Mon. Not. R. astr. Soc. (1988) **235**, 1343–1359

An analytic theory of self-propagating star formation

Thomas Neukirch* *Astronomisches Institut, Ruhr-Universität Bochum, Postfach 10 21 48, 4630 Bochum 1, FRG*

Johannes V. Feitzinger *Sternwarte der Stadt Bochum, Castroper Str. 67, 4630 Bochum, FRG &Astronomisches Institut, Ruhr-Universität Bochum,Postfach 10 21 48, 4630 Bochum 1, FRG*

Accepted 1988 June 8. Received 1988 April 27; in original form 1988 January 29

Summary. An analytic theory of self-propagating star formation is derived. A basic assumption is the smallness of the propagation length scale compared with typical galactic length scales. A system of non-linear partial differential equations results. It is shown that in contrast to other models based on so-called *reaction–diffusion equations* it is possible not only to include details of the physics of the interstellar medium, but also the dynamics of the galaxy consistently. A simple example is investigated to demonstrate the ability of the theory to produce spiral structure. An interesting analogy of the percolation phase transition of computer models and the first instability of the analytic theory is found.

1 Introduction

Almost a decade ago the concept of the formation of galactic structures by stochastic self-propagating star formation (SSPSF) was developed (Mueller & Arnett 1976; Gerola & Seiden 1978). These and almost all subsequent investigations of self-propagating star formation were done by computer simulations (Seiden, Schulmann & Gerola 1979; Comins 1981; Gerola, Seiden & Schulmann 1980; Seiden, Schulmann & Feitzinger 1982; Freedman & Madore 1982; Seiden 1983) or based on the specific assumptions underlying the computer models (Schulmann & Seiden 1982).

The first analytic approach to self-propagating star formation was made by Cowie & Rybicki (1982) and by Balbus (1984), but they assume a wavefront to exist and do not include any feedback mechanisms in their theory. Wavefronts were shown to exist under certain special circumstances in SSPSF-simulations by Freedman, Madore & Mehta (1984).

Closely related to the theory of self-propagating star formation is the theory of the interstellar medium (ISM). Several investigations showed that in addition to the well-known static three-phase state (McKee & Ostriker 1977), time-dependent states may exist, as cyclic phase

*Present address: Institut für theoretische Physik IV, Ruhr-Universität Bochum, Postfach 10 21 48, 4630 Bochum 1, FRG.

changes or burst-like runaway states (Habe, Ikeuchi & Tanaka 1981; Ikeuchi & Tomita 1983; Ikeuchi, Habe & Tanaka 1984). Besides these pure ISM-models, where only the gas phases are considered, a few authors investigated the back reaction of the gas phase on star formation (Shore 1981; Bodifée & de Loore 1985). They found the same periodic or burst-like time structures as the investigators of the pure ISM-models.

These models only considered the time development of the ISM, resp. stellar phase and did not take care of spatial couplings which might lead to propagating phenomena. The mathematical structure of these models was the same as that of models of chemical reactions with dissipative structures (e.g. Nicolis & Prigogine 1977). In these chemical reaction models spatial coupling is introduced by adding a diffusion term to each equation. The resulting set of equations is called a system of reaction–diffusion equations (RD-equations). Self-organized spatial structuring is then possible. Therefore several authors introduced additional diffusion terms into their model equations (Shore 1983; Nozakura & Ikeuchi 1984; Feitzinger 1985). Ferrini, Marchesoni & Shore (1985) give a derivation of the diffusion term from a master equation.

Whereas Nozakura & Ikeuchi (1984) investigated a pure ISM-model, Shore (1983) and Feitzinger (1985) developed models with diffusion of stellar matter. All authors find wave-like phenomena in their models, but only Shore (1983) and Feitzinger (1985) propose that their models represent analytical models of self-propagating star formation. Nozakura & Ikeuchi (1984) mention the resemblance of what they call a ‘trigger wave’, but they emphasize the difference between matter diffusion and the ‘diffusion’ of star formation, which is characteristic of self-propagating star formation.

As a theoretical approach to self-propagating star formation, the RD-equation models have several shortcomings, which we will discuss in the next section. The aim of this paper is to present a new approach to get the correct spatial coupling terms for the description of self-propagating star formation in the limit of small propagation length scales.

The outline of this paper is as follows. In Section 2 we show the inconsistency of the simple RD-equation approach with the basic ideas of self-propagating star formation. Based on this discussion we develop in Section 3 the basic equations of the new analytical model. In Section 4 a concrete example is investigated and in Section 5 we compare our approach with the results of the SSPSF-computer simulations and give a short discussion of the new approach.

2 Basic considerations

As a starting point let us summarize the main features of the proposed RD-models (Shore 1983; Nozakura & Ikeuchi 1984; Feitzinger 1985). The mass density of a galaxy is divided up into several distinct components. The mass density of the component i is denoted by ρ_i. In the above-mentioned RD-models it is postulated that each component obeys a non-linear partial differential equation:

$$\frac{\partial \rho_i}{\partial t} + \nabla \cdot (\mathbf{v} \rho_i) = F_i(\boldsymbol{\rho}) + D_i \Delta \rho_i, \tag{2.1}$$

where D_i is a constant diffusion coefficient, Δ is the Laplace operator, $\mathbf{v}$ is the bulk velocity of the matter flow and $F_i(\boldsymbol{\rho})$ is (in the simplest case) a non-linear *function* of the density vector $\boldsymbol{\rho} = (\rho_1, \ldots, \rho_n)$. This set of equations describes the mass balance between the different components. The physical meaning of the terms in equation (2.1) is the following:

$\nabla \cdot (\mathbf{v} \rho_i)$ describes the density change due to the bulk motion of matter which in spiral galaxies is differential rotation.

$F_i(\boldsymbol{\rho})$ describes the creation and destruction of stars and the heating or cooling of gas.

The $D_i\Delta\rho_i$-term was introduced to couple neighbourhood space points and to allow for wave-like phenomena which are well known to occur in similar sets of equations, e.g. in theoretical chemistry and biology (Nicolis & Prigogine 1977; Haken 1977, 1983). A possible justification for this is given by Ferrini *et al.* (1985). As we will discuss shortly, a simple diffusion in space is not a correct description of self-propagating star formation.

To complete the discussion of the RD-models we state the further assumption usually made that the total density distribution is stationary and therefore the dynamics of the galaxy may be neglected. The same assumption is made in all SSPSF-simulations. Since the densities $\rho_i(\mathbf{r})$ should add up to give the total density $\rho(\mathbf{r})$, we may eliminate one equation from the system (2.1):

$$\rho(\mathbf{r})=\sum_{i=1}^{n}\rho_i(\mathbf{r}). \tag{2.2}$$

If we impose the constraint of total mass conservation, $\rho(\mathbf{r})$ should obey the equation of continuity

$$\frac{\partial\rho}{\partial t}+\nabla\cdot(\mathbf{v}\rho)=\nabla\cdot(\mathbf{v}\rho)=0, \tag{2.3}$$

where the first equality sign is due to the assumption of stationarity. (2.3) is fulfilled identically, if ρ is taken to be axisymmetric and $\mathbf{v}$ is purely rotational motion about the axis of symmetry. This is an appropriate assumption for spiral galaxies. If we now sum up system (2.1), using (2.2) and (2.3), we get:

$$\sum_{i=1}^{n}(F_i(\boldsymbol{\rho})+D_i\Delta\rho_i)=0. \tag{2.4}$$

Since the F_i describe the creation or destruction of different sorts of matter and since the total mass should be conserved and furthermore the galaxy is assumed to be in a state of *dynamical equilibrium*, we should have

$$\sum_{i=1}^{n}F_i(\boldsymbol{\rho})=0 \tag{2.5}$$

which means that there is no net creation or destruction of matter. Equation (2.5) immediately leads to

$$\sum_{i=1}^{n}D_i\Delta\rho_i=0. \tag{2.6}$$

We have to conclude that either we have at least one negative diffusion coefficient or that one of the components is forced to flow in a way to compensate the diffusion gains or losses of the other components in a certain space region in order not to invalidate the assumption of stationarity for the total mass density. Both possibilities are not acceptable. It is therefore not consistent for RD-models to fix the total density distribution *a priori*. If matter transport by diffusion is allowed, the dynamics of the galaxy has to be included explicitly.

There are further difficulties for RD-models of self-propagating star formation. Stars are formed out of interstellar gas (molecular clouds). In the SSPSF-models the stars which have formed at one point *trigger* star formation at a neighbourhood point by shock waves etc. They do not *move* to the point where the new star formation event occurs. Therefore self-propagating star formation is in a certain way *non-local.* In contradiction to this, the RD-models are local in the sense that the stellar matter first moves to the neighbourhood space points and triggers star formation there. This is not characteristic of self-propagating star formation. The only motion of the stars in the SSPSF-models is differential rotation. Any other motion will lead to contradictions with the assumptions of axisymmetry and stationarity of the total mass density.

Another unnatural feature of the RD-models is that the star formation would propagate even to places which are completely void of gas, whereas propagation should only occur to places with enough gas. This is of course due to the modelling of *matter flow* instead of modelling the *flow of star formation.*

So we come to the conclusion that RD-equations in this simple form are not capable of describing self-propagating star formation correctly. A new approach is necessary, which will be developed in the following section.

3 The new approach

We mentioned above that the basic assumption of self-propagating star formation is that the formation of stars at a point $\mathbf{r}$ at the time t is triggered by stars which were at the point $\mathbf{r}-\delta\mathbf{r}$ at the time $t-\delta t$. The distance $|\delta\mathbf{r}|$ and the delay time δt are not independent, but give the average spreading velocity $\mathbf{v}_\mathrm{T}$ of the trigger mechanism:

$$\mathbf{v}_\mathrm{T}=\frac{\delta\mathbf{r}}{\delta t}. \tag{3.1}$$

As a starting point of the mathematical discussion, we again take the system of equations (2.1), but we now drop the diffusion terms $D_i\Delta\rho_i$, which led to difficulties. We then get:

$$\frac{\partial\rho_i}{\partial t}+\nabla\cdot(\mathbf{v}\rho_i)=F_i[\boldsymbol{\rho}], \tag{3.2}$$

where $F_i[\boldsymbol{\rho}]$ means that F_i is not necessarily a *function* of its argument, but may have a more complicated dependence on $\boldsymbol{\rho}$. It is convenient to concentrate on the equation which shall describe the propagation of star formation. Let the densities of the participating components be ρ_s for the stellar density and ρ_g for the gas density, where gas density means the density of the gas component in which stars are formed. We will not specify this any further. The following discussion may be easily extended to as many components as suitable. It is merely a matter of convenience to take only two components for the following discussion. The corresponding F will be called F_s.

First we split F_s up into a part F_sc which describes the creation of stars and into a part F_sd which describes the destruction of stars:

$$F_\mathrm{s}=F_\mathrm{sc}-F_\mathrm{sd}, \tag{3.3}$$

where F_sc and F_sd are positive definite. The interesting part for star formation is F_sc; let us therefore discuss this part. In the RD-models the assumption was made that the F_i were functions of the densities $\rho_i(\mathbf{r},t)$:

$$F_i=F_i[\boldsymbol{\rho}(\mathbf{r},t)]. \tag{3.4}$$

As stated above the basic idea of self-propagating star formation is the triggering of star formation at location $\mathbf{r}$ at time t by stars which resided at $\mathbf{r}-\delta\mathbf{r}$ at time $t-\delta t$. Therefore F_s will depend on $\rho_s(\mathbf{r}-\delta\mathbf{r}, t-\delta t)$. As we are dealing with a mean-field theory, we have to add up all influences of those stars over all space and all former times. Obviously, the more distant in space and time the less influence a star or a stellar association will have on the star formation process at location $\mathbf{r}$. To take this into account we introduce a weight function $\Phi(\delta\mathbf{r}, \delta t)$ which goes to zero very fast for growing $|\delta\mathbf{r}|$ and δt. This is in analogy to the SSPSF-models, where it is assumed that only nearest neighbour cells have an influence on the star formation process in a specific cell and that this influence lasts only for one time step. The following ansatz is then made for F_{sc}:

$$F_{sc}(\mathbf{r}, t)=\int_{-\infty}^{\infty} d\delta\mathbf{r} \int_{0}^{\infty} d\delta t\, \Psi(\rho_s(\mathbf{r}-\delta\mathbf{r}, t-\delta t), \rho_g(\mathbf{r}, t))\cdot\Phi(\delta\mathbf{r}, \delta t), \tag{3.5}$$

where Ψ is a *function* of its arguments. This is the point where stochasticity enters. Φ may be considered as the probability density per unit time of a unit amount of stellar matter at place $\mathbf{r}-\delta\mathbf{r}$ at the time $t-\delta t$ to trigger star formation at the place $\mathbf{r}$ at time t. Φ is therefore taken to be positive definite and normalized to unity:

$$\int_{-\infty}^{\infty} d\delta\mathbf{r} \int_{0}^{\infty} d\delta t\, \Phi(\delta\mathbf{r}, t)=1. \tag{3.6}$$

Ψ specifies how much stars are formed from the gas density $\rho_g(\mathbf{r}, t)$ due to a certain stellar density ρ_s at place $\mathbf{r}-\delta\mathbf{r}$ at time $t-\delta t$.

In this ansatz the non-local character of self-propagating star formation is explicit. We get an integro-differential equation for the development of the stellar density. We assume that the trigger mechanism (e.g. shock waves) expands spherically symmetrically in the absence of interstellar matter. Therefore Φ will only depend on the propagation distance $|\delta\mathbf{r}|$. The direction of propagation of star formation will be determined by the dependence of Ψ on the gas density ρ_g.

Since in normal galaxies the maximum propagation length scale will be small compared with a galactic radius, we may expand $\Psi[\rho_s(\mathbf{r}-\delta\mathbf{r}, t-\delta t), \rho_g(\mathbf{r}, t)]$ into a Taylor series in $\delta\mathbf{r}$ and δt, keeping only the first order in δt and the second order in $\delta\mathbf{r}$:

$$\Psi[\rho_s(\mathbf{r}-\delta\mathbf{r}, t-\delta t), \rho_g(\mathbf{r}, t)]=\Psi[\rho_s(\mathbf{r}, t), \rho_g(\mathbf{r}, t)]$$

$$-\frac{\partial\Psi}{\partial\rho_s}[\rho_s(\mathbf{r}, t), \rho_g(\mathbf{r}, t)]\cdot\left\{\delta\mathbf{r}\cdot\nabla\rho_s+\delta t\frac{\partial\rho_s}{\partial t}\right\}+\frac{1}{2}\delta\mathbf{r}\cdot\mathbf{A}\cdot\delta\mathbf{r}, \tag{3.7}$$

where $\mathbf{A}$ is the matrix with the elements A_{ij} defined by:

$$A_{ij}=\frac{\partial^2\Psi}{\partial\rho_s^2}\frac{\partial\rho_s}{\partial x_i}\frac{\partial\rho_s}{\partial x_j}+\frac{\partial\Psi}{\partial\rho_s}\frac{\partial^2\rho_s}{\partial x_i\partial x_j} \tag{3.8}$$

taken at $\delta\mathbf{r}=\delta t=0$.

Inserting (3.7) into (3.5) we see that all first-order terms in $\delta\mathbf{r}$ and all second-order terms with $i\neq j$ vanish, because an odd function is integrated over the complete $\delta\mathbf{r}$-space. Introducing the abbreviations:

$$T=\int_{-\infty}^{\infty} d\delta\mathbf{r} \int_{0}^{\infty} d\delta t\, \delta t\, \Phi \tag{3.9}$$

$$D=\int_{-\infty}^{\infty} d\delta\mathbf{r} \int_{-\infty}^{\infty} d\delta t(\delta\mathbf{r})^2\Phi \tag{3.10}$$

and using (3.6), we get for F_{sc}:

$$F_{\mathrm{sc}}\left[\rho_{\mathrm{s}}(\mathbf{r},t), \nabla\rho_{\mathrm{s}}(\mathbf{r},t), \frac{\partial\rho_{\mathrm{s}}}{\partial t}(\mathbf{r},t), \Delta\rho_{\mathrm{s}}(\mathbf{r},t), \rho_{\mathrm{g}}(\mathbf{r},t)\right]$$

$$=\Psi[\rho_{\mathrm{s}}(\mathbf{r},t),\rho_{\mathrm{g}}(\mathbf{r},t)]-T\frac{\partial\Psi}{\partial\rho_{\mathrm{s}}}\frac{\partial\rho_{\mathrm{s}}}{\partial t}+D\frac{\partial^2\Psi}{\partial\rho_{\mathrm{s}}^2}(\nabla\rho_{\mathrm{s}})^2+D\frac{\partial\Psi}{\partial\rho_{\mathrm{s}}}\Delta\rho_{\mathrm{s}}. \tag{3.11}$$

Since we do not know Φ explicitly, we cannot calculate the numerical values of T and D from (3.9) and (3.10). T and D have a simple physical interpretation. T is the mean value of δt, as (3.9) shows. Since δt represents the time necessary for the propagation of star formation for one event, T is a typical value of this time and therefore is the timescale of self-propagating star formation. Compared to SSPSF-computer models T should be of the order of one timestep of these models.

D is the mean value of $(\delta\mathbf{r})^2$. Since $|\delta\mathbf{r}|$ is the range of a single star formation event, $D^{1/2}$ may be interpreted as the average propagation length scale, which in comparison to the SSPSF-computer models would be of the order of one cell size of these models.

Given the average length scale and the average time scale, the average propagation velocity may be defined as:

$$v_{\mathrm{prop}}=\frac{D^{1/2}}{T}. \tag{3.12}$$

If we insert (3.11) into (3.2) and solve for $\partial\rho_{\mathrm{s}}/\partial t$, we get for ρ_{s} the non-linear evolution equation:

$$\frac{\partial\rho_{\mathrm{s}}}{\partial t}=\frac{1}{1+T(\partial\Psi/\partial\rho_{\mathrm{s}})}\left\{-\nabla(\mathbf{v}\rho_{\mathrm{s}})+\Psi-F_{\mathrm{sd}}+D\frac{\partial\Psi}{\partial\rho_{\mathrm{s}}}\Delta\rho_{\mathrm{s}}+D\frac{\partial^2\Psi}{\partial\rho_{\mathrm{s}}^2}(\nabla\rho_{\mathrm{s}})^2\right\}. \tag{3.13}$$

The new spatial coupling terms have the general form:

$$\frac{D(\partial\Psi/\partial\rho_{\mathrm{s}})}{1+T(\partial\Psi/\partial\rho_{\mathrm{s}})}\Delta\rho_{\mathrm{s}}+\frac{D(\partial^2\Psi/\partial\rho_{\mathrm{s}}^2)}{1+T(\partial\Psi/\partial\rho_{\mathrm{s}})}(\nabla\rho_{\mathrm{s}})^2. \tag{3.14}$$

The first of these terms is reminiscent of a diffusion term, but it is to be noticed that the diffusion coefficient in general depends on both densities. As discussed above propagation to regions without gas should be inhibited. Even though we have not yet specified Ψ explicitly this requirement leads to very general restrictions on Ψ. The terms representing star formation including the spatial coupling terms have to vanish for zero gas density to fulfil this requirement. Therefore, if $\rho_{\mathrm{g}}=0$:

$$\Psi(\rho_{\mathrm{s}},0)=0, \tag{3.15}$$

$$\frac{\partial\Psi}{\partial\rho_{\mathrm{s}}}(\rho_{\mathrm{s}},0)=0, \tag{3.16}$$

$$\frac{\partial^2\Psi}{\partial\rho_s^2}(\rho_s, 0)=0. \tag{3.17}$$

In this case in (3.13) only the usual differential rotation term and the stellar destruction term are left.

Due to total mass conservation we have:

$$\sum_{i=1}^{n} F_i=0. \tag{3.18}$$

There must then be at least one further component which obeys a non-linear evolution equation similar to (3.13). Obviously this will be the gas component from which stars are formed. There may of course be further components which will be affected by the propagation phenomenon, e.g. H I-gas heated by shock waves or clouds which are formed by the 'snow-plough'-effect of shock waves. Therefore in more elaborate models we will have a system of equations like (3.13) supplemented by equations for the components not affected by propagating star formation. The validity of the expansion will depend on the smallness of the parameter ξ defined as:

$$\xi^2=\frac{D}{R^2}, \tag{3.19}$$

where R is a typical length scale of the galaxy under consideration. Therefore if the propagation length scale and the length scale of the galaxy are comparable, higher-order terms may become important in (3.7) or the expansion may not be valid at all and (3.5) has to be used in (3.3). This should not be the case for normal galaxies but might be the case for dwarf galaxies.

It is possible to give up the assumption of a stationary mass distribution. We would then have to supplement the current set of equations describing the mass balance between the components by another set describing momentum balance, energy balance and the Poisson equation for the inclusion of self gravitation. Such systems of equations without the inclusion of self-propagating star formation were given by Scalo & Struck-Marcell (1984) for the description of the gas phase of a galaxy and by Chiang & Prendergast (1985) for both the gas and the stellar phase. In this way it would in principle be possible to get a completely self-consistent model of a galaxy with self-propagating star formation, although the mathematical problem would of course be very difficult.

4 A simple example

To illustrate some aspects of the new approach let us discuss a very simple example. Consider a circular two-dimensional disc galaxy with a constant surface density σ and radius R. Let the total surface density be split up into two components only: the gas density σ_g and the stellar density σ_s. σ_s and σ_g add up to give σ:

$$\sigma_s(\mathbf{r}, t)+\sigma_g(\mathbf{r}, t)=\sigma=\text{const}. \tag{4.1}$$

If we solve (4.1) for either σ_g or σ_s, we may eliminate this density from the evolution equation. Therefore we need only one evolution equation in this simple example. The other density is given by (4.1).

The assumptions of this example are of course oversimplifications. But our aim here is not to model a galaxy in every single aspect, but to demonstrate that our approach is capable of including propagating star formation into a consistent picture of a galaxy, which may be derived in the future. This simple model will also shed some light on the first SSPSF-simulations which were based on similar assumptions.

Eliminating σ_g, equation (3.13) now reads:

$$\frac{\partial \sigma_s}{\partial t}=\frac{1}{1+T(\partial\Psi/\partial\sigma_s)(0)}\left\{-\nabla\cdot(\mathbf{v}\sigma_s)+\Psi(\sigma_s,\sigma-\sigma_s)-F_{sd}(\sigma_s)+D\frac{\partial\Psi}{\partial\sigma_s}(0)\,\Delta\sigma_s+D\frac{\partial^2\Psi}{\partial\sigma_s^2}(\nabla\sigma_s)^2\right\}. \tag{4.2}$$

As a boundary condition we take a vanishing stellar density at R:

$$\sigma_s(R)=0. \tag{4.3}$$

To complete the problem we have to specify the functions Ψ, F_{sd} and the velocity field $\mathbf{v}(r,\varphi)$. (3.15)–(3.17) become:

$$\Psi(\sigma_s,0)=0, \tag{4.4}$$

$$\frac{\partial\Psi}{\partial\sigma_s}(\sigma_s,0)=0, \tag{4.5}$$

$$\frac{\partial^2\Psi}{\partial\sigma_s^2}(\sigma_s,0)=0. \tag{4.6}$$

If we neglect spontaneous star formation for simplicity, we also have:

$$\Psi(0,\sigma_g)=\Psi(0,\sigma-\sigma_s)=0 \tag{4.7}$$

since there should be no induced star formation if the stellar density is zero. Furthermore Ψ should be positive for positive densities:

$$\Psi\geq 0 \quad \text{for} \quad \sigma_s,\sigma_g\geq 0. \tag{4.8}$$

Since F_{sd} represents the destruction of stars, F_{sd} should vanish if the stellar density is zero:

$$F_{sd}(0)=0. \tag{4.9}$$

F_{sd} does not depend on the gas density for obvious reasons and it will be positive for positive σ_s. With these quite general, but physically reasonable assumptions on Ψ and F_{sd}, we may now look for solutions of (4.2). One obvious solution is the stationary homogeneous solution

$$\sigma_s=\sigma_{s0}\equiv 0 \tag{4.10}$$

which means that there is only gas in the 'galaxy'. This is a possible starting point for galactic evolution. We now investigate the stability of solution (4.10). Setting:

$$F(\sigma_s)=\Psi(\sigma_s,\sigma-\sigma_s)-F_{sd}(\sigma_s) \tag{4.11}$$

and linearizing (4.2) about (4.10), we get:

$$\sigma_s(r,\varphi,t)=\sigma_{s0}+\sigma_{s1}(r,\varphi,t) \tag{4.12}$$

$$\frac{\partial\sigma_{s1}}{\partial t}=\frac{1}{1+T(\partial\Psi/\partial\sigma_s)(0)}\left\{-\nabla\cdot(\mathbf{v}\sigma_{s1})+\frac{dF}{d\sigma_s}(0)\,\sigma_{s1}+D\frac{\partial\Psi}{\partial\sigma_s}(0)\,\Delta\sigma_{s1}+\dots\right\}. \tag{4.13}$$

With the normal mode ansatz:

$$\sigma_{s1}=\hat{\sigma}_{mk}(r)\exp(\lambda_{mk}t)\exp(\mathrm{i}m\varphi) \tag{4.14}$$

we get the ordinary differential equation:

$$\lambda_{mk}\hat{\sigma}_{mk}=\frac{1}{1+T(\partial\Psi/\partial\sigma_s)(0)}\left\{-\mathrm{im}\Omega(r)\,\hat{\sigma}_{mk}+\frac{dF}{d\sigma_s}(0)\,\hat{\sigma}_{mk}\right.$$
$$\left.+D\frac{\partial\Psi}{\partial\sigma_s}(0)\left[\frac{d^2\hat{\sigma}_{mk}}{dr^2}+\frac{1}{\mathrm{r}}\frac{d\hat{\sigma}_{mk}}{dr}-\frac{m^2}{r^2}\hat{\sigma}_{mk}\right]\right\}, \tag{4.15}$$

where a purely rotational velocity field has been assumed and $\Omega(r)$ is the angular velocity. (4.15) together with the boundary condition (4.3) is an eigenvalue problem for the eigenvalue λ_{mk}. Instability will occur, if at least one of the eigenvalues λ_{mk} has a positive real part $\Re(\lambda_{mk})$.

Let us write (4.15) in a more convenient form:

$$\frac{d^2\hat{\sigma}_{mk}}{dr^2}+\frac{1}{r}\frac{d\hat{\sigma}_{mk}}{dr}-\frac{m^2}{r^2}\hat{\sigma}_{mk}-\frac{\mathrm{im}\Omega}{D(\partial\Psi/\partial\sigma_s)(0)}\hat{\sigma}_{mk}=\omega_{mk}\hat{\sigma}_{mk} \tag{4.16}$$

with

$$\omega_{mk}=\frac{1}{D(\partial\Psi/\partial\sigma_s)(0)}\left[\left(1+T\frac{\partial\Psi}{\partial\sigma_s}(0)\right)\lambda_{mk}-\frac{dF}{d\sigma_s}(0)\right]. \tag{4.17}$$

In the Appendix it is shown that ω_{mk} has always a negative real part for arbitrary $\Omega(r)$. It is also shown that ω_{01} has the largest real part. Since

$$\lambda_{mk}=\frac{D(\partial\Psi/\partial\sigma_s)(0)\,\omega_{mk}+(dF/d\sigma_s)(0)}{1+T(\partial\Psi/\partial\sigma_s)(0)} \tag{4.18}$$

and $(\partial\Psi/\partial\sigma_s)(0)\geq 0$ as well as D, $T>0$, we have $\Re(\lambda_{mk})>0$ only if $(dF/d\sigma_s)(0)>0$, because $\Re(\omega_{mk})$ is negative. As ω_{01} has the biggest real part, $\Re(\lambda_{01})>0$ first of all. If $(dF/d\sigma_s)(0)$ becomes large, we may also have $\Re(\lambda_{mk})>0$ with $m=1, 2, \ldots$. This depends of course on the rotation curve. So in our simple model the first axisymmetric mode is the most unstable one. The eigenvalues and eigenfunctions of all axisymmetric modes may be calculated explicitly without knowledge of $\Omega(r)$. For $m=0$, (4.15) reduces to Bessel's equation and therefore the eigenfunctions are:

$$\hat{\sigma}_{0k}=J_0\left(j_{0k}\frac{r}{R}\right) \tag{4.19}$$

where $J_0(x)$ is the Bessel function of the first kind and j_{0k} is the kth zero of $J_0(x)$. For the eigenvalues λ_{0k} we get:

$$\lambda_{0k}=\frac{(dF/d\sigma_s)(0)-D(\partial\Psi/\partial\sigma_s)(0)(j_{0k}/R)^2}{1+T(\partial\Psi/\partial\sigma_s)(0)}. \tag{4.20}$$

For some special rotation curves $v_\varphi(r)=r\Omega(r)$, we are able to obtain all eigenfunctions explicitly. There is a mathematical coincidence between our linearized equation (4.15) and the linearized equation discussed by Shore (1983, p. 204); therefore the solutions look similar.

There is of course no coincidence between the parameters occurring in the differential equations. Furthermore we impose homogeneous Dirichlet boundary conditions at a finite radius which leads to a discrete spectrum of eigenvalues. The resulting criteria for instability are therefore different from those of Shore (1983).

For the trivial case of no rotation $\Omega=0$ we just get Bessel functions of higher order and similar eigenvalues as (4.20) with the corresponding zeros j_{mk} replacing j_{0k}.

For rigid rotation we get the same eigenfunctions as for no rotation but the eigenvalues are now complex. Let Ω be the angular velocity, then

$$\lambda_{mk}=\frac{(dF/d\sigma_s)(0)-D(\partial\Psi/\partial\sigma_s)(0)(j_{mk}/r)^2-\mathrm{i}m\Omega}{1+T(\partial\Psi/\partial\sigma_s)(0)}. \tag{4.21}$$

The third and most interesting case is that of a flat rotation curve $v_\varphi=\text{const}=v_0$. (4.16) may then be solved by confluent hypergeometric or Kummer's functions (Abramowitz & Stegun 1965). The eigenfunctions are:

$$\hat{\sigma}_{mk}=(2B_{mk}r)^m\exp(-B_{mk}r)\,M(a_{mk},b_m,2B_{mk}r), \tag{4.22}$$

where $M(a, b, z)$ is the confluent hypergeometric function,

$$B_{mk}=\omega_{mk}^{1/2}, \tag{4.23}$$

$$a_{mk}=m+\frac{1}{2}+i\frac{mv_0}{2DB_{mk}(\partial\Psi/\partial\sigma_s)(0)}, \tag{4.24}$$

$$b_m=2m+1. \tag{4.25}$$

We get for λ_{mk}:

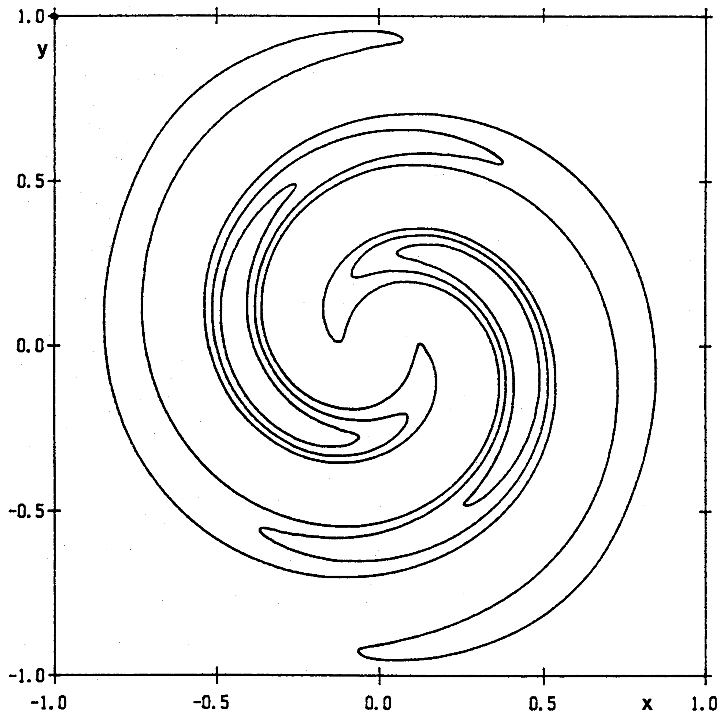
$$\lambda_{mk}=\frac{(dF/d\sigma_s)(0)+D(\partial\Psi/\partial\sigma_s)(0)(z_{mk}/2R)^2}{1+T(\partial\Psi/\partial\sigma_s)(0)}, \tag{4.26}$$

where z_{mk} is the kth zero of the confluent hypergeometric function $M(a_{mk}, b_m, 2B_{mk}r)$ and may be complex. Equation (4.15) was integrated numerically with the shooting method (e.g. Press *et al.* 1986). Several normal modes are shown in Fig. 1 for $m=2$ and different values of the parameter

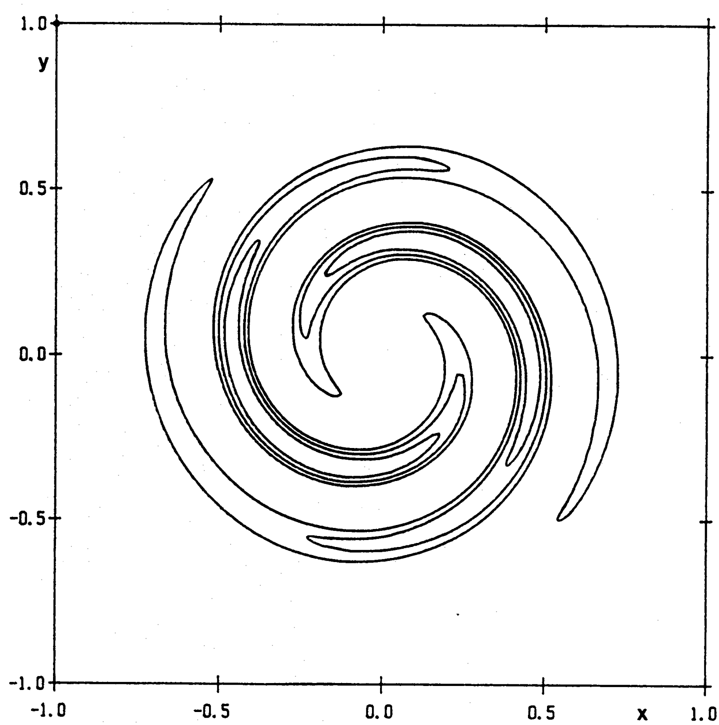
$$K=\frac{v_0R}{D(\partial\Psi/\partial\sigma_s)(0)} \tag{4.27}$$

so that the pictures show a sequence of increasing rotational velocity v_0. The resulting modes have a clear spiral shape. The equidensity contours wind up and are compressed with increasing velocity. We remind the reader that we chose a constant density distribution as the starting point of our analysis, so the shape of the contours will be altered if a more realistic total density distribution is chosen. Nevertheless the possible structures obtained here are not too far from structures obtained with more realistic galaxy models and density wave theory (*cf.* Fig. 2).

It has to be mentioned that in this constant density model, we will never have an isolated $m=2$ mode instability, because if an $m=2$ mode is unstable, there is at least one unstable $m=1$ mode and one unstable $m=0$ mode, too. So we will possibly have some difficulties to get symmetric two-armed structures in a constant density model. On the other hand, there remain

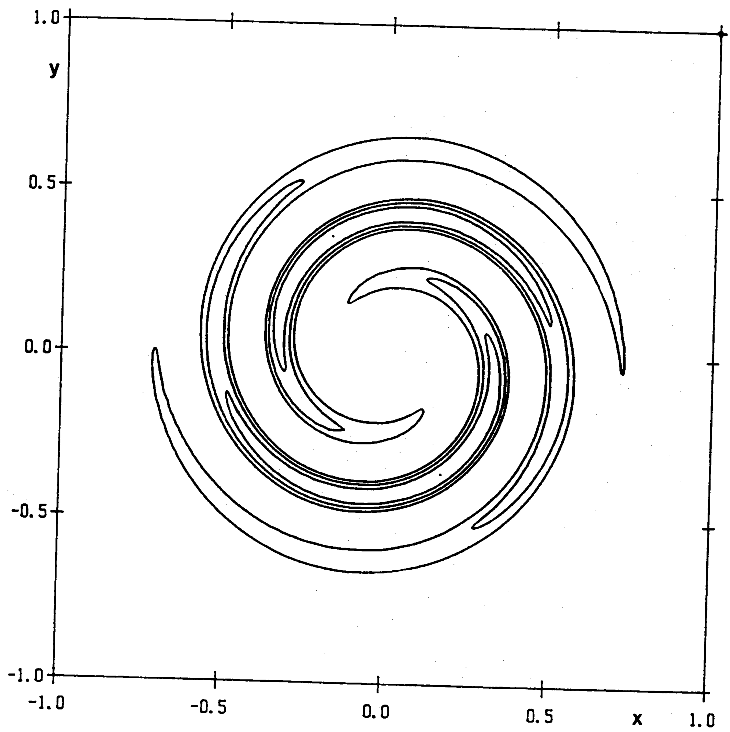


(a)



(b)

Fig. 1. Density contours of the real part of normal modes for $m=2$ are shown for the following values of K: (a) $K=100$; (b) $K=200$; (c) $K=300$.



(c)

Figure 1 – *continued.*

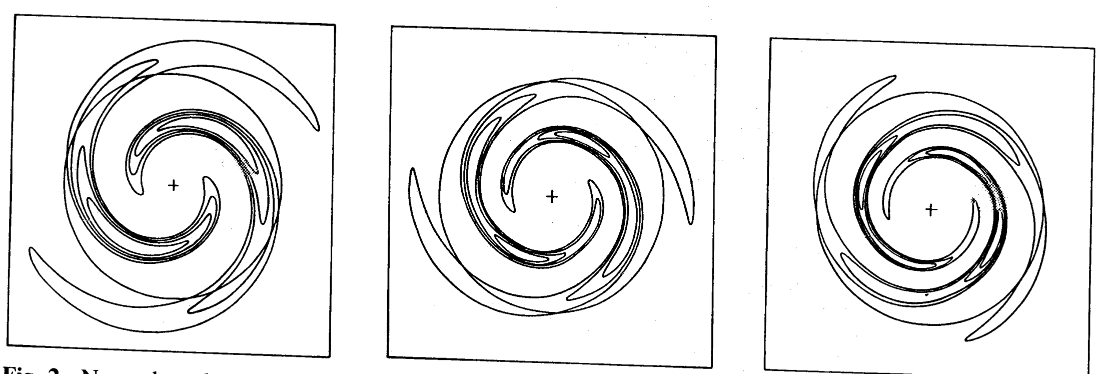


Fig. 2. Normal modes of galactic models in density wave theory (taken from Haass, Berlin & Lin 1982).

a lot of further possibilities which may lead to symmetric structures in more elaborate models or in the nonlinear regime, but that will be a matter of future work.

5 Comparison with SSPSF-simulations and discussion

In the SSPSF-simulations it was found that up to a critical propagation probability P_c, star formation is negligible (e.g. Seiden & Gerola 1982). If the propagation probability exceeds P_c, star formation sets in. Schulman & Seiden (1982) explained this feature of the SSPSF-models as a percolation phase transition. The model discussed in the last section has similar properties. As long as $A \equiv dF/d\sigma_s(0)$ does not exceed a critical value A_c, the solution $\sigma_s = 0$ is stable and star formation remains negligible. Increasing A over A_c leads to an instability and star formation will start. Therefore we may compare the percolation phase transition of the SSPSF-models and the $m=0$-instability of our analytical model. Having done this we are now able to give a possible answer to the question why the SSPSF-simulations do not produce two-armed spirals. The reason may be the correspondence of the phase transition to an axisymmetric instability. The structures which are produced in the simulations are due to the stochasticity of the SSPSF-models. If several similar runs were averaged, there would be no structures left, except axisymmetric ones. A non-axisymmetric instability would lead to a structured disk even if several runs were averaged over. It is also clear why the structures seen in the computer models are more pronounced near the phase transition. It is well known that fluctuations are largest near a phase transition. So the structures seen in the SSPSF-simulations are just the stochastic noise near the percolation phase transition.

No further instabilities were found in the simple constant density SSPSF-models (*cf.* Seiden & Gerola 1982). This is not the case for our model which indicates non-axisymmetric instabilities for growing A. This may be due to the fact that local mass conservation is not always guaranteed in the SSPSF-simulations and therefore condition (4.1) is violated in those models. Another reason may be the existence of a stable stationary non-homogeneous axisymmetric solution of (4.2), which corresponds to the average solution of the SSPSF-simulations.

It should be investigated whether it is possible to get non-axisymmetric instabilities in the SSPSF-models if (4.1) is explicitly incorporated. An indication of non-axisymmetric instabilities was found in the more elaborate SSPSF-model of Seiden *et al.* (1982). It may be valuable to translate their model into our analytical approach.

In this paper we have given a new theoretical formulation for theories of self-propagating star formation. The resulting analytical description of self-propagating star formation allows the inclusion of the galactic dynamics in a self-consistent way. It is also possible to distinguish between the importance of the influence of different processes in the ISM on the evolution of a galaxy.

The investigation of a simple model shows that it is in principle possible to get unstable spiral shaped modes of star formation. The authors hope that more elaborate models will give firm criteria to decide whether the dynamics of a galaxy or the local processes in the ISM are responsible for the formation of spiral structure.

References

Abramowitz, M. & Stegun, I. A., 1965. *Handbook of Mathematical Functions*, Dover, New York.
Balbus, S. A., 1984. *Astrophys. J.*, **277**, 550.
Bodifée, G. & de Loore, C., 1985. *Astr. Astrophys.*, **142**, 297.
Chiang, W. H. & Prendergast, K. H., 1985. *Astrophys. J.*, **297**, 507.
Comins, N. F., 1981. *Mon. Not. R. astr. Soc.*, **194**, 169.
Cowie, L. L. & Rybicki, G. B., 1982. *Astrophys. J.*, **260**, 504.
Feitzinger, J. V., 1985. *The Milky Way Galaxy*, p. 559, Reidel, Dordrecht, Holland.
Ferrini, F., Marchesoni, F. & Shore, S. N., 1985. *Adv. Chem. Phys.*, **62**, 493.
Freedman, W. L. & Madore, B. F., 1983. *Astrophys. J.*, **265**, 140.
Freedman, W. L., Madore, B. F. & Mehta, S., 1984. *Astrophys. J.*, **282**, 412.
Gerola, H. & Seiden, P. E., 1978. *Astrophys. J.*, **223**, 129.
Gerola, H., Seiden, P. E. & Schulman, L. S., 1980. *Astrophys. J.*, **242**, 517.
Haass, J., Bertin, G. & Lin, C. C., 1982. *Proc. natn. Acad. Sci. U.S.A.*, **79**, 3908.
Habe, A., Ikeuchi, S. & Tanaka, Y. D., 1981. *Publs. astr. Soc. Japan*, **33**, 23.
Haken, H., 1977. *Synergetics*, Springer, Berlin.
Haken, H., 1983. *Advanced Synergetics*, Springer, Berlin.
Ikeuchi, S. & Tomita, H., 1983. *Publs. astr. Soc. Japan*, **35**, 56.
Ikeuchi, S., Habe, A. & Tanaka, Y. D. 1984. *Mon. Not. R. astr. Soc.*, **207**, 909.
McKee, C. F. & Ostriker, J. P., 1977. *Astrophys. J.*, **218**, 148.
Mueller, M. W. & Arnett, W. D., 1976. *Astrophys. J.*, **210**, 670.
Nicolis, G. & Prigogine, I., 1977. *Self-Organization in Nonequilibrium Systems*, Wiley, New York.

Nozakura, T. & Ikeuchi, S., 1984. *Astrophys. J.*, **279,** 40.
Press, W. H., Flannery, B. P., Teukolsky, S. A. & Vetterling, W. T., 1986. *Numerical Recipes*, Cambridge University Press, Cambridge.
Scalo, J. M. & Struck-Marcell, C., 1984. *Astrophys. J.*, **276,** 60.
Schulman, L. S. & Seiden, P. E., 1982. *J. stat. Phys.*, **27,** 83.
Seiden, P. E., 1983. *Astrophys. J.*, **266,** 555.
Seiden, P. E. & Gerola, H., 1979. *Astrophys. J.*, **233,** 56.
Seiden, P. E. & Gerola, H., 1982. *Fundam. Cosmic Phys.*, **7,** 241.
Seiden, P. E., Schulman, L. S. & Feitzinger, J. V., 1982. *Astrophys. J.*, **253,** 91.
Seiden, P. E., Schulman, L. S. & Gerola, H., 1979. *Astrophys. J.*, **232,** 702.
Shore, S. N., 1981. *Astrophys. J.*, **249,** 93.
Shore, S. N., 1983. *Astrophys. J.*, **265,** 202.

Appendix A

We separate the operator

$$\mathbf{L}_{\mathrm{m}}=\frac{d^2}{dr^2}+\frac{1}{r}\frac{d}{dr}-\mathrm{im}\Theta(r) \tag{A1}$$

into a hermitian part $\mathbf{L}_{\mathrm{m}}^{\mathrm{H}}$ and an anti-hermitian part $\mathbf{L}_{\mathrm{m}}^{\mathrm{A}}$:

$$\mathbf{L}_{\mathrm{m}}^{\mathrm{H}}=\frac{d^2}{dr^2}+\frac{1}{r}\frac{d}{dr}-\frac{m^2}{r^2} \tag{A2}$$

$$\mathbf{L}_{\mathrm{m}}^{\mathrm{A}}=-\mathrm{im}\Theta(r), \tag{A3}$$

where we used the scalar product

$$\langle f, g\rangle=\int_0^{\mathrm{R}} \bar{f}(r)\, g(r)\, r\, dr \tag{A4}$$

for the definition of hermitian and anti-hermitian and $\bar{f}$ is the complex conjugate of f. Since $\mathbf{L}_{\mathrm{m}}^{\mathrm{H}}$ has a complete set of orthonormal eigenfunctions f_{mj}^{H} which may be chosen to be real valued, we may expand the eigenfunctions χ_{mj} of $\mathbf{L}_{\mathrm{m}}$ into a series of f_{mj}^{H}:

$$\chi_{mj}=\sum_k c_{kj} f_{mk}^{\mathrm{H}} \tag{A5}$$

with complex coefficients c_{kj}. The eigenvalue ω_{mj} of the operator $\mathbf{L}_{\mathrm{m}}$ may then be written as follows:

$$\omega_{mj}=\frac{\int_0^{\mathrm{R}} \bar{\chi}_{mj}\mathbf{L}_{\mathrm{m}}\chi_{mj}\, r\, dr}{\int_0^{\mathrm{R}} \bar{\chi}_{mj}\chi_{mj}\, r\, dr}. \tag{A6}$$

Inserting (A5) and using (A2), (A3) and the orthogonality of f_{mj}^{H}, we get:

$$\omega_{mj}=\frac{\Sigma_k |c_{kj}|^2\omega_{mk}^{\mathrm{H}}+I}{\Sigma_k |c_{kj}|^2}, \tag{A7}$$

where

$$\omega^{\mathrm{H}}_{mk} = -\frac{j^2_{mk}}{R^2} \tag{A8}$$

are the eigenvalues of the hermitian operator $\mathbf{L}^{\mathrm{H}}_{\mathrm{m}}$, which are negative. I is defined by:

$$I = \sum_{j,n} \bar{c}_{kj} c_{kn} \int_0^{\mathrm{R}} f^{\mathrm{H}}_{mj} \mathbf{L}^{\mathrm{A}}_{\mathrm{m}} f^{\mathrm{H}}_{mn}\, r\, dr. \tag{A9}$$

Using the anti-hermiticity of $\mathbf{L}^{\mathrm{A}}_{\mathrm{m}}$, we get:

$$I = -\sum_{j,n} \bar{c}_{kj} c_{kn} \int_0^{\mathrm{R}} (\mathbf{L}^{\mathrm{A}}_{\mathrm{m}} f^{\mathrm{H}}_{mj})\, f^{\mathrm{H}}_{mn}\, r\, dr.$$

$$= -\bar{I}. \tag{A10}$$

Therefore, denoting the real part of a complex number z by $\Re(z)$, we get:

$$\Re(I) = 0 \tag{A11}$$

and

$$\Re(\omega_{mk}) = \frac{\Sigma_j |c_{kj}|^2 \omega^{\mathrm{H}}_{mj}}{\Sigma_j |c_{kj}|^2} < \omega^{\mathrm{H}}_{m1} < 0 \tag{A12}$$

(A12) is valid for arbitrary m and since

$$\ldots < \omega^{\mathrm{H}}_{m+11} < \omega^{\mathrm{H}}_{m1} < \cdots < \omega^{\mathrm{H}}_{01} < 0 \tag{A13}$$

(Abramowitz & Stegun 1965) we see that if non-axisymmetric ($m \neq 0$) modes are unstable, at least one unstable mode of every $n < m$ exists, too.

Appendix B

Let us give a concrete example to illustrate the ideas of Section 4. We will rely on the SSPSF-models as closely as possible in order to be able to compare our results with the results of the computer simulations. Let $\Psi(\sigma_{\mathrm{s}}, \sigma_{\mathrm{g}})$ be defined as:

$$\Psi = a\sigma_{\mathrm{s}}\sigma_{\mathrm{g}} \tag{B1}$$

and

$$F_{\mathrm{sd}} = b\sigma_{\mathrm{s}} \tag{B2}$$

where a and b are constants. The dependence of Ψ on the gas density is the same as in the simple SSPSF-model of Seiden & Gerola (1979). The explicit dependence of Ψ on the stellar density has no analogy in the SSPSF-models. The simplest computer simulations use a constant propagation probability. We believe it to be more realistic to take Ψ to depend on the stellar density, too. F_{sd} models the vanishing of stars, e.g. by supernova events and is for simplicity taken to be proportional to the stellar density. It is obvious that F_{sd} will not depend on the gas density. As in the SSPSF-model by Seiden & Gerola (1979), we neglect spontaneous star formation. We here only investigate the constant rotation curve case.

Equation (4.2) now reads:

$$\frac{\partial \sigma_s}{\partial t}=\frac{1}{1+Ta(1-\sigma_s)}\left\{-\frac{v_0}{r}\frac{\partial \sigma_s}{\partial \varphi}+a\sigma_s(1-\sigma_s)-b\sigma_s+Da(1-\sigma_s)\,\Delta\sigma_s\right\}. \tag{B3}$$

The linearization of (B3) reads:

$$\frac{\partial \sigma_{s1}}{\partial t}=\frac{1}{1+Ta}\left\{-\frac{v_0}{r}\frac{\partial \sigma_{s1}}{\partial \varphi}+(a-b)\,\sigma_{s1}+Da\Delta\sigma_{s1}+\dots\right\}. \tag{B4}$$

Introducing the dimensionless variables

$$x=\frac{r}{R}, \tag{B5}$$

$$\xi^2=\frac{D}{R^2}, \tag{B6}$$

$$A=Ta, \tag{B7}$$

$$B=Tb, \tag{B8}$$

$$\eta=\frac{v_0}{v_{\text{prop}}}, \tag{B9}$$

$$\Lambda=T\lambda, \tag{B10}$$

and the normal mode ansatz (4.14), we finally get the equation:

$$\left[\frac{d^2}{dx^2}+\frac{1}{x}\frac{d}{dx}-\frac{m^2}{x^2}-i\frac{mn}{A\xi}\frac{1}{x}\right]\hat{\sigma}_{mk}=\omega_{mk}\hat{\sigma}_{mk} \tag{B11}$$

with

$$\omega_{mk}=\frac{(1+A)\,\Lambda_{mk}-A+B}{A\xi^2}. \tag{B12}$$

The solutions of (B11) which vanish for $x=1$ are given by

$$\sigma_{mk}=(2\omega_{mk}^{1/2}x)^m\exp(-\omega_{mk}^{1/2})\,M(a_{mk},b_m,2\omega_{mk}^{1/2}x) \tag{B13}$$

with

$$a_{mk}=m+\frac{1}{2}+i\frac{m\eta}{2\xi A^{1/2}} \tag{B14}$$

$$b_m=2m+1. \tag{B15}$$

The parameter K of equation (4.27) turns out to be

$$K=\frac{\eta}{\xi A}. \tag{B16}$$

Since K is the only parameter in the numerical integration of (B11), we conclude that the shape of the normal modes does only depend on K.

The shape of the normal modes is therefore determined by the ratio of rotational velocity to propagation velocity, by the ratio of the propagation length scale to the galactic length scale and by the ratio of the propagation time scale to the star formation time scale.