

III

WARPS

Alar Toomre
Department of Mathematics
Massachusetts Institute of Technology

The large-scale warps of disk galaxies remain a challenge for theorists. Recent speculations about their origins or longevity have tended to focus on the possible benefits of halos. After some remarks about modes, this review will do likewise. It will conclude that any applause even there can as yet be only scattered rather than thunderous.

1. WHY NOT MODES?

Much of the sad state of this subject can be traced to the old rumor by Hunter and Toomre (1969, hereafter HT) which alleged that no isolated thin disk without an implausibly sharp outer edge admits discrete $m = 1$ modes of bending, let alone any that would permit it to remain warped for very long in a steadily precessing integral-sign shape. In retrospect, I have often wished this rumor were false. Life would indeed have been sweeter for warp theorists if HT had confirmed rather than contradicted Lynden-Bell's (1965) idea that the self-gravity of bent disks might enable them to avoid the differential precession to oblivion that was emphasized already by Kahn and Woltjer (1959). Faced with all the reports from radio astronomers since the mid-1970's that warped outer disks are remarkably common even in galaxies without close companions, we could now be smiling "we told you so" instead of scrambling around looking for excuses!

Could the mode idea have been right after all, and its dismissal too hasty? Anyone tempted to ponder this important question should remember that HT found plenty of discrete modes when their cold, self-gravitating disks were chosen sufficiently sharp-edged. What was it about the outer edges that made such a drastic difference?

To understand that sensitivity in modern terms, start with a ring of test particles in orbit near the periphery of a galaxy, each bobbing up and down with frequency $\kappa(r)$ about the plane $z = 0$ while circling with angular speed $\Omega(r)$.² Much like Lindblad's kinematic waves in that same plane, any $\sin m\theta$ or $\cos m\theta$ vertical distortions of this ring will then advance in longitude θ either with the fast pattern speed

$$\Omega_F(r) = \Omega(r) + \kappa_z(r)/m \quad (1)$$

or else with the slow pattern speed

$$\Omega_S(r) = \Omega(r) - \kappa_z(r)/m \quad . \quad (2)$$

In the $m = 1$ context where the ring merely tilts as if rigid, the first type of motion is essentially an Eulerian nutation or wobble; as Lynden-Bell pointed out, it can occur even when a ring is flabby. The second kind of $m = 1$ motion will probably seem more familiar, especially when $\kappa_z > \Omega$, as must be true near any oblate object that provides a restoring torque back toward $z = 0$; it represents the slow retrograde drift or "regression" of the line of nodes of our tilted gyroscope.

To build a mode in a disk consisting of many rings, one requirement is obvious: Somehow or other, these rates of regression or nutation must be adjusted to exact constancy from radius to radius. In practice, this means using the gravity from neighboring rings to augment κ_z — since there can be little doubt that any local self-gravity tends to increase the vertical stiffness, and the more so the more corrugated we imagine the disk to be. The mathematics can get tedious when dealing with warps as gradual as the integral signs that are our prime concern, but as usual it simplifies greatly in the short-wave or WKBJ approximation. Then the self-attraction of a corrugated thin disk of projected density μ and radial wavenumber k can easily be shown to boost its effective κ_z^2 by the amount $2\pi G \mu |k|$, or to modify the precession rates to

$$\Omega_P(r;k) = \Omega(r) \pm \frac{1}{m} [\kappa_z^2(r) + 2\pi G \mu(r) |k|]^{1/2} \quad . \quad (3)$$

As this formula indicates, the $m = 1$ rates of regression or nutation of outlying material in a thin disk can probably be increased enough by local forces to keep pace with the more rapidly precessing interior. However, we also see that such insistence may well penalize a disk with excessive corrugations in any broad outer zone of low density.

HT used essentially the above reasoning to conclude that one should expect distinct and well-behaved $m = 1$ modes of bending only in those unlikely disks whose reciprocal density $1/\mu$ remains integrable all the way to a postulated sharp outer edge. Much as I wish otherwise, this old argument still seems valid and it is buttressed, as we are about to see, by an even more blatant difficulty caused by the group velocity

$$c_g = m \partial \Omega_P / \partial k = \text{sgn}(k) \cdot \pi G \mu(r) / (\Omega_P - \Omega) \quad . \quad (4)$$

In plain English, equation (4) cautions that all local properties of a bending wave of trailing spiral planform ($k > 0$) tend to drift radially inward if that wave happens to be slow ($\Omega_P < \Omega$) and outward if it is fast ($\Omega_P > \Omega$), whereas for leading waves those senses of propagation are just the opposite. Now it may seem exotic to talk of spiral warps when mainly interested in stable modes which themselves must exhibit straight nodal

lines and hence no spiral preference. But such talk is less farfetched than it sounds. Even decent bending modes — like various bar modes in the plane — can be regarded as superpositions of leading and trailing waves in equal measure. And granted that, it is only sensible to ask what processes can possibly replenish those constituent waves that drift radially despite our labors to fine-tune their pattern speeds.

The answer is easy enough for any proposed slowly-regressing $m = 1$ mode. Its inward-drifting trailing waves will in due course presumably transmit/reflect from the interior as outward-drifting leading waves, and thus they will indeed replenish briefly the leading waves that exist now. But the crisis will come as those new waves continue to travel outward with the speed c that decreases roughly like the density $\mu(r)$: Unless there is a sharp g edge nearby to reflect them back as trailing waves to complete the cycle, we see that it can take literally forever for such wave information to drift out more and more sluggishly, never to return. A similar fate awaits a would-be fast mode — just swap the words leading and trailing — as illustrated vividly by the impulse response shown in Figure 8 of HT. In this mechanical sense, the outer parts of any gently tapered disk must therefore act very much like an absorbing beach to all $m = 1$ waves that might have sought to slosh up and down within it.

Modes aside, some beautiful examples of periodically forced bending waves that lend themselves almost ideally to WKBJ reasoning were reported recently by Shu, Cuzzi and Lissauer (1982). The foremost of these is a slow $m = 4$ trailing spiral wave pattern that propagates inward from the radius where it continues to be excited thanks to a 5:3 resonance with a slightly inclined satellite. The name of that satellite is Mimas, and the locale is the A ring of Saturn.

2. FOUR RECENT SUGGESTIONS

In the absence of clearcut modes, at least four ideas have surfaced in print lately seeking to explain the observed warps of galaxies with the active help or passive acquiescence of halos. The "active" ideas have stressed two possible sources of fresh excitation, whereas the "passive" speculations have focused more on ways to reduce the bothersome rates of differential precession. Let me deal first with the activists.

2.1 Mathieu instability

Binney (1978, 1981) pointed out that a steadily revolving bar-like halo or a genuine central bar may spontaneously excite z -oscillations in a galactic disk near certain resonant radii, via the classic phenomenon known as Mathieu instability. What matters crucially here is that some significant portion of the vertical stiffness contributed by that bar or halo should appear to a given star or gas atom to vary periodically with time at almost exactly twice the intrinsic frequency κ_z (ignoring the local self-gravity for the moment) with which that item and its neighbors would prefer to bob up and down. If this resonance condition can be met,

then it will be possible for each vibrating particle to acquire just such a phase of oscillation that its vertical spring rate will be stronger than average whenever it approaches the plane $z = 0$, and weaker than average when it recedes. Barring losses, the net result should be a steady gain of amplitude — much as achieved by an energetic child on a swing.

The condition for Mathieu resonance at radius r follows from noting that the bar-like potential needs to advance or recede relative to the orbiting particle by exactly two cycles, or by one full turn, during one intrinsic vertical period. Hence that relative speed must itself equal $\pm \kappa_z$, or the imposed bar or halo must revolve in space simply with one or the other of the elementary $m = 1$ speeds $\Omega \pm \kappa_z$ appropriate to that radius. Assuming the resulting amplitudes to be independent of θ , it is easy to see that those motions again correspond to the aforesaid fast nutation and slow regression, respectively.

Unfortunately, the addition of local self-gravity to this incomplete story uncovers what may well be a fatal flaw. I am referring here not to my own distaste for postulating any retrograde-moving triaxial halos, nor to the lack of evidence as yet that the resulting wave propagation (here presumably outward, for fast and slow waves alike) even from a fast bar resonance can yield warps resembling those observed. My worry is simply this: Unlike that production of fresh $m = 4$ bending waves by Mimas — from scratch at a known rate — the Mathieu process manages only to amplify by a finite and modest factor per revolution whatever $m = 1$ vibrations already exist near its resonance radius. And if the group velocity always carries waves away and never returns them, there may soon be precious little left to be amplified. Thus, I am afraid what seemed "a major breakthrough" to Saar (1979) in the last IAU review of this subject may yet fizzle away to nothing.

2.2 Flapping instability

Flapping instabilities as such have been known for a long time, but until recently they were not thought relevant to the observed major warps of galaxies. It was presumed that they arise only in circumstances that are either too small-scale — like the two-stream or "hose" instability recently revisited by Bertin and Casertano (1982), who insist that the scales may be sizable after all — or else too preposterous, such as the pair of superposed and contra-rotating Maclaurin disks which HT (p.752) knew to be unstable for $m \geq 2$. Both of these related instabilities seemed to demand a fairly stiff coupling between the oncoming streams, to ensure that they deform vertically almost like one.

Bertin and Mark (1980) were the first to recognize that such lateral stiffness is unnecessary, and that even a single thin sheet of matter that slides rapidly through a hot background of stars, neutrinos, or whatever may gravitationally amplify its corrugations — somewhat like a flag flapping in the wind. Their original derivation of this important result still frightens me, but by another route I concur that such amplification is indeed possible, at least at moderate wavelengths. I also agree that

the local speeds of any growing corrugations must be intermediate between those of the disk and the spheroid/halo. And lastly I commend Bertin and Mark for realizing at once that their amplification — like the Mathieu kind just discussed — is convective, and thus requires wave feedback.

Alas, I doubt for two reasons that even these interesting disk-halo interactions have anything to do with the observed $m = 1$ warps:

1. It remains very much to be demonstrated that warps as gentle in the vital θ direction as our $m = 1$ tilted rings can amplify in this manner. One should not forget that even those furiously contra-rotating Maclaurin disks cited by HT showed no $m = 1$ instabilities.

2. As Bertin and Mark understood, their need for steady feedback postulates a discrete $m = 1$ mode that rotates forward more rapidly than the spheroid/halo but more slowly than the bulk of the disk. Earlier we saw why discrete ring-tilting modes remain so elusive, and also why any slowly advancing $m = 1$ mode would amaze gyroscopists. Evidently Bertin and Mark thought otherwise, judging from all their talk about corotation radii and a "type of bending mode different from that of HT". But their reasoning there rested on a blunder. In their dispersion relation (II-3), they simply forgot the frequency κ_z due to distant matter.

Ironically, HT themselves (p.762) reported that some $m \geq 2$ modes of bending remain discrete even in disks with very blurry edges. Hunter (1969), who studied them extensively, noted that those modes generally do drift forward with modest speeds Ω_P . Hence I think they consisted essentially of slow multi-armed waves trapped between the center and the radius at which $\Omega_P = \Omega - \kappa_z/m$. Hunter and I have long wondered quietly whether such modes or their z -kin might be partly responsible for the minor ripples known to exist in our Galaxy interior to the Sun — and well they might be, if the Bertin-Mark process indeed manages to amplify them.

2.3 Nearly spherical potentials

Turning now to two proposals where the halo remains passive, I shall be briefer since they are more self-explanatory. For instance, on the premise that massive halos might be almost spherical, Tubbs and Sanders (1979) cautioned that one should not worry too much about differential precession. They stressed that the gaseous warps of galaxies like NGC 5907 (Sancisi 1976) or M83 (Rogstad, Lockhart and Wright 1974) occur far enough outside their optical disks that it remains at least arguable that such shapes could still be the (mangled?) relics of very ancient warps.

I think Tubbs and Sanders had a marginally valid point with those two galaxies and some others. To recapture the gist of their argument, remember first that those gas disks are huge: Out there near $r = 40$ kpc in NGC 5907, for instance, one revolution alone takes about 1.3×10^9 years. Even so, one probably needs to insist that any nodal regression rates be almost ten times slower; otherwise in 10^{10} years the rings at, say, $r = 30, 35, 40$ and 45 kpc would have drifted hopelessly out of

phase. Tubbs and Sanders noted correctly that the quadrupole potential from the known stellar disks is unlikely to create much trouble at such radii, and they also knew that any outlying dark matter needed to produce the relatively flat rotation curves of those galaxies cannot remotely be as flattened as the gas disks lest it cause too much regression. Assuming for simplicity that the halo potential is constant on similar spheroids of axis ratio c/a , it is easy to confirm that this tolerable flattening is indeed rather small: In that special case, the vertical frequency κ_z about the equatorial plane is everywhere just a multiple a/c of the local angular speed Ω , and it follows from the regression rates $\dot{\Omega}_S = \Omega - \kappa_z$ that the potentials had better be round to about one part in ten. Halo ^z densities, to be sure, can be somewhat more oblate, but even they ought probably not exceed E2 shapes to leave some margin for disk self-gravity. In short, uncomfortable maybe, but not impossible.

The situation remains grim, however, with the outer warp of our own Galaxy — and by analogy probably also those of M31 and M33. Simply the far-field formulas cited by Tubbs and Sanders imply that the regression rates $\dot{\Omega}_S$ in a pure exponential disk should amount to -0.147 , -0.102 and -0.074 times $\Omega(r)$ at radii equal to 4, 5 and 6 scale lengths. At those rates and radii, I reckon that rings of test particles would regress some 130° , 65° and 35° , respectively, every 10^9 years, assuming that the orbital period equals 2.5×10^8 years at the distance of 3 scale lengths roughly appropriate to the Sun. And these are but underestimates, ignoring all extra nuisance from local gravity. Hence I still maintain that the survival of our fairly clean and hardly spiral warp would become remarkable in less than 10^9 years if we continue to pretend that our Galaxy consists only of a disk, and after at most 3×10^9 years if we adopt the most generous disk-halo composite. What a pity that the near-perpendicular orientation and the large tip velocity of the Magellanic Stream now exclude even the LMC from coming to our aid.

2.4 Increasingly oblate halos

Petrou (1980) added a nice logical twist to the reasoning by Tubbs and Sanders. She urged that, rather than deplore any pronounced oblateness of halos, we should actually welcome it — provided such flattening increases outward rapidly enough to compensate for the decreasing $\Omega(r)$ by a steady increase of the relative stiffness κ_z/Ω about the imagined common equatorial plane of the disk and the halo. ^z Petrou reminded us all that the goal is not to avoid the gyroscopic regression as such, but only to make sure that its speed $\kappa_z - \Omega$ is almost identical at every radius. She also illustrated with some ^z examples that this goal can in principle be met handily for several known huge warps including that of NGC 5907.

As implied above, I nonetheless doubt that even Petrou's idea can ensure for 10^{10} years the survival of the relatively close-in HI warp of our Galaxy against the torques from our minimal known stellar disk. Even if it can, there remains one philosophical worry: Kinematic theories concerned only with survival do not explain why warps should have begun with their nodal lines so neatly aligned from radius to radius.

3. STEADY FORCING BY A TILTED HALO?

To end on an upbeat note, let me conclude with a few tentative words about another idea involving halos that has long struck me (a) almost as outrageous astronomically, and yet (b) almost as delightful dynamically, as it would be to postulate that an invisible ring of massive satellites encircles a disk galaxy in a nearby and inclined orbit. The tidal force of that ring would certainly ensure that the central galaxy gets and stays warped about a single (though precessing) line of nodes. And even better, as HT (p.768) noted already, such steady forcing would yield much the same results regardless of whether the victim disk has a sharp or fuzzy edge.

Alternatively, as a classic example, imagine that our precessing Earth were surrounded by a disk of debris even more extensive than that around Saturn. Ever since Laplace it has been known that the equilibrium shape of such a disk would be dominated nearby by our equatorial bulge, and far away by the Sun and the Moon — with the warped outcome shown in Figure 1. Notice the shift in allegiance that would occur near $r_{\text{crit}} \approx 7.7$ Earth radii: It is so rapid because the ratio of the outer versus quadrupole torques varies here like r^5 for equal angles.

Along with Dekel and Shlosman (this volume), who had the same notion, I have been wondering lately whether something analogous might not still afflict disk galaxies that lacked the good manners to be born with spin vectors exactly parallel to the short axes of the oblate massive halos in which they may now be embedded. As said above, I used to regard such tilted geometries as unlikely indeed. However, the recent eloquence of several cosmologists to the effect that halos probably formed first, and galaxies slightly later by crashing gaseous infall, has persuaded me that it is at least conceivable that massive disks deep within halos started out misaligned from their hosts typically by some tens of degrees. And even without logic, desperate men try many things!

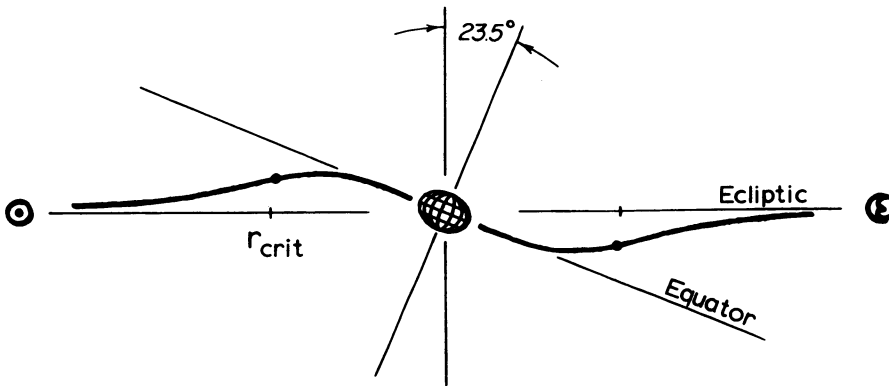


Figure 1. Cross-section through a hypothetical disk around the Earth.

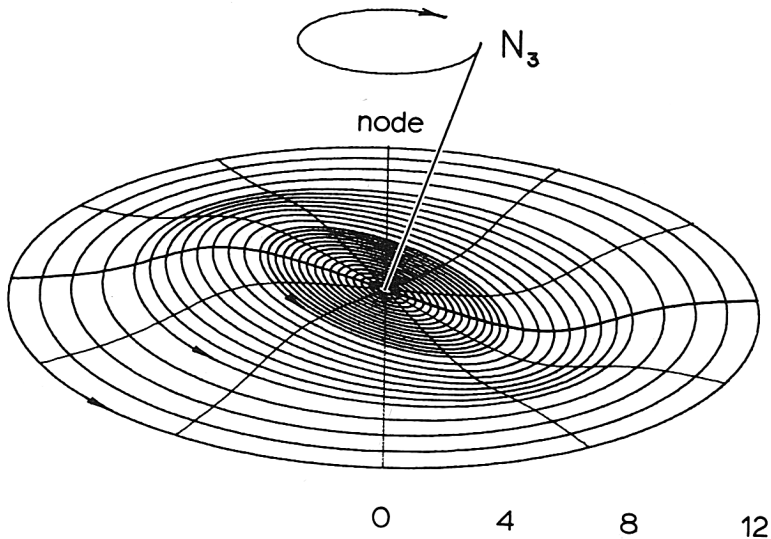


Figure 2. Perspective view of a tilted exponential disk regressing slowly under the tidal influence of a distant massive ring imagined horizontal. The disk slopes 1-in-3 at its center, and its first four scale lengths are shaded darkest. The forcing was taken to consist only of the linear part of the tidal field; its strength matches that of a ring of unit disk mass and of a radius equal to 10 disk scale lengths.

One key question about this kind of proposal is how long the central disk can thereafter remain tilted despite dynamical friction against the doubtless flabby halo. Another serious concern is whether the tidal force from a plausibly massive and flattened halo can compete deeply enough with torques from the disk — i.e., whether the halo can impose a warp already from about 4 scale lengths outward, as in this Galaxy.

In their own paper, Dekel and Shlosman offer some early encouragement on the first topic, whereas for this review I concentrated on the second. I took an exponential disk and divided it into some 50 equally-spaced rings, ranging from quite massive ones in the interior to virtual test-particle rings in the outer half. Next I calculated all the inter-ring torques. Then I exposed this collection to a linear tidal force from that imagined massive distant ring of $(R/10)^3$ times the disk mass, where R is its radius in disk scale lengths — and proceeded to search for the simplest shape-preserving warp as an eigenvalue problem involving the drift rate, which proved to be a tiny -0.018 times $\Omega(3)$. The corresponding warp shape is shown in Figure 2. I think it attests that the idea has some merit — but that is all I can or ought to claim here.

This work was supported in part by grant MCS 78-04888 from the NSF, and of course also by many old discussions with Chris Hunter.

REFERENCES

- Bertin, G., and Casertano, S.: 1982, *Astron.Astrophys.* 106, 274.
 Bertin, G., and Mark, J.W.-K.: 1980, *Astron.Astrophys.* 88, 289.
 Binney, J.: 1978, *M.N.R.A.S.* 183, 779.
 Binney, J.: 1981, *M.N.R.A.S.* 196, 455.
 Hunter, C.: 1969, *Studies in Appl.Math.* 48, 55.
 Hunter, C., and Toomre, A.: 1969, *Astrophys.J.* 155, 747. [= HT]
 Kahn, F.D., and Woltjer, L.: 1959, *Astrophys.J.* 130, 705.
 Lynden-Bell, D.: 1965, *M.N.R.A.S.* 129, 299.
 Petrou, M.: 1980, *M.N.R.A.S.* 191, 767.
 Rogstad, D.H., Lockhart, I.A., and Wright, M.C.H.: 1974, *Astrophys.J.* 193, 309.
 Saar, E.: 1979, in IAU Symposium No. 84, "The Large-Scale Characteristics of the Galaxy" (Dordrecht: Reidel), p.513.
 Sancisi, R.: 1976, *Astron.Astrophys.* 53, 159.
 Shu, F.H., Cuzzi, J.N., and Lissauer, J.J.: 1982, *Icarus*, in press.
 Tubbs, A.D., and Sanders, R.H.: 1979, *Astrophys.J.* 230, 736.

DISCUSSION

BERTIN: From preliminary numerical work, Casertano and I have hints that a D1 mode of the good kind needed for "flapping instability" may survive in models that have a smooth disk density in the outer regions. We hope to confirm this pretty soon.

TOOMRE: Bravo, if you can do it. And also don't forget those discrete, slow $m > 1$ modes found long ago by Hunter.

SANDERS: I would like to stress that the large warps shown yesterday by Dr. Sancisi exist mainly in the tenuous gas beyond the stellar disk, and not in the self-gravitating stellar disk itself. This certainly suggests to me that we may treat the warp as a collection of test particles; there is no need to consider self-gravity in the warps. Moreover, we also see in several cases — for instance, NGC 5907 and 628 — that the gaseous warps begin precisely at the truncated edge of the exponential disk, or where the gravitational potential becomes rapidly spherically symmetric and the differential precession approaches zero. I submit that these external galaxies where we actually observe the light distributions in the disks and the morphology of their warps may be telling us more about warp dynamics than our own galaxy where our picture is more restricted.

TOOMRE: You know I like to peek at observations myself, but I also think any theoretical review worth its salt needs to challenge theorists mostly on their own ground — and sometimes even on their own Galaxy.

SANDERS: I also have a question about the final model you discussed. If there are two preferred planes — a galaxy disk and an outer plane perhaps established by a tilted halo — would you expect the gaseous warp to bend back toward the galaxy plane at larger radii, as it seems to in one or two well-known cases?

TOOMRE: Indeed I wouldn't. If you are thinking of NGC 4762, that sure has looked to me for years like some fairly mature but transient tidal shape. My only real worry has been that its apparent neighbor, NGC 4754, has that hefty excess redshift of around 500 km/s .

SANDAGE: How much mass must the tilted halo have, relative to the disk, for this model of yours and of Dekel and Shlosman to produce the warp?

TOOMRE: Probably at least three times the disk mass. And that's counting only within those ten or so disk scale lengths.

SANDAGE: Are you then not unhappy that the best cases where optical warps are suggested — M33 and NGC 628 — are late Sc galaxies where the optical halo is absent? For M33, I believe that no dark halo is suggested even by the rotation curve.

TOOMRE: Yes, I am unhappy. But as you heard, I am even less happy about most of the alternatives.

NELSON: All the models that you have described to us involve thin stellar disks with no velocity in the z-direction. However, the warps are mainly observed in the gas disks of galaxies, and there the vertical pressure forces are comparable to the gravitational forces. Do you not think that hydrodynamical models for warp dynamics should be considered?

TOOMRE: Sorry about neglecting those models here. Frankly, I could not bring myself to believe that they are desperately relevant, and I had all those other things to discuss.

NELSON: My point is that if you include the gas pressure and the finite thickness, other wave modes appear for which the differential precession can be much smaller (M.N.R.A.S. 177, 265, 1976; 196, 557, 1981).

TREMAINE: In view of the interest in hot thin disks, could you tell us about work you have done on the firehose instability in such systems?

TOOMRE: The main thing I did in 1966 or thereabouts was to calm myself down! I had first thought that such instabilities might set in even with minor anisotropy, but then I spent several weeks calculating them (mostly numerically) via the collisionless Boltzmann equation in a non-rotating $\text{sech}^2 z$ sheet of stars with different Gaussian distributions of random velocities in the vertical and horizontal directions. And I was sorry to find that this hose instability ceases already when the vertical rms speed exceeds a mere 30 per cent of the horizontal rms speed. That is where I lost interest, and why I never even published properly.