UC Irvine
UC Irvine Previously Published Works

Title
Origin of Stars and Planets: The View from 2021

Permalink
https://escholarship.org/uc/item/89p7w9hk

Author
Trimble, Virginia

Publication Date
2002

DOI
10.1007/10856518_63

Copyright Information

Peer reviewed

Origin of Stars and Planets: The View from 2021

Virginia Trimble

Dept. of Physics, Univ. of California, Irvine, CA 92697, USA
and
Dept. of Astronomy, Univ. of Maryland, College Park, MD 20742, USA

Abstract. What (or who) made the Earth, and how? What made the Sun, and are there others? Thus mankind has wondered for eons. Not surprisingly, the answers have evolved a good deal over the decades, centuries, and millenia. So, perhaps more importantly, have the meanings of the questions. Here is an attempt to explore, "as far as thought can reach", both back into the past and forward into the future, the territory defined by such questions, illuminated by light (and infrared) from the present workshop. If you decide to come along on the quest, it is probably worth carrying in your baggage the related, if strange-sounding, question, is star formation important?

1 Introduction

Thirty-some years ago, Jesse L. (for Leonard) Greenstein, who will appear again in a few pages, instructed the women graduate students in astronomy at Caltech not to use their middle names or initials, on the grounds that the space would soon be needed for our "maiden" names or initials, when we married and adopted our husbands' surnames. (His wife was born a Kittay.) It will not surprise you to hear, if you know any of us, that Judith G. (for Gamora) Cohen, Donna E. (for Etta) Weistrop, and I all immediately began using our middle names and initials on any and all possible occasions. A decade or two later, it dawned on us that there were not really enough astronomers in the world to make this necessary, and (as you can check by looking at our recent papers), we mostly stopped using said middle initials, as incidentally did Allan Sandage, who has not inserted R. (for Rex) in many years.

Nevertheless, it remains true that I was born Virginia L. (for Louise) Trimble, and so share the initials of the VLT, whose present and future accomplishments we gathered to celebrate. Whether this is other than the purest of coincidences[1], only The Shadow (no middle initial, but J. for Jansen would not be inappropriate) knows. It was clear from the chair's introduction of this presentation that the correspondence had at least been noted by the SOC.

An alternative explanation for my being asked to give this talk can be found on a postcard [4], which says, "You can get a great reputation for wisdom just by telling people what they already know". This is precisely the function of the concluding speaker or reviewer of a conference, and one which I have been exercising for at least 25 years [18] or perhaps 50, if you count appearances as narrator in the annual Christmas pageant at Toluca Lake Grammar School.

---

[1] The cube of 26 is, after all, a large number. On the other hand, at least three astronomical entities share the acronym AIC [20].

2   The Beginnings

"Vaya'as Elohim et shney hame'orot 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 hakokhavim. And also the stars. Vayehi erev, vayehi boker, yom r'vi-i. And there was evening, and there was morning. A fourth day." This, from the author of the first chapters of Genesis, is the earliest published description of star formation.

There are three things to be noted about this scenario. First, all the stars will be the same age. Second star formation is not really a separate, important phenomenon, but is more or less incidental to the creation of the universe and the formation of the solar system. We will encounter both ideas again. Third, stars did not appear until the fourth day, and so a four day conference is barely long enough to solve the associated problems.

There is a contrasting tradition, at least as old, in which star formation is an on-going process, most familiar perhaps in the form of Greek myths, in which various heroes and demi-gods are placed among the stars as a reward for meritorious behavior. The ancient Egyptians also held this view, with the souls of Pharaohs being entitled to ascend to the heavens and join the stars, and shafts out of their tombs were sometimes provided for the purpose of facilitating their journey to the Indestructable, or circumpolar, constellations [2]. Frequently-reproduced texts describe the events, e.g. *Sinuhe*, "The King of Upper and Lower Egypt (Sehetepibre) flew up to heaven.", and, a sort of converse, in the *Story of the Shipwrecked Sailor*, "A star fell, and they (75 serpents, children and brothers and a little daughter) went up in flames through it."

The Middle Egyptian word for star, conventional pronunciation sebah, (like writers of Hebrew and Arabic they generally left you to guess the vowels), consists of five glyphs, the individual ones for s, b, and aleph, then a five-pointed star, with one point straight up and looking very much like a modern child's drawing, and last the determinative meaning sun, which is a circle with a dot in it, like the medieval astrological symbol and our subscript in solar mass, $M_\odot$, etc. The short form is just the star, sometimes with the dotted circle at the center, but if the idea "stars are suns" ever occurred to any of them, I have not seen it. Other words with the same pronunciation but different meanings often include the star. Thus it developed into a triliteral, implying sound. Such words have an additional determinative; a simplified house for sebah = door, a man carrying a stick in threatening fashion for sebah = teach, and a sort of arc and compass for sebah = surveying instrument.

We can, therefore, identify two questions with very long pedigrees. First, is star formation an on-going or one shot process? This was not finally settled until about 1950. And, is star formation important? If you think the answer to this is a hearty, unambiguous "yes", take a look at how many pages it occupies in the reference and review volumes of Cox [7], Lang [13], and Bahcall and Ostriker [3] and in issues of *Annual Reviews* since the paper of Shu et al. [17] appeared.

There is a comparably old question connected with the formation of planets (explored from Greek times down to the present by Brush [5] and references therein), phrased most succinctly as monadic vs. dyadic. That is, did our (and

by extension other) solar systems arise from a single entity like the Swedenborg/Kant/Laplace rotating nebula or out of some interaction between two or more previously existing entities, as in the Buffon/Chamberlin-Moulton/Jeans-Jeffreys-Lyttleton tidal encounter scenarios?

In case you might have forgotten, some form of tidal encounter picture was the main-stream, respectable point of view from the 1890s until the 1940s. How did they manage to get it so wrong, to the point where Russell et al. [14] (in the textbook that introduced English speakers to astronomy for two generations) put "The Nebular Hypothesis" in small print and "The Hypothesis of Dynamic Encounter" in large type? They were misled by the concentration of angular momentum in planetary orbits around a slowly-rotating sun, and concluded that "the angular momentum of the planets must have been put into the system from outside". In other words, there was important physics – magnetic fields and the solar wind – missing from their world view.

If half of all three-sigma results turn out eventually to be wrong (which is roughly so), it is very frequently just because there is physics missing from the discussion, and John Bahcall has used this as part of his private definition of what is really meant by a 3σ result – if it is wrong, we have left out something important. The next section looks at the sense in which this might still be a problem in understanding planet and star (vs. galaxy and large scale structure) formation.

Incidentally, the monadic/dyadic issue is still with us, in the form of the question of the extent to which cloud contraction and star formation are triggered. White described observational evidence for triggering in S185 and the Eagle. Triggering by a nearby supernova was assumed in the talk by Adams on the sun's parent star cluster. But poster presenter Zavagno mentioned that she had been warned off giving too much emphasis to triggered star formation in her discussion of Sharpless 128, on the grounds that it was not currently fashionable.

3   Missing Physics

There seem to be four issues here, three of which affect calculations of star and planet formation one way or another. First is physics that you already know is likely to be important, but do not or cannot yet include in your code for lack of space or time. Magnetic fields in N-body simulations (of nearly anything) are an obvious example. Bate's talk addressed the problem of dynamic range: it takes many particles to watch one protostar and its disk form, limiting substructure in his calculations to things bigger than about 1 AU and clouds to at most 50 $M_\odot$. Thus he forms no 100 $M_\odot$ stars. Heating and cooling of gas by dissipation and by the turn-on of the first stars are often also not treated adequately in simulations of formation of galaxies and large scale structure, where the dynamic range problem is also severe, with smaller galaxies sometimes represented by only about one particle. Among the ways of tackling missing physics in this sense are (a) deliberately make your code worse in resolution or some other parameter and see how much difference it makes, (b) draw on observations for

your initial conditions (talks by Alves, Kramer, Lombardi, Walmsley), and (c) wait for computers to become faster.

Second is essential physics that is currently missing from human knowledge, not just from your calculation. I believe star formation does not suffer from this. In contrast, models of large scale structure and streaming must begin with assumptions about the kinds and amounts of dark matter, dark energy, topological singularities, and so forth.

Third is the curse of the variable or adjustable parameter, sometimes including the initial conditions for a simulation. Gamow is supposed to have said that, with five parameters, you could fit an elephant. This is roughly true if you think of a Fourier decomposition of the outline. Burkert, speaking on formation of binary stars, remarked that, historically, what you got out was what you put in, and this was to a certain extent still the case.

All scenarios for galaxy formation have this problem in spades – there are infall and outflow of primordial or enriched gas, star formation as a function of gas density, composition, and whatever else matters, the initial mass function and its dependence on gas density, composition, etc., and, at some level (especially if you care about nucleosynthesis by Type Ia supernovae) binary star populations to be considered. Notice that, in principle, solutions to some of the problems in understanding star formation discussed at the workshop would be, in turn, improved input for galactic calculations.

Fourth is dimensionality. Klessen noted that star formation logically proceeds from 3-d (turbulence) to 2-d (disks) to 1-d (main sequence stars). In practice, the physicist's first approximation is the spherical cow. The next step is, of course, the bipolar cow (Fig. 1), and Shu et al. [17] noted as one of the principal advances of the decade preceding their review the recognition of the ubiquity of bipolar outflows in young stellar objects. The talks by Bourke, Bacciotti, Garcia,
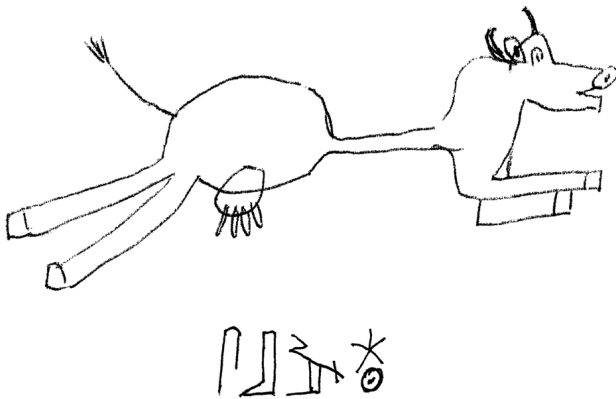
Fig. 1. The bipolar cow (F.C. Adams) jumps over the hieroglyphs for star (in a modern Egyptological hand, V. Trimble)

and Romaniello and many of the posters made clear that such outflows are even more in evidence (and less purely bipolar) when your angular resolution and wavelength coverage improve. The thrust of Li's talk was that it is possible to model what we see, assuming suitable amounts of rotation and magnetic fields, but that this is not quite the same as understanding what causes the outflows, or, in my opinion, how they know how much rotation and field need to be removed, and when, to leave the range of ZAMS stars we see (Wring, Hogerheijde).

Disks are the other extreme of two-dimensionality. The talks of Brandner, Lada, Launhardt, and Wilner showed that disks are like belly buttons, that is, everybody has one when young, but they are not all the same shape.

Finally, it came as a surprise to hear from Clarke that disks may be more difficult to get rid of on the right time scale than they were to form in the first place. Her discussion left many of us uncertain about whether this was a case of "all of the above" (we will meet more shortly) or a case of physics missing from the calculation, in the sense of the dominant process not yet having been identified. Some of this physics may be preserved in our own Kuiper belt (talk by Jewitt). Or perhaps there is an octupole cow lurking out there somewhere.

4   Some More Recent History

William Herschel, beginning in 1789 and perhaps guided by Kant's and later Laplace's nebular hypothesis, arranged the various sorts of nebulae he had observed and drawn into what he regarded as an evolutionary sequence, beginning with the diffuse and/or planetary nebulae, going on to bright/dense stellar nebulae, then to blurry or hazy stars and clusters, and finally to isolated stars and clusters. Well, he had some of the details wrong, but he was sure that star formation was an on-going process that observers could reasonably study, and so essentially right.

Another William, Huggins, who in 1865 was the first to see emission lines from some of Herschel's nebulae (and thus prove them to be truly diffuse, not just unresolved star clusters) began to gum up the works when he declared that Orion and other bright nebulae could not be Herschel's "nebulous fluid" from which stars could be built, because he could see only a hydrogen, a nitrogen, and one other line, while uniformity of nature required the presence of all the elements found in the sun, earth, and stars, dominated by calcium, sodium, iron, and so forth. That the strongest lines did not always come from the most abundant elements was not clarified until the 1925 work of Cecilia Payne on spectra of K giants.

Then, strangely it now seems to us, came an epoch in which on-going star formation ceased to be regarded as a part of the real universe or as a topic on which serious astronomers should work. It came in the wake of Hubble's 1929 linear relation between galaxy distances and velocities, implying a cosmic age close to 2 billion years, about the same as that of the solar system, and about the same as the time a typical star might plausibly live on "subatomic energy". And, from many things having similar measured ages to the dictum that everything

was the same age seems to have been a short step. Thus one finds many of the great and good of the astronomical community in the 1920s and 30s and even later making snide remarks about the whole subject of star formation. (Most of the references are in Trimble [19], where also is to be found a discussion of the discoveries of general interstellar extinction, H I, and all that led the community back to something like Herschel's view.) Here are some of my favorites:

"In former times, when theories of cosmic evolution were taken somewhat more seriously than they are at present … " (Baker 1932)

"A good many arguments weigh against the hypothesis of recent creation" (of S Dor, Jeans 1929)

The alternative of on-going star formation is "not very attractive", and bright stars live on some much more generous source of energy than synthesis of heavy elements. (Atkinson 1936)

"Most stars were created at the beginning of the universe some 2 or 3 thousand million years ago." (Goldberg and Aller 1943)

" … origin of stars and stellar systems at the epoch of catastrophe" (Bok 1936)

"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, Dugan and Stewart 1938 edition)

Hoyle and Lyttleton were led specifically to consider accretion on to a previously existing solar type star as being responsible for the more massive ones. This is foreshadowed in Eddington's 1926 *Internal Constitution of the Stars*. Hoyle and Lyttleton (1939 etc) [11] are still cited not because anyone still supposes that this is where OB stars come from, but because their formulae are useful for estimating accretion rates in interacting binary systems and such.

A last published quote from Jesse Greenstein in 1951 (in the Hynek astrophysics compendium) says, "Like the accretion hypothesis, the condensation hypothesis now seems suitable only for creation of a few exceptional objects rather than all the stars." And he leaves the problem unsolved.

In a preliminary effort to find out more about this epoch, I wrote to Fred Hoyle and to Lyman Spitzer, asking for their "takes" on the period. Hoyle wrote (9 Jun 1997) saying, "You ask why in 1939 we did not think of ongoing star formation. Because at the time the astronomical world did not believe in an interstellar medium containing hydrogen." And he went on to mention that, soon after, he and his colleagues had other things (like radar and proximity fuzes to worry about.

Spitzer responded (November 26, 1996), "My keen interest in the problem (star formation) dates from 1939 … supergiant stars must be younger than any likely age for the Galaxy … dusty clouds … were the stellar birthplace. Later that year, I submitted a paper to which these theoretical suggestions provided an introduction … Several astronomers objected to this introduction as too speculative, and much to my regret the paper was published without (it)." To my subsequent letter asking who the objectors had been, he responded that it was not appropriate to say because one of them was still living. I thought it likely

that he meant Greenstein, but had not attempted to verify this. At any rate, the original, "too speculative", version of Spitzer's 1939 paper now appears in his collected works (edited by him and J.P. Ostriker), though it is now too late to ask any further questions of Spitzer. A phone call to Greenstein this afternoon (13 May 2001) yielded the information that it wasn't him and he hasn't a clue who it might have been.

5   The Poster Session

More than 60 presentations, with authors ranging from Apai to Ziegler, covered almost as many topics, including some just a little outside the main thrust of the conference, occasionally indeed way out. My favorites included:

- The youngest protostar? Well possibly not, but Belloche et al. indicated that IRAM 04191 is no older than the neolithic revolution.
- The most distant dust that reaches us? Krivova suggested that certain radar-detected meteoroids represent dust from the disk of Beta Pictoris, expelled by a Jupiter-like protoplanet.
- The longest correlation length? Ibrahimov displayed light curves of FU Orionis and V1057 Cygni and showed that they share several symmetries (one only obvious when you turn the light curve upside down).
- The most deviant IMF? Stolte presented evidence that the young galactic center cluster, the Arches, is particularly rich in high mass stars (or deficient in low mass ones).

6   Current Conditions: Observational

I will take the (by no means original) point of view that the status of a field like origins of stars and planets can be described by the questions that the community is trying to answer. Many of these, as Mark McCaughrean made clear in his multi-faceted (and multi-media) introduction, have been alive and well for 20 years and more, in the case of stars. Names in parentheses are those of speakers who presented some part of an answer or a better-posed version of the question.

6.1   Planets

Some of my "questions about planets" would be regarded by others as part of the definition of a planet. The definition assumed here is things less than about 10 times the mass of Jupiter, currently in orbit around stars.

Statistical questions (Mayor, Neuhäuser, Hatzes): How many are there as a function of planet mass, composition (gas, ice, rock, metal) and orbit parameters (a, e, and i to the equator of the parent star; angles involving the node are not a property of orbits but only of our point of view)? How are the several properties of planets correlated with each other (where the solar system example

includes correlation of composition with semi-major axis)? What are the minor constituents of systems like, including comets, asteroids, and moons (Schulz)?

Host star questions: What are their masses, compositions, locations in the Milky Way, rotation periods, magnetic fields, kinematics, and ages, and how are these correlated with the sorts of planets they have? Do the averages and distributions of host properties differ from those of stars in general? It was once advertized that OBA stars still had their birth quotas of angular momentum, while most GKM stars had formed planets. I think no one would still argue this.

System questions: Are multiple planets the norm or the exception? How typical is our mix of Jovian and terrestrial? Are all planet systems coeval with their stars? (No – one pulsar has some, but it remains a unique example, despite accurate timing data for many pulsars.) Are all planet systems roughly co-planar? You don't expect them to be, said a theorist during discussion, even if they formed that way, because of subsequent interactions (Monin).

Evolution: can we see disks giving way to planets? How? And what do they look like? (Brandner, Launhardt, Looney, Wilner, Wolf).

Individual planet questions: Is chemical differentiation common among Jupiter-mass planets outside the solar system? Do other terrestrial planets have non-equilibrium atmospheres? This last is, of course, perilously close to "Is there extraterrestrial life?".

Of these, the issues of initial co-planarity and chemical differentiation are part of some experts' definitions of planets, as is the question of whether there are free-range ones in Orion or elsewhere (Barrado y Navascués, Roche).

Data reduction: How do you turn measured fluxes, line profiles (etc.) into masses, temperatures, compositions, and all the rest? This is important both for planets and for low mass stars (Baraffe, Guenther, Testi).

6.2   Stars

Statistical questions: What is the IMF? What is it correlated with (gas composition, turbulence, density, cloud mass, presence/absence of spiral arms or other observable triggers ...) and how? (Bontemps, Eislöffel, Zinnecker).

Binary questions: How many of them are there? (Bouvier, Joergens). What is the initial distribution of semi-major axis, eccentricity, and mass ratio? (Sterzik). How are these correlated among themselves and with gas composition, density, degree of ordering of magnetic fields, etc.?

Evolution questions: Which of these features can already be discerned among pre-main-sequence stars and young stellar objects, or even cloud cores, and how do statistics, binary populations, and planetary companions change as stars age? (Chini, Greene, Jayawardhana, Kaper, Monin, Nürnberger, Petr-Gotzens, Prusti).

Collective questions: How efficient is star formation in a given gas mass? Not very, noted Shu et al. [17] as the second major surprise in the decade leading up to their work. How does that efficiency depend on composition, presence or absence of triggering, and all the rest?

Cosmic questions: What was the total star formation rate as a function of time or redshift and environment? What are the entities in which most of it occurred at each epoch? How have the statistical, binary, and collective properties changed with time? Here we need to know not only about star formation per se, but about the assemblage of gas clouds dense enough to participate in the process. Blitz made it clear that we have no theoretical understanding of GMC formation, while Alves and Lombardi pointed out that observations can at least tell us what conditions the assembly process needs to aim at to make clouds like real ones.

7   Current Conditions: Theoretical

It is easy to make insulting remarks about theorists (e.g. the back end of the bipolar cow), so in the interests of fairness, I quote two theorems and their corollaries.

A. Redman's theorem: A competent theorist can explain any set of observations using any theory. Longair's corollary: In some cases he need not even be competent.

B. Redman's anti-theorem: A competent observer, given a large enough telescope, can find an object that matches any theoretical prediction and another which falsifies it. Corollary: In these days of virtual observatories, she may not even need a telescope.

7.1   Planet Questions

If we accept that planets form in circum (proto)stellar disks, then what are the dominant processes, gravitational instability, coagulation of dust to planetesimals and beyond, something else? (Wuchterl)

What patterns of post-formation dynamical evolution are possible? Which are common? Migration, it seems, including in our own system (Jewitt). Expulsion? Mutual tilting of orbits?

If free-range planets exist, were they formed that way and how, or liberated? (Reipurth)

7.2   Star Questions

What (all) supports molecular clouds against premature dynamical collapse? All perhaps of turbulence, magnetic fields, rotation, and thermal pressure?

Is contraction/collapse always, sometimes, or never triggered? By what? At what stage do dense cores form, and is the dominant process fragmentation or something else?

What makes the IMF? Turbulence is a relatively new player in this game.

What makes binaries and is responsible for the observed distribution of separations (from two stellar radii to a tenth of a parsec: at least we know what is responsible for the end points, just as we know what sets the upper and lower

limits to masses of single stars)? What determines the initial distributions of eccentricities from 0 to 1 and of mass ratios from 1 to 0, and how do they evolve? How are the initial values of these correlated?

How do clouds, on all scales from entire giant molecular assemblages down to individual Bok globules and disks, get rid of excess angular momentum and magnetic fields, and how do they know when to stop, since even very young stars do not rotate at break-up or have fields of Megagauss? Is collimated, bipolar, magnetic outflow the only game in town? What about ambipolar diffusion?

7.3   All of the Above?

Several of the items in the two previous sections and others mentioned during the workshop were phrased as A vs. B (sometimes vs. C). That is, which description of the observations is most accurate? Which physical process is responsible or dominant? Such dichotomization and the defense of an alternative by its propounders are a customary and often useful part of the scientific landscape. But I believe we have heard about several cases where the correct answer has to be "both please" (Winnie the Pooh), "all of the above" (multiple choice on the GRE), "some of them are and some of them aren't" (Franklin Pangborn), or "that depends".

Clarke tentatively put disk destruction into this class. I suspect that whether you see rotation or turbulence in molecular clouds depends a good deal on the length scale you examine and whether you reduce your data to the rest frame of a chap living inside the cloud near a dense core. Another example must surely be the IMF. Is it determined by fragmentation, mergers, when accretion stops, the efficiency of outflows? Surely they all matter.

Fritz Zwicky was a great believer in "all of the above", which he called the morphological method. He had little to say about star formation. The strongest contrast I can think of is with the work of Viktor A. Ambartsumian, who had very definite views on star formation which could not be combined with any of the ideas discussed at the meeting. He was indeed among the first to revive the idea of on-going star formation after its period of disrepute. But he decided very early (see Ambartsumian [1] for a retrospective) that stars were born by expansion out of very dense, "prestellar" matter, so that planetary nebulae were an early stage, not a late one, and expanding OB associations a natural product. The idea colored work in Armenia and, to a certain extent, throughout the USSR for decades, and without knowing about it one is puzzled to find Shklovskii in non-technical books working very hard to demonstrate that planetaries are old. Chapter 6 of Shklovskii and Sagan [16] should also be read keeping that history in mind, and it is appropriate to quote from them, "It was once believed that all the stars in the heavens were formed at about the same time, several billions of years ago. But there are now a number of lines of evidence that stars are being formed continuously, by condensation of the interstellar gas and dust."

On-going vs. one-shot star formation and expansion from pre-stellar matter vs. condensation from gas and dust are cases where "both, please" is not a

possible resolution. We think, however, we now know the right answer in both those cases!

7.4   Potentially Useful Analogies

It is much easier to tell a story after it is over than while you are in the middle of it. In this spirit, some events of early 20th century astronomy may be enlightening. The astrophysicist of 1900 would surely have said that the most important unsolved problem was the source of stellar energy. Well he would have said it in German, but that is another story. The answer, nuclear reactions, led directly to considerations of nucleosynthesis in stars, from Atkinson and Houtermans (1929) to B^2FH [6]. The nuclear reactions, their energetics and products, and how they interface with stellar structure and evolution can be regarded as largely solved problems.

In 1920, a (again perhaps the) most divisive observational issue was superficially about distance scales, but really about the existence of external galaxies. Curtis [8] said yes, and Shapley [15] said no, in a 1920 debate that is still remembered by cognoscenti. There was a definite observational answer, coming from Hubble [12] in the form of the discovery of Cepheid variables in NGC 6822 and M31. The answer then led directly to Hubble's use of Cepheids as an extragalactic distance indicator, calibrating all the rest, and so to the discovery of his linear distance-velocity relation, that is, the expansion of the universe. Notice that the customary inclusion of a K term in studies of stellar motions means that, if the galaxy had been expanding, it would have been discovered in time to save Einstein from his cosmological constant! Oh, Hubble's answer was "yes", in case you weren't sure.

It seems plausible that there is a corresponding key question for planet and star formation whose answer will lead us in some new, interesting direction. Unfortunately, I do not know what it is (or I would be working on it instead of writing this manuscript). But it is probably not, "Are there other solar systems?" or "How do stars form?".

8   The View from 2021

"I find it difficult to make predictions, especially about the future", so supposedly said Enrico Fermi. In fact, only the future is safe. Predictions about the past, variously called hindcasts, retroposals, and postdictions (the words came from Thomas Gold, an expert), are all too easily falsified.

Easiest to summarize is the computational front, where astronomers will continue to benefit from smaller hardware and bigger software driven primarily by commercial and military goals (though occasionally our algorithms, especially from radio interferometry, are borrowed back the other way into telecommunications, medicine, and so forth). This means, of course, more and faster, including many more ways to link, ask, and communicate. Perhaps there will be ubiquitous PixelPrintPads (pronounced "3-Ps") against which you touch your thumb

to log into anywhere, from anywhere, and can carry on by talking, scribbling, or even feeding in punch cards. Reading of brain waves is probably a couple more revolutions ahead.

It is still, however, a good bet that you and your students will continue to strain computing capacity with magnetic fields, gas flows, dust, multi-wavelength, and dynamic range problems. Twenty years is, after all, only about 13 Moore's law doubling times, or 10^4, and the star formation code that now stretches from a 50 $M_\odot$ cloud down to 1 AU details won't yet quite reach, simultaneously, a 10^5 $M_\odot$ cloud and planet sizes. Cleverness will, therefore, still be required and rewarded. And, of course, you will still spend more time debugging then running, because you are a scientist, not an accountant.

On the observational front, someone will probably be writing just about the last proposal for observing time on NGST (McCaughrean), if it is launched a little late (what else is there) into a 10–15 year mission. SOFIA should still be available, but SIRTF (Meyer) long gone, unless it has not yet been launched. Other things relevant to origins of planets and stars that you might reasonably find in space are SIM and GAIA (American and European astrometric missions, somewhat oriented to planet searches); conceivably TPF (Terrestrial Planet Finder, the next generation) funded if not flown; Astro-F (which, like all Japanese missions will get a proper name only when it is safely in orbit to do its infrared photometry and spectroscopy); and HerschelFIRST (André). SAFIR, the Single Aperture Far InfraRed Observatory, or whatever it morphs into, might be under construction. Potentially available to look for X-rays as a signature of young stars (Montmerle), and surely renamed, will be Spectrum-X, MAXIM, Const-X, and perhaps others.

On the ground, there will have been some sacrifices, probably the Lick 120″ and plausibly some of the four-meter class instruments at less than ideal sites, perhaps the Greenbank Telescope (and, of course, NSF will still be looking for a partner to operate the 12-meter Kitt Peak millimeter telescope). But the total area of glass covered by aluminum or silver or – surely this is an area of materials science ready for exploitation with a substance that reflects all wavelengths penetrating to ground and not in need of frequent recoating? – will still be larger, and larger per astronomer, than at any time in the past. Already with us, or nearly so, and good for at least 20 more years are the VLT and its outriggers (Lagage, Malbet, Montmerle, Moorwood, Richichi, Siebenmorgen), Kecks and their sideKecks (Boden), Gemini, Subaru, the Large Binocular Telescope, GranTeCan, Magellan, HET, SALT, and the MMT (for single-mirror telescope).

At longer wavelengths, the VLA should long since have become the Expanded VLA (followed presumably by the Grossly Expanded VLA), and ALMA (Henning, van Dishoeck) be a major, multi-user facility, well-tuned to many aspects of star formation near and far. If the Square Kilometer Array does not yet have a filler aperture adding up to a square kilometer, it has the advantage of being accomplishable in steps (unlike most space missions).

What about the next step from 10-meter apertures up to 30, 50, or 100? A Giant Segmented Mirror Telescope is the second highest priority in the US

National Research Council decade report for the 2000's. Competing concepts in the US include MAXAT (Gemini consortium), ELT (Univ. of Texas), and CELT (Univ. of California and Caltech). There is the Swedish XLT and the largest of all, OWL (Gilmozzi). Will they all happen? Not by 2021, and, because the main showstopper for both CELT and OWL (Gilmozzi; G.A. Chanan, private communication) is money, perhaps not at all except as the "AOL-OWL", that is, with major private sponsorship. I believe the two syllables of that compound are pronounced identically, but I am not sure.

All of these, from 6-meters on up, are being planned with some form of active and adaptive optics (which is like setting your windmill to blow the breeze back the way it came, before it gets to you, but which also enormously improves sensitivity by reducing noise within an image as well as angular resolution) and some 4-meters are being retrofitted. Many also form or will form parts of interferometers, whose analysis is like being led blindfolded to the top of a mountain and then being required to describe not only the entire mountain but its Fourier Transform. This is even more true in optical than in radio astronomy (the description came from the later Peter Scheuer), because there are not, so far (and may never be) optical superheterodynes that let you keep dividing your supply of photons in half and still have the same number you started with, limiting optical interferometers to a dozen or so baselines. NGST is, in contrast, like throwing a butterfly up at 20 or 30 g and then expecting it to spread its wings and fly away.

Theory is even harder to project into the future, because the retooling time (and also the half life) can be very short. It seems unlikely that anyone will be doubting that the dominant physics in star and planet formation is gravity and electromagnetism, but this does not seem particularly helpful.

It may be a bit more helpful to remember that there have been in the past several cycles, in which angular momentum (à la Laplace) gave way to gravitational processes (à la Chamberlin/Moulton and Jeans), which gave way to dynamically important magnetism (à la Mestel and Woltjer and Parker), and back to angular momentum again (Larson, Shu) and gravitational instability (Boss, for planets) and magnetic fields (Mouschovias) again, and so forth, as the bandwagon process of the year. This will surely happen again.

At each return, naturally, the considerations become more sophisticated (or anyhow harder for a novice to understand). Thus last year a couple of my Maryland colleagues published a paper tabulating no fewer than 11 kinds of instabilities in accretion disks, most of which have not yet been applied to the problems of planet and binary formation. The process of the year for star formation in 2001 is probably turbulence, and it could come and go several times, though I would bet about 50-50 that the process of the year in 2021 will be one not currently in the inventory, because it hasn't yet been thought of at all or hasn't been recognized as relevant.

Another theoretical territory comprises chemical bonds, stickiness, fluffiness, and other "materials" properties of matter (Bockelée-Morvan, Dominik, Wright)

that must surely be important in planet formation, and here is one jurisdiction in which Space Station experiments may actually be a useful guide to theory.

Finally, any science progresses by a continuous cycling around data, the models they inspire, the predictions of those models, collection of confirming or conflicting data, improvement of the models, deducing the consequences of the improved models, and so forth. You can enter the cycle at any point (notice that numerical simulations often belong to the "deducing the consequences of ideas" stage). Some phases of the scientific process have the stark elegance of chess. Others – and studies of star and planet formation appear to be at this stage – are more like mud wrestling.

9   Changing the Course of History

Just as only the future is safe when making predictions, so in the case of influencing the course of history, only the future is possible. What can we do (beyond the obvious of producing the best science possible with the resources already in hand) to ensure the future health of work on origins of stars and planets, astronomy as a whole, or indeed scientific research in general? Diamond [9] has synthesized an enormous sweep of human history and prehistory, identifying a handful of factors that have led to widely different rates of economic progress in the range of human tribes, clans, nations, and empires on the several inhabited continents. Four of these are directly applicable to science, to wit (1) surplus food or productivity, (2) population size, (3) ease of diffusion of ideas and artifacts, and (4) concentration of power.

Surplus productivity may seem too obvious even to mention. Societies of hunter-gatherers build very few large telescopes; we and our graduate students are fed on grain and grapes that we do not grow; and the number of members of the International Astronomical Union in a country is a fairly good proxy for gross national product. If you take the number of IAU members in a country and divide first by total population and then by per capita annual income in dollars, you get an index on which India, Japan, and Portugal rank nearly equal at 1.3–1.5 $\times 10^{-10}$ and the two Chinas are nearly even at $8 \pm 0.4 \times 10^{-11}$. Indeed 44 of the 62 adhering countries fall between 0.8 and 8 $\times 10^{-10}$ on this scale. The outliers on the high side are former portions of the Soviet Union and its Eastern European neighbors (as are many of the "high normals") who got into the habit of living scientifically above their incomes during the cold war (plus Vatican City at R = 2.9×10^{-8}!).

The outliers on the low side are (a) large and still climbing out of poverty, with science not yet very high on the list of priorities (like Malaysia and Indonesia) and (b) not necessarily poor so recently, but with little recent tradition of scientific research of any kind (Algeria, Iran). And the 120 or so UN member nations that do not adhere to the IAU are, on average, both considerably poorer and considerably smaller than the ones that do. A better normalization might include both per capita income above some threshold and population over some threshold (the quota for Vanuatu and St. Kitts and Nevis is rather less that one astronomer

per country, and considerably less than one per island), which leads us toward the second item.

I am not quite sure what you are supposed to do with these numbers except to recognize that if you vote simultaneously for slower total economic growth (for whatever reason, however meritorious), lower taxes, and more scientific research, you are asking your fellow country persons to do the improbable, if not the impossible.

Second, population size, is already a more complex variable and leads to an "action item". If all human beings fall in a single gaussian curve of ingenuity (which may well be true) and there were no benefits or penalties from interactions, then the creation of new scientific ideas should be linearly proportional to numbers working in a field, independent of the size of the interacting community. This is the equivalent of star formation rate being linearly proportional to gas density and is almost certainly wrong for human creativity. Mozart and van Rijn (though not Vermeer) flourished among unusually large numbers of other not-quite-so-good composers and painters.

What is the real relationship? (a) A threshold, above which productivity scales as some power of population size? This is the philosophy behind the concept of "critical mass" that university departments are forever trying to achieve for their new research initiatives. It seems to mean either "two more than we have now" or "two more than you have in your group". (b) Next simplest is a power law, $P(n) \propto n^x$ and x larger than unity, like a Schmidt law for star formation. (c) Or there could be several thresholds for different levels of innovation, each followed by some proportionality. These are the optimistic ideas. Alternatively, however, the truth may be more like (d) a broken power law, with the benefits of interaction decreasing as the population becomes too large for each member to know all the others. Or (e) indeed at some point perhaps there is a saturation, where, like rats in a too-crowded cage, we spend most of our time biting each other (i.e. writing and reviewing proposals) and very little on eating and reproducing, let alone thinking how to improve the cage, or, like the servants in a very large household, we spend almost all our time looking after each other (writing letters of recommendation, sitting on committees) and almost none serving the mistress, astronomy.

Many American astronomers feel that we are already perilously close to condition (e) and so, I suspect, do other national and transnational communities. We all know colleagues who have had 30 or 60 or even 100 PhD students in their careers (60+ for Dennis Sciama; 100 for the mathematical physicist Elliott Montroll, who also had 10 biological children) and it may be time for our admiration and envy of such colleagues to be tinged with caution.

The numbers in Table 1 are provided for you to try testing your own feelings or hypotheses, if you feel so inclined. The total American astronomical community (based on AAS membership with minor corrections) is about three times the number shown as IAU members. This may not be typical. The smallest populations represented at the workshop seem to have been Hungary, Ireland, Portugal, and Uzbekistan.

Table 1. Astronomy as a measure of GNP

Country	Members	R	Country	Members	R
Algeria	3	2.5×10^{-11}	Japan	448	1.4×10^{-10}
Argentina	90	2.5×10^{-10}	Korea RP	51	8.1×10^{-11}
Armenia	31	3.3×10^{-9}	Latvia	8	7.9×10^{-10}
Australia	191	4.8×10^{-10}	Lithuania	12	7.9×10^{-10}
*Austria	31	1.8×10^{-10}	Macedonia	0	—
**Belgium	88	3.7×10^{-10}	Malaysia	7	3.0×10^{-11}
Bolivia	0	—	Mexico	83	1.0×10^{-10}
Brazil	109	1.0×10^{-10}	**Netherlands	167	8.7×10^{-10}
Bulgaria	50	1.5×10^{-9}	New Zealand	26	4.1×10^{-10}
Canada	199	3.0×10^{-10}	Norway	22	1.8×10^{-10}
Chile	46	2.2×10^{-10}	Peru	1	8.9×10^{-12}
China, PR	368	8.6×10^{-11}	*Poland	117	4.2×10^{-10}
China, ROC	23	7.4×10^{-11}	**Portugal	17	1.3×10^{-10}
Croatia	13	6.2×10^{-10}	Rumania	37	3.1×10^{-10}
*Czech Rep.	71	6.4×10^{-10}	Russia	344	5.0×10^{-10}
**Denmark	52	4.2×10^{-10}	Saudi Arabia	11	5.3×10^{-11}
Egypt	39	1.3×10^{-10}	*Slovak Rep.	27	5.8×10^{-10}
*Estonia	22	2.4×10^{-9}	South Africa	46	1.7×10^{-10}
*Finland	37	3.6×10^{-10}	*Spain	204	3.6×10^{-10}
**France	609	4.6×10^{-10}	**Sweden	95	5.4×10^{-10}
Georgia	19	2.4×10^{-10}	**Switzerland	70	4.0×10^{-10}
**Germany	488	2.9×10^{-10}	Tajikistan	8	6.9×10^{-9}
*Greece	89	6.4×10^{-10}	Turkey	53	1.4×10^{-10}
*Hungary	41	5.4×10^{-10}	UK	535	4.3×10^{-10}
Iceland	4	7.8×10^{-10}	Ukraine	119	9.5×10^{-10}
India	227	1.4×10^{-10}	Uruguay	6	2.0×10^{-10}
Indonesia	13	1.3×10^{-11}	USA	2235	2.7×10^{-10}
Iran	15	3.9×10^{-11}	Uzbekistan	8	1.3×10^{-10}
Ireland	33	4.9×10^{-10}	Vatican City	5	2.9×10^{-8}
Israel	45	9.6×10^{-10}	Venezuela	11	5.8×10^{-11}
**Italy	409	3.3×10^{-10}			

R = Number of members/(population $\times$ per capita income)

Membership data from IAU records just prior to 2000 General Assembly

Population and economic data from The New York Times Almanac (1999)

* Participants in A&A	** Also ESO members
2617 IAU members	1998 IAU members

Third of the conditions considered by Diamond is the ease of diffusion of various things. His examples include (a) cultivars, easy in longitude, much harder in latitude because the climate changes, and also hard across barriers of deserts, mountain ranges, and dense forests, hence some of the advantages Eurasia has had over the pre-Columbian Americas and pre-Vascan Africa, and (b) technologies, easier across wide continents than isthmuses, still harder across oceans, and often requiring a minimum population for maintenance. You will immediately think of some scientific/astronomical analogies, like the sudden implementation of adaptive optics when the barriers of military classification fell (even though the idea came originally out of the astronomical community and wheels were slowly being reinvented within it). Another case is the largely independent development of ideas about both high energy radiation mechanisms and cosmology (dark matter, origin of large scale structure, limits on neutrino masses) east and west of the Oder-Neisse line.

The action item under this third heading is, clearly, to tell everybody anything that you think might be of use to them. This unfortunately contradicts the action item from heading two, which is not to drown your colleagues in your preprints, emails, and all the rest.

Fourth and last is the issue of concentration of power. Even more than items 2 and 3 this is a two-edged sword. Only a community with a great deal of centralized decision making can declare "we will build and deliver an atomic weapon" or "we will put a man on the moon in this decade" and make it happen. Although the primary drivers were quite different, science including astronomy has benefitted at some level from these, even study of formation of stars and planets, since at least we now know that the oldest moon rocks are a bit younger than the meteorites but a bit older than the oldest earth rocks, and quite a lot about some kinds of shock propagation. But converse examples are also easy to accumulate. Ming China began a series of open ocean voyages in the early 15th century, to which a new government faction brought a screeching halt in 1433. Contract the case of Columbus who (in case you had forgotten) had already been turned down by three funding agencies in three different statelets before he approached Ferdinand and Isabella. The Japanese decision to stop gun production in the late 1600s was also made to stick because of a centralized government isolated from competition.

It is not in any sense that the West has been immune from such errors. The words Luddite and saboteur are our very own. Other examples [10] are edicts forbidding the use of Arabic numerals by Florentine bankers (1299), the use of machines by tailors in Cologne (1397) and ribbon looms in Danzig (1579). But, because there were and are many European states, treading many different paths, competition soon forced adoption of the more efficient technology. Just knowing that there is another, better way of doing things can be a powerful incentive to change. Contrast the delayed (but not much delayed) electrification of London street lights around 1888 (the Gas company objected) with the persistence of the QWERTY keyboard on which we write. Yes, there are almost certainly better

arrangements, but have you ever seen one? Whole books have been written about the subject.

What astronomical lessons can we learn here? Chances are, you will agree with some but not all of the following:

1. Britain perhaps performed a public service by staying out of ESO as long as possible and carrying on with independent development of 4-meter class telescopes and focal plane instrumentation, though not necessarily a service to her own astronomical community.

2. Very few things of the CELT/ELT/OWL classes will be built, so they will probably not be the best possible of their type. Electric cars would work a lot better if there had been as many manufacturers and generations of them as of internal combustion vehicles.

3. An historic strength of the US scientific community has been the multiplicity of funding sources and decision processes, including several large states (especially California and Texas) and several government agencies (NASA, NSF, DOE, and DOD and its ancestors the Army, Navy, and Air Force Offices of Scientific Research). Thus the possibility of restricting large projects to NASA should be Viewed With Alarm.

4. The number of different journals in which a fairly respectable astronomer can publish has fallen enormously over the years, with the loss of something like 100 independent observatory publications that existed near the turn of the previous century, the decline of national academy publications (at least for astronomy; remember Hubble's law first appeared in the Proceedings of the US National Academy of Sciences), and, more recently, the mergers that led to A&A, and the demise of QJRAS, Irish Astronomical Journal, Comments on Astrophysics, and others. While this has presumably yielded economies of scale, it has also concentrated the power of saying yeah or nay into fewer editorial hands and encouraged the remaining ones to adopt more nearly uniform standards.

The action item that comes out of this fourth, concentration of power, consideration could be phrased as "celebrate diversity" (a singularly politically correct sentiment) or "don't always vote for mergers".

Acknowledgements

y Navascués, Rosina Iping, and Dick Ballard. Conference participants included native speakers of more than 20 languages (not to mention broken English, the international language of modern science, according to Zeldovich) who kindly provided for my final overhead suitable expressions of farewell and hopes for speedy next meetings in them all. Unfortunately not even Springer has that complete a set of type fonts, and so only "Auf Wiedersehen, até à próxima, and Ankh Wdjah Seneb, the Ancient Egyptian, May you Live, Prosper, and be Healthy".

References

1. Ambartsumian, V.A. 1960. QJRAS 1, 152
2. Badawy, A.M. 1954. History of Egyptian Architecture (Gizah), p. 153
3. Bahcall, J.N. & Ostriker, J.P. (eds.) 1997. Unsolved Problems in Astrophysics, Princeton Univ. press
4. Brilliant, A. 1985. Pot Shots No. 3253, Brilliant Enterprises, Santa Barbara
5. Brush, S.G. 1996. A History of Modern Planetary Physics, Cambridge Univ. Press
6. Burbidge, E.M., Burbidge, G.R., Fowler, W.A. & Hoyle, F. 1957. RMP 29, 547
7. Cox, A.N. (ed.) 2000. Allen's Astrophysical Quantities, Fourth Edition, Springer
8. Curtis, H.D. 1921. Bull. NRC 2, 194
9. Diamond, J. 1998. Guns, Germs, and Steel, W.W. Norton
10. Economist, The 1999. Issue of 31 December
11. Hoyle, F. & Lyttleton, R.A. 1939. Proc. Cam. Phil. Soc. 35, 405, 595, 608
12. Hubble, E.P. 1924. NY Times, 23 November, p. 6
13. Lang, K.R. 1980. Astrophysical Formulae, 2nd Ed., Springer
14. Russell, H.N., Dugan, R.S., & Stewart J.Q. 1926. Astronomy (Boston: Ginn & Co.)
15. Shapley, H. 1921. Bull. NRC 2, 171
16. Shklovskii, I.S., Sagan, C. 1966. Intelligent Life in the Universe, Holden-Day
17. Shu, F.H., Adams, F.C., & Lizano, S. 1987. ARA&A 25, 23
18. Trimble, V. 1975. RMP 47, 877
19. Trimble, V. 1997. In S.S. Holt & L.G. Mundy (eds.) Star Formation Near and Far, AIP Conf. Ser. 393, p. 15
20. Trimble, V. & Aschwanden, M. 2001. PASP, 113 (Sept. issue)