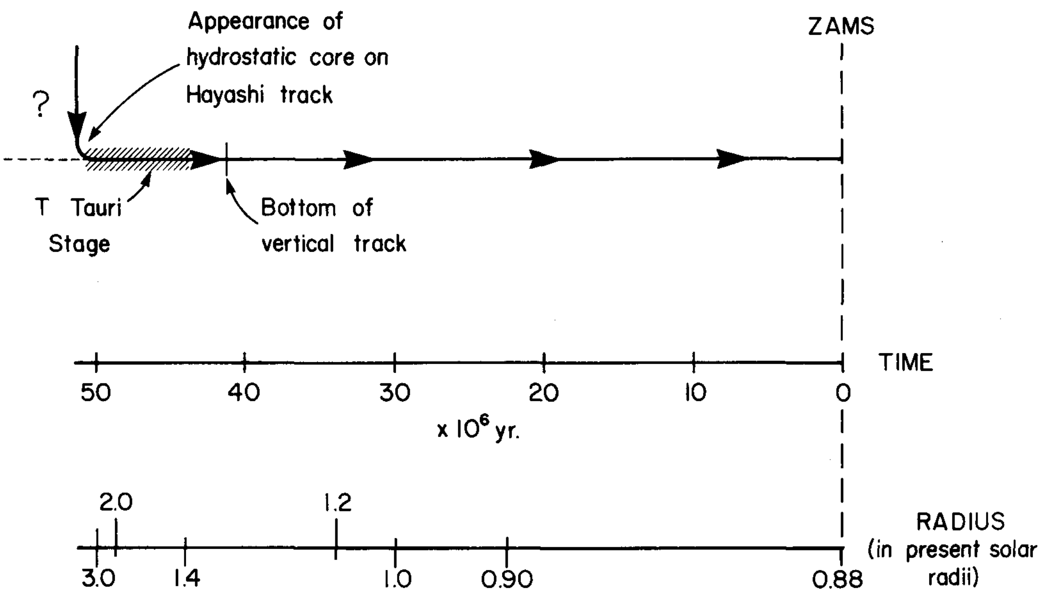

Fig. 2. Schematic representation of the movement of a 1 $M_\odot$ star along its theoretical Hayashi track (upper solid line). The time scale of contraction to the main sequence and the change in radius are marked on the center and lower scales.

https://doi.org/10.1017/S1539299600004962 Published online by Cambridge University Press

Let us now turn to the predictions of stellar models. The observed concentration of T Tauri stars in that portion of the H-R Diagram occupied by the evolutionary tracks of contracting stars (Cohen and Kuhi 1979) is one of the reasons for identifying these objects with stars of up to about 3 solar masses in the early hydrostatic phases of contraction toward the main sequence. Fig. 2 displays the theoretical pre-main sequence history of a 1.0 $M_\odot$ star as a track in the H-R Diagram. The solid line represents the Hayashi track of the hydrostatic core, which it is now believed will become visible only when its opaque, collapsing, outer dust envelope is completely accreted. Thus a real star, in the sense of what we can detect optically from afar, will in effect enter upon this track at some point that is determined by the circumstances of the later stages of accretion, and probably by the aspect angle as well. These are matters of which we know very little; the particular entry point to the Hayashi track indicated in Fig. 2 represents only an illustrative, spherically symmetric calculation by Stahler et al. (1981). But if the sun evolved in this way, when it first became optically luminous it had already shrunk far inside the present orbit of Mercury (at 83 $R_\odot$).

At this time, it is believed that the sun was surrounded by a flattened disk of gas and dust that had in effect, been left behind during the slow shrinkage of the original interstellar cloud as its core accreted to form what was to be the sun. The early low-density history of this disk, and of its subsequent agglomeration into the planetary system are the subjects of an extensive theoretical literature, which I shall not attempt to describe. For the moment, let us simply assume the existence of such a disk, having a radius of a few hundred a.u.

There is a troubling feature about the present planetary system, regarded as the residue of that primeval solar nebula, namely that the total angular momentum vector of the planetary orbital motion is now inclined at 6° to that of the sun's rotation. Yet in the case of the Galilean satellites of Jupiter, which must have originated in a similar way during the contraction of proto-Jupiter, that obliquity is very small: the inclination of their orbits to the equatorial plane of the planet are all less than 0°.5. In the case of the inner satellites of Saturn, no inclinations are greater than 1°.5. A possible explanation of the non-zero solar obliquity may lie in a consequence of the meteoritic evidence that I have described: namely that the sun was born not alone, but in a rich star cluster. This has revived encounter theories of the origin of the planetary system, as put forward by Jefferies and Jeans about 60 years ago, and more recently by Woolfson (1978). One important point made by Woolfson (and later by Kobrick and Kaula 1979) must be taken seriously in the present context: that in such a star cluster, close encounters between members may be frequent enough to have profound effects upon the 'solar nebulae' presumably carried by most or all of those young stars. The difficulty in being very specific about such effects has in the past been the lack of information on star densities in very primitive clusters.

Fig. 3. The Trapezium Cluster of the Orion Nebula, photographed in near-infrared light. The original negative was obtained with the 120-inch reflector, Lick Observatory.

https://doi.org/10.1017/S1539299600004962 Published online by Cambridge University Press

Some such information is now available for one interesting case: the concentration of faint pre-main sequence stars around the Trapezium in the Orion Nebula (Fig. 3). That concentration is revealed to us thanks to the OB stars of the Trapezium, which have cleared the dust and gas out of a small volume of about 0.1 pc^3 in the near side of that molecular cloud. The number of stars in that pocket corresponds to a space density of 500-600 stars pc^{-3} down to $M_V = 6.6$, or brighter than about 1/5 the solar luminosity (Herbig 1982). This is about 2 orders of magnitude higher than the central star densities of older galactic clusters. Since the velocity dispersion in the Trapezium Cluster is known to be about 5 km sec^{-1} from astrometric studies (Strand 1958, Fallon 1975), it is possible to calculate the average interval between encounters of the optically detectable members of the Trapezium Cluster. For passages within 100 a.u., that interval is only a few 10^6 yr. Therefore if the group survives for the representative lifetime of 10^8 yr that is inferred from cluster statistics (Wielen 1971), then essentially all its members will have suffered a near-miss at a separation comparable with the diameter of the present planetary system.

One cannot be sure that the sun formed in a cluster as dense as the Trapezium. But if so, such a close passage would surely have a major effect upon not only the gross structure of the disks that by that time have formed around each star, but upon the thermal history of the solids within those solar nebulae. It could be that the non-zero solar obliquity is an artifact of such an encounter, an idea first proposed by Mottman (1977).

Direct observations of young stars show that there are other violent events, beside such close encounters, which take place on the stars themselves that will certainly interfere with the tranquil consolidation of the inner solar nebula. The most spectacular are the so-called Fuor eruptions, the best example being that displayed by the former T Tauri star V1057 Cygni which flared up in 1969 and now, 13 years later, is in decline back toward its former subdued state (Fig. 4). There is some reason to think that such outbursts are repetitive, at intervals of perhaps 10^4 yr. The evidence now on hand indicates that, at the peak of such an eruption, the star is surrounded by an envelope of ejected gas in which the opacity mimics a hot photosphere having a radius of 12-15 $R_\odot$, much larger than the maximum size of the same star during its original Hayashi-style contraction (Fig. 2). Such a flareup would vaporize dust within the inner solar nebula, and the high-velocity gas shells that the star is observed to eject at the same time would certainly scour out of the same volume any material that was not already consolidated into large lumps. It has recently been suggested (Mundt 1981, Feigelson 1982) that among the T Tauri stars lesser events of this kind are even more frequent. In view of all this violent activity that either we observe or can anticipate around very young stars, it is not surprising that the stony meteorites, regarded as witnesses of those long-ago times in our own history, tell such a story of being shocked, or remelted, or accreted from a complex mix.

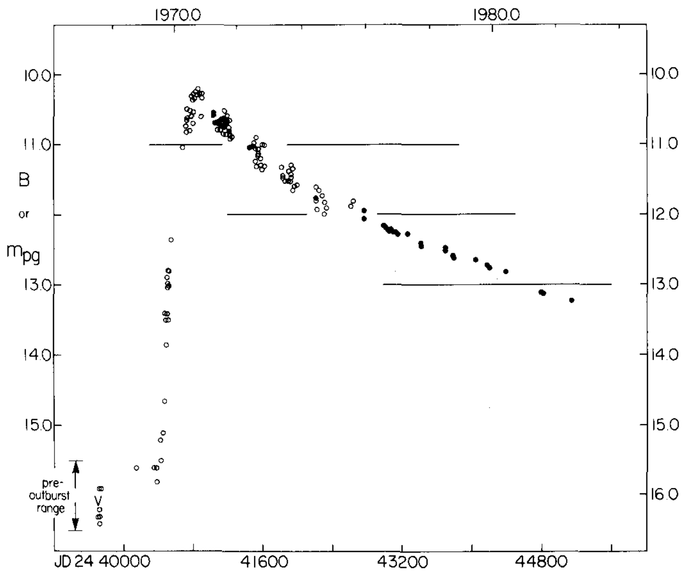

Fig. 4. The light curve of V1057 Cyg, from shortly before the 1969 rise to maximum until mid-1982. The ordinate is photoelectric B magnitude (filled circles) or photographic m_{pg} (open circles). The curve is based on all published data, together with unpublished Lick observations by R. P. S. Stone.

Let us now try to connect these results from stellar astronomy with the early history of the local part of the solar system through the chronology of the lunar surface, as reconstructed largely through returned lunar samples.

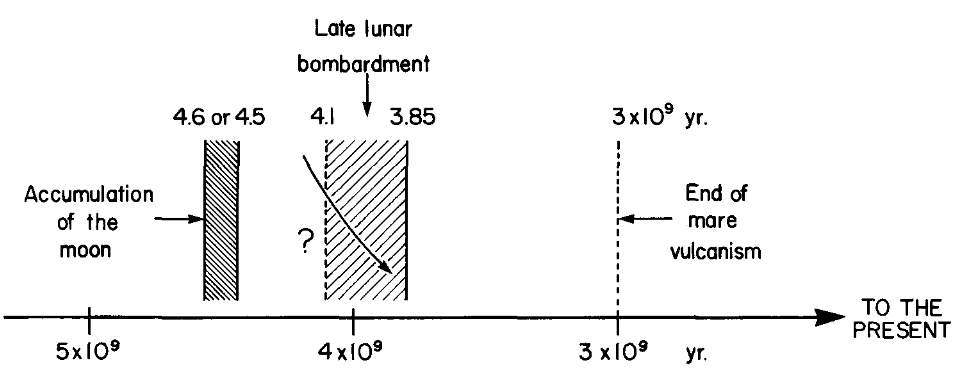

Fig. 5. Schematic representation of the history of the lunar surface, as inferred from lunar samples.

Fig. 5 shows the essential features: a brief period of accumulation, with a period of heavy bombardment of the surface which effectively ceased about 3.8 x 10^9 yr ago. (There is some disagreement whether this bombardment consisted of more than one spike of heavy infall, or a continuous tapering-off as suggested in this Figure.) The time of onset of the bombardment is not directly determinable on account of the heavy churning-up of the lunar surface in the later stages. One would like to connect this history, whose time scale and zero is set by the decay constants of long-lived radioactivities, with the history of the sun as predicted by solar evolutionary models. I am told that the age of the sun, as determined by a fit to its present T_e and radius, is indeterminate to at least 1 x 10^9 yr, largely on account of uncertainties in composition. If one makes the usual assumption that the sun arrived on the ZAMS 4.5 x 10^9 yr ago, it is then possible to superimpose

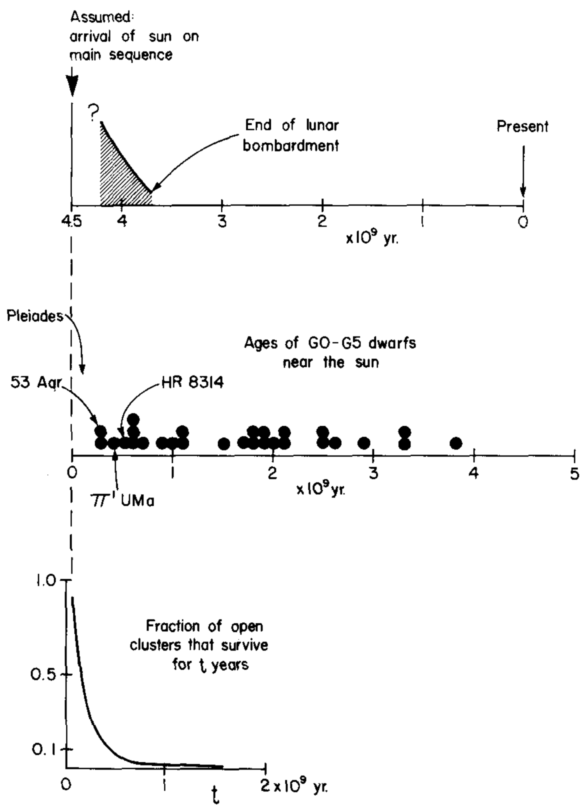

Fig. 6. Top: the chronology of the lunar surface, from Fig. 5.
Middle: the ages of G0-G5 main sequence stars in the solar neighborhood as inferred from their surface lithium abundances, with zero point set at their arrival in the main sequence. These two diagrams are shifted horizontally into coincidence by the assumption that the sun arrived on the main sequence 4.5 x 10^9 yr ago.
Bottom: the time scale of disintegration of galactic clusters after Wielen (1971).

https://doi.org/10.1017/S1539299600004962 Published online by Cambridge University Press

the lunar history upon the ages of solar-type stars in the solar neighborhood referred to that zeropoint (Fig. 6), as inferred from their surface lithium abundances (Duncan 1981). One should not be surprised that even the youngest of these stars are not members of young star clusters: the lowest curve of Fig. 6 shows how rapidly such clusters dissolve.

One sees that if the lunar history is applicable to other stars 2
of about 1.0 $M_{\odot}$, then the consolidation and cleanup of a solar nebula continues for about 7 x 10^8 yr, long after the star has reached the main sequence. In other words, there are some nearby stars that are young enough that something interesting, in this sense, may still be happening near them.

Is there any chance that at distances of 10-20 parsecs we could detect any direct or indirect signs of this activity, or at least of the survival of a solar nebula? Several possibilities occur to one:
1) Large quantities of circumstellar dust should be detectable by its infrared emission, or by its scattered light. These are well-known phenomena in the case of pre-main sequence stars, but as far as I know, no critical polarization study or infrared examination has been made of young stars on the main sequence.
2) Could there be direct detection of any of the phenomena of cleanup of the 'solar' nebula; for instance, might the impact of asteroidal-size bodies on planetary surfaces be observed? It is estimated that the projectile which produced Mare Imbrium arrived with a kinetic energy of about 10^{33} ergs (Wetherill 1975), but direct detection of such flashes would seem unlikely: they would be too rare, and the amount of detectable radiant energy released quite uncertain.
3) One thinks of the large quantities of icy material that may exist in any new planetary system, of the kind that is often postulated to have supplied our own comet cloud. In our own case, an icy comet suitably near the sun produces a coma of H and OH from the photodissociation of evaporated H_2O by ultraviolet sunlight. The much more intense ultraviolet radiation field of a young star should make the process more effective. The presence of an H_2O maser at T Tauri shows that gaseous H_2O does exist near some young stars. But an optical search has been made for photodissociated OH around T Tauri stars, and none was found.

Therefore, we have as yet no direct proof that such circumstellar debris survives around young main sequence stars. But very clearly, the stars themselves carry for a long time some memory of their turbulent youth. Fig. 7 shows, on the left of the ZAMS line, the pre-m.s. track of a 1.0 $M_{\odot}$ star, and on the right (with the time axis compressed by a factor 50) the 4.5 x 10^9 yr since — we assume — the sun arrived on the main sequence. There is a clear decline in the level of detectable mass ejection during passage through the T Tauri stage, but whether that can properly be connected with the sun's present mass loss via the solar wind, at a level 10^6 times lower, is uncertain. It is a pity that no direct spectroscopic means has been found to detect mass ejection from solar-type stars during the greater part of their lives. One does see

https://doi.org/10.1017/S1539299600004962 Published online by Cambridge University Press

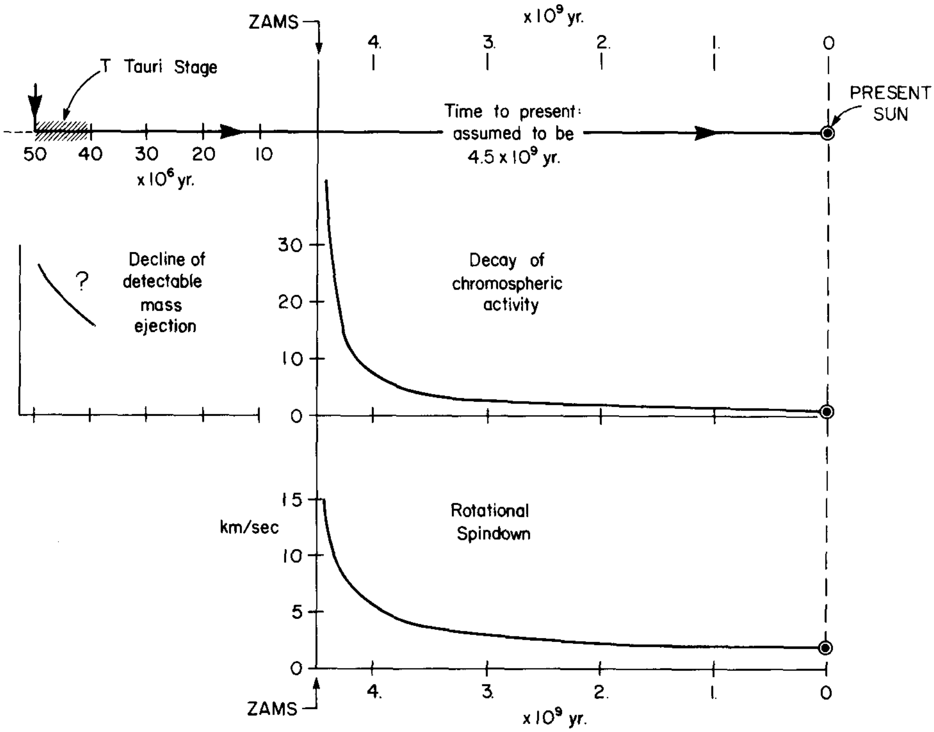

Fig. 7. The decay with time of properties of solar-type stars. The section to the left represents contraction and decline in the level of detectable mass ejection prior to arrival on the main sequence. The section to the right (with time scale compressed by a factor 50) shows the decline with age in the level of surface activity of solar type stars (from the ultraviolet lines of chromosphere and transition region), and of the equatorial velocity of rotation. In both of the latter diagrams, there is substantial scatter of individual stars around the average curves shown.

a clear decline of the speeds of axial rotation of solar-type stars as they age, which is believed to be evidence for magnetic braking through the transfer of the star's angular momentum to ejected material (Soderblom 1982). There is an equally striking statistical decline in the surface activity of stars as they age. Figure 7 also shows the (average) decay with time of the energy emitted in ultraviolet chromospheric and transition-region emission lines. In this case, it appears that the rise toward younger ages carries across the ZAMS and connects with the line intensities observed in the pre-main sequence T Tauri stars (Boesgaard and Simon 1982).

https://doi.org/10.1017/S1539299600004962 Published online by Cambridge University Press

The extension of such curves backward in time tempts one to conjecture upon the consequences of such an active sun upon the primitive earth. This enhanced ultraviolet radiation, at the level observed in the T Tauri stars, would certainly have had an effect upon the terrestrial atmosphere. If there had been an original atmosphere of material accreted from the solar nebula, and hence essentially of solar composition, Sekiya et al. (1981) have calculated that the heating through reactions that begin with the photodissociation of H_2O would have been sufficient to dissipate that atmosphere in about 20 x 10^6 yr (if there had been no significant screening by interplanetary material), that is long before the sun reached the main sequence. Whether such an original H-rich atmosphere existed or not, it is believed that an atmosphere would have formed by outgassing from the planet. Canuto et al. (1982) have investigated the response of such a H_2O and CO_2-rich atmosphere to the flux of the sun in the 1200-2000 Å region, estimating that from the T Tauri stage through arrival on the main sequence the ultraviolet excess declined from about 10^3 to perhaps 50 times its present value. The result would be that the concentrations of O_2 and O_3 at all altitudes would by 10^3 to 10^4 times higher than at the present. This would have major implications for the geochemistry of the primitive earth as well as for biology, because it would provide free oxygen long before photosynthetic organisms appeared, and because the ozone would screen primitive organic material from the lethal solar ultraviolet.

Just a word upon what is clearly the basic assumption of this Discourse. Obviously, I believe that there is no scientific reason to claim that the circumstances which produced our planetary system were highly unusual or unique. Hence it is fair to attempt to connect the planetary and meteoritic record with the processes and activities that we see taking place during the early evolution of solar-type stars. The event that took place here 4 1/2 billion years ago was governed by the same physics and chemistry and dynamics that operate in the Taurus clouds and in the Orion Nebula tonight, and the overall astronomical conditions should not have been significantly different.

Yet it would be unrealistic to claim that this scenario can be in any sense firm or final, on account of the rapid and inexorable advance of knowledge. In explanation of having even attempted such a synthesis, may I simply repeat the words of a famous Athenian of the Fifth Century B.C. who, in comment upon a historical record just as flawed and incomplete and arguable as is the astronomical, said:[3] "we can rest satisfied with having proceeded upon the clearest data, and having arrived at conclusions as exact as can be expected in matters of such antiquity."

FOOTNOTES

1. Despite the detection of low-luminosity pre-main sequence members of the expanding associations Per OB2 and Lac OB1 (Herbig 1954a, 1954b).

2. Mottmann (1977) suggested that the 'solar nebula' of a passing member of the sun's star cluster was the source of the late lunar bombardment. The duration of such an encounter would be so short that only a brief spike of infall would be expected; under the circumstances assumed here for encounters in the Orion Trapezium cluster, a duration of no more than about 100 years would be expected. Mottmann's proposal also requires that the cluster and at least one other solar nebula in it survive for 7 x 10^8 yr. Wielen's (1971) results suggest that only about 5% of ordinary galactic clusters exist as such for so long.

3. Thucydides, The Peloponnesian War (Modern Library transl.), book I, 21.

REFERENCES

Begemann, F. 1980, Rep. Prog. Phys. 43, 1309.
Blitz, L. and Shu, F. H. 1980, Astrophys. J. 238, 148.
Boesgaard, A. M. and Simon, T. 1981, 2nd Cambridge Workshop on Cool Stars, Stellar Systems and the Sun (Smithsonian Astrophys. Obs., Spec. Report 392), 2, 161.
Boulanger, F., Stark, A. A. and Combes, F. 1981, Astr. Astrophys. 93, L1.
Canuto, V. M., Levine, J. S., Augustsson, T. R. and Imhoff, C. 1982, Nature 296, 816.
Clayton, D. D. 1979, Space Sci. Rev. 24, 147.
Cohen, M. and Kuhi, L. V. 1979, Astrophys. J. Suppl. 41, 743.
Cohen, R. S., Cong, H., Dame, T. M. and Thaddeus, P. 1980, Astrophys. J. 239, L53.
Duncan, D. K. 1981, Astrophys. J. 248, 651.
Fallon, F. W. 1975, Stellar Motions in the Orion Nebula Cluster (Thesis, Univ. Florida).
Feigelson, E. D. 1982, Icarus, in press.
Herbig, G. H. 1954a, Publ. Astr. Soc. Pacific 66, 19.
Herbig, G. H. 1954b, Astr. J. 59, 342.
Herbig, G. H. 1982, Proc. N. Y. Acad. Sci., in press.
Kobrick, M. and Kaula, W. M. 1979, Moon and Planets 20, 61.
Mottmann, J. 1977, Icarus 31, 412.
Mundt, R. 1981, 2nd Cambridge Workshop on Cool Stars, Stellar Systems, and the Sun (Smithsonian Astrophys. Obs., Spec. Report 392), 2, 18.
Reeves, H. 1978, Protostars and Planets (ed. T. Gehrels), p. 399; this paper expands upon the original proposal of 1972. See also B. M. P. Trivedi: 1977, Astrophys. J. 215, 877.
Schramm, D. N. 1974, Ann. Rev. Astr. Astrophys. 12, 383.
Sekiya, M., Hayashi, C. and Nakazawa, K. 1981, Prog. Theor. Phys. 66, 1301.

https://doi.org/10.1017/S1539299600004962 Published online by Cambridge University Press

Soderblom, D. R. 1982, Astrophys. J. Suppl., in press.
Stahler, S. W., Shu, F. H. and Taam, R. E. 1980, Astrophys. J. 241, 637.
Strand, K. 1958, Astrophys. J. 128, 14.
Wetherill, G. W. 1975, Proc. 6th Lunar Sci. Conf., p. 1539.
Wielen, R. 1971, Astr. Astrophys. 13, 309.
Woolfson, M. M. 1978, Origin of the Solar System (ed. S. F. Dermott) pp. 179, 199; these are reviews which contain references to the original papers.