Star Formation History and IMF in the Galaxy

GERARD GILMORE\textsuperscript{1} \\
& ROSEMARY WYSE\textsuperscript{2}

\textsuperscript{1} Institute of Astronomy, Cambridge, UK. \\
\textsuperscript{2} Department of Physics & Astronomy, The Johns Hopkins University, Baltimore, USA.

Abstract

The rate of star formation during early evolution of the Milky Way can be derived from abundance ratios of old stars. For the Galactic halo, the small scatter in and constancy of \([O/Fe]\) through effectively all the halo formation phase requires the star formation rate to have been within a factor of a few of the mean rate, some \(10 \, M_\odot \, yr^{-1}\), at all times. Similar limits apply to the Galactic disk. Constancy of \([O/Fe]\) also requires the massive star IMF to have been approximately constant over the whole abundance range for which significant numbers of stars exist. 

The halo gas is lost quickly to the central regions of the Galaxy. The halo does not pre-enrich the disk, invalidating the concept of a "Solar Cylinder". Rather, there is a "Solar Sausage". Metal enriched gas fountaining from the disk provides the "halo" gas seen as QSO absorption lines. To make this consistent with observed QSO absorption line properties requires the disk to have formed early, perhaps by \(z \gtrsim 2\).

1 Theoretical Preconceptions

There are three natural characteristic timescales which may be associated with the halo of a galaxy like the Milky Way. The first of these is the cooling time, which is the time for radiative processes to remove the internal energy of a cloud. Defining the cooling rate per unit volume to be \(n^2 \Lambda(T)\), where \(n\) is the particle number density, and where the form of \(\Lambda(T)\) includes the contributions from free-free, bound-free, and bound-bound transitions, and thus is an implicit function of chemical abundance as well as being an explicit function of temperature \(T\), gives

\[
t_{\text{cool}} = \frac{3nkT}{n^2\Lambda(T)}. \tag{1}
\]

The second natural timescale is the gravitational free fall collapse time of the system, which is the time it would take for the system to collapse upon itself if there were no pressure support. This depends only on the mean density of the system, and is

\[
t_{ff} \sim 2 \times 10^7 n^{-\frac{1}{2}} \, yr. \tag{2}
\]

The third timescale is the collision time between gas clouds in the halo, in the case where one imagines a halo to be made of \(N_{cl}\) independent clouds. This timescale is then a measure
of the longevity of a cloudy halo against disruptive (and dissipative) cloud-cloud collisions. Particularly for the present discussion, it is a measure of the timescale on which QSO absorption line clouds can remain associated with the Galactic halo. The (absorption line) cloud collision time is \( \text{(York et al. 1986)} \mid 14 \)

\[
T_{\text{coll}} \sim \frac{1}{N_{\text{cl}} v \sigma} \sim 5 \times 10^8 \left( \frac{R}{50 \text{kpc}} \right)^3 N_{\text{cl}}^{-1} \left( \frac{v}{100 \text{kms}^{-1}} \right)^{-1} \text{yr.} \tag{3}
\]

For the parameters of the halo of the Milky Way all three timescales are similar.

\[
t_{\text{cool}} \sim t_{\text{ff}} \sim t_{\text{coll}} \lesssim 1 \text{Gyr.} \tag{4}
\]

By comparison, the natural timescale for disk formation is \( \sim 4 \text{Gyr} \). This is deduced from the requirement that the disk mass must have fallen in from \( \sim \lambda^{-1} \alpha_{\text{disk}} \text{kpc} \), with \( \alpha_{\text{disk}} \) the disk radial scale length (\( \sim 4 \text{kpc} \)) and \( \lambda \) the disk angular momentum parameter (\( \lambda = \sim 0.07 \)) together with simple calculations of the rate at which loosely bound material falls into a growing density perturbation in the early Universe. Thus theoretical prejudice leads to an expectation that Galactic halo formation was rapid, and could be complete by redshifts \( \sim 2-3 \), but disk formation should have continued until later times, \( z \lesssim 1 \).

2 Observational Preconceptions

QSO absorption line studies (e.g. Pettini, this meeting; Wolfe, this meeting; Steidel and Sargent 1992)\[10 \] indicate that a 'typical' chemical abundance has risen to \( \sim \frac{1}{10} \text{ solar} \) by a redshift \( \sim 2 \), and remains at at least that level thereafter. For redshifts \( \lesssim 2 \) Bergeron (this meeting) shows that the cross sections of the galaxies associated with absorption systems are \( \sim 50 \text{kpc} \) to redshifts at least as small as \( \sim 0.4 \). That is, QSO absorption lines show galactic halos to be moderately to considerably metal enriched at early times, but then to retain metal enriched gas for many Gyr. This second result is in disagreement with the timescale expectations noted above. We return to this point later.

3 The Star Formation History of the Milky Way Galaxy

3.1 Where did the stars form?

3.1.1 The Outer Halo: The spatial distribution of the outer galactic halo is poorly known. Available constraints come from modelling the kinematics of stars in the Solar Neighbourhood, and from in situ surveys. Each method has its strengths and weaknesses.

3.1.2 A. Deduction from local kinematics. There is a relationship between stellar kinematics at some location, large scale potentials and large scale spatial density distributions. which is described by the collisionless Boltzmann equation. Gilmore, Wyse, and Kuijken (1989) [3] have used this equation to establish the relationship between rotational velocity about the Galactic Centre (effectively angular momentum), velocity dispersion radially out from the Galactic Centre (effectively radial pressure), and the density profile of the corresponding stellar tracer population. This is shown in Figure 1. In this figure, stellar samples binned in [Fe/H] are plotted, and show some tendency for the lowest abundance data to cross
Figure 1: The relation between radial velocity dispersion $\sigma_r$ and asymmetric drift $V_{rot}$ of samples of old stars in the Galaxy.

lines of constant density profile. Although this result is confused by possible selection effects (Norris and Ryan 1991),[8] and so is of uncertain statistical significance, it is present in all current stellar samples. If real, it implies that the first stars formed during dissipational collapse of the Galactic halo.

3.1.3 B. In Situ Determinations Direct determination of the outer halo density profile is complicated by small number statistics and distance uncertainties. A recent relevant result based on the spatial distribution of BHB stars has been derived by Arnold and Gilmore (1992)[1]. For a sample of BHB stars within $\sim 15$ kpc of the Galactic Centre they derive a best fit power law index of the density law $\alpha = 3.4 \pm 0.2$, while for a more distant sample the result is $\alpha = 2.7 \pm 0.1$. [In both cases an axis ratio $q = c/a \leq 0.5$ is preferred.] These density profiles are in good agreement with the local kinematic results, in suggesting a shallow envelope in the outer Galaxy, and again marginally support the hypothesis that star formation began during dissipational collapse of the Galactic halo.

This result, albeit of low significance, means that the first star formation could have begun when the Galaxy first turned round from the general Hubble flow, and so is consistent with the first star formation occurring at very high redshifts. Even if star formation did occur during dissipation of the outer halo, this timescale deduction is of course not required. Nonetheless it does mean that in principle Galactic star formation can have occurred at the redshifts when the first QSOs are seen.
3.1.4 The Inner Halo The very central regions of the halo are very highly dissipated. The bulge seen by IRAS (see for example Harmon and Gilmore 1988) has half-light radius of only a few hundred pc, smaller than the vertical extent of the disk. The relationship of these stars to the rest of the Galaxy remains unclear. They may be a central component of the old disk, a central component of the bulge, or a discrete structure (merger remnant?). The main optical bulge in the central few kpc of the Galaxy remains very poorly studied. Its properties are mainly deduced as yet by analogy with other galaxies (Wyse and Gilmore 1988). However, it does seem to be moderately metal enriched (> solar), moderately old, and to have a significant rotational contribution to its support against gravity. It is equally well explained as being the remnant of the first disk-like structure in the Galaxy, destroyed during later accretion, or as being evidence for a kinematic-chemical abundance gradient in the inner halo.

In either case the relevant result for the present discussion is that a metal enriched stellar population formed from moderately dissipated (the optical bulge) and from extremely highly dissipated (the IR bulge) gas at relatively early times.

3.2 When did the Stars Form?

While absolute stellar age determinations remain difficult, differential measurements to determine the range of stellar ages in a narrow range of metallicities are more reliable. Such studies have been made for globular clusters (see Chaboyer et al. 1992 for a recent summary) and for field stars (Schuster and Nissen 1989). For globular clusters there is now a well established age range of ~ 3 Gyr, with the metal poor and inner clusters all being consistent with the same (old) age, and some of the outer clusters being younger. A similar result applies to local metal-poor field stars - an age range of a few Gyr is possible, though most of the subdwarfs are consistent with a narrow spread of (old) ages.

An adequate summary of the age and abundance information relevant for this discussion is that all halo star formation took place in ~ 3 Gyr some ~ 15 Gyr ago, with most in ~ 1 – 2 Gyr. The resulting abundance distribution is peaked at ~ 1/30 the solar value.

Similar studies of disk stars however show a range of ages from nearly those of halo stars to stars forming now. The abundance of the ISM at the solar Galacto-centric radius was within a factor of 2 of solar some ~ 10 Gyr ago, and has remained near solar since then.

3.3 How Rapidly Did the Stars Form?

The most direct limits on the star formation rate (SFR) in the early Galaxy come from chemical abundance considerations. Both the total abundance [Fe/H] and the individual element ratios are relevant.

3.3.1 Chemical Evolution of the Galactic halo The halo stars at and beyond the solar galactocentric distance are adequately described as having a gaussian abundance distribution with mean [Fe/H] ~ -1.5 dex, and dispersion ~ 0.5 dex. Hartwick (1976) showed that this distribution, for an IMF like that of the disk (see below) requires that ~ 90% of the gas mass must have been lost. Wyse and Gilmore (1992) considered the angular momentum distribution of the various definable stellar populations in the Galaxy to show that the "lost" halo gas is most likely the precursor of the Galactic bulge.

This has an important consequence for chemical evolution models of the local disk which use the concept of the "Solar cylinder" - all stars in a column through the Sun are considered.
A more physically well-founded sample should use a truncated cylinder, which should also curve as it rises away from the Plane, to allow for radial diffusion of stellar orbits. This is better described as a “Solar Sausage”. The implication of this is that the chemical precursors of the Galactic disk population are still in the disk – the metal-weak thick disk discussed by the Mt Stromlo group may be of relevance here.

3.3.1 Element ratios and Star Formation Rates Chemical element ratios are the most sensitive tracer of past star formation rates. This follows naturally from the different lifetimes of stars which produce different elements the ‘alpha’ elements, especially oxygen, Ca, Mg, and Ti for present purposes, are created in short-lived (≤ few × 10^8 yrs) massive stars. Thus, the amount of oxygen created is a measure of the current star formation rate of massive stars. Other elements, including much of the iron, are created in longer-lived systems (Type I supernovae), which therefore reflect a time-averaged past SFR in their production rate. The element ratios in the ISM therefore reflect the ratio of the present SFR (oxygen) to the past averaged SFR (Fe). This ratio is preserved in newly forming stars, so that we can measure it now for the halo formation phase.

A detailed description of this type of model is presented by Wheeler, Sneden and Truran (1989) for our Galaxy, and explicitly in terms of limits on the past histories of star formation rates by Gilmore and Wyse (1991). The important observation is that the element ratio [O/Fe] is constant at a value ~ 0.55 throughout the whole halo abundance range. The most recent observational data extend this range to [Fe/H] = -2.8, beyond all but a few per cent of the mass of the halo (Gilmore, Edvardsson, Gustafsson and Nissen, 1992). Remarkably, the data show no detectable cosmic scatter in [O/Fe] over the whole range -3 ≤ [Fe/H] ≤ -1, with a limit of ≤ 0.1 dex resulting. Note that this range covers effectively the entire stellar mass in the halo.

The constancy and small scatter in this element ratio provide severe constraints on both the stellar initial mass function (IMF) and the SFR during halo formation. The constant value requires each forming star to see a well-mixed mass-averaged IMF of yield (see below). Thus, the mixing must have had time to be efficient, and the lowest mass stars contributing to the yield (~ 10 M⊙) must have had time to evolve. This sets an upper limit on the SFR which is hard to quantify, but based on guestimates of mixing times, local sound speeds, the extent to which SN occur in groups, and so on, the maximum SFR is ≤ 5 times the mean SFR. The mean SFR for the halo is ~ 10 M⊙ yr^-1, to create ~ 10^{10} M⊙ in stars in ≤ 10^9 yr. Thus the proto-Galactic halo formation phase is unlikely to have been a starburst. Similar analyses for the Galactic disk, based on the systematic radial trends in [O/Fe] vs [Fe/H] and the very small scatter at any radius, similarly preclude any significant period in the disk with SFR more than a factor of a few higher than the present (radially-dependent) value (Edvardsson et al. in preparation).

3.3.3 Element Ratios and the Halo Stellar IMF The numerical value of the [O/Fe] ratio for halo stars is determined by the stellar yields as a function of mass (and [Fe/H]) and by the relative number of stars as a function of mass, the IMF. Wyse and Gilmore (1992) have quantified the limits on the IMF provided by the abundance distribution of halo stars. Changes in the IMF are more reliably derived than is the slope, due to uncertainties in the theoretical yields. The Wyse and Gilmore calculation is summarised in Table 1. The result is that a systematic change in the value of [O/Fe] ~ 0.1, a realistic maximum allowed by observations, restricts the slope of the high mass IMF to have changed by less than ~ 0.4. Note that this limit includes any systematic metallicity-dependence of the yield. It is also
interesting that the best match to the observations is with an IMF similar to that in the disk today. That is, remarkably, there is no evidence for any systematic change in the stellar initial mass function over the mass range \(100 M_\odot \gtrsim M \gtrsim M_\odot\) over 3 orders of magnitude in metallicity, \(-3 \lesssim [\text{Fe/H}] \lesssim 0\).

4 The QSO Absorption Line History of the Milky Way

From the discussion above we see that the Milky Way formed most of its halo stellar population in \(\lesssim 1 - 2\) Gyr, that these stars formed with a standard IMF and a relatively slow SFR, and that most of the halo gas was lost, probably to the central bulge, on the same timescale. This star formation began and concluded at high (\(\gtrsim 2\)) redshift. The disk very rapidly became enriched to within a factor \(\sim 2\) of solar abundance, and has maintained that abundance and an approximately constant SFR and IMF since it first formed.

By contrast, QSO absorption lines show a large and stable cross-section for gaseous halos, of order 50 kpc. The resolution of this inconsistency is provided by worms, chimneys, and fountains. Metal enriched disk gas is blown from the disk to occupy the halo. Equation 3 showed that the lifetime of this gas is short, so that continual replenishment is required, in agreement with expectation from a steady disk SFR. The small scatter in element ratios at a radius in the Galactic disk requires that this gas either returns to its origin, or falls to the outer disk and flows in radially. That is, either it keeps its angular momentum, or loses it all. For consistency, with both direct age data and with the evolution of QSO absorption lines, the disk must have been in place and enriched at redshifts \(\sim 2\), somewhat earlier than some models prefer. Additionally, galactic halos are viable explanations of QSO metallic line absorption systems only in that they are temporary repositories of disk gas. While it is in the halo, the metal-enriched disk gas resists the temptation to form stars in situ, at least at a rate sufficient to leave detectable young stars in the halo today.

References