THE ASTROPHYSICAL JOURNAL, 433:303–312, 1994 September 20

THE CO/SiO RADIATIVE INSTABILITY IN COOL STAR ATMOSPHERES REVISITED

M. CUNTZ[1] AND D. O. MUCHMORE[2]
Received 1993 December 6; accepted 1994 March 22

ABSTRACT

We revisit the formation of radiative instabilities in cool star atmospheres and compare our results with those given by Muchmore, Nuth, & Stencel. We have considered the combined influence of CO and SiO molecules and have computed models for a grid of effective temperatures and geometrical dilution factors for the stellar radiation. Our results are based on the analysis of the energy balance of gas elements with prescribed thermodynamic properties. Our results show that radiative instabilities are most likely primarily caused by CO, whereas SiO is expected to play only a minor role, except when the CO density is reduced compared to LTE values or the CO band can be assumed to be optically thick. The onset of radiative instabilities is expected to be strongly modified when dynamic phenomena such as stochastic shocks are present. Our results provide strong evidence that dust formation can most likely not occur via a radiative instability alone. Therefore, we present a revised version of the Muchmore et al. dust formation paradigm, which also considers hydrodynamic cooling. The new paradigm is particularly relevant in cases where dust is formed relatively close to the stellar photosphere.

Subject headings: circumstellar matter — instabilities — stars: late-type — molecular processes — stars: atmospheres

1. INTRODUCTION

In an earlier paper, Muchmore, Nuth, & Stencel (1987, hereafter MNS) have presented theoretical evidence that radiative instabilities due to SiO molecules can occur in the outer atmospheres of cool giants and supergiant stars, rapidly reducing the temperature by as much as 1000 K. The work of MNS was preliminary, but it was evident that their results would have important implications for the physics of outer stellar atmospheres, such as (1) the formation of molecules and dust, (2) the occurrence of bifurcated (i.e., two-component) cool star chromospheres, (3) the occurrence of SiO maser emission, and (4) outer atmospheric dynamic phenomena. Maciel (1976, 1977), Elitzur, Brown, & Johnson (1989), and Jørgensen & Johnson (1992) have discussed the dynamic role of molecules in outer stellar atmospheres including their possible role in generating mass loss. The physics of SiO masers was the topic of two reviews by Elitzur (1992) and Stencel (1993). Muchmore (1990) has assessed molecular equilibria and disequilibria in red giant atmospheres.

Molecular cooling instabilities are also considered to be the major cause for the two-component structure of chromospheres in the Sun and other stars. Observational and theoretical results for the Sun were given by Ayres (1981), Ayres & Testerman (1981), Kneer (1983), Deming et al. (1984), Muchmore & Ulmschneider (1985), Muchmore (1986), Ayres, Testerman, & Brault (1986), Mauas, Avrett, & Loeser (1990), and Athay & Dere (1990). Results for late-type stars other than the Sun, particularly for Arcturus (K1.5 III), were given by Heasley et al. (1978), Ayres (1986), Tsuji (1988), Cuntz & Muchmore (1989), and Wiedermann & Ayres (1990). The latter authors analyzed CO fundamental lines near 4.7 μm for several G and K giant stars and showed that all these stars must have bifurcated upper atmospheres consisting of "hot" and "cool" components. Many years ago, Glassgold & Langer (1976) discussed cooling instabilities in CO clouds, which are important for evolution and fragmentation processes. In a very recent paper, Le Bourlot et al. (1993) present a detailed study of thermal bifurcation in dark interstellar clouds due to the nonlinear behavior of the interstellar chemical reaction rates, which leads to a low- and a high-ionization phase of the gas.

The result that the outer atmospheres of late-type stars, particularly cool giants and supergiants, are most likely radiatively unstable was also used to interpret recent "nonstandard" observational results. In a recent study Drake et al. (1992) reported results from a radio continuum observation monitoring program for Betelgeuse (M2 Iab). They found variabilities in the form of dips at different wavelength, which were uncorrelated and behaved stochastically. The timescales ranged from 10^6 to 10^7 s. Drake et al. argued that the observations indicate thermal instabilities in the wind region, which can lead to sudden molecule formation and therefore to the observed drastic decrease in the degree of hydrogen ionization. Clayton et al. (1992) recently discussed a series of *IUE* observations together with visible and infrared data of the RCB stars RY Sgr, R CrB, and V854 Cen. The authors concluded that dust is formed relatively close to the stellar surface, which is inconsistent with traditional stellar atmosphere models. Evidence that dust formation close to the stellar surface can also occur in Betelgeuse was given by Bester et al. (1991). Brugel et al. (1991) reported observational and theoretical results for L_2 Puppis, a Mira star, obtained during 1985–1987. These results provided evidence that patches of the stellar surface alternate between a quiet state (with little dust, an extended warm atmosphere, and low atmospheric TiO absorption) and a dusty wind state (with a cooler atmosphere and higher atmospheric TiO absorption). This result also indicates hydrochemical instabilities and associated atmospheric velocity structures.

The results of MNS are quite intriguing but are nevertheless preliminary. The reason is severalfold. First, MNS have treated

[1] High Altitude Observatory and Advanced Study Program, National Center for Atmospheric Research, P.O. Box 3000, Boulder, CO 80307-3000. The National Center for Atmospheric Research is sponsored by the National Science Foundation. Also Joint Institute for Laboratory Astrophysics, University of Colorado and National Institute of Standards and Technology.
[2] 6909 Weedin Place NE A304, Seattle, WA 98115.

solely SiO band opacity together with a continuum opacity in the optically thin approximation, while ignoring frequency-dependent effects of other molecules such as CO. Second, MNS have treated only a very small set of effective temperatures and only cases with a dilution factor 0.5 for the stellar radiation field. As continuum opacity MNS used a Rosseland mean opacity based on the table of Kurucz (1979), which was widely extrapolated to lower temperatures and pressures. Better continuum opacities have meanwhile become available, which are expected to affect the results. We now have considered the combined influence of CO and SiO molecules and have computed models for a larger grid of effective temperatures and geometrical dilution factors for the stellar radiation. The radiation in the SiO band was assumed to be optically thin in all models. In the case of the CO band we also considered the optically thin limit and also explored the potential impact of optical depth effects by considering reduced CO cooling rates. For O, C, and Si we assumed solar abundances. For the continuum opacity, we use the version of Alexander, Augason, & Johnson (1989). We also note that Cuntz & Muchmore (1989) have meanwhile investigated time-dependent effects of acoustic wave heating and molecular cooling in the outer atmosphere of Arcturus (K1.5 III) and have treated CO, SiO, as well as a continuum opacity also including radiative transfer effects. Cuntz & Muchmore found that the radiative energy losses occurred preferably in CO, whereas SiO as well as the continuum appeared as a source of *radiative heating*. This result was entirely unexpected and should be viewed as a further reason to revisit radiative instabilities in cool stars.

We should not rule out entirely that molecules other than CO and SiO also impact the onset of radiative instabilities. These molecules must fulfill two criteria: First, the molecules must be abundant. Second, the molecular spectrum must consist of distinct bands which lie at largely different wavelengths or, alternatively, the spectrum must contain a large number of lines which produce a considerable temperature variation of the cooling rate. In all other cases, the molecules can only impact the generation of radiative instabilities in conjunction with other constituents such as H^-. We note that H_2O molecules might possibly match the above-mentioned criteria. In any case, our paper is not intended to ultimately "solve" the problem of the initiation of radiative instabilities; nevertheless we hope to make a significant contribution to this topic.

The paper is structured as follows: In § 2, we discuss the behavior of the CO/SiO radiative cooling function and the instabilities which occur in or near LTE. In § 3, we discuss timescales for some of the competing processes which may invalidate the LTE assumption. We also present estimates for the influence of dynamic phenomena such as shock waves with long and short periods. In § 4, we apply our results to models of Betelgeuse and present a new dust formation paradigm. Our conclusions are given in § 5.

2. PHYSICS OF RADIATIVE INSTABILITIES

Ayres (1981) and Kneer (1983) have first emphasized that the potential instability in the cooling function for an optically thin gas arises from the competing influence of a continuum opacity, notably H^-, and molecular opacity, particularly that of CO. This phenomenon depends upon (*i*) the presence of strong vibration-rotation bands in the infrared and (*ii*) the very strong temperature dependence of molecular concentrations in some ranges of temperature.

For our study we take the simplest possible approach, a local stability criterion strictly applicable only to optically thin gas. As in Cuntz & Muchmore (1989) we use a three-frequency point approximation and introduce temperature-dependent frequency weights:

$$w_i(T) = \int_{\lambda_{\min,i}}^{\lambda_{\max,i}} B_\lambda(T) d\lambda / B(T) , \qquad (1)$$

where the integration intervals are from 4.29 to 6.2 μm for CO, referred to as $w_1(T)$, and from 7.5 to 10.0 μm for SiO, referred to as $w_2(T)$. The weight for the continuum band is then $w_3(T) = 1 - w_1 - w_2$, representing all other frequencies. For our study we take the molecular constants of Huber & Herzberg (1979a, b). For CO we take the partition function of Kurucz (1983) and for SiO we take the partition function of Rossi, Maciel, & Benevides-Soares (1985). For the continuum we use the Rosseland mean opacity table of Alexander et al. (1989). At low pressures, where tabulated values were not available, we extended the table by assuming $\partial\kappa/\partial p = 0$ for the opacity per gram. For further information, see also Muchmore & Ulmschneider (1985) and MNS (note that the denominator in eq. [1] of MNS should have been $4J + 2$).

The cooling function ϕ is then defined as $\phi = \sum_{i=1}^{3} \kappa_i (S_i - J_i)$, where S is the source function, and J is the mean intensity. We assume the gas to be optically thin, in LTE, and that it is irradiated by a star with an effective temperature $T_{\rm eff}$, so that the mean intensity is given by $J_\nu = W B_\nu(T_{\rm eff})$. W is the spherically symmetric dilution factor for the stellar radiation field given by Mihalas (1978, p. 120). Also note that this approach is a significant improvement over the work of MNS, who only treated SiO molecules together with the continuum band and only considered the case $W = 0.5$.

Assuming $S_\nu = B_\nu$ and introducing the functions $w_i(T)$ the cooling function becomes

$$\begin{aligned}\phi = {} & \kappa_{\rm CO}[B(T)w_1(T) - W \cdot B(T_{\rm eff})w_1(T_{\rm eff})] \\ & + \kappa_{\rm SiO}[B(T)w_2(T) - W \cdot B(T_{\rm eff})w_2(T_{\rm eff})] \\ & + \kappa_{\rm cont}[\![B(T)[1 - w_1(T) - w_2(T)] \\ & - W \cdot \{B(T_{\rm eff})[1 - w_1(T_{\rm eff}) - w_2(T_{\rm eff})]\}]\!] , \end{aligned} \qquad (2)$$

which is illustrated in Figure 1 for two different effective temperatures, with a range of dilution factors and densities. For O, C, and Si we use solar abundances.

The cooling function above is strictly valid only when both the CO and SiO bands are optically thin. Mauas et al. (1990) have extensively discussed the failure of this approach when optical depths become appreciable and found that even for a high-gravity star like the Sun, the approximation is reasonable for chromospheric layers and above. For giant stars, on the other hand, with large pathlengths, the CO band can easily become optically thick in the inner parts of a circumstellar envelope (cf. Cuntz & Muchmore 1989). This significantly changes the scenario which had been suggested by Kneer (1983) and MNS. Cooling by CO and by SiO occur at different places in a steadily expanding atmosphere. Each can only affect the energy balance significantly in those layers where $\tau \leq 1$, but where it is still appreciable. Figure 2 presents individual contributions to the cooling curve for two specific selections of parameters. In this case, CO completely dominates for $T < 4000$ K and the continuum dominates at higher T. SiO is insignificant compared to CO. The reason is that the abundance of Si is an order of magnitude smaller than that for CO and $w_1(T)/w_2(T)$ is typically about 5, caused by the behavior of the Planck function in the CO and SiO bands, while the opacity per molecule is very similar.

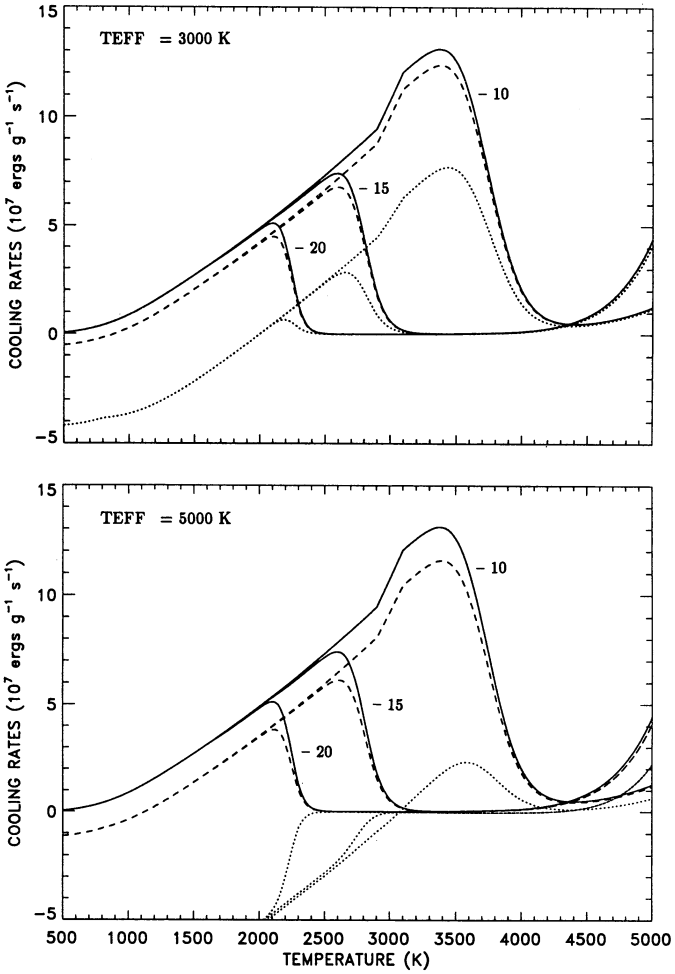

FIG. 1.—Radiative cooling functions considering CO, SiO, and continuum opacity at various densities and geometrical dilution factors for $T_{\rm eff}$ = 3000 K and 5000 K. The curves are labeled with the logarithm of the density. The geometrical dilution factors considered are for 1 R_* (*dotted lines*), 2 R_* (*dashed lines*), and ∞ (*solid lines*).

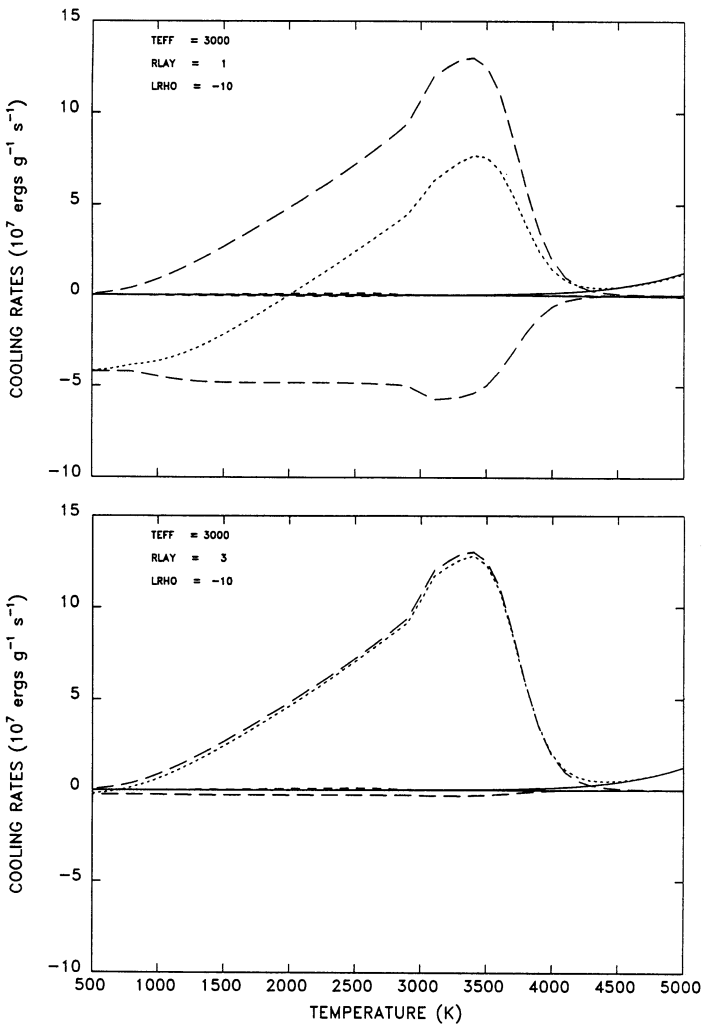

FIG. 2.—The different contributions to the radiative cooling functions for $T_{\rm eff}$ = 3000 K and log $\rho = -10$. The upper figure shows the results for 1 R_* and the lower figure for 3 R_*. The long-dashed lines indicate the heating and cooling in CO, the short-dashed lines (almost invisible) indicate the heating and cooling in SiO, and the solid lines indicate the heating and cooling in the continuum. The dotted lines show the total radiative cooling functions.

In Figure 3 we evaluate the domains of radiative instability defined as $\partial\phi/\partial T < 0$ in the temperature-density plane for different effective temperatures. We find that there is always a main "CO radiative instability island" most pronounced at low densities and relatively high effective temperatures. The influence of the geometrical dilution factor is found to be less important, except for $T_{\rm eff} \gtrsim 4500$ K and at distances $R \lesssim 1.3$ R_*. For small dilution factors, the differences due to different effective temperatures are insignificant, as expected. For large dilution factors, the temperature bifurcation depends also on the stellar effective temperature, particularly in the very low density regime. For $W = 0.5$, log $\rho = -15$, and $T_{\rm eff} = 3000$ K, we find a bifurcation of temperature between 2650 K and 3360 K. For $T_{\rm eff} = 4000$ K, the bifurcation is between 2750 K and 3370 K, and for $T_{\rm eff} = 5000$ K, it is between 3130 K and 3710 K. We also investigated the reason for the low-T and high-T boundaries of the main CO instability islands. We found that the low-T boundary is determined by the interplay among CO band absorption, CO band emission, and continuum absorption. CO absorption is most important at the high end of the effective temperature range considered, whereas continuum absorption dominates at the low end of the effective temperature range. The high-T boundary of the main CO instability islands is given by the interplay between CO band emission and continuum emission. Also note that our study completes an earlier study of Muchmore & Ulmschneider (1985), who have discussed the regions of thermal instability for $T_{\rm eff} = 5000$ K and $T_{\rm eff} = 6000$ K at conditions appropriate to the solar temperature minimum.

In some cases, additional regions of instability appear. These regions are attributable to the behavior of the continuum opacity, which shows that radiative instabilities are not always necessarily a consequence of frequency-dependent effects. In the neighborhood of $T = 2500$ K, $d\kappa_{\rm cont}/dT < 0$, due to the contributions of H_2O and TiO (Alexander, Johnson, & Rypma 1983; Alexander et al. 1989). However, the boundaries of these regions are poorly determined; the table of continuum opacities does not have adequate resolution to form good numerical derivatives for some regimes of T and p. In particular, there is a gross discontinuity in the table's behavior near log $T = 3.3$ and log $p = -2.5$. Near log $\rho = -5$, a further regime of insta-

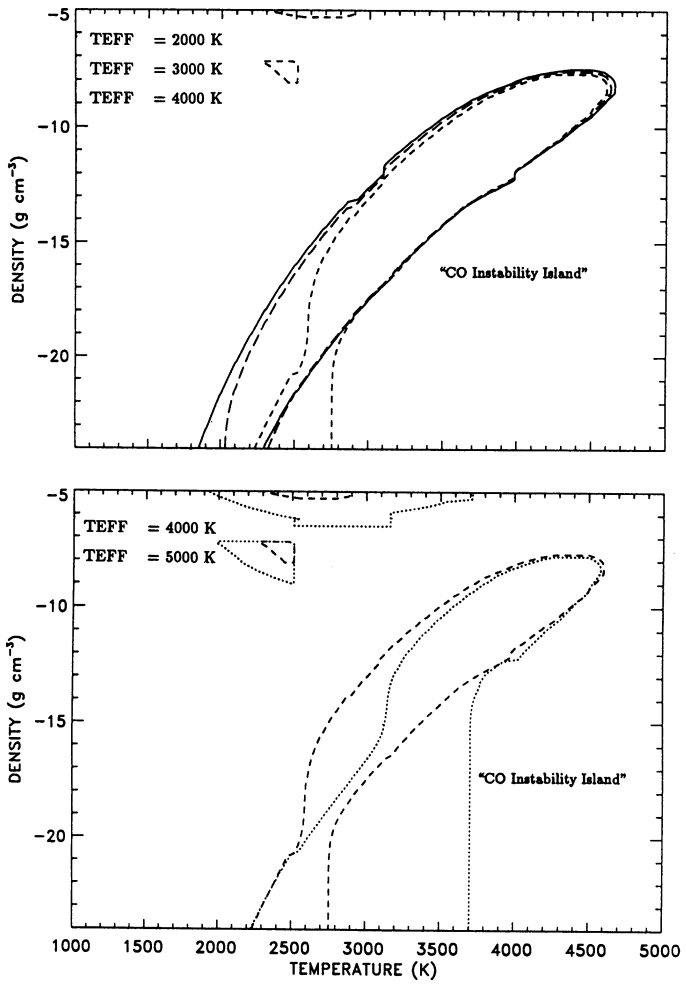

FIG. 3.—Domains of radiative instability in the temperature-density plane for different effective temperatures considering CO, SiO, and continuum opacity. The upper figure shows the results for 2000 K (*solid lines*), 3000 K (*long-dashed lines*), and 4000 K (*short-dashed lines*), whereas the lower figure shows the results for 4000 K (*short-dashed lines*) and 5000 K (*dotted lines*). In all cases the geometrical dilution factor is $W = 0.5$.

bility appears for $T_{\rm eff} \gtrsim 4000$ K. Here $\partial\phi/\partial T$ is dominated by the term $(\partial\kappa_{\rm cont}/\partial T)[B(T)w_3(T) - WB(T_{\rm eff})w_3(T_{\rm eff})]$. At the low-T end of this region, $WB(T_{\rm eff})w_3(T_{\rm eff}) \gg B(T)w_3(T)$; at the high-$T$ end, these terms become comparable, which is the condition responsible for the high-T boundary. The low-T boundary occurs as the Planck function shifts toward the optical band and continuum emission replaces CO absorption as the dominant contributor to $\partial\phi/\partial T$.

We now discuss the limiting case in which CO band opacity does not exist or is drastically reduced in size. This can occur either due to reduced CO particle densities caused by (time-dependent) non-LTE reaction rates or by an extreme Si overabundance/C underabundance, or, alternatively, in cases when the CO band is sufficiently optically thick, whereas the SiO band is optically thin. This case has been studied in detail by MNS. They used for the continuum opacities the table of Kurucz (1979), which is strictly valid only in the temperature range between log $T = 3.5$ and 5.0, and in the pressure range between log $p = -2.0$ and 6.0. As a consequence the table of Kurucz was widely extrapolated. This procedure was in part justified as the SiO band opacity often largely exceeded the continuum opacity. On the other hand, as the boundaries of the radiative instability regions are defined as $\partial\phi/\partial T = 0$, a very uncertain value for the continuum opacity can nevertheless affect the results. This indeed happened. In Figure 4 we reevaluate the domains of radiative instability in the temperature-density plane for different effective temperatures. As continuum opacity we use the version given by Alexander et al. (1989). We also find that there is always a main "SiO radiative instability island" most pronounced at low densities and relatively high effective temperatures. For $W = 0.5$, log $\rho = -15$, and $T_{\rm eff} = 2000$ K, we find a bifurcation of temperature between 1980 K and 2380 K. For $T_{\rm eff} = 3000$ K, the bifurcation is between 2040 K and 2400 K, and for $T_{\rm eff} = 4000$ K, it is between 2250 K and 2750 K. These results show that the bifurcation due to SiO occurs at much lower temperatures than the bifurcation due to CO previously discussed. Our results are nevertheless substantially different from those of MNS. MNS also found main SiO radiative instability islands, but found also regions of "thermal runaway" below temperatures of $\simeq 1200$ K. These regions do not exist (or are drastically reduced in size) when CO molecules are included.

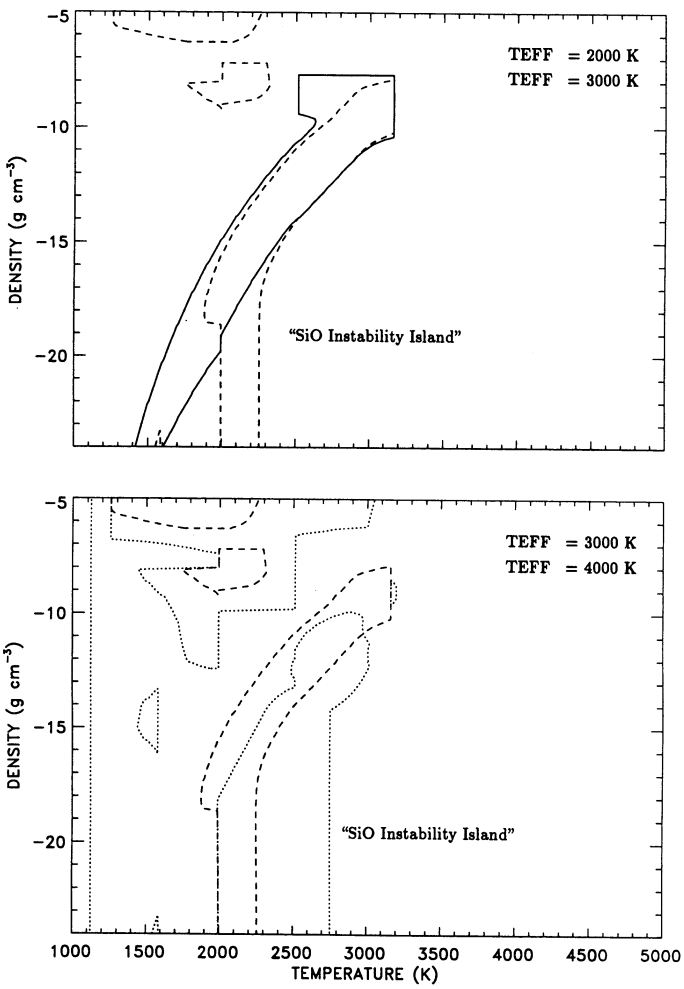

FIG. 4.—Domains of radiative instability in the temperature-density plane for different effective temperatures considering SiO and continuum opacity, but no CO. The upper figure shows the results for 2000 K (*solid lines*) and 3000 K (*dashed lines*), whereas the lower figure shows the results for 3000 K (*dashed lines*) and 4000 K (*dotted lines*). In all cases the geometrical dilution factor is $W = 0.5$.

MNS found that the regions of thermal runaway are relatively large and most pronounced at relatively low effective temperatures. This latter result, however, was not caused by the extrapolation of the Kurucz opacities, but by an error in the definition of the temperature-dependent frequency weight functions.

We also have explored the transition between the regions of radiative instability due to CO and those caused by the joint influence of CO and SiO. We computed a grid of models reducing the CO LTE number density by $f_{CO} = 0.1$, 0.02, and 0.004, respectively. Our results are given in Figure 5. These results are also applicable in the case when CO LTE number densities are maintained, but the CO cooling rate is reduced due to optical depth effects.

3. TIMESCALES

The actual behavior of gas in an expanding circumstellar envelope depends on more than just the radiative cooling rate discussed in § 2. Whether gas is actually cooled (heated) in this way depends on whether there is adequate time, on whether chemical reactions proceed fast enough to maintain LTE in molecular abundances, and on whether hydrodynamic effects occur which significantly alter the temperature. MNS discussed some of these issues; here we present similar estimates for our more general case. We also present estimates of timescales associated with dynamic phenomena, including shock waves of both short and long periods.

Since many ideas in this section are the same as those presented by MNS, but using updated data, the descriptions are abbreviated whenever possible.

3.1. *Radiative, Excitation, and Chemical Timescales*

The radiative cooling time is given by

$$t_{\rm cool} = -\frac{3}{8\pi}\frac{R}{\mu}\int_{T_1}^{T_{\rm RE}} \phi(T, \rho)^{-1}\, dT\,. \qquad (3)$$

Here T_1 denotes the temperature just above the onset of molecular cooling (LTE) and $T_{\rm RE}$ denotes the radiative equilibrium temperature. (All other symbols have there usual meaning.) We also included the geometrical factor of 4π, which was not used in the estimates of MNS. For the calculation of $t_{\rm cool}$ we adopted $T_{\rm eff} = 3500$ K and $W = 0.5$. We found an increase in $t_{\rm cool}$ with decreasing density, changing from 1×10^4 to 1×10^6 s for a density decrease from log $\rho = -10$ to -17.5 (see Table 1). These results depend only marginally on the effective temperature considered. As a test case we used $T_{\rm eff} = 2000$ K and found that $t_{\rm cool}$ changes from 4×10^4 to 3×10^5 s for a density decrease from log $\rho = -12.5$ to -17.5, whereas for log $\rho = -10$, $t_{\rm cool}$ remains the same.

For cooling to actually proceed on this timescale requires that other conditions be fulfilled. Chemical reactions must be fast enough to keep molecular abundances near LTE. This

TABLE 1

CO TIMESCALES (s)

log ρ	$t_{\rm exc}$	$t_{\rm cool}$	$t_{\rm chem}$
−10	1×10^{-9}	1×10^4	2×10^6
−12.5	3×10^{-7}	6×10^4	5×10^8
−15	1×10^{-4}	2×10^5	2×10^{11}
−17.5	3×10^{-2}	1×10^6	5×10^{13}
−20	1×10^1	...	2×10^{16}

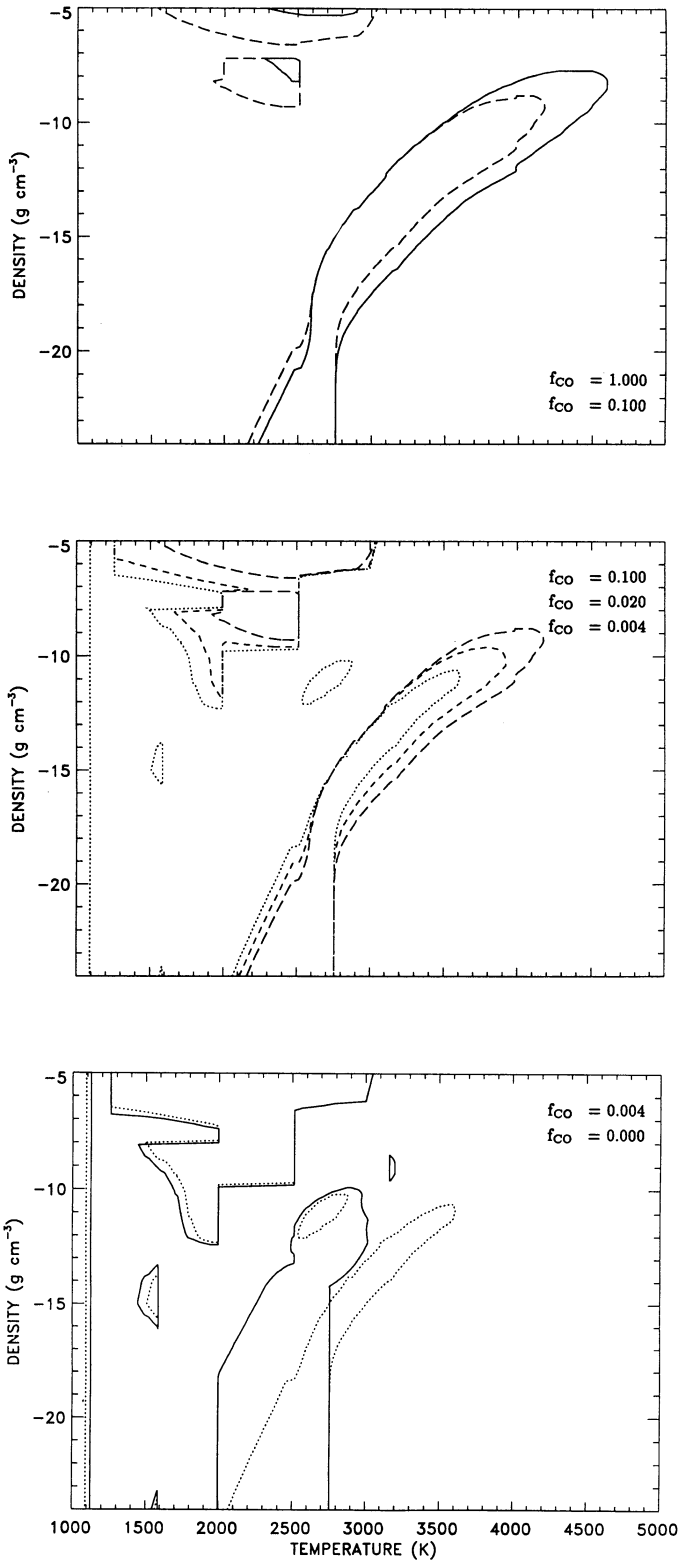

FIG. 5.—Domains of radiative instability in the temperature-density plane for $T_{\rm eff} = 4000$ K and $W = 0.5$ considering CO, SiO, and continuum opacity, where the CO number densities are systematically reduced by factors f_{CO}. The upper figure shows the results for $f_{CO} = 1.0$ (*solid line*) and 0.1 (*long-dashed line*). The middle figure shows the results for $f_{CO} = 0.1$ (*long-dashed line*), 0.02 (*short-dashed line*), and 0.004 (*dotted line*), whereas the lower figure shows the results for $f_{CO} = 0.004$ (*dotted line*) and 0.0 (*solid line*).

condition will be satisfied in the inner parts of a steadily expanding envelope. It will fail farther outward because of low densities and will also fail in the presence of shock waves, if the shocks are strong enough to dissociate molecules and if the period between shocks is not much longer than the reassociation time. A further condition for the cooling mechanism to work is that collisional excitation of molecular IR bands must be fast enough to redistribute energy absorbed in the continuum band. Note that this last condition will always be satisfied in the density regime explored.

The collisional timescale for excitation of the fundamental IR band of SiO was estimated by MNS to be $t_{\rm exc} \sim 10^{-10}\,\rho^{-1}$ g cm^{-3} s at $T = 2500$ K. A similar estimate for CO can be found following Ayres & Wiedemann (1989). The collisional deexcitation rate is then given by

$$C_{10} = P_{\rm H}\,\frac{\exp\,(B_{\rm H} - A_{\rm H}\,T^{-1/3})}{1 - \exp\,(-\Theta/T)} \qquad (4)$$

for collisions with hydrogen atoms. $P_{\rm H}$ is the partial pressure for hydrogen, $A_{\rm H} = 3.0$ and $B_{\rm H} = 18.1$ are the Landau-Teller constants, and $\Theta = 3122$ K is the excitation temperature. For $T = 2500$ K we thus get $t_{\rm exc} = 1/C_{10} \simeq 10^{-19}\,\rho^{-1}$ cm^3 g^{-1} s.

The timescale for chemical reactions, i.e., the formation of CO, is much more difficult to approximate. MNS, who discussed the formation of SiO, relied on the direct reaction Si + O → SiO and used a rate constant of $k_1 = 10^{-17}$ cm^3 s^{-1}. Adopting the same concept for the formation of CO results in a rate constant of $k_1 \simeq 2 \times 10^{-17}$ cm^3 s^{-1} (Dalgarno, Du, & You 1990). As chemical timescale we thus get

$$t_{\rm chem} = \ln 2/(k_1\,N_0) = 2 \times 10^{-4}\,\rho^{-1}\ {\rm cm}^3\ {\rm g}^{-1}\ {\rm s}\,. \qquad (5)$$

The results are given in Table 1. We note, however, that this concept is questionable since in outer atmospheres of cool stars molecules are typically formed in a more complex way. The actual number density of CO is found by solving complicated networks of chemical reaction chains. In the case of oxygen-rich circumstellar envelopes, detailed computations have been given by Gail & Sedlmayr (1986), Glassgold & Huggins (1986), Mamon, Glassgold, & Omont (1987), and Beck et al. (1992), among others. In all cases it was found that $t_{\rm chem}$ is much shorter than that given in Table 1. This is particularly true in systems in which molecular reaction rates are modified by the presence of a chromospheric (= UV) radiation field (e.g., Glassgold & Huggins 1986, Beck et al. 1992).

3.2. *Hydrodynamic Timescales*

The microscopic timescales should be compared with the hydrodynamical timescales. This is simple in cases of radially moving gas in which no time-dependent effects occur. There are then two timescales associated with the expansion of the envelope: the wind acceleration timescale given by $t_{\rm acc} = (\partial u_m/\partial r)^{-1}$, where u_m denotes the mean (time-averaged) velocity of the envelope, and the expansion timescale, or dilution time, which is $t_{\rm dil} = r/2u_m$. The latter one corresponds in principle to $t_{\rm dyn}$ used in MNS. Of the two, $t_{\rm dil}$ is more relevant since it determines the rate of change of the density.

In the case of shocks which propagate through an expanding envelope, other timescales also become relevant. There is, for instance, the repetition time of the shocks which is t_P. In cases of monochromatic shock waves, t_P can be identified with the shock wave period. Shocks are usually described by the Mach number M_s, which is

$$M_s = \frac{U_s - u_1}{c_1}\,, \qquad (6)$$

where U_s is the shock velocity, u_1 is the preshock gas velocity, and c_1 is the preshock sound velocity. The postshock velocity u_2 is given by

$$u_2 = \frac{\Theta - 1}{2\Theta}\,M_s c_1\,. \qquad (7)$$

Also note that c_1/c_m is given by

$$\frac{c_1}{c_m} = \sqrt{\frac{1+\Theta}{1+\Phi}}\,, \qquad (8)$$

where $\Phi = p_2/p_1$ and $\Theta = \rho_2/\rho_1$ denote the ratios of pressure and density across the shock, respectively. The subscript m indicates quantities representative of the mean atmosphere, i.e., time-averaged conditions. For more information, see Gail, Cuntz, & Ulmschneider (1990). Note, however, that we now have defined the ratio c_m/c_1 somewhat differently. We also assumed that the shocks propagate through gas with an adiabatic exponent γ. In the case of constant μ and γ, Φ and Θ are functions of the shock strength M_s only (e.g., Landau & Lifshitz 1975). Quasi-analytic expressions for cases of variable μ and γ have been given by Nieuwenhuijzen et al. (1993) and others.

A further timescale which is relevant in the presence of shocks is the time for quasi-adiabatic cooling (= *hydrodynamic refrigeration*) referred to as $t_{\rm hvd}$. $t_{\rm hvd}$ is the characteristic time for changes in the volume *between* the shocks due to the combined effects of the overall wind expansion and the shock waves. It is given by

$$t_{\rm hyd} = \left(\frac{\partial u}{\partial r}\right)^{-1} = \left(\frac{1}{t_{\rm acc}} + \frac{1}{\langle P\rangle}\,\frac{2u_2}{c_m}\right)^{-1}\,, \qquad (9)$$

where $\langle P\rangle$ is the mean wave period. In our quantitative estimates given in Table 2 we have chosen an envelope appropriate to Betelgeuse, for which we adopted the wind model of Hartmann & Avrett (1984) as a mean atmosphere. The value of $t_{\rm hyd}$ depends on the mean sound speed, the shock strength, and the wave period. It is also not density sensitive. Since $t_{\rm hyd}$ refers to changes of hydrodynamic and thermodynamic quantities *between* the shocks, the relevant timescale is therefore min (t_P, $t_{\rm hyd}$). We assumed values for the wave periods based on the following considerations: We take advantage of the fact that the inner part of the Betelgeuse envelope is chromosphere-like and probably dominated by heating through short-period shock waves generated in the underlying convection zone. Such waves have a frequency spectrum with an estimated

TABLE 2

HYDRODYNAMIC TIMESCALE (s)

				$t_{\rm hyd}$	
log ρ	$t_{\rm dil}$	$t_{\rm acc}$	t_P	$M_s = 1.5$	$M_s = 5.0$
−10.....	6×10^{12}	4×10^{10}	2×10^{6}	4×10^{6}	1×10^{6}
−12.5.....	2×10^{10}	2×10^{8}	2×10^{6}	4×10^{6}	1×10^{6}
−15.....	1×10^{8}	4×10^{7}	2×10^{6}	3×10^{6}	1×10^{6}
−17.5.....	3×10^{8}	4×10^{9}	...	...	...
−20.....	5×10^{9}	...	3.6×10^{7}	7×10^{7}	3×10^{7}

maximum at about 2×10^6 s (Cuntz 1992), which varies proportional to gravity g^{-1}, which is, however, uncertain. This period was used for the estimates at log ρ of -10, -12.5, and -15. Farther out in the envelope, short-period waves are unimportant and shocks with longer periods dominate. In these regions, wave periods corresponding to stellar pulsation modes are appropriate. For the tabulated values at log $\rho = -20$ we adopted a period of 3.6×10^7 s ($\simeq 1.15$ yr), which was identified by Dupree et al. (1987) both in photospheric and chromospheric studies. Note that evaluating hydrodynamic timescales in a mean (time-averaged) atmosphere would produce results which are meaningless.

Note that equation (9) is strictly applicable only in cases in which no major shock-shock interaction occurs. In monochromatic wave models $\langle P \rangle$ can also be identified with t_P. In models with significant shock-shock interaction the situation is more complex: *both* the mean wave period $\langle P \rangle$ *and* the post-shock velocity u_2 tend to increase with atmospheric height (see, e.g., Cuntz 1987, 1992). Nevertheless, equation (9) is still useful in this case in order to get an order-of-magnitude estimate for the prevailing hydrodynamic timescale.

5. AN APPLICATION TO ATMOSPHERE MODELS OF BETELGEUSE

Effects of radiative instabilities are most obvious when applied to existing outer atmosphere models of cool giants and supergiants. As tutorial examples we make use of models of Betelgeuse given by Hartmann & Avrett (1984) and Cuntz (1992). Figure 6 shows domains of radiative instability in the temperature-density plane appropriate to Betelgeuse together with theoretical models. We used 3900 K as effective temperature as proposed by Tsuji (1989) and used $W = 0.5$ and $W = 0$ as geometrical dilution factors. Note that neither the difference due to the two different geometrical dilution factors considered nor the uncertainty in the effective temperature of Betelgeuse impact our results (see also Fig. 3 showing the effect of different effective temperatures on the position of the instability islands). Similar to Figure 2 of MNS we show the Alfvén wave–driven wind model of Hartmann & Avrett (1984) in Figure 6*a*. We find that the Hartmann & Avrett wind model crosses the main CO instability island 3 times. MNS also found that the Hartmann & Avrett model crosses the SiO instability island 3 times, which is, however, located at somewhat lower temperatures. The main difference with MNS, however, is the following: MNS found that the Hartmann & Avrett model also penetrates the SiO cooling instability region located at temperatures below $\simeq 1200$ K. When assumed that also CO molecules are present, we find that this region does not exist or is drastically reduced in size. As a consequence, a "thermal runaway" is now impossible or unlikely to occur—contrary to MNS, who argued that the SiO cooling instability in this area could potentially lead to the formation of dust. This was an extremely significant finding, because most dust-driven wind theories lack a mechanism which carries a sufficient amount of matter to the dust formation radius.

Figures 6*b* and 6*c* show two time steps of the time-dependent ab initio model of Cuntz (1992) for the outer atmosphere of Betelgeuse. The first time step is given by a monochromatic wave model (almost no shock-shock interaction), whereas the second time step is given by a stochastic wave model (strong shock-shock interaction). The time span between the models is 6.69×10^7 s. Both models look extremely complicated when plotted in the temperature-density plane. We see that for a

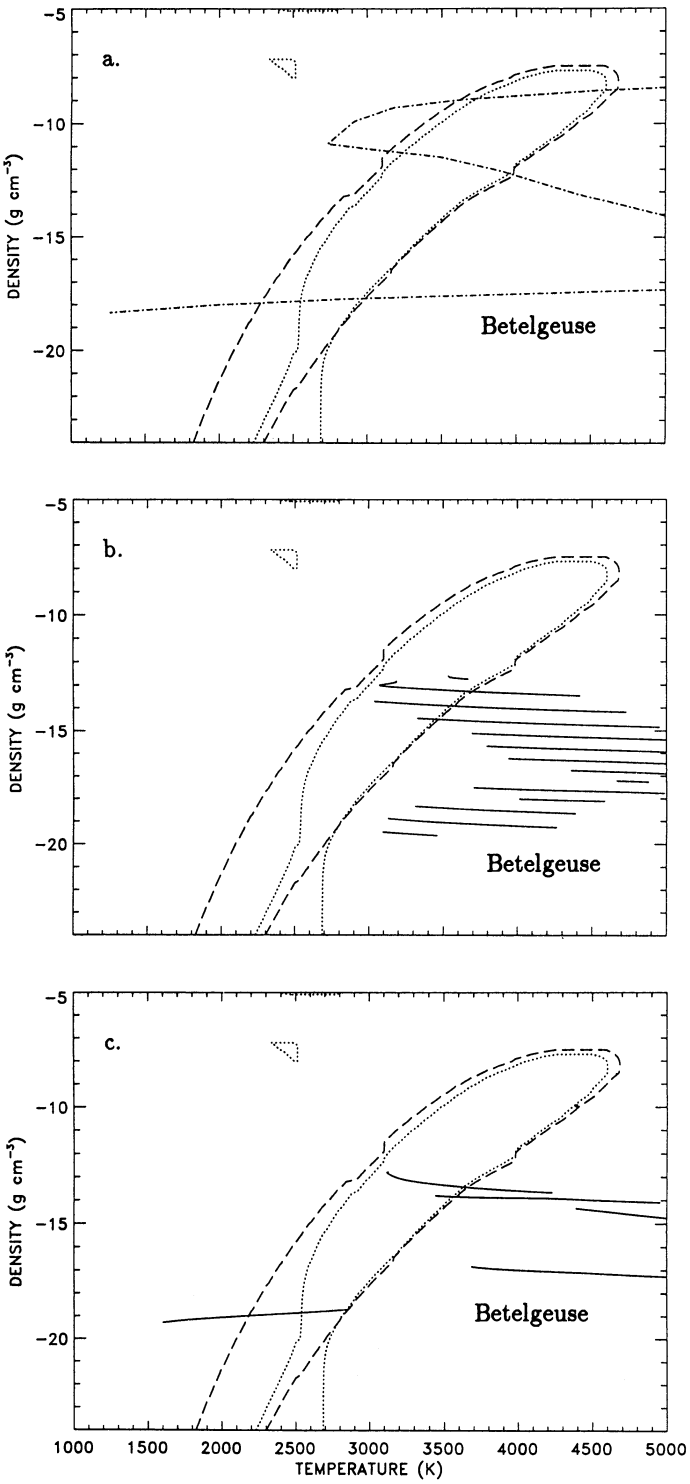

FIG. 6.—Domains of radiative instability in the temperature-density plane appropriate to Betelgeuse for $W = 0.5$ (*dotted lines*) and $W = 0.0$ (*dashed lines*) based on CO, SiO, and continuum opacity together with theoretical models from the literature. Fig. 6*a* shows the wind model of Hartmann & Avrett (1984) (*dash-dotted line*), whereas Figs. 6*b* and 6*c* show two time steps of the time-dependent ab initio model of Cuntz (1992) (*solid lines*). The time span between the models is 6.69×10^7 s. The first time step is given by a monochromatic wave computation (almost no shock-shock interaction), whereas the second model is produced by stochastic waves (strong shock-shock interaction). In Figs. 6*b* and 6*c* reading from top to bottom, the solid lines represent increasing height in the models; the breaks in the lines show the positions of shocks.

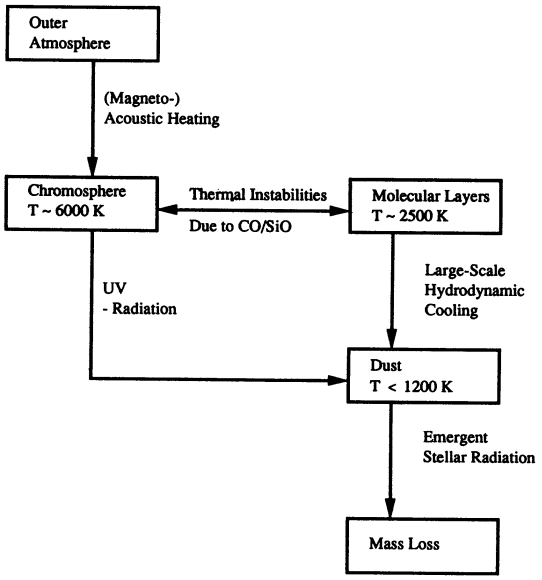

FIG. 7.—The revised paradigm for dust formation in cool evolved stars

fixed temperature many different values for the realized atmospheric density exist, which is ultimately a consequence of the atmospheric shock structure. Both models penetrate the CO radiative instability island many times. However, in the monochromatic wave model (Fig. 6*b*) the repetition time of the shocks is only $\simeq 2 \times 10^6$ s, so we cannot expect that a radiative instability would have time to develop. The model given in Figure 6*c* relies on stochastic waves. This model is characterized by a complicated hydrodynamic and thermodynamic structure with time-dependent episodes of energy and momentum deposition. The direction of the flow alternates between infalling and outflowing motions depending on the shock strength and the radiation-hydrodynamic history of the flow. The prevailing hydrodynamic timescales in this model are typically larger than those in the monochromatic case. The models are also characterized by quasi-adiabatic cooling (or hydrodynamic refrigeration), which reduces the temperature by 1000–1500 K below the radiative equilibrium temperature at a length scale of several pressure scale heights. The associated timescales range from 2×10^6 s for small-scale atmospheric motions to 5×10^7 s,[3] which is the characteristic time between episodic inflow and outflow events. These timescales are consistent with observational constraints (see summary of Querci & Querci 1986). Similar timescales were also seen in a recent study of Drake et al. (1992), who found evidence for thermal instabilities in radio continuum observations. Also note that the models of Cuntz (1992) treat only energy losses (and gains) due to H^- and Mg II, Ca II, and Fe II emission lines and do not include molecules at all. Nevertheless, temperatures as low as 1200 K are reached. This occurs without the influence of radiative instabilities, but it can be expected that large-scale quasi-adiabatic cooling *and* radiative instabilities might act in a synergic manner.

Based on all these results we present a new paradigm for dust formation (see Fig. 7), which is particularly relevant in cases where dust is formed relatively close to the stellar photosphere. The paradigm also considers hydrodynamic cooling. It also takes into account the effect of ionizing chromospheric UV radiation, which can drastically enhance the efficiency of dust formation. A detailed discussion of this effect in the case of Betelgeuse was recently given by Beck et al. (1992). A previous version of the dust formation paradigm based on observational constraints and also considering the effect of ionizing UV radiation has been given by Stencel, Carpenter, & Hagen (1986). A paradigm which also takes into account the impact of stellar evolution processes was given by Cuntz & Stencel (1992).

[3] Note that this value was incorrectly given in Cuntz (1992) as 5×10^6 s.

5. CONCLUSIONS

We have revisited the formation of radiative instabilities in cool star atmospheres due to CO, SiO, and continuum opacity and compared our results with those of MNS. Our results are based on the analysis of the energy balance of gas elements with prescribed thermodynamic properties. The radiation in the SiO band was assumed to be optically thin. In the case of the CO band we also considered the optically thin limit and explored the potential impact of optical depth effects by considering reduced CO cooling rates. For O, C, and Si we assumed solar abundances. The impact of relevant timescales has been evaluated. We also discussed possible effects associated with the dynamics of long-period and short-period shock waves. Our major findings include the following:

1. Radiative instabilities exist for a broad range of thermodynamic conditions, which are present in cool star atmospheres. In many cases, they are caused mainly by CO molecules and to a much smaller extent by SiO.
2. In the framework of our study we found that this conclusion is altered only when the CO cooling rate is substantially reduced. This can occur either due to a reduced CO particle density caused by (time-dependent) non-LTE reaction rates or by an extreme Si overabundance/C underabundance, or, alternatively, in cases when the CO band is sufficiently optically thick, whereas the SiO band is optically thin. This latter case is probably the most important exception.
3. We cannot rule out entirely that molecules other than CO and SiO also impact the onset of radiative instabilities. These molecules must fulfill two criteria: First, the molecules must be abundant. Second, the molecular spectrum must consist of distinct bands which lie at largely different wavelengths or, alternatively, the spectrum must contain a large number of lines which produce a considerable temperature variation of the cooling rate. A molecule which possibly matches these criteria is H_2O, which is an important constituent of atmospheres of O-rich stars later than M5.
4. Note that molecules which do not meet these criteria can nevertheless provide a significant contribution to the overall atmospheric cooling. Note, however, that they can impact the generation of radiative instabilities only when other constituents such as H^- are present. These molecules can therefore be treated as part of the continuum opacity.
5. We have evaluated the domains of radiative instability in the temperature-density plane for different stellar effective temperatures assuming optically thin radiation in the CO and SiO bands, LTE number densities for CO and SiO, and solar abundances for O, C, and Si. We found that there is always a main CO instability island most pronounced at low densities and relatively high effective temperatures. The influence of the geometrical dilution factor is found to be less important, except for $T_{\rm eff} \gtrsim 4500$ K and at distances $R \lesssim 1.3\ R_*$. In some cases, additional regions of instability appear, which are attributable to the behavior of the continuum opacity. This latter result suggests that radiative instabilities are not always necessarily a consequence of frequency-dependent effects.

6. We compared our results with those given by MNS. MNS have treated the interplay between optically thin continuum and SiO band opacity and also found regions of radiative instability. They found main SiO radiative instability islands" and regions of thermal runaway below temperatures of $\simeq 1200$ K, which also depend on the effective temperature of the star. These results can still be viewed as relevant, when CO molecules exist but the CO band opacity is substantially reduced. Note, however, that MNS have given the regions of thermal runaway somewhat incorrectly in part due to the extrapolation of an early version of the Kurucz opacity to very low temperatures and pressures.

7. The regions of "thermal runaway" do not exist (or are drastically reduced in size) when CO molecules are included (assuming LTE number densities and that the CO band radiation is treated in the optically thin limit). This result provides strong evidence that dust formation cannot occur via radiative instabilities alone. As an alternative mechanism we point to large-scale quasi-adiabatic hydrodynamic cooling, which was studied in model computations. We also suggest that cooling via radiative instabilities and cooling via hydrodynamic effects can act in a synergic manner.

8. Even in cases, in which SiO molecules are unimportant for generating radiative instabilities, SiO molecules can still be considered as a tracer for "very cool" atmospheric layers. This result should be considered as relevant for the interpretation of SiO maser emission and its relationship to the onset of dust formation.

9. We have applied our results to atmospheric models of Betelgeuse including those given by Cuntz (1992), which are based on stochastic shock waves. The models of Cuntz show temperatures 1000–1500 K below the radiative equilibrium temperature at a length scale of several pressure scale heights due to large-scale quasi-adiabatic hydrodynamic cooling. The associated timescales range from 2×10^6 s for small-scale atmospheric motions to 5×10^7 s, which is the characteristic time between episodic inflow and outflow events. We note that these timescales are consistent with observational constraints (see summary of Querci & Querci 1986). Similar timescales were also seen in a study of Drake et al. (1992), who found evidence for thermal instabilities in radio continuum observations of Betelgeuse. Also note that the timescale for hydrodynamic cooling is not density sensitive, which is extremely helpful in high-density regimes.

10. We point out that shock-shock interaction in a field of short-period waves is a crucial feature for radiative instabilities to occur. The reason is that shock-shock interaction tends to produce shock repetition times, which are occasionally much larger than the radiative cooling time of the gas.

11. Based on all these results we present a new paradigm of dust formation, which is particularly relevant in cases where dust is observed to form relatively close to the stellar photosphere. Note, however, that some of our conclusions are not final. It would be important to compute better models for stellar atmospheres, which self-consistently treat effects such as thermal bifurcation, the propagation and interaction of stochastic shocks, and thermodynamic effects associated and interaction with nonequilibrium chemistry.

12. As a first step toward this goal, it would be interesting (*i*) to also include other molecules besides CO and SiO in a frequency-dependent manner, (*ii*) to also study the impact of improved continuum opacities given by Kurucz (1994), and (*iii*) to self-consistently include thermal instabilities in recent ab initio cool star chromosphere models.

Note added in manuscript.—In a very recent paper, Wiedemann et al. (1994) presented an observational study of strong vibration-rotation lines of CO near 4.6 μm, which can serve as a diagnostic tool for the thermal conditions in late-type stars near and above the temperature minimum region in chromospheric models. Wiedemann et al. studied eight stars with luminosity class II–III and spectral type G2–K5. They found that the CO models feature a steady decrease in temperature with height in contradiction to chromospheric models, which are based on the analysis of the Ca II and Mg II emission lines. The authors conclude that thermal bifurcation generated by radiative instability effects might be able to reconcile these contradicting scenarios. In another paper, Carpenter et al. (1994) analyzed the O I and C I resonance lines in the UV spectrum of α Orionis (M2 Iab) obtained with the Goddard High Resolution Spectrograph. The authors also give results from preliminary model computations for the line formation processes. These results provide preliminary evidence that dust is present within the chromosphere, which would indicate gross inhomogeneities which are possibly a consequence of radiative instability effects.

M. C. acknowledges financial support through the Advanced Study Program at the National Center for Atmospheric Research and through NASA grant NAGW-2904 to the University of Colorado.

REFERENCES

Alexander, D. R., Augason, G. C., & Johnson, H. R. 1989, ApJ, 345, 1014
Alexander, D. R., Johnson, H. R., & Rypma, R. L. 1983, ApJ, 272, 773
Athay, R. G., & Dere, K. P. 1990, ApJ, 358, 710
Ayres, T. R. 1981, ApJ, 244, 1064
———. 1986, ApJ, 308, 246
Ayres, T. R., & Testerman, L. 1981, ApJ, 245, 1124
Ayres, T. R., Testerman, L., & Brault, J. W. 1986, ApJ, 304, 542
Ayres, T. R., & Wiedemann, G. R. 1989, ApJ, 338, 1033
Beck, H. K. B., Gail, H.-P., Henkel, R., & Sedlmayr, E. 1992, A&A, 265, 626
Bester, M., Danchi, W. C., Degiacomi, C. G., Townes, C. H., & Geballe, T. R. 1991, ApJ, 367, L27
Brugel, E. W., Willson, L. A., Bowen, G. H., & Magalhães, M. 1991, unpublished poster paper, presented at Cool Stars, Stellar Systems, and the Sun, Seventh Cambridge Workshop
Carpenter, K. G., Robinson, R. D., Judge, P. G., Ebbets, D. C., & Brandt, J. C. 1994, in Cool Stars, Stellar Systems, and the Sun, Proc. Eighth Cambridge Workshop, ed. J.-P. Caillault, in press
Clayton, G. C., Whitney, B. A., Stanford, S. A., & Drilling, J. S. 1992, ApJ, 397, 652
Cuntz, M. 1987, A&A, 188, L5
Cuntz, M. 1992, in Cool Stars, Stellar Systems, and the Sun, Proc. Seventh Cambridge Workshop, ed. M. Giampapa & J. Bookbinder (ASP Conf. Ser., 26), 383
Cuntz, M., & Muchmore, D. O. 1989, A&A, 209, 305
Cuntz, M., & Stencel, R. E. 1992, in Cool Stars, Stellar Systems, and the Sun, Proc. Seventh Cambridge Workshop, ed. M. Giampapa & J. Bookbinder (ASP Conf. Ser., 26), 451
Dalgarno, A., Du, M. L., & You, J. H. 1990, ApJ, 349, 675
Deming, D., Hillman, J. J., Konstiuk, T., Mumma, M. J., & Zipoy, D. M. 1984, Sol. Phys., 94, 57
Drake, S. A., Bookbinder, J. A., Florkowski, D. R., Linsky, J. L., Simon, T., & Stencel, R. E. 1992, in Cool Stars, Stellar Systems, and the Sun, Proc. Seventh Cambridge Workshop, ed. M. Giampapa & J. Bookbinder (ASP Conf. Ser., 26), 455
Dupree, A. K., Baliunas, S. L., Guinan, E. F., Hartmann, L., Nassiopoulos, G. E., & Sonneborn, G. 1987, ApJ, 317, L85
Elitzur, M. 1992, ARA&A, 30, 75
Elitzur, M., Brown, J. A., & Johnson, H. R. 1989, ApJ, 341, L95

Gail, H.-P., Cuntz, M., & Ulmschneider, P. 1990, A&A, 234, 359
Gail, H.-P., & Sedlmayr, E. 1986, A&A, 166, 225
Glassgold, A. E., & Huggins, P. J. 1986, ApJ, 306, 605
Glassgold, A. E., & Langer, W. D. 1976, ApJ, 204, 403
Hartmann, L., & Avrett, E. H. 1984, ApJ, 284, 238
Heasley, J. N., Ridgway, S. T., Carbon, D. F., Milkey, R. W., & Hall, D. N. B. 1978, ApJ, 219, 970
Huber, K. P., & Herzberg, G. 1979a, Constants of Diatomic Molecules (New York: Van Nostrand Rheinhold)
———. 1979b, Molecular Spectra and Molecular Structure, 4 (New York: Van Nostrand Rheinhold)
Jørgensen, U. G., & Johnson, H. R. 1992, A&A, 265, 168
Kneer, F. 1983, A&A, 128, 311
Kurucz, R. 1979, ApJS, 40, 1
———. 1983, private communication
———. 1994, in IAU Colloq. 146, Molecules in the Stellar Environment, Lecture Notes in Physics, 428, ed. U. G. Jørgensen (Berlin: Springer), 282
Landau, L. D., & Lifshitz, E. M. 1975, Fluid Mechanics (London: Pergamon)
Le Bourlot, J., Pineau des Forêts, G., Roueff, E., & Schilke, P. 1993, ApJ, 416, L87
Maciel, W. J. 1976, A&A, 48, 27
———. 1977, A&A, 57, 273
Mamon, G. A., Glassgold, A. E., & Omont, A. 1987, ApJ, 323, 306
Mauas, P. J., Avrett, E. H., & Loeser, R. 1990, ApJ, 357, 279
Mihalas, D. 1978, Stellar Atmospheres (San Francisco: W. H. Freeman & Company)
Muchmore, D. O. 1986, A&A, 155, 172
———. 1990, in Cool Stars, Stellar Systems, and the Sun, Proc. Sixth Cambridge Workshop, ed. G. Wallerstein (ASP Conf. Ser., 9), 39
Muchmore, D. O., Nuth, J. A., III, & Stencel, R. E. 1987, ApJ, 315, L141 (MNS)
Muchmore, D. O., & Ulmschneider, P. 1985, A&A, 142, 393
Nieuwenhuijzen, H., De Jager, C., Cuntz, M., Lobel, A., & Achmad, L. 1993, A&A, 280, 195
Querci, M., & Querci, F. 1986, in Cool Stars, Stellar Systems, and the Sun, Proc. Fourth Cambridge Workshop, ed. M. Zeilik & D. M. Gibson (Berlin: Springer), 492
Rossi, S. C. F., Maciel, W. J., & Benevides-Soares, P. 1985, A&A, 148, 93
Stencel, R. E. 1993, in Conference on Astrophysical Masers, ed. A. Clegg & A. Nedoluha (Berlin: Springer), 449
Stencel, R. E., Carpenter, K. G., & Hagen, W. 1986, ApJ, 308, 859
Tsuji, T. 1988, A&A, 197, 185
———. 1989, in The Evolution of Peculiar Red Giant Stars, ed. H. R. Johnson & B. Zuckerman (Cambridge: Cambridge Univ. Press), 81
Wiedemann, G., & Ayres, T. R. 1990, in Cool Stars, Stellar Systems, and the Sun, Proc. Sixth Cambridge Workshop, ed. G. Wallerstein (ASP Conf. Ser. 9), 158
Wiedemann, G., Ayres, T. R., Jennings, D. E., & Saar, S. H. 1994, ApJ, 423, 806