SLAC–PUB–4670
July 1988
[AS]

Role Of High-Altitude Tracers In The Measurement Of The Galactic Disk Potential

ANDREW GOULD*

Stanford Linear Accelerator Center
Stanford University, Stanford, California 94309

ABSTRACT

The role of high-altitude ($z \gg 200$pc) tracer stars in determining the Galactic disk potential is systematically analyzed. It is shown that such tracers cannot, by themselves, be used to measure the missing mass in the disk. Up to 60% of the disk mass may be in unobserved components with only a 20% discrepancy between the measured high-altitude disk potential and what could be expected from the observed components alone. In addition, a sample of $\sim$ 1000 stars is necessary to measure the disk potential to even 20% 1-σ accuracy. However, high-altitude tracers may be used to constrain the high-velocity dispersion components of the unobserved matter. Consequently, while high-altitude tracer observations are important, they should not be given undue emphasis in tracer observation programs.

Submitted to *Astrophysical Journal*

* Work supported by the Department of Energy, contract DE–AC03–76SF00515.

1. Introduction

The determination of the density profile of the Galactic disk in the neighborhood of the Sun is a problem with a substantial history going back to Oort (1932) and even before. The basic method for making this determination is to measure the number density $\nu(z)$ of some isothermal population of tracer stars. [The population is assumed isothermal if its velocity dispersion $\overline{v^2}(z)$ is independent of height, within errors.] In principle one may simply "read off" the potential $\psi(z)$ from the density and velocity dispersion data using the one-dimensional thermodynamic relation

$$\frac{d\psi(z)}{dz} = -\overline{v^2}\ \frac{d\ln\nu(z)}{dz}, \tag{1.1}$$

or its integral form

$$\psi(z) - \psi(0) = -\overline{v^2}\ \ln\frac{\nu(z)}{\nu(0)}. \tag{1.2}$$

These equations are special solutions of the collisionless Boltzmann equation. In practice, there may not be enough data from any one star population to make a reliable estimate. It would be preferable to fit the potential to density data from several tracer populations (which are typically at different temperatures.) Alternatively, one might compare the potentials as fit to several tracer populations. An additional constraint comes from Poisson's equation

$$\frac{d^2\psi(z)}{dz^2} = 4\pi G[\rho(z) + \rho_{\rm eff}]; \qquad \rho(z) = \rho_A(z) + \rho_B(z), \tag{1.3}$$

and its boundary conditions

$$\psi'(0) = \psi(0) = 0. \tag{1.4}$$

Here ρ_A and ρ_B are the sum of the *mass* densities of all observed and unobserved components of the disk,

$$\rho_A(z) \equiv \sum_{\text{ob}} A_i \exp[-\psi(z)/\langle v^2 \rangle_i]; \qquad \rho_B(z) \equiv \sum_{\text{unob}} B_j \exp[-\psi(z)/\langle v^2 \rangle_j], \quad (1.5)$$

and ρ_{eff} accounts for effects due to the large scale structure of the Galaxy. The A_i in equation (1.5) are the densities of the observed components (stars, dust etc.) of the Galactic disk as measured at the Galactic plane. The $\langle v^2 \rangle_i$ are the corresponding dispersion velocities of these components. In a one-dimensional problem, the distribution can always be represented as a sum of Gaussian components. It is generally found that each star type (or other disk material) is well represented as a single isothermal component. However, in principle, it is possible that some disk material would have to be represented as the sum of several isothermal components. The B_j and $\langle v^2 \rangle_j$ are the densities and velocity dispersions of the unobserved components. The parameter ρ_{eff} is known from theory (see §IV),

$$\rho_{\text{eff}} \sim .08 \pm .04 \rho_A(0), \quad (1.6)$$

but only within fairly wide errors. Much of this paper will be concerned with understanding the influence of the uncertainty in ρ_{eff} on measurements of the disk potential. For now, however, the important thing about ρ_{eff} is that it is very strongly believed to be a constant over the distance scales I am considering. (More precisely, the first correction to the leading constant term is quadratic and small enough that it does not significantly alter the results of my calculations.) In §IV I will discuss exactly how ρ_{eff} depends on the large-scale structure of the

Galaxy and in what sense it is an "effective density." The first boundary condition in equation (1.4) comes from the (assumed) symmetry of the disk potential about $z = 0$. The second is a convention. The integrated column density of the observed plus unobserved components is called Σ.

Note that while the A_i and $\langle v^2 \rangle_i$ are observed quantities, the combined density of these components at finite altitudes, $\rho_A(z)$ $(z > 0)$, is *not* an observed quantity. It depends on $\psi(z)$ which in turn depends on the unobserved components. Thus, $\rho_A(z)$ cannot be calculated from a knowledge of the densities and velocity dispersions of the components observed at the plane. Moreover, the various observed components can be directly observed only up to some definite finite altitudes which depend on their intrinsic luminosities. This is why I have given the density profile of the components observed at the plane the neutral label "$\rho_A(z)$" rather than, say, "$\rho_{\rm obs}(z)$".

Bahcall (1984a, 1984b) has carried out two studies in which he solved the combined Poisson-Boltzmann equation using as sources all observed matter distributions plus various trial unobserved distributions. The resulting potentials were then compared to the potential as determined by applying equation (1.2) to tracer density data. In one study, Bahcall used F stars as tracers. In the other he used K giants. In each case, the tracer samples extended up to 200 to 600 pc, only about one to three times the scale height of the disk. He found that if one assumes no unobserved matter, then the column density, Σ, is roughly 52 $M_\odot {\rm pc}^{-2}$ for $\rho_{\rm eff} = 0$ and 48 $M_\odot {\rm pc}^{-2}$ for $\rho_{\rm eff} = 0.1\rho_A(0)$. However, the resulting potential gives a poor fit to the tracer data, one outside his "95% confidence" interval. Thus he was compelled to fit the tracer data by trying various combi-

nations of unobserved components. Typically he found that to get good fits, he had to assume that about half the disk matter was unobserved.

Recently, Kuijken and Gilmore (1987) have argued that the column density of the disk and its unobserved fraction are best determined from tracer populations which rise high above the disk.* In this region, where the disk itself has vanishing density, the potential should just be linear (plus a term reflecting the large-scale Galactic structure),

$$\psi(z) = -\Delta\psi + Kz + Fz^2 \qquad (D \gg D_\star), \tag{1.7}$$

where K and F are constants which are related to Σ and $\rho_{\rm eff}$ by

$$K = 2\pi G\Sigma, \qquad F = 2\pi G\rho_{\rm eff}, \tag{1.8}$$

and $D_\star \sim 200$ pc is the scale height of the disk. [The constant $\Delta\psi$ is necessary to satisfy the boundary condition (1.4) at the origin.] Kuijken and Gilmore measured K using a sample of 530 K dwarf stars as high up as 2000 pc, well above the $\sim$ 200 pc scale height of the disk. They found that their measured value of K (or Σ) was consistent with the value which would be inferred if one assumed no missing matter (that is 48 $M_\odot {\rm pc}^{-2}$). Thus they concluded that there was no missing mass, apparently contradicting Bahcall's result.

Kuijken and Gilmore are continuing their survey and other observers are also undertaking surveys of high altitude tracers. These surveys offer a perspective

* Private conversations have convinced me that Kuijken and Gilmore are not alone in holding this view.

on the disk potential which is as different from that of low-altitude tracer data as a photograph of Andromeda is from one of the Milky Way. I believe that as with any new perspective, one should proceed cautiously, heavily discount previously acquired intuition, and conduct a thorough reanalysis of the problem from the new viewpoint. Thus, in this paper, I propose to analyze what can be learned from high-altitude tracers and with how great an accuracy. For the bulk of the paper, I will consider only what can be learned from high-altitude data *alone*. At the end, I will briefly address the question of how high and low-altitude data may be combined, and what can be learned from the combination which cannot be adduced from each sample considered separately. However, a systematic analysis of this more complicated problem is left to a later paper.

A complete description of the missing matter would require a determination of all the B_j and $\langle v^2 \rangle_j$ in equation (1.5). However, at the present time, such a complete determination would appear to be a distant prospect. A more realistic short term goal would be to answer the following three questions:

i. Is there a significant amount of unobserved matter?

ii. If so, what is its total density?

iii. What is its velocity dispersion?

Although the above three questions may seem unambiguous, they actually conceal radically conflicting viewpoints on the notion of missing matter. To different people, the question "Is there missing matter?" may mean one of at least three things:

1. Is there a significant density of matter in the neighborhood of the Sun which

we do not observe?

2. Is the column density of the disk significantly larger than would be inferred from the disk components observed in the neighborhood of the Sun?

3. Of all the mass in the disk, is a significant fraction due to disk components which are not observed in the neighborhood of the Sun?

These questions may appear to be just clever rewordings of one another, but each refers to a completely different concept. More importantly, measurements which help answer (or definitively answer) one of these questions may have little or nothing to say about the others. Thus, it is crucial to give a mathematical definition to each of these ideas of "missing matter" and to determine how they are related.

I define the following three dimensionless quantities corresponding respectively to the above three definitions,

$$\Delta_\rho \equiv \frac{\rho_B(0)}{\rho(0)} = 1 - \frac{\rho_A(0)}{\rho(0)}, \tag{1.9}$$

$$\Delta_\Sigma \equiv 1 - \frac{\Sigma_{\rm min}}{\Sigma}, \tag{1.10}$$

$$\Delta_T \equiv \frac{\Sigma_B}{\Sigma} = 1 - \frac{\Sigma_A}{\Sigma}. \tag{1.11}$$

where

$$\Sigma_A \equiv \int_{-\infty}^{\infty} dz \rho_A(z), \qquad \Sigma_B \equiv \int_{-\infty}^{\infty} dz \rho_B(z), \qquad \Sigma \equiv \Sigma_A + \Sigma_B, \tag{1.12}$$

and $\Sigma_{\rm min}$ is the minimum possible column density of the disk consistent with the components observed at the plane.

I consider these quantities in turn. The first, Δ_ρ, is a measure of the missing matter which is at the plane. Clearly if there is a lot of unobserved, low velocity-dispersion matter, Δ_ρ will be large. On the other hand there could be substantial quantities of unobserved matter, but if its velocity dispersion were high, very little of this matter would be at the plane and hence Δ_ρ would be low. Thus, while Δ_ρ provides useful information about the disk potential, it cannot, by itself, resolve the question of missing matter. The determination of Δ_ρ is, in principle, straightforward. On the one hand, $\rho_A(0)$ is directly observed. On the other, $\rho(0)$ can be found by fitting the low-altitude potential to a parabola

$$4\pi G\rho(0) = -2F + \left.\frac{d^2\psi(z)}{dz^2}\right|_{z\ll D_\star}, \tag{1.13}$$

where $D_\star \sim 200$pc is the scale height of the disk. One possible problem with this procedure is that F is not known perfectly. However, from equation (1.6), it is clear that this uncertainty plays a minor role. More important is the difficulty in finding tracers with a large enough density and small enough velocity dispersion to be sensitive to the potential near the plane. In any event, Δ_ρ will not figure prominently in this paper.

In a way, Δ_Σ is the mirror opposite of Δ_ρ: it is very sensitive to high velocity-dispersion dark matter, but very insensitive to dark matter with low velocity dispersion. This relationship, which is not at all obvious, will be analyzed in detail in §II. Thus, like Δ_ρ, it provides useful information about the potential, but cannot, by itself, resolve the question of missing mass. There is another sense in which Δ_Σ is the mirror opposite of Δ_ρ: it will play an extremely prominent role in this paper. This is because high-altitude tracers *naturally* measure

Δ_Σ. High-altitude tracers measure a single number, K, or equivalently, Σ, the column density. (They also yield information about F, but that is irrelevant to the present discussion.) This number must then be compared with some other *single number* which characterizes the known matter distribution. The only such number is $\Sigma_{\rm min}$, the minimum column density consistent with the observed components. The comparison of these two numbers [eq. (1.10)] is Δ_Σ.

I cannot emphasize too strongly that this situation is very different from that produced by medium and low-altitude tracers. Using a general set of tracers, one could in principle map the potential with arbitrary accuracy. Of course there will be a whole range of potentials which are consistent with any given finite data set. Some of these potentials can be ruled out on physical grounds. For example, the second derivative of the potential is the mass density, and this cannot be negative. But, by using the Poisson constraint as Bahcall (1984a, 1984b) did, one can rule out many more potentials which would otherwise be consistent with the data. In particular, none of the components described by the $\langle v^2 \rangle_i$ in equation (1.5) can have coefficients less than A_i. Thus the Poisson constraint contains an enormous amount of information. However, But the higher one gets in the potential, the more components die out, and the less significant this constraint becomes. Finally, at very high altitudes, the Poisson constraint provides only one piece of information, namely that the linear term in the potential must be at least as great as $2\pi G\Sigma_{\rm min}$. Thus, $\Sigma_{\rm min}$ may be regarded as the high-altitude rump of the Poisson constraint.

This residual constraint is very weak. As I show in §II, there are potentials which just barely satisfy it and yet have a tremendous amount of missing matter.

The procedure for determining $\Sigma_{\rm min}$ is as follows. If one could be sure that there were no unobserved components (and that one knew $\rho_{\rm eff}$), one could numerically determine the total column density from equations (1.3), (1.4), and (1.5). The quantity so calculated is $\Sigma_{\rm min}$. It is easy to convince oneself (again assuming that $\rho_{\rm eff}$ is known) that

$$\Sigma \geq \Sigma_{\rm min} \geq \Sigma_A, \tag{1.14}$$

where equality holds if and only if there is no missing mass. Thus, $\Sigma_{\rm min}$ is indeed the minimum possible column density consistent with the observed components.

The third quantity, Δ_T, corresponds to my own naive view of what is meant by "missing matter" in so far as it can be described by a single parameter: It is the fraction of the disk mass which is in unobserved components. In addition, this parameter is sensitive to both high and low-velocity dispersion dark matter. For high-velocity dark matter, it just as sensitive Δ_Σ. For low-velocity dark matter, it is not quite as sensitive as Δ_ρ, but is much more sensitive than Δ_Σ. Therefore, in this paper I will use Δ_T to quantify the fraction of the disk matter which is dark.

However, Δ_T does have an important drawback; it is not easily measured. It is defined in terms of Σ_A, which is *not* an observed quantity. Recall that Σ_A is the integral of $\rho_A(z)$ which depends on the potential and thus on the unobserved components. The only way to determine Σ_A is to measure the whole potential, using tracers and the Poisson constraint. This determination is a difficult problem (Bahcall 1984a, 1984b). Fortunately, the difficulties in measuring Δ_T do not translate into difficulties in carrying out the analysis of this paper.

In §II, I begin by analyzing what one learns about the missing mass (Δ_T) by measuring K (and hence Δ_Σ). I show that if Δ_Σ is large then Δ_T must also be large, but that the reverse is not true. If the missing mass has a low velocity dispersion, then well over half the mass may be in unobserved components even though there is not much discrepancy between Σ and $\Sigma_{\rm min}$. In particular I find that Δ_T may be as high as 60% when Δ_Σ is only 20%. Thus, one may use measurements of Δ_Σ to demonstrate the existence of missing matter, but not to rule it out.

In §III, I calculate how well one may measure K, if nothing is assumed about F. (In other words, how well may K and F may be measured simultaneously.) I find that the (1-σ) statistical error in the measurement of K in these circumstances is

$$\frac{\delta K}{K} \sim \left(\frac{75}{N}\right)^{\frac{1}{2}}, \tag{1.15}$$

where N is the number of stars observed above the disk. With 350 such stars, the error is $\sim 45\%$. With 2000 stars it is $\sim 20\%$. If this result is combined with the results of §II, it is clear that high-altitude tracers cannot (by themselves) put any serious constraint on the missing matter, Δ_T, unless astronomical amounts of data are taken.

In §IV and §V, I examine the possibility that the problems encountered in §III can be avoided by making a theoretical estimate of F. In §IV I calculate F theoretically and give an estimate of the error in this calculation. In §V, I show how this theoretical determination may be combined with observational data to yield a single estimate of K. I find that the uncertainty in this combined

calculation is substantially reduced relative to equation (1.15), but is still large,

$$\frac{\delta K}{K} \sim 27\% \qquad (N = 350). \tag{1.16}$$

At $N = 1000$, the uncertainty is still $\sim 20\%$.

Thus, the viewpoint that the missing matter is best measured by observing a tracer population well above the disk, does not stand up to examination. Despite the fact that the potential has a simple form in this region, only a very crude determination of the missing mass is possible. In avoiding the region where the potential has a complicated form, one also "avoids" the region where one can apply the powerful Poisson constraint. As I will show in later paper, this constraint plays a key role in determining the missing matter.

The conclusions of §III and §V must be modified by the following more refined analysis. Recall that it is not Σ itself that gives information about the missing matter, but Δ_Σ. This latter quantity is related to Σ by a "known quantity", $\Sigma_{\rm min}$. However, the procedure for inferring $\Sigma_{\rm min}$ from the observed components involves assumptions about F. This means that if an error is made in measuring (or estimating) F in §III (or §IV), then there will also be an error in the minimum column density. This will in turn induce an error in Δ_Σ. In §VI, I calculate this error and show that its sign is such as to *reduce* the effects discussed in §III and §V, but that its magnitude is relatively small. Another way of saying the same thing is that an error in F will induce an error in K which may be divided into two pieces. The first (smaller) piece has no effect on Δ_Σ, while the second has a

direct effect. I call the first piece "the irrelevant error in K", $(\delta K)_{\rm irr}$, and find

$$\frac{(\delta K)_{\rm irr}}{\delta K} \sim 25\%, \tag{1.17}$$

One rather disappointing result of §V is that even very large numbers of star counts do not significantly improve the accuracy of the measurement of F (or $\rho_{\rm eff}$). Recall that this quantity is known from theory only to within 50%. At $N = 2000$, this uncertainty drops somewhat, but only to $\sim 40\%$. Since F is an important parameter of the large-scale structure of the Galaxy, one might hope that these high-altitude measurements would shed some light on it. In §VII, I investigate a back-door route to this goal. Measurements of K made in and near the disk depend only very weakly on F. Therefore, if these measurements can beat down the error in K to low levels, then the high-altitude measurements can give important information about F. In §VII, I estimate how well the present data ($N = 350$) can constrain F if K is known to some specified accuracy. I find that substantial improvements can be made in the accuracy of F if the error in K (as determined from low-altitude measurements) can be reduced $\lesssim 15\%$. Further improvements are possible if more high-altitude stars are observed.

The central focus of this paper is the use of high-altitude tracer data in the measurement of the disk mass, *in general.* While the paper was stimulated by the work of Kuijken and Gilmore, its application is not restricted to any particular data set. Thus, in the bulk of the paper I treat the Kuijken and Gilmore data set as an example and do not attempt to analyze it in detail. However, Kuijken and Gilmore have used their data to make the strong claim that there is no evidence for missing matter in the disk. I believe that a number

of subtle considerations conspire to undercut this claim. In §VIII, I show that the Kuijken and Gilmore data tend to indicate that there is a significant quantity of missing matter, although the uncertainties are far too large to make a firm claim.

I conclude that tracer observation programs should not be designed to maximize the number of high-altitude tracers. Star for star, these tracers require more telescope time and are more prone to systematic errors. The hope that these drawbacks of high-altitude tracers might be compensated by their assumed exceptional value in determining the missing mass is illusory.

2. Δ_T vs Δ_Σ: Missing Mass As Measured By K

In this section, I investigate the extent to which missing mass can be calculated by measuring K, the linear term in the high altitude potential. To simplify the discussion, I will assume throughout this section that $F = 0$. I will deal with problems posed by a finite F in subsequent sections. With this assumption $\Sigma_{\rm min}$ is known so that [by eq. (1.10)], a measurement of K directly implies an estimate of Δ_Σ. Thus, the problem is reduced to analyzing how much missing mass (Δ_T) is implied by a given measurement of Δ_Σ. In general, one expects that this will depend on the velocity dispersion of the missing mass. It may also depend on the degree to which the observed distribution deviates from being isothermal. Of course it may also depend in a subtle way on the precise composition of both the observed and unobserved components. However, in this section I am not interested in such subtleties. I wish only to find how the relationship between Δ_T and Δ_Σ varies with the gross characteristics of the system. I will therefore consider a

set of disk models each of whose components (whether observed or unobserved) has one of the three velocity dispersions, v_l^2, v_m^2, or v_h^2, which satisfy the following extreme relations

$$v_l^2 \ll v_m^2 \ll v_h^2. \tag{2.1}$$

The total column density (observed plus unobserved) of these components will be designated Σ_l, Σ_m, and Σ_h. No assumption will be made about the relative size of the column densities. However, I will assume that the density at the plane of the medium-velocity component is much greater than that of the high-velocity component,

$$\rho_m(0) \gg \rho_h(0). \tag{2.2}$$

Condition (2.2) will almost always be satisfied in the limit (2.1), so this is not a significant additional assumption. (Note that since the system is collisionless, the different components can coexist at different temperatures. These components are assumed to interact with one another only gravitationally.)

Remarkably, the Poisson-Boltzmann equation may be solved exactly in the limits (2.1) and (2.2). The solution, given below, is derived in Appendix A.

$$\frac{d\psi}{dz} = K\tanh\Big(\frac{z}{D} + \tanh^{-1}\frac{K_{lm}}{K}\Big) - K_{lm}\left[1 - \tanh\Big(\frac{z}{D_{lm}} + \tanh^{-1}\frac{K_l}{K_{lm}}\Big)\right], \tag{2.3}$$

$$4\pi G\rho(z) = \frac{K}{D}\mathrm{sech}^2\Big(\frac{z}{D} + \tanh^{-1}\frac{K_{lm}}{K}\Big) + \frac{K_{lm}}{D_{lm}}\mathrm{sech}^2\Big(\frac{z}{D_{lm}} + \tanh^{-1}\frac{K_l}{K_{lm}}\Big) + K_l\delta(0), \tag{2.4}$$

$$4\pi G\rho_m(0) = \frac{K^2 - K_{lm}^2}{2v_h^2}; \qquad 4\pi G\rho_h(0) = \frac{K_{lm}^2 - K_l^2}{2v_m^2}; \tag{2.5}$$

where

$$K_l \equiv 2\pi G\Sigma_l, \quad K_m \equiv 2\pi G\Sigma_m, \quad K_h \equiv 2\pi G\Sigma_h,$$
$$K_{lm} \equiv K_l + K_m; \qquad K \equiv K_l + K_m + K_h, \tag{2.6}$$

$$D \equiv \frac{2v_h^2}{K}; \qquad D_{lm} \equiv \frac{2v_m^2}{K_{lm}}. \tag{2.7}$$

Note that the solution is exact only in the limit (2.1). It might also be regarded as the first term in a systematic expansion of the solution around $v_l^2 = 0$, $v_h^2 = \infty$. However, from the standpoint of this paper, the leading term is all that is important.

Using equation (2.5), the relationship between Δ_T and Δ_Σ may be evaluated for a variety of models. The procedure is as follows. Choose a model described by observed components K_l, $\rho_m(0)$, and $\rho_h(0)$. Equation (2.5) gives the inferred column densities K_m and K_h (assuming no missing mass) hence the total (minimum) column density, $K_{\rm min} = K_l + K_m + K_h$. [In accord with eq. (2.6), I use K's and Σ's interchangeably.] Now I assume that there is enough unobserved matter to produce a Δ_Σ of a given magnitude (say 20%). According to the definition of Δ_Σ [eq. (1.10)], this implies that the total (observed plus unobserved) column density is

$$K = (1 - \Delta_\Sigma)^{-1} K_{\rm min}. \tag{2.8}$$

I now suppose that all the unobserved mass is in a component with one of the 3 velocity dispersions, say for example, v_m^2. Again using equation (2.5), I solve for the new *total* density of the medium component, $\rho_m^{\rm tot}(0)$. This solution also gives the new *total* column density for the medium component, $K_m^{\rm tot}$. From these and

the definition of Δ_T [eq. (1.11)], I calculate the missing mass fraction,

$$\Delta_T = \frac{K_m^{\rm tot}}{K}\left(1 - \frac{\rho_m(0)}{\rho_m^{\rm tot}(0)}\right). \tag{2.9}$$

I consider models whose observed components have a total velocity dispersion v_m^2 [i.e., $K_l v_m^2 = K_h(v_h^2 - v_m^2) \sim K_h v_h^2$], and work in the limit (2.1). Then for unobserved components with high, medium, and low velocity dispersions, I find

$$\begin{aligned}
&\Delta_T = \Delta_\Sigma \qquad \text{(high)}, \\
&\Delta_T = \Delta_\Sigma(2 - \Delta_\Sigma)[1 + \tau(1 - \Delta_\Sigma)]^{-} \qquad \text{(medium)}, \\
&\Delta_T = [1 - (1 - \tau)(1 - \Delta_\Sigma)^2]^{\frac{1}{2}} - \tau(1 - \Delta_\Sigma) \qquad \text{(low)},
\end{aligned} \tag{2.10}$$

where

$$\tau \equiv K_l/K \tag{2.11}$$

parameterizes the deviation of the observed distribution from a isothermal one. Before analyzing the general (non-isothermal) case, consider the limit where the observed distribution is a pure medium velocity component. For definiteness, I will evaluate Δ_T when $\Delta_\Sigma = 20\%$. In this isothermal limit ($\tau = 0$), equation (2.10) reduces to

$$\begin{aligned}
&\Delta_T = \Delta_\Sigma \to 20\% \qquad \text{(high)} \\
&\Delta_T = \Delta_\Sigma(2 - \Delta_\Sigma) \to 36\% \qquad \text{(medium)} \\
&\Delta_T = [\Delta_\Sigma(2 - \Delta_\Sigma)]^{\frac{1}{2}} \to 60\% \qquad \text{(low)}.
\end{aligned} \tag{2.12}$$

Thus, for high-velocity unobserved matter, Δ_Σ reflects Δ_T perfectly. However, if the unobserved matter is low-velocity, then it may be present in huge (majority)

quantities, while changing the column density only marginally. The reason for this is that if a large quantity of unobserved low-velocity matter is present, it raises the potential everywhere, thereby greatly reducing the column densities of the observed components relative to what was inferred on the basis of the observed distributions alone. From this explanation, it is clear that if the observed distribution is highly non-isothermal (and hence has a large low-velocity component) an additional low-velocity unobserved component will give rise to *greater* increase in the column density compared to the isothermal case. This is confirmed by Figure 1 in which I have plotted Δ_T against τ for fixed Δ_Σ and for each of the three types of unobserved matter. In order to assess what value of τ is most appropriate for our own disk, I give a more general definition which reduces to equation (2.11) in the special case of above models:

$$\tau \equiv \frac{\sum_i K_i \mid \langle v^2 \rangle_i - \overline{v^2} \mid}{2\overline{v^2} \sum_i K_i}; \qquad \overline{v^2} \equiv \frac{\sum_i K_i \langle v^2 \rangle_i}{\sum_i K_i}, \tag{2.13}$$

where the K_i are the inferred column densities for each of the observed components. Using this definition and Bahcall's (1984a) values for the A_i and $\langle v^2 \rangle_i$, I find that for the observed components of the Milky Way disk, $\tau \sim 0.25$.

3. Determination Of K And F From Tracer Data

In this section, I imagine that velocity dispersion and density data of a tracer distribution are measured in a region $L_1 < z < L_2$ with the aim of determining K. I will assume $L_1 \gg D_\star$ so that virtually the entire mass of the disk is below the observations and the derivative of the potential is very well approximated by

$$\frac{d\psi}{dz} = K + 2Fz. \tag{3.1}$$

When I make a numerical calculation, I will take $L_1 = 400$pc and $L_2 = 2000$pc. The upper limit is in conformity with the upper limit of the Kuijken and Gilmore data. The choice of the lower limit requires some justification. Using Bahcall's data (1984a) one finds (assuming no missing matter) that $\sim 75\%$ of the disk mass is below 400 pc (and above -400 pc). Thus the form (3.1) is not yet really valid at 400 pc. In fact, to get above 97% of the disk mass, one must go to $\sim$ 1000 pc. This latter number would then appear to be a more realistic lower limit. However, while the potential is not the *simple* function (3.1) of two parameters at 400 pc, it can to a very good approximation be written as a more complicated function of two parameters. This function would take account of the (nearly) exponentially decaying mass density of the high-velocity components which are measured at the Galactic plane. On the other hand, at altitudes well below 400 pc, the potential cannot be written as function of two parameters. In principle, then, one should set the lower limit at 400 pc, but use the more complicated potential. However, since I am interested in estimating the *uncertainties* in K and not K itself, such complexities would obscure the analysis without improving

the results. The idealization (3.1) with the given limits is therefore an appropriate one for the task at hand.

In general, one must assume that the tracer population is not isothermal. For example, the velocity dispersion of the Kuijken and Gilmore K dwarf sample rose from ~ 20 km s^{-1} near the plane to ~ 40 km s^{-1} at 2000 pc. This contrasts sharply with the situation faced by Bahcall. Both his F dwarf and K giant samples were found to be isothermal within errors. One possible reason for the difference is that K dwarfs are simply less isothermal. However, the non-thermal elements of any population will become more and more pronounced at higher elevations for the simple reason that the low-velocity components become suppressed. It may well be that if the F dwarfs were followed up to 2000 pc, their non-isothermal components would be more apparent. In any event, if one wishes to use high-altitude tracers, one must be prepared to analyze non-isothermal distributions. This requires that equation (1.1) be generalized. The appropriate generalization is given as equation (4-36) in Binney and Tremaine (1987),

$$\frac{d\psi(z)}{dz} = -\overline{v^2}(z)\,\frac{d\ln\nu(z)}{dz} - \frac{d\overline{v^2}(z)}{dz}, \tag{3.2}$$

which may be integrated to yield

$$\nu(z)\overline{v^2}(z) = \int_{z'=z}^{\infty} dz' \frac{d\psi(z')}{dz'}\nu(z'). \tag{3.3}$$

To find the potential from density and velocity-dispersion data one has essentially two choices. One may fit the density data to some parameterized curve and fit the velocity-dispersion data to another, and then read off the potential from equation

(3.2). Alternatively, one may fit the density data to some parameterized curve and then use equation (3.3) to fit the potential to another parameterized curve. In this latter fit, the velocity-dispersion data are the points which are being fit. In general, I believe that the first procedure is probably preferable because it does not involve any assumption about the density in the region beyond the data. However, in my model calculation I will use the second procedure. One reason for doing this is that in this asymptotic regime, the derivative of the potential is known to have a very simple form [given by eq. (3.1)] with only two parameters. If the first method were used, a *third fit* would have to be done to fit the potential to this form after the velocity dispersion and density were fit. This would make the entire calculation less clear. A second reason is that, as a practical matter, the unknown density in the region beyond $L_2 = 2000$ pc plays very little role.

In this section and §V I will simulate the analysis of data using the procedure described above in order to estimate the uncertainties in the calculations. While the procedure is valid for either isothermal or non-isothermal tracers, I will assume that the underlying distribution is isothermal. My reason for this is that it makes the simulation analytically tractable. Since I am interested only in calculating the *uncertainties* in the results and not the results themselves, this assumption should not cause any serious problems. (Notice, however that I have been careful to use a procedure of data analysis which is valid for a non-isothermal population. If my simulation relied on a procedure which was valid only for isothermal tracers, this would cause serious problems.) Whenever making a numerical calculation, I will use

$$\overline{v^2} = (25 \text{ km s}^{-1})^2 \tag{3.4}$$

This is roughly the velocity dispersion of the Kui ken and Gilmore sample of K dwarfs at a scale height.

The net result of this simulation will be the variances and covariance of K and F, expressed as a covariance matrix. Since K and F are dimensionful, the elements of this matrix will also be dimensionful, a.id moreover will have different dimensions from one another. In order to make the elements of the covariance matrix dimensionless, I introduce the dimensionless basis vector,

$$(K/K_\star, F/F_\star) \qquad \text{(Standard Basis Vector)}, \tag{3.5}$$

where

$$F_\star \equiv 2.0 \times 10^{-4} (\text{km s}^{-1})^2 \text{pc}^{-2}; \qquad K_\star \equiv 1.3 (\text{km s}^{-1})^2 \text{pc}^{-1}. \tag{3.6}$$

The choice of the normalization factors $K_\star$ and $F_\star$ is somewhat arbitrary. I have taken $K_\star$ as the minimum value of K (K_{min}) given the observed matter in the solar neighborhood. I have chosen $F_\star$ to be the expected value of F if $K = K_\star$. This expected value is calculated in the next section. With this choice of basis, the diagonal elements of the covariance matrix will be the (square of) the fractional uncertainties in K and F. Note that the scale associated with this normalization,

$$\frac{K_\star}{F_\star} \sim 6.5 \text{ kpc}, \tag{3.7}$$

is of the same order as the Galactocentric distan e, $R \sim 8.5$ kpc, and the radial scale of the disk, $h \sim 3.5$ kpc. As I will make clear in the next section, this is

related to the fact that the disk and halo contributions to the Galactic rotation curve are of the same order at the position of the Sun.

Consider then, a series of n measurements of ν and $\overline{v^2}$ taken at positions z_i, $D_\star \ll L_1 < z_i < L_2$. First, I will mimic the determination of $\nu(z)$ from the number density data. Since, the distribution is isothermal, the density will be given by

$$\nu(z) = \exp(-A_0 - A_1 z - A_2 z^2), \tag{3.8}$$

where A_0 is an overall normalization constant, $A_1 = K/\overline{v^2}$, and $A_2 = F/\overline{v^2}$. I will thus try to fit the logarithm of the density to a quadratic, $a_0 + a_1 z + a_2 z^2$. The variance of a_1 will then translate directly into one source of error in determining K. Proceeding with a standard linear fitting analysis (see for example Press *et al.* 1986) one finds

$$\chi^2(a_0, a_1, a_2) = \sum_{i=1}^{n} \frac{1}{\sigma_i^2}\{\ln[\nu_{\rm ob}(z_i)] - (a_0 + a_1 z_i + a_2 z_i^2)\}^2, \tag{3.9}$$

where $\nu_{\rm ob}$ is the observed density derived from star counts. I will consider only statistical (and not systematic) errors in the measurement of $\nu_{\rm ob}(z_i)$. Then the variance is just the inverse of the star count in the z_i bin,

$$\sigma_i^2 = [\nu_{\rm ob}(z_i)\alpha z_i^2 (L_2 - L_1)/n]^{-1}, \tag{3.10}$$

where α is the spherical angle of the observation cone. (For simplicity of exposition I take the discrete volume element to be $\alpha z_i^2 \Delta z_i$ throughout this paper, with α taken to be independent of z. I also assume that the bins are large enough to

ignore the difference between "n" and "$n-1$" in the usual variance formulae.) Thus, equation (3.9) becomes

$$\chi^2(a_0,a_1,a_2) = \int_{z=L_1}^{L_2} \alpha z^2 dz\ \nu_{\rm ob}(z)\{\ln[\nu_{\rm ob}(z)] - (a_0 + a_1 z + a_2 z^2)\}^2. \quad (3.11)$$

Differentiating with respect to a_j yields the matrix equation

$$b^{\nu}_{jk} a_k = d_j \quad (3.12)$$

where

$$b^{\nu}_{jk} = \int_{L_1}^{L_2} \alpha z^2 dz\ \nu_{\rm ob}(z) z^{j+k} \qquad d_j = \int_{L_1}^{L_2} \alpha z^2 dz\ \nu_{\rm ob}(z) z^j [\ln(\nu_{\rm ob}(z)]. \quad (3.13)$$

The covariance matrix for the a_k is given by the inverse of b^{ν}_{jk}. (See Appendix B.) I will assume, as is almost certainly the case, that

$$\frac{F\overline{v^2}}{K^2} \ll 1. \quad (3.14)$$

Then b^{ν}_{ij} may be evaluated

$$b^{\nu}_{ij} = N A_1^{-(i+j)} \frac{f(i+j+2)}{f(2)}, \quad (3.15)$$

where

$$f(k) \equiv k! \sum_{l=0}^{k} \left[\frac{(KL_1/\overline{v^2})^l}{l!} \exp(-KL_1/\overline{v^2}) - \frac{(KL_2/\overline{v^2})^l}{l!} \exp(-KL_2/\overline{v^2}) \right], \quad (3.16)$$

and N is the number of stars in the sample. The covariances of a_1 and a_2 contribute directly to the covariances of K and F. This is clear because, if the

velocity dispersion $\overline{v^2}(z)$ were known precisely at every point, then K and F could be read off directly from a_1 and a_2. Thus, by evaluating and inverting b^ν_{ij}, and then converting to the standard basis [eq. (3.5)], one obtains the contribution to the covariance from the measurement of $\nu(z)$. I have carried out this calculation using the above standard values of $\overline{v^2}$ and the L_i, and taking $K = K_\star$. I find that the (dimensionless) covariance matrix is given by

$$c^\nu_{ij} = \frac{1}{N}\begin{pmatrix} 0.3 & -3.5 & 8.9 \\ -3.5 & 42.3 & -112.4 \\ 8.9 & -112.4 & 307.7 \end{pmatrix} \qquad (i,j = 0,1,2). \tag{3.17}$$

The zeroth row and column in the above equation are the covariances of the overall constant term in the potential. This term is physically unobservable and hence irrelevant. Its covariances are of interest only because they illustrate the degree to which the determination of physical observables depends on the irrelevant normalization of the tracer density. The relevant terms in equation (3.17) are

$$c^\nu_{ij} = \frac{1}{N}\begin{pmatrix} 42.3 & -112.4 \\ -112.4 & 307.7 \end{pmatrix} \qquad (i,j = 1,2). \tag{3.18}$$

Thus, for $N = 350$ (roughly the number of stars above 400 pc in the Kuijken and Gilmore sample), the uncertainty in K from the density measurement alone is $\sim (42/N)^{1/2} \sim 35\%$.

The uncertainties induced by the fitting of $\overline{v^2}$ are also significant. These are independent of the errors induced by the fitting of ν, so I am justified in assuming that $\nu(z)$ is completely known for purposes of this specific calculation. I adopt a trial potential $\psi(z) = a_0 z + a_1 z^2$, and make a prediction for $\overline{v^2}(z)$ based on

equation (3.3). Thus

$$\chi^2(a_0,a_1)=\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\left\{\overline{v_{\rm ob}^2}-[\nu(z_i)]^{-1}\int_z^\infty dz'(a_0+2a_1z')\nu(z')\right\}^2. \tag{3.19}$$

In this case,

$$\sigma_i^2=2(\overline{v^2})^2[\nu(z_i)\alpha z_i^2(L_2-L_1)/n]^{-1}, \tag{3.20}$$

so that the inverse of the covariance matrix is given by

$$b_{ij}^v=\frac{1}{2(\overline{v^2})^2}\int_{L_1}^{L_2}\frac{\alpha z^2dz}{\nu(z)}\left[\int_z^\infty dz'(2z')^i\nu(z')\right]\left[\int_z^\infty dz'(2z')^j\nu(z')\right]. \tag{3.21}$$

As in the previous evaluation, I assume the condition (3.14). Then b_{ij}^v may be evaluated

$$\begin{aligned}
b_{00}^v=&\frac{N}{2(\overline{v^2})^2},\quad b_{01}^v=b_{10}^v=\frac{N}{K\overline{v^2}f}(8+8x+4x^2+x^3)e^{-x}\Big|_{KL_2/\overline{v^2}}^{KL_1/\overline{v^2}},\\
b_{11}^v=&\frac{2N}{K^2f}(38+38x+19x^2+6x^3+x^4)e^{-x}\Big|_{KL_2/\overline{v^2}}^{KL_1/\overline{v^2}},\\
f\equiv&(2+2x+x^2)e^{-x}\Big|_{KL_2/\overline{v^2}}^{KL_1/\overline{v^2}}.
\end{aligned} \tag{3.22}$$

The dispersion fit contribution to the covariances of K and F is given by the inverse of this matrix. I have evaluated this matrix in my standard dimensionless basis and using my standard values of K, L_i, and $\overline{v^2}$ and find,

$$c_{ij}^v=\frac{1}{N}\begin{pmatrix}31.4 & -58.4\\ -58.4 & 116.0\end{pmatrix}. \tag{3.23}$$

According to the "resistor rule" for covariances (see Appendix B), covariance matrices from measurements carried out "in series" add directly, while covariance

matrices from measurements carried out "in parallel" add inversely. (A set of measurements of a vector is said to be carried out "in parallel" if each may be used separately to determine the vector. Conversely, it is said to be carried out "in series" if the entire set is necessary to make the determination.) Using this rule, I find that the (dimensionless) covariance matrix for the entire measurement process is

$$c_{ij}^{\nu v} = \frac{1}{N} \begin{pmatrix} 74 & -171 \\ -171 & 424 \end{pmatrix}. \tag{3.24}$$

Thus, the 1-σ uncertainty in the determination of K is $\sim (74/N)^{1/2} \sim 45\%$ for the case $N = 350$. This is in addition to any systematic errors which arise from the measurement procedure.

Equation (3.24) has two salient features. First, the diagonal elements (variances) are very large compared to the values one would naively expect ($\sim 1/N$). Second, the dimensionless variance of F is much larger than that of K. Since the large variance of K is an essential limiting factor in our ability to use high-altitude tracers to measure the disk potential, it is important to understand the physical origin of these features.

The large variances are a direct result of the large covariance (off diagonal elements). That is, K and F have an unexpectedly large correlation coefficient,

$$\sim \frac{-171}{(74 \cdot 424)^{1/2}} \sim -0.97. \tag{3.25}$$

Physically, this arises from the fact that it is difficult to distinguish between the effects of the quadratic and linear terms in the potential. This problem is

exacerbated by the fact that at the lower limit, $L_1 = 400$ pc, the effective density term in the potential is already fairly large. Recall that it was necessary to set the lower limit this high in order to insure that the potential could be written as a known function of two variables. *Thus, this large covariance is connected in a fundamental way to the attempt to avoid the details of the potential by getting above the disk.* To judge the impact of having to set L_1 so high, I evaluate the covariance matrix for the (hypothetical) case $L_1 = 0$, $L_2 = \infty$. For this case, all equations may be evaluated by hand and one quickly finds

$$c_{ij}^{\nu v} = \frac{1}{N}\begin{pmatrix} 16 & -23 \\ -23 & 38 \end{pmatrix}; \qquad [L_1 = 0, \quad L_2 = \infty]. \tag{3.26}$$

Thus a factor of ~ 5 in the variance of K is due to the fact that no observations from the region where the effective density term is negligibly small can be used.

The large value of the variance of F relative to that of K is, in a sense, an artifact of the normalization, (3.6). The scale of the normalization [eq. (3.7)] is 6.5 kpc, while the intrinsic scale of the measurement is $\overline{v^2}/K \sim 480$ pc. If I had normalized K and F according to this latter scale, the variance of F would have been ~ 200 times lower. This is why I say that the large variance in F is an artifact of the normalization. Nevertheless, this difference in scales is a real effect, and it is one which can be used to advantage as I will show in the next section.

4. Theoretical Relation Between K and F

The fractional variance in F ($\sim 424/N$) found in the last section is so large that a typical measurement may easily yield an absurd value of F. For example, if $N = 350$, then there is a high probability that F will be measured to be negative. While F is not precisely known, it may be taken as positive with considerable confidence. One should really include a theoretical constraint on F to prevent it from wandering so far afield. Since the covariance of K and F is large, this constraint on F will produce an indirect constraint on K. This section is therefore devoted to a theoretical calculation of F. It turns out that what is constrained theoretically is not F alone, but a linear combination of F and K. In the next section I will show how this constraint may be incorporated into the evaluation of K, and estimate the degree to which the uncertainty is thereby reduced.

As I will show, the theoretical uncertainty in F is very large ($\sim 50\%$) when measured in units of $F_\star$. Since these are the natural units for measuring F, one might at first be led to believe that not much would come of such a "constraint." However, recall the point made above that the variance of F in equation (3.24) was small in units natural to the measurement but large in units of $F_\star$. Thus, the ratio of distance scales acts as sort of a lever which allows one to drastically reduce the uncertainty in one arena by making a fairly crude calculation in another. Unfortunately, as I will demonstrate in the next section, this reduction is not as great as one would like.

What physical effects, then, contribute to F? First, there are effects which relate to large scale structure of the Galactic potential. Second, there are effects

which arise from changes in the velocity ellipsoid as one moves away from the disk. I consider these in turn.

The potential above the Galactic plane depends not only on the vertical mass distribution of the disk but also on the large scale structure of the Galaxy. Stars which are at some height z directly above the Sun are a distance $R + z^2/2R$ from the center of the Galaxy. (Here R is the solar Galactocentric distance.) Thus there is a quadratic term in the potential due to the Galactic halo and bulge, $\psi_{\rm hb}$,

$$\psi_{\rm hb}(z) = \phi_{\rm hb}(R + z^2/2R) - \phi_{\rm hb}(R) \sim z^2 \frac{\phi_{\rm hb}'(R)}{2R}, \tag{4.1}$$

where $\phi_{\rm hb}(r)$ is the (spherical) potential due to the halo and the bulge. Comparison of the above equation with the definition of F shows,

$$F = \frac{\phi_{\rm hb}'(R)}{2R}. \tag{4.2}$$

Note that $\phi_{\rm hb}'(R)$ is just the centripetal force (due to the halo and the bulge) on the Sun [really, on the Local Standard of Rest (LSR)] as it travels in its circular orbit about the Galactic center. This force plus the radial force due to the disk potential, $\phi_{\rm disk}'(R)$, must balance the centrifugal force on the LSR, V^2/R,

$$\frac{V^2}{R} = \phi_{\rm hb}'(R) + \phi_{\rm disk}'(R). \tag{4.3}$$

I model the disk density as falling off exponentially with radial distance. Then (Binney and Tremaine 1987)

$$\phi_{\rm disk}'(R) = 2\pi G\Sigma g(R/2h); \qquad g(x) \equiv e^{2x}x[{\rm I}_0(x){\rm K}_0(x) - {\rm I}_1(x){\rm K}_1(x)], \tag{4.4}$$

where h is the scaling radius of the disk, Σ is the disk column density at the

Sun, and the I and K are standard Bessel functions. Combining equations (4.2) through (4.4) and the definition of K gives

$$F = \frac{V^2}{2R^2} - \frac{K}{2R} g(R/2h). \tag{4.5}$$

The parameters in the above equation have the values (Binney and Tremaine 1987)

$$R \pm \Delta R = 8.5 \pm 1.0 \text{ kpc}; \quad h \pm \Delta h = 3.5 \pm 0.5 \text{ kpc}; \quad V \pm \Delta V = 220 \pm 15 \text{ km s}^{-1}. \tag{4.6}$$

Combining equations (4.5) and (4.6) gives

$$\frac{F}{F_\star} + .68 \frac{K}{K_\star} = 1.68 \pm .47 \tag{4.7}$$

where $F_\star$ and $K_\star$ are given by equation (3.6). (Recall that I chose $F_\star$ to be the expected value of F when $K = K_\star$.) The uncertainty in equation (4.7) was evaluated by considering ΔR, Δh, and ΔV to be independent, and evaluating the variance in F at $K = K_\star$,

$$\begin{aligned}(\Delta F)^2 = &\left(\frac{V}{R^2}\right)^2 (\Delta V)^2 + \left[\frac{K_\star}{4h^2} g'(R/2h)\right]^2 (\Delta h)^2 + \\ &\left[\frac{V^2}{R^3} + \frac{K_\star}{4hR} g'(R/2h) - \frac{K_\star}{2R^2} g(R/2h)\right]^2 (\Delta R)^2.\end{aligned} \tag{4.8}$$

It is of some interest that the ΔR term is by far the dominant one in the above expression. The contribution of the ΔV term is $\sim$ 3 times lower and that of the Δh term is $\sim$ 8 times lower.

A second source of effective density comes from the (possible) ways the velocity ellipsoid changes as a function of z. The most important such change is a rotation. At the Galactic plane, the dispersion velocity (squared) in the r direction is about 4 times that in the z direction. If the velocity ellipsoid rotates so that the long axis of this ellipsoid always points toward the center of the Galaxy, then for stars high above the plane, the observed doppler shift partially reflects this faster radial component. Thus, the velocity dispersion tends to increase with height. According to equation (3.2), an increasing velocity dispersion mimics the effect of a more slowly decreasing density. It thus makes us think that the potential is rising more slowly than it actually is. A positive F causes the potential to rise more rapidly than it would otherwise. Thus, this rotation may be characterized as a negative contribution to F. There is most likely some rotation of the velocity ellipsoid, but whether it is enough to point the long axis toward the Galactic center, or half that much, or some other amount is a matter of speculation. I therefore introduce a parameter β which is unity for the case of full rotation and zero for the case of no rotation. I confess almost complete ignorance as to the value of β and assign it to be

$$\beta = .7 \pm .3, \tag{4.9}$$

so that the entire range of reasonable possibilities is contained within 1-σ. It is easy to show that the corrections to F from this effect are given by

$$F_{\rm ellip} = -3\beta\frac{\overline{v^2}}{R^2}, \tag{4.10}$$

where $\overline{v^2}$ is the velocity dispersion of the tracers. Note that this effect varies from tracer to tracer depending on the velocity dispersion. It will also be a function

of z if the tracers are not isothermal. I will estimate this effect for my model isothermal tracers with velocity dispersion of 25 km s^{-1},

$$\frac{F_{\rm ellip}}{F_\star} \sim .08 \pm .04. \tag{4.11}$$

Note that the entire effect is small compared to the *uncertainty* in equation (4.7). I will therefore ignore it in the remainder of this paper.

To a good approximation, then, F (or $\rho_{\rm eff}$) may be regarded as an effect of the large-scale structure of the Galactic potential. Indeed, from equation (4.2), one may write

$$\rho_{\rm eff} = \frac{M_{\rm hb}(R)}{4\pi R^3}, \tag{4.12}$$

where $M_{\rm hb}(r)$ is the total mass of the halo and bulge interior to r. This equation makes clear in what sense $\rho_{\rm eff}$ is an "effective density." [It was on the basis of eq. (4.12) that I made the claim at the beginning of this section that F is known to be positive.]

It is clear from equation (4.5) why $K_\star/F_\star$ is of the same general order as R and h. These latter two quantities are the only lengths in the equation. The only dimensionless parameter entering this equation which could differ appreciably from unity is the ratio of the disk to the halo contributions to the centripetal force. As it happens, these two quantities are of the same general order. Thus $K_\star/F_\star$ must be of the same order as R and h.

5. Combined Theoretical and Experimental Estimate of K and F

It is straightforward to combine two independent measurements of a vector (see Appendix B). The covariance matrix is found by adding the contributing covariance matrices inversely. Thus, the joint (dimensionless) covariance matrix, $c^{\nu v+gs}$, is given by

$$c^{\nu v+gs} = [b^{\nu v} + b^{gs}]^{-1} = \left[N \begin{pmatrix} .205 & .0826 \\ .0826 & .0356 \end{pmatrix} + \frac{1}{.47^2} \begin{pmatrix} .68^2 & .68 \\ .68 & 1 \end{pmatrix} \right]^{-1}, \quad (5.1)$$

where $b^{\nu v}$ is the inverse of $c^{\nu v}$ [eq. (3.24)], and b^{gs} is the "inverse covariance matrix" associated with the Galactic structure analysis made in the last section. I have placed the words "inverse covariance matrix" in quotations because it is not the inverse of any matrix; it is degenerate. Nevertheless, it can be obtained in the standard way by writing down the χ^2 associated with equation (4.7), and differentiating with respect to the vector, $(K/K_\star, F/F_\star)$.

While it is easy to evaluate $c^{\nu v+gs}$ for any particular value of N, $c^{\nu v+gs}$ is not a simple function of N. There are three ways of dealing with this. First, the matrix may be evaluated for a particularly relevant case. Second, its elements may be graphed as functions of N. Finally, some limiting features which are independent of N may be analyzed analytically. I consider each of these in turn.

For the case $N = 350$, $c^{\nu v+gs}$ is found to be

$$c^{\nu v+gs} = \begin{pmatrix} .0738 & -.1378 \\ -.1378 & .3181 \end{pmatrix} \sim \frac{1}{350} \begin{pmatrix} 26 & -49 \\ -49 & 111 \end{pmatrix}. \quad (5.2)$$

Comparison of equations (3.24) and (5.2) shows that there is a dramatic improvement in the uncertainty of K (and a still more dramatic one in the uncertainty of

F) when the Galactic structure argument is applied. However, the uncertainty in K is still $\sim 27\%$.

Figure 2 shows the uncertainty in $K/K_\star$ as a function of the number of stars observed. Values are calculated both with (solid) and without (dashes) the theoretical constraint [eq. (4.7)]. Also shown in this figure is the uncertainty in $F/F_\star$, from the combined determination. This curve is fairly constant over the broad range where the uncertainty is dominated by its theoretical component. Even for $N = 800$ stars, the uncertainty in F is $\sim 50\%$. Can this uncertainty be substantially reduced, short of making a very large number of observations? The answer is, in principle, yes. I will discuss this further in §VII.

Finally, some intuitive understanding of the role of the theoretical constraint can be gained by posing the question: If an error is made in the theoretical estimate of F, how large will the resulting error in K be (assuming statistical uncertainties can be ignored)? Of course the qualification about statistical uncertainties never really applies. In the limit of small statistics this uncertainty is too large to be ignored and in the limit of large statistics one would just use the results of §III and would not bother with the theoretical constraint. Nevertheless, the calculation is instructive because it demonstrates the *limits* of the use of the theoretical constraint.

Suppose, then, that a theoretical estimate F is made for the quadratic parameter in equation (1.7), but the actual potential is described by a different quadratic parameter, f,

$$\psi'(z) = k + 2fz \qquad \delta F \equiv F - f. \tag{5.3}$$

I will assume that actual velocity dispersion, $\overline{v^2}$ and actual density

$$\nu(z) = A \exp\left(-\frac{kz + fz^2}{\overline{v^2}}\right), \tag{5.4}$$

are measured in each of n measurements. (That is, I will ignore statistical errors.) I will assume that a trial potential $\tilde{\psi}'(z) = K + 2Fz$ is varied in K to minimize χ^2. To answer the question, what is the induced error, $\delta K \equiv K - k?$, I follow my familiar procedure:

$$\chi^2(K) = \sum_{i=1}^{n} \frac{1}{\sigma_i^2}\left\{\overline{v^2} - [\nu(z_i)]^{-1} \int_z^\infty dz'(K + 2Fz')\nu(z')\right\}^2. \tag{5.5}$$

Using the same variance, (3.20), equation (3.3), and the definitions of δK and δF, this becomes

$$\chi^2(K) = \frac{1}{2(\overline{v^2})^2} \int_{L_1}^{L_2} \frac{\alpha z^2 dz}{\nu(z)}\left[\int_z^\infty dz'(\delta K + 2\delta F z')\nu(z')\right]^2. \tag{5.6}$$

Differentiating with respect to K gives

$$0 = \int_{L_1}^{L_2} \frac{\alpha z^2 dz}{\nu(z)}\left[\int_z^\infty dz'\nu(z')\right]\left[\int_z^\infty dz'(\delta K + 2\delta F z')\nu(z')\right]. \tag{5.7}$$

Again employing the condition (3.14), one finds

$$\frac{\delta K}{K} = -\frac{2\delta F\overline{v^2}}{K^2}\left[\frac{(6 + 6x + 3x^2 + x^3)e^{-x}\Big|_{KL_2/\overline{v^2}}^{KL_1/\overline{v^2}}}{(2 + 2x + x^2)e^{-x}\Big|_{KL_2/\overline{v^2}}^{KL_1/\overline{v^2}}}\right] \sim 17\%, \tag{5.8}$$

where I have used the standard values for $K, \overline{v^2}$, and the L_i, and equation (4.7) for the estimate of the error, δF.

Note that if K were 17% higher than the 1.3(km s^{-1})^{2}pc^{-1} which I have used in the above estimate, then the entire range for the theoretically predicted F [eq. (4.7)] would be moved lower by 0.23×10^{-4}(km s^{-1})^{2}pc^{-2}. Taking this coupling into account, I find that a 1-σ error in the estimate of F produces an error of $\sim 22\%$ in K. As expected, this error is below the ($\sim 27\%$) error found when statistical uncertainties were incorporated into the analysis.

6. Covariance Of $\Sigma_{\rm min}$ and F

In the preceding section I showed that F and K have a significant covariance, whether F is estimated purely from theory, or whether the experimental evidence for the relationship between K and F is weighed against the theoretical evidence. Regardless of the method of determination, if F is overestimated then K (or equivalently, Σ) will be underestimated. However, in this case $\Sigma_{\rm min}$ will also be underestimated. Since Δ_Σ depends on the difference between these two quantities, the error will be somewhat less than was indicated above. In Appendix A I show that if F is overestimated by δF, then the minimum column density $\Sigma_{\rm min}$ will be underestimated by

$$\frac{\delta\Sigma_{\rm min}}{\Sigma_{\rm min}} = -\frac{4\delta F v_m^2}{K_{\rm min}^2}\left[\ln\left(\frac{2}{1+\tau}\right) + \frac{1}{2}\tau\right], \tag{6.1}$$

where τ is thermal deviation defined by equation (2.11), and v_m is the mass-weighted velocity dispersion of the disk. Note that equation (6.1) has the limiting forms

$$\begin{aligned} \frac{\delta\Sigma_{\rm min}}{\Sigma_{\rm min}} &= -\frac{4\delta F v_m^2}{K_{\rm min}^2}(\ln 2 - \tau/2), \quad (\tau \ll 1); \\ \frac{\delta\Sigma_{\rm min}}{\Sigma_{\rm min}} &= -\frac{4\delta F v_m^2}{K_{\rm min}^2}\left[\frac{1}{2} + \frac{(1-\tau)^4}{8}\right], \quad (\tau \sim 1), \end{aligned} \tag{6.2}$$

for the extreme isothermal and non-isothermal observed distributions respectively. Thus, $\delta\Sigma_{\rm min}$ depends only very weakly on how isothermal the distribution is. I define the "irrelevant portion of δK" as the value of $\delta K_{\rm min}$ when τ and v_m take on their Milky Way values: $\tau \sim 0.25$; $v_m \sim 18{\rm km\ s^{-1}}$.

$$\frac{(\delta K)_{\rm irr}}{K} \equiv -4\delta F \frac{v_m^2}{K^2}\left(\ln\frac{2}{1.25}+\frac{.25}{2}\right) \sim -0.09\frac{\delta F}{F_\star}. \tag{6.3}$$

How important is this correction? To answer this question, I evaluate the ratio of $(\delta K)_{\rm irr}$ to the quantity δK found in the last section [eq. (5.8)] for the case when F was estimated purely theoretically,

$$\frac{(\delta K)_{\rm irr}}{\delta K} = -2\left(\ln\frac{2}{1.25}+\frac{.25}{2}\right)\frac{v_m^2}{\overline{v^2}}\left[\frac{(2+2x+x^2)e^{-x}\Big|_{KL_2/\overline{v^2}}^{KL_1/\overline{v^2}}}{(6+6x+3x^2+x^3)e^{-x}\Big|_{KL_2/\overline{v^2}}^{KL_1/\overline{v^2}}}\right] \sim -25\%, \tag{6.4}$$

where I have made the evaluation with my standard values of the L_i, and $\overline{v^2}$. Note in particular that equation (6.4) is independent of δF as befits a ratio of two quantities which are each linear in δF. Equation (6.4) demonstrates that this correction is a small effect and in no way undercuts the conclusions of the previous sections.

7. Using Constraints on K To Determine F

As I discussed in §V, high-altitude tracers cannot, by themselves, give very much information about F. The theoretical uncertainty in F is of order 50%, and only observation of several thousand tracers can significantly improve on this (see Figure 2). This problem is rooted in the large covariance between F and K. If K could somehow be constrained, then the high-altitude tracer data would become very sensitive to F.

Where could one find such an F-independent constraint on K? Measurements of K made in and near the disk do not depend strongly on F because the effective density term in the potential is small. This implies that the covariance between F and K from these measurements may, to a first approximation, be ignored. Of course, the disk potential is complicated and this means that K will be covariant with the several parameters which are required to represent this potential. Thus, determining K accurately from these measurements will not necessarily be easy. I will not inquire into these difficulties in this paper. I will simply assume that K has been measured to some specified accuracy,

$$\mathrm{var}(K/K_\star) = \gamma^2; \qquad \text{(Assumed Measurement From Disk)}, \tag{7.1}$$

and that this measurement has negligible covariance with F. I will then ask how well high-altitude tracers can measure F given such a constraint.

The "inverse covariance matrix" associated with this measurement (in my standard basis) is

$$b^{\mathrm{disk}} = \frac{1}{\gamma^2}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}. \tag{7.2}$$

The combined covariance matrix is then (by the resistor rule),

$$c^{\rm comb} = \left[b^{\rm disk} + b^{\nu v} + \left\{ \begin{matrix} 0 \\ 1 \end{matrix} \right\} b^{\rm gs} \right]^{-1}, \tag{7.3}$$

where the quantity in braces is 1 if one wishes to include the Galactic structure argument in the estimate of K and F, and is 0 if one wants to consider the tracer data independently of that argument. I have calculated the uncertainty in F for various values of γ and N and for both of the above cases. My results are given in Figure 3. The solid line shows the uncertainty in F as a function of γ for $N = 350$ and assuming that the Galactic structure constraint is included in the analysis. The dashed curves all assume that the galactic structure constraint has been ignored. They assume $N = 350$, 1050, 1750, and 2450 stars. Notice that if the error in K can be reduced below 15%, then all these curves show a significant improvement in the uncertainty of F relative to the 50% [eq. (4.7)] given by Galactic structure argument. The more high-altitude tracers are observed, the better F is measured. For example, if $N \sim 1000$ and the uncertainty in K can be reduced to $\sim 10\%$, then the uncertainty in F can be brought down to $\lesssim 30\%$.

8. Some Comments On The Analysis Of Kuijken And Gilmore

In a widely circulated preprint, Kuijken and Gilmore (1987) analyzed a sample of 530 K dwarf tracers which they had observed. They fit the density and velocity dispersion to curves and then used these results to fit the potential to a three parameter family of curves

$$\psi(z) = K[(z^2 + D^2)^{1/2} - D] + Fz^2. \tag{8.1}$$

[The formulae which they used to relate the potential to the density and velocity dispersion can be shown to be equivalent to my eq. (3.2), although this is not obvious at first glance.] Kuijken and Gilmore found that the fitting of equation (8.1) was insensitive to choice of D over the range $D \lesssim 300$pc. They attributed this to the fact that their data contained predominantly high-altitude stars and thus was naturally insensitive to this "low-altitude" parameter. (Actually, as I will detail at the end of this section, the situation is somewhat more complicated. The parameterization (8.1) destroys valuable information about the potential. The parameter D is *intrinsically* both difficult to fit and physically meaningless. These problems become worse the more evenly balanced is the data between high and low-altitude tracers.) For now the important point is that Kuijken and Gilmore obtained no more information about the potential with their three-parameter fit to all the data than they would have from a two-parameter fit to the high-altitude data. This makes it possible to compare their results with the theoretical analysis I gave in §III through §V.

Holding F fixed, Kuijken and Gilmore varied K to find the best fit. For two particular fixed values of F they found,

$$K = 1.4K_\star \quad (F \equiv 0); \qquad K = 1.0K_\star \quad (F \equiv 2F_\star). \tag{8.2}$$

(These results include slight adjustments for the velocity ellipsoid effects discussed in §IV.) By linear interpolation, one may estimate the "best fit" for K given an intervening choice of fixed F,

$$K/K_\star = 1.4 - 0.2F/F_\star \qquad \text{(best fit for fixed } F\text{)}. \tag{8.3}$$

They found that between the two trials shown in equation (8.2), the former actually gave the better fit. However, it was grossly out of line with their theoretical estimate of F,

$$\frac{F}{F_\star} + .85\frac{K}{K_\star} = 2.8 \pm 0.5 \qquad \text{(KG)}, \tag{8.4}$$

(The error bar is my reconstruction from their argument.) They thus used the latter value in equation (8.2) which was in conformity with their estimate (8.4). Kuijken and Gilmore then estimated the uncertainty in K, apparently by getting a best fit using fixed values for F at the extreme limits of equation (8.4). Their net result was $K = .96K_\star \pm 17\%$ which, they noted, is consistent with no missing matter. Kuijken and Gilmore emphasized that while their result was not strictly inconsistent with missing matter (particularly of low velocity dispersion) it did tend to rule out significant amounts of medium and high velocity-dispersion dark matter.

There are several subtle considerations which tend to undermine these conclusions. First, equation (8.4) substantially overestimates F. At $K = K_\star$, it gives roughly twice the value I calculated [eq. (4.7)]. Part of this difference derives from the fact that the Galactic parameters used by Kuijken and Gilmore ($R = 7.8 \pm .7$ kpc, $h = 4.5 \pm 1$ kpc) differ from those of equation (4.6). However, if the KG parameters are plugged into my equations (4.5) and (4.8), then one finds

$$\frac{F}{F_\star} + .5\frac{K}{K_\star} = 2.0 \pm .46 \qquad \text{(KG corrected)}, \tag{8.5}$$

roughly halfway between the KG estimate, (8.4), and my estimate (4.7). The remainder of the difference is accounted for by the fact that Kuijken and Gilmore

mistakenly took the halo contribution to F to be the *local* density of the halo, rather than its *effective* density, $M_{\rm halo}(R)/4\pi R^3$. [See eq. (4.12).] Had they fixed F according to my value [eq. (4.7)], or according to their own Galactic parameters [eq. (8.5)], they would have found [by eq. (8.3)] $K \sim 1.2K_\star$ or $K \sim 1.1K_\star$ respectively.

Second, when combining two estimates of a two-component vector, one must be careful to weight each appropriately, that is, by its inverse covariance matrix. When this weighting is done, the central value of F will not be same as its central value as determined by one of the estimates alone, but will be intermediate between the two central values. Thus Kuijken and Gilmore should not have fixed F at its central theoretical value, but allowed it vary to maximize the combined likelihood. This procedure (which I simulated in §V) would have driven the central value of F down toward the better fit in equation (8.2). By equation (8.3), this would have driven the central value of K up. This is a modest effect because the theoretical constraint on F, (4.7), is much tighter than the experimental constraint, (3.24).

Third, the error bars are larger than Kuijken and Gilmore expect. The same procedure for combining two estimates mentioned in the paragraph above, also gives the combined error. As I showed in §V, this error is substantially larger than that given by fixing F at the limits of the theoretical constraint and fitting for K. Specifically, I found that the combined uncertainty was 27% rather than 17%.

While I am not in possession of Kuijken and Gilmore's raw data, the above considerations lead me to suspect that a more careful analysis of them would

yield, $K = 1.25K_\star \pm 27\%$. This is still, as Kuijken and Gilmore claim, consistent with no missing mass. However, it is equally consistent with $K = 1.5K_\star$. By equations (2.10), this latter value corresponds to 33% high-dispersion, 48% medium-dispersion, or 65% low-dispersion missing matter. Thus, these data are not sufficient to either demonstrate or rule out the existence of substantial amounts of missing matter of any type.

I turn now to an examination of the parameterization, (8.1). In the limits of high and low z, this equation becomes

$$\psi(z) = \left(\frac{K}{2D} + F\right)z^2 \quad (z \ll D_\star); \qquad \psi(z) = -KD + Kz + Fz^2 \quad (z \gg D_\star). \tag{8.6}$$

The first equation is just the generic low-z behavior of any physical potential satisfying the boundary condition (1.4). From equation (1.13), one finds that it relates the parameter D to the total density at the plane, $\rho(0)$,

$$K/D = 4\pi G\rho(0). \tag{8.7}$$

The second equation is just the generic high-z potential of any physical potential. From equation (1.7), one finds that it relates the parameter D to the "potential offset", $\Delta\psi$,

$$KD = \Delta\psi. \tag{8.8}$$

The question is, for a given K can a single value of D reasonably satisfy both equations (8.7) and (8.8)? To answer this question, one must first understand what is the physical significance of $\Delta\psi$ and what measurements are sensitive to it.

At first sight it may appear that $\Delta\psi$ has no physical significance in as much as a potential is determined physically only up to an overall constant. However, this constant is already set by the boundary condition (1.4). Evaluating $\Delta\psi$ [by integrating eq. (2.3)], one finds that it is closely related to the mass-weighted velocity dispersion of the disk, v_m^2,

$$\Delta\psi = v_m^2 \left[2 \left(\ln \frac{2}{1+\tau} + \frac{\tau}{2} \right) \right]. \tag{8.9}$$

Here τ is the thermal deviation defined by equations (2.11) and (2.13). Note that the factor in square brackets is the same very weak function of τ which appeared in the formula for $\delta\Sigma_{\rm min}$ [eq. (6.1)]. For extreme isothermal and non-isothermal disks it is respectively 1.39 and 1.00. Consider then a class of disks of fixed K and v_m^2, but variable τ. As the thermal deviation increases, $\Delta\psi$ falls slowly while $\rho(0)$ rises rapidly. Thus, there is no reason to believe that one D can satisfy both equations (8.7) and (8.8). I give two examples. If the disk were composed entirely of its observed components as catalogued by Bahcall (1984a), then D would have a value of 240 pc by equation (8.7), and of 283 pc by equation (8.8). Kuijken and Gilmore believe that substantially more low velocity-dispersion interstellar medium is observed (13 $M_\odot {\rm pc}^{-2}$) than does Bahcall (5 $M_\odot {\rm pc}^{-2}$). If the disk were composed only of the observed components in this model, then D would have values of 146 pc and 231 pc according to the respective formulae.

Measurement of "D" in the first sense [$\rho(0)$] requires low velocity-dispersion tracers so that they are very sensitive to the behavior of the potential at low altitudes. Measurement of "D" in the second sense ($\Delta\psi$) requires relatively high velocity-dispersion tracers with a significant sampling near the plane. The

low-altitude sampling is necessary to normalize the tracer density and the high velocity dispersion is necessary to map the potential at high altitudes. Any particular sample will satisfy these conflicting criteria to various degrees and will consequently be sensitive to none, one, or both of these D's. However, even by combining several very good tracer samples at different velocity dispersions, one can never hope to pin down "D", nor would such a measurement have any significance if one could.

9. Conclusions

With present statistics, high-altitude tracer data cannot be used to seriously constrain the missing mass in the disk. The $N \sim 350$ stars observed so far allow the disk column density to be measured with a 1-σ error of $\sim 27\%$. If 1000 stars were observed, this would allow a determination within 20%. For $N > 1000$, the uncertainty declines more slowly than $N^{-1/2}$. At $N = 2000$, for example, the uncertainty is still $\sim 17\%$. Uncertainties of this magnitude do not allow any significant limit to be placed on the missing mass. This is because a constraint on the column density translates directly into a constraint on the amount of unobserved matter in high-velocity components, but leaves a lot of room for unobserved matter in low-velocity components. For example, if the column density were measured as equal to the density expected on the basis of the observed components alone (with an uncertainty of 20%), it would still be possible for the majority of the mass of the disk to be in unobserved components. Thus, high-altitude tracers cannot, by themselves, be used to constrain the missing matter unless a very large number of stars are observed. On the other hand,

they promise to be quite useful in constraining the amount of matter in high-velocity components. This is important because these are just the components to which low altitude tracers are least sensitive.

If K can be measured sufficiently well (within $\sim 15\%$) by low-altitude measurements, then it will be possible to use high-altitude tracers to determine F to much greater accuracy than it is presently known

These results have important implications for the planning of tracer observations. If high-altitude tracers actually played a decisive role in the determination of the missing disk mass, then an exceptional effort should indeed be made to observe a large number of these stars, despite the additional telescope time required and the additional risk of systematic errors inherent in such observations. However, as I have shown, these stars provide information on only one piece of the complicated puzzle of disk missing mass. Therefore, while some observations of high-altitude tracers are worthwhile, the principal focus should be on medium and low-altitude tracers.

A full analysis of how to balance these various observations must be left to a later paper.

ACKNOWLEDGEMENTS

I would like to thank John Bahcall, Michael Peskin, and Paul Schechter for their many critical comments and discussions. In addition, George Blumenthal, Joel Primack, Scott Tremaine, and Chris Wendt brought a number of important points to my attention.

APPENDIX A: Some Exactly Solvable Models

In this Appendix, I solve the 3-component disk model described in §II. I also calculate the reduction in the column density due to a small effective density $\rho_{\rm eff}$ (or F). The definitions of all quantities are the same as in §II,

$$K_{lm} \equiv K_l + K_m; \qquad K \equiv K_l + K_m + K_h, \tag{A.1}$$

$$D \equiv \frac{2v_h^2}{K}; \qquad D_{lm} \equiv \frac{2v_m^2}{K_{lm}}. \tag{A.2}$$

I begin by solving a simpler two-component model, where $K_h = 0$. In the limit $v_l^2 \ll v_m^2$, all the mass from the low-velocity component will be below the mass of the medium velocity component. Thus, the former may be taken to be a *surface density*. All of the matter at finite z will then have velocity dispersion v_m^2. The boundary conditions on the potential become

$$\psi(0) = 0; \qquad \psi'(0) = K_l. \tag{A.3}$$

Since the velocity dispersion is a constant, equation (3.3) may be written,

$$\rho(z)v_m^2 = \int_z^\infty dz' \frac{d\psi(z')}{dz'} \nu(z'). \tag{A.4}$$

Poisson's equation may be written

$$4\pi\rho(z) = h'(z), \tag{A.5}$$

where

$$h(z) \equiv \psi'(z) - 2Fz \tag{A.6}$$

is the derivative of that portion of the potential which is due to disk material (as

opposed to the effective density). Note that the definition of K_{lm} implies

$$h(\infty) = K_{lm}. \tag{A.7}$$

Combining equations (A.5) and (A.6) gives

$$\begin{aligned} h'(z)v_m^2 &= \int_z^\infty dz'\ h(z')h'(z') + 2F\int_z^\infty dz'\ h'(z')z' \\ &= \frac{1}{2}[h(\infty)^2 - h(z)^2] + 2F\int_z^\infty dz'\ h'(z)z'. \end{aligned} \tag{A.8}$$

For the case $F = 0$, one can immediately write down the solution to this equation which satisfies the boundary conditions (A.3),

$$h(z) = K_{lm}\tanh\left(\frac{z}{D_{lm}} + \tanh^{-1}\frac{K_l}{K_{lm}}\right). \tag{A.9}$$

Thus the density at the plane is just

$$4\pi G\rho_m(0) = h'(0) = \frac{K_{lm}}{D_{lm}}\mathrm{sech}^2\left(\tanh^{-1}\frac{K_l}{K_{l\,n}}\right) = \frac{K_{lm}^2 - K_l^2}{2v_m^2}. \tag{A.10}$$

To solve the problem for finite but small F, I write

$$h(z) = K_{lm}\tanh\left(\frac{z}{D_{lm}} + \tanh^{-1}\frac{K_l}{K_{lm}}\right) + F\xi(z). \tag{A.11}$$

I will want to find the effect of this effective potential term on the column density *given that the density at the plane is still given by equation* (A.10). This implies

the boundary conditions,

$$\xi(0) = \xi'(0) = 0. \tag{A.12}$$

Substituting equation (A.11) into equation (A.8) gives

$$\begin{aligned}\frac{\xi'(z)}{2D_{lm}} = &\left[\xi(\infty) - \xi(z)\tanh\left(\frac{z}{D_{lm}} + \tanh^{-1}\frac{K_l}{K_{lm}}\right)\right. \\ &\left. + 2\int_z^\infty z' d\tanh\left(\frac{z'}{D_{lm}} + \tanh^{-1}\frac{K_l}{K_{lm}}\right)\right] + O(F).\end{aligned} \tag{A.13}$$

Using equations (A.12) and (A.13), one may evaluate $\xi(\infty)$,

$$\xi(\infty) = -2\int_0^\infty z d\tanh\left(\frac{z}{D_{lm}} + \tanh^{-1}\frac{K_l}{K_{lm}}\right) = -2D_{lm}\ln\left(\frac{2K_{lm}}{K_{lm}+K_l}\right). \tag{A.14}$$

It would now be possible to substitute equation (A.14) back into equation (A.13) and solve for $\xi(z)$. However, I am interested only in determining the effect of F on the column density and not on the details of the potential. By equation (A.7), this effect is

$$\delta K_{lm} = F\xi(\infty) = -2FD_{lm}\ln\left(\frac{2K_{lm}}{K_{lm}+K_l}\right). \tag{A.15}$$

I now turn to the three component problem, first with $F = 0$. Since $\rho_h(0) \ll \rho_m(0)$ and $v_h^2(0) \gg v_m^2(0)$, the entire high-velocity distribution may be considered to be above the medium and low-velocity distributions. That is, in this limit, the entire column density K_{lm} may be considered to be a surface density. This

means that one may simply make the substitutions

$$K_l \rightarrow K_{lm}; \quad K_{lm} \rightarrow K; \quad D_{lm} \rightarrow D, \tag{A.16}$$

into equation (A.9), and write down the high-velocity part of the potential. Combining this result with equation (A.9) yields equation (2.3),

$$\frac{d\psi}{dz} = K \tanh\Big(\frac{z}{D} + \tanh^{-1}\frac{K_{lm}}{K}\Big) - K_{lm}\Big[1 - \tanh\Big(\frac{z}{D_{lm}} + \tanh^{-1}\frac{K_l}{K_{lm}}\Big)\Big], \tag{A.17}$$

Lastly, I calculate the effect of a small F on the total column density of the three-component model. The reduction of the low and medium-velocity components is given by equation (A.15). If the influence of F is for the moment ignored, this reduced value of K_{lm} leads to an increased value of $K_h \equiv K - K_{lm}$ because the high-velocity component has the same plane density, but is more weakly held to the plane. This effect (which I denote δK_1) may be evaluated by considering equation (A.10),

$$(K + \delta K_1)^2 - (K_{lm} + \delta K_{lm})^2 = K^2 - K_{lm}^2, \tag{A.18}$$

or

$$\delta K_1 = \frac{K_{lm}}{K}\delta K_{lm}. \tag{A.19}$$

Equation (A.19) takes account of both the increase in K_h and the decrease in K_{lm}. Finally, the effect of F on the high-velocity distribution (which I denote δK_2) is the same as its effect on the medium-velocity distribution (with an appropriate

substitution of parameters),

$$\delta K_2 = -2FD \ln\left(\frac{2K}{K + K_{lm}}\right). \tag{A.20}$$

Combining both terms gives

$$\delta K = \delta K_1 + \delta K_2 = -\frac{4Fv_m^2}{K}\left[\frac{v_h^2}{v_m^2}\ln\left(1 + \frac{K_h}{2K - K_h}\right) + \ln 2 - \ln\left(1 + \frac{K_l}{K_{lm}}\right)\right]. \tag{A.21}$$

If one holds the total velocity dispersion fixed at v_m^2,

$$K_l v_m^2 = K_h(v_h^2 - v_m^2) \sim K_h v_h^2, \tag{A.22}$$

and works in the limit $v_m^2 \ll v_h^2$, then equation (A.21) reduces to

$$\delta K = -\frac{4Fv_m}{K}\left[\ln\left(\frac{2}{1+\tau}\right) + \frac{1}{2}\tau\right], \tag{A.23}$$

where

$$\tau \equiv \frac{K_l}{K}. \tag{A.24}$$

This completes the proof.

APPENDIX B: The Resistor Rule For Covariances

In this Appendix, I prove the resistor rule for covariance matrices mentioned in §III. Before doing so, however, I briefly review the procedure for calculating covariance matrices.

Let y_k be a series of independent measurements of a function taken at positions x_k with variances σ_k^2. (In this simplified treatment I will assume that the positions x_k have negligible uncertainty.) Suppose that these measurements are to be fit to a linear combination of p trial functions, $f_i(x)$. Then χ^2 is given by

$$\chi^2(a_1 ... a_p) = \sum_{i,k} \frac{1}{\sigma_k^2}[a_i f_i(x_k) - y_k]^2 = \sum_{i,j} a_i b_{ij} a_j - \sum_i 2a_i d_i + h, \qquad \text{(B.1)}$$

where

$$b_{ij} \equiv \sum_k \frac{f_i(x_k) f_j(x_k)}{\sigma_k^2}, \quad d_i \equiv \sum_k \frac{f_i(x_k) y_k}{\sigma_k^2}, \quad h \equiv \sum_k \frac{y_k^2}{\sigma_k^2}, \qquad \text{(B.2)}$$

Note that b_{lm} is the covariance matrix of the d_l,

$$\begin{aligned} \langle d_l d_m \rangle - \langle d_l \rangle \langle d_m \rangle &= \sum_{k,q} \frac{f_l(x_k) f_m(x_q)}{\sigma_k^2 \sigma_q^2} \left(\langle y_k y_q \rangle - \langle y_k \rangle \langle y_q \rangle \right) \\ = \sum_{k,q} \frac{f_l(x_k) f_m(x_q)}{\sigma_k^2 \sigma_q^2} \sigma_q^2 \delta_{kq} &= \sum_k \frac{f_l(x_k) f_m(x_k)}{\sigma_k^2} = b_{lm}. \end{aligned} \qquad \text{(B.3)}$$

The first step used the definition of the d_l and the fact that the $f_l(x_k)$ are constants. The second used the fact that the y_k are independent with variances σ_k^2. Setting the derivative of equation (B.1) with respect to a_l equal to zero, one

obtains the best fit (which minimizes χ^2),

$$a_i = \sum_j c_{ij} d_j, \qquad c \equiv b^{-1}. \tag{B.4}$$

One may show that c is indeed the covariance matrix by direct evaluation.

$$\begin{aligned} \mathrm{covar}(a_i a_j) \equiv \langle a_i a_j \rangle - \langle a_i \rangle \langle a_j \rangle &= \sum_{l,m} c_{il} c_{jm} \left[\langle d_l d_m \rangle - \langle d_l \rangle \langle d_m \rangle \right] \\ &= \sum_{l,m} c_{il} c_{jm} b_{lm} = c_{ij}. \end{aligned} \tag{B.5}$$

In equation (B.5), the second step follows from equation (B.4) and the fact that the c_{ij} are constants. The third step follows from equation (B.3), and the fourth from the definition of c.

I turn now to the resistor rule. First, suppose a set of n independent determinations of the parameters a_i are made "in parallel". These may each separately be described by quantities b_{ij}^r, d_i^r, and h^r which are defined analogously to the quantities in equation (B.2). Then

$$\chi^2 = \sum_{i,j} a_i B_{ij} a_j - \sum_i 2 a_i D_i + H, \tag{B.6}$$

where

$$B_{ij} \equiv \sum_r b_{ij}^r, \qquad D_i \equiv \sum_r d_i^r, \qquad H \equiv \sum_r h^r. \tag{B.7}$$

Since the d_i^r, d_j^r each separately satisfy equation (B.3), and since d_i^r is independent

of d_j^s for $r \neq s$,

$$\left\langle d_i^r d_j^s \right\rangle - \left\langle d_i^r \right\rangle \left\langle d_j^s \right\rangle = b_{ij}^r \delta^{rs}. \tag{B.8}$$

It follows that the D_i and B_{ij} satisfy their own version of equation (B.3),

$$\langle D_i D_j \rangle - \langle D_i \rangle \langle D_j \rangle = \sum_{r,s} \left\langle d_i^r d_j^s \right\rangle - \left\langle d_i^r \right\rangle \left\langle d_j^s \right\rangle = \sum_{r,s} b_{ij}^r \delta^{rs} = B_{ij}. \tag{B.9}$$

Thus, defining $C \equiv B^{-1}$ and using the (upper case) analog of equation (B.5), above, one finds that B_{ij} is the inverse covariance matrix for the a_i. This proves one half of the resistor rule, that covariances from measurements made "in parallel" add inversely. [Note that in the above derivation, I did not assume that the individual b_{ij}^r were invertible. I assumed only that they reflected independent measurements and satisfied eq. (B.3). Thus, in particular, the argument applies to degenerate constraints like eq. (4.7) and its corresponding "inverse covariance matrix" b^{gs}, eq. (5.1), even though the latter is not invertible.]

The other half of the resistor rule says that covariances from measurements made "in series" add directly. Suppose that a vector a_i is determined jointly from two independent sets of measurements, x_k^1 and x_l^2. If for any particular measurement these vectors differ from their means values by δx_k^1 and δx_l^2, then a_i will differ from its mean value by

$$\delta a_i = \sum_k \frac{\partial a_i}{\partial x_k^1} \delta x_k^1 + \sum_l \frac{\partial a_i}{\partial x_l^2} \delta x_l^2. \tag{B.10}$$

I define c_{ij}^1 to be the covariance of a_i ($\langle \delta a_i \delta a_j \rangle$) when it is assumed that the uncertainties of the second measurement can be ignored (that is, when $\partial a_i / \partial x_l^2$

is set to zero.) Similarly, I define c^2_{ij} to be the covariance of a_i when the uncertainties in the first measurement are ignored. Finally, I define C_{ij} to be the covariance of a_i when both uncertainties are taken into account. From equation (B.10) and these definitions one finds that

$$c^1_{ij} \equiv \sum_{k,k'} \frac{\partial a_i}{\partial x^1_k} \frac{\partial a_j}{\partial x^1_{k'}} \left\langle \delta x^1_k \delta x^1_{k'} \right\rangle , \qquad c^2_{ij} \equiv \sum_{l,l'} \frac{\partial a_i}{\partial x^2_l} \frac{\partial a_j}{\partial x^1_{l'}} \left\langle \delta x^2_l \delta x^2_{l'} \right\rangle , \tag{B.11}$$

and

$$\begin{aligned} C_{ij} \equiv & \sum_{k,k'} \frac{\partial a_i}{\partial x^1_k} \frac{\partial a_j}{\partial x^1_{k'}} \left\langle \delta x^1_k \delta x^1_{k'} \right\rangle + \sum_{k,l'} \frac{\partial a_i}{\partial x^1_k} \frac{\partial a_j}{\partial x^2_{l'}} \left\langle \delta x^1_k \delta x^2_{l'} \right\rangle \\ & + \sum_{k',l} \frac{\partial a_i}{\partial x^1_{k'}} \frac{\partial a_j}{\partial x^2_l} \left\langle \delta x^1_{k'} \delta x^2_l \right\rangle + \sum_{l,l'} \frac{\partial a_i}{\partial x^2_l} \frac{\partial a_j}{\partial x^1_{l'}} \left\langle \delta x^2_l \delta x^2_{l'} \right\rangle \\ = & c^1_{ij} + 0 + 0 + c^2_{ij}. \end{aligned} \tag{B.12}$$

The last step in equation (B.12) follows from equation (B.11) and the independence of x^1_k and x^2_l. This completes the proof.

REFERENCES

1. Bahcall, J. N. 1984, *Ap. J.*, **276**, 169.

2. Bahcall, J. N. 1984, *Ap. J.*, **287**, 926.

3. Binney, J. and Tremaine, S. 1987, *Galactic Dynamics* (Princeton: Princeton University Press).

4. Kuijken, K. and Gilmore, G. 1987, *Nature*, submitted.

5. Oort, J. H. 1932, *Bull. astr. Insts, Neth.*, **6** 249.

6. Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T. 1986, *Numerical Recipes*, (Cambrigde: Cambridge University Press).

FIGURE CAPTIONS

1) Missing mass (Δ_T) as a function of thermal deviation τ. Shown are cases where the unobserved components have high (dots), medium (dashes), and low (solid) velocity dispersions. A column density discrepancy (Δ_Σ) of 20% is assumed. Note that τ for the Milky Way disk is about .25.

2) Fractional uncertainties in K and F as functions of star count (N). For K, uncertainties are shown both with (solid) and without (dashes) the theoretical constraint [eq. (4.7)]. For F, the fractional uncertainty (dots) includes the theoretical constraint.

3) Fractional uncertainty in F assuming that K can be independently measured to a fractional accuracy γ. The solid curve assumes a star count of 350 and includes the theoretical constraint. The dashed curves ignore the theoretical constraint and assume star counts of 350, 1050, 1750 and 2450 respectively.

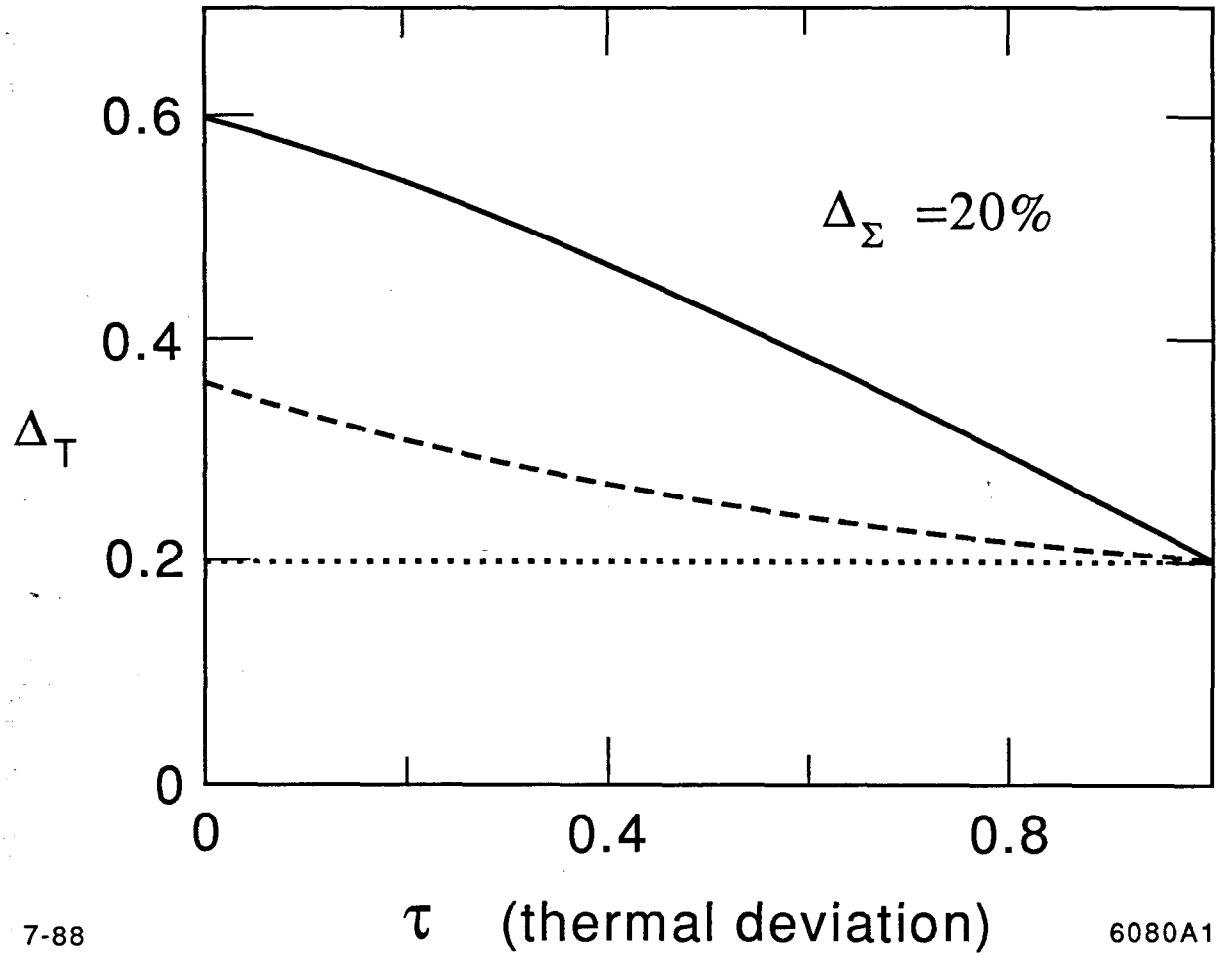
0.6
0.4
0.2
0
Δ_T
Δ_Σ =20%
0
0.4
0.8
τ (thermal deviation)
7-88
6080A1

Fig. 1

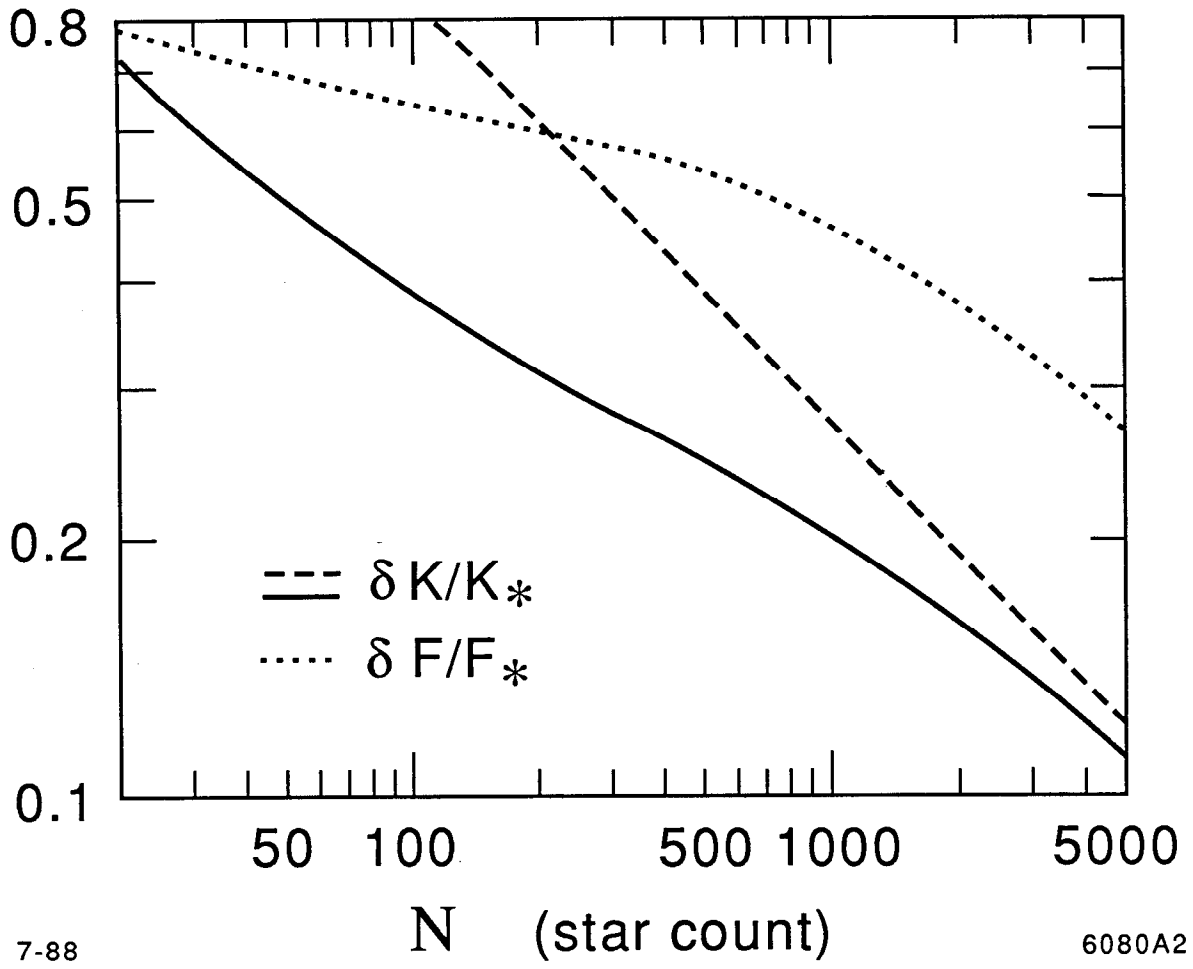

Fig. 2

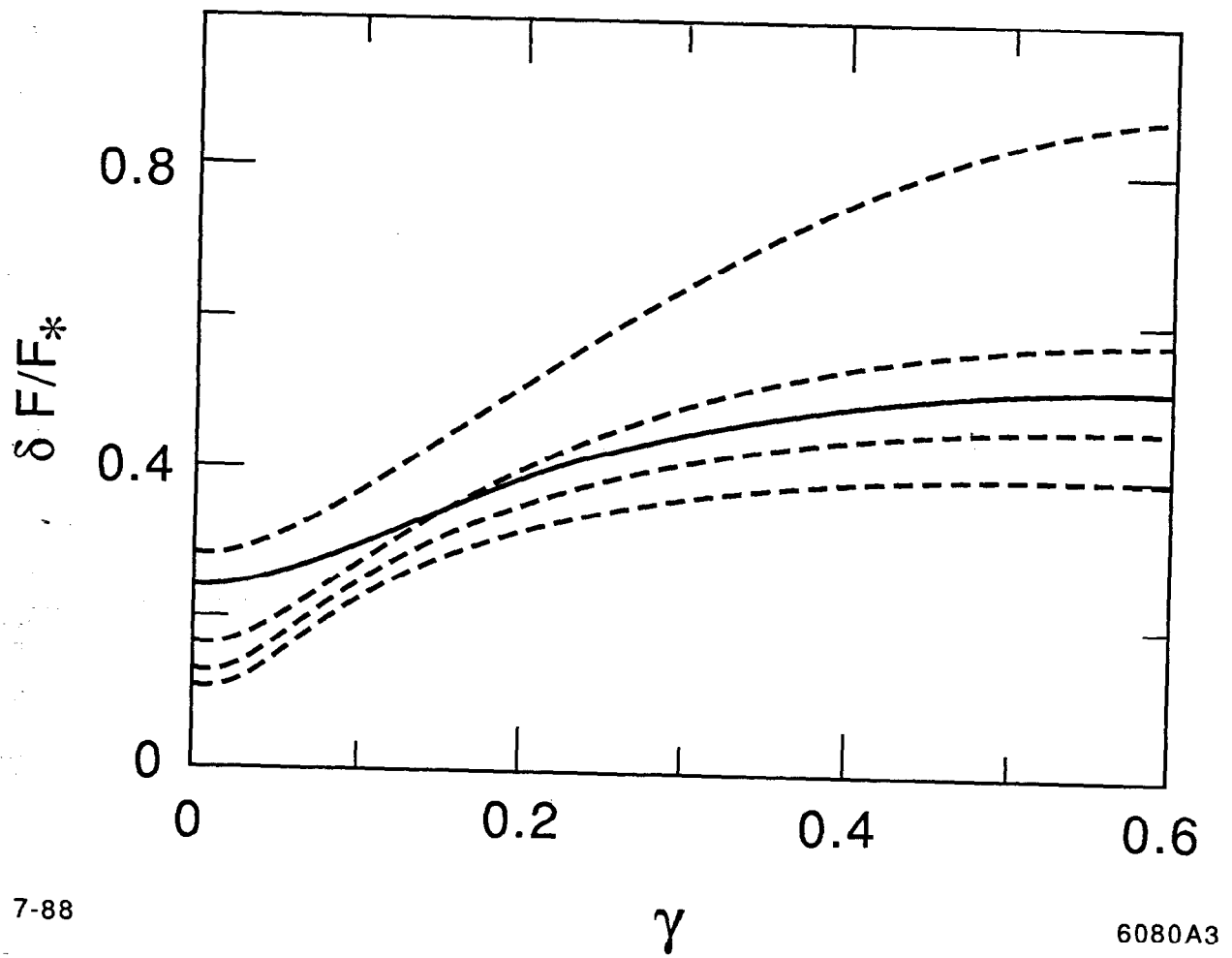

Fig. 3