Summary: Star Formation Near and Far

Richard B. Larson

Yale Astronomy Department, Box 208101, New Haven, CT 06520-8101

In her historical introduction to this conference, Virginia Trimble reminded us that the study of star formation is a subject with recent origins, and that just fifty years ago it was still considered very speculative to suggest that stars are presently forming from interstellar matter in the Milky Way. In my concluding remarks, I shall try to summarize with a few brief and personal comments where I think we have come in the past fifty years in understanding star formation, as illustrated by the many interesting contributions that we have heard at this meeting. The meeting included a remarkable range of topics and scales, from planets to the universe at large redshifts, and it will not be possible for me to summarize everything of importance that was covered; it will be a challenge even to mention all of the main topics that were discussed.

Let me begin by updating, from a personal perspective, Virginia's historical introduction. The way was prepared for modern studies of star formation by the advances in our understanding of stellar evolution that took place in the 1950s, which made it clear that the Milky Way contains stars and star clusters of all ages and that one cannot avoid the conclusion that stars are presently forming. Dark clouds, now known to be molecular, were discussed during this period as possible stellar birth sites, and the T Tauri stars were identified with some confidence as newly formed pre-main-sequence stars. By the 1960s, when I came on the scene as a graduate student at Caltech, the time was ripe for an attack on the problem of protostellar collapse using the techniques of numerical hydrodynamics that had been developed earlier for nuclear weapons work. I was guided in this direction by my thesis advisor Guido Münch (who, coincidentally, was at the same time also Virginia Trimble's thesis advisor), and I calculated the collapse of a spherical protostar, that being what was then feasible. I initially regarded this calculation as an academic exercise, since theoretical discussions of star formation had already strongly emphasized the importance of magnetic fields and rotation, as well as the possibility of continuing fragmentation. At about the time when I was completing this work, however, another Caltech student, Eric Becklin, discovered an infrared source in Orion which he suggested might be a protostar, and since his observation agreed qualitatively with what I had predicted, we were both very encouraged to believe that protostars like those calculated might actually exist. Theory and observation subsequently made great strides, and the study of star formation was soon no longer purely a matter for speculation.

Almost thirty years have passed since then, and we now possess a great wealth of information on star formation, as we have seen at this conference. The development that has impressed me most is the recent avalanche of high-quality data produced by many modern instruments, among which the Hubble Space Telescope has been especially notable. Many speakers showed spectacular and beautiful photographs of regions of star formation that are unprecedented in the amount of detail they reveal, and I believe that it is no exaggeration to say that observations like these open up a completely new window on the universe. Thanks to them, the science of astronomy, which previously had to be content with fuzzy images and data of generally poorer quality than those of other sciences, can now boast of data whose quality and detail begin to rival those of such Earth-bound sciences as geology, meteorology, and perhaps even biology. As a theorist, I am humbled by the realization that the observations have now far surpassed the limited ability of theory to explain what is being observed, and I believe that this forces a change in the way in which we try to approach an understanding of things. The hope of predicting everything from some kind of "fundamental theory" seems to me a mirage, and I think that instead we now have to focus on the phenomenology and work toward trying to understand it bit by bit. One wouldn't think of trying to predict all of the complex phenomena of geology, for example, from any fundamental theory, and I am sure that star formation, in all its aspects, is no less complicated a subject than the formation of geological landforms.

I have tried to organize my remarks around some of the main questions that one might ask about star formation. Some of these questions were introduced at the opening of the conference by Silk. The first obvious question is "Where do stars form?", and the answer that has been known for many years is that stars form in molecular clouds, or more specifically, in the dense gravitationally bound cloud cores that were discussed here by Myers. Immediately, however, we are confronted with the fact that these clouds are extremely complex in their structure and dynamics, so that we cannot yet provide a good answer to the next obvious question, "How do star-forming clouds evolve?". Molecular clouds appear to be dominated kinematically by a chaotic state of motion often called "turbulence", even though magnetic fields are believed to play an important role and the observed motions may be partly wavelike. The observed turbulent motions contribute importantly to supporting clouds against gravity, at least on the larger scales, and the talk by Ostriker and the posters by Vázquez-Semadeni and collaborators showed that important progress is being made in understanding, with the help of numerical simulations, how MHD turbulence can help to support molecular clouds and account for some of their properties. I think, however, that it has been a myth that molecular clouds are long-lived structures that are maintained close to equilibrium for many dynamical times; such long lifetimes are not supported by the available direct evidence, based on the ages of associated young stars and clusters, which suggests cloud lifetimes that are not much longer than the dynamical time. Most molecular clouds and condensations may therefore be fairly transient structures, like the structures obtained in the numerical simulations which often are more like transient fluctuations than coherent, long-lived “objects”.

How is the collapse of star-forming clumps initiated? A widely accepted view has been that self-gravitating clumps gradually separate from a magnetic field by slow ambipolar diffusion until a configuration approximating a singular isothermal sphere is formed, which then begins to collapse from the inside out. However, given the highly turbulent state and the apparently short lifetimes of molecular clouds, it is not clear that there is sufficient time for such a slow ambipolar diffusion process to operate, and it seems likely to me that most star-forming clumps are produced by some more dynamical mechanism like turbulent compression; self-gravitating clumps are indeed found to be produced in this way in the numerical simulations mentioned above. Local dissipation of the turbulent motions that provide part of the support against gravity may then play an important role in initiating the collapse of these clumps, as suggested in the poster by Goodman et al.

How does the collapse subsequently proceed? New results presented at this meeting suggest that inside-out collapse may not occur, even in the standard ambipolar diffusion picture, since the posters by Basu and by Safier et al. both show that in even this case a singular isothermal sphere is never closely approached; instead, magnetic support becomes unimportant at a relatively early stage and a rapid collapse begins from the outside in. Myers presented evidence that in at least one well-studied cloud core, collapse has already begun on a large scale before any central object has formed, contrary to the prediction of the inside-out collapse model. Thus, it appears that revision may be needed to the popular but idealized model in which a self-similar inside-out collapse results in a constant rate of accretion onto a central object; more realistic calculations now indicate that the accretion rate must start out high and then decrease with time. An important implication of this result is that the collapse process cannot really be scale-free, as is often assumed, but instead it must necessarily depend on initial and boundary conditions.

How does rotation affect the collapse? In the simplest case of uniform collapse, rotation should lead to the formation of a disk, whereas in the more realistic case of a centrally condensed collapsing clump, one would expect the formation of a central object surrounded by a disk. Such protostellar disks are indeed now known to be very common in regions of star formation, and it is becoming possible to study their properties systematically, as we heard from Mundy and Sargent. Both inflow and rotation are often observed, and there is evidence for evolution from an early infall-dominated stage to a later rotation-dominated stage. The masses of most of the observed disks are however estimated to be rather small, less than 0.01 solar masses, and it is not yet clear how many of them might form planetary systems like our own. As was reviewed by Hartigan, there is much evidence that material continues to be accreted from these protostellar disks onto their central stars. Theoretical efforts to understand the evolution of such disks have been dominated by the concept of an “accretion disk” and have sought to find transport mechanisms capable of driving an accretion flow. The physics of the various possible transport processes remains poorly understood, however, and we heard from Stone that one of the original ideas, namely that fluid turbulence might drive an accretion flow, does not work because such turbulence tends to transport angular momentum inward rather than outward. The possibility remains that MHD turbulence may serve better to drive an accretion flow, but more work is needed to demonstrate this.

What about the possibility of continuing fragmentation? In principle, one could imagine that a collapsing clump might fragment hierarchically into smaller and smaller objects until fragmentation is stopped by the increase of opacity at a mass less than 0.01 solar masses. However, this clearly is not what happens in most cases, and theory and observation now seem to be converging on an understanding that fragmentation is strongly limited in centrally condensed collapsing clumps and typically produces only two objects, i.e. a binary system. Two stars is the minimum number that could take up any significant fraction of the initial angular momentum, so it is perhaps not surprising that, as Mathieu reminded us, the basic unit of star formation is typically a binary system. The presence of a binary companion clearly prevents the existence of a disk with a size comparable to the binary separation, but it apparently does not prevent the existence of disks with much smaller or much larger sizes. Multiple systems are also commonly observed, especially in the densest star-forming environments, and Whitworth presented numerical simulations illustrating how multiple fragmentation might occur in a realistically complex situation involving protostellar interactions in a star-forming cloud.

Most stars also form in larger groupings or clusters of some kind, so it is important also to understand how star formation might occur in clusters. We heard a whirlwind of interesting theoretical ideas on this subject from Lin and Clarke, but it would be an impossible task to try to summarize them all here. Clearly, much further work and testing will be needed to establish which of these ideas are most useful in helping us to understand the observations. There is beginning to be a rich phenomenology on the subject of cluster formation, and since the processes involved are obviously very complex, being closely related to the evolution of massive molecular cloud cores, I think that theoretical work will have to be guided increasingly by the phenomenology, focusing again on trying to build up an understanding of it bit by bit.

What are the predicted properties of newly formed stars? The prediction that protostars should first be observable as infrared sources, and the good agreement that has been achieved between predicted and observed infrared spectra, have been regarded as triumphs of the theory, although it is not yet clear whether the infrared spectra can be used to diagnose the detailed structure of protostellar envelopes and disks. Theoretical models also predict continuing infall of residual gas onto the central star or disk, but at this meeting we heard more about another phenomenon that was not predicted at all theoretically, namely the jets that all newly formed stars seem to produce at an early stage of their evolution; Mundy and Hester showed HST pictures of the detailed structure of some of these jets. Echoing the words of a famous physicist when confronted with a new particle that did not fit the then-established scheme of things, we might ask "Who ordered jets?" Jets and other outflows are in fact much more conspicuous properties of newly formed stars than is any evidence for continuing gas infall, and for some years they seemed to contradict the theoretical models in this respect. Now it is clear that infall, outflow, and rotation are all present simultaneously in many cases, and that continuing infall effects can indeed be seen when one looks hard enough. Even though we still do not understand in any detail the origin of the jets, we can no longer doubt, after seeing the striking HST picture of the HH30 jet emerging from the center of a protostellar disk, that jets originate from the innermost parts of circumstellar disks close to the central star, and that they are collimated along the rotation axis from the outset. It seems almost certain that the origin and collimation of these jets somehow involve magnetic fields twisted by the rotation of the disk, and that the basic energy source is provided by accretion from the disk onto the central star. The fact that the observed jets are not uniform but contain strings of bright knots, suggesting that they have been emitted in spurts, suggests that the accretion process may itself be sporadic, perhaps related to the FU Orionis flareups that are apparently experienced recurrently by most young stars.

How do newly formed stars influence their birth clouds? Much attention has been devoted to this subject, and some of the effects of star formation on molecular clouds are indeed dramatic. Jets and the associated molecular outflows can blow sizable holes in star-forming clouds, and they may provide a significant energy source for the turbulence in these clouds, although it is not yet clear how much coupling there is between outflows and overall cloud dynamics. Another effect that is clearly of major importance is that the ionizing radiation from newly formed hot stars can rapidly evaporate away the remaining gas in a star-forming cloud, etching away in the process the less dense parts of the cloud and revealing the previously obscured denser parts. The already widely publicized HST photographs of M16 (the "Eagle Nebula") presented here by Hester clearly show both the expected ionized outflows and the intricate structure of the remaining dense parts of the cloud, including a number of small clumps apparently caught in the act of forming stars. Clearly some of the processes of stellar birth are being revealed here, but much remains to be understood about the complex structures that are seen.

My next question is one of broad importance, and it was addressed by several speakers: "What is the stellar Initial Mass Function?" I believe that we have learned two basic facts about the IMF that we can now state with some confidence. First, as we heard in the talks by Kulkarni and Basri, brown dwarfs are clearly not very common objects, even though they can be detected and identified readily with present techniques. Even if the number of stellar objects per unit mass interval remains approximately constant down into the brown dwarf regime, as suggested by Basri, the number per unit logarithmic mass interval then drops strongly below 0.1 solar masses, and the IMF defined in this way peaks at around a few tenths of a solar mass. The total amount of mass per unit logarithmic mass interval peaks at a larger mass around one solar mass, and we can now say definitely that not only is there relatively little mass at the top end of the IMF, there is even less at the bottom end. This implies what seems to me the most basic fact about star formation, namely that there is a characteristic stellar mass of the order of one solar mass, and that most of the mass that forms stars goes into stars with masses of this order. The second basic fact about the IMF, which has long been known, is that at masses above the mass where it peaks it has an approximately power-law tail toward higher masses. As we heard from Heap and Richer, the slope of this power law seems to be very similar in different locations, at least wherever good direct star-count data exist; in fact, most recent studies have yielded results consistent with the original Salpeter law. A possible exception pointed out by Heap is that the IMF appears to be flatter at the center of the 30 Doradus cluster, suggesting that the formation of massive stars has been favored there. Also, as noted by Richer, there is evidence that the IMF for Galactic halo stars may be flatter than the Salpeter law.

What determines the IMF? Those who regularly attend meetings on star formation know that for some years there have been two competing views on the question of what determines the characteristic stellar mass. One view, based on the scale-free inside-out collapse model, is that a forming star continues to grow in mass by accretion at a constant rate until the accretion process is terminated by a wind whose onset is controlled by internal stellar physics; the other, which I have favored, is that stellar masses depend instead on cloud properties, and that the characteristic mass is closely related to the Jeans mass calculated from the typical temperature and total pressure in molecular clouds. I remark here that if the collapse process is not scale-free but depends on initial and boundary conditions, as discussed above, then stellar masses must necessarily depend at least in part on cloud properties. In this case, it is possible that the mass at which the IMF peaks might be different in different circumstances.

On scales larger than those of individual clusters or molecular cloud cores, much additional interesting structure is seen in regions of star formation, as was reviewed by Humphreys. Often several young clusters are seen in the same star-forming region, and a more dispersed population of young stars is also present. Sometimes there is evidence for multiple episodes of star formation, and it is possible that some of the later episodes are triggered by the effects of earlier ones in sweeping up and compressing the remaining gas into expanding shells, as discussed by Surdin. I was particularly impressed again with some of the new data that we saw, especially the spectacular pictures of the giant H II regions 30 Doradus and NGC 604 and their associated star-forming complexes. Here, especially in 30 Doradus, an enormous range of phenomena is observed in the same region of space – massive molecular clouds, infrared protostars, luminous young clusters, ionized flows, windblown bubbles, supernova remnants, hot X-ray emitting gas, and multiple large shells, sometimes interacting with each other – making this region a real happy hunting ground for students of star formation and interstellar matter, as well as a good object lesson in the true complexity of the real world.

Concerning the global history of star formation in galaxies, a question of long standing has been whether star formation rates vary in a smooth and regular way with time, or whether their variation is more discontinuous and characterized by discrete episodes of major star forming activity. The star formation histories of spiral galaxies were reviewed by O'Connell, who noted that the photometric properties of both the disks and the bulges of these systems show well-defined trends suggestive of a smooth evolution regulated by overall galactic properties such as Hubble type. Whatever the detailed mechanisms involved, it would not be surprising if galactic star formation rates were to depend on dynamical timescales and hence on Hubble types in the sense that the galaxies of later Hubble type, which tend to have longer dynamical timescales, turn their gas into stars more slowly, as the observations indicate. The possible controlling factors for star formation were discussed by Skillman, who noted that the occurrence of star formation is always correlated with the presence of a high surface density of gas in galaxies, and that often but not always, significant star formation is observed only where the gas surface density exceeds the critical value required for large-scale gravitational instability to occur in the gas layer. Possible dynamical influences were reviewed by Kenney, who noted that strong spiral waves or bars can play an important role in redistributing the gas in galaxies and causing it to pile up in preferred locations such as nuclear rings. He also noted that since the critical surface density required for gravitational instability is high in the nuclear regions of galaxies, one might expect that if this critical surface density were exceeded as a result of gas inflows, a very high star formation rate, i.e. a starburst, would result.

What causes starbursts in galaxies? As we heard from Hibbard, the most extreme starbursts are found in the central regions of the ultraluminous infrared galaxies, and these galaxies are invariably very peculiar morphologically and are almost certainly strongly interacting or merging systems. As was noted above, the observations suggest that the basic requirement for an extremely high star formation rate is a very high gas surface density. Merger simulations show that sufficiently extreme central concentrations of gas are indeed produced when galaxies merge. High gas surface densities and high star formation rates are sometimes also seen in the regions of overlap between interacting galaxies, and luminous young star clusters seem to be produced in abundance in starburst regions. Heckman discussed the cosmogonical importance of starbursts, and pointed out that they account for a significant fraction of the formation of massive stars in the universe and hence for a significant fraction of the heavy-element production. If nuclear starbursts also produce low-mass stars in the numbers expected for a normal IMF, they must play an important role in the formation of galactic bulges. Outflows from starburst regions may also be important in enriching the surrounding intergalactic medium in heavy elements.

Star formation rates in high-redshift spiral galaxies were reviewed by Madau, who also discussed the overall cosmic history of star formation. The inferred cosmic star formation rate increases strongly with increasing redshift up to at least $z \sim 1$, but this increase does not appear to continue to much higher redshifts, and the evidence for galaxies with $z > 2.5$ suggests lower total star formation rates at these early times (although these early values are strictly lower limits.) The overall cosmic SFR may therefore start out low at very early times, rise to a maximum at a redshift somewhere between 2.5 and 1, and then drop strongly toward smaller redshifts. Integrating this inferred cosmic SFR over time yields roughly the total number of stars presently seen. The striking implication of this result is that we can now observe directly, by looking at these large redshifts, most of the star formation that has ever occurred in the universe, at least in the presently dominant spiral galaxies. It is interesting to note, as Virginia Trimble pointed out, that the apparent peak in the cosmic star formation rate at a redshift of around 2 is reminiscent of the peak in AGN activity that also occurred at about the same time; it is not implausible that these two results could be related, and that both could perhaps reflect an epoch when gas was condensing at a maximum rate into galaxies.

What about the star formation history of elliptical galaxies? We heard contrasting presentations on this subject from Worthey and Dressler. Worthey reviewed studies of the present stellar content of elliptical galaxies, and showed that it is now possible to make some progress with the old problem of separating age and metallicity effects; the mean stellar ages found in this work vary considerably between galaxies, suggesting that many elliptical galaxies have had a long and complex formation history. Dressler disagreed, pointing out that normal-looking elliptical galaxies are seen in about their present numbers in intermediate-redshift clusters, and that these intermediate-redshift elliptical galaxies have red colors indicating that they are already old and therefore must have formed at redshifts significantly greater than 2. Studies of the evolution of the galaxy content of clusters, using Hubble types derived from HST images, show that instead of the elliptical galaxy population having changed, the basic change with time has been the replacement of spiral galaxies in the intermediate-redshift clusters with S0 galaxies in low-redshift clusters. Thus, spiral galaxies appear to evolve into S0s in clusters, just as had originally been proposed to account for the "Butcher-Oemler effect". The disagreement between Worthey and Dressler may not be as large as it seems, however, since Dressler studied only galaxies in giant clusters while Worthey studied galaxies that are mostly not in such clusters, and the latter systems may indeed have had a more extended history of star formation.

Finally, what about star formation in dwarf galaxies? These systems are of interest because they appear relatively simple and primitive, and they could represent or resemble the building blocks of larger galaxies; they also contain relatively large amounts of dark matter. From Hatzidimitriou we heard that the nearby irregular galaxies show evidence for a complex history of star formation characterized by episodes of activity interspersed with inactive periods, a behavior aptly described as "gasping star formation". Given the general lack of organization of these galaxies and their susceptibility to disturbances of internal or external origin, it may not be surprising that their star formation history is also irregular, but it is certainly not yet well understood. From Smecker-Hane we heard that even the tiny dwarf spheroidal galaxies, which are presently gas-free systems, show evidence in their color-magnitude diagrams for multiple episodes of star formation separated by dead periods. The dead periods could not plausibly have been caused by the loss of all of the gas from these small galaxies, since some gas must have remained to fuel subsequent episodes of star formation; possibly the reason for the inactive periods is that the gas surface density became too low during these periods to support star formation.

I shall end my remarks on a much more speculative note with a topic that was mentioned only in passing at this meeting by Silk, but that could become an active subject of discussion at future meetings on star formation. We have talked at length about the formation of the visible stars in the universe, but what about the dark matter? If the recently announced results for microlensing by putative dark objects in the Galactic halo are correct, and if it should indeed be true that part or possibly even all of the dark matter in the halo is in the form of "machos" with masses in the stellar mass range above 0.1 solar masses, then we would have a completely new problem of understanding the origin of these stellar-mass machos. Among known types of objects in the stellar mass range, hydrogen-burning stars have long since been ruled out as dark-matter candidates, and white dwarfs may be excluded too because if most of the dark matter were in the form of white dwarfs, some of them should have been seen. That leaves neutron stars and black holes, both presumably being the collapsed remnants of massive stars formed at very early times. If most of the mass in the universe is really in the dead remnants of massive stars, then most of the star formation that has ever occurred has not yet been directly seen, and it must have produced only massive stars at very high redshifts. Theoretical work that is beginning to appear in the literature suggests possible reasons, based on the expected high temperatures at early times, why the formation of metal-free "Population III stars" at very high redshifts might have produced only massive stars. If there was indeed an early pre-galactic era when massive Population III stars were forming at a high rate, then clearly there is a whole new universe, quite different from the familiar one, still waiting to be discovered and explored at very high redshifts!

Whether or not such speculations have merit, it is certain that much remains to be learned about star formation at all epochs, and that there will be exciting times ahead as we continue to explore the subject and extend our reach. I look forward to attending many more stimulating meetings such as this one has been.