SPIRAL STRUCTURE IN GALAXIES: A DENSITY WAVE THEORY BASED ON GLOBAL MODES

G. Bertin
Dipartimento di Fisica
Universita' degli Studi di Milano
Milan, Italy

A paper dedicated to Professor C.C. Lin

The Lin-Shu Symposium: 50 Years of Spiral Density Waves
Beijing, Tsinghua University, June 24-28, 2013

I first met Frank Shu and Ron Allen in Bures-sur-Yvette (with Jan Oort, and many others) in 1974.

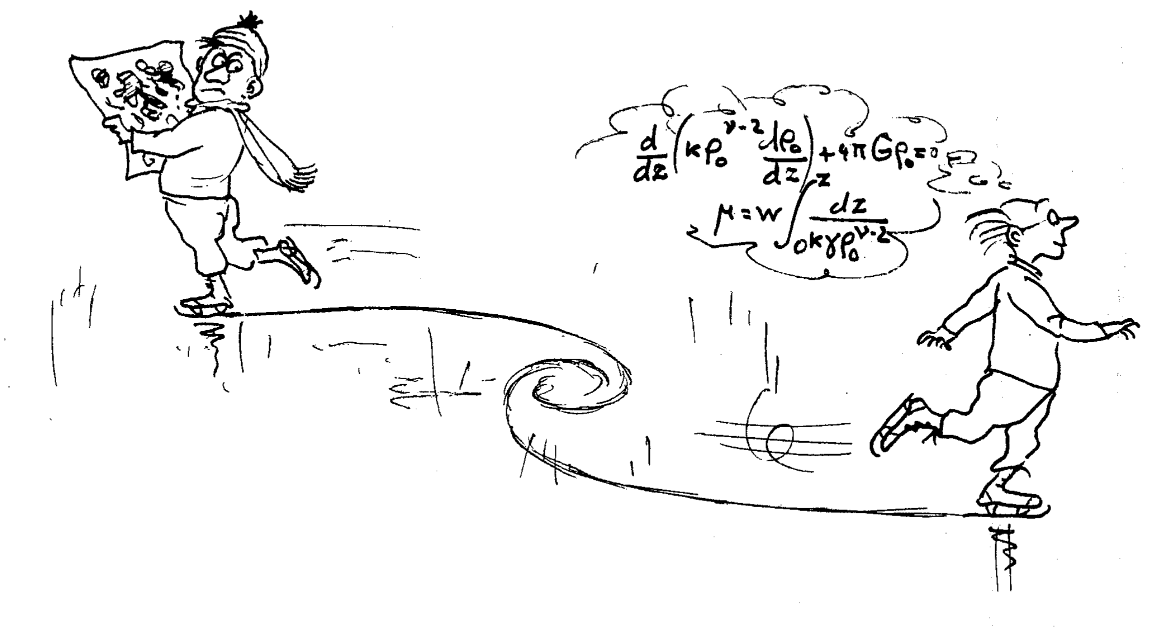


Illustration de J.P. PETIT, participant au Colloque

I first met C.C. Lin at lunch, at the MIT Faculty Club on Monday September 8th, 1975.

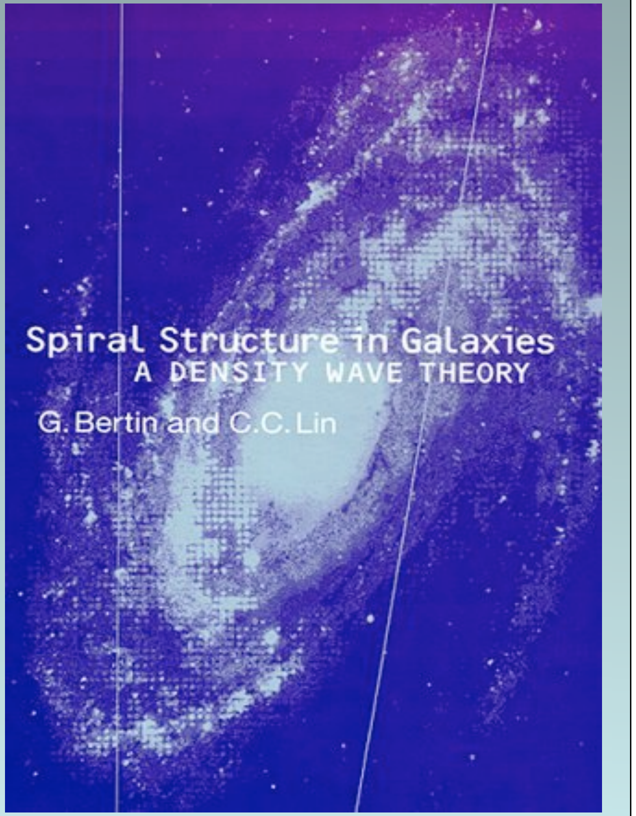
What started then as a postdoc visit turned out to develop into a 20+ year collaboration.

The Density Wave Theory, for the explanation of spiral structure in galaxies (and with following applications to the context of self-gravitating accretion disks), is one of the key achievements in Astrophysics.

A semi-empirical approach based on the working hypothesis of Quasi-Stationary Spiral Structure in galaxies has originated an impressive number of quantitative successful observational tests that have attracted the interest of the astronomical community in the last five decades.

The development of a successful and internally consistent theory has required the solution of a number of challenging conceptual problems at the frontier of Astrophysics and of Applied Mathematics.

1964 - 1966	Density Wave Theory just born: Lin&Shu
1976	Key steps toward a theory of (self-excited) global modes
1986	CCL's 70th birthday. Symposium at MIT (1987). Key steps toward a unified theory of normal and barred spiral structure
1996	Publication of the Monograph (MIT Press)
2006	CCL's 90th birthday. Symposium in Beijing
2013	Lin-Shu Symposium in Beijing



Spiral Structure in Galaxies
A DENSITY WAVE THEORY
G. Bertin and C.C. Lin

OUTLINE

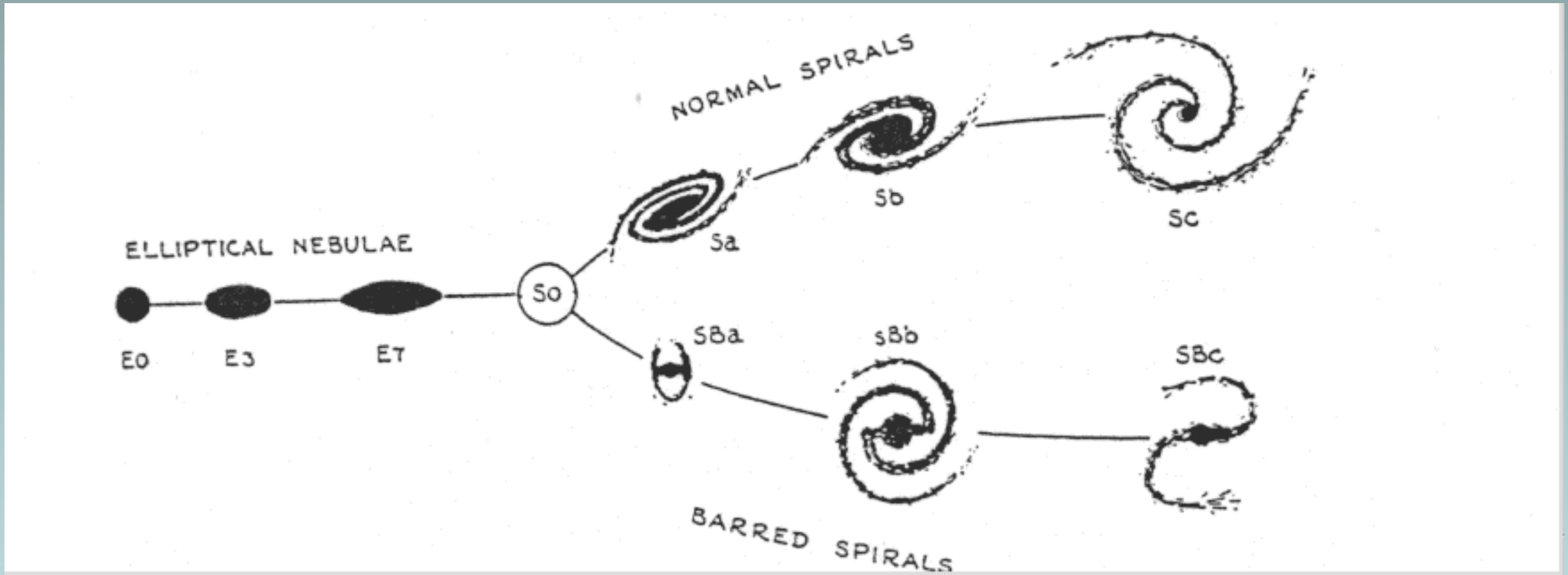
PART I – Morphology of spiral galaxies, the problem of spiral structure, QSSS as a working hypothesis
(~ the first fifteen years....)

PART II – Theory of spiral structure in galaxies
(~ the following fifteen years....)

PART III – Recent results: light disks, prominent spiral arms in the outer gaseous disks
(~ the last fifteen years....)

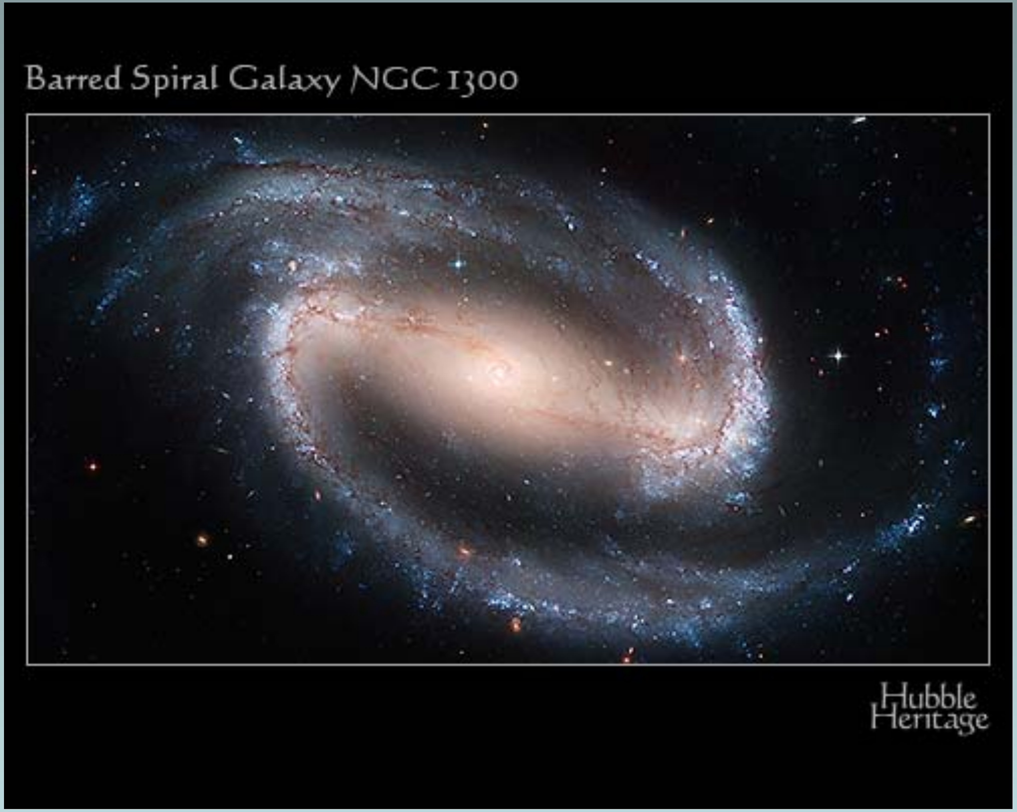
PART I. MORPHOLOGY OF SPIRAL GALAXIES, THE PROBLEM OF SPIRAL STRUCTURE, QSSS AS A WORKING HYPOTHESIS

Spiral structure and bars in galaxies.
Origin of the Hubble morphological sequence?

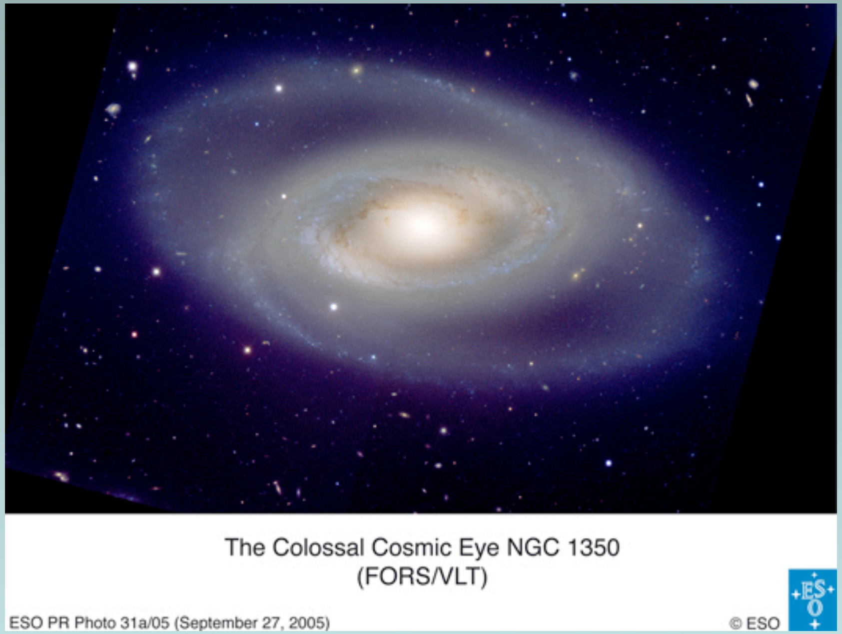


Hubble 1926

Barred spirals and Normal spirals



M81 (NGC 3031)
grand design

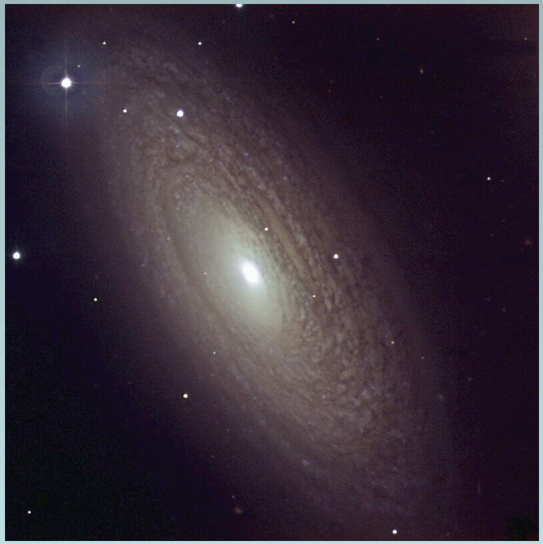


The Colossal Cosmic Eye NGC 1350
(FORS/VLT)

ESO PR Photo 31a/05 (September 27, 2005)

© ESO

NGC 2841
flocculent



M74
(NGC 628)
Sc

NGC 5364 Sc

NGC 2403 Sc

M100
(NGC 4321)
Sc

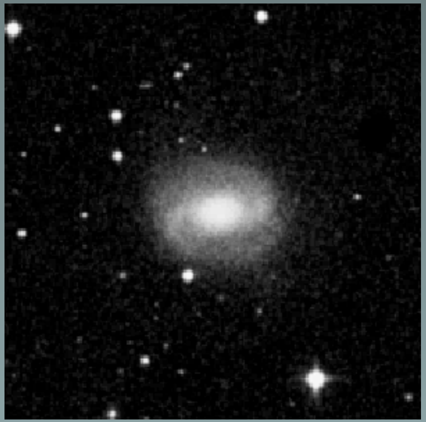
NGC 2997 Sc

NGC 309 Sc

Multiple arms
(especially in the visible)

NGC 2859 SB0

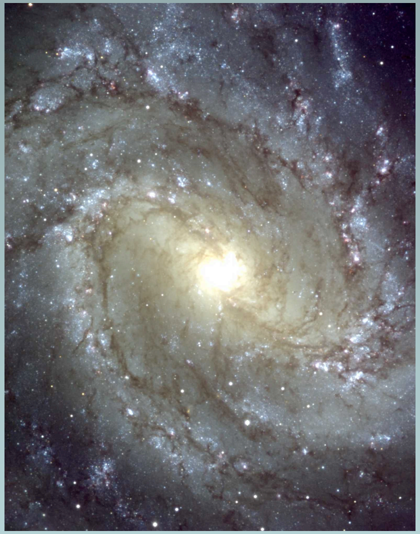
NGC 7743 SBa



NGC 5383 SBb

NGC 1398 SBb

NGC 1097 SBb



M83 (NGC 5236)
SBc

M51
(NGC 5194 + NGC 5195)

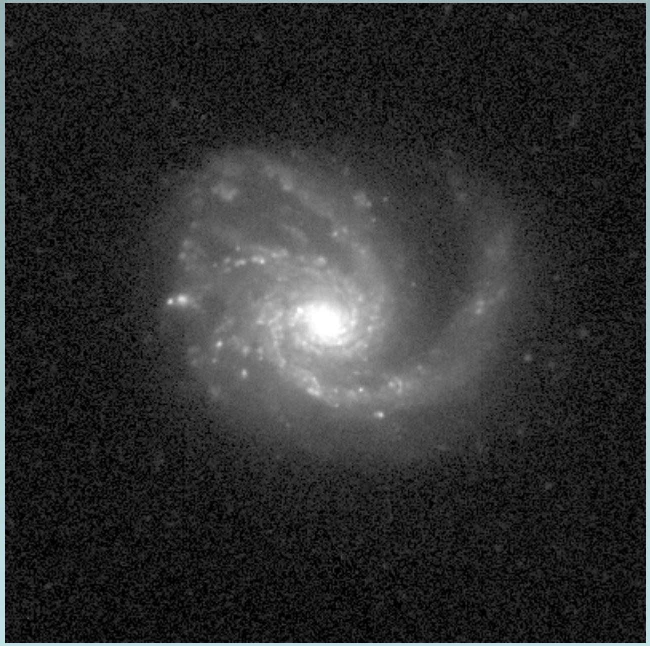
Tidal interaction?

NGC 1637 Sc

lopsidedness

M101 (NGC 5457) Sc

M99 (NGC 4254) Sc



ESO 510-G13

NGC 5907

Edge-on views

NGC 4565

NGC 891

KINEMATICS, DIFFERENTIAL ROTATION, EPICYCLES

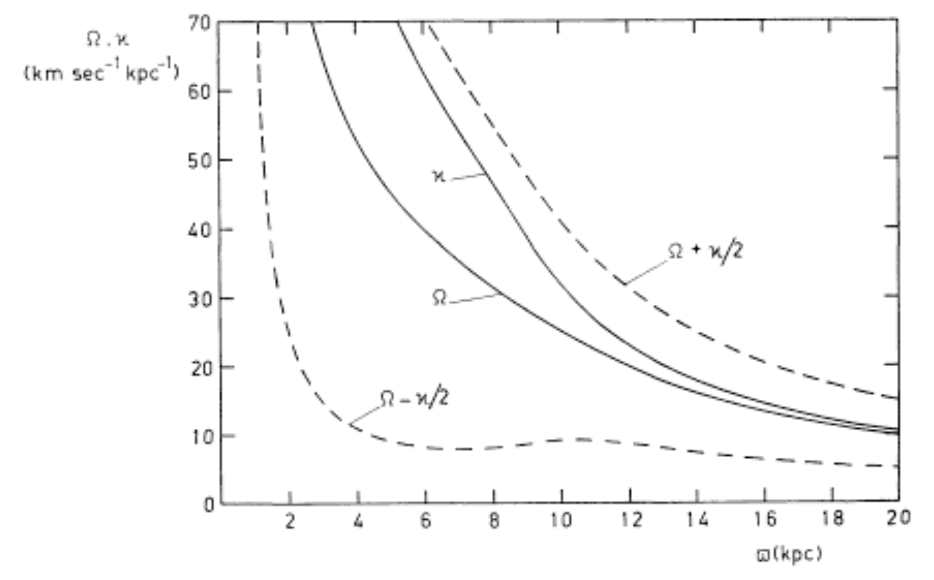
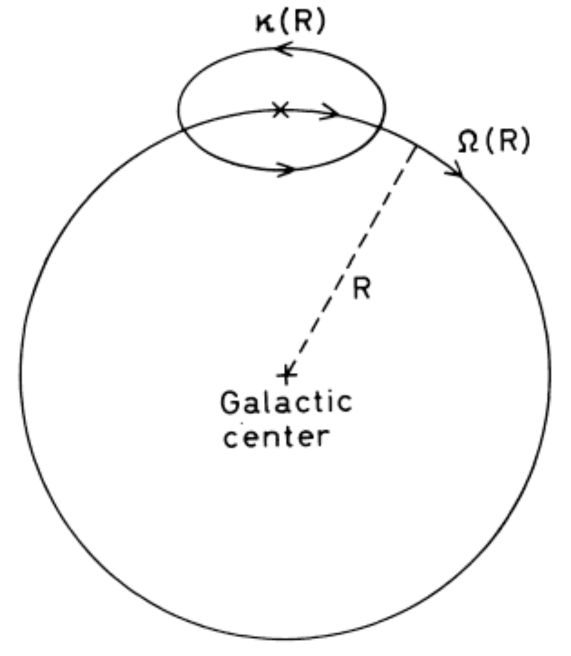


FIG. 1. Angular velocity Ω, epicyclic frequency κ, and related quantities for the 1965 Schmidt model of the Galaxy.



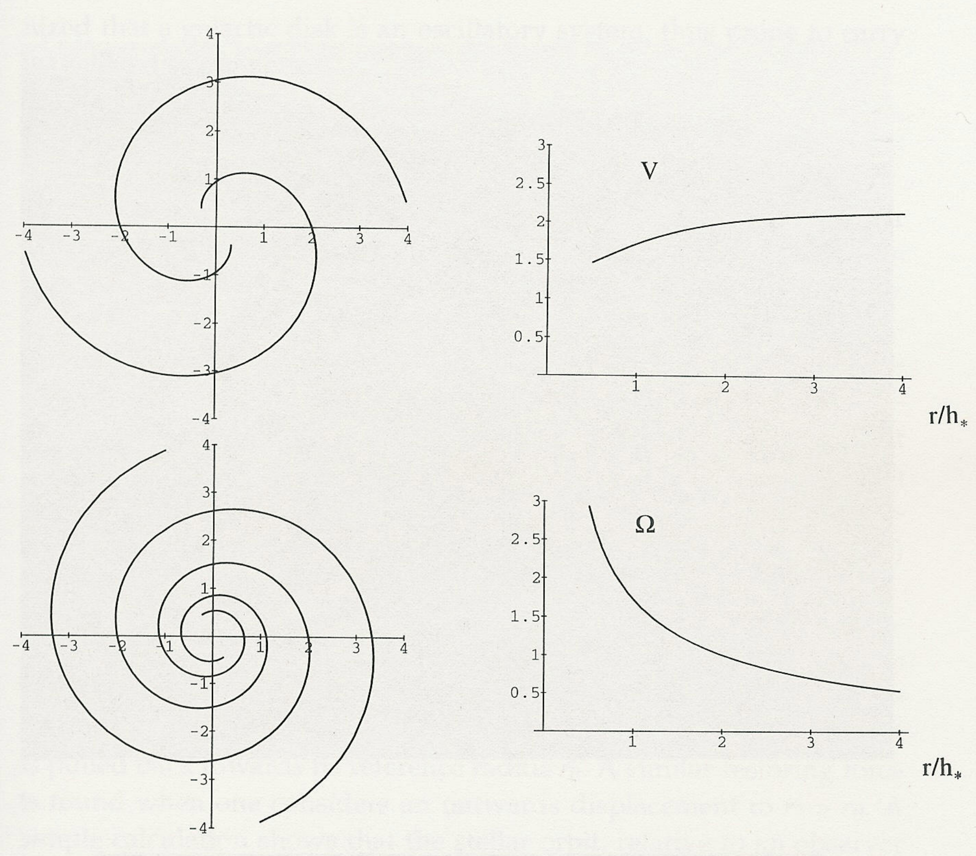
$$\kappa^2 = 4\Omega^2\left(1+\frac{1}{2}\frac{d\ln\Omega}{d\ln r}\right); \Omega^2 = \frac{1}{r}\frac{d\Phi}{dr}$$

κ epicyclic frequency
Ω differential rotation

(note that, in the 1960s, dark halos did not exist!)

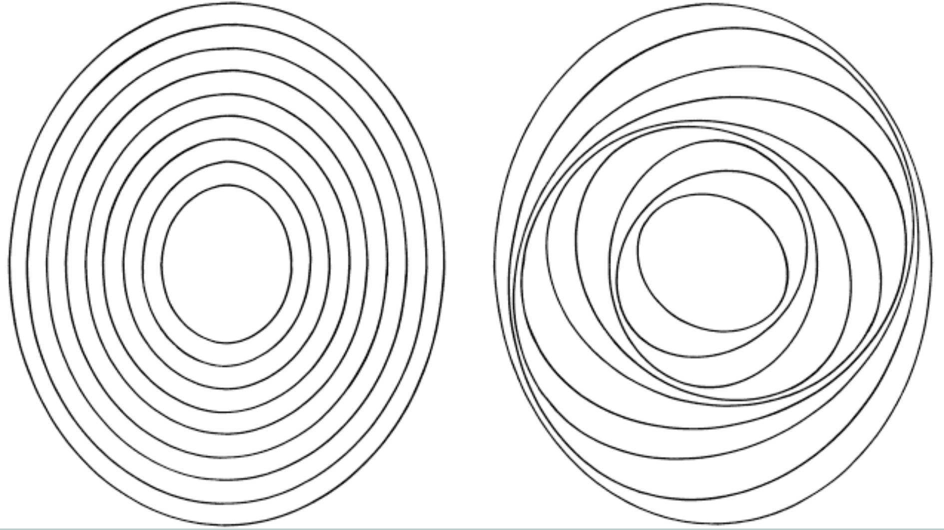
THE WINDING DILEMMA

Differential rotation would rapidly (in a couple of turns) stretch any material arm into a tightly wound spiral structure:
How can we reconcile this fact with the observations of so many galaxies with open arms?



LINDBLAD’S KINEMATIC WAVES

B. Lindblad’s kinematic waves: a precursor of the density wave theory



A set of closed orbits seen in a frame rotating at $\Omega_p = \Omega - \kappa/2$

THE PROBLEM OF SPIRAL STRUCTURE (Oort, 1961 conference at IAS, Princeton, attended by CCL)

Focus on grand design, that is, regular, large-scale spiral structure:

- Primarily gaseous or primarily stellar arms? (observations in different wavebands will tell us; in fact, a decisive confirmation came from K-band data starting with the 1990s)

- How did spiral structure originate?
- How does it persist once it has originated?

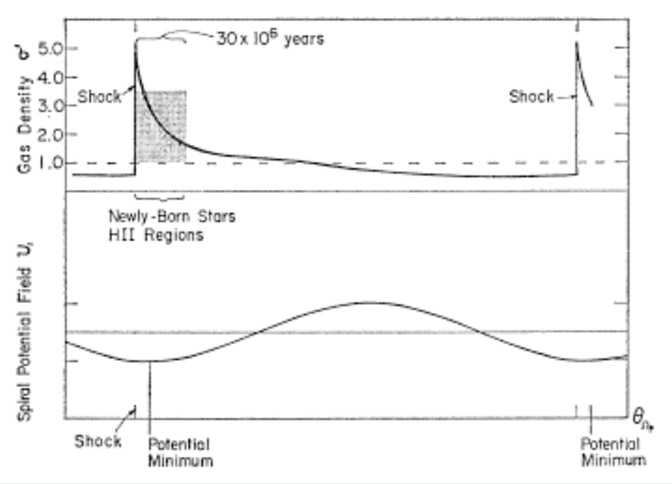
THE PROBLEM OF SPIRAL STRUCTURE

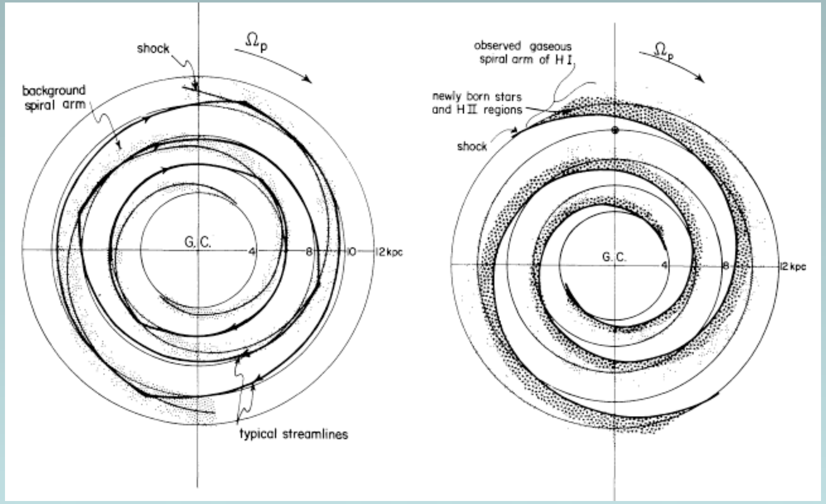
In relation to the large scale:

- **Why certain spirals are barred and others are not?**
- **How do we explain the different degrees of regularity in the observed spiral structure?**
- **How do we explain flocculent spiral structure?**
- **Why trailing structure?**
- **Why is grand design generally two-armed?**
- **Why do we often see different coexisting morphologies?**
- **How do we explain the Hubble morphological classification?**
- **What sets the amplitude of the observed structure?**

QUASI-STATIONARY SPIRAL STRUCTURE AS A WORKING HYPOTHESIS: THE SHOCK-WAVE PATTERN

Because the rotation of the disk is differential, a quasi-stationary pattern will be moving supersonically with respect to the interstellar medium over most of the galaxy disk. If its amplitude is large enough, shocks will be generated and these, in turn, will favor star formation.





W.W. Roberts 1969

Visser's study of M81 (PhD Thesis 1978)

From the properties of the pattern observed in the optical and fitted by a (short) density wave rotating at a suitable pattern speed Ω_p, the kinematics of the shocked HI gas can be computed and compared in detail with the observed motions measured by 21 cm radio observations.

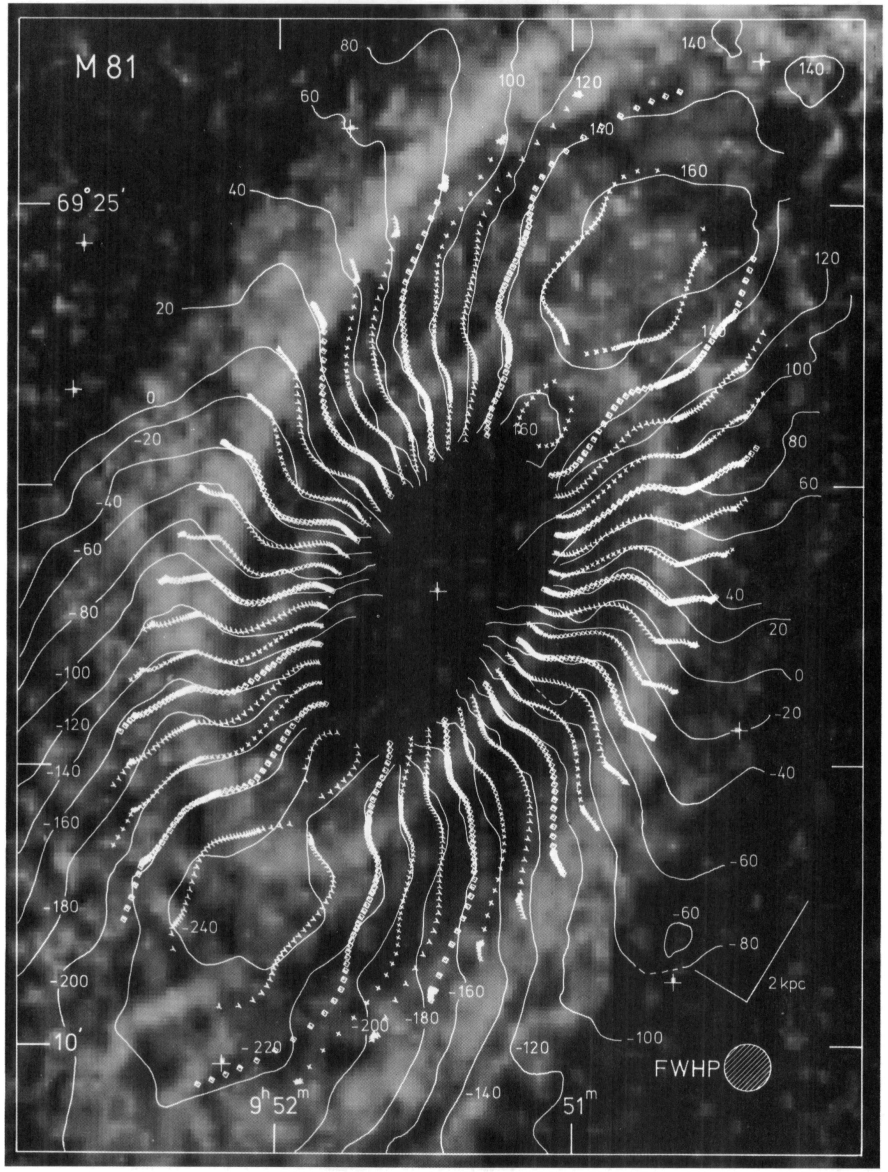
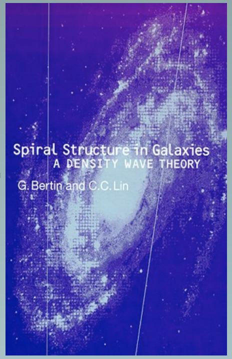


Fig. 5. The radial-velocity field of the final model (symbols) together with the observed velocity field (full and dashed lines) at an angular resolution of 50″, superimposed on a radiograph of the density distribution of the atomic hydrogen at 25″ resolution. See also the caption of Fig. 4

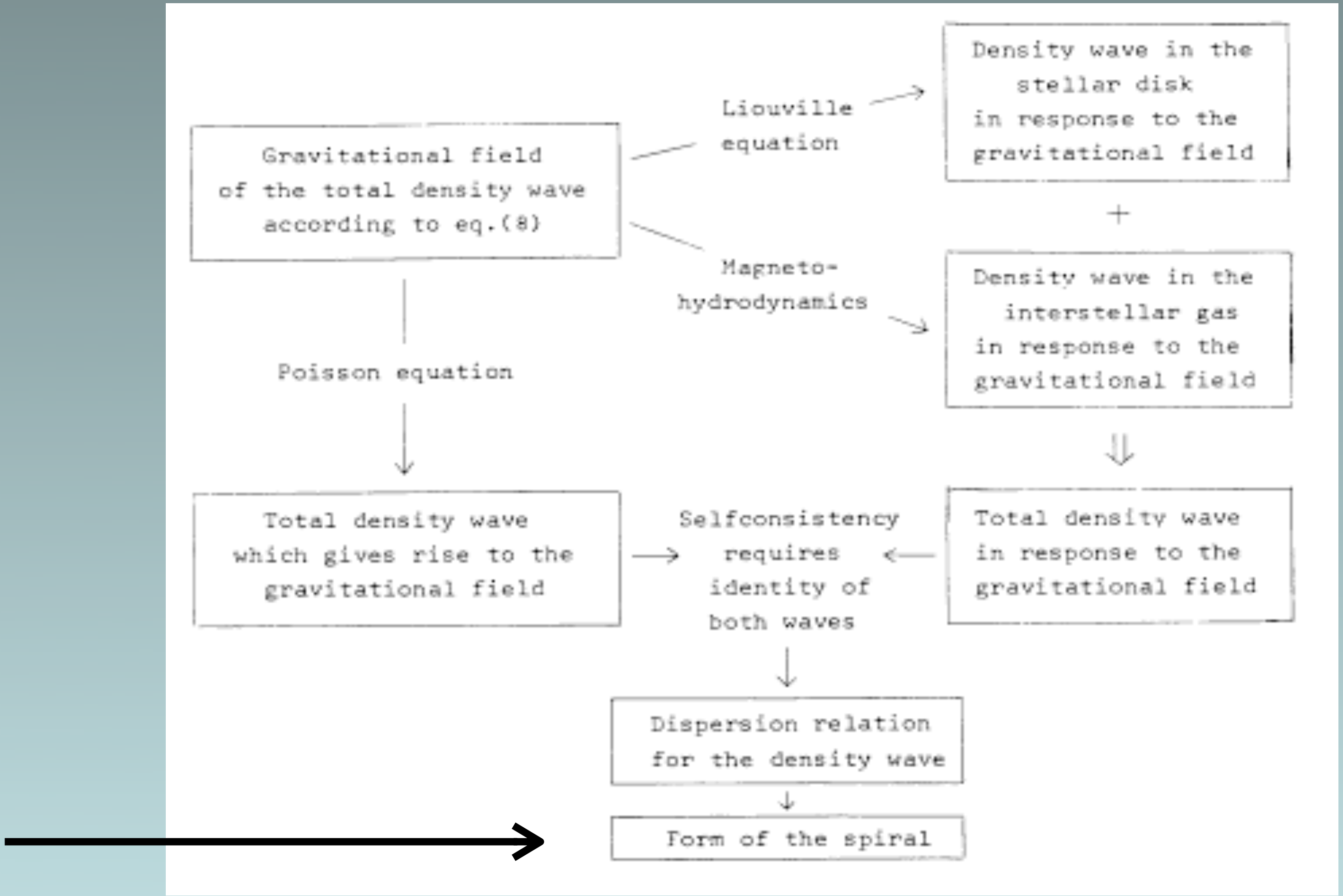
Visser 1980

PART II – THEORY OF SPIRAL STRUCTURE IN GALAXIES

FOUR DYNAMICAL SCENARIOS

	Quasi-stationary	**Rapidly evolving**
Internal origin	Discrete spectrum; one or few self-excited modes 	Continuous spectrum, regenerative spiral structure **P.O. Lindblad 1960; Goldreich & Lynden-Bell 1965**
External origin	Discrete spectrum, damped modes. Orbiting satellite?	Continuous spectrum. "One-shot" tidal interactions **("swing") A. Toomre 1980**

DENSITY WAVES



C.C. Lin & F.H. Shu 1964, 1966

DENSITY WAVES IN A THIN FLUID DISK

$$\sigma_1 = \bar{\sigma} \exp[i(\omega t - m\vartheta + \Psi(r))]$$

$$k(r) \equiv \frac{d\Psi(r)}{dr}; |rk| >> 1, m = O(1)$$

Dispersion Relation:

$$(\omega - m\Omega)^2 = \kappa^2 + c^2 k^2 - 2\pi G\sigma \, |k|$$

Doppler-shifted frequency

Angular momentum conservation

pressure

Jeans

LOCAL STABILITY

Dispersion relation studied as $\omega = \omega$ (k)

$$Q = \frac{c\kappa}{\pi G \sigma} \geq 1$$

Condition for local stability against axisymmetric perturbations

For Q < 1, there is a range of wavelengths for which density waves are unstable.

SEMI-EMPIRICAL APPROACH

Dispersion relation studied as k = k(r;m,Ω_p)

m	**Number of arms**
$\omega = m\Omega_p$	**Pattern frequency**

By assigning these quantities, under the assumption that spiral structure is quasi-stationary, the dispersion relation is used to determine the pitch-angle of (to "draw") spiral arms:

$$\tan i = \frac{m}{rk}$$

Short-wave-branch. Empirically, corotation in the outer disk

GROUP PROPAGATION?

$$c_g = -\frac{\partial \omega}{\partial k}$$

Short waves (used for application to observations) bound to disappear quickly, propagating inwards, in the direction of the galaxy center………

A. Toomre 1969; a "feedback" would resolve the problem C.C. Lin 1970

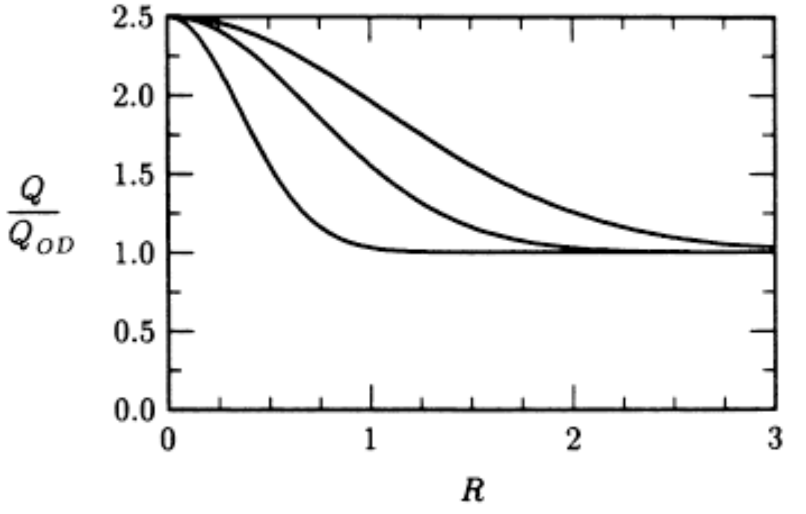
GAS, STARS, SELF-REGULATION

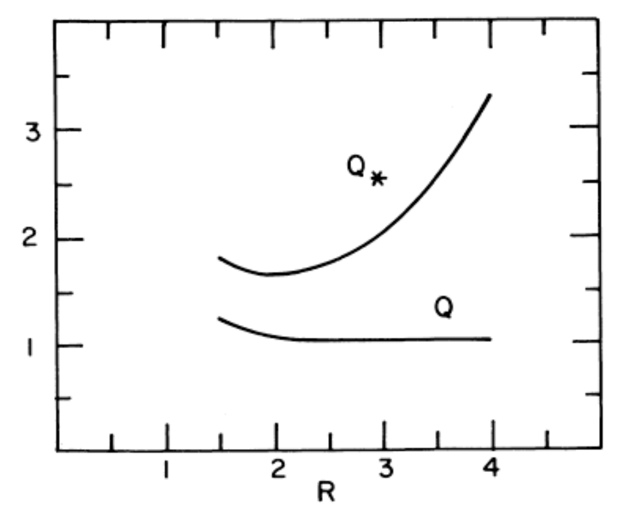
Gas is a <u>cold dissipative</u> component

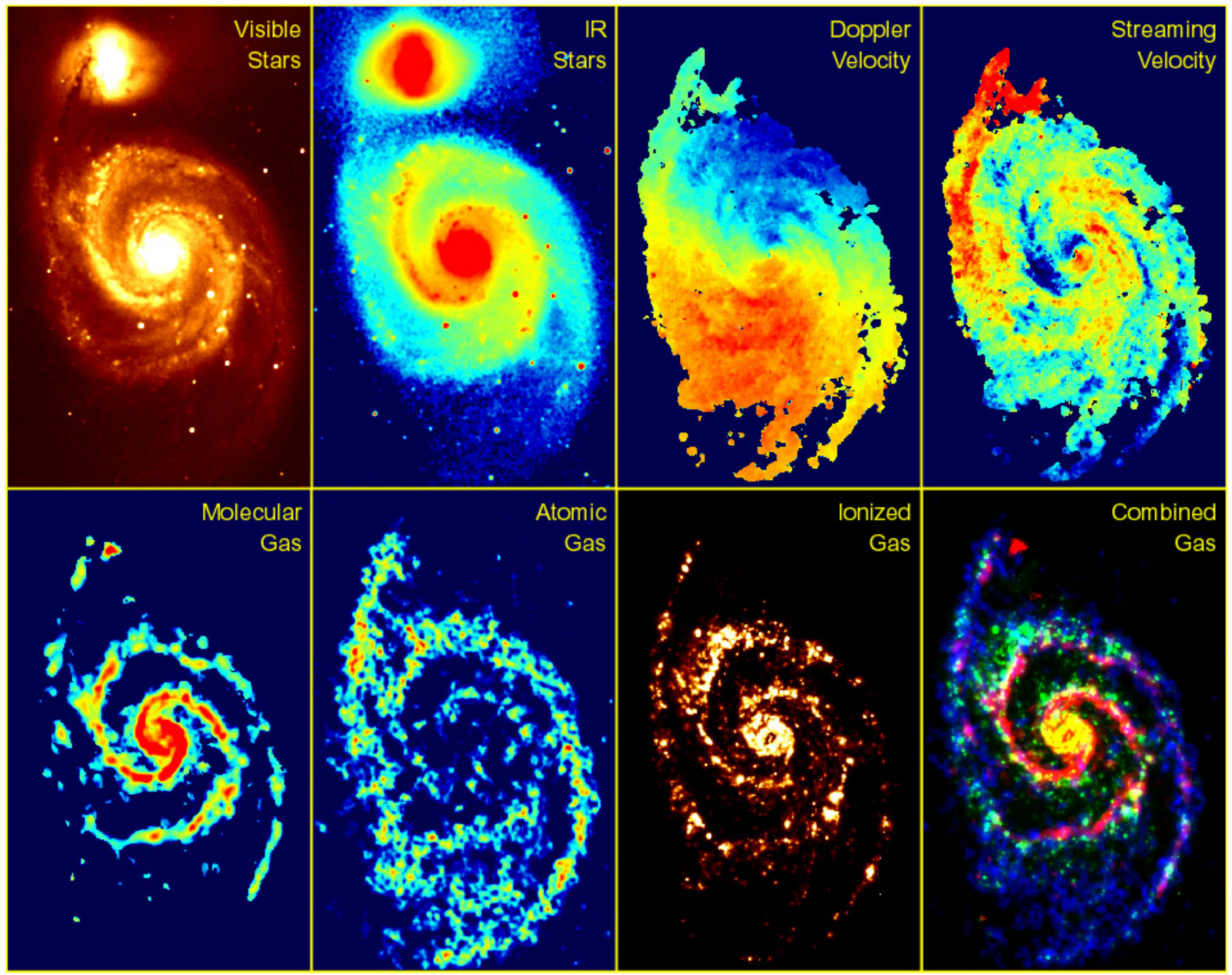
Gas is <u>diffuse</u>, subject to <u>small-scale</u> spiral activity

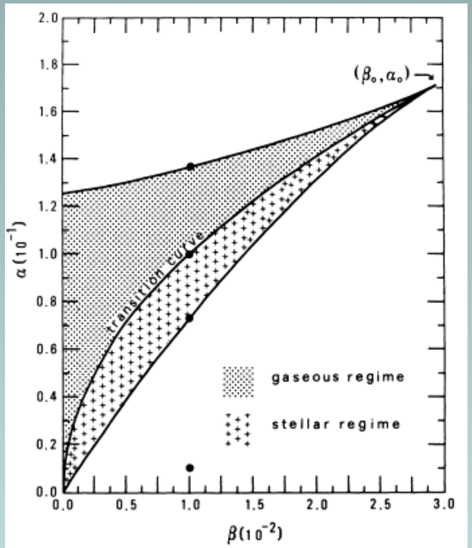
A small amount of gas makes the effective Q much smaller than Q$_*$, <u>fueling Jeans instability</u>

A sufficient amount of gas enforces a "<u>thermostat</u>" (Q close to marginal stability, <u>self-regulation</u>)







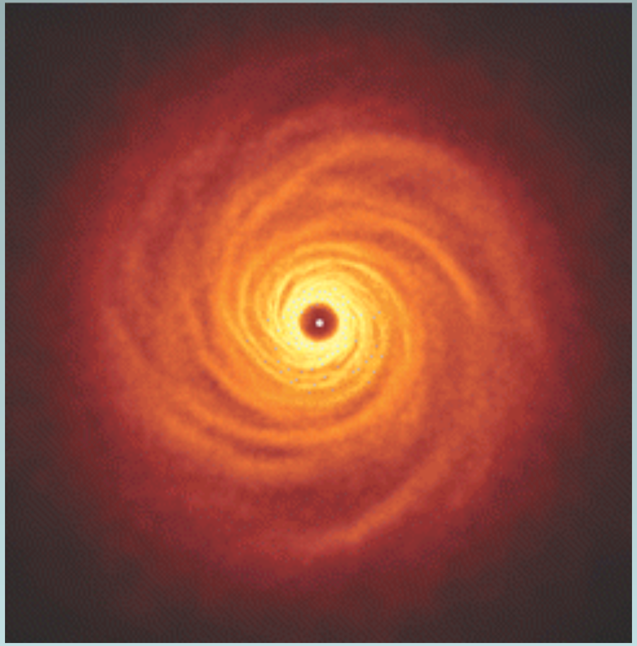


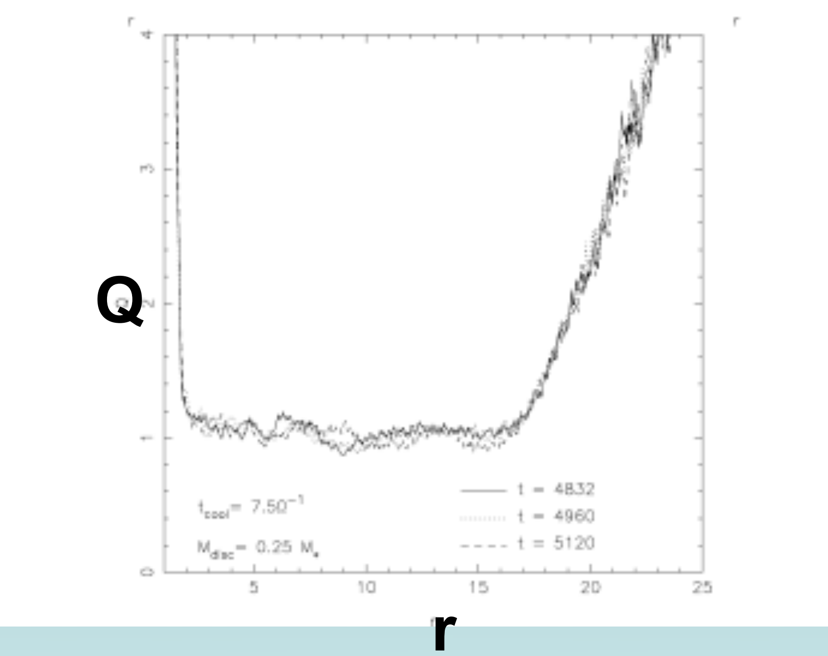
Bertin, Romeo A&A 1988

Credits:
CO: BIMA array; S. Vogel, T. Helfer, and BIMA SONG team
Halpha: Palomar Maryland-Caltech Fabry-Perot: S. Vogel, R. Gruendl, R. Rand
IR: Kitt Peak: R. Gruendl, S. Vogel
HI: VLA: A. Rots

Self-regulation for a moderately heavy disk $M_{disk} = M_{star}/4$; SPH simulations of a protostellar self-gravitating disk

(Lodato & Rice 2004)





ORIGIN OF GLOBAL SPIRAL STRUCTURE

Feedback (density waves refracted back from the galaxy central regions; "Q-barrier")

+

Overreflection at corotation

=

Self-excited global modes, i.e. "standing waves" (no radial propagation!)

$$\sigma_1 = \overline{\sigma}(r)\exp[i(\omega t - m\vartheta)]$$

eigenfunction eigenfrequency

OVERREFLECTION AND PROPAGATION DIAGRAMS

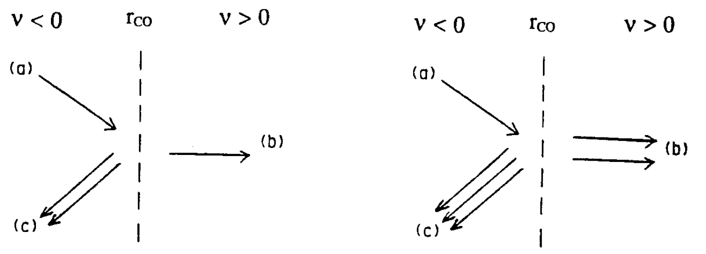


Figure 10.2
Process of overreflection illustrated for two cases of different levels of local stability at corotation. *Left:* Case of local marginal stability (see figure 10.1 or left frames of figures 9.1 and 10.3). *Right:* Case of a system locally unstable at corotation (see middle frames of figures 9.1 and 10.3); for a description, see also the caption to figure 4.1).

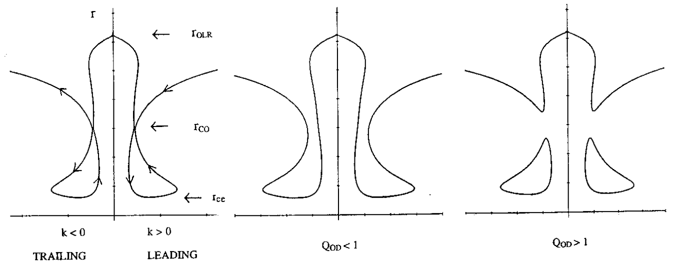


Figure 9.1
Propagation diagrams for the quadratic dispersion relation. The left frame illustrates the properties of the diagram in detail for a model that might support a tightly wound spiral mode. The radial coordinate is on the vertical axis, starting with $r = 0$. The radial wave number k is on the horizontal axis. Note the properties of the wave branches in the vicinity of the marked locations (r_{co}, r_{OLR}, and r_{ce}). Arrows denote direction of group propagation along each wave branch. The middle and right frames illustrate the qualitative change induced in the diagram by changing the value of Q_{OD} below or above the "marginal" value of 1.

Waves have <u>negative energy density inside corotation</u>; <u>trailing</u> modes are excited by carrying angular momentum outwards.
Normal (Type I) overreflection:
long into stronger short wave
J. W-K. Mark 1974

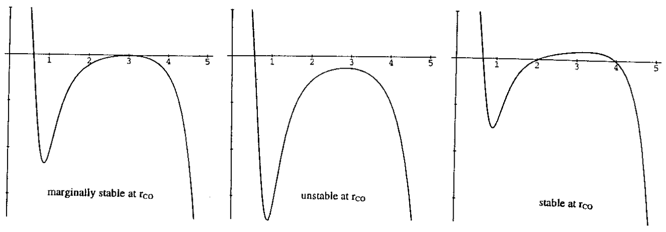
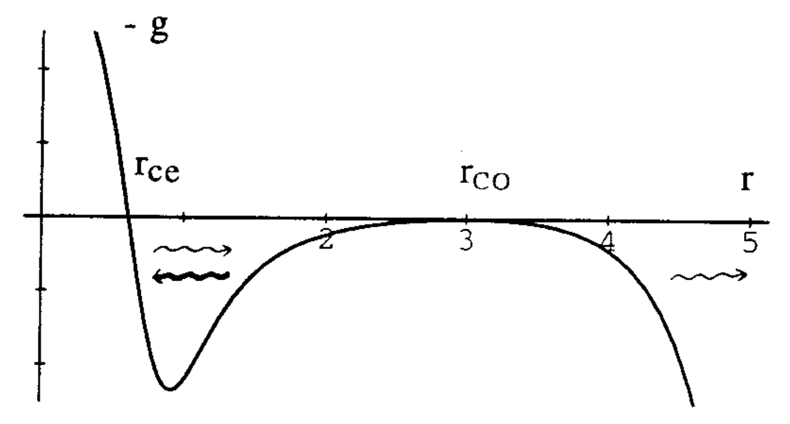


Figure 10.3
Function $(-g)$ can be seen like the potential for a particle with zero energy. The plots given here actually reproduce the function indicated in eq. (10.2) (regime A) for the same models used to draw the propagation diagrams in figure 9.1, following the same sequence from left to right; there is a small difference, in that here the departures from unity of Q_{OD} have been slightly exaggerated, in order to better appreciate the changes induced by modifying the conditions of local stability at corotation ($r_{co} = 3$, as in figure 10.1). This figure can also be used to describe possible cases related to regime B (section 10.2.2).

A TWO-TURNING POINT PROBLEM (a Schroedinger-type equation)

$$\frac{d^2u}{dr^2} + g(r;\omega)u = 0$$



Boundary conditions:

Outgoing wave (at large radii)
Evanescent wave (at small radii)

Simple turning point at r_{ce} (~ bulge radius)
Double turning point at r_{co} (corotation)

feedback overreflection

(for normal spiral modes Lau, Lin, Mark 1976; Bertin, Lau, Lin, Mark, Sugiyama 1977)

“RESONANT CAVITY”

Discrete spectrum, Bohr-Sommerfeld
“global Dispersion Relation”

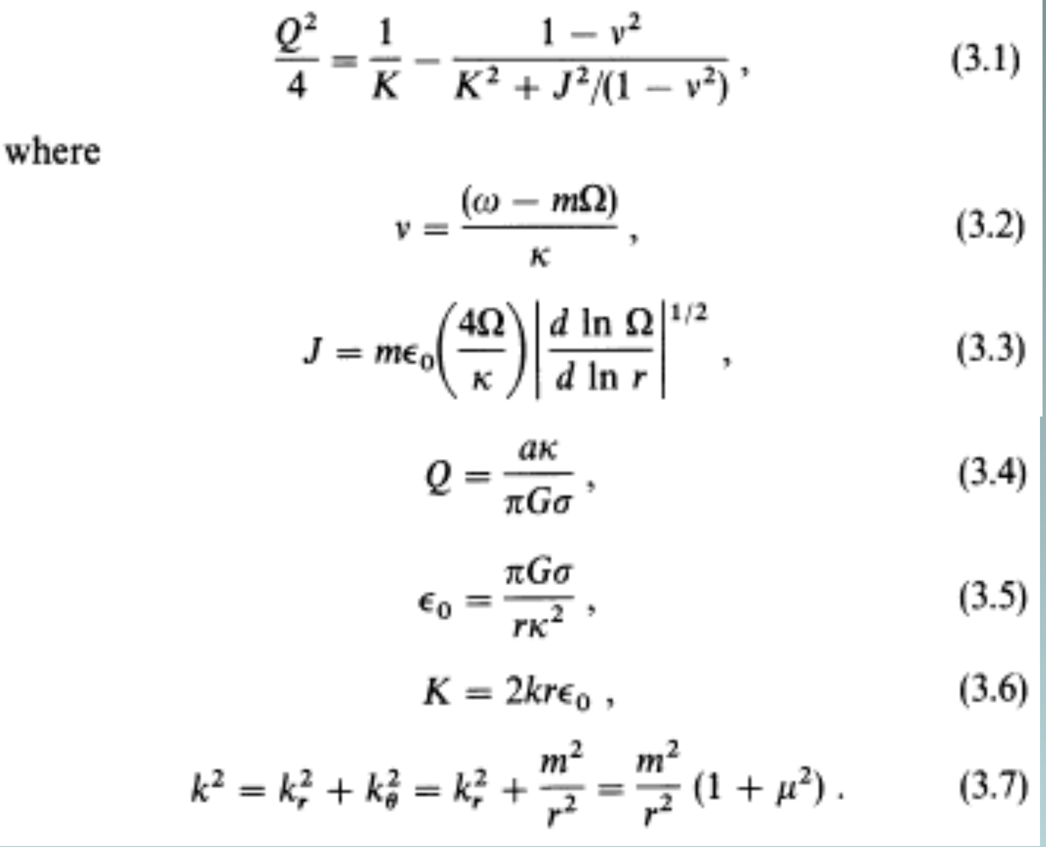
$$\oint k(r;\omega)dr = (2n+1)\pi + i\ln\sqrt{2}$$

$$\oint k(r;\omega_R)dr = (2n+1)\pi$$

$$\gamma\oint\frac{dr}{|\partial\omega/\partial k|} = \gamma\tau = \ln\sqrt{2}$$

CUBIC DISPERSION RELATION:

a unified theory of normal and barred spirals



$$\frac{Q^2}{4} = \frac{1}{K} - \frac{1-\nu^2}{K^2 + J^2/(1-\nu^2)}\,, \tag{3.1}$$

where

$$\nu = \frac{(\omega - m\Omega)}{\kappa}\,, \tag{3.2}$$

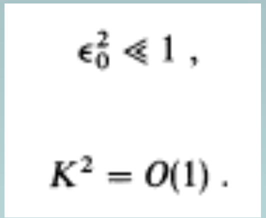
$$J = m\epsilon_0\left(\frac{4\Omega}{\kappa}\right)\left|\frac{d\ln\Omega}{d\ln r}\right|^{1/2}\,, \tag{3.3}$$

$$Q = \frac{a\kappa}{\pi G\sigma}\,, \tag{3.4}$$

$$\epsilon_0 = \frac{\pi G\sigma}{r\kappa^2}\,, \tag{3.5}$$

$$K = 2kr\epsilon_0\,, \tag{3.6}$$

$$k^2 = k_r^2 + k_\theta^2 = k_r^2 + \frac{m^2}{r^2} = \frac{m^2}{r^2}(1+\mu^2)\,. \tag{3.7}$$



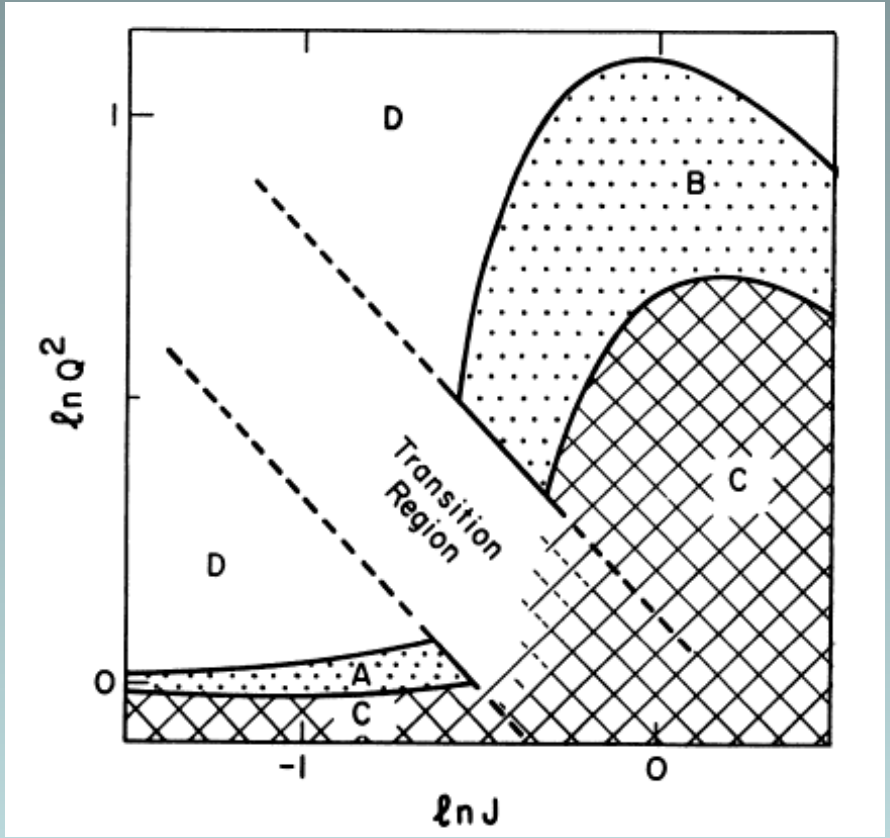
$$\epsilon_0^2 \ll 1\,,$$

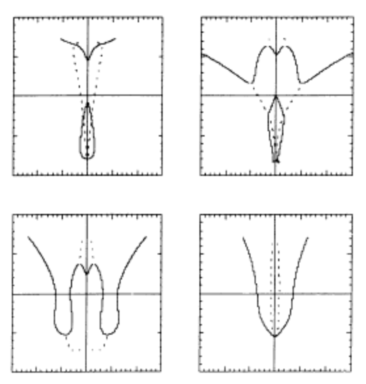
$$K^2 = O(1)\,.$$

Bertin 1983; Bertin, Lin, Lowe 1984; Bertin, Lin, Lowe, Thurstans 1989

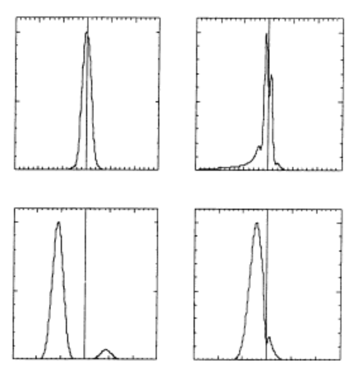
NORMAL AND BARRED SPIRAL MODES

hot disk





heavy disk



In region A, Type I overreflection, of <u>long into stronger short trailing</u> wave

In region B, Type II overreflection, of <u>leading into stronger trailing</u> wave

MODE PROTOTYPES

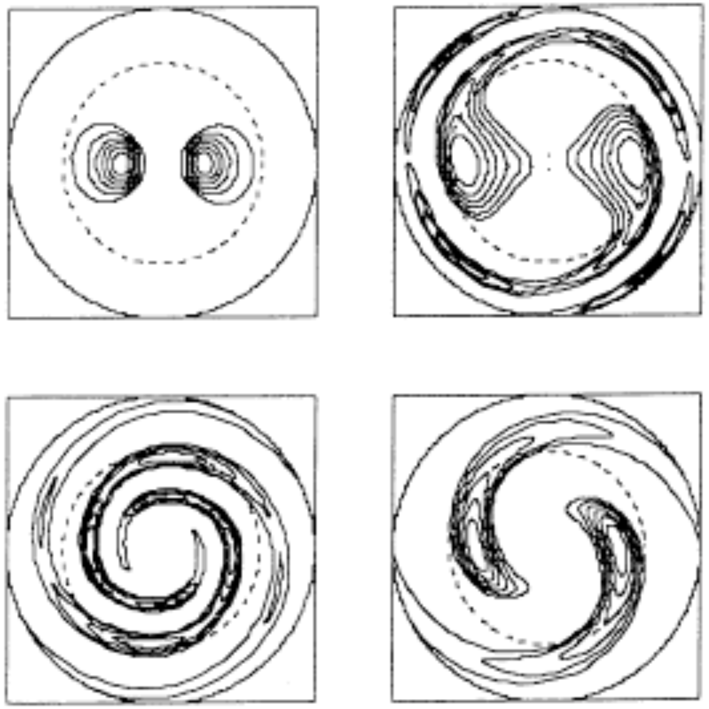
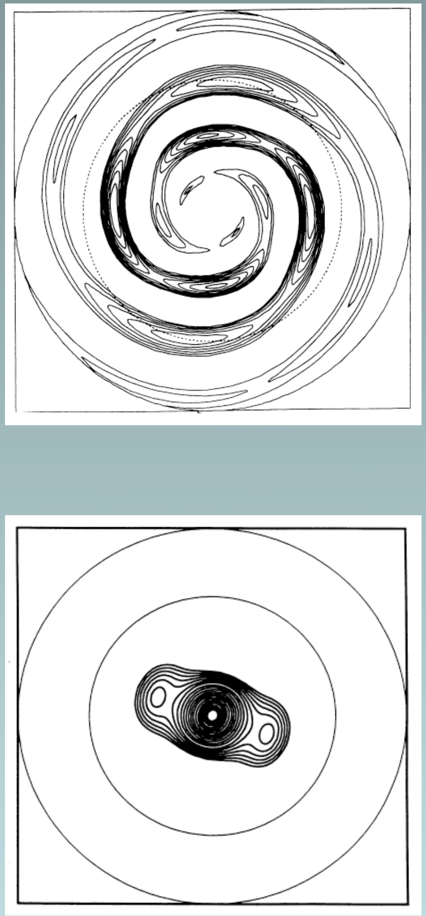


FIG. 11.—Mode prototypes. Four key morphological types are compared: SB0, SB(s), and S, all with moderate growth; a violently unstable S mode at the low right corner. The dynamical properties of these modes are discussed in Paper II.



ILR INHIBITS HIGHER MODES

The Inner Lindblad Resonance
in the stellar disk absorbs incoming waves completely (Mark 1971); this interrupts the cycle at the basis of the resonant cavity for modes with higher m and higher n.
Only few modes, with m = 1, 2 are viable.

For the gas, the process is far less efficient, and multiple armed structure can be generated.

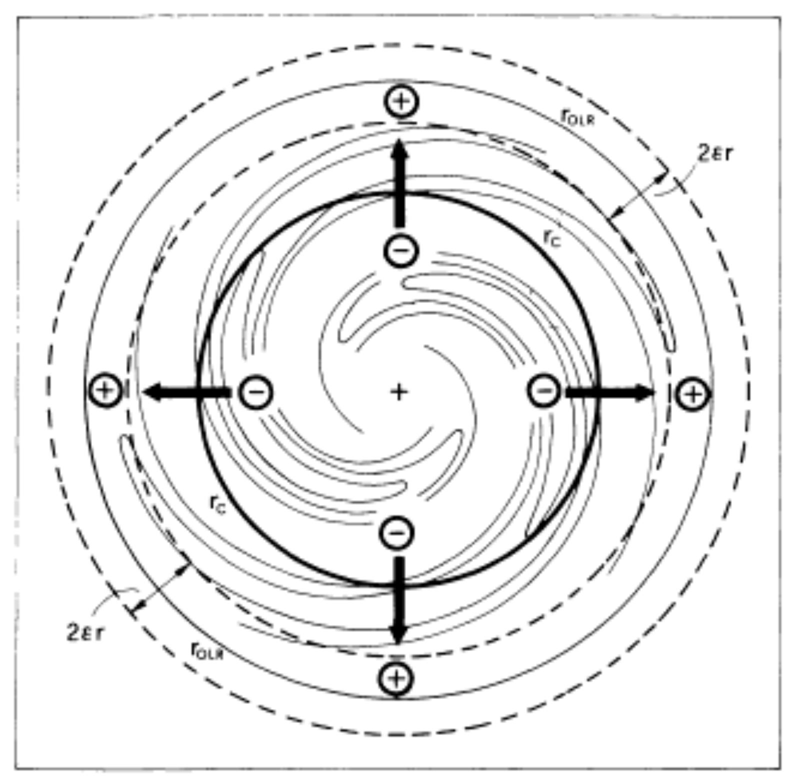
NON-LINEAR SATURATION THROUGH SHOCKS

(Kalnajs 1972; Roberts & Shu 1972)

There is an interesting problem of gas consumption if we wish to argue that the mechanism of self-regulation is long-lasting.

ANGULAR MOMENTUM TRANSPORT

The outwards transport of angular momentum, which is at the basis of the excitation of spiral structure, will lead to a secular evolution of the galaxy disk.



Bertin 1983

BARRED MODES, AMPLITUDE MODULATION, COROTATION RADIUS

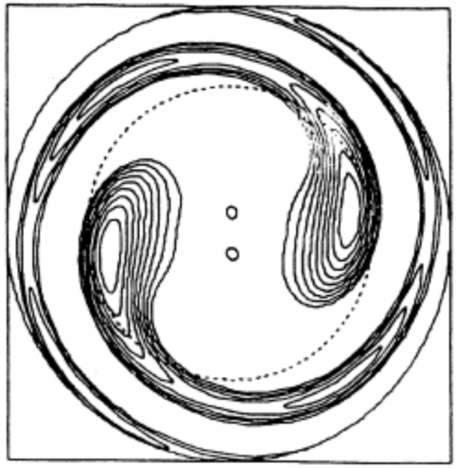


FIG. 1——Positive density contours for a barred mode from the survey reported by Bertin et al. (1989a). The structure of the mode (pitch angle, gaps, and arm shape) closely resembles the barred spiral structure found in NGC 1300 [compare with Fig. 1(f) of Elmegreen et al. 1992a]. The dotted circle is the corotation circle.

THE MORPHOLOGY CLASSIFICATION

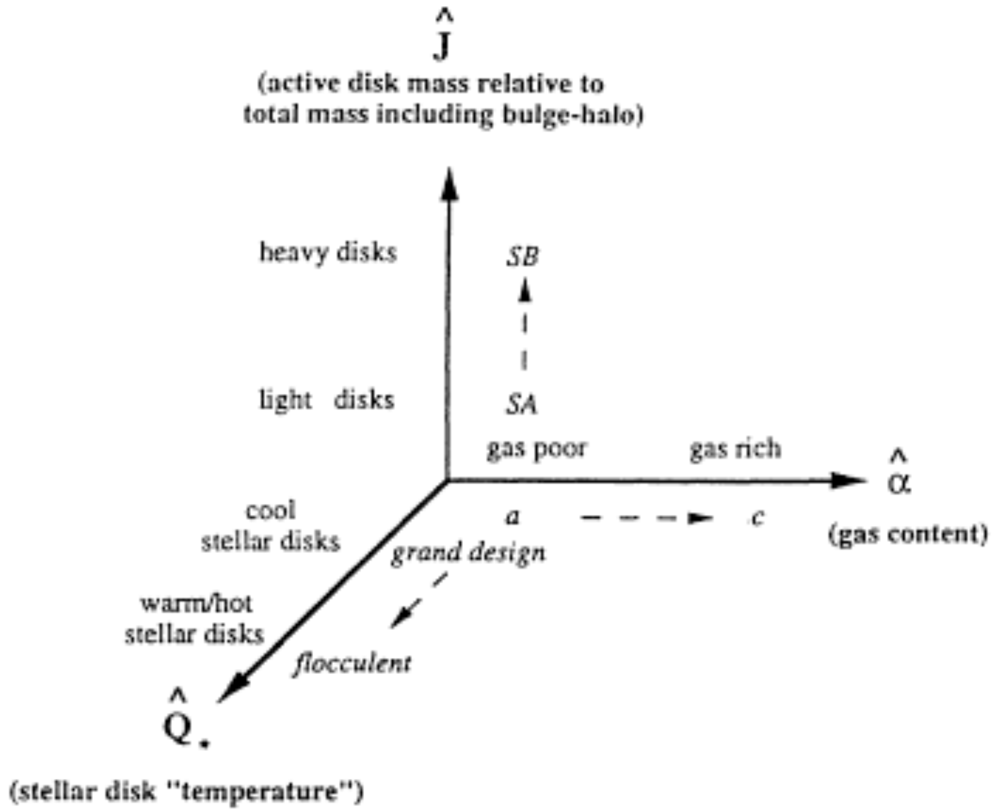


Fig. 1. Framework for the classification of the morphologies of spiral galaxies on the basis of their intrinsic modal characteristics (see Bertin 1991)

K-BAND OBSERVATIONS*: THE DECISIVE “PROOF”

* In the near-infrared, at 2.1μ, they let us “see” the underlying stellar disk; became available starting in the early 1990s

- Large scale spiral arms are a density wave in the stellar disk
- Grand design is very frequent and generally two-armed
- Even when its amplitude is large, the density perturbation is smooth and sinusoidal, thus the theory of linear modes can be applied to the observed morphologies.
- Multiple-armed spiral structure is mostly a Population I (“gas”) phenomenon.

Block & Wainscoat 1991; Zaritsky, Rix & Rieke 1993; Block, Bertin, et al. 1994; …

FREQUENTLY ASKED QUESTION

So, “swing theory” or “modal theory”?

Swing is a mechanism, not a theory. Swing is overreflection of leading into trailing waves; as such, it is automatically incorporated in the modal theory of bar modes. Normal spiral modes are excited by long into short wave overreflection.

The rapidly evolving/external origin scenario (continuous spectrum, no unstable modes, “one-shot” tidal scenario) has not been demonstrated to be viable and is unlikely to be the general explanation of grand design spiral structure.

CONCLUSIONS OF PART II

- Semi-empirical approach is the successful approach

- Role of gas initially underestimated, as a passive component only. Grand-design normal spiral structure requires gas (self-regulation) to be long-lasting

- Role of dark halos. Pure disks would naturally generate barred spiral structure

- Normal and barred spiral structure within a unified theory

- If grand-design spiral structure were not so frequent, the tidal scenario could be thought as a plausible possibility. In the K-band two-armed grand-design structure is ubiquitous; thus the tidal scenario is not viable

PART III
SOME RECENT RESULTS

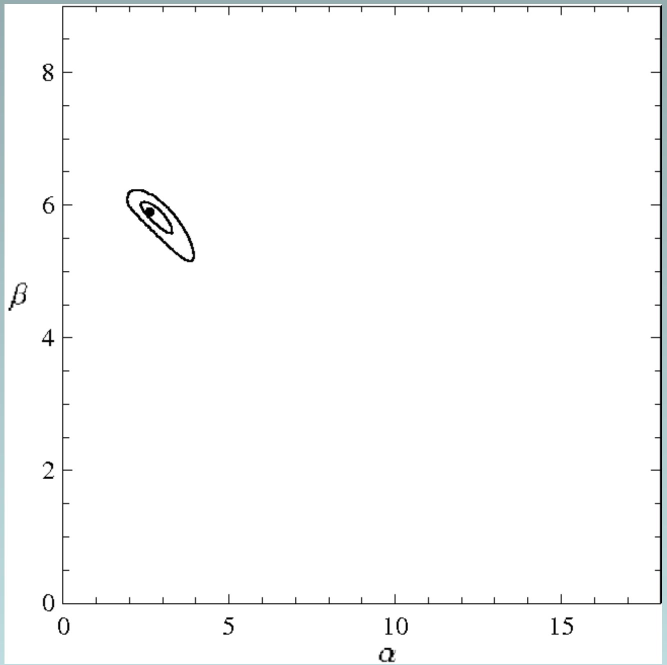
There remains a common belief, that, to support grand design spiral structure, galaxy disks must be heavy. Here are two recent results in favor of light disks.

(1) Self-consistent disk-halo decomposition of a rotation curve (Amorisco & Bertin 2010, A&A).

“Live” isothermal halo embedding an axisymmetric disk (see also Monet et al. 1981). General results are in favor of relatively light disks.

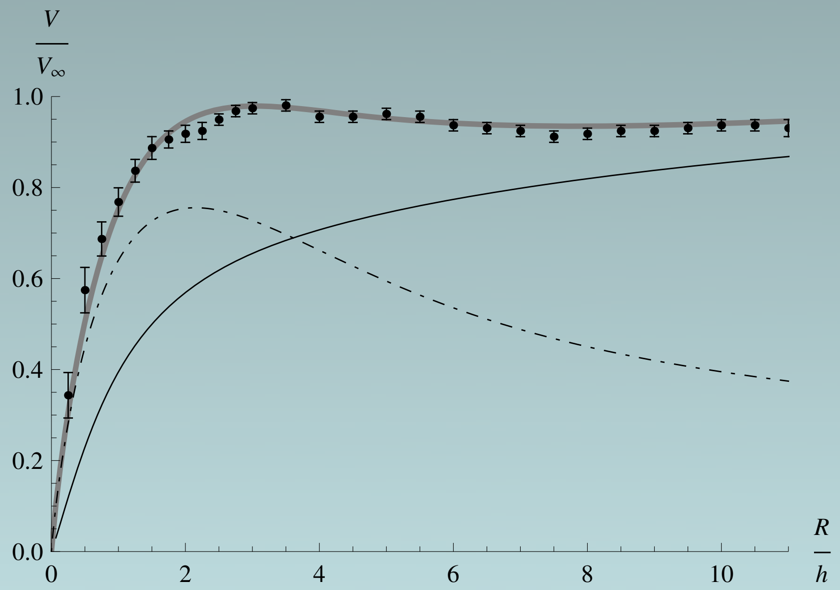
[Note that light disks are also found by the *Disk Mass Project* (Bershady et al. 2011, ApJL; for UGC 463, see Westfall et al. 2011, ApJ).]

Self-consistent disk-halo decomposition of the rotation curve of NGC 3198. Halo assumed to be isothermal, with shape and density distribution determined by disk-halo interaction. The best fit has a lighter disk than the maximum-disk.



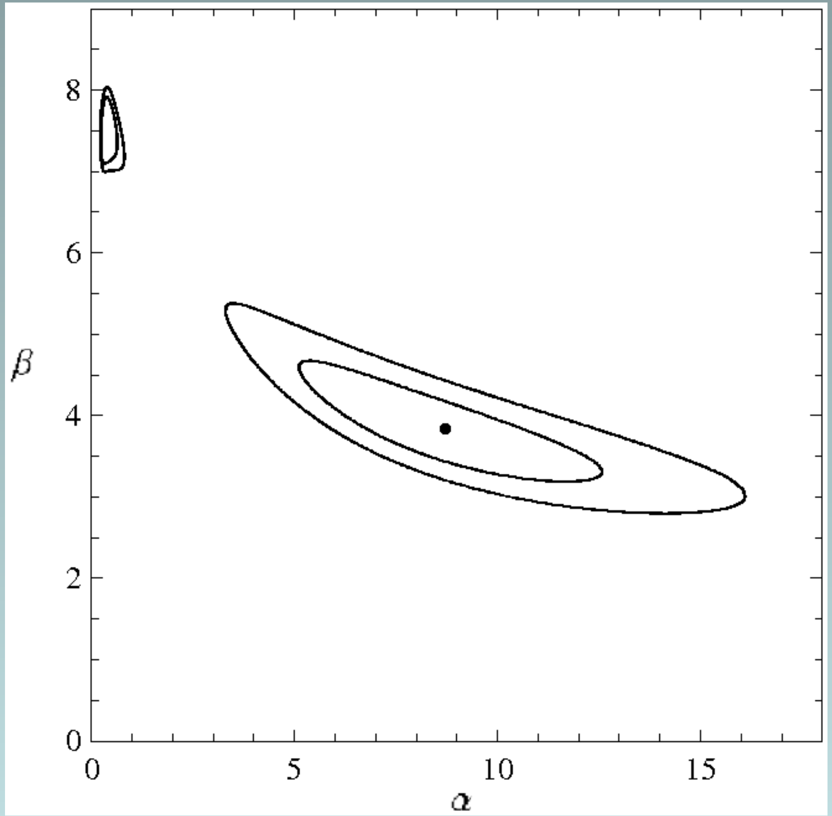
max disk

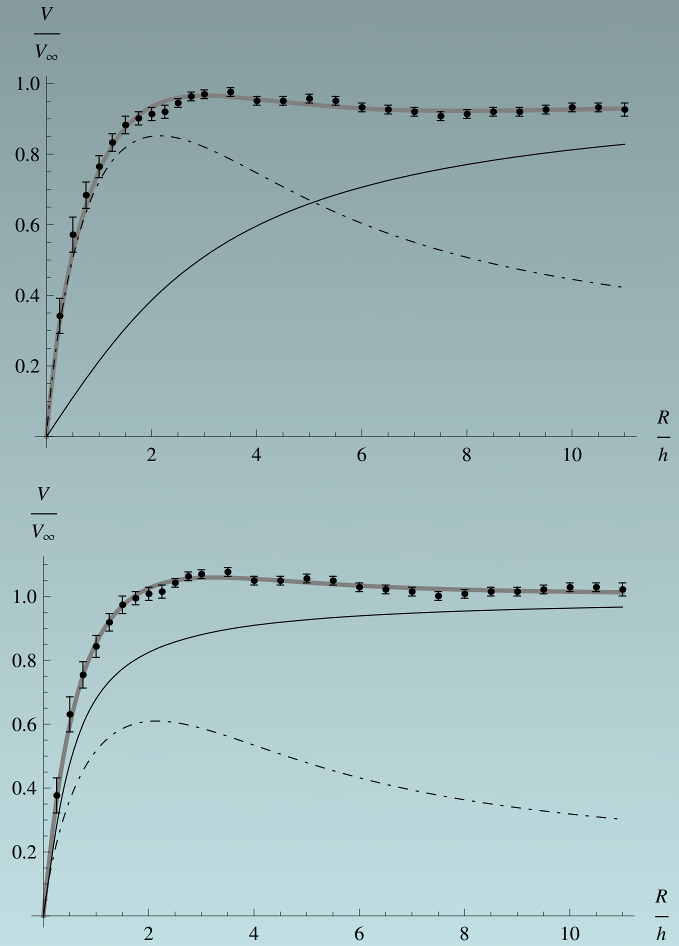
heavy halo



Parametric non-consistent decomposition. Disk-halo degeneracy. Note the maximum-disk solution.

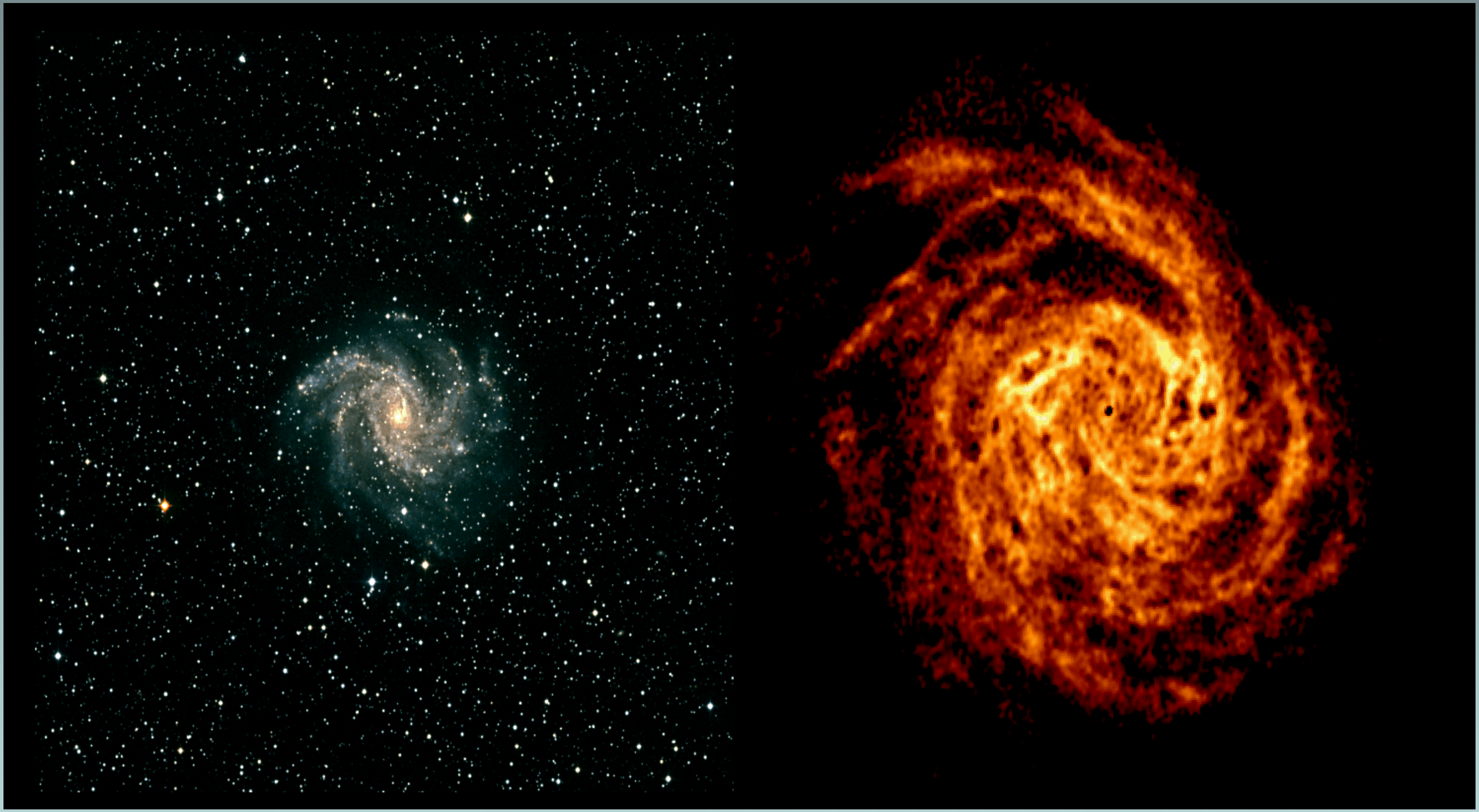
max
disk





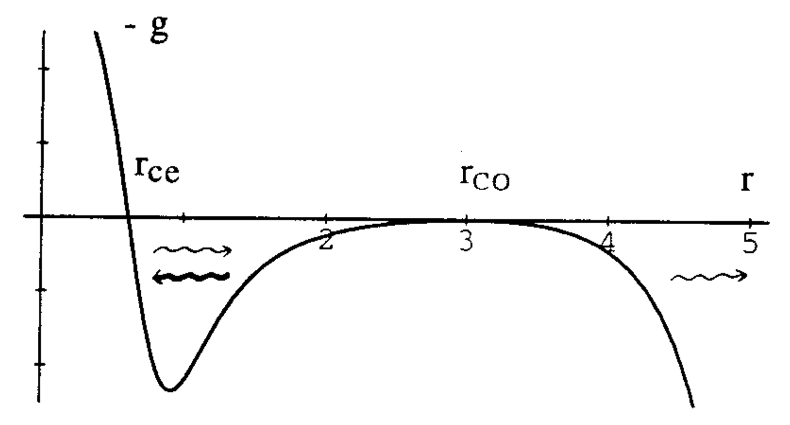
heavy
halo

(2) Prominent spiral arms in the gaseous outer disks (Bertin & Amorisco 2010, A&A).



NGC 6946 R. Boomsma, F. Fraternali, T. Oosterloo, R. Sancisi, J. van der Hulst 2002

Large scale arms well beyond the optical disk



$$\frac{d^2u}{dr^2} + g(r;\omega)u = 0$$

Outgoing wave at large radii

In the gas the density wave is partly transmitted beyond the Outer Lindblad Resonance: then its amplitude is governed by the conservation of wave action (Bertin & Amorisco 2010, A&A), with a sort of tsunami in the outer disk. A quantitative test is proposed.

DENSITY WAVE THEORY: FUTURE PROSPECTS & CONCLUSIONS

- Re-examine the shock scenario in the light of the new picture of the Interstellar Medium within the theory of Global Spiral Modes.

- Re-examine the other classical tests in the light of the theory of Global Spiral Modes and of the newly acquired observing capabilities.

- Self-regulation, non-linear evolution, and statistical studies of spiral morphologies at finite z.

The Density Wave Theory, for the explanation of spiral structure in galaxies (and with following applications to the context of self-gravitating accretion disks), is one of the key achievements in Astrophysics.

A semi-empirical approach based on the working hypothesis of Quasi-Stationary Spiral Structure in galaxies has originated an impressive number of quantitative successful observational tests that have attracted the interest of the astronomical community in the last five decades.

The development of a successful and internally consistent theory, started with the seminal papers by Lin and Shu about 50 years ago, has required the solution of a number of challenging conceptual problems at the frontier of Astrophysics and of Applied Mathematics.