# The Formation and Evolution of the Milky Way

## *The distribution of the chemical elements in our galaxy serves as a "fossil record" of its evolutionary history*

Cristina Chiappini

Our galaxy is a highly evolved entity. Not merely a random assortment of stars, like so many grains of sand on a beach, it is an elegant structure that shows both order and complexity. We know that the Milky Way is a spiral disk galaxy, similar to many others we see in the sky. This surprisingly beautiful shape is so common among galaxies that the universe almost seems to delight in building them. The end product is especially remarkable in the light of what is believed to be the starting point: nebulous blobs of gas. How the universe made the Milky Way from such simple beginnings is not altogether clear. The task of unraveling this mystery has been cast to astronomers, such as myself, who attempt to construct models of the Galaxy's evolution based on its present appearance.

These models need to account for not only the large-scale gravitational forces involved in assembling the Galaxy, but also the chemical composition of its primary components, the stars. It turns out that the chemistry of the stars holds clues to how the Galaxy was made and how it has changed through time. The gas blobs that evolved into the Milky Way consisted merely of hydrogen and helium (and a smattering of lithium), the elements that were created in the Big Bang. All the other elements were literally created by the stars. Unlike the medieval alchemists, the stars can actually transmute one element into another—they are prodigious chemical factories. Nevertheless, even today hydrogen and helium make up about 98 percent of the normal matter in the universe. It's the distribution of the elements that make up the final 2 percent that makes all the difference to studies of galactic evolution.

The most recent models of our galaxy's chemical evolution actually need to incorporate many other observed properties as constraints. These include the density of gas in various parts of the disk, the rate at which stars are born and die, refined measures of the Sun's chemical composition, and the rate at which the elements are produced by the stars, among many others. Astronomers love constraints because without them a model is little more than hand-waving conjecture. The tricky part is coming up with a successful model that incorporates as many constraints as possible.

Although the development of new technologies has improved the quality of the observations and so refined the constraints on astronomers' models, we are still far from a complete understanding of our galaxy's evolution. Like our galaxy, the field itself is still evolving. Here I provide an overview of how astronomers attempt to uncover our galaxy's past, and I introduce a new model that accounts for some of the most recent observations.

### The Anatomy Lesson

As we arbitrarily divide the human body into a torso with a head and limbs, so we can conceptually separate the Galaxy into various components. The flying-saucer shape—consisting of the central *bulge* and the spiral disk—is only the most obvious part of the Galaxy *(Figure 2)*. The spiral disk itself can be subdivided into a *thin disk*, which rises about 1,000 light-years above and below the galactic midplane, and a *thick disk*, which extends to about 3,500 light-years on either side of the plane. The relative flatness of our galaxy is evident when one considers that the galactic disk is generally thought to be about 120,000 light-years across. Our sun resides in the thin disk about 28,000 light-years from the galactic center.

Not seen in any photograph of a spiral galaxy is the spherical *halo* that completely surrounds the disk and the bulge. This is partly because the vast bulk of the halo consists of *dark matter*—material of unknown composition that cannot be seen, but whose presence is deduced by its strong gravitational influence. The halo does have a stellar component, often referred to as the *stellar halo*, but it is simply too dim to be seen as a distinct structure. Lying within the halo, however, are structures that can indeed be seen in

---

*Cristina Chiappini is a research scientist at the Osservatorio Astronomico di Trieste. Aside from her theoretical work on the Milky Way's formation, she also observes planetary nebulae to trace their contribution to the Galaxy's chemical evolution. When she's not doing research, she enjoys dancing and traveling, especially between Italy, Brazil and the United States. Chiappini's address: Via G.B. Tiepolo 11, Trieste - TS, 34131, Italy. Internet: chiappini@ts.astro.it*

Malin/IAC/RGO

Figure 1. Milky Way look-alike galaxies NGC 1232 *(above)* and NGC 891 *(right)* offer a glimpse of how our galaxy would appear if viewed from a distance of several million light-years. The different colors evident in the face-on view of NGC 1232 indicate the existence of separate stellar populations that compose the central bulge *(reddish yellow)* and the spiral disk *(blue)*. The edge-on view of NGC 891 reveals the extreme flatness of spiral-disk galaxies, which typically have an aspect ratio of about 1:30. The flying-saucer shape of a spiral galaxy such as the Milky Way and its distinct stellar populations can be explained by theoretical models of its formation and chemical evolution. (NGC 1232 image courtesy of the European Southern Observatory.)

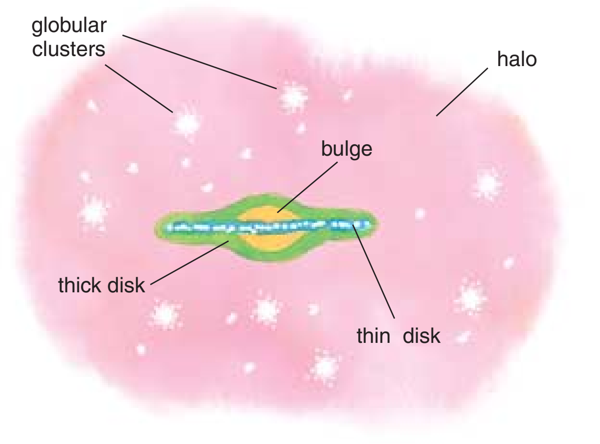

**Figure 2. The Milky Way galaxy has several subregions containing stellar populations that can be distinguished by their chemical composition and orbital dynamics *(see Figure 3)*. The spiral disk (about 120,000 light-years across) consists of an inner thin disk and a slightly fatter thick disk. A large spheroidal halo, (at least 300,000 light-years across) containing both stars and stellar collections called globular clusters, surrounds the disk. On average, the stars of the thin disk are rich in elements heavier than helium (the so-called metals), whereas the halo, bulge and thick-disk stars are metal poor. These distinctions offer clues to the Galaxy's evolution.**

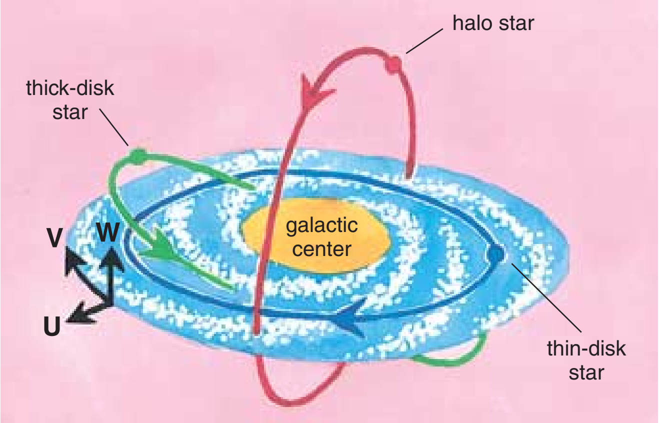

**Figure 3. Kinematic properties of stars—their orbital velocities around the galactic center—differ for the various subregions of the Galaxy. Thin-disk stars tend to have a high rotational velocity *(V)* but a low vertical velocity *(W)*. Thick-disk stars have slightly higher vertical velocities, whereas halo stars tend to have the highest vertical velocities and almost no rotational velocity. Since a star maintains the velocity of the gas in which it formed, the primordial gas that evolved into the different subregions must have had unique origins. A star's radial velocity *(U)*, away from the galactic center, can be added to W and V to describe its orbit.**

the telescope: spheroidal collections of stars known as *globular clusters*. About 200 globular clusters are known, and they appear to be some of the oldest objects in the Galaxy.

It took many decades of careful study to tease apart the various regions of the Milky Way, and the process of dissecting out fine-scale subregions continues even today. One of the reasons it's so difficult is that we cannot measure the properties of all the stars in the Galaxy—they are simply too far away. For the most part astronomers can only give close scrutiny to stars in the solar neighborhood. In evolutionary models this region is generally conceived as a cylinder centered on the Sun, with a radius of 3,000 light-years and "infinite" height (so it includes parts of the thick disk and the halo). Of course, we can't view the stars at the most extreme distances in this cylinder. Even so, we do get a closer view of some thick-disk stars and halo stars because some of them happen to be passing through the thin disk during our era as they orbit the Galaxy.

In fact, the different orbits of the stars, their *kinematic* properties, provide a crucial distinction between stars that belong to different regions of the Galaxy *(Figure 3)*. The kinematics of a star is defined by three velocity components: its rotational velocity *(V)* in the direction of galactic rotation, its vertical velocity *(W)* perpendicular to the galactic plane and its radial velocity *(U)* away from the galactic center. So, for example, stars in the thin disk, such as the Sun, have a small vertical velocity and tend to stay in the galactic plane, whereas thick-disk stars have slightly larger vertical velocities, and halo stars tend to have the largest vertical velocities (and almost no rotational velocity). A star's kinematic properties are one of the ways that astronomers can recognize an interloper from another part of the Galaxy.

As it happens, the stars in the halo and the disk differ in other ways as well and so are said to belong to different stellar populations. The idea of stellar populations was first conceived by the German-born astronomer Walter Baade in 1944. He had been studying the Andromeda galaxy and noticed that the spiral arms are populated by blue stars, which he called, plain enough, *population I*. In contrast, the other parts of the galaxy—the central bulge, the halo and the globular clusters—consist of red stars, which he called *population II*. Although Baade's scheme has since been refined to include various intermediate populations, at the time it served to revolutionize the study of the stars and helped to trigger the modern era of research in stellar evolution and star formation.

Baade's scheme was successful because it held a fundamental truth about what the stars were made of and how they came to be. An analysis of the two populations revealed that population I stars tend to be relatively rich in elements heavier than helium—which astronomers refer to as *metals*—whereas population II stars, especially those in the halo and the globular clusters, are relatively poor in metals.

A star's metallicity is determined by spectroscopic measures of its surface, and it is thought to represent the chemical composition of the gas cloud that collapsed to form the star. (Astronomers can also measure the metallicity of ambient gas clouds directly.) In general, the metallicity of a star is defined by the abundance of iron (Fe) compared with its hydrogen (H) abundance. This relation is normalized to the solar abundance on a logarithmic scale:

$$[\mathrm{Fe/H}] = \log(\mathrm{Fe/H}) - \log(\mathrm{Fe/H})_{\mathrm{Sun}}$$

The most metal-poor star ever observed in our galaxy is located in the halo. It is old and has a metallicity [Fe/H] of about –4.0, or about 10,000 times less than the Sun! That it happens to be an ancient star is not a coincidence. When it was born the stellar chemical factories were only just beginning to start operation, so there was simply not a great abundance of metals that could be incorporated into the star. As a general rule [Fe/H] increases with time so old objects are more metal poor than young ones.

How stars make "metals" is now reasonably well understood. For the most part they are formed by a chain of fusion of lighter elements. There are actually several astrophysical processes involved, each of which forms a different assortment of elements *(Figure 4)*. How and when these processes take place is largely dependent on the mass of a star.

The lightest stars, some having merely one-tenth the mass of our sun, live the longest, potentially for many billions of years. In contrast, the heaviest stars, weighing up to 150 solar masses, have comparatively brief lives, on the order of a few million years. The distinction is crucial because it is primarily at the end of its life that a star makes its contribution of newly synthesized elements to the Galaxy.

The mass of a star determines not only its lifespan, but also the types of chemical elements it will contribute to the interstellar gas that will form the next generation of stars. Since the very-low-mass stars can be as old as the Galaxy itself (about 14 billion years), they contribute very little to the chemical evolution of the Galaxy. Low- and intermediate-mass stars, such as our sun, die by ejecting an outer envelope of material into the interstellar medium—forming structures known as planetary nebulae—mostly containing helium-4, carbon and nitrogen (see "The Shapes of Planetary Nebulae," July–August 1996). The most massive stars (more than eight solar masses) end their lives in a more violent way, exploding as type II supernovae. These stars enrich the Galaxy with several elements, but mainly with oxygen and other so-called *alpha elements*—neon, magnesium, silicon and sulfur—which are formed by the fusion of alpha particles (helium-4 nuclei).

There is another type of exploding supernova that also seeds the Galaxy with elements. This is the type Ia supernova. This explosion involves a binary system in which a white dwarf star and an intermediate-mass star (a red giant) orbit each other (see "White Dwarf Stars," November–December 2000). The two stars are so close to each other that the white dwarf gradually pulls a considerable amount of material from the outer envelope of the expanding red giant. At a certain point the white dwarf will acquire so much mass that it collapses under its own weight and produces an explosion that blasts the bulk of its material into the interstellar medium—mostly in the form of iron, but also some sulfur, silicon and calcium. Such explosions contributed about 70 percent of the iron we see today in the Galaxy.

The rate at which iron is produced in the Galaxy depends on the masses of red giants in these binary systems and their total numbers. The heaviest intermediate-mass stars (about 8 solar masses) can reach the red-giant phase within 30 million years, whereas stars like the Sun will take about 10 billion years. But since stars of different

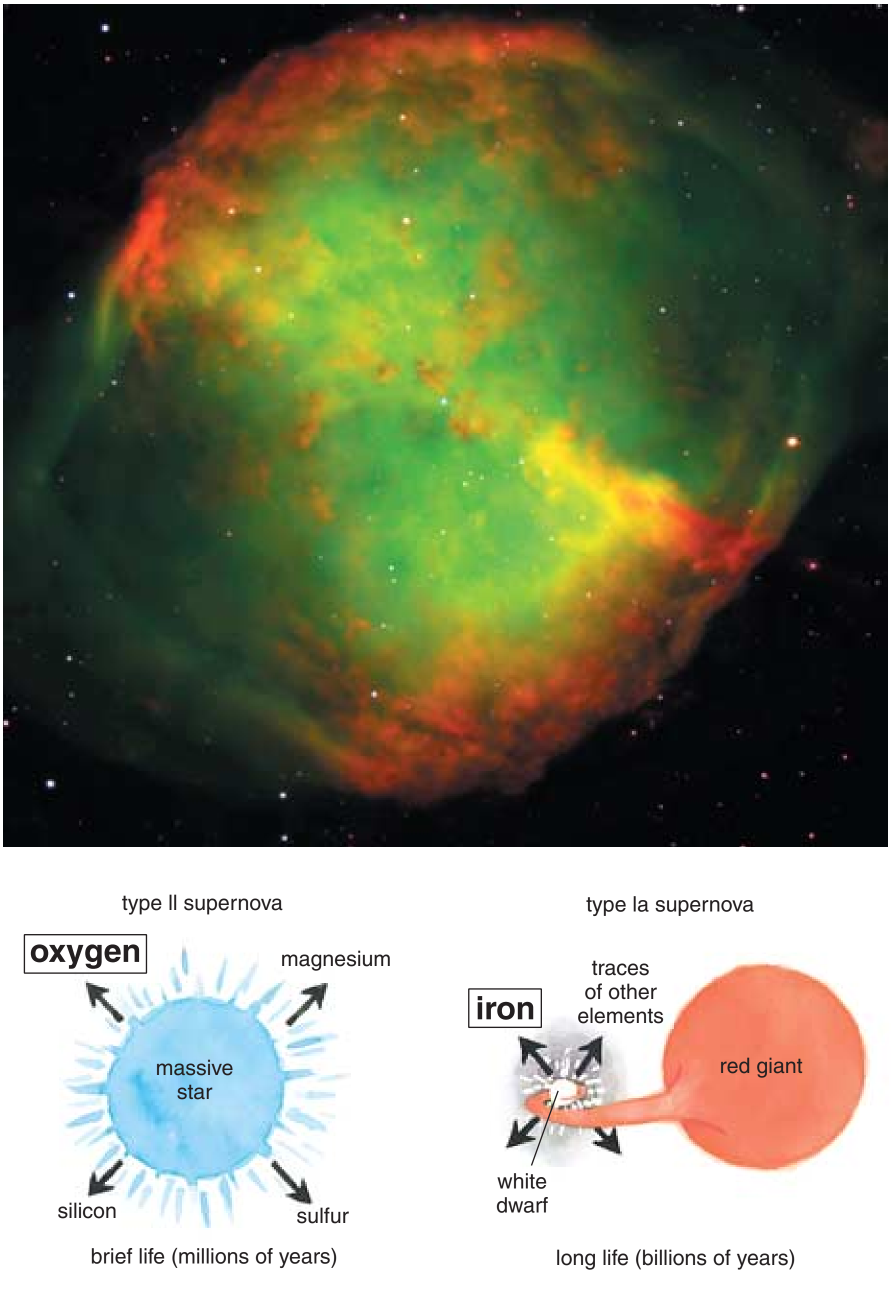

**Figure 4. Three astrophysical processes contribute the bulk of the chemical elements to the interstellar medium. Low- and intermediate-mass stars, such as our sun, cast off most of their elements (notably carbon and nitrogen) near the end of their lives in gassy exhalations called "planetary nebulae"** ***(top, Dumbbell Nebula, M27).*** **Massive stars** ***(lower left)*** **end their brief lives in type-II supernova explosions, seeding the Galaxy with many elements, especially the "alpha" elements, such as oxygen. A type Ia supernova** ***(lower right)*** **detonates when a white dwarf star collects a critical amount of mass from a giant companion. Type Ia supernovae are responsible for about 70 percent of the iron in the Galaxy and typically require about one billion years for binary systems to mature before they explode. Because each of these processes takes place on a different timescale, the relative abundance of different chemical elements in a particular region of the Galaxy offers clues to the rates of star formation and the region's evolutionary history** ***(see Figure 5).*** **(M27 image courtesy of the European Southern Observatory.)**

masses aren't produced in equal numbers, astronomers must also consider the *initial-mass function*—the probability that a newborn star will have a certain mass. For example, a star with the mass of the Sun is about 150 times more common than a star of 30 solar masses. When all of these factors are taken into consideration, it turns out that iron enrichment is a relatively slow process. The typical binary system needs to age for a billion years before a type Ia supernova explodes.

These three processes hold an important key to understanding the evolution of the Milky Way precisely because they occur on very different timescales.

### A Cosmic Clock

Let's consider the rate at which the elements are produced in the Galaxy. The interstellar medium will be enriched faster in elements produced by short-lived stars (that is, the most massive ones) and more slowly in those elements produced essentially by type Ia supernovae and the low- and intermediate-mass stars. So the ratio of two elements—such as oxygen (O) and iron—that are returned to the interstellar medium on different timescales can be used as a "clock" when compared to the general metallicity [Fe/H] of that part of the Milky Way. By measuring specific abundance ratios in stars from different parts of the Galaxy, astronomers can discover how fast the metal enrichment proceeded and the timescale over which the region was formed.

A little thought should reveal that early in the evolution of our galaxy the primary sources of iron (the type Ia supernovae) had yet to make the bulk of their contribution because it takes a good billion years for most of these systems to reach maturity. On a plot of [O/Fe] versus [Fe/H], we would expect the early history of the Galaxy to have a nearly flat relation between oxygen and iron (forming a "plateau") since these elements are, at first, created at the same rate inside type II supernovae *(Figure 5).* However, when the consummate iron producers, the type Ia supernovae, start to make the bulk of their contribution, the [O/Fe] ratio should drop (as the denominator increases), and so the slope of the line decreases.

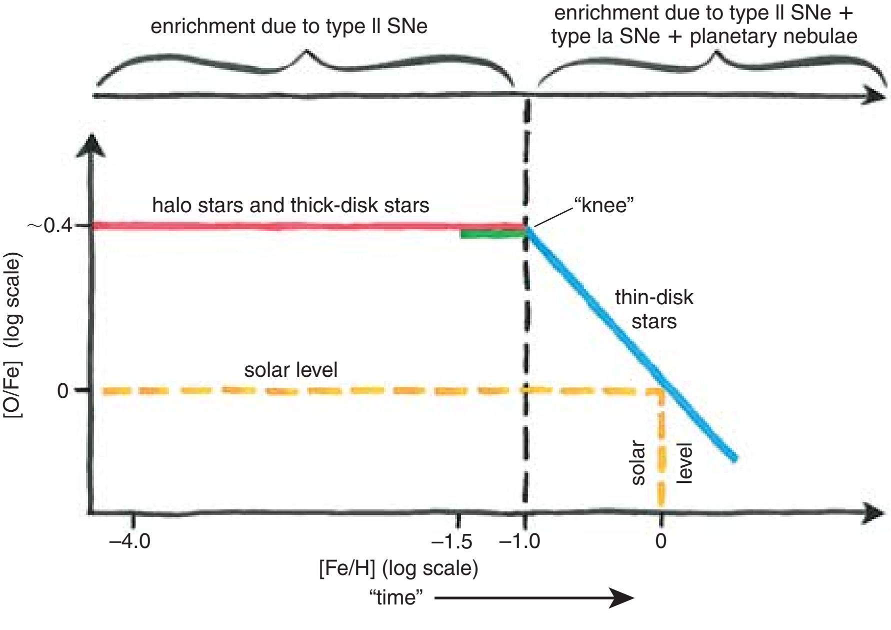

**Figure 5. The general metallicity of the Galaxy—as measured by the abundance of iron (Fe), compared with hydrogen (H)—increases with time *(abscissa)* and so serves as a basis for comparing the relative abundances of two elements (such as oxygen (O) and iron; *ordinate)* that are created on different timescales. A plot of these quantities reveals a "plateau" of metal-poor stars (metallicity less than –1) that drops at a "knee" as the relative proportion of iron in the Galaxy increases. Since type Ia supernovae (SNe) are the primary source of iron, astronomers believe that the "knee" occurred about one billion years after the Galaxy began to form *(see Figure 4)*. The halo stars (*red line)* and some of the thick-disk stars (*green line)* tend to occupy the "plateau," whereas thin-disk stars *(blue line)* occupy the descending slope. These observations suggest that the halo and part of the thick disk were formed in the first billion years of the Galaxy's evolution, and the thin disk formed later.**

Now we come to an interesting observation: On a graph that plots this relation, the plateau is occupied by stars of the halo and the thick disk, whereas the descending slope consists of stars in the thin disk. The point at which the slope starts to fall, the "knee," is a critical indicator of when most of the type Ia supernovae started to enrich the interstellar medium with iron. Since it takes one billion years for this to happen, we know that the halo must have formed within the first billion years of the Galaxy's life, whereas the thin disk in the solar vicinity formed more slowly. The few bulge stars that have been measured also reside on the graph's plateau, suggesting that the bulge too formed early in the Galaxy's history.

So the halo is old and populated by old, metal-poor stars. But where are the halo's young stars—those massive blue stars (Baade's population I) that richly populate the thin disk? They're missing because the building material, the gas, needed to make a star has been used up. The galactic disk, on the other hand, appears to have plenty of gas left and is still a place of vigorous stellar birth (see "The Formation of Star Clusters," May–June 1998 and "Protostars," July–August). This distinction is a key ingredient in models of the Galaxy's chemical evolution.

It's not fully understood why the galactic disk is still in its active phase of star formation. One of the ideas that has been proposed is that gas *infall* onto the galactic disk continues to provide a source of fuel.

The contrary notion, often called the *simple model,* holds that the Milky Way is a closed box in which gas neither enters nor leaves—there is no infall. The simple model has been rejected because of an observation, one which astronomers refer to as the *G-dwarf Problem*. As the name suggests, G-dwarfs are small stars, and because of their low mass they can live for many billions of years—some date to the earliest era of the Milky Way's formation. If the galactic disk has maintained a constant mass from the beginning, there should be a fairly large number of metal-poor G-dwarf stars in the solar vicinity simply because there was a large amount of metal-poor gas from which the stars could be made early in the Galaxy's history. In reality, there are relatively few metal-poor G-dwarf stars. One way to solve this problem is to assume that the galactic disk originally had less mass than it does now. With time it acquired more mass from the infall of gas, a phenomenon that I'll consider in greater detail below.

The simple model is effectively a straw man that was made to be knocked down by the G-dwarf problem. Realistic models of the Galaxy's evolution are considerably more sophisticated.

### Making the Milky Way

The granddaddy of galactic-formation models was conceived in the early 1960s by three astronomers: Olin Eggen, Donald Lynden-Bell and Allan Sandage. Their 1962 publication played a seminal role in the field and is now simply referred to by the authors' initials: *ELS*. The ELS model was based on the relative velocities and chemical compositions of stars in populations I and II. As I described earlier, the population I stars are relatively rich in metals, and they follow orbits in the plane of the galactic disk. In contrast, the metal-poor population II stars in the halo follow elliptical orbits that cut across the plane of the Milky Way.

These distinctions could be explained, said ELS, by the way in which the Galaxy formed *(Figure 6)*. According to ELS, the Milky Way began as a spherical cloud of gas—a protogalaxy—that was born collapsing toward its center. The original gas was poor in metals, and so stars formed as the cloud was collapsing would also be metal poor. These newly made stars maintained the kinematic properties of the gas in the collapsing cloud, and so followed eccentric orbits around the center of the Galaxy, forming the population II stars of the halo and the globular clusters. As the cloud contracted, some of its energy would have been lost to heat in a *dissipative collapse*. The rotational speed of the collapsing cloud would also increase due to the conservation of angular momentum (which is a function of rotational velocity and radius). Such changes would induce the cloud to collapse preferentially along its rotational axis, so that it would become progressively flatter—and thus form a disk. The gas in the flattened disk would be enriched in metals produced by supernovae from the first generation of stars. Like their counterparts in the halo, stars formed in the flattened disk would preserve the metallicity and kinematics of the gas at the time of their birth, and so form the population I stars. All of this took place within 300 million years according to ELS.

In the decades that followed, a number of observations indicated that the Galaxy could not have formed in such a rapid collapse. The ELS model, as originally proposed, could not be right. One notable alternative was suggested by the American astronomers Leonard Searle and Robert Zinn in 1978. Searle and Zinn had been studying the globular clusters in the galactic halo and noticed a wide discrepancy in the metallicity of these objects. According to their metallicities, some globular clusters appeared to be significantly older than others. The spread in the globular clusters' ages meant that they could not have been formed in the relatively brief timescale proposed by ELS.

Instead of a single-cloud collapse, Searle and Zinn proposed that the halo of the Milky Way formed by the aggregation of many cloud fragments, each of which may have already formed stars and globular clusters *(Figure 7)*. Since the fragments had independent evolutionary histories, they could form objects of varying ages. In some sense the Searle and Zinn model has been confirmed by observations that show that small, or "dwarf," galaxies continue to collide with the Milky Way to this day. These dwarf galaxies may have evolved from the cloud fragments that failed to become part of the Milky Way early in its evolution. The Sagittarius dwarf galaxy, which was discovered in 1996, appears to be just such a fragment. Over the course of billions of years it oscillates back and forth through the galactic plane, and with each pass it loses some of its mass. In time it will be completely consumed.

Other authors have proposed various *serial* and *parallel* models of the Galaxy's formation. In a serial model, the Galaxy forms as a continuous process during a single infall event. The halo represents the early phases of the process, and the disk forms only after the halo is completed. The ELS model is sequential in this manner, except that everything is formed very quickly. In contrast, parallel models assume that the various galactic components started forming at the same time from the same gas, but then evolved at different rates according to their respective star-formation histories.

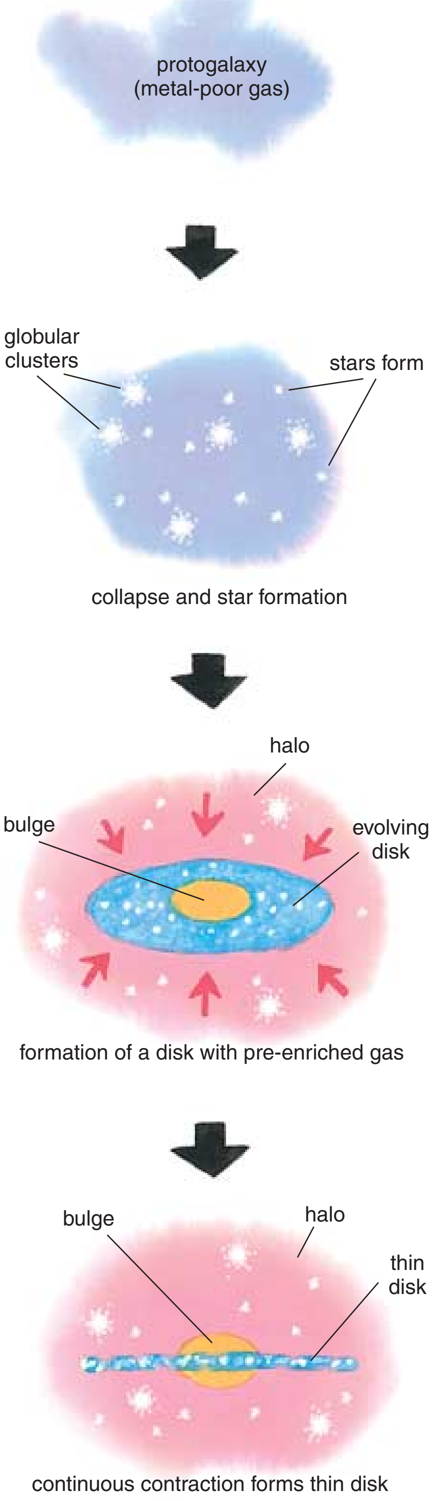

**Figure 6. The "ELS" model holds that the Milky Way formed from the rapid collapse of a single cloud of gas. Stars formed early in the collapse maintained the dynamics of the metal-poor gas and so now travel around the Galaxy in elliptical orbits within the halo. As the cloud collapsed *(red arrows)* preferentially along its rotational axis, it formed a disk that had been enriched with the metals produced by the early generations of halo stars.**

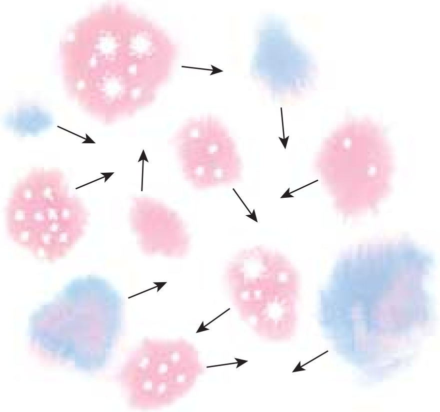

**Figure 7. The Searle and Zinn model proposes that the Milky Way formed from an aggregation of several cloud fragments. This model helps to explain the observed differences in the metallicity of globular clusters in the galactic halo. Since each of the cloud fragments had independent histories, some may have evolved more than others, and so have produced objects of greater metallicity.**

## A Halo-Disk Discontinuity?

New observations suggest that none of the early models holds a complete explanation of how the Milky Way was made. In particular, models such as ELS suggest that the formation of the disk involved a smooth dissipational collapse of the halo. Such models also assume a continuous evolutionary transition in the formation of the thick disk and the thin disk. It appears, however, that our galaxy's formation was neither smooth nor continuous.

According to Rosemary Wyse of Johns Hopkins University and Gerard Gilmore of the Institute of Astronomy in the United Kingdom, the halo and the thin disk are distinct entities that could not have formed from a single cloud of gas. They base their distinctions on the angular momenta of the Galaxy's stellar populations. They show that the halo and the bulge tend to consist of low-angular-momentum stars, whereas the thick disk and the thin disk typically contain stars with a high angular momentum. Since angular momentum is conserved, these distinctions reflect the intrinsic characteristics of the parent gas from which the stars evolved. So these galactic components must have originated from separate clouds of material with different angular momenta.

There is also evidence that the rate of star formation has not been continuous in the Galaxy's history. Observations by Raffaele Gratton, of the Astronomical Observatory of Padova, Italy, and his colleagues, suggest that the rate of star formation decreased suddenly in the solar neighborhood fairly early in the Galaxy's evolution. Gratton and his colleagues studied the relative chemical abundances of iron compared with two alpha (α) elements (oxygen and magnesium) for stars in the halo, the thick disk and the thin disk. At a certain point in the Galaxy's history, as measured along an [α/H] timeline, there appears to be a "gap" during which almost no alpha elements were produced *(Figure 8)*. This is evident as a sudden increase in [Fe/α] while [α/H] remains constant. The identity of the stars on either side of the gap suggest that star formation effectively stopped after the formation of the halo/thick disk (which are both very old) but before the thin disk formed.

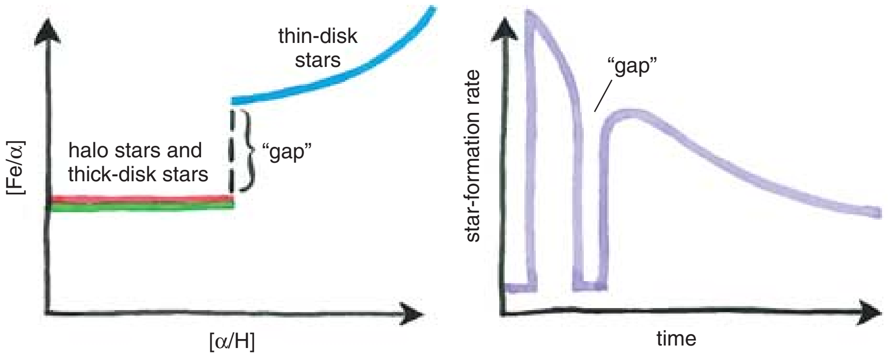

**Figure 8. A "gap" in the relative abundances of iron and the alpha (α) elements *(left)*, such as oxygen, is interpreted as a period during which the star-formation rate in the Galaxy decreased *(right)*. This is because the alpha elements are produced by the type II supernovae, which are effectively indicators of the star-formation rate. The gap in star formation appeared to occur after the halo and the thick disk literally "ran out of gas," and did not increase again until newly accreted gas settled down to form the thin disk.**

The duration of this gap can also be deduced. Since the alpha elements are produced by the type II supernovae, which are the explosions of short-lived stars, their rate of production is effectively a measure of the star-formation rate. On the other hand, the quantity of iron actually increased during this time because the binary systems that produced the type Ia supernovae were created long before the gap in star formation. Given the typical maturation period of type Ia supernovae, the data suggest that the gap lasted no more than a billion years.

By studying the kinematics of these same stars, Gratton's team identified three distinct populations. One population made up the halo, part of the thick disk and perhaps the bulge stars (which originated from the dissipative collapse of part of the halo). Another population of stars made up the thin disk, which resulted from an extreme dissipative collapse of the disk. And the third population consisted of a relatively small number of stars in the thick disk that had a unique origin. This third population of metal-poor stars (with [Fe/H] less than –1.0) probably formed in satellite galaxies and was then added to the Milky Way during the gap in star formation. In this view the thick disk actually has two components.

Other scientists have also found that the thick disk and the thin disk are kinematically distinct. Timothy Beers of Michigan State University and Jesper Sommer-Larsen of the University of Copenhagen studied the kinematics and composition of a large sample of metal-poor stars. Their analysis suggests that the most metal-poor component of the thick disk had its origin in a major accretion event.

How can we explain such an event? In one scenario, a satellite galaxy collided with the galactic disk when the thin disk was still mostly gaseous. The thin disk was heated—some of its matter was scattered—by the collision, and this formed the metal-poor part of the thick disk. The bulk of the thin disk resettled into the midplane and formed a "new" thin disk. The timing of this event is thus constrained by the age of the oldest star in the thin disk, which in the solar neighborhood is about 10 billion years. It's possible that the gap in star formation noted by Gratton and his coworkers was a result of this collision, and this marks the disk-halo discontinuity.

All told, the collective evidence now suggests that low-angular-momentum material formed the stellar halo and the bulge in a rapid, dissipative collapse (especially in the innermost regions) in the manner suggested by ELS. Mergers with dwarf galaxies *(à la* Searle and Zinn) would also have contributed to the halo, but most of the major mergers would have happened before the formation of the thin disk. In fact, since the thin disk is quite fragile, its existence implies that mergers could not have contributed more than a few percent of its mass in the past five billion years. In stark contrast to ELS, the thin disk evolved independently of the halo, from gas with a high angular momentum. The first thin disk was thickened by the last major galactic merger about 10 billion years ago. As more gas was added to the disk, it settled into a new thin disk—the one in which our sun formed.

These ideas are quite new and still hotly discussed.

### A Two-Infall Model

In the light of these recent observations, my colleagues and I have developed a new model that seeks to account for the distribution of stars seen in the halo and the thin disk. Our *two-infall* model assumes that an initial collapse formed the halo (and probably part of the thick disk). Star formation in the halo continued until the gas density dropped below a certain threshold. In our model the halo "runs out of gas" because of an extremely efficient rate of star formation—the number of stars formed per unit of time was considerably higher than it is now. Gas lost by the halo accumulates in the center and so forms the bulge. After the halo forms and star formation ceases, a second infall event forms the thin disk. This event was either a result of a merger with a small galaxy, or perhaps due to the longer time required for material with a high angular momentum to fall. As suggested by Wyse and Gilmore, the evolution of the halo and the disk are almost entirely independent.

Our model also predicts an age for the formation of the thin disk. Until now, astronomers knew it took longer than one billion years to make the thin disk, but we didn't know whether the bulk of what we observe in the solar neighborhood was formed in two billion years or eight billion years. The

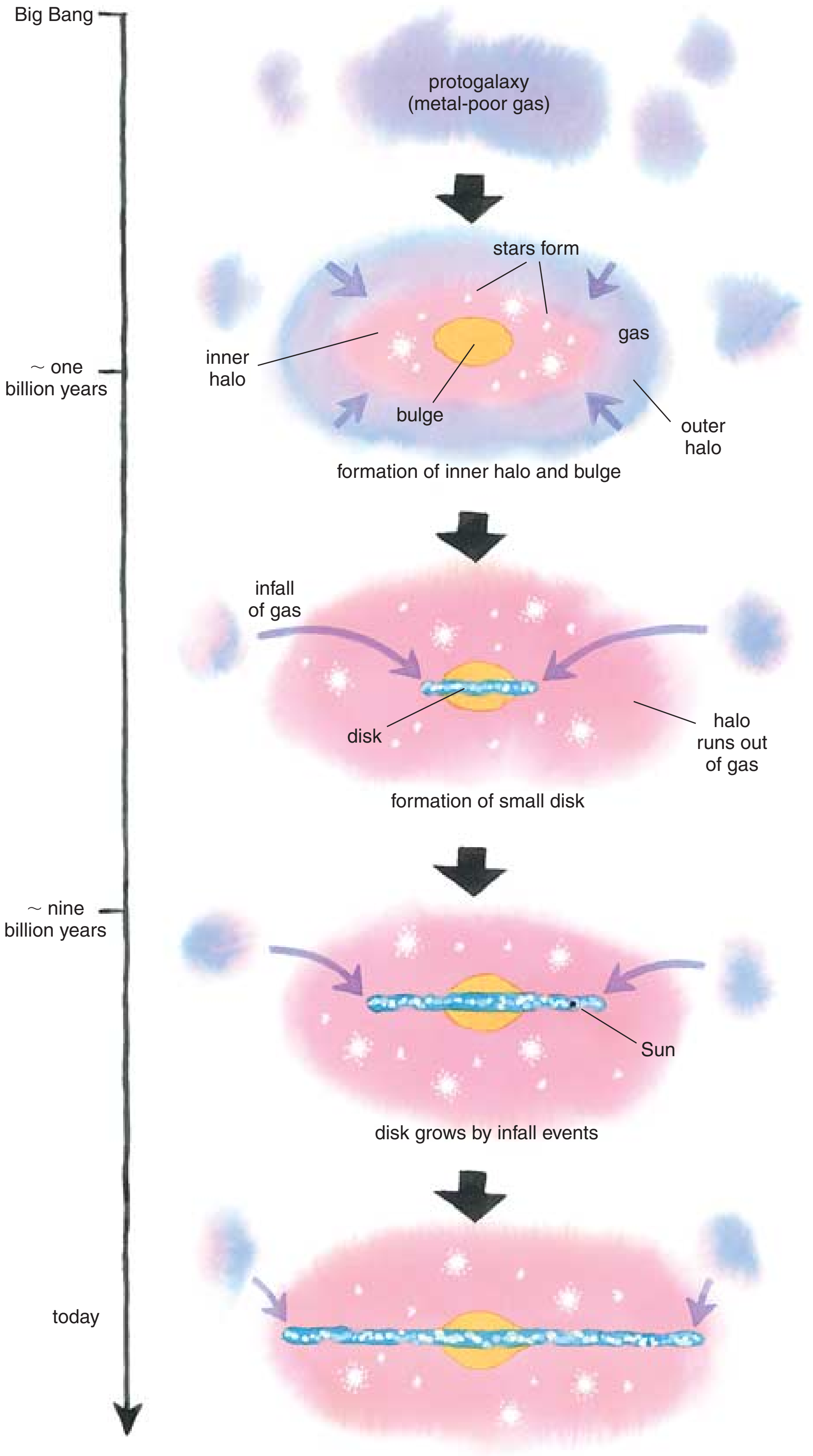

**Figure 9. The "two-infall" model proposes that the halo and the disk of the Milky Way formed from separate bodies of gas at different times in the Galaxy's evolution. The halo and the bulge formed first from a metal-poor gas cloud with low angular momentum in the first one billion years. In contrast, the disk formed later by the infall of high-angular momentum gas. The disk also appears to be evolving "inside-out," with the central-most regions forming first. The formation of the solar neighborhood, which began about 10 billion years ago, was completed when the disk was about seven billion years old, whereas the outer parts of the disk continue to grow even today with the infall of extragalactic gas clouds.**

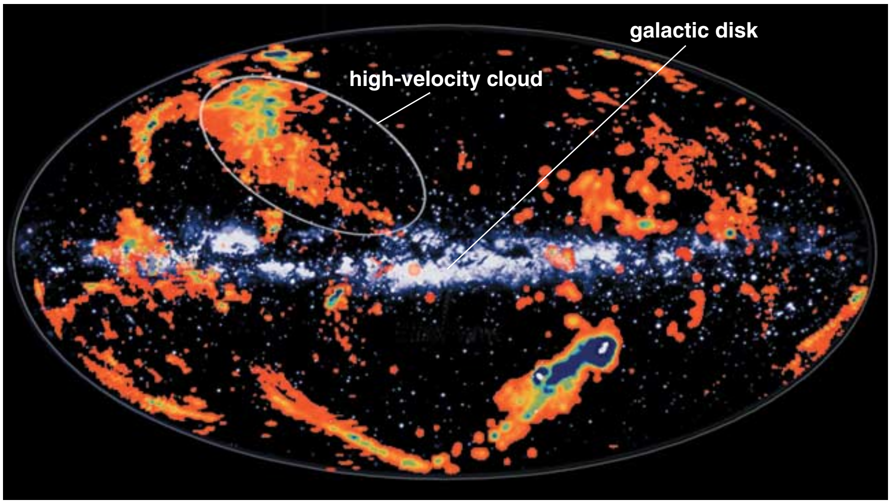

**Figure 10. High-velocity clouds, consisting of nebulous blobs of gas, are falling onto the galactic disk from the galactic halo. One interpretation holds that these clouds are evidence that infalling primordial matter maintains the rate of star formation in the galactic disk. Observations suggest that such clouds are indeed replenishing the Galaxy's gas supply at a pace that explains the current rate of star formation in the solar neighborhood—about one new star every year. (Image courtesy of Bart Wakker, University of Wisconsin-Madison, and NASA.)**

most reliable way to assess the age of the thin disk is to use the constraint imposed by the metallicity distribution of G-dwarfs in the solar neighborhood.

A certain number of G-dwarf stars are born in every generation of star formation, so the metallicity of each new generation increases as the interstellar medium is enriched by the stellar deaths of previous generations. Since their lifetimes are so long, they should all still be observable. With the help of chemical-evolution models that include factors such as the star-formation rate and the return of metals to the interstellar medium, the metallicity distribution of the G-dwarfs allows us to infer the rate at which the Galaxy is enriched and thus how long it took to complete the formation of the thin disk at the Sun's position.

Until 1995, chemical-evolution models had based their results on G-dwarf metallicity distributions that were originally published in 1975 (and then slightly modified as recently as 1991). Based on these distributions, the old models usually indicated that the thin disk in the solar vicinity was formed within three billion years. Since this was not much longer than the timescale assumed for the formation of the halo (around one billion years), the old models justifiably assumed that the disk formed from the halo gas.

After 1995, however, two independent groups of scientists revised the G-dwarf metallicity distributions with more precise data, based on new observations and modern spectroscopic techniques. The predictions of the older models do not fit the newer data. Furthermore, one of the problems with models that assumed the disk formed from the halo gas was that it overestimated the halo-to-disk mass ratio. The observed value is about 1:20, whereas the one-infall models usually predicted something on the order of 1:5. The problem is solved if we allow a longer timescale for the formation of the thin disk.

The two-infall model allows for this, of course, because the addition of mass to the disk can occur much later during the second infall event. In fact, based on the new G-dwarf metallicity distributions, our model suggests that it took seven billion years to complete the formation of the thin disk in the Sun's vicinity. This is considerably longer than any previous model has suggested, and it indicates that the disk could not have been formed from the halo gas, but formed mainly from extragalactic gas.

The evolutionary histories of the disk and the halo are indeed independent—for the most part. Some evidence indicates that the thin disk did not form all at once. Chemical abundances in different parts of the disk suggest that there is a radial metallicity gradient, so that the inner regions of the disk are older than the outer parts. This suggests an "inside-out" formation of the thin disk. My colleagues and I have explored the significance of this metallicity gradient considering the star-formation rate and the radial distribution of stars and gas in the Galaxy. Our results show that the outer parts of the disk are indeed relatively poor in metals, and this suggests that metal-poor halo gas has contributed to the formation of the outer regions of the disk. In contrast, the inner disk appears to have evolved independently of the halo. We propose that the outermost parts of the disk are still being formed, and what we see in the outer regions of the disk may actually be a mixture of halo and disk components.

If the outer disk is indeed still forming, we might even be able to observe the infall of gas clouds right now. Leo Blitz of the University of California, Berkeley, argues that this may be the significance of *high-velocity clouds*. These objects, which are basically blobs of gas, have been known to astronomers for more than four decades. Their velocities indicate that they are indeed falling toward the disk, but not everyone agrees on their significance. Some astronomers believe that the high-velocity clouds were originally ejected from the disk during supernova explosions—producing phenomena known as *galactic fountains*—and are now returning to the disk. One way to distinguish between these two possibilities is to measure the chemical abundances of the high-velocity clouds. If these objects represent the infall of primordial gas, they should be metal poor; if they are the products of galactic fountains, they should be quite rich in heavy elements. Whatever their origins, high-velocity clouds are indeed merging with our galaxy, replenishing its gas supply at a pace sufficient to produce about one new star every year—roughly the observed rate in the solar vicinity.

## The Future

We obviously need more observations before we can refine our models of the Milky Way's evolution. For one thing, we are still uncertain about the formation timescale for the thin disk outside the solar neighborhood because we lack precise observational constraints. One of the more promising approaches in this regard is now on the horizon, however. The element deuterium (a hydrogen isotope consisting of a proton and a neutron) is a very sensitive chemical marker of the gas consumption in a given locale. All of the

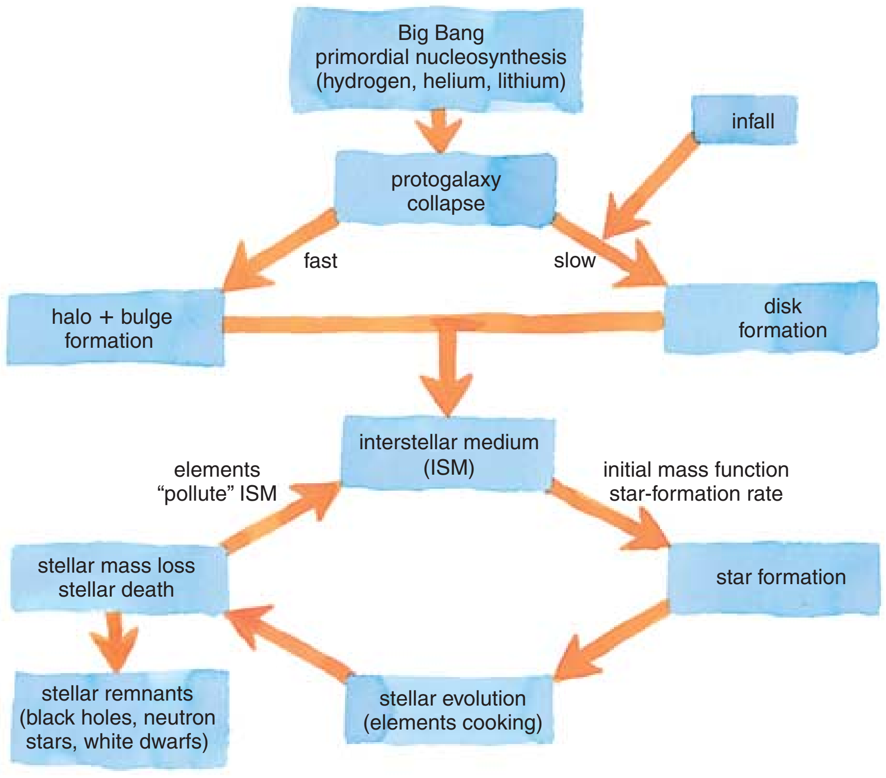

**Figure 11. Many factors must be considered in the construction of models that explain the chemical evolution of our galaxy. Not all of these factors are fully understood at this time, which limits the "resolution" of the models—how much of the Galaxy's evolution they can explain. The development of new observatories, however, promises to refine our measures of the Galaxy's chemical and kinematic fine structure, and so our understanding of the processes involved in its evolution.**

deuterium in the universe was created in the Big Bang and none has been released into the interstellar medium since, but because it's consumed inside stars its abundance steadily decreases in direct proportion to the rate of star formation. A measure of its abundance throughout the Galaxy would give us an idea of how fast the inner and outer parts of the disk have evolved.

At the moment the only measure of deuterium in the Galaxy outside the solar neighborhood was that recently reported for the galactic center by Donald Lubowich, of the American Institute of Physics in New York, and his colleagues. They found the deuterium abundance to be the lowest ever measured, about nine times less than that found in the solar neighborhood. This result is consistent with the inside-out formation hypothesis, but as yet we have no measure of the deuterium abundance in the outer regions of the disk. This may soon change, however, with the data collected by the Far Ultraviolet Spectroscopic Explorer (FUSE) satellite. The FUSE satellite, which is now in orbit, is measuring the abundance of deuterium throughout the Galaxy.

Farther in the future, the GAIA project, which is slated to be launched by the European Space Agency by 2012, holds great promise for our attempts to solve the puzzle of the Milky Way's formation. The GAIA satellite will be taking a massive stellar census, measuring the positions, motions and chemical compositions of more than a billion stars. It will, in effect, provide a three-dimensional map of our galaxy with unprecedented accuracy and resolution.

For now, however, a full understanding of our galaxy's evolution remains elusive. The current state of affairs was nicely summed up recently by Sandage: "The study of origins is the art of drawing sufficient conclusions from insufficient evidence."

## Bibliography

Barbuy, B., S. Ortolani, E. Bica and S. Desidera. 1999. The metal-rich bulge globular cluster NGC 6401. *Astronomy and Astrophysics* 348:783–788.

Burton, W. B., and R. Braun. 1999. Clouds from near and far. *Nature* 402:359–361.

Buser, R. 2000. The formation and early evolution of the Milky Way Galaxy. *Science* 287:69–74.

Chiappini, C., F. Matteucci and R. Gratton. 1997. The chemical evolution of the Galaxy: The two-infall model. *The Astrophysical Journal* 477:765–780.

Chiappini, C., F. Matteucci, T. Beers and K. Nomoto. 1999. The earliest phases of galaxy evolution. *The Astrophysical Journal* 515:226–238.

Chiappini, C., F. Matteucci and D. Romano. 2001. Abundance gradients and the formation of the Milky Way. *The Astrophysical Journal* 554:1044–1058.

Croswell, K. 1995. *The Alchemy of the Heavens*. New York: Anchor Books.

Eggen, O. J., D. Lynden-Bell and A. R. Sandage. 1962. Evidence from the motions of old stars that the galaxy collapsed. *The Astrophysical Journal* 136:748–766.

Gilmore, G., and R. F. G. Wyse. In press. The thick disk-halo interface. In "Dynamics of Star Clusters and the Milky Way," ASP Conference Series 228, ed. S. Deiters, B. Fuchs, A. Just, R. Spurzem and R. Wielen. http://xxx.lanl.gov (astro-ph/0104242).

Gratton, R., E. Carretta, F. Matteucci and C. Sneden. 2000. Abundances of light elements in metal-poor stars. IV. [Fe/O] and [Fe/Mg] ratios and the history of star formation in the solar neighborhood. *Astronomy and Astrophysics* 358:671–681.

Lubowich, D. A., *et al.* 2000. Deuterium in the galactic centre as a result of recent infall of low-metallicity gas. *Nature* 405:1025–1027.

Matteucci, F. 2001. *The Chemical Evolution of the Galaxy*. New York: Kluwer.

Pagel, B. E. J. 1997. *Nucleosynthesis and Chemical Evolution of Galaxies*. Cambridge: Cambridge University Press.

Searle, L., and R. Zinn. 1978. Compositions of halo clusters anad the formation of the galactic halo. *The Astrophysical Journal* 225:357–379.

Wyse, R. F. G., and G. Gilmore. 1995. Chemistry and kinematics in the solar neighborhood: Implications for stellar populations and for galaxy evolution. *The Astronomical Journal* 110:2771–2787.