ANNEXE

ANNEXE 1: ASTROPHYSICAL JOURNAL, 483, 464-474, 1997

LEBLANC F. AND HUBERT D., A GENERALIZED MODEL FOR THE PROTON EXPANSION IN ASTROPHYSICAL WINDS. I. THE VELOCITY DISTRIBUTION FUNCTION REPRESENTATION, ASTROPHYSICAL JOURNAL, 483, 464-474, 1997.

THE ASTROPHYSICAL JOURNAL, 483:464–474, 1997 July 1

A GENERALIZED MODEL FOR THE PROTON EXPANSION IN ASTROPHYSICAL WINDS. I. THE VELOCITY DISTRIBUTION FUNCTION REPRESENTATION

F. LEBLANC AND D. HUBERT
Département de Recherche Spatiale, CNRS URA 264, Observatoire de Paris
Received 1996 October 1; accepted 1997 January 28

ABSTRACT

We construct a new approach to model the velocity distribution function (VDF) for the protons in stellar atmosphere expansions or planetary polar winds. The generalized Grad method of construction is used, and comparisons with the bi-Maxwellian polynomial expansion model are made in applications to the solar wind in the context of the measurements made by the *Helios* probes between 0.3 and 1 AU. A fitting procedure based on a sum of two Maxwellian functions is used to check the convergence property of both polynomial expansions and to calculate the predicted polynomial expansion profiles along the magnetic field orientation for typical proton VDFs in the solar wind. The generalized model is better adapted than the bi-Maxwellian polynomial expansion function to reproduce the long-tail features of a majority of the observed proton VDFs; moreover, our model does not display negative values of the VDF, contrary to the bi-Maxwellian expansion for normalized heat flux larger than unity. A 16 moment approximation, which corresponds to a third order of development, allows us to provide an associated set of generalized transport equations better closed than the equivalent system associated with a bi-Maxwellian polynomial expansion.

Subject headings: plasmas — solar system: general — solar wind — stars: mass loss

1. INTRODUCTION

The solar corona is the source region of the solar wind, a supersonic flow composed of electrons, protons, and heavy ions that carves the heliosphere. The polar wind is a plasma flow of electrons, protons, and O^+ ions, escaping from the polar terrestrial ionosphere along the magnetic field lines. Prototypes of stellar atmospheres expanding into the interstellar space, they expand from collision-dominated regions, through transition regions, to collisionless regions. To model such plasmas in their collision-dominated regions, an isotropic hydrodynamic approach (Chapman 1965; Parker 1958) is generally well adapted because the collisions between particles are sufficiently numerous to thermalize the flow. It consists of the resolution of a system of transport equations for the macroscopic parameters. In collisionless regions, kinetic models (Chamberlain 1960; Lemaire & Scherer 1971) are much more adapted than fluid models. Indeed, particles in such a medium move nearly independently of the others. These methods generate exospheric models based on a microscopic description of the flow. But in transition regions, both models fail because, with a Knudsen number (which is defined as the ratio of the particle mean free path to the density scale height) roughly equal to one (Shizgal 1977), interactions between particles are sufficient to alter pure collisionless approaches and not enough to justify a pure collision-dominated state. For instance, the proton temperature anisotropy is equal to unity for all the hydrodynamic models, while the kinetic approaches provide proton temperature anisotropy $(T_\parallel/T_\perp)_p$ between 10 and 900 (Lemaire & Scherer 1973), whereas observations in the solar wind reveal values of 2 ± 1.

Therefore, another approach is needed to model plasma expansions in these intermediate regions. Grad (1958) has proposed a solution of the Boltzmann equation, the so-called moment approach, in which the distribution function is a polynomial expansion with a weight function, which is the local equilibrium Maxwellian distribution function. Used in near–local-equilibrium states, it is based on the resolution of an associated system of transport equations for the velocity moments of the distribution function. Thus, Cuperman, Weiss, & Dryer (1980; 1981) have derived a set of transport equations up to the fourth order of the velocity moments for spatially inhomogeneous, anisotropic, multispecies, and spherically symmetric systems of particles obeying an inverse-square law of interactions. They have calculated the velocity collisional transfer by the Fokker-Planck operator based also on a local Maxwellian polynomial expansion. But for plasma flows in states far from the local equilibrium, such as, for instance, with large temperature anisotropies, these microscopic descriptions seem not to be available. Then, an improved solution is to take into account as first principles the properties of the nonequilibrium system. Chew, Goldgerger, & Low (1956) were the first to derive transport equations based on a bi-Maxwellian species distribution function. Transport effects such as heat flux and viscosity for Coulomb collisions have been included by Chodura & Pohl (1971) and Oraevskii, Chodura, & Feneberg (1968). Finally, Demars & Schunk (1979) extended this development to arbitrary degrees of ionization and built the 16 moment set of transport equations (using velocity moments until order 3 in gyrotropic microscopic states). They have shown in applications to the solar and polar winds (Demars & Schunk 1989, 1990, 1991) the great interest of such models that provide a microscopic and a macroscopic description in relatively good agreement with observations.

Grad's solution cannot describe states with large temperature anisotropies because such a polynomial expansion based on an isotropic Maxwellian function is only allowed to generate small deviations from isotropic states. In the same way, a polynomial expansion with a bi-Maxwellian function as a weight factor cannot generate large deviations from the bi-Maxwellian distribution function. For instance, it cannot be used to describe plasma flows that are characterized by large, magnetic field–aligned heat flux because a bi-Maxwellian function is not well adapted to generate such

heat flux. Several authors have tried to determine the fields of application of the Demars & Schunk (1979) model. Palmadesso, Ganguli, & Mitchell (1988) have found unstable waves modes when this model is used for large-scale magnetospheric-ionospheric dynamics and have defined a threshold that depends on the magnetic field–aligned heat flux $q_{\|p}$ normalized by the free-streaming heat flux (which is the thermal energy of ions $P_{\|p}$ times the magnetic field–aligned thermal speed $V_{\mathrm{th}\|p}$) and is $q_{\|p} \leq 0.44(P_{\|} V_{\mathrm{th}\|})_p$. Above this limit the bi-Maxwellian expansion cannot correctly describe such plasmas. Cordier (1994a, 1994b) has studied the hyperbolic nature of the system of transport equations built from a 16 moment polynomial expansion for collisionless plasmas and has obtained very similar limits: $q_{\|p} \leq 0.46(P_{\|} V_{\mathrm{th}\|})_p$. More recently, a general analysis of the microscopic description was developed by Levermore, Morokoff, & Nadiga (1995). It consists of studying the realizability and positivity of any model of velocity distribution functions (VDFs). They have built methods to make sure that the velocity moments and the VDF remain physically realistic. The best illustration of these rules is given by Gombosi & Rasmussen (1991), who has shown that the 16 moment VDF has negative values for too high values of the magnetic field–aligned heat flux; Robineau, Blelly, & Fontanari (1996) have stressed the hypothesis of fluid description and concluded that the transport of energy should contribute only to a small part of the local thermal energy creation. A different criterion has been defined by Mintzer (1965). He analyzed the nature of the convergence of any polynomial expansion generalizing the Grad approach and has obtained a necessary condition on the choice of the weight function in comparison with the exact solution. Then, Hubert (1985) has shown the nonconvergence of a bi-Maxwellian polynomial expansion describing a microscopic model characterized by a long tail, typical of the observed proton velocity distribution of the solar wind.

Considering all these analyses, the results of Demars & Schunk (1990) are open to some criticisms because they have elaborated models for the solar wind proton VDF (Demars & Schunk 1990) from *Helios 2* solar probe data (Marsch et al. 1982). Indeed, this probe has revealed the existence of three main types of three-dimensional proton VDF contours: an isotropic state, an anisotropic state displaying a long tail in the magnetic field orientation, and a double-humped VDF oriented in the same way and associated with large values of temperature anisotropy and heat flux. Furthermore, the heat flux value is higher than the limit defined by previously cited authors [in $q_{\|p} \leq a(P_{\|} V_{\mathrm{th}\|})_p$] with the coefficient a being larger than unity in most of these observed proton distributions, and the bi-Maxwellian polynomial expansion provides unphysical VDF profiles.

This paper is the first part of a work whose purpose is to establish a generalized microscopic solution as well as a generalized fluid theory for the description of time-dependent, spatially nonhomogeneous, anisotropic, and multispecies space plasma winds. Our approach is directly inspired by the generalized polynomial solution of the Boltzmann equation established by Mintzer (1965), in which the most important point for application is the choice of the zeroth-order approximation. Indeed, it is as important that a good approximation can be obtained when the expansion is truncated after a small number of terms as it is for the expansion to converge when the number of terms approaches infinity; moreover, the zeroth-order approximations have to be sufficiently simple so as not to give a complex associated system of coefficients of the polynomial expansion (or, equivalently, a sufficiently simple associated system of the general transport equations). Thus, our approach will be to model the characteristics of the solar wind proton VDF in order to select a zeroth-order approximation from an exact solution of a similar but simplified problem. A similar step has been used by Hubert (1983) in the field of non-Maxwellian auroral phenomena, whose results provide good comparison with respect to Monte Carlo simulations (Hubert & Barakat 1990; Barakat & Hubert 1990).

Section 2 is devoted to the analysis of typical properties of the proton VDF observed in solar wind. Section 3 deals with the mathematical structures of the bi-Maxwellian polynomial expansion. In § 4 we develop our new solution for the microscopic description of the proton in solar and polar winds. Section 5 is a brief discussion of the merits of this approach. A conclusion is presented in § 6.

2. CHARACTERISTICS OF THE PROTON VELOCITY DISTRIBUTION FUNCTION OF THE SOLAR WIND

2.1. *Macroscopic Features*

Marsch et al. (1982) have provided observations of the radial evolution of the macroscopic parameters of the protons between 0.3 and 1 AU. Using the data of the *Helios 2* probe, the evolution of the total temperature and the heat flux has been determined for different speeds of the protons. The total temperature T_p evolves faster for low-speed winds than for high-speed winds. It follows a radial evolution as $R^{-\alpha}$, where $0.76 \leq \alpha \leq 1.33$, and R is the distance to the Sun (Schwenn & Marsch 1991). The proton temperature anisotropy $(T_{\|}/T_{\perp})_p$ increases with heliocentric distance; it evolves between 1 and 2 and is always larger than unity at 1 AU. Furthermore, Marsch has shown that at 0.3 AU, the proton perpendicular temperature $T_{\perp p}$ could be twice as large as the proton parallel temperature $T_{\|p}$. Indeed, $T_{\perp p}$ decreases by 2 orders of magnitude from 0.3 to 1 AU. The evolution of the total heat flux Q_p is also proportional to $R^{-\alpha}$, where $3.8 \leq \alpha \leq 4.7$. The heat flux is larger for slow winds than for fast ones. Finally, the perpendicular heat flux $q_{\perp p}$ is much smaller than the parallel heat flux $q_{\|p}$.

For heliocentric distances larger than 1 AU, data of several probes provide us macroscopic parameter radial variations also (Gazis 1984). The density, as expected, varies as R^{-2}, but the proton temperature evolves slower (following a $R^{-\alpha}$ law with $0.3 \leq \alpha \leq 0.7$; see Maksimovic 1995) than an adiabatic expansion, as it was the case between 0.3 and 1 AU. Gazis, from Voyager 1 data, has provided an estimation of the radial component of the heat flux. He has suggested that Q_p follows a radial evolution as $R^{-2.44}$. The differences between observations made in the regions less distant than 1 AU from the Sun and beyond are usually explained by a heating of the solar wind caused by shocks or interplanetary compression zones mainly observed beyond 1 AU. Moreover, Liu et al. (1995) have explained that somewhere near 10 AU, the observations of the macroscopic parameters are perturbed by the high-speed winds that mix with slow solar winds. Then, at large heliocentric distances, the analysis of the total proton temperature T_p and of the total heat flux Q_p have to take into

account new cooling/heating mechanisms and interpenetrating processes.

To sum up, for all heliocentric distances, observations reveal a nonadiabatic expansion of an anisotropic plasma composed of particles that interact weakly. We will focus our attention on the evolution of the temperature anisotropy and the heat flux.

2.2. *Microscopic Features*

The study (Marsch et al. 1982; Marsch 1982) of the proton VDF has revealed some specific aspects of the solar wind microscopic state. It has shown a wide variety of nonthermal features: gyrotropic contours with respect to the magnetic field direction, a core and halo shape of the proton VDF along the magnetic field, a temperature anisotropy of the core always as $(T_{\|}/T_{\perp})_{\rm core} \leq 1$ while $(T_{\|}/T_{\perp})_{\rm total} \geq 1$ almost everywhere, large values of the magnetic field–aligned heat flux, and even resolved double peaks. The observations have shown three main types of three-dimensional VDFs. The first class is composed of isotropic VDFs and is mainly observed in interplanetary magnetic field sector boundaries. Most of the VDF profiles develop a long tail in the magnetic field orientation (see Fig. 2, below). The third class displays a surprising profile, which reveals a magnetic field–aligned second peak (see Fig. 1).

Observed at all heliocentric distances, this type of shape represents 20% of all the measurements made by the *Helios 2* probe, according to Marsch et al. (1982). Furthermore, a correlation between the Alfvén speed and the relative speed between the two peaks persists during low and high streams, suggesting an Alfvén waves regulation phenomena (Montgomery et al. 1976). The double-humped proton distribution function is a stable feature of the solar wind proton VDF, related to large values of the magnetic field–aligned heat flux.

The crosses in Figures 1 and 2 are a cut through the proton VDF cases I and K (obtained from Marsch data: Marsch et al. 1982) along the magnetic field direction. The case K was measured at an heliocentric distance of 0.39 AU and a solar wind speed of 494 km s^{-1}, and the case I at 0.54 AU and 618 km s^{-1}.

In order to characterize the microscopic state, Feldman (1979) fits the observations with a sum of two Maxwellian functions. In Figures 1 and 2 we show results of such fitting. The dashed and dotted lines fit the core and the halo respectively with a Maxwellian function, while the solid line is the sum of these two Maxwellian functions. Tables 1 and 2 give the value of the fitting variables obtained in the two cases I and K and the derived macroscopic parameters. It also provides the values determined by Marsch et al. (1982) for

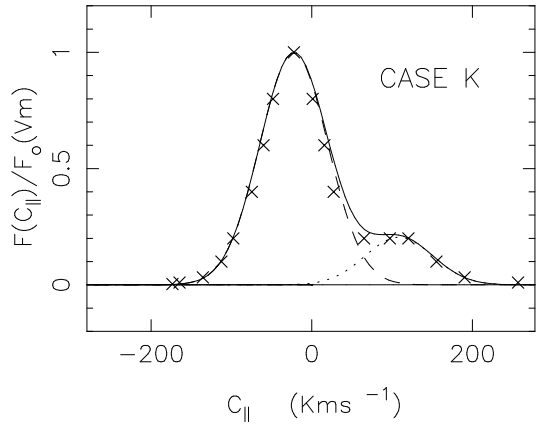

FIG. 1.—Magnetic field–aligned profile of case K issued from Marsch data and fitting by a sum of two Maxwellian functions plotted in the random velocity space. The crosses represent the *Helios* probe measurements, the solid line is the sum of two Maxwellian functions, and the dashed and dotted lines are the independent representation of the two Maxwellian functions composing the sum.

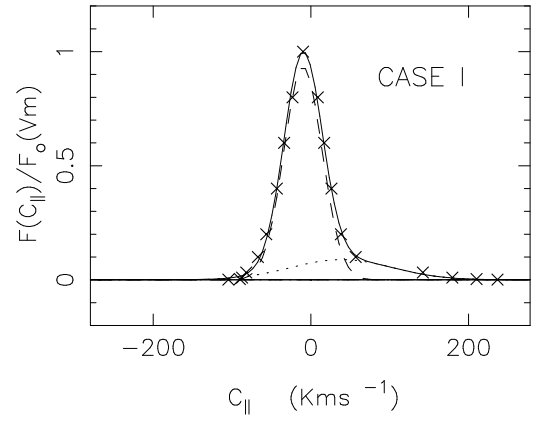

FIG. 2.—Magnetic field–aligned profile and fitting of case I: the plotting format is the same as that for Fig. 1.

TABLE 1

FITTING OF CASE I

FITTING		VELOCITY MOMENTS		
Parameter	Value	Parameter	Fitting Data	Marsch Data
$n_1/(n_1+n_2)$	0.8	n_p (cm^{-3})	14	10.4
T_1 (K)	71,130	V_m (km s^{-1})	9.9	...
T_2 (K)	500,480	$T_{\|\|p}$	204,420	267,000
U_{12} (km s^{-1})	49	$q_{\|\|p} \times 10^{-13}$ (kg m^3 s^{-3})	6.2	2.88
		$n_p m_p \langle c_{\|\|}^4 \rangle \times 10^{-8}$ (kg m^4 s^{-4})	3.27	...
		$n_p m_p \langle c_{\|\|}^5 \rangle \times 10^{-3}$ (kg m^5 s^{-5})	4.16	...
		$n_p m_p \langle c_{\|\|}^6 \rangle \times 10^{2}$(kg m^6 s^{-6})	8.08	...

NOTE.—Results of a sum of two Maxwellian functions fitted to the profile along the magnetic field of case I. The terms n_1, T_1 defined the first Maxwellian function of the sum and T_2, U_{12} the second; U_{12} is the relative velocity between the two Maxwellian functions; $n_2 = 1 - n_1$, T_1, and T_2 are temperatures; V_m is the average velocity relative to the velocity of the maximum phase space density; $T_{\|p}$ is the parallel temperature; and $q_{\|p}$ is the parallel heat flux.

TABLE 2

FITTING OF CASE K

FITTING		VELOCITY MOMENTS		
Parameter	Value	Parameter	Fitting Data	Marsch Data
$n_1/(n_1 + n_2)$	0.83	n_p (cm^{-3})	7.82	26.2
T_1 (K)	217,638	V_m (km s^{-1})	22.2	...
T_2 (K)	215,327	$T_{\|p}$	509,706	566,000
U_{12} (km s^{-1})	131	$q_{\|p} \times 10^{-13}$ (kg m^3 s^{-3})	3.47	6.19
		$n_p m_p \langle c_\|^4 \rangle \times 10^{-7}$ (kg m^4 s^{-4})	9.92	...
		$n_p m_p \langle c_\|^5 \rangle \times 10^{-1}$ (kg m^5 s^{-5})	1.05	...
		$n_p m_p \langle c_\|^6 \rangle \times 10^{+3}$ (kg m^6 s^{-6})	2.25	...

NOTE.—The variables are defined as in Table 1.

comparison. It is important to underline that in our fitting we have only used a contour plot along one direction to calculate the macroscopic parameters, while Marsch has used three-dimensional contour plots. Furthermore, the value of $q_{\|p}$ that we give as the Marsch value is calculated as $q_{\|p} = q_p/(2n_p)$, where q_p is the heat flux density given in the Marsch data, and n_p is the density.

As shown in Figures 1 and 2, a sum of two Maxwellians fits these profiles very well along the magnetic field direction. With this model, we are able to calculate the velocity moments at any order. We have established a great sensitivity of the profile in relation to the velocity moment values. Indeed, a variation of 10% of the heat flux value can generate a second peak in a profile modeled with a sum of two Maxwellian functions. The fitting has to be very accurate for regions of large velocity in order to obtain the best estimate of the velocity moments.

Figures 3*a* and 3*b* are obtained with the Marsch (1982) macroscopic parameters. They represent values of $q_{\|p}$ as a function of the parallel thermal velocity, $V_{\text{th}\|p} = (2k_B T_{\|p}/m_p)^{1/2}$ times pressure term $P_{\|p} = k_B T_{\|p}$ (k_B is the Boltzmann constant, and m_p is the proton mass), which is the parallel free-streaming heat flux. The double-peaked VDF measurements (Fig. 3*b*) are separated into other cases. The solid line represents the relation $q_{\|p} = (P_\| V_{\text{th}\|})_p$. It corresponds to the upper limit defined in the paper of Palmadesso et al. (1988), above which unphysical behaviors

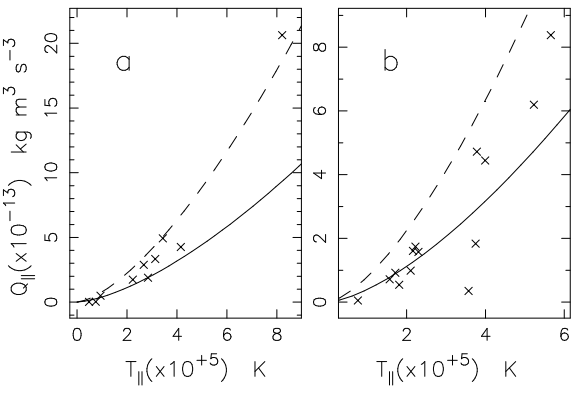

FIG. 3.—Representation of $q_{\|p}$ as a function of $(P_\| V_{\text{th}\|})_p$ for Marsch values: (*a*) the data without a resolved double peak; (*b*) the double-humped cases. The crosses represent the *Helios* probe measurements, the solid line denotes the $q_{\|p} = (P_\| V_{\text{th}\|})_p$ relation, and the dashed line denotes the constraints of construction of a generalized polynomial expansion.

(for instance, negative values of the proton VDF) appear in the model of Demars & Schunk (1979). Indeed, Palmadesso et al. (1988) have defined the domain of application of the solution of Demars & Schunk (1979) by the condition $q_{\|p} \leq a(P_\| V_{\text{th}\|})_p$, where $a = 0.44$ according to the study of the dispersion relation issued from a 16 moment model, and they have stressed that a is equal to unity for dynamic fluid plasma simulations. We have retained this last value for a as the less restrictive criterion. Although it is not a statistical analysis because of the small number of three-dimensional VDF profiles in Marsch's (1982) paper, it nevertheless gives an indication of the validity of the 16 moment model in solar wind. For instance, in cases I and K, $q_{\|p}/(P_\| V_{\text{th}\|})_p$ is respectively equal to 1.7 and 1.3. Therefore, it seems that in most of these cases, the bi-Maxwellian polynomial expansion cannot be used to model solar wind proton VDF. Furthermore, we find no relation between the normalized parallel heat flux value and the existence of the double peak, if we compare Figure 3*a* with Figure 3*b*.

The lack of measurements in transitional regions of the polar wind does not give us the possibility to provide a similar analysis. But a few Monte Carlo simulations (Barakat, Bargouthy, & Schunk 1995; Wilson 1992) have supplied the evolution of the predicted proton VDF profiles. Thus, taking into account their results, we note that the typical proton VDF profiles in the polar wind are similar to the typical proton VDF shapes in the solar wind. Moreover, measurements in the solar wind are exclusively made in the transition region far from the Sun, while a number of measurements have been made in the collision-dominated region of the polar wind, but few are made in higher regions. Thus, the study of the polar wind has a complementary aspect to the study of the solar wind from the point of view of the measurements of the plasma parameters of astrophysical winds.

3. THE BI-MAXWELLIAN POLYNOMIAL EXPANSION TO MODEL THE PROTON VELOCITY DISTRIBUTION FUNCTION

3.1. *General Algebraic Rules of Construction*

First developed by Grad (1958), a polynomial expansion based on a weight function presents great interest for solving the Boltzmann equation. The weight function in the original Grad approach is the local Maxwellian equilibrium distribution function. His mathematical structure has been generalized for far-from-equilibrium states by Mintzer (1965). This latter has defined the general form of an expansion in velocity space around a zeroth-order distribution

function, which has to be chosen as a good approximation in some limits of the exact solution of the Boltzmann equation (in order to ensure the rapid convergence of the expansion). The VDF is expressed as follows:

$$f(\boldsymbol{r}, \boldsymbol{v}, t) = f^0(\boldsymbol{r}, \boldsymbol{v}, t) \sum_{q=0}^{+\infty} a_q(\boldsymbol{r}, t) M_q(\boldsymbol{v}) , \quad (1)$$

where f^0 is the zeroth-order function, the M_q are a complete set of functions in velocity space (but may also depend on space and time), and the coefficients a_q act as the generalized state variables.

In order to construct the simplest set of equations for the a_q, we define an orthonormalization rule,

$$\int_V f^0(\boldsymbol{r}, \boldsymbol{v}, t) M_p(\boldsymbol{v}) M_q(\boldsymbol{v}) d\boldsymbol{v} = \delta_{pq} , \quad (2)$$

which involves

$$a_q(\boldsymbol{r}, t) = \int_V f(\boldsymbol{r}, \boldsymbol{v}, t) M_q(\boldsymbol{v}) d\boldsymbol{v} , \quad (3)$$

where V is the velocity space. The equations for a_q are obtained by multiplying the Boltzmann equation by M_q and integrating the result in velocity space. If only the first few velocity moments are useful, then a solution is to form the set M_q from the complete set of monomials 1, $\boldsymbol{v}$, $\boldsymbol{vv}$, $\boldsymbol{vvv}$, ... (using a dyadic notation). Then, the moments of order m will linearly derive from the coefficients a_q with $q \leq m$. But the equation for a_q involves coefficients of at least order $q + 1$. Therefore, it is necessary to make some assumptions for the coefficients higher than those that we are solving. As for the Grad choice, Mintzer has suggested to truncate the expansion, that is, higher coefficients a_q than a given order m are set equal to zero. Finally, we obtain a closed set of equations for the coefficients a_q. In the same way, an equivalent system of equations for the velocity moments of lower order than m (or for the equivalent macroscopic parameters) can be obtained by multiplying the Boltzmann equation with v^q for q varying from 0 to m and by integrating in velocity space. The closure assumption is determined with the expression of f, which depends on $(a_q)_{q=0,m}$ and consequently on the velocity moments of lower order than m. Such a limiting condition, that is, to know a priori f, is essential in determining the expression of the collisional transfers for non-Maxwellian potential forces and in closing the set of transport equations in spatially nonhomogeneous systems.

Lowell (1967) has defined the convergence properties of the polynomial expansion constructed from equations (1), (2), and (3). We review them, as this point is seldom discussed, before embarking on long calculations that may not lead to a good approximation whatever the number of terms kept in the expansion. The coefficients a_q of the expansion (eq. [3]) minimize the integral:

$$I = \frac{1}{n_p} \int_V \frac{[f^E(\boldsymbol{r}, \boldsymbol{v}, t) - f^A(\boldsymbol{r}, \boldsymbol{v}, t)]^2}{f^0(\boldsymbol{r}, \boldsymbol{v}, t)} d\boldsymbol{v} , \quad (4)$$

where f^E is the exact solution and f^A the approximation. It is equivalent to minimizing the mean square error in an expansion of $F = f^{0-1/2} f^E$. Thus, the velocity space regions in which the $f^{0-1/2} f^A$ would be near $f^{0-1/2} f^E$, following this criterion, could not exactly correspond to the regions where the exact solution f^E would be well approximated by f^A. Therefore, particular care must be taken when we choose f^0. A natural criterion for the convergence of the polynomial expansion, which would have been that $\int_V [f^E(\boldsymbol{r}, \boldsymbol{v}, t) - f^A(\boldsymbol{r}, \boldsymbol{v}, t)]^2 d\boldsymbol{v}$ must be minimized, is useless because it prevents an easy construction of the equations for the coefficients a_q (Mintzer 1965). Then, a necessary condition derived from equation (4) is that the integral $\int_V [f^E(\boldsymbol{r}, \boldsymbol{v}, t)]^2 / [f^0(\boldsymbol{r}, \boldsymbol{v}, t)] d\boldsymbol{v}$ exists. One condition for the definition of this integral is that

$$\lim_{|\boldsymbol{v}| \to \infty} \frac{f^E(\boldsymbol{r}, \boldsymbol{v}, t)}{[f^0(\boldsymbol{r}, \boldsymbol{v}, t)]^{1/2}} = 0 . \quad (5)$$

In what follows we use the condition defined in equation (5) to know whether the polynomial expansion is divergent.

3.2. *Bi-Maxwellian Function as a Weight Factor*

Several authors have generalized Grad's method for far-from-equilibrium states with a bi-Maxwellian function as a weight factor. First applied by Chew et al. (1956) for a collisionless anisotropic plasma, the approach was generalized by Demars & Schunk (1979) to a large class of space plasmas. They have provided a system of transport equations for the first 16 velocity moments and have defined a microscopic description as in equation (1). The mathematical definition of the bi-Maxwellian weight function is

$$f^0_{bM}(\boldsymbol{r}, c_\parallel, c_\perp, t) = n_p \left(\frac{m_p}{2k_B \pi T_{\parallel p}} \right)^{1/2} \frac{m_p}{2k_B \pi T_{\perp p}} \times \exp \left\{ -\frac{m_p}{2k_B T_{\perp p}} c_\perp^2 - \frac{m_p}{2k_B T_{\parallel p}} c_\parallel^2 \right\} , \quad (6)$$

where n_p is the density, $c_\parallel$ is the random thermal velocity parallel to the magnetic field ($\boldsymbol{c_s} = \boldsymbol{v_s} - \boldsymbol{u_s}$, where $\boldsymbol{u_s}$ is the species average drift velocity), and the perpendicular velocity $c_\perp$ is a vector of two dimensions (we consider gyration-dominated plasmas), $T_{\parallel p} = m_p \langle c_{\parallel p}^2 \rangle / k_B$ and $T_{\perp p} = m_p \langle c_{\perp p}^2 \rangle / 2k_B$, where $\langle A \rangle = \int_V f(\boldsymbol{r}, \boldsymbol{v}, t) A d\boldsymbol{c_p}$. The polynomials generated by the weight function defined in equation (6) and the orthonormalized rule (see eq. [2]) are the Hermite polynomials along the parallel velocity denoted H_i and the associated Laguerre polynomials of order zero along the perpendicular velocity denoted L_j (Jancel & Kahan 1966). Therefore,

$$f^A_{bM}(\boldsymbol{r}, c_\parallel, c_\perp, t) = f^0_{bM}(\boldsymbol{r}, c_\parallel, c_\perp, t) \times \left\{ 1 + \sum_{i,j=1}^{m} a_{ij}(\boldsymbol{r}, t) H_i(c_\parallel) L_j(c_\perp) \right\} \quad (7)$$

is the approximation expansion of order m with $i + j \leq m$. In the 16 moment development, m is equal to 3.

In the Demars & Schunk (1990) investigations, the derived polynomial expansion should be able to reproduce all the profiles of the VDF observed in the solar wind and expected in the polar wind. Their model should also be able to provide the macroscopic parameter evolution in good agreement with observations. But some criticisms have been made on their conclusions. Palmadesso et al. (1988) have solved the dispersion relation issued from a 16 moment model and obtained unstable waves when $q_{\parallel p} / (P_\parallel V_{\text{th}\parallel})_p$ is above about 0.44. However, in their study they did not consider the collisional terms, and, consequently, they overestimated this criterion by neglecting dissipative terms. But they pointed out that in any case when running dynamic fluid plasma simulations, unstable behaviors appear when

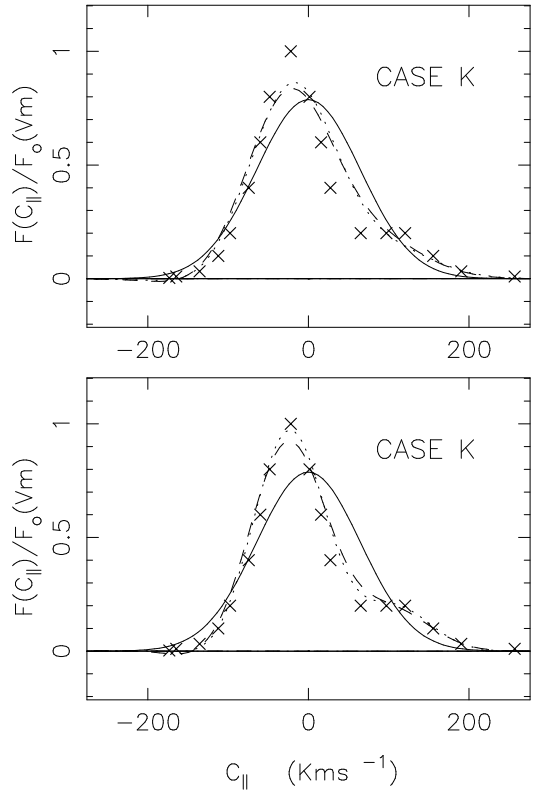

FIG. 4.—Polynomial expansion based on a bi-Maxwellian weight factor, centered on the average velocity for case K issued from Marsch data. The crosses represent the *Helios* probe measurements in the magnetic field direction. The solid line is the zeroth-order function defined in § 3.2 (see eq. [6]). In the top panel, the dashed line corresponds to order 3 of development (see eq. [7]) and the dotted line to order 4, while the bottom panel shows orders 5 and 7, respectively.

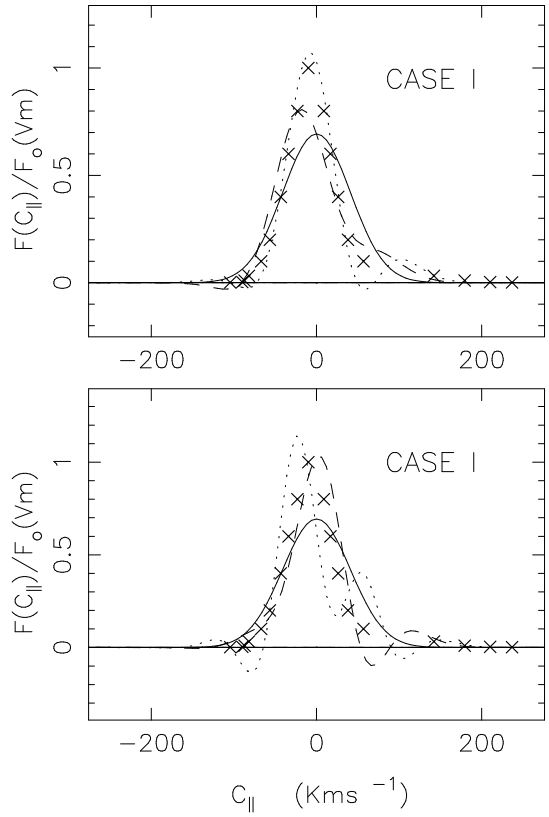

FIG. 5.—Polynomial expansion based on a bi-Maxwellian weight factor for case I: the plotting format is the same as that for Fig. 4.

$q_{\|p}/(P_{\|}\, V_{\mathrm{th}\|})_p$ is on the order of 1. Indeed, in that case, the polynomial expansion Φ in $f^A_{bM} = f^0_{bM}(1 + \Phi)$ reaches a value of the order of the unity, so that the VDF may be in the process of breaking down (Palmadesso et al. 1988). Thus, when Demars & Schunk (1990) tried to reproduce contour plots of a typical proton VDF in the solar wind, in order to obtain a double-peaked shape, they increased the total proton heat flux value determined by Marsch within the important experimental uncertainty, and they arbitrarily fixed the anisotropy heat flux, which was not given in Marsch data. Therefore, their conclusions, to a certain extent, fail because they did not respect the limitation criterion on the parallel heat flux. Thus, the second peak appears simultaneously with negative values of the proton VDF. In conclusion, it seems that the contour plots with a double peak that they have obtained are due to a mathematical artifact of the polynomial expansion, as pointed out as early as 1985 by Hubert (Hubert 1985), stressing consequences on instabilities.

The convergence criterion of a polynomial expansion defined in equation (5) is not easy to apply because of the unknown f^E. But if we use the representation of f^E defined in § 2.2, then we are able to check such a criterion. We model the observed VDF by the sum of two Maxwellian distribution functions: $f(v) = n_1 f_{\mathrm{Max}1}(T_1, v) + n_2 f_{\mathrm{Max}2}(T_2, v)$. In regions of large velocity, one of the Maxwellian functions of the sum is much larger than the other, which therefore can be neglected and is the function that fits the core. The convergence criterion (see eq. [5]) $f_{\mathrm{Max}}(c_{\|})/(f^0_{bM})^{1/2}(c_{\|}) \to 0$ when $c_{\|} \to \pm\infty$ is then equivalent to the condition $T_{\mathrm{Max}} < 2T_{\|}$. Tables 1 and 2 give the fitting values (the Maxwellian function number 2 fits the halo, the number 1 the core); then, we can note that this required condition is not verified in case I for large positive velocity modulus, as $T_{\mathrm{Max}} = T_2 = 500{,}480$ K and $2T_{\|} = 408{,}840$ K.

In Figures 4 and 5 we have plotted *Helios* probe observations denoted as cases K and I, respectively, in the parallel velocity plane (*crosses*) and the corresponding bi-Maxwellian polynomial expansion of order 0 (*solid line*), 3 (*dashed line*), and 4 (*dotted line*) in the top panel and 0 (*solid line*), 5 (*dashed line*), and 7 (*dotted line*) in the bottom panel. We note the evident nonconvergence of the model in case I and the slow convergence in case K (in that case we have established the convergence criterion and that the apparition of the oscillation is very sensitive to the values of the velocity moment). The convergence criterion defined in equation (5) seems to be realistic according to these two figures.

Furthermore, the negativity of the distribution function is in contradiction with the Demars & Schunk (1979) assumption on the calculation of the collisional transfers, i.e., the polynomial expansion Φ must remain lower than unity.

4. GENERALIZED POLYNOMIAL EXPANSION OF THE PROTON VELOCITY DISTRIBUTION FUNCTION IN POLAR AND SOLAR WINDS

Mintzer (1965) has recommended the choice of a weight function (which is also the zeroth-order approximation)

near the expected exact solution. Then, for applications to solar wind, an interesting idea for a microscopic model would be an expansion with a weight factor, such as an asymmetric function involving large parallel heat flux. Hubert (1985), being inspired by the Whealton & Woo (1971) exact solution of the BGK equation for the moderate ionized plasma subjected to an electrostatic field, built a polynomial expansion with a weight function defined as

$$f_G^0(\boldsymbol{r}, c_\parallel, c_\perp, t) = n_p \frac{m_p}{4k_B \pi T_{\perp p} D^*} \times \exp\left\{-\frac{m_p}{2k_B T_{\perp p}} c_\perp^2 - \frac{c_\parallel + D^*}{D^*} + \frac{1}{E^*}\right\} \times \operatorname{erfc}\left\{E^{*1/2}\left(\frac{1}{E^*} - \frac{c_\parallel + D^*}{2D^*}\right)\right\}, \quad (8)$$

where $c_\parallel$, $c_\perp$, and $T_{\perp p}$ were defined in § 3, and erfc is the complementary error function (Abramovitz & Stegun 1964). D^* and E^* are chosen such that the exact first-order velocity moments $\langle c_\parallel^m \rangle$ and $\langle c_\perp^l \rangle$ are equal to those provided by the zeroth-order approximation, that is,

$$\int_V c_\parallel^m f_G^0(\boldsymbol{r}, \boldsymbol{c}, t) d\boldsymbol{c} = \int_V c_\parallel^m f^E(\boldsymbol{r}, \boldsymbol{c}, t) d\boldsymbol{c} \text{ for } m \le 3 ,$$
$$\int_V c_\perp^l f_G^0(\boldsymbol{r}, \boldsymbol{c}, t) d\boldsymbol{c} = \int_V c_\perp^l f^E(\boldsymbol{r}, \boldsymbol{c}, t) d\boldsymbol{c} \text{ for } l \le 2 . \quad (9)$$

Then, we obtain the following relations:

$$D^* = \left(\frac{\langle c_{p\parallel}^3 \rangle}{2}\right)^{1/3} ,$$
$$E^* = \frac{2m_p D^{*2}}{k_B T_0} , \quad (10)$$
$$T_0 = T_{\parallel p} - \frac{m_p D^{*2}}{k_B} ,$$

and the formal expression for f_G^A in which f_G^0 is completely defined:

$$f_G^A(\boldsymbol{r}, c_\parallel, c_\perp, t) = f_G^0(\boldsymbol{r}, c_\parallel, c_\perp, t) \times \left\{1 + \sum_{i,\, j=1}^{n} a_{ij}(\boldsymbol{r}, t) M_i(c_\parallel) L_j(c_\perp)\right\}, \quad (11)$$

where $(L_j)_{j=1,n}$ are Laguerre polynomials associated with the perpendicular part (dependent on $c_\perp$) of f_G^0. $(M_i)_{i=1,n}$ is an orthonormalized polynomial set defined in the $c_\parallel$ velocity space and associated with the parallel part of f_G^0, and n is the order of development with $i + j \le n$. With this function we are able to calculate the velocity moments at any order.

Such a f_G^0 has a long suprathermal tail in the magnetic field direction and displays an anisotropic temperature. We follow the advice of Mintzer (1965) by choosing a zeroth-order function as a known solution of equations that approximate the real physical phenomena. Indeed, in the polar wind Barakat et al. (1995) considered a flow of H^+ ions through a background of O^+ ions (with constant average velocity and temperature) subjected to an electrostatic field and velocity-dependent Coulomb collisions. They obtained long-tailed VDF profiles that become double-humped VDF profiles when the number of collisions decrease in the transition region. In the solar wind a similar evolution of the VDF of the protons was obtained by Livi & Marsch (1987) by studying the evolution of proton species subjected to self-Coulomb collisions described by the BGK operator. Therefore, the relaxation of protons in a background of neutrals subjected to an electrostatic field and a constant collision frequency (which is the problem solved by Whealton & Woo 1971) can be considered as the first-order approximation of processes that provide typical observed VDFs.

Furthermore, the definition of f_G^0 proposed by Hubert imposes that $q_{\parallel p}$ be less than 2 times the parallel free-streaming flux (derived from the condition that the effective temperature T_0 has to be positive). This limitation is represented in Figure 3 by the dashed lines in the context of the observations made by the probe *Helios* in the solar wind. Most of the measurements are under the broken curve, which means that f_G^0 is almost everywhere defined. A similar approach for the study of non-Maxwellian ion states in the auroral latitudes has been developed by Hubert (1983). The polynomial expansion is based on the exact solution of the BGK equation for a moderate ionized plasma subjected to an electrostatic field perpendicular to the magnetic field. Hubert & Barakat (1990) and Barakat & Hubert (1990) have shown the good convergence properties of this polynomial expansion when compared to Monte Carlo simulations.

As the convergence property defined for a bi-Maxwellian function, an equivalent condition to equation (5) is obtained by substituting for large velocity modulus, f^E with the sum of two Maxwellian distribution functions defined in § 2.2. We have determined an equivalent condition as $T_{\mathrm{Max}} \le 2T^0$, where T^0 is defined in equation (10), and T_{Max} is the larger temperature between the two temperatures T_1 and T_2 of the sum of the two Maxwellian functions. For case I, $2T^0 = 91{,}105$ K and $T_{\mathrm{Max}} = T_2 = 500{,}480$ K for large negative values of the parallel velocity, and in case K, $2T^0 = 484{,}182$ K and $T_{\mathrm{Max}} = T_1 = 217{,}638$ K. However, for these regions we have observed that the function that fits the halo overestimates the profile of case I in large negative velocity regions; the other Maxwellian function is more adapted (in this case, the criterion of convergence is verified: $T_{\mathrm{Max}} = T_1 = 71{,}133$ K). Nevertheless the long-tail part, in positive velocity regions, of the function is more interesting because it characterizes the non-Maxwellian behaviors of the VDF. Thus, for positive values of the parallel velocity, the convergence criterion is always verified. Indeed, the condition of convergence is equivalent to

$$\lim_{c_\parallel \to +\infty} \frac{\exp\left[-(m_p/2k_B T_{\mathrm{Max}})c_\parallel^2 + (c_\parallel/D^*)\right]}{\operatorname{erfc}\left\{E^{*1/2}[(1/E^*) - (c_\parallel + D^*)/2D^*]\right\}^{1/2}} = 0 ,$$

which is always verified because

$$\lim_{c_\parallel \to +\infty} \operatorname{erfc}\left[E^{*1/2}\left(\frac{1}{E^*} - \frac{c_\parallel + D^*}{2D^*}\right)\right] = 2 .$$

Therefore, the necessary convergence condition defined in § 3 is verified in both cases.

We show in Figures 6 and 7 the corresponding polynomial expansion function f_G^A obtained with velocity moments of Tables 1 and 2 in the same way as in Figures 4 and 5. We have used for the construction of f_G^A the same order m of development and the assumption of gyration-dominated plasmas as in § 3. The crosses represent the *Helios* probe measurements. The top panel of Figures 6 and

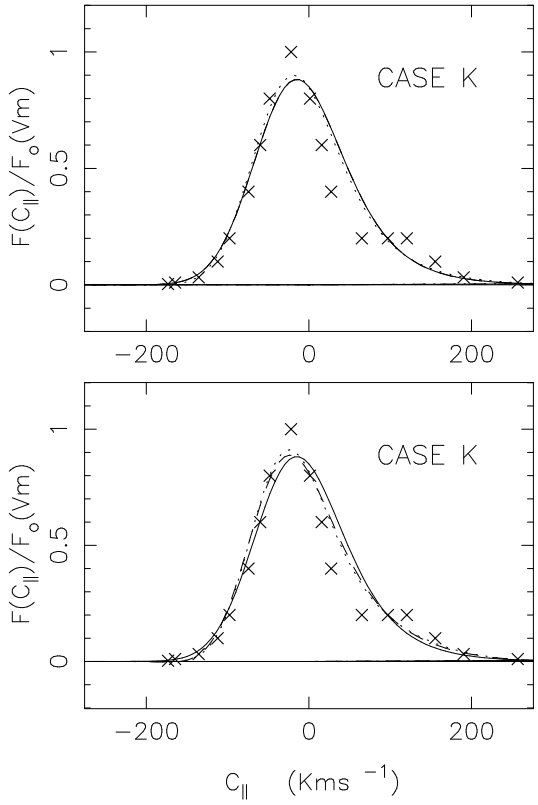

FIG. 6.—Polynomial expansion based on the generalized weight function defined in § 4, centered on the average velocity for case K issued from Marsch data. The crosses are the *Helios* probe measurements in the magnetic field direction. The solid line is the zeroth-order function defined in § 4 (see eq. [8]). In the top panel, the dashed line corresponds to order 3 of development (see eq. [11]) and the dotted line to order 4, while the bottom panel shows orders 5 and 7, respectively.

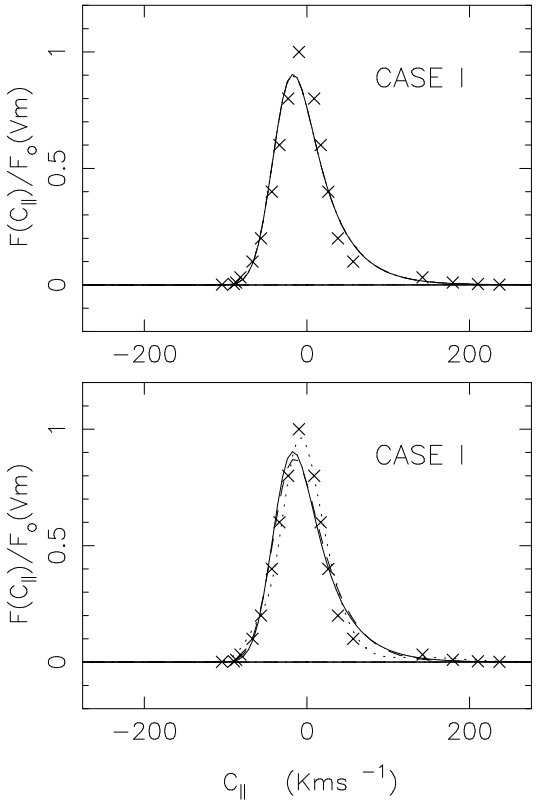

FIG. 7.—Polynomial expansion based on the generalized weight function defined in § 4 for case I: the plotting format is the same as that for Fig. 6.

7 represents the 0, 3, and 4 orders of the generalized polynomial expansion, and the bottom panel represents the orders 0, 5, and 7.

Comparing Figure 6 to Figure 4, we note that case K seems to be better approached with a bi-Maxwellian polynomial expansion than with a generalized approximation for order m larger than 4. Indeed, in Figure 6, a double peak never appears, but neither do negative values, as observed in Figure 4. Both developments seem to converge slowly and to need orders larger than 3 to fit accurately the observed distribution function.

In case I, the divergence of the bi-Maxwellian polynomial expansion is obvious. With a generalized polynomial expansion the convergence is slow, but with development of order 7 the polynomial expansion approaches very accurately the exact profile.

Another query into the nature of a polynomial expansion f^A of given order p is to compare the velocity moments of larger order than p calculated from f^A, to their exact values.

We have calculated the velocity moments of order 4, 5, and 6 in the magnetic field direction from the bi-Maxwellian expansion (see eq. [7]) and from the generalized expansion (see eq. [11]) limited at the third order. Table 3 shows the accuracy for cases I and K in comparison with the exact values of the velocity moments obtained by

TABLE 3

COMPARISON OF VELOCITY MOMENTS

Parameter	CASE I		CASE K	
	Generalized	Bi-Maxwellian	Generalized	Bi-Maxwellian
$(\langle c_\parallel^4\rangle^A - \langle c_\parallel^4\rangle^E)/\langle c_\parallel^4\rangle^E$ (%).....	4.7	56.2	30.5	10.5
$(\langle c_\parallel^5\rangle^A - \langle c_\parallel^5\rangle^E)/\langle c_\parallel^5\rangle^E$ (%).....	19.9	35.5	89.7	39.6
$(\langle c_\parallel^6\rangle^A - \langle c_\parallel^6\rangle^E)/\langle c_\parallel^6\rangle^E$ (%).....	32.7	85.1	163	17.6
$\langle c_\parallel^4\rangle^A/[3(\langle c_\parallel^2\rangle^E)^2]$	2.21	...	1.55	...
$\langle c_\parallel^6\rangle^A/[15(\langle c_\parallel^2\rangle^E)^3]$	9.6	...	4.2	...

NOTE.—Comparison between the velocity moments of orders 4, 5, and 6 of the approximation f^A (superscript A) for a polynomial expansion of order 3 (see eqs.[7] and [11]) with respect to the moment value of the exact function (superscript E) defined in Tables 1 and 2. The ratios between moments of order 4 and 6 of f^A and the corresponding value for a Maxwellian function defined with the exact density and parallel temperature are also given.

the fitting defined in § 2.2. Their non-Maxwellian nature (in comparison with the theoretical value of the velocity moments of a Maxwellian function defined with the exact parallel temperature) is also provided (last two rows in Table 3). The accuracy of the moment of order 4 is very good in case I for a generalized polynomial expansion and is better estimated than for a bi-Maxwellian one. Furthermore, the study of their non-Maxwellian features allows us to conclude that the velocity moments of higher order than p are more and more "non-Maxwellian" and contribute to the formation of the long tail in the VDF profile. In case K, velocity moments are better estimated with a bi-Maxwellian polynomial expansion despite the negative values of the distribution function. Globally, we can hope for a better macroscopic description of protons in solar wind, because our assumption of velocity moments closure is better adapted to the long-tailed VDF, which is a feature of a large majority of solar wind proton VDF shapes.

5. DISCUSSION

We have chosen for our study two typical profiles, cases I and K, from *Helios* probe measurements. These VDFs reproduce the main features of the proton distribution function observed in the solar wind but are also typical distribution functions of moderate ionized plasmas with a constant collision frequency (Hubert 1985) or with soft interactions (Skullerud 1984) and of Monte Carlo simulations applied to polar wind (Barakat et al. 1995; Wilson 1992). A sum of two Maxwellian functions fits with good precision the cuts through the three-dimensional distribution along the magnetic field direction; in particular, this fitting seems to be well adapted to the long-tailed profiles.

The long-tailed profiles follow in a number of kinetic states an $\exp(-\beta v^{\gamma+1})$ law with β constant, for instance, in an ion swarm experiment $-0.5 \leq \gamma \leq 1$ (Skullerud 1984) or with $\gamma = 0$ for the solution of the BGK equation with a constant collisional frequency (Whealton & Woo 1971). Then, when these tails are fitted with a sum of two Maxwellian functions [roughly following an $\exp(-\beta v^2)$ law for large velocity values], we overestimate the decreasing of the function. Consequently, the criterion of convergence will be easier to verify. In fact Skullerud (1984) has shown that in the case of a polynomial expansion based on a Maxwellian function, the convergence would never occur if the accurate solution falls off asymptotically slower than any Gaussian. In the same way, Hubert (1985) has clearly shown the difficulty of the convergence of a bi-Maxwellian polynomial expansion. He has constructed such a polynomial expansion to approximate the solution of a BGK equation for the moderately ionized plasmas subjected to an electrostatic field, and he has concluded that whatever the value of $q_{\|p}/(v_{\mathrm{th}\|} P_{\|})_p$ might be, the criterion of convergence could not be verified. He has emphasized that higher orders in polynomial expansion do not necessarily give a better approximation.

The slowness of the convergence is another aspect revealed by Figures 4, 5, 6, and 7. In Figures 4 and 5, displaying the bi-Maxwellian polynomial expansion, oscillations and negative values appear as soon as the third order. But they do not appear in Figures 6 and 7 because the zeroth-order generalized function damps such unphysical behaviors. Indeed, in Figure 8 we show the polynomial expansion of order 3 based on the generalized function (see eq. [11]) with a solid line, and on the bi-Maxwellian function (see eq. [7]) with a dashed line. We note that the two profiles are very similar. A third-order level of microscopic description cannot reproduce accurately the profiles of the observed VDF. But it is important to stress that the generalized distribution function at the third order does not display negative values, contrary to the bi-Maxwellian polynomial expansion at the same order. This property is certainly important for the evaluation of the velocity collision transfer.

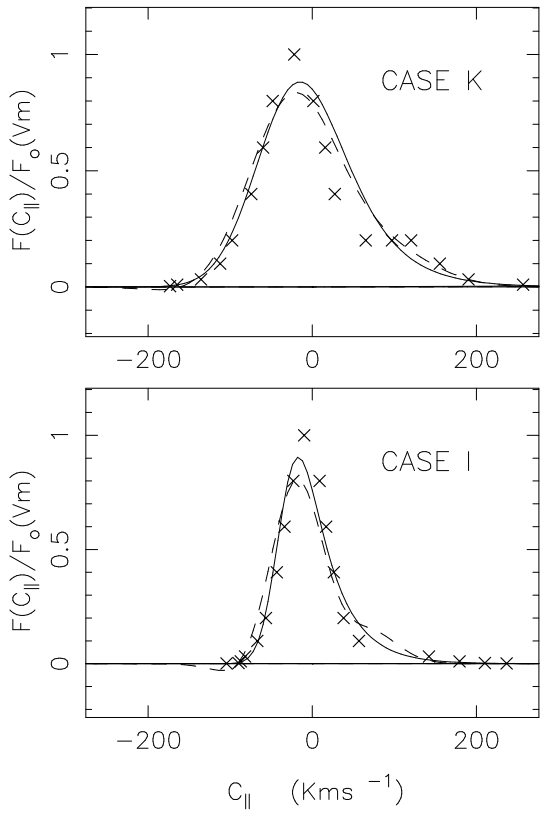

FIG. 8.—Polynomial expansions of order 3 based on the generalized weight function defined in § 4, eq. (11) (*solid line*), and on the bi-Maxwellian weight function defined in § 3, eq. (7) (*dashed line*), centered on the average velocity for cases K (*top*) and I (*bottom*) issued from Marsch data. The crosses are the *Helios* probe measurements in the magnetic field direction.

Another point, already emphasized, is the closure of the associated transport equations to a given order. Table 4 shows the precision obtained on the velocity moments of order $p+1$ calculated with an approximated polynomial expansion function of order p, for p equal to 3, 4, 5, and 6. There is no evident increase of the accuracy of the derived velocity moments of order $p+1$, except for the bi-Maxwellian expansion in case K. Nevertheless, we note that there is a good estimation of $\langle c_{\|}^4 \rangle$ from the generalized expansion of order 3 in case I, as well as for $\langle c_{\|}^5 \rangle$ at a fourth order of development. In case K the results relating to the closure of the transport equations show that a bi-Maxwellian polynomial expansion seems to be more adapted than a generalized one.

The double-humped functions represent 20% of all the observations made in the solar wind by the *Helios 2* probe according to Marsch et al. (1982). Their real physical origins have not yet been well determined and still constitute an open field of research. Nevertheless, if we want to reproduce their form with a polynomial expansion, then we need to

TABLE 4

PERCENTAGE ERROR ON VELOCITY MOMENTS

	CASE I		CASE K	
Order of Development	Generalized	Bi-Maxwellian	Generalized	Bi-Maxwellian
3	4.7	56.2	30.5	10.5
4	21.7	35.5	74.1	39.6
5	42.5	27.5	53	12
6	24.7	26.5	73	7.4

NOTE.—Calculation of the percentage error on velocity moments at order $p + 1$ for a bi-Maxwellian polynomial expansion and a generalized polynomial expansion of order $p = 3$, 4, 5, and 6.

find another weight function and probably at least a fourth-order polynomial expansion. The f_G^0 profile is too different from a double-peaked shape to be well adapted according to Mintzer's advice.

Other approaches exist. For instance, Eu (1980) and Chen & Eu (1982) have developed a modified moment method in order to reconcile irreversible thermodynamics for system far-from-equilibrium states with the solution of the Boltzmann equation. In this work the microscopic description is defined as $f_p(\boldsymbol{r}, \boldsymbol{v}, t) = \exp\{-\beta(\boldsymbol{r}, \boldsymbol{v}, t)H_p(\boldsymbol{r}, \boldsymbol{v}, t)\}$, where H_p is a polynomial of third order, and $\beta = 1/(k_B T_p)$. To overcome mathematical difficulties in the definition of the entropy density production, the method of cumulant expansion is used (Clause & Balescu 1982). In a similar attempt Levermore (1995) has suggested the use of distribution functions $f_p = \exp\{P(v)\}$, where $P(v)$ is a polynomial of even degree. But Gombosi et al. (1994) has stressed the difficulties in obtaining an explicit closed-form expression of P in terms of the velocity moments of f_p, for degree of P higher than 2. However, the great interest of Levermore's approach, as in Eu's proposition, is to define an entropy balance equation and to secure the positivity of f_p and the realizability criterion. We have tried to verify whether such a function defined with a polynomial P of degree 4 (the lowest order to model the double-humped function) was able to fit cases I and K. This was easily obtained in case K, but without good precision, and we have met difficulties in fitting profile case I. Another method has been proposed by Cuperman, Weiss, & Dryer (1983) and Cuperman et al. (1987). They have constructed a VDF from the minimization of Boltzmann's H-function, subjected to the constraints provided by the set of the selected components of the approximation (seven or nine macroscopic parameters). Although applied to spherically symmetric systems, this model provides almost all the typical profiles of the solar wind and is able, according to the authors, to generate an improved associated system of transport equations. The sum of two Maxwellian functions in the velocity space parallel to the magnetic field, multiplied by a Maxwellian function in the perpendicular velocity space, should also be considered. But it remains difficult to use it because the derived system of transport equations is composed of 57 variables, as the velocity moments up to order 5 are needed to define the sum of two Maxwellian functions.

6. CONCLUSION

The observations of the *Helios* probe in solar wind between 0.3 and 1 AU have revealed the main aspects of the proton VDF: a magnetic field–aligned suprathermal tail and, for about 20% of them, the presence of a second peak. These nonthermal features have been studied by several authors. Up to now, the most sophisticated solution proposed to model these proton VDFs is a polynomial expansion with a bi-Maxwellian as the weight factor, associated with a system of transport equations constructed following the Grad method. This model provides temperature anisotropies similar to observations, contrary to the collisionless approaches or to the fluid models.

But the choice of the microscopic description is not satisfied according to criticisms that have determined the field of application of such a model. In fact, the polynomial expansion is able to model only exact solutions that are near the weight function. We have established that the 16 moment bi-Maxwellian polynomial expansion does not converge for the most typical case of a proton VDF in solar wind. In this case, the system of transport equations does not respect mathematical rules such as the hyperbolicity condition. Moreover, the limited order polynomial expansion can generate negative values. Therefore, the collisional terms are not well estimated, and the closure assumption is not appropriate. We have also shown that an increase of the order of development does not improve the model.

Therefore, we propose a new approach to model the proton microscopic state in the solar and polar winds (or, more generally, for stellar atmosphere expansion) and to derive generalized transport equations. It is also a polynomial expansion, but it is based on a function derived from the exact solution of the BGK equation for the moderate ionized plasma subjected to an electrostatic field. The advantage of this representation is an intrinsic asymmetry of the weight function, which displays a suprathermal tail in the magnetic field direction, that is, a profile close to 80% of the observed VDF. Furthermore, the heat flux $q_{\|p}$ that is used in the definition of the weight function is not involved in the construction of the polynomial part of the expansion. Therefore, the limitation on the intensity of $q_{\|p}$ derives only from the rules of construction of this generalized weight function. But in the solar wind it seems not to be a difficulty because almost all the observations are in agreement with this limitation. Consequently, no negative value of the VDF appears with large values of the parallel heat flux, as it was the case with a bi-Maxwellian polynomial expansion of order 3. This new solution verifies the Mintzer criterion of convergence in all the observed typical cases, contrary to the bi-Maxwellian polynomial expansion, and, consequently, it is better adapted for the determination of macroscopic parameters when we solve the associated system of transport equations.

The double-humped VDF needs to use more sophisticated zeroth-order models for a multimoment approach or kinetic models such as those in Livi & Marsch (1987). Indeed, these authors have combined the action of the large-scale interplanetary magnetic field and of the collisional scattering and have obtained double-humped profiles.

Next developments will be the construction of the set of transport equations associated with the generalized polynomial expansion and the study of its properties.

REFERENCES

Abramovitz, M., & Stegun, I. A. 1964, Handbook of Mathematical Functions (New York: Dover)
Barakat, A. R., Bargouthy, I. A., & Schunk, R. W. 1995, Geophys. Res. Lett., 22, 1857
Barakat, A. R., & Hubert, D. 1990, Ann. Geophys., 8, 697
Chamberlain, J. W. 1960, ApJ, 131, 47
Chapman, S. 1965, Smithson. Contrib. Astrophys., 2, 1
Chen, M., & Eu, B. C. 1982, J. Chem. Phys., 77, 2696
Chew, G. F., Goldgerger, M. L., & Low, F. E. 1956, Proc. R. Soc. London A, 236, 112
Chodura, R., & Pohl, F. 1971, Plasma Phys., 13, 645
Clause, P.-J., & Balescu, R. 1982, Plasma Phys., 24, 1429
Cordier, S. 1994a, Math. Models Meth. Appl. Sci., 4, 625
———. 1994b, Math. Models Meth. Appl. Sci., 4, 647
Cuperman, S., Weiss, I., & Dryer, M. 1980, ApJ, 239, 345
———. 1981, ApJ, 251, 297
———. 1983, ApJ, 273, 363
Cuperman, S., Yatom, H., Dryer, M., & Lewis, D. 1987, ApJ, 314, 404
Demars, H. G., & Schunk, R. W. 1979, J. Phys. D, 12, 1051
———. 1989, Planet. Space Sci., 37, 85
———. 1990, Planet. Space Sci., 38, 1091
———. 1991, Planet. Space Sci., 39, 435
Eu, B. C. 1980, J. Chem. Phys., 73, 2958
Feldman, W. C. 1979, Sol. Syst. Plasma Phys., 1, 321
Gazis, P. R. 1984, J. Geophys. Res., 89,775
Gombosi, T. I., Groth, C. P. T., Roe, P. L., & Brown, S. L. 1994, Phys. Fluids, submitted
Gombosi, T. I., & Rasmussen, C. E. 1991, J. Geophys. Res., 96, 7759
Grad, H. 1958, Handb. Phys., 12
Hubert, D. 1983, Planet. Space Sci., 31, 119
———. 1985, J. Phys. D, 18, 1521
Hubert, D., & Barakat, A. R. 1990, Ann. Geophys., 8, 687
Jancel, R., & Kahan, J. 1966, Electrodynamics of Plasma (New York: Wiley)
Lemaire, J., & Scherer, M. 1971, J. Geophys. Res., 76, 7479
———. 1973, Rev. Geophys. Space Phys., 11, 427
Levermore, C., Morokoff, W. J., & Nadiga, B. T. 1995, Phys. Fluids, submitted
Levermore, C. D. 1995, J. Stat. Phys., submitted
Liu, S., Marsch, E., Livi, S., Woch, J., Wilken, B., von Steiger, R., & Goeckler, G. 1995, Geophys. Res. Lett., 22, 2445
Livi, S., & Marsch, E. 1987, J. Geophys. Res., 92, 7255
Lowell, H. H., Jr. 1967, Phys. Fluids, 10, 35
Maksimovic, M. 1995, Ph.D. thesis, Univ. Paris VII
Marsch, E. 1982, NASA CP Sol. Wind, 5, 355
Marsch, E., Mulhauser, K. H., Schwenn, R., Rosenbauer, H., Pillip, W., & Neubauer, F. M. 1982, J. Geophys. Res., 87, 52
Mintzer, D. 1965, Phys. Fluids, 8, 1076
Montgomery, M. D., Gary, S. P., Feldman, W. C., & Forslund, D. W. 1976, J. Geophys. Res., 81, 2743
Oraevskii, V., Chodura, R., & Feneberg, W. 1968, Plasma Phys., 10, 819
Palmadesso, P. J., Ganguli, S. B., & Mitchell, H. G., Jr. 1988, Geophys. Monogr. Ser., 44, 133
Parker, E. N. 1958, ApJ, 128, 664
Robineau, A., Blelly, P.-L., & Fontanari, J. 1996, J. Atmos. Terr. Phys. 58, 257
Schwenn, R., & Marsch, E., eds. 1991, Physics of Inner Heliosphere, Part 2: Particles, Waves, and Turbulence (Berlin: Springer)
Shizgal, B. 1977, Planet. Space Sci., 25, 203
Skullerud, H. R. 1984, J. Phys. B, 17, 913
Whealton, J. H., & Woo, S.-B. 1971, Phys. Rev. A, 6, 2319
Wilson, G. R. 1992, J. Geophys. Res., 97, 10,551

ANNEXE 2: PLANETARY SPACE SCIENCE, 49, 645-656, 2001.

LEBLANC, F. AND R.E JOHNSON, SPUTTERING OF THE MARTIAN ATMOSPHERE BY SOLAR WIND PICK-UP IONS, PLANET. SPACE SCI., 49, 645-656, 2001.

PERGAMON

Planetary and Space Science 49 (2001) 645–656

Planetary and Space Science

www.elsevier.nl/locate/planspasci

Sputtering of the Martian atmosphere by solar wind pick-up ions

F. Leblanc *, R.E. Johnson

Engineering Physics and Department of Astronomy, University of Virginia, 102, Charlottesville, VA 22903, USA

Received 18 May 2000; received in revised form 27 November 2000; accepted 1 December 2000

Abstract

Monte Carlo models are used to describe the interaction between the incident pick-up ions and the Mars neutral atmosphere. The sputtered population inside the corona and escaping particles are described using a modified 3D test particle model, whereas, the heating effect due to the incident flux is described using a 2D Direct Simulation Monte Carlo. These results show that the standard 1D models overestimate the sputtering yield (by 15–25%), when corrected for coronal ejection. It is also shown that the exobase altitude can depend on sputtering. Two epochs of Mars history are simulated. For an epoch suggested to be about 2.5 Gyr ago, sputtering is at least as important as dissociative-recombination for populating the corona and the heating due to the pick-up ion flux is of the same order as the EUV and UV heating of the thermosphere. For this epoch and present solar minimum conditions we present the distribution in the density and the energy of the sputtered particles in Mars corona. In particular we show that the polar and dusk regions are the most dense regions and the shadow of the dawn and polar regions the most energetic regions.

1. Introduction

Most scenarios for the evolution of Mars (Pepin, 1994; Carr, 1999) assume that water was a stable element on the surface in an earlier era (Zuber et al., 2000). This implies that Mars was wetter, warmer and had a thicker atmosphere than today. It is not yet understood when or how the water disappeared. The various scenarios for the loss of water agree on the following main phases of the atmospheric evolution. Before 4.4 Gyr from present, strong hydrodynamic escape, heavy bombardment of meteorites and outgassing have been used to describe Mars atmospheric evolution (Brain and Jakosky, 1994). The heavy bombardment may have lasted until 3.7 Gyr ago ending about the same time that the magnetic field of Mars turned off (Acuña et al., 1999, 1998; Carr, 1999). Impact bombardment and early hydrodynamic escape is estimated to have caused a loss of roughly 90% of the primitive Martian volatiles. However, impact surface features show a strong degradation due to erosional processes (Jakosky and Jones, 1997) suggesting that water was still present at the end or after the heavy bombardment period.

Carr (1999) has suggested that a 1–0.5 bar thick atmosphere should have been present after the heavy bombardment period. He also emphasized that one of the main factors which could explain the present much thinner 5 mb atmosphere is the sputtering of the atmosphere by pick-up ions, first described by Luhmann and Kozyra (1991). Photochemical escape (McElroy, 1972; Zhang et al., 1993a) and carbonate recycling (CO_2 deposited in a near surface reservoir) could also contribute. Jakosky and Jones (1997) concluded from the analysis of the present isotope ratios that 85–90% of Ar, C, H and N may have been lost by sputtering. The purpose of this work is to describe the 3D nature of the sputtering process in two epochs: the present solar minimum and an epoch suggested to be 2.5 Gyr ago to the present. The ultimate goal is to calculate whether or not sputtering determined the CO_2 and oxygen composition of the atmosphere and contributed to the inferred loss of 0.95–0.45 bar of atmosphere (Carr, 1999).

Whereas, the sputtering of Io's atmosphere was a well studied process (e.g., see Johnson, 1994; Smyth and Combi, 1988), the sputtering of the Martian atmosphere by *pick-up ions* was first described by Luhmann and Kozyra (1991). On the dayside, solar EUV and photo-electrons ionize a fraction of the neutral atmosphere. Because of the absence of a large permanent magnetic field, these newly ionized particles are picked-up by the partially deflected field frozen in the solar wind. They are accelerated along gyroradial trajectories in the Martian tail direction and some of them reimpact the

* Correspondence address: Service d'Aeronomie du CNRS, 91371 Verrieres le Buisson Cedex BP 3, France. Tel.: +33-1-64474303; fax: +33-1-69202999.

E-mail address: francois.leblanc@aerov.jussieu.fr (F. Leblanc).

0032-0633/01/$ - see front matter
PII: S0032-0633(01)00003-4

neutral atmosphere and collide with neutral particles with sufficient energy to generate new ejecta and ballistic particles. The net loss of atmosphere due to this mechanism has been estimated in several papers (Jakosky et al., 1994; Kass and Yung, 1995, 1996; Johnson and Liu, 1996; Johnson et al., 2000). But these have all been 1D models and the potential feedback processes been treated approximately (Johnson and Luhmann, 1998). That is, the sputtered neutral particles increase the corona density where they can be ionized and can form additional pick-up ions liable to reimpact the neutral atmosphere. Johnson and Luhmann (1998) estimated the density of sputtered particles in the corona and concluded the feedback process could significantly increase the loss of atmosphere due to the sputtering during the period from the time the Martian magnetic field decayed to the present. The principal purpose of this work is to use a 3D Monte Carlo simulation to test the validity of the 1D models of atmospheric sputtering. Therefore, we provide a new estimate of the average number of ejected neutrals generated by an incident particle. A second purpose is to re-calculate the importance of the coronal density enhancement produced by sputtering. The 3D approach allows us to treat all incident angles for the pick-up ions and to determine the horizontal structure of the sputter-produced corona.

In Section 2 we describe the Monte Carlo models and in Section 3 we quantify the effect of sputtering in two epochs. These results depend critically on the pick-up ion fluxes initially estimated by Zhang et al. (1993a) but the calculated yields can be applied when more accurate pick-up ion fluxes are available. In the last section, we present our conclusions.

2. Monte Carlo simulations

In order to model the interaction of the incident pick-up ions with the Martian neutral particles we developed a 3D Monte Carlo (MC) simulation which follows the incident particles when they penetrate the neutral atmosphere and describes the cascade of collisions and recoil particles that are generated. We also use a 2D Direct Simulation Monte Carlo (DSMC, see Bird, 1994) to describe more accurately the effect of this interaction on the atmosphere below the exobase. Both approaches are based on the same model. Since the number of real particles in the atmosphere or in the incident flux is too large, in the Monte Carlo approach one assumes that a simulated particle represents a large number of real particles. In this work, each tracked particle typically represents between 10^{25} and 10^{28} real particles. This number, called the weight of the particle, is an important parameter of the simulation since it also fixes the density of simulated recoil particles due to the flux of incident particles. These in turn determine the accuracy with which the cascade of energy due to an incident particle is described. Rather than use different weights for these particles we break the problem into two pieces because the majority of the particles are slow and it is primarily the few fast ones that populate the corona. In the 3D Monte Carlo simulation, only particles with a kinetic energy above a certain energy are followed in order to limit the computing time and low energy particles, roughly, an energy lower than 0.1 eV, are not followed in this model. The main consequence of such simplification is that the MC model does not self-consistently describe the heating of the neutral atmosphere due to the bombardment by the keV ions. In the present epoch, this assumption is justified by the fact that UV and EUV heating are more efficient. Therefore, for the present period, we use the model of atmospheres developed by Zhang et al. (1993a), Bougher and Roble (1991) and Bougher et al. (1999). However, for the period 2.5 Gyr ago, the sputter heating can have an important effect on the altitude of the exobase, mainly in the subsolar regions (e.g., Pospiezalska and Johnson, 1992, 1996; Wong and Johnson, 1995). This increase of the exobase radius implies that a larger surface area of Mars atmosphere is submitted to the flux of pick-up ions and it can change the position of the population above the ionopause. We determine the heating of the atmosphere by the energetic incident particles using a 2D-DSMC code. In this model, all the simulated particles are followed. Starting with the Zhang et al. atmosphere (1993a) for the earlier epochs, the results of this simulation define the neutral atmosphere for the 3D Monte Carlo model. The latter is then used to determine the characteristics of the sputtered population (coronal density enhancement and sputtering yield).

The incident pick-up ions, mainly O^+ and H^+ (Luhmann and Kozyra, 1991; Brecht 1997a, b), are efficiently neutralized before reaching the exobase, which is defined as the altitude above which collisions are unlikely (Luhmann and Kozyra, 1991; Johnson et al., 2000). Indeed, the collision cross section for charge exchange is typically one order larger than that for "knock-on" collisions (collision which generates significant exchange of momentum). Our Monte Carlo models then describe high-energy incident neutral particles moving only under the effect of the Mars gravity field and interacting with low-energy neutrals in the Martian atmosphere. We follow these particles above an altitude from the Mars surface such that any incident particle reaching this altitude has only a small effect. Such particles generate only recoil particles which will be thermalized quickly and do not contribute to the heating near the exobase, to the corona or to the total escape (Johnson et al., 2000).

In both MC and DSMC approaches, the collisions are binary. Such collisions can happen between two fast particles (recoil or incident particles) or between a fast particle and a background atmospheric particle. The number of collisions is determined by our choice of maximum impact parameter, 2.5 Å . This impact parameter was chosen to optimize the sputtering calculation as discussed in Johnson et al. (2000). It corresponds to an energy transfer greater than 50% of the energy of the incident particle in the case of a 0.07 eV particle, greater than 5% in the case of a particle of 2 eV and greater than 0.5% for a 20 eV particle. We then neglect the collisions which are associated with small energy transfer.

The particle tracking is described following the algorithm of Bird (1994). The domain where we follow the particles is divided into cells and the motion of the particles is described as follows:

- At a time t of the simulation, we calculate the displacement of the fast particles for a chosen time step $\mathrm{d}t$. $\mathrm{d}t$ is chosen in order that the fastest particle moves less than the smallest distance between two boundaries of the smallest cell.
- At $t + \mathrm{d}t$, in each cell, the probability that a collision occurs between fast/fast or fast/background atmospheric particles is determined from the volume of the cell, the effective collision cross section for fast/fast and fast/background particles, respectively, the densities in the cell of both types of particles, the average relative speed between the particles and the time step $\mathrm{d}t$. The average relative speed between particles is sampled during the simulation. From the computed probability, the average number of pairs which collide is determined. Pairs of colliding particles are selected randomly and a test on their relative speed is made to reproduce the average number of collisions which should statistically happen between the time t and $t + \mathrm{d}t$ in each cell (Bird, 1994). If a collision is selected, the velocities of the two particles after a collision is calculated using a universal potential of interaction (Johnson et al., 2000; Ziegler et al., 1985). An electronic energy loss (Johnson et al., 2000; Firsov, 1959) is also included which takes into account the energy lost by the colliding particles due to interaction with the electrons of each impactors. When all the collision calculations which should occur have been performed, all the particles are moved ballistically for a new period $\mathrm{d}t$ without encountering further collisions.

The particles are followed and the collisions are described as long as these particles remain between the lower boundary (140 km for present time and 150 km for earlier epoch) and an arbitrary altitude (400 km for present time and 600 km for earlier epochs) chosen higher than the exobase altitude. When they cross the lower boundary, they are suppressed, since we consider that at such an altitude these particles are quickly thermalized. When they cross the upper boundary, they become ballistic particles which are subject only to the Martian gravity field. These coronal particles are followed until they either again cross the upper boundary of the domain that we consider as collisional where they are reintegrated inside the population of colliding fast particles, or until they reach a distance of $2R_{\mathrm{m}}$ from the Mars surface above which we suppress them. Indeed, these particles represent only a small part of the coronal particles and have a non-negligible probability for ionization or drag by the solar wind before returning to $2R_{\mathrm{m}}$. Actually, around $1/3$ of them are ionized by photo-ionization, charge exchange or electron impact before coming back below $2R_{\mathrm{m}}$, but their contribution to the corona density when they come back in it is small and is neglected in the following results.

The main difference between the 3D-MC simulation and the 2D-DSMC simulation is that low-energy particles (typically energy below 0.1 eV) are assumed to be frozen in the former. Indeed one of the goals of this work is to determine the importance of the potential feedback process: ionization of the neutral particles sputtered into the corona. This requires a 3D approach because Mars' rotation changes significantly the corona. A full 3D-DSMC which includes the coronal particles and describes the heating of the thermosphere is computationally too expensive at present. Luhmann and Kozyra (1991) show that the newly ionized particles which can reimpact and efficiently sputter the neutral atmosphere mainly originated from altitudes much higher than the exobase in order for them to be accelerated sufficiently before colliding with atmospheric particles. Zhang et al. (1993a) in an overly simple field model assume that the ionization mainly takes place above the ionopause. These authors considered photo-ionization, charge exchange and electron impact, with the last process being the most efficient process for ionization of the O atom in the Martian corona. Therefore, we need to describe accurately the enhancement of the corona density above the ionopause. For this reason, in the 3D MC code, we only follow the particles which have enough energy to reach such an altitude. In the present epoch, the ionopause is estimated to be roughly at 300 km at the subsolar point and 700 km in the pole (Zhang et al., 1993a; Hodges, 2000). For the earlier epoch, the ionopause altitude should be similar to the present one since the atmosphere is thicker than at present (roughly one-order higher density at 300 km) but with a 1–3 times more intense solar wind (Johnson and Luhmann, 1998; Zhang et al., 1993a). The energy threshold, we used in this simulation, is defined such that a particle of lower kinetic energy than some fraction of the escaping energy is not followed. This threshold fixes the average number of recoil particles which have to be followed at each time step, which is directly linked to the computing time of the simulation. This in turn determines an altitude above which the 3D aspects of the corona are well described.

3. The 1EUV and 3EUV epochs

Zhang et al. (1993a) have defined three different periods for the solar activity history. For each period the flux of solar EUV radiation is given in terms of the present EUV flux at solar minimum (1EUV). The 6EUV period was assumed to correspond to a very early epoch of the solar system with an EUV flux 6 times greater than the present. During this period Mars may have had a field (Acuña et al., 1999) and if not feedback processes dominate the interaction between pick-up ions and Mars' atmosphere. We will then not consider this period in this present work. The 3EUV flux is assumed to correspond roughly to the period

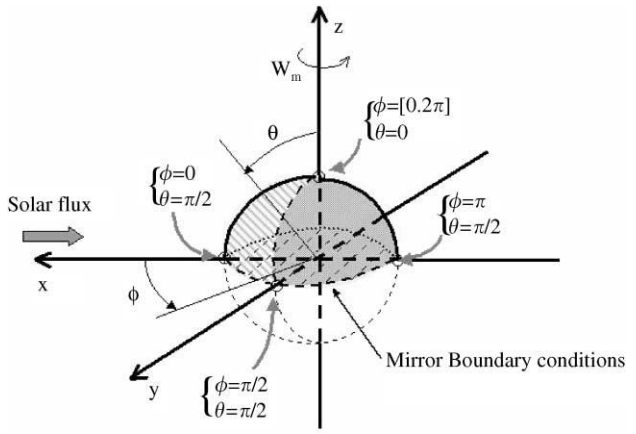

Fig. 1. Coordinate system. The half-upper part is the domain of calculation, the clear dashed part is the domain bombarded by the incident flux and the darker part the non-bombarded one.

around 2.5 Gyr ago at a time at which the magnetic field had already decayed leaving a remanent field not strong enough to protect the Mars atmosphere from the incident pick-up ions. In this paper, the effect of the pick-up ions on the atmosphere is calculated at 1EUV and 3EUV.

The real flux of pick-up ions is fully 3D according to Brecht (1997a, b). However, for comparison with earlier results, we used the over-simplified model defined by Zhang et al. (1993a) and used by Luhmann et al. (1992). This flux is assumed to be uniform across a surface perpendicular to the Mars/Sun axis. The particles are all assumed to be O^+ ions with 1 keV energy. Any contribution of H^+ is neglected because these particles have a small sputtering effect since their mass is small compared to the average mass of the atmospheric particles (Luhmann and Kozyra, 1991). However, the proton flux may add considerable heat (Kallio and Luhmann, 1997; Brecht, 1997b) and will be described in future work. The flux used in this simulation is equal to 5×10^5 ions cm^{-2} s^{-1} for the present epoch at solar minimum (1EUV) and 9×10^7 ions cm^{-2} s^{-1} for the period starting 2.5 Gyr ago (3EUV). The later is larger than present solar maximum conditions.

Since the flux of pick-up ions used in this simulation is symmetric around the Sun/Mars axis, we will simulate half the atmosphere of Mars and will introduce a mirror boundary condition to describe the whole atmosphere (Fig. 1). A 3D model is needed because calculating the coronal density due to the sputtered population implies that several Mars days have to be simulated since the coronal particles have velocity of the order of Mars' rotation. Therefore, Mars' rotation has been introduced. The coordinate system (x, y, z) is centered on Mars with x always directed towards the Sun and z with the same orientation as Mars' rotation vector. In this paper, we neglect the 25° of obliquity of the Mars rotation axis with respect to the ecliptic plane. Indeed, over one year the average angle between the rotation axis and the Solar incident flux is equal to 90°. A test including the obliquity of Mars has been performed and has provided no significant changes compared to the case without it. We also defined a spherical coordinate system (r, θ, ϕ) where r is the distance from the center of Mars, θ is the angle between the vector considered and the z axis and ϕ is the angle between the x axis and the projection of this vector on the Mars orbital plane (xy). θ varies from 0 at the pole to $\pi/2$ at the equator and ϕ varies from 0 at the subsolar point (12 a.m. Mars local time) to 2π. $\phi = \pi$ corresponds to 12 p.m. Mars local time. $\theta = \pi/2$ corresponds to the plane defining the mirror boundary condition (Fig. 1).

The space around Mars in which we follow the particles is divided into cells which are distributed exponentially in the radial (r) direction (in order that each cell contains roughly the same number of atmospheric particles) and equally distributed in the latitude (θ) and longitude (ϕ) directions, respectively, 20, 20 and 80 cells for the radial, latitudinal and longitudinal directions in the domain penetrated by the ions. Above this domain until $2R_m$ the space is covered by a grid distributed in the same way and with $50 \times 20 \times 80$ cells. The calculation of the density of fast particles and the collision rate in the 3D MC model is based on the assumption that a stationary state of the ballistic and escaping particles is reached after a few Mars rotations (typically, three rotations is enough to reach such a stationary state). The time spent and the collisions made by fast particles inside a cell are accumulated. The final density of particles and rate of collision are then calculated at the end of the simulation by dividing by the total time of the simulation and by converting the result into a real density by taking into account the volume of the cells and the weight of the simulated particles.

3.1. *1EUV*: *present solar minimum conditions*

The Martian atmosphere in the present epoch (solar minimum) has been characterized by Mars Thermospheric Global Circulation Model (MTGCM) 3D upper atmosphere–ionosphere model of Bougher et al. (1991, 1999) and by Zhang et al. (1993a). The MTGCM model includes the effect of the UV and EUV solar flux, the chemistry for CO_2, CO, N_2, O_2, Ar and O molecules and O_2^+, CO_2^+ and O^+ ions and the global dynamics between the dayside/nightside for solar maximum and minimum. Zhang et al. (1993a) use a 1D model which describes the density and temperature profiles of the atmosphere based on a 1D two stream model (Nagy and Cravens, 1988). In this paper, we use the density profiles of Zhang et al. (1993a) and the temperature profiles of Bougher et al. (1999) to define the neutral atmosphere. These profiles are in agreement with the measurements made by the Martian probes and by airglow measurements. A dayside/nightside dependence of the temperature is introduced following the work of Bougher et al. (1999). We simplify the description of the neutral atmosphere by assuming that a molecule of CO_2 is equal to three individual atoms of oxygen and then by considering

that the atmosphere consists only of O atoms. This assumption is discussed in Johnson et al. (2000) and in the corona for the 1EUV epoch O oxygen dominates CO_2 above 205 km (Zhang et al., 1993a). The corresponding density and temperature are deduced from CO_2 and O profiles. This last assumption implies that the loss estimate presented in this work is an upper bound of the real one (Johnson et al., 2000).

The 3D collisional domain described for the fast recoils is between 140 and 400 km in altitude, above which particles are in ballistic trajectories. The background Martian atmosphere consists of 5×10^{33} real particles. For this simulation, we used a weight factor equal to 2×10^{25} and followed 2898 incident particles simulating the bombardment of the atmosphere. The energy threshold used is $\sim$ 0.07 eV (or $\sim$ 700 K). For comparison, the exospheric temperature on the dayside as determined by Bougher et al. (1999) is equal to 209 K. This energy means that the corona is accurately described above 320 km since we follow all the particles which have enough energy to reach such altitude. This estimate is close to the ionopause altitude, 300 km, calculated in Zhang et al. (1993a).

Because collisions above 200 km primarily affect the lower energy recoils, the standard estimate of the exobase altitude applies, $n/(\sigma_d H) = 1$, where n is the density, σ_d is the momentum transfer cross section between the fast and the atmospheric particles and H is the scale height of the neutral atmosphere. For the potential used here and for (0.2, 20 eV) particles, which is the range of energy of the coronal fast particles, σ_d is equal to $\sim 7 \times 10^{-16}$ cm^2 (Johnson et al., 2000). On the dayside, $H = 22$ km and the exobase altitude is $\sim$ 180 km (172 km according to Zhang et al., 1993a). Actually most of the fast recoil particles reaching 400 km in altitude have a radial speed around 2.0 km s^{-1} at 200 km in altitude which corresponds to an energy of 0.4 eV. Fig. 2 presents the average number of collisions that a fast recoil particle of radial speed 2.0 km s^{-1} at a given altitude makes in our model between this altitude and 400 km. The solid line without a symbol has been obtained by averaging over the half sphere, the circles correspond to the polar regions ($\phi = [0, 2\pi]$ and $\theta = \pi/2$), the cross line to the subsolar region ($\phi = 0$ and $\theta = \pi/2$) and the square line to the nightside region ($\phi = \pi$ and $\theta = 0$). From Fig. 2 an escaping particle of radial speed 2.0 km s^{-1} will make 0.3 energy transfer collisions before reaching 400 km losing, on the average, $\sim$ 20% of its energy, roughly consistent with the definition of an exobase. The estimate of the number of collisions above a given altitude varies across Mars being smaller in the polar and nightside regions which is due to the position and angle distribution of the primary impacting particles.

Fig. 3 presents the number of collisions of the fast recoils (> 0.07 eV) cm^{-3} s^{-1} (gray lines with opened symbols) and the density of recoils (dark lines with filled symbols) at three different positions in the atmosphere as a function of the altitude. The three positions are indicated using the same legend as in Fig. 2. It is seen that below 160 km the fast recoil population decreases due to rapid thermalization. A comparison between the three gray lines shows that most of the collisions happen in the subsolar regions below 200 km with a peak at 180 km for the polar and nightside regions. The dark solid lines indicate that the density of recoils in the polar regions is as high as in the subsolar regions for altitudes greater than 250 km. This illustrates the fact that the sputtered particles in the subsolar regions are redistributed in the direction of the nightside hemisphere. That is, the

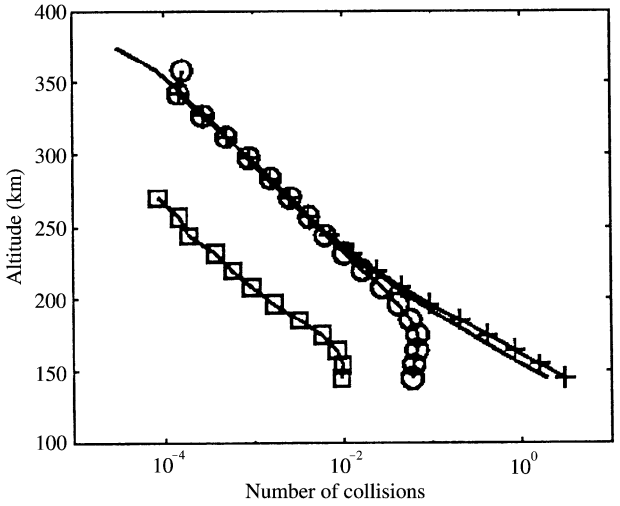

Fig. 2. Present epoch (solar minimum-3D MC simulation): Number of collisions made by a recoil particle of radial speed 2.0 km s^{-1} moving from the indicated altitude to 400 km. Solid line without symbol: average on the half-sphere of the domain of calculation. Line with crosses: subsolar regions. Line with circles: polar regions. Line with squares: midnight regions.

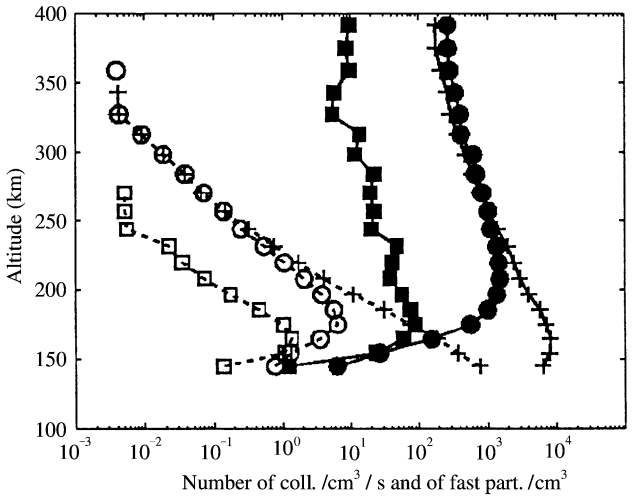

Fig. 3. Present epoch (solar minimum-3D MC simulation): Number of collisions per second and cm^3 (gray lines with opened symbols) and density of recoil particles per cm^3 (dark lines with filled symbols) as a function of the altitude. The different lines correspond to subsolar point (crosses), polar (circles) and nightside (squares) regions.

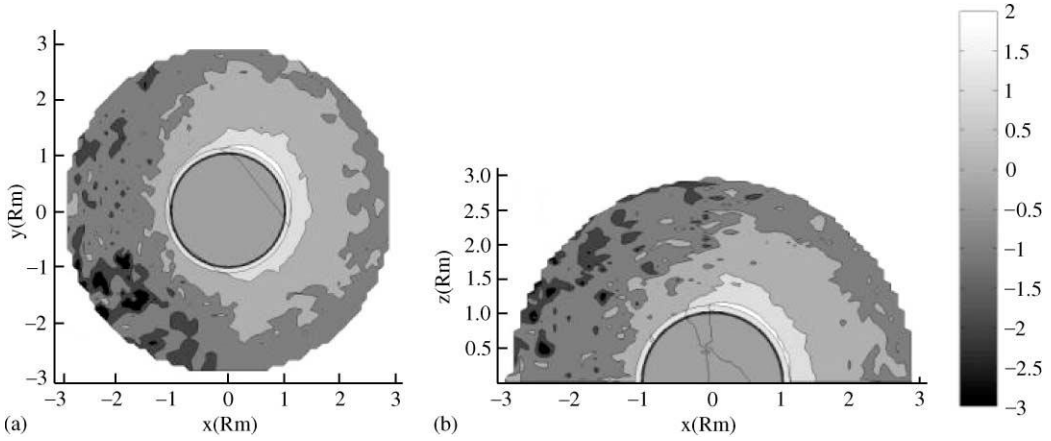

Fig. 4. Present epoch (solar minimum-3D MC simulation): Density of sputter produced coronal particles. The scale bar is in base 10 logarithm of the density in cm^{-3}. (a) In the equatorial plane. (b) In a plane perpendicular to the equatorial plane along the Mars/Sun axis. Mars is represented by the full and half circle in the center of each figure. Mars' rotation vector is directed towards the reader in Fig. 4a and towards the positive z axis in Fig. 4b.

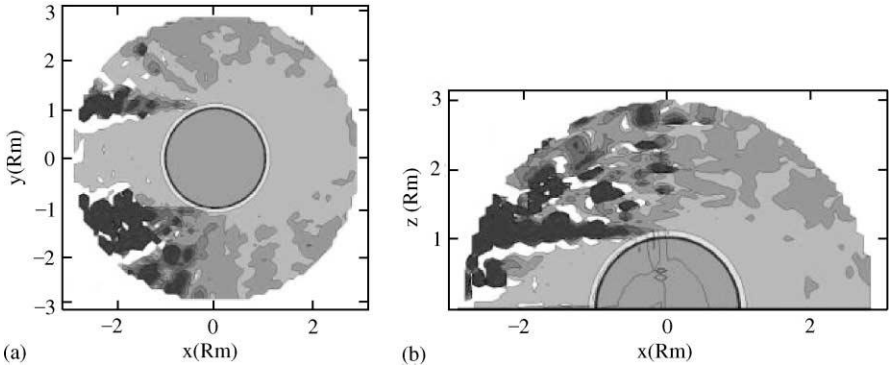

Fig. 5. Present epoch (solar minimum-3D MC simulation): Average energy of sputter produced coronal particles. (a) In the equatorial plane. (b) In a plane perpendicular to the equatorial plane along the Mars/Sun axis. The clearest regions correspond to energy lower than 2 eV (less than the escaped energy), the following darker one to energy larger than 2 eV and lower than 4 eV. The darkest regions corresponds to energy from 4 eV to more than 40 eV.

density in the nightside is due to ballistic particles which return to the atmosphere.

The simulation was also used to obtain the sputtering yield, which is the ratio of the total number of escaping particles to the total number of incident ones reaching the exobase. This was found to be equal to 2.9 (8387 simulated particles escaped due to 2898 incident particles). For comparison, Johnson et al. (2000) used a 1D model with the same scattering function, and obtained ~ 3.9 but added in ~ 1.3 due to the single collision ejection in the corona, directly treated here. In fact, we find that although a large proportion of the escaping particles come from atmospheric regions bombarded by particle with tangential incident angle (from regions around $\phi = \pm\pi/2$ and $\theta \in [\pi/5, \pi/2]$), the analytic correction made by Johnson et al. (2000) is too large.

The calculation also gives the enhancement in the corona density due to the sputtering of the neutral atmosphere by pick-up ions. Fig. 4 presents the density in the corona in two planes: the (xy) plane in Fig. 4a and the (xz) plane in Fig. 4b. The order of magnitude is in good agreement with the 1-D analytic results of Johnson and Luhmann (1998) for the subsolar regions and low altitudes. For present solar minimum conditions, the enhancement of the density due to the sputtering is negligible compared to dissociative-recombination of ionospheric O_2^+ (Zhang et al., 1993b; Kim et al., 1998). But Fig. 4a shows that due to the rotation of Mars the coronal density is not spherically symmetric and is somewhat larger in the dusk regions than in the dawn regions. The results also show that in the polar regions (Fig. 4b) and in the dusk regions (Fig. 4a) the enhancement in the density due to the sputtering has to be taken into account in order to accurately estimate the direct sputter loss from these regions.

Fig. 5 presents the average local kinetic energy of the sputtered particles inside the corona for the same two planes as in Fig. 4. The average energy in the corona (between the exobase and $2R_m$) is 15 eV per particle with the escaping particles having an average energy equal to 24 eV and the bound ballistic particles of average energy equal to 0.26 eV. Out of the equatorial regions, the high-energy

particles are in the "shadow" of the polar regions and the ballistic ones populate the subsolar regions and the tail. In the equatorial regions, the high-energy particles are mainly in the "shadow" of the dawn and dusk regions. The high-energy particles are more spread in the dawn regions than in the dusk regions because of Mars' rotation. This confirms the discussion above that sputtering efficiently produces energetic recoils and escaping particles when incident particles penetrate the atmosphere with angles close to tangential. It emphasizes, therefore, the importance of accurately describing these regions. The absence of a correlation between Figs. 4 and 5 indicates that the main component of the corona is the low energy particles. Compared to the results of Zhang et al. (1993a) and Kim et al. (1998), the population in the corona due to the sputtering might be distinguished by an in situ measurement from the population produced by the dissociative-recombination of O_2^+ ions, for which the energies are lower than 5 eV.

3.2. *3EUV*: $\sim$ 2.5 Gyr *ago*

3.2.1. *Description of the atmosphere heating due to the sputtering*

In order to first describe the heating of the neutral atmosphere due to the bombardment by pick-up O^+ ions we used a 2D DSMC model without any energy threshold below which particles are not followed. 6×10^6 simulated particles were followed each representing 10^{19} atmospheric atoms, considered all to be O atoms as discussed. Since a 2D model does not allow us to take into account Mars' rotation, we limit our simulation to 5.6 h (a 1/4 of one Mars' rotation) in order to use a realistic total incident flux. The 2D is convenient since heating due to the bombardment can be obtained with fewer incident particles than needed to estimate the loss of atmosphere or the enhancement of the coronal density. Indeed, the diffusion of the energy inside the atmosphere is determined by all the atmospheric particles whereas only few sputtered particles reach the exobase altitude for each incident particle. 100 incident particles are simulated which is enough to determine the difference between the polar and equatorial regions due to the geometry of the incident flux. The results are applicable above the exobase, but do not provide an accurate description of the full corona ($\sim 2R_m$). To obtain such results several days have to be simulated to take into account Mars' rotation.

The domain used in this simulation corresponds to the plane $\theta = \pi/2$ in Fig. 1 (plane xy). We limit the calculation to a region between $\phi = 0$ (subsolar regions) and $\phi = 4\pi/5$ (nightside regions) and between 150 and 600 km with ballistic particles tracked up to $2R_m$. Since the incident flux is symmetric around a Sun/Mars axis, we only describe half the atmosphere. The collisional domain between 150 and 600 km and $\phi = 0$ and $\phi = 4\pi/5$ is divided into 30×20 cells, respectively, in the radial and latitudinal directions. Mirror conditions are used at $\phi = 0$ and $\phi = 4\pi/5$. A mirror condition at $\phi = 4\pi/5$ has no influence on the result since no significant effect due to the sputtering are observed on the nightside of the atmosphere. Actually, the heat flux on the dayside is null largely before $\phi = 4\pi/5$. The particles which cross either 150 km or $2R_m$ from the surface are suppressed. The particles between 600 km and $2R_m$ from the surface are assumed to be ballistic and do not collide until they return to 600 km. Boundary conditions deduced from the solar maximum condition solution of Bougher et al. (1999) are imposed at 150 km. A temperature from 280 K on the dayside to 200 K on the nightside and a constant density of 10^{10} cm^{-3} of O atoms are used to define these boundary conditions. These bottom boundary conditions allow us to take into account the EUV and UV heatings which occur below 150 km. The initial atmosphere is deduced from the 3EUV solution of Zhang et al. (1993a). This choice does not significantly influence the results.

Figs. 6a and b present the profiles of temperature after 5.6 h of bombardment. Fig. 6a presents the profiles of the temperature with respect to the altitude r along the 5 lines plotted in Fig. 6b, which also gives the 2D distribution of the temperature. The large gray half-sphere in Fig. 6b is Mars' surface whereas, the half-dark ring is the Martian atmosphere below 150 km in altitude which is not influenced by the incident flux. Above this ring, the calculated profile of the temperature is plotted. On the left side of the Fig. 6b (dayside), temperatures higher than 300 K (up to $\sim$ 350 K) are obtained, whereas, on the right side around $\phi = 4\pi/5$ a nearly isothermic temperature profile of 200 K is found, confirming that this side of the atmosphere is not influenced for the bombardment geometry assumed here. On the dayside, there is a variation of $\sim$ 80 K going from 150 to 600 km. The heating produced by the incident flux is seen to be of the same order or higher than the one produced by the EUV and UV heating for the 3EUV epoch as predicted by Zhang et al. (1993a). Indeed, we obtained a dayside exospheric temperature of 350 K whereas these authors found a temperature of 330 K. The heating due to the bombardment extends below 220 km in the subsolar regions (in good agreement with Luhmann and Kozyra, 1991).

Fig. 7 provides the profiles of the density vs. the altitude for different regions of the atmosphere. The different lines correspond to the legend used in Fig. 6a. The squares (nightside) correspond to the isothermic profile shown in Fig. 6a. The density in the subsolar region, represented by the crosses, is more than an order of magnitude higher at 600 km than the one predicted profile (solid line without symbol, Zhang et al., 1993a). From Figs. 6a and 7, we can estimate the exobase altitude again using $n/(\sigma_d H) = 1$. For the sputtered particles (0.2, 20 eV) the corresponding momentum transfer cross section is equal to 7×10^{-16} cm^2. The scale height H deduced from the temperature profile is equal to 28 km on the dayside and to 47 km on the nightside and the theoretical exobase altitude from the previous definition is then equal to about 300 km on the dayside (216 km according to Zhang et al., 1993a) and around

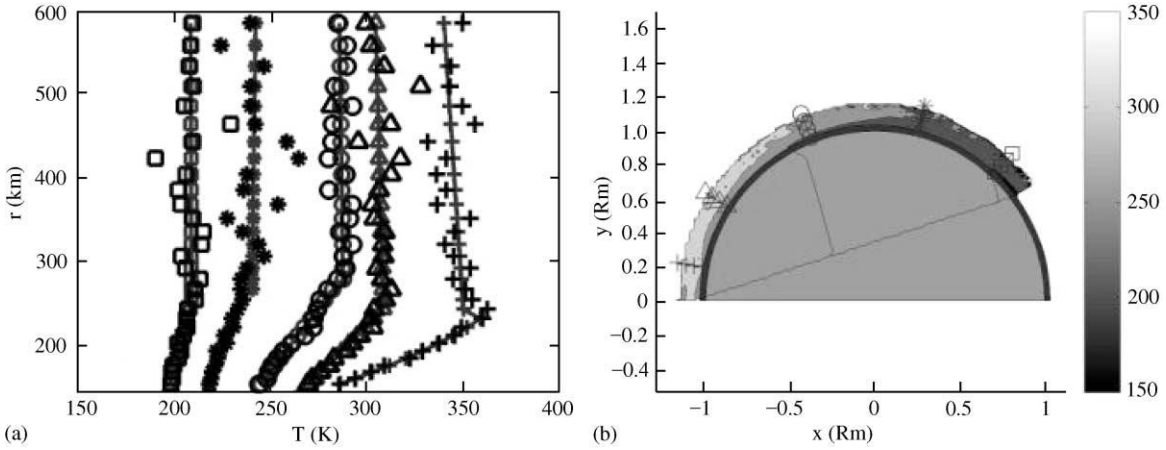

Fig. 6. Earlier epoch (2D DSMC simulation): (a) Temperature profiles of the neutral atmosphere with respect to the altitude. The different lines are profiles obtained at different latitudes of the Mars atmosphere as indicated in (b) (from left to right at $\phi = \pi/20$, $\pi/5$, $2\pi/5$, $3\pi/5$ and $3\pi/4$). (b) 2D temperature profiles of the neutral atmosphere (in K).

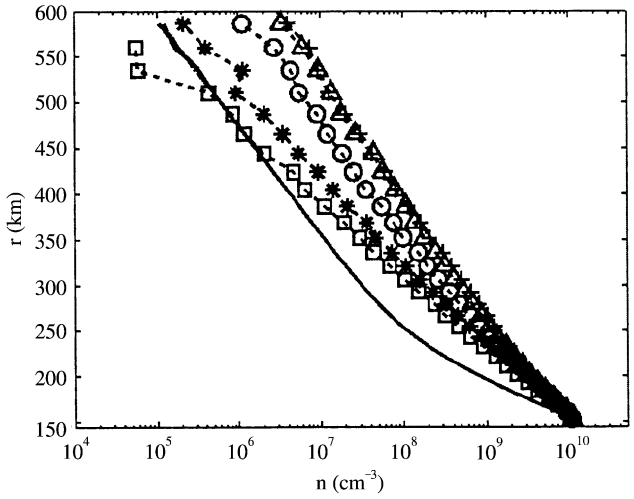

Fig. 7. Earlier epochs (2D DSMC simulation): Density profiles of the neutral atmosphere. Same legends as in Fig. 6. The solid line without symbol is the profile obtained by Zhang et al. (1993a) for the same epoch.

260 km on the nightside. Global high wind speeds are associated with the difference in density from day to night sides. The region of highest latitudinal speed is at 600 km just before the polar region and corresponds to winds of 156 m s^{-1}. For comparison, Bougher et al. (1999) have obtained for solar maximum wind speed of 260 m s^{-1} in the same region. The sputtering process at this epoch is, therefore, nearly as efficient as the EUV and UV flux in determining the properties of the exobase region and could be also important at solar maximum.

3.2.2. *The sputtered population*

A model of the Mars atmosphere heated by the incident flux in the 3EUV epoch was described in the previous subsection. In order to fully describe the corona up to $2R_m$ a 3D calculation is needed. For night to dayside, we used the profiles of the density and temperature which are interpolated from the solar maximum condition profiles of Bougher et al. (1999) for the nightside and the dashed lines with crosses in Figs. 6a and 7 for the dayside. On the nightside, the estimate made by Bougher et al. (1999) is more accurate than the results obtained above which ignore heat loss other than by atom transport and escape, whereas, on the dayside the heating due to the sputtering dominates. As for the 1EUV case, here 3.0 rotations of Mars have been simulated to reach a stationary state in terms of coronal density and yield. The background atmosphere consists of 2×10^{34} O particles and the weight factor used in this simulation is 3×10^{27}. 3112 incident particles were followed and the energy threshold used here is 0.11 eV (or $\sim$ 1100 K). For these tracked recoils the coronal density is accurately estimated in this simulation above 610 km. Below this altitude, the coronal density can be extrapolated from the results obtained in the previous subsection.

Fig. 8 gives the number of collision of an energetic recoil particle moving with a radial speed of 2.35 km s^{-1} from a given altitude to 600 km. The same legends defined in Fig. 2 are used. Based on the same maximum impact parameter as in the 1EUV simulation, Fig. 8 shows that a fast recoil particle of radial speed 2.35 km s^{-1} will make in an average 0.64 collisions between 300 and 600 km. The 2.35 km s^{-1} speed is the average speed of the ballistic particles reaching 600 km and corresponds to a kinetic energy of 0.46 eV. Around 0.2 eV is needed for a particle to go from 300 to 600 km and 0.26 eV is the average energy of the ballistic particles inside the corona. For the universal potential (Ziegler et al., 1985) and the cutoff chosen (2.5 Å), above 350 km 0.46 eV atoms will make 0.2 collisions with atmospheric atoms losing on the average 10% of its energy. From the exobase altitude

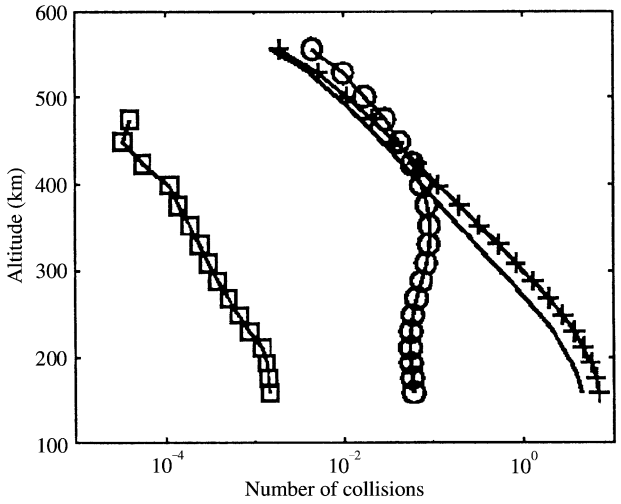

Fig. 8. Earlier epoch (3D MC simulation): Number of collision made by a recoil particle of radial speed 2.35 km s^{-1} in function of the altitude. Same format as in Fig. 2.

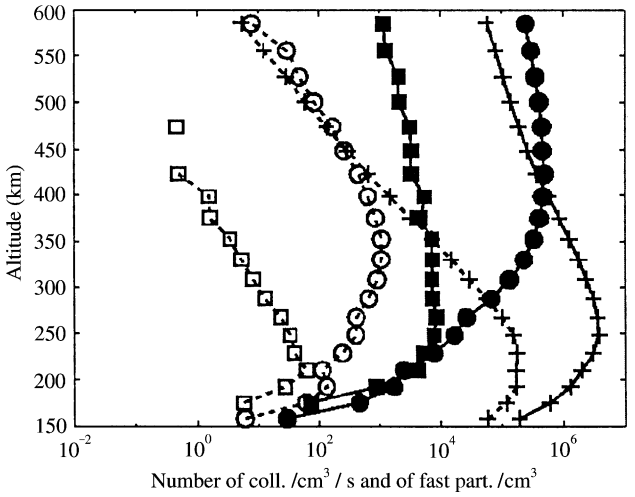

Fig. 9. Earlier epoch (3D MC simulation): Number of collision per second and cm^{-3} and density of recoil particles in cm^{-3} in function of the altitude. Same format as in Fig. 3.

(216 km) estimated by Zhang et al. (1993a), an atom of radial speed of 2.35 km s^{-1} will make on the average 3.0 collisions and lose all its energy before reaching 600 km.

In Fig. 9 the fast recoil densities are associated with the subsolar (crosses), polar (circles) and nightside (squares) regions. As in Section 3.1, the collisions between fast and atmospheric particles mainly occur in the subsolar regions. The remarks made for the density profiles in the subsolar and polar regions apply here also. The regions above the Martian poles are the most populated regions in the upper corona. The densities are not only higher than in the 1EUV but ballistic particles from the dayside contribute $\sim 10^3$ cm^{-3} at 610 km on the nightside. The density profile of fast particles thermalizes rapidly at low altitudes, so that the regions influenced by the pick-up ion bombardment are above 180 km.

The calculated yield is equal to 3.6 (corresponding to 11256 escaping for 3112 incident). The 1D model of Johnson et al. (2000) gave a yield of ~ 2.9 corrected to ~ 4.2 for coronal ejection. Again the correction to take into account the single collision sputtering was overestimated. As in the 1EUV case, the origin of the ejected atoms is in a large part from regions bombarded by grazing particles.

Figs. 10 and 11 present the density and energy of the coronal population issued from the bombardment of the atmosphere above 610 km in altitude. As for Figs. 4 and 5, cuts in the planes (xy) and (xz) are provided. The morphology of the coronal density is very similar to the 1EUV case (Fig. 5), but the density is from 2 to 3 order higher in the corona in this epoch. The maximal density of fast particles at 600 km is 10^5 cm^{-3} in the plane (xz) (Fig. 10b) and corresponds to polar regions as shown in Fig. 9. In the plane (xy) (Fig. 10a) the maximum is $10^{5.7}$ cm^{-3} and corresponds to dusk regions. This is much larger than the analytic 1D calculation made by Johnson and Luhmann (1998) who found $10^{4.5}$ cm^{-3} at 600 km in the subsolar region using the Zhang et al. (1993a) atmosphere. It is also consistent with the results obtained in previous subsection (Fig. 7) showing density of 10^6 cm^{-3} at 600 km on the dayside. The 3D picture of the corona density allows us to determine the density in the nightside and to obtain a more realist dependence of the density on altitude. For instance, Johnson and Luhmann (1998) overestimate the density by an order of magnitude at $2R_m$, whereas we find a higher density in the equatorial dusk regions at 600 km as shown in Fig. 10a. The density inside the corona due to the sputtered particles is shown here to be of the same order of magnitude as the density of oxygen due to the dissociative recombination of the O_2^+ (Zhang et al., 1993b), and is even higher than that estimate in the polar and dusk regions. Fig. 11 provides the local average energy of the coronal particles. As for Fig. 5, the asymmetry of the energy distribution is due to Mars' rotation. The average energy for the recoil particles inside the corona is equal to 9 eV, with an average energy for the escaping population equal to 13 eV in (xz) and to 9 eV in (xy) and for the ballistic population equal to 0.26 eV. As for the 1EUV case, the highest energy atoms are in the shadow of the dawn and polar regions and the lowest energetic particles populate the tail and the subsolar regions. The average energy at the 3EUV epoch is lower than at the 1EUV epoch because the energy of the incident particles is more easily spread in the denser atmosphere. The lack of a correlation between Figs. 10 and 11 indicates that the lowest energy particles are the main population in the corona. As for the 1EUV case, the average energy of the population sputtered in the corona is significantly higher than that from dissociative-recombination < 5 eV for present solar maximum conditions (Kim et al., 1998). It may be possible to distinguish the neutral coronal population produced by the sputtering from the dissociative-recombination corona during solar maximum, which is roughly an intermediate case between the 1EUV and the 3EUV epochs. The neutral

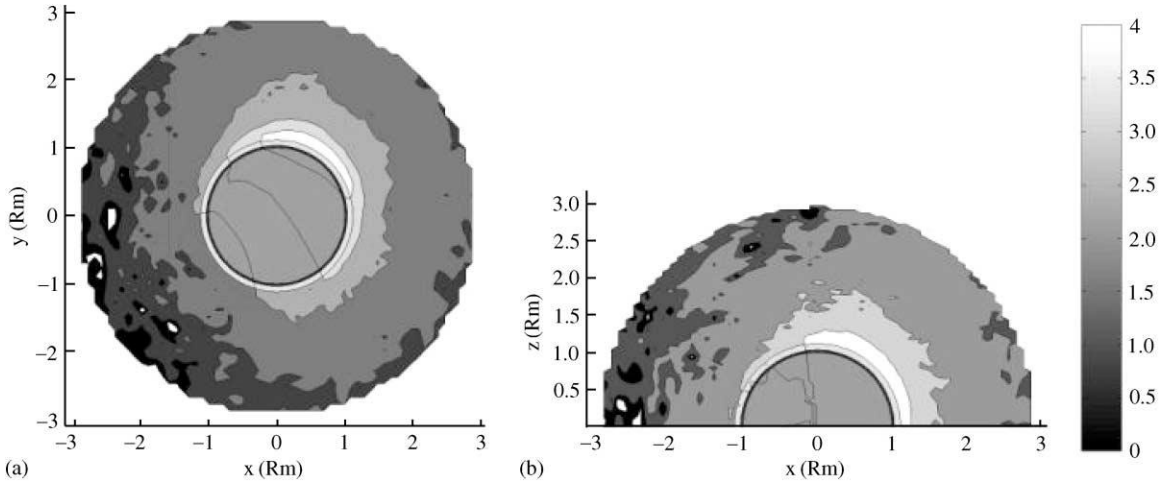

Fig. 10. Earlier epoch (3D MC simulation): Density of sputter produced coronal particles. The scale bar is in base 10 logarithm of the density in cm^{-3}. (a) In the equatorial plane. (b) In a plane perpendicular to the equatorial plane along the Mars/Sun axis.

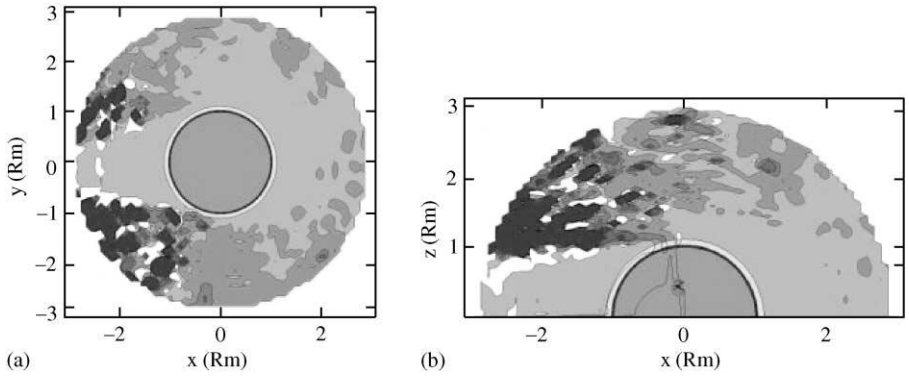

Fig. 11. Earlier epoch (3D MC simulation): Average energy of sputter produced coronal particles. (a) In the equatorial plane. (b) In a plane perpendicular to the equatorial plane along the Mars/Sun axis. Same format as in Fig. 5.

coronal density is roughly doubled by the sputter produced population compared to the estimations made by Zhang et al. (1993a). If we suppose that the density of ionized particles is roughly also doubled, the ionopause altitude is then roughly increased by $H \ln 2 \simeq 28$ km where H is the scale height (~ 47 km on the dayside). In the end, the density of ionized coronal particles above the ionopause will be roughly the same but the ionopause altitude higher, so that the total number of ionized particles will be larger. Moreover, since the sputter produced particles are on the average more energetic than those from dissociative-recombination, the sputtering becomes more important with increasing ionopause altitude.

4. Concluding remarks

We present in this paper the first 3D picture of the effect of the solar wind pick-up ions on the atmosphere of Mars. Two periods of Mars history are considered: one, present solar minimum and the other 3 times that EUV flux suggested to be ~ 2.5 Gyr ago. The process described in this paper is the sputtering of the atmosphere by ions picked up by the interplanetary magnetic field frozen inside the solar wind (Luhmann and Kozyra, 1991). It has been shown in previous studies (Jakosky et al., 1994; Kass and Yung, 1995; Johnson et al., 2000) that such process could significantly contribute to the Mars atmospheric loss. Johnson and Luhmann (1998) have shown that a potential feedback process associated with sputtering has also to be taken into account to accurately estimate the total loss of atmosphere. The sputtering requires development of a 3D simulation describing the trajectories of the incident ions around Mars (Brecht, 1997a, b) and the interaction of these particles penetrating the Martian atmosphere.

To compare with previous 1D studies, we used the model of incident flux of pick-up ions by Luhmann et al. (1992). For present solar minimum conditions, the heating due to the sputtering is small so we only consider the populations sputtered by the incident particles reaching altitudes higher

than the ionopause. Above this altitude these neutral particles can be ionized (Zhang et al., 1993b) and then can be accelerated by the solar wind fields and removed from Mars or they can be accelerated and reimpact the atmosphere. This bombardment can cause additional atoms to escape the gravity field of Mars. For the earlier epoch, the heating due to the incident particles is not negligible compared to the EUV and UV heating and is simulated with a full 2D Direct Simulation Monte Carlo model. For both epochs, the sputtered population reaching the corona above the ionopause is described by using a 3D Monte Carlo model.

We showed that in the 3EUV epoch the exobase altitude, here estimated as the altitude where the mean free path and the density scale height are equal, is around 300 km, significantly higher than the $\sim$ 216 km estimate of Zhang et al. (1993a) who only considered the UV and EUV heating. This increase is due to the heating by the incident pick-up ion flux and is a lower bound since the H^+ heating has been neglected. We also calculated the number of collisions per recoil particles in the corona showing that the collisions should not be ignored above the nominal exobase. As an example, above 180 km which is the exobase altitude calculated by Zhang et al. (1993a) for present solar minimum conditions, a sputtered particle loses by collisional energy transfer with the background around 20% of its energy before reaching the ionopause. For the earlier epoch, on the dayside, collisions can be neglected only above 350 km.

For both epochs we provide an estimate of the yield: the rate between the number of escaping neutral O particles to the number of incident ones. We found for present solar minimum a total yield of 2.9 which is less than the 3.9 estimate Johnson et al. (2000) and for the 3EUV epoch a total yield of 3.6, less than the 4.2 yield estimate of Johnson et al. (2000). This difference can be explained by our more complete description of the geometry of the corona. Indeed using the 3D model we are able to accurately calculate the effect of the grazing ions which are very efficient at producing escape. This contribution was overestimated by Johnson et al. (2000). The loss of particles s^{-1} during these two epochs are 6.5×10^{23} O atoms/s in the 1EUV and 1.3×10^{26} O atoms/s during the 3EUV epoch. These values are slightly larger than the estimate made by Jakosky et al. (1994) who found 3×10^{23} O atoms/s in the 1EUV and 10^{26} O atoms/s during the 3EUV epoch. To this loss should be added the much larger loss due to direct O^+ pick-up ion loss and to exospheric O atom loss (Luhmann et al., 1992).

The coronal density calculated in this work confirms the importance of sputtering (Johnson and Luhmann, 1998). At present solar minimum, the enhancement of the coronal density due to the sputtered population is negligible compared to dissociative-recombination of O_2^+ (Kim et al., 1998; Hodges, 2000) except in the polar and dusk equatorial regions. In the earlier epoch, this density is of the same order in the corona or greater in the polar and dusks equatorial regions than the density predicted by Zhang et al. (1993a). The number of lost O^+ pick-up ions should therefore be underestimated in Zhang et al. (1993a), Luhmann et al. (1992) and Luhmann (1997). Moreover, the energy range of the sputtered particles compared to the energy range of the O atoms produced by dissociative-recombination of O_2^+ (Kim et al., 1998; Hodges, 2000) is larger. Dissociative-recombination of the O_2^+ ions provides coronal O atoms of energy lower than 5 eV whereas the sputtering of the atmosphere generates coronal particles of average energy around 15 eV at present solar minimum and of 9 eV in the 3EUV epoch. This difference is also important for the feedback process (Johnson and Luhmann, 1998) since it implies that these sputter produced particles populate larger distance from Mars than the dissociative–recombination process where they are more readily ionized.

Acknowledgements

This work was supported by the NASA's Planetary Atmosphere Program.

References

Acuña, M.H., et al., 1998. Magnetic field and plasma observations at Mars: initial results of the Mars Global Surveyor mission. Science 279, 1676–1680.

Acuña, M.H., et al., 1999. Global distribution of crustal magnetization discovered by the Mars Global Surveyor MAG/ER experiment. Science 284, 790–793.

Bird, G.A., 1994. Molecular Gas Dynamics and the Direct Simulation of Gas Flows. Clarendon Press, Oxford, England.

Bougher, S.W., Engel, S., Roble, R.G., Foster, B., 1999. Comparative terrestrial planet thermosphere 2. Solar cycle variation of global structure and winds at equinox. J. Geophys. Res. 104, 16,591–16,611.

Bougher, S.W., Roble, R.G., 1991. Comparative terrestrial planet thermosphere 1. Solar cycle variation of global mean temperature. J. Geophys. Res. 96, 11,045–11,055.

Brain, D.A., Jakosky, B.M., 1994. Atmospheric loss since the onset of the Martian geologic record: combined role of impact erosion and sputtering. J. Geophys. Res. 103, 22,689–22,694.

Brecht, S.H., 1997a. Hybrid simulations of the magnetic topology of Mars. J. Geophys. Res. 102, 4743–4750.

Brecht, S.H., 1997b. Solar wind proton deposition into the Martian atmosphere. J. Geophys. Res. 102, 11,287–11,294.

Carr, M.H., 1999. Retention of an atmosphere on early Mars. J. Geophys. Res. 104, 21,897–21,909.

Firsov, O.B., 1959. A qualitative interpretation of the mean electron excitation energy in atomic collisions. Zh. Eksp. Teor. Fiz. 36, 1517–1523 (Sov. Phys. JETP (Engl. Transl.) 36, 1076–1086, 1959) (in Russian).

Hodges, R.R., 2000. Distributions of hot oxygen for Venus and Mars. J. Geophys. Res. 105, 6971–6981.

Jakosky, B.M., Jones, J.H., 1997. The history of the Martian volatiles. Rev. Geophys. 35, 1–16.

Jakosky, B.M., Pepin, R.O., Johnson, R.E., Fox, J.L., 1994. Mars atmospheric loss and isotopic fractionation by solar-wind induced sputtering and photochemical escape. Icarus 111, 271–288.

Johnson, R.E., 1994. Plasma-ion sputtering of an atmosphere. Space Sci. Rev. 69, 215–253.

Johnson, R.E., Liu, M., 1996. The loss of atmosphere from Mars. Science 274, 1932.

Johnson, R.E., Luhmann, J.G., 1998. Sputter contribution to the atmospheric corona on Mars. J. Geophys. Res. 103, 3649–3653.

Johnson, R.E., Schnellenberger, D., Wong, M.C., 2000. The sputtering of an oxygen thermosphere by energetic O^+. J. Geophys. Res. 105, 1659–1670.

Kallio, E., Luhmann, J.G., 1997. Charge exchange near Mars: the solar wind absorption and energetic neutral atom production. J. Geophys. Res. 102, 22,183–22,197.

Kass, D.M., Yung, Y.L., 1995. Loss of atmosphere from Mars due to solar-wind induced sputtering. Science 268, 697–699.

Kass, D.M., Yung, Y.L., 1996. The loss of atmosphere from Mars: response. Science 274, 1932–1933.

Kim, J., Nagy, A.F., Fox, J.L., Cravens, T.E., 1998. Solar cycle variability of hot oxygen atoms at Mars. J. Geophys. Res. 103, 29,339–29,342.

Luhmann, J.G., 1997. Correction to "The ancient oxygen exosphere of Mars: implications for atmosphere evolution" by Zhang et al. J. Geophys. Res. 102, 1637.

Luhmann, J.G., Johnson, R.E., Zhang, M.H.G., 1992. Evolutionary impact of sputtering of the Martian atmosphere by O^+ pick-up ions. Geophys. Res. Lett. 19, 2151–2154.

Luhmann, J.G., Kozyra, J.U., 1991. Dayside pick-up oxygen in precipitation at Venus and Mars: spatial distribution, energy deposition and consequences. J. Geophys. Res. 96, 5457–5467.

McElroy, M.B., 1972. Mars: an evolving atmosphere. Science 175, 443.

Nagy, A.F., Cravens, T.E., 1988. Hot oxygen atoms in the upper atmospheres of Venus and Mars. Geophys. Res. Lett. 15, 433.

Pospiezalska, M.K., Johnson, R.E., 1992. Plasma heating of Io's atmosphere. Geophys. Res. Lett. 19, 949–952.

Pospiezalska, M.K., Johnson, R.E., 1996. Monte Carlo calculations of plasma ion-induced sputtering of an atmosphere: SO_2 ejected from Io. J. Geophys. Res. 101, 7565–7573.

Pepin, R., 1994. Evolution of the Martian atmosphere. Icarus 111, 289–304.

Smyth, W.H., Combi, M.R., 1988. A general model for Io's neutral gas clouds. II Application to the sodium cloud. Astrophys. J. 328, 888–918.

Wong, M.C., Johnson, R.E., 1995. A three-dimensional azimuthally symmetric model of the atmosphere of Io, 1., Photochemistry and the accumulation of a nightside atmosphere. J. Geophys. Res. 101, 23,243–23,254.

Zhang, M.H.G., Luhmann, J.G., Bougher, S.W., Nagy, A.F., 1993a. The ancient oxygen exosphere of Mars: implication for atmosphere evolution. J. Geophys. Res. 98, 10,915–10,923.

Zhang, M.H.G., Luhmann, J.G., Nagy, A.F., Spreiter, J.R., Stahara, S., 1993b. Oxygen ionization rates at Mars and Venus: relative contributions of impact ionization and charge exchange. J. Geophys. Res. 98, 3311–3318.

Ziegler, J.F., Biersack, J.P., Littmark, V., 1985. The Stopping and Ranges of Ions in Solids. Pergamon Press, New York.

Zuber, M.T., et al., 2000. Internal structure and early thermal evolution of Mars from Mars Global Surveyor, Topography and Gravity. Science 287, 1788.

ANNEXE 3: ICARUS, 164, 261-281, 2003.

LEBLANC F., AND JOHNSON R.E., MERCURY'S SODIUM EXOSPHERE, ICARUS, 164, 261-281, 2003.

ACADEMIC PRESS

Available online at www.sciencedirect.com
SCIENCE DIRECT®

Icarus 164 (2003) 261–281

ICARUS

www.elsevier.com/locate/icarus

Mercury's sodium exosphere

F. Leblanc [a,*] and R.E. Johnson [b]

[a] *Service d'Aéronomie du CNRS, Verrières-Le-Buisson, France*
[b] *University of Virginia, Charlottesville, VA 22903, USA*

Received 18 December 2002; revised 18 April 2003

Abstract

Mercury's neutral sodium exosphere is simulated using a comprehensive 3D Monte Carlo model following sodium atoms ejected from Mercury's surface by thermal desorption, photon stimulated desorption, micro-meteoroid vaporization and solar wind sputtering. The evolution of the sodium surface density with respect to Mercury's rotation and its motion around the Sun is taken into account by considering enrichment processes due to surface trapping of neutrals and ions and depletion of the sodium available for ejection from the surfaces of grains. The change in the sodium exosphere is calculated during one Mercury year taking into account the variations in the solar radiation pressure, the photo-ionization frequency, the solar wind density, the photon and meteoroid flux intensities, and the surface temperature. Line-of-sight column densities at different phase angles, the supply rate of new sodium, average neutral and ion losses over a Mercury year, surface density distribution and the importance of the different processes of ejection are discussed in this paper. The sodium surface density distribution is found to become significantly nonuniform from day to night sides, from low to high latitudes and from morning to afternoon because of rapid depletion of sodium atoms in the surfaces of grains mainly driven by thermal depletion. The shape of the exosphere, as it would be seen from the Earth, changes drastically with respect to Mercury's heliocentric position. High latitude column density maxima are related to maxima in the sodium surface concentration at high latitudes in Mercury's surface and are not necessarily due to solar wind sputtering. The ratio between the sodium column density on the morning side of Mercury's exosphere and the sodium column density on the afternoon side is consistent with the conclusions of Sprague et al. (1997, Icarus 129, 506–527). The model, which has no fitting parameters, shows surprisingly good agreement with recent observations of Potter et al. (2002, Meteor. Planet. Sci. 8, 3357–3374) successfully explaining their velocity and column density profiles vs. heliocentric distance. Comparison with this data allows us to constrain the supply rate of new sodium atoms to the surface. We also discuss the possible origins of the strong high latitude emissions (Potter and Morgan, 1990, Science 248, 835–838; 1997a, Adv. Space Res. 19, 1571–1576; 1997b, Planet. Space Sci. 45, 95–100; Sprague et al., 1998, Icarus 135, 60–68) and the strong variations of the total content of the sodium exosphere on short (Potter et al., 1999, Planet. Space Sci. 47, 1441–1449) and long time scales (Sprague et al., 1997, Icarus 129, 506–527).

Keywords: Mercury; Atmospheres; Dynamics; Surfaces; Planets

1. Introduction

Since its discovery by Potter and Morgan (1985), a large set of observations of Mercury's sodium exosphere has been accumulated but a comprehensive description of this atmosphere is not available. The suggested mechanisms for producing Mercury's neutral sodium exosphere are photon stimulated desorption (McGrath et al., 1986; Yakshinskiy and Madey, 1999), thermal desorption (Hunten et al., 1988; Madey et al., 1998; Yakshinskiy et al., 2000), sputtering by impacting solar particles (Johnson and Baragiola, 1991; Potter and Morgan, 1997a; Killen et al., 2001), micro-meteoroid vaporization (Morgan et al., 1988; Hunten et al., 1988; Cintala, 1992) and chemical sputtering (Potter, 1995). The sputtering of the surface by reimpacting magnetospheric pick-up ions (Cheng et al., 1987; Ip, 1993) is probably less important in term of total production of exosphere sodium but acts mainly on nightside unlike the other processes (Delcourt et al., 2003). The modeling of Mercury's exosphere has mostly been carried out by treating these desorption processes separately and then comparing the results to observations for a narrow set of orbital positions. Consequently there has been considerable disagreement on the morphology and content of the exosphere and on the dominant desorption processes. Here we show that these processes must

* Corresponding author.
E-mail address: francois.leblanc@aerov.jussieu.fr (F. Leblanc).

0019-1035/$ – see front matter
doi:10.1016/S0019-1035(03)00147-7

be treated concurrently since they not only populate the exosphere but deplete the sodium in the surface layers of grains as Mercury very slowly rotates while it orbits the sun. Therefore, we model here the exosphere/surface layer population including all of the desorption processes above except that due to magnetospheric ions and chemical sputtering.

Ion sputtering is mainly due to solar particles that penetrate Mercury's magnetosphere along open magnetic field lines (Kabin et al., 2000; Killen et al., 2001; Ip and Kopp, 2002). These ions impact most often at high latitudes with keV/amu energy and eject mainly neutral particles by momentum transfer collisions and electronic excitations in the upper monolayers of surface grains (Johnson, 1990). Photon stimulated desorption is due to solar photons impacting the surface on the dayside, ionizing the lattice locally and repulsively ejecting neutral sodium from the first few monolayers of grains on Mercury's surface (Madey et al., 1998; Yakshinskiy and Madey, 1999). Thermal desorption (also called evaporation) is due to the thermal agitation of adsorbed atoms or molecules and leads mainly to neutral ejection (Yakshinskiy et al., 2000). Meteoroid bombardment, which creates the regolith and mixes the surface, also produces exospheric atoms and molecules by vaporization and melting of the impactor and the impacted surface. These processes are discussed in more detail in a number of recent reviews (Killen and Ip, 1999; Stern, 1999; Johnson, 2002b).

McGrath et al. (1986) suggested that photon stimulated desorption and thermal desorption were the dominant surface ejection/desorption processes. However Potter and Morgan (1990) and Potter et al. (1999) observed strong emission brightness enhancements at high latitudes which have been suggested to be due to solar wind sputtering. These authors concluded therefore that solar wind sputtering could produce from 0 to 32% of Mercury sodium exosphere (Killen et al., 2001). However significant differences between morning and afternoon sodium emission brightness for a period of several Mercury's years have been interpreted as an exosphere mainly produced by thermal desorption (Sprague, 1992a, 1992b; Sprague et al., 1997). The Na implanted into the surface regolith at night would be released through diffusion in Mercury's grain in the morning or afternoon at high latitudes (Sprague, 1990a). Killen and Morgan (1993b) criticized later this point arguing that diffusion in Mercury grain was too slow to explain the observed exosphere. But, recent laboratory measurements show thermal desorption is a very efficient ejection process as soon as the surface temperature reaches more than 400 K (Yakshinskiy et al., 2000). As a consequence, on the hot surface of Mercury (subsolar point temperature between 575 and 700 K), thermal desorption should lead to a rapid depletion of sodium in the first few monolayers of grains on the dayside surface (on time scale much shorter than one Mercury hour). This depletion, therefore, places strong limits on the other mechanisms of ejection (Hunten and Sprague, 2002) in particular on the photon stimulated desorption and solar wind sputtering processes. The sinks of the sodium exosphere are neutral loss by escape, which is enhanced by the solar radiation pressure (Smyth, 1986), reabsorption by the surface (Shemansky and Broadfoot, 1977; Hunten and Sprague, 1997) and solar photo-ionization of exospheric neutral particles, accelerated by the convection electric field associated with the solar wind (Ip, 1987).

In this paper, a 3D Monte Carlo model is used to simulate the contribution to Mercury's sodium exosphere due to the four main surface desorption mechanisms. We first describe in detail our approach for calculating the sodium exosphere and the sodium in upper few monolayers of the surface grains (Section 2). A self consistent approach is developed for the different processes of ejection taking into account the rapid depletion of available sodium in Mercury's surface with increasing zenith angle. We then provide insights on the sodium density distribution in Mercury's surface (Section 3.1), on the morphology of Mercury's exosphere with respect to heliocentric position and phase angle (Section 3.2) and on the morning/afternoon column density asymmetries (Section 3.3). Finally, we compare our model to recent measurements by Potter et al. (2002) which provided detailed information on the shape of the sodium cloud (Section 4). In Section 5, we discuss supply and loss rates (5.1) and give some insights into the possible origins of the high latitude emissions and of the temporal variation of the total content of sodium atoms of Mercury's exosphere (5.2). In Section 6 we summarize our conclusions.

2. Mercury sodium exosphere

The 3D Monte Carlo model of Mercury sodium exosphere used here has been adapted from a model successfully applied to the description of the Europa sodium exosphere (Leblanc et al., 2002; Johnson et al., 2002a). In the present version of this model, several million test-particles are followed around Mercury. A test-particle is a simulated particle which represents a large number of real sodium atoms. Typically this number, called the weight of the test-particle, is between 10^{20} to 10^{26}. Because the atmosphere is nearly collisionless, sodium atoms produced by thermal desorption, photon stimulated desorption, solar wind ion sputtering and micro-meteoroid vaporization can each be tracked separately.

Thermal desorption, photon stimulated desorption and solar wind ion sputtering are here assumed to act roughly on the same population of sodium bound in Mercury's surface as discussed below. Madey et al. (1998) have suggested that absorbed exospheric sodium atoms which stick to the surface can become bound at regular surface sites. Therefore, they do not have substantially lower surface binding energies than intrinsic surface sodium atoms and their subsequent desorption behavior should not be substantially different. Here we assume that any sodium atom in the top few monolayers of a grain on Mercury's surface can be ejected either by thermal

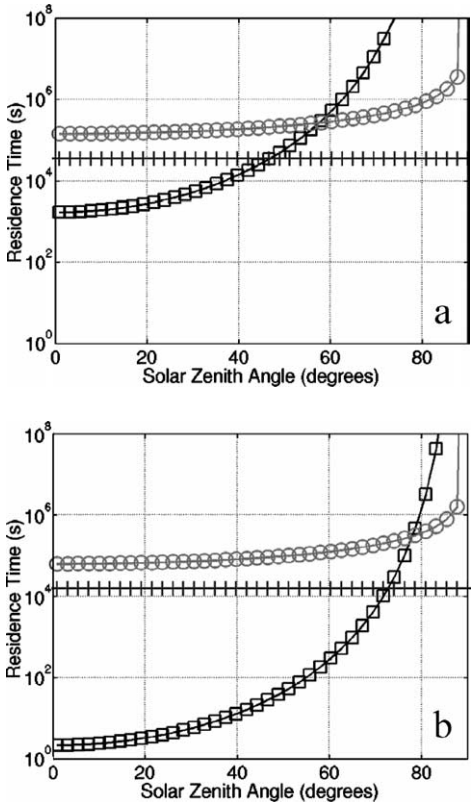

Fig. 1. Time of residence of a sodium atom in Mercury's upper surface versus solar zenith angle. Circles: submitted to photon stimulated desorption. Squares: submitted to thermal desorption. Crosses: submitted to solar wind sputtering. (a) At Mercury's aphelion. (b) At Mercury's perihelion.

desorption, by photon stimulated desorption or by solar wind ion sputtering. There are in fact a variety of binding sites for sodium in a silicate surface including very deep strongly binding sites, particularly on a surface that is damaged by radiation (Yakshinskiy et al., 2000). Sodium bound below the surface layers need to reach the surface layers by thermal or radiation-enhanced diffusion (McGrath et al., 1986; Potter et al., 2000).

We plotted the residence time for sodium atoms in Mercury's surface at aphelion in Fig. 1a and at perihelion in Fig. 1b. The time of residence is displayed with respect to solar zenith angle (0° subsolar point and 90° terminator/pole) for solar wind sputtering (crosses), for thermal desorption (squares) and for photon stimulated desorption (circles). We use the definition of the ejection rates for each of these processes described in Sections 2.1–2.3. For thermal desorption we use a distribution centered around an average binding energy equal to 1.85 eV as discussed below. This corresponds to the most probable value (see Section 2.1). Solar wind sputtering is much more sensitive to the structure of the open magnetic field lines at Mercury than on the solar zenith angle as shown in Fig. 1. These field lines roughly correspond to zenith angles between 20° to 70° in our model. Fig. 1 shows that below 50° solar zenith angle at the aphelion and below 70° solar zenith angle at the perihelion the time of residence of a sodium atom in Mercury's surface is up to several orders smaller for thermal desorption than for the other processes of ejection. On Mercury's dayside thermal desorption strongly limits the ability of photon stimulated desorption and solar wind sputtering to eject exosphere sodium atoms (Hunten and Sprague, 2002). Thermal desorption, photon stimulated desorption and solar wind sputtering are not fully competitive since solar wind sputtering can act on the nightside (Kabin et al., 2000; Ip and Kopp, 2002) and sputter-enhanced diffusion by energetic ions and electrons can supply the surface with sodium atoms from depth (McGrath et al., 1986; Potter et al., 2000). This supply process could enhance the role of energetic solar particles. It will not be explicitly treated in this work but included in estimates of the net supply rate.

The source due to micro-meteoroid impact differs from those discussed above because meteoroids contain sodium atoms which can be directly ejected into the exosphere on impact vaporization. Impact vaporization of surface material can also contribute as an endogenic source of sodium atoms even from the tightest binding sites and the meteoroids can, in principal, add sodium to the surface. For simplicity, we assume in this paper that micro-meteoroid vaporization is fully independent of the other ejection processes contributing only exospheric sodium atoms coming from the fully vaporized meteoroid. Meteoroid impact also replenishes the surface by mixing as discussed below. In subsequent work, a more detailed model of the effects of meteoroids will be considered.

The supply of sodium atoms to Mercury's exosphere/surface layer eventually limits Mercury's sodium loss and is important in determining the total content of sodium atoms in Mercury's exosphere/surface layer. Micro-meteoroid vaporization, discussed above, is one external source. As described in Section 2.4, this flux corresponds to an average supply of the order of 3.5×10^{23} Na/s. To this supply can be added other sources: meteoroid impact mixing of grains at the surface replacing depleted grains with fresh grains known as gardening (McGrath et al., 1986; Cintala, 1992), creation of new regolith by large meteoroid impacts (Killen and Morgan, 1993a) and thermal (Killen and Morgan, 1993a) or sputter-enhanced diffusion in grains (McGrath et al., 1986; Johnson and Baragiola, 1991; Potter et al., 2000). Since, the effect of concentration and temperature gradients is not well known, estimating the rate with which new sodium atoms are brought to a grain surface where they are able to be ejected into the exosphere is problematic at present. Moreover the surfaces of grains can be oppositely charged by photon and electron fluxes (Jurac et al., 1995) affecting the amount of sodium available for desorption. Photon bombardment can result in positive charging and, therefore, inward sodium diffusion, whereas low energy electrons can cause negative charging and, therefore, outward diffusion of sodium (Madey et al., 2002). Although, this can also affect the availability of sodium near the subsolar point and the poles the magnitude of the effect is at present poorly constrained by laboratory measurements. In some

previous work on Mercury's sodium exosphere (Ip, 1990; Killen et al., 2001) the total supply rate was assumed to be fast enough to maintain a constant density at the surface. In order to bracket the possible supply rates we will calculate the Mercury's exosphere for three cases: no density depletion in the surface, a supply of new sodium atoms two times the supply due to micro-meteoroid vaporization and a supply rate equal to the meteoroid delivery rate only.

The time step of the simulation is the time which separates the calculation of two successive positions of a test-particle (typically less than 0.1 s real time). It is calculated as the time needed for a particle with the escape velocity (~ 4.2 km/s at Mercury surface) to move a distance less than 300 m. Comparing to known trajectories, we show that using such a time step the trajectory of a test-particle is accurately calculated with the central force algorithm used in this model (Leblanc et al., 2002). This algorithm has also been checked by comparison with a fourth order Runge Kutta algorithm.

The initial sodium concentration c_{Na} in Mercury's surface is set to 0.0053 with a surface density of 7.5×10^{14} atoms/cm^2 (Killen et al., 2001). These parameters define the initial number of sodium atoms (3×10^{30} Na) put inside the exosphere/surface layer at the beginning of the simulation. A grid of 500×60 cells with equal surface area is defined to describe Mercury's surface density distribution. At each time step and for each cell a spatial position in this cell is chosen randomly. The probabilities of a test-particle to be ejected at this position by thermal desorption, photon stimulated desorption and solar wind sputtering are then calculated. The way each of these probabilities is calculated is described in Sections 2.1–2.3. In random order, we then test whether or not each process of ejection could lead to the ejection of a particular test-particle. This is realized by comparing a number randomly calculated between 0 to 1 to the probability of ejection associated with each process. If ejection occurs, a test-particle is created at the selected position with a velocity vector for which the size and direction are calculated from the distribution associated with the selected mechanism of ejection (see Sections 2.1–2.3). The surface density in the cell is then decreased by the weight of the ejected test-particle. The value of this weight depends of the number of particles in the cell and on a chosen maximum value, chosen to limit the computing time. The weight of a test-particle does not depend on the process of ejection and changes from one test-particle to another. We repeat this set of tests by selecting random positions in the cell until all of the particles initially in a cell have been tested during this time step. Each cell of the surface grid is tested in the same way. The sodium produced by meteoroid vaporization is treated independently as described in Section 2.4.

The motion of a test-particle in the environment of Mercury is driven by the gravity of Mercury and the Sun and by the solar radiation pressure (Smyth, 1986) which depends on the instantaneous heliocentric radial velocity of each atom (Leblanc et al., 2002). Since sodium atoms are ejected from Mercury's surface mainly as neutrals (Madey et al., 1998), magnetic and electric effects are not considered. Collisions with other particle are also neglected. During its motion, a test-particle can either be ionized by solar photons or can reach 50 R_M (Mercury radii) from Mercury or can reimpact the surface. For a test-particle ionized below 10 km in altitude the effect of the solar wind electric field of convection is such that it most likely reimpacts the surface. Therefore, we assumed it is reabsorbed by the surface at the same longitude and latitude (the total number of particles in the corresponding cell is then increased by the weight of the test-particle). Test-particles ionized above 10 km are assumed to escape Mercury and, therefore, are suppressed. Test-particles reaching 50 R_M are also assumed to escape and are suppressed. An accurate treatment of the ion recycling and also of the magnetospheric ion sputtering effect will be the subject of a forthcoming paper.

A test-particle which reimpacts the surface can either stick at the surface or be reemitted. The capability of an exospheric sodium atom to stick or to hop at Mercury's surface depends mainly on the surface temperature and porosity (Johnson, 2002b). The sticking properties of a sodium atom impacting a SiO_2 surface have been measured by Yakshinskiy et al. (2000). From their measurements we can extrapolate a function $\mathrm{Stick} = A \times \exp(B/T_s)$ where Stick represents the probability of a sodium atom to stick to a surface at a temperature T_s. A and B are calculated in order to fit the measurements: $\mathrm{Stick} = 0.5$ at $T_s = 250$ K and $\mathrm{Stick} = 0.2$ at $T_s = 500$ K. The porosity of the regolith, Por, increases the sticking efficiency (Johnson, 2002b). The effective sticking probability of a sodium atom impacting Mercury's surface is $\mathrm{Stickeff} = \mathrm{Stick}/[1 - (1 - \mathrm{Stick}) \times \mathrm{Por}]$. If $\mathrm{Por} = 0$, the measurements of Yakshinskiy et al. (2000) for a flat surface are obtained, whereas if $\mathrm{Por} = 1$, all the impacting particles stick in the regolith. In this work we set $\mathrm{Por} = 0.8$ (Johnson, 2002b). When a test-particle sticks its weight is added to the number of particles in the cell which it impacts. When it does not stick we reemit it with an energy determined by the local surface temperature. That is, in the regolith a particle interacts several times with the surface before being re-emitted (Killen et al., 1990) and therefore it is quickly thermalized. In earlier works (Ip, 1990; Smyth and Marconi, 1995a; Killen et al., 2001), the sticking probability and a thermal accommodation coefficient were estimated from experiments for flat surfaces.

We simulated Mercury sodium exosphere/surface layer for more than six Mercury's years, i.e., more than three diurnal cycles (one Mercury's year = 87.97 Earth days and one Mercury's day or diurnal cycle = 58.65 Earth days). This length of time is chosen in order that the results of the simulation are independent of the initial surface density distribution. In this first simulation we use large weights (giving low statistics) in order to obtain a rough description of the surface density distribution at any heliocentric position during one whole Mercury's year. The use of large weights reduces the computing time but generates picture of

the sodium exosphere not well suited for comparison with detailed observations (in particular with Potter et al., 2002). Therefore, in a second set of simulations we use the surface density and exosphere distributions obtained by this first simulation to initialize the simulation but use much smaller weights. In the second set of simulations, we simulate at least 500 h of Mercury motion about the Sun in advance of a chosen heliocentric position of Mercury. In such a case, good statistics are obtained for the exospheric density around Mercury up to 50 R_M from Mercury's surface. Such times are much larger than the typical time for ionization of a sodium atom at Mercury's orbit (between 1.4 and 3.3 h). This is also much larger than the time it takes a 1 km/s ejecta from Mercury's surface to reach the limit of the domain of the simulation (50 R_M). A velocity of 1 km/s is smaller than the minimum velocity needed by a sodium particle to reach 50 R_M from Mercury (Ip, 1993; Smyth and Marconi, 1995a).

The sodium exosphere density is calculated by integrating the density on a 3D grid for a time interval around a chosen heliocentric position of Mercury in order to reduce the noise/signal ratio (Leblanc et al., 2002). This integration is based on the assumption that during a sufficiently small time (10 Earth hours, $\sim 1/210$ of one Mercury year), the sodium exosphere density at any position around Mercury does not vary significantly. This assumption is realist considering the time scales for the exospheric sources. This length of time is also much larger than the period associated with solar wind sputtering (half an hour, see Section 2.3).

The rotation of Mercury on itself is taken into account at each time step. The position and the heliocentric velocity vector of Mercury are also calculated by solving the Kepler laws at each time step. In this paper, we will use the true anomaly angle (TAA) to indicate the position of Mercury with respect to the Sun (see Fig. 2). The TAA is the angle between the Sun–Mercury direction when Mercury is at its perihelion and the Sun–Mercury direction. That is, TAA is equal to 0° when Mercury is at its perihelion and is equal to 180° when Mercury is at its aphelion.

The solar photon flux reaching Mercury varies in $1/R_{\mathrm{Hel}}^2$ where R_{Hel} is the heliocentric distance of Mercury. The lifetime of a neutral sodium atom before photo-ionization is proportional to the solar photon flux and is set to 3.3 h at Mercury aphelion (Smyth and Marconi, 1995a). A 3 times larger value for the photo-ionization lifetime has been suggested by Cremonese et al. (1997) but will not be used here. Sodium atoms inside Mercury's shadow are not ionized by solar photons. In this region, charge exchange or electronic impact can in principal ionize sodium atoms. Because of very small electron and ion densities (Lundin et al., 1997) in Mercury's magnetosphere, these ionizations rates are much smaller than photo-ionization and are not considered. The solar photon radiation pressure is adapted from the formula used in Leblanc et al. (2002). Its intensity varies with respect to sodium atom heliocentric distance and radial velocity and is null in Mercury's shadow.

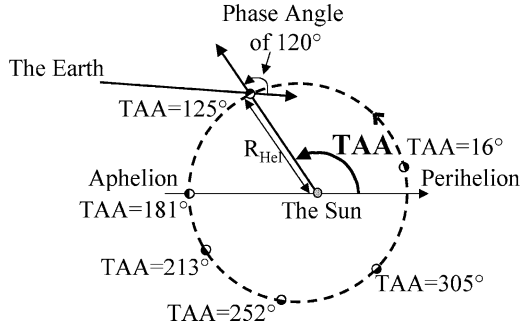

Fig. 2. Mercury position around the Sun at different True Anomaly Angle (TAA), adapted from Smyth and Marconi (1995a). The axis of rotation of Mercury around the Sun and of Mercury around itself are identical and perpendicular to the figure towards the reader.

The temperature of the surface T_s depends on the solar zenith angle and on the solar photon flux intensity:

- On the dayside: $T_s(\text{longitude}, \text{latitude}) = T_0 + T_1 \times (\cos(\text{latitude}) \times \cos(\text{longitude}))^{1/4}$ with $T_1 = 100$ K and T_0 varying from 600 K at the perihelion to 475 K at the aphelion. This is deduced from Butler (1997). It has also been parametrized in order to reproduce the subsolar point temperature variation with respect to Mercury's heliocentric distance from ~ 575 K at the aphelion to ~ 700 K at the perihelion (Hale and Hapke, 2002).
- On the nightside: $T_s = 100$ K.

The latitude varies from $-90°$ to 90° with negative values corresponding to the Southern hemisphere and positive values to the Northern hemisphere and the longitude varies from $-180°$ to 180° where $-90°$ to 90° corresponds to the dayside. The subsolar point is placed at latitude and longitude equal to zero.

The calculated densities around Mercury are integrated to obtain line-of-sight column densities at different phase angles (see Fig. 2 for the definition of the phase angle). Atoms inside Mercury's shadow or hidden by Mercury from the observer are excluded as they cannot be observed in emission. The net emission in rayleighs is not provided since the optical depth for sodium around Mercury is needed, which is beyond the scope of this work.

2.1. Thermal desorption

Thermal desorption has been neglected in most of the studies of the sodium exosphere (Ip, 1993; Killen and Morgan, 1993a; Smyth and Marconi, 1995a; Morgan and Killen, 1997; Killen et al., 2001) although its importance has been underlined by Sprague (1990a, 1992a, 1992b) and Hunten and Sprague (1997, 2002). Killen et al. (2001) assumed it can act only on physi-sorbed. Moreover it only produces 0.03 to 0.05 eV sodium atoms (Yakshinskiy et al., 2000) whereas the escape energy at Mercury's surface is equal to 2.07 eV. Such sodium atoms remain close to the surface (do not have enough energy to reach more than 60 km

from the surface) and, therefore, do not contribute significantly to Mercury sodium loss or to the emission brightness above 60 km. However it has been suggested by Hunten and Sprague (1997, 2002) that thermal desorption can strongly limit the amount of sodium ejected by photo stimulated desorption or solar wind sputtering (Fig. 1). As soon as the surface temperature reaches more than 400 K in the early morning of Mercury, thermal desorption can quickly deplete the sodium in the surface layers of grains that is available for desorption. This would be most efficient on Mercury's dayside surface in the absences of sources.

Yakshinskiy et al. (2000) have published results of laboratory study of thermal desorption of sodium atoms from thin SiO_2 films. The rate of ejection of sodium atoms into the gas phase (number of sodium atoms per second) can be expressed with the following formula: $\tau_{\text{thd}} = \nu c_{\text{Na}} \exp(U/k_{\text{B}}T_{\text{s}})$ where ν is the vibrational frequency in the surface ($\sim 10^{13}$ s^{-1}), U is the binding energy, k_{B} the Boltzmann constant, c_{Na} the sodium concentration in the surface and T_{s} the surface temperature. Following Yakshinskiy et al. (2000) multiple binding sites with energies from 1.4 to 2.7 eV with a most probable value equal to $U = 1.85$ eV provide the best fit to their measurements, as discussed earlier. Hunten and Sprague (2002) suggest that a value closer to 1.4 eV is more appropriate for the surfaces of the Moon and Mercury. Here we consider a gaussian distribution for U defined between 1.4 and 2.7 eV with a most probable value equal to 1.85 eV (Yakshinskiy et al., 2000). For any of these values, τ_{thd} is highly dependent on the surface temperature such that below 350 K the ejection rate is close to zero. The probability of ejection of a test-particle from Mercury's surface during a time step is the product of the time step of the simulation by τ_{thd} and divided by c_{Na}. This probability does not depend on the local concentration c_{Na}. It is taken into account via the number of test-particles inside each cell, which is equal to the local surface density multiplied by the surface area of the cell and divided by the weight. The energy distribution of sodium atoms ejected by thermal desorption is that for particles fully accommodated to the surface. It is a Maxwellian Boltzmann flux distribution (Brinkman, 1970; Smith et al., 1978): $f(E,\theta) = 2\cos(\theta) \times E/(k_{\text{B}}T_{\text{s}})^2 \exp(-E/k_{\text{B}}T_{\text{s}})$ where E is the energy of the ejecta and θ is the angle between the velocity vector of the ejecta and the normal to the surface.

2.2. Photon stimulated desorption

The intensity of the flux of ejected particle by photon stimulated desorption varies with respect to the flux of photon at Mercury's orbit and therefore with respect to the square of Mercury heliocentric distance. At Mercury's surface it is also proportional to the cosine of the zenith angle on the dayside whereas this flux is null on the nightside. By fitting to observations of Mercury sodium exosphere, Killen et al. (2001) estimated a cross section for photon stimulated desorption $Q = 1.4 \times 10^{-21}$ cm^2 using an average photon flux equal to $1.5 \times 10^{14}/R_{\text{Hel}}^2$. The solar averaged cross section for photon stimulated desorption measured in laboratory by Yakshinskiy and Madey (1999) is equal to $Q = 3.0 \times 10^{-20}$ cm^2. This has been obtained for a flat surface and needs to be reduced to take into account the porosity of the regolith. Reduction factors of four to ten have been discussed for ion sputtering when the ejecta stick with unit efficiency (Hapke and Cassidy, 1978; Johnson, 1989). Because the sodium atoms do not stick with unit efficiency and photon and ion-induced desorption have different dependences on the incident angle, in this paper we use a factor of 3 reduction (Yakshinskiy, personal communication) and therefore use $Q = 1.0 \times 10^{-20}$ cm^2. The rate of ejection of sodium atoms from Mercury's surface can be calculated as the product of the UV flux for photon energies $\geqslant 5$ eV (Yakshinskiy and Madey, 1999; Killen et al., 2001), times the cosine of the zenith angle, the cross section Q, and the concentration c_{Na}. As in Section 2.1 we deduce the probability of ejection of a test-particle from Mercury's surface by photon stimulated desorption during one time step as the product of this rate with the time step of the simulation and divided by c_{Na}.

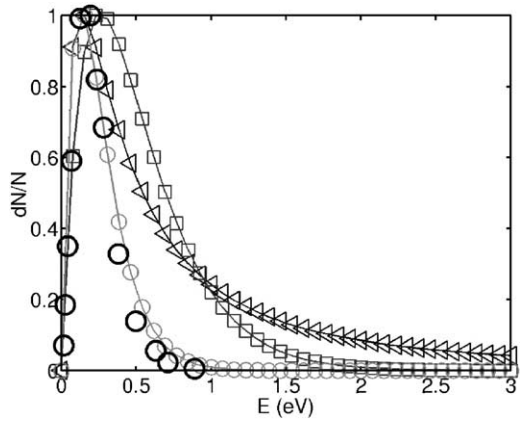

Fig. 3. Energy distribution of sodium atoms ejected from Mercury's surface normalized by its maximum. Square line: micro-meteoroid vaporization. Triangle line: solar wind sputtering. Circle line: photon stimulated desorption. Dark circle symbols: measurements of Yakshinskiy and Madey (1999).

The energy distribution of sodium atoms ejected from lunar type surface by photon stimulated desorption has also been measured (Yakshinskiy and Madey, 1999). We fit that data with a Maxwellian Boltzmann flux distribution $f(E,\theta) = 2\cos(\theta) \times E/(k_{\text{B}}T)^2 \exp(-E/k_{\text{B}}T)$ where E is the energy of the ejecta and θ is the angle between the velocity vector of the ejecta and the normal to the surface. We obtain a rather good fit for a temperature T equal to 1500 K as shown in Fig. 3 where the circle line represents the fit and the dark circle symbols represent the measurements of Yakshinskiy and Madey (1999).

2.3. Solar wind sputtering

Solar wind ion impact has been shown to have highly variable temporal and spatial distributions at Mercury's sur-

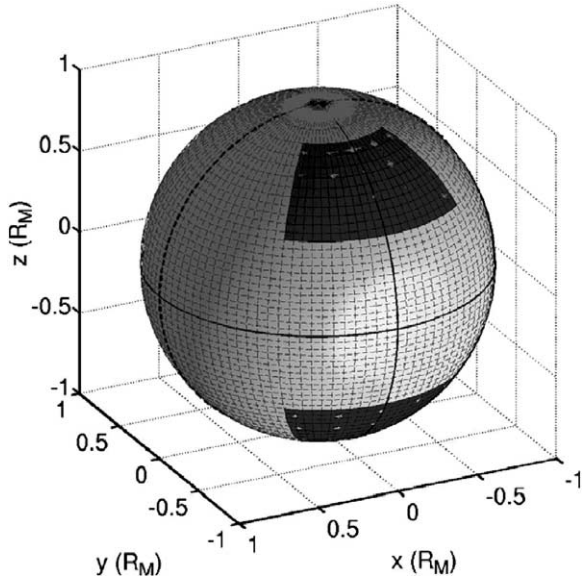

Fig. 4. Spatial distribution of the solar wind impacts at Mercury's surface. The sphere represents Mercury's surface, the Sun is in the $-Y$ direction and the North pole in the $+Z$ direction. The dashed dark line represents the terminator. The total impacted area is symmetric with respect to the subsolar point (intersection of the two solid dark lines at $X = 0$ and $Y = -1$). In our model, the dark surfaces represent the maximal surfaces impacted by solar wind particle. Such surfaces decrease and increase with a period of half an hour from their maximal area (shown on the figure) down to their respective center (at longitude 0° and latitudes ±40°).

face (Kabin et al., 2000; Killen et al., 2001; Sarantos et al., 2001; Ip and Kopp, 2002). This distribution is dependent on the Interplanetary Magnetic Field (IMF) orientation which is highly variable at Mercury's orbit (Luhmann et al., 1998). In Fig. 4, we show a case where solar wind ions penetrate Mercury's magnetosphere along open magnetic field lines and impact a region with area equal to 7.3×10^{16} cm^2 roughly 20% of Mercury dayside surface. This example has been chosen because it is similar to the November 20th case that Killen et al. (2001) reported as the largest fraction of Mercury's surface bombarded by solar wind ions. This is one possible bombarded region which we assume is a source only on the dayside despite the fact that ions can also reach the nightside (Kabin et al., 2000; Ip and Kopp, 2002). Such a bombarded area is most likely non-symmetric with respect to the equator and to the subsolar longitude (Sarantos et al., 2001). Here it is assumed to be symmetric to simplify our discussion. Since solar wind sputtering is variable in time (Sarantos et al., 2001; Killen et al., 2001), we simulate this by varying the surface area bombarded by solar wind ions. We vary the surface from that shown in Fig. 4 down to a negligibly small surface concentrated at the center of each area (at longitude 0° and latitude ±40°). This corresponds to an extreme situation with no solar wind sputtering which should be close to the case of a strong North Bz IMF component (Sarantos et al., 2001; Ip and Kopp, 2002). For simplicity, the variation of the surface area is assumed to follow a cosine law with a time period of half an hour which is of the order of the variation of the Bz IMF component at Mercury (Luhmann et al., 1998). At any time, the whole bombarded surface the two area is kept symmetric with respect to the subsolar point.

The energy distribution of sodium atoms ejected from an Na_2SO_4 surface bombarded by energetic ions has been measured (Wiens et al., 1997). They reported that the best fit of the energy distribution for the ejecta is a Thompson Sigmund distribution of the type $f(E,\theta) = 2 \times \cos(\theta) \times 2EU/(E+U)^3$ with a binding energy $U = 0.27$ eV. Such an energy distribution is displayed in Fig. 3 by the solid line with triangles.

The efficiency of solar wind particles for ejecting sodium atoms from Mercury's surface after long term bombardment can be estimated by using laboratory measurements for the rate of erosion of a silicate surface times the sodium concentration (Flavill et al., 1980). The average yield (the number of ejected particle per impacting solar wind particle, all species considered) suggested by Killen et al. (2001) is equal to 0.15 atoms/ion multiplied by the atomic concentration. This does not take into account the porosity. Hapke and Cassidy (1978) and Johnson (1989) suggest reduction factors of the order of 0.1 and 0.25 respectively for atoms that stick efficiently, and Morgan and Killen (1997) use factors between 0.2 and 0.3. In this paper, we will use an effective yield accounting for porosity equal to 0.06 atoms/ion which is similar to the value suggested by Johnson and Baragiola (1991). Here we use the average yield above and will later consider a distribution of binding sites as the new measurements become better understood.

A typical slow solar wind ion has a drift velocity of 400 km/s. The solar wind density varies roughly as the inverse of the square of the heliocentric distance. We scale the solar wind density at Mercury's position to the mean density in the slow solar wind observed at the Earth orbit, around 10 particles/cm^3 (Schwenn and Marsch, 1991). This leads to a solar wind density of 60 particles/cm^3 at 0.42 AU from the Sun. We use the same relation as Killen et al. (2001) between bombarded area at Mercury's surface and area at Mercury's magnetopause with open magnetic field lines. These authors suggested a factor 4 of increase of this area from Mercury's surface up to the magnetopause. This estimate of the flux of solar wind particles reaching the surface is an upper limit since it neglects any screening effects due to bow shock and magnetopause crossings (Leblanc et al., 2003). The rate of sodium atoms ejected from Mercury's surface by solar wind sputtering is equal to the product of the yield times the solar wind flux reaching the surface times the concentration c_{Na}. From this rate, as in Section 2.1 we deduce the probability of ejection of a test-particle from Mercury's surface by solar wind sputtering during one time step as the product of this rate with the time step of the simulation divided by the concentration c_{Na}. The ejection of sodium atoms produced by solar wind sputtering is assumed to have the same rate everywhere in the bombardment area.

2.4. Micro-meteoroid vaporization

Morgan et al. (1988) were the first to describe the micro-meteoroid vaporization effect at Mercury. These authors assumed a uniform bombardment of the surface. Killen and Ip (1999) reported the observations at the Earth of a two times larger flux on the dawn side (leading side) with respect to the flux on the dusk side (trailing side). They suggested that this asymmetry should be similar at Mercury. In the following, we will assume that micro-meteoroid bombardment generates a flux ejected from Mercury's surface at the center longitude of the leading side (the dawn side) that is two times larger than the flux ejected at the center longitude of the trailing side (the dusk side). This difference is described here as a power law in longitude with the power chosen to fit the above constraints.

Eichhorn (1978) measured the temperature of the vapor produced by hypervelocity particle impacts on solid target and reported values between 2500 and 5000 K. According to OKeefe and Ahrens (1977), the mean speed of the ejecta is roughly half the value of the sound speed. Knowing that the speed of sound in a silicate is between 4 and 6 km/s gives values in agreement with Eichhorn's (1978) measurements. Moreover, Wilson et al. (1999) deduced from observations of the Moon's sodium exosphere during the Leonid Meteor Shower, a speed for sodium ejecta between 2.1 to 2.4 km/s. However, the typical temperature of the vaporized material should be smaller than the measured temperature (Smyth and Marconi, 1995b). In this paper, we use an energy distribution of the sodium ejecta produced by micro-meteoroid vaporization equal to a Maxwellian flux distribution with a temperature equal to 3000 K. That is, we use a most probable speed of sodium ejecta equal to 1.5 km/s and an average speed equal to 2.1 km/s (see square line in Fig. 3).

As stated in Section 2, micro-meteoroid vaporization is assumed to act independently of the three other mechanisms of ejection. It is also not dependent on the local sodium concentration. Following Killen et al. (2001), we consider that micro-meteoroid bombardment at Mercury's orbit evolves with respect to Mercury's distance to the Sun following a power law in $R_{\mathrm{Hel}}^{1.9}$. These authors provide different values of the flux of regolith vaporized by micro-meteoroid bombardment at the perihelion of Mercury based on different theoretical works. We consider an average value equal to a vaporization rate of 5×10^{-15} g/(cm^2 s). Using the putative concentration of sodium in Mercury's regolith for the concentration in micro-meteoroid we deduce that at the perihelion, 5×10^{23} Na/s are ejected by micro-meteoroid bombardment and scale this flux intensity to other Mercury heliocentric distances following a law in $1/R_{\mathrm{Hel}}^{1.9}$. In our model, a test-particle is ejected from the surface by this process in each time interval having a length defined as the weight associated with micro-meteoroid vaporization divided by the flux of sodium atoms ejected by micro-meteoroid vaporization.

3. Mercury exosphere/surface layer

3.1. Mercury's surface layer

In order to get a picture of the surface concentration during one Mercury year, we simulate more than six Mercury years, that is, more than three diurnal cycles. We note the surface density distribution at different heliocentric positions during the last rotation of Mercury around the Sun. The length of this simulation is enough to obtain a solution independent of the initial conditions. It is also enough to reach a steady state between loss and supply averaged over the last year. The only direct sodium supply to Mercury's exosphere/surface layer we considered in this first simulation is the one due to micro-meteoroid vaporization. The initial density in the surface layer was assumed to be uniform and equal to 4×10^{12} Na/cm^2.

In Plate 1, we plot the surface density in Na/cm^2 at different TAA during the last Mercury's year of the simulation (corresponding to the six positions of Mercury displayed in Fig. 2). In each panel of Plate 1, strong inhomogeneities in the surface density appear from night to day sides, from high to low latitudes and from morning to afternoon. On the dayside, the density of the sodium in the surface layers that is available for desorption is highly depleted. This depletion is due to the short residence time due to thermal desorption near the subsolar point as displayed in Fig. 1 (solar zenith angle close to 0°). This time is controlled by the surface temperature (Fig. 1), and the dayside depletion covers a smaller area at the aphelion than at the perihelion. On the nightside, the highest surface density is reached in polar regions just before the dawn terminator. Nightside polar regions absorb more sodium than nightside equatorial regions because of lower surface temperatures in the model used, and sodium atoms readily migrate from day to night sides over the poles. These high surface concentrations are almost as efficiently depleted during the day as the equatorial regions, especially close to the perihelion, and, therefore, lead to local maxima in Mercury's Na exosphere (see Section 3.2).

The changes in the surface density distribution (in particular in the nightside) differ significantly during Mercury's motion from perihelion to aphelion (TAA from 0° to 180°) from the changes occurring during Mercury's motion from aphelion to perihelion (TAA from 180° to 360°) despite similar solar conditions. This is seen by comparing panels at TAA = 125° and at TAA = 252°. The principal difference is that Mercury's heliocentric radial velocity is opposite in sign. As a consequence, two sodium atoms ejected from Mercury's surface at two symmetric positions with the same initial velocity in Mercury's frame have opposite radial heliocentric velocities. A sodium atom with a negative radial heliocentric velocity is subject to a decreasing solar radiation pressure up to a null value, whereas a sodium atom with a positive radial heliocentric velocity is subjected to an increasing solar radiation pressure. As a consequence, as seen in Plate 1, the recycling from day to night sides is

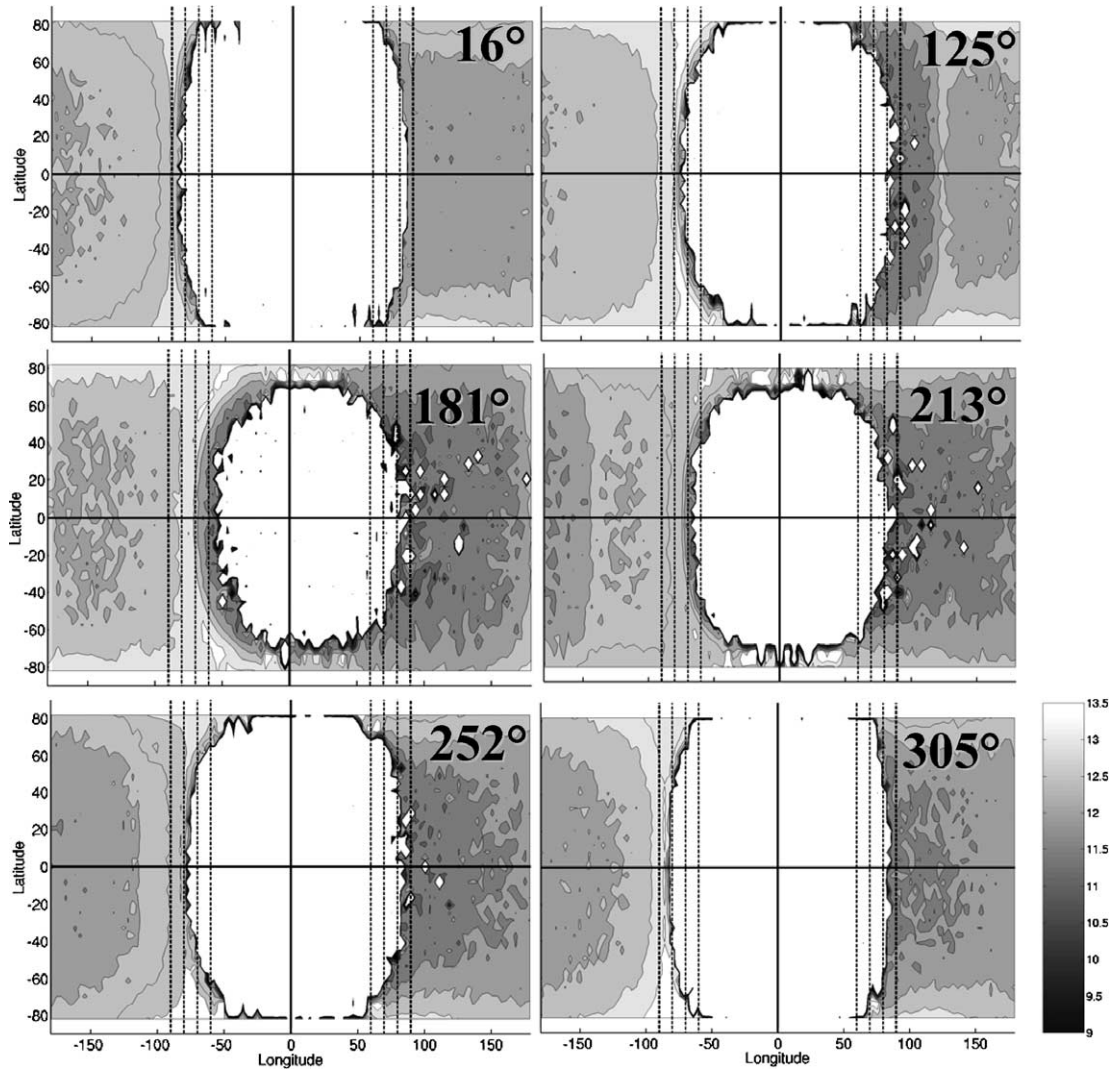

Plate 1. Surface density distribution calculated at different Mercury's heliocentric positions (in Na/cm^2). The corresponding TAA is indicated on each panel. Mercury subsolar point is at the center of each panel, the dark horizontal line represents the equator line, and the dark vertical line the subsolar line. Dayside is between $\pm 90°$ in longitude (thickest dashed vertical lines). Thin vertical dashed lines at $\pm 80°$, $\pm 70°$, and $\pm 60°$ in longitude are also plotted. Logarithmic scale bar is indicated at the right bottom.

different for these two parts of Mercury's motion. In general, the depletion of the surface density starts very early in the morning, the earliest depletion being when Mercury is close to perihelion (TAA = 16° and 305° Plate 1) where surface temperature is the highest. At 181°, the depletion on the dayside is the smallest and the nightside density slowly increases.

Figure 5 shows the evolution of the total number of sodium atoms in Mercury's surface layer with respect to TAA (solid line with value divided by 100). This value evolved from a minimum around TAA = 225° equal to 1.95×10^{30} Na to a maximum around TAA = 145° equal to 2.72×10^{30} Na. These values are smaller than the initial number we put inside Mercury's exosphere/surface layer (3×10^{30} Na). It shows the independence of our final result with respect to this initial condition and the strong dependence of our result on the supply rate. Using the same surface

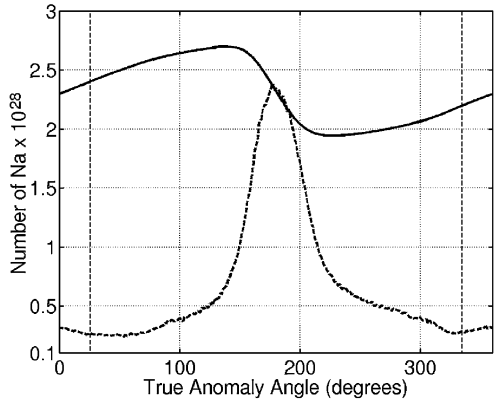

Fig. 5. Total number of sodium atoms with respect to TAA. Solid line: in Mercury's surface. Dashed line: in Mercury's exosphere. The total number of sodium in Mercury's surface has been divided by 100. The two vertical lines indicate the position of the inversion of the Sun rotation around Mercury.

density, 7.5×10^{14} atoms/cm^2, we found a sodium concentration in Mercury's surface that varies between 0.005 and 0.003. Figure 5 is also a good illustration of the difference between the two portions of Mercury's motion around the Sun since there is a significantly larger number of sodium atoms in Mercury's surface between TAA $= 0°$ and TAA $= 180°$ than between TAA $= 180°$ and TAA $= 360°$. The strongest variation in the total content of Mercury's surface is correlated with the largest variation in the total content of the exosphere (dashed line in Fig. 5) as it will be discussed in the next section.

3.2. Mercury exosphere

In Plate 2, we display Mercury's sodium exosphere at the same six heliocentric positions of Mercury as in Plate 1. In each panel, the column density distribution around Mercury (circle dashed line) is displayed as it would be seen from the Earth with a phase angle of 120° (see the example in Fig. 2 of the geometry of the observation at TAA $= 125°$). The sodium hidden by Mercury from the observer has not been included, nor has the sodium in Mercury's shadow. A phase angle of 120° corresponds to the morning side of Mercury (Mercury's rotation vector is in the $+Z$ direction in Plates 2 and 3). The terminator is indicated on each panel by a dashed line.

In order to properly compare our results with observations we use a linear scale in Plate 3 rather than the logarithmic scale used in Plate 2. These latter plots can be compared directly with the published images of Mercury's exosphere obtained from the Earth with the same phase angle (see Section 4). Plate 2 provides a general picture of Mercury's exosphere (in particular the size of the exosphere and its tail) whereas Plate 3 details the maxima of the column density usually reported in observations from the Earth. Plate 2 shows that close to the aphelion/perihelion, the solar radiation pressure tends to zero and, therefore, the sodium exosphere is less elongated. By contrast, around 252° and 125° Mercury's tail is largest. Due to the changing radiation pressure, the tail elongation with time is different in going from perihelion to aphelion from the elongation with time in going from aphelion to perihelion as explained in the previous section.

Plate 3 shows that at all heliocentric positions of Mercury high latitude maxima in the column density occur. These are due to the maxima in the Na concentration in Mercury's surface which are formed on the nightside at high latitudes (Plate 1) and are not due to solar wind sputtering. In order to check the origin of these maxima, we performed another simulation similar to the one described in Section 2 but in which solar wind sputtering contribution was suppressed. The results of this simulation provided similar images of Mercury's Na exosphere as the ones displayed in Plates 2 and 3. Therefore the high emission brightness at high latitudes previously attributed to solar wind sputtering effect can be due to maxima in the surface Na concentration at high latitudes on the nightside (Plate 1).

It is somewhat surprising that the total amount of sodium in the exosphere is equal to $\sim 2.4 \times 10^{28}$ Na at aphelion and $\sim 0.27 \times 10^{28}$ Na at perihelion (Fig. 5 dashed line). The minimum in the total content of the exosphere is reached around TAA $= 50°$ and TAA $= 330°$ (Fig. 5) and the maximum at the aphelion. As shown in Fig. 5, the difference between aphelion and perihelion is not simply correlated to the total surface content but is more closely correlated to the Sun's apparent rotation around Mercury. This determines the speed with which the enriched nightside surface passes the dawn terminator into the dayside. This speed is highest at the perihelion and minimum at the aphelion. Actually the rotation of the Sun around Mercury changes direction at TAA $= 335°$ and TAA $= 23°$ (vertical dashed lines in Fig. 5). Another factor which contributes to enhance the total content of Mercury's exosphere at aphelion with respect to perihelion is the almost two times lower ionization frequency of Na atoms at aphelion than at perihelion. Finally, Fig. 5 shows also that the total content of Mercury exosphere plus Mercury's surface layer is not constant along one Mercury's year but is maximal at TAA $= 140°$ with 2.72×10^{30} Na and minimal at TAA $= 225°$ with 1.96×10^{30} Na. Indeed loss and supply rates are equal when averaged over one Mercury's year but can be significantly different at particular heliocentric positions of Mercury.

Figure 6 displays the percentage of the total amount of Na atom which has been ejected from Mercury's surface by each of the four processes considered in this work. Thermal desorption produces from 60 to 95% of the Na ejecta, whereas photon stimulated desorption ejects between 4 to 18% of the Na ejecta, meteoroid vaporization between 0.8 to 30% and solar wind sputtering between 0.1 to 0.7%. Meteoroid vaporization depends only on Mercury's heliocentric distance and therefore is minimum at the aphelion. Photon stimulated desorption depends on Mercury's heliocentric distance and on its competition with thermal desorption. Therefore, it is minimum at perihelion and aphelion. Solar wind sputtering shows a significant maximum at the perihelion despite a decrease of the solar wind density by more than a factor two. Indeed in our model, we limit the sputtering by solar wind particles to dayside high latitudes which are significantly less depleted at the aphelion than at the perihelion. Solar wind sputtering is crucially dependent on the surface concentration of the bombarded surface and, therefore, on the efficiency of thermal desorption. Because of this it contributes more to the exosphere at the aphelion than at the perihelion. Solar wind sputtering also displays a dependence on Mercury's distance to the Sun as illustrated by the two minima at TAA $= 120°$ and 260°. As stated previously and illustrated in Fig. 6, solar wind sputtering is a small net contributor to Mercury's Na exosphere except around the perihelion. Indeed we compared Plate 2, TAA $= 181°$ with the plot issued from the run without solar wind sputtering. In the latter case, there is no clear extension of the Na

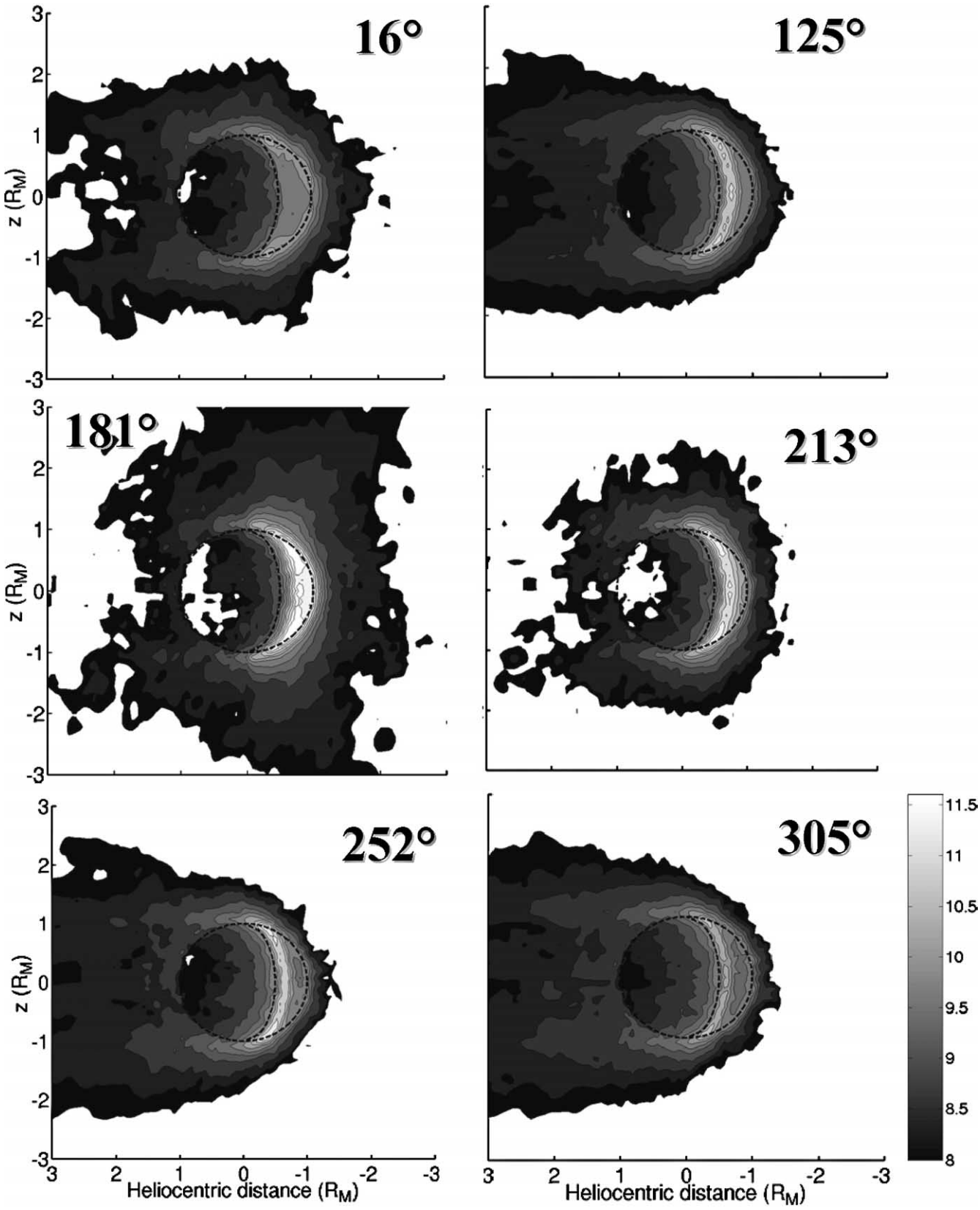

Plate 2. Sodium column density (in Na/cm^2) distribution in Mercury's exosphere at different Mercury's heliocentric positions (at TAA indicated on each panel) as seen for a phase angle of 120°. $+Z$ value corresponds to the North pole. Dashed circle line represents the limit of Mercury surface and the dashed line inside each circle the terminator at such phase angle. The Sun is placed at the right side of Plate 2 (negative heliocentric distances). Logarithmic scale bar is indicated at the right bottom.

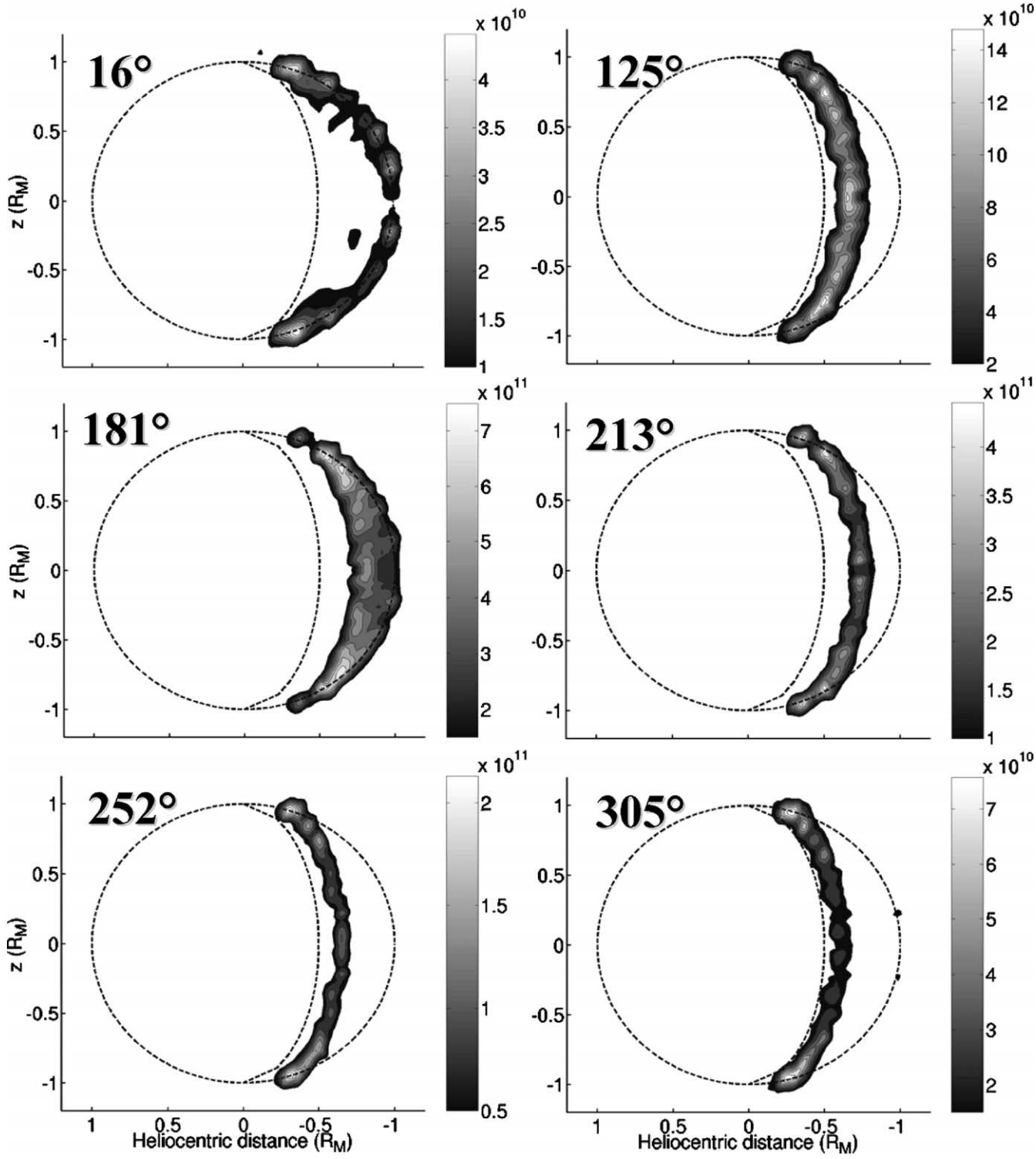

Plate 3. Same as Plate 2 but with a linear scale. Contrary to Plate 2 different linear scale bars have been used as indicated at the right of each panel.

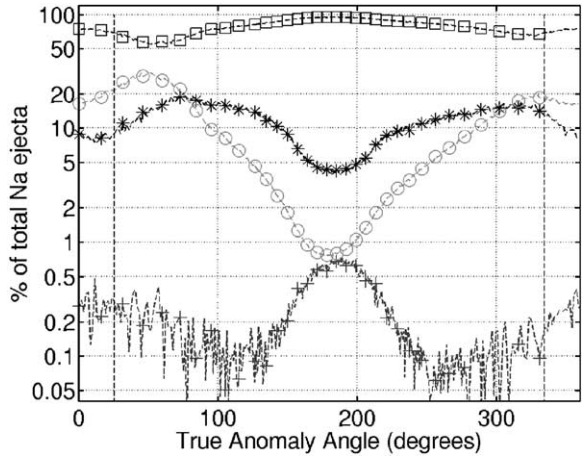

Fig. 6. Relative contribution to the total amount of Na atom ejected from Mercury's surface for the four processes of ejection considered in this paper with respect to TAA. Square line: thermal desorption. Star line: photon stimulated desorption. Circle line: meteoroid vaporization. Cross line: solar wind sputtering. The two vertical lines indicate the position of the inversion of the Sun rotation around Mercury.

cloud around $Z = \pm 2$ R_M and $X = +1$ R_M like that seen in panel TAA $= 181°$ Plate 2. Sputtering is the most energetic ejection (Fig. 3), a property which is partially erased by the effect of the solar radiation pressure except at the perihelion and aphelion. Close to the perihelion it represents around $1/5$ of the energetic Na produced at Mercury's surface (by photon stimulated desorption, meteoroid vaporization and solar wind sputtering, see Fig. 6). At any other TAA, its role is minor with respect to photon stimulated desorption and meteoroid vaporization (Fig. 6) and is reduced by solar radiation pressure.

Another illustration of the importance of the dayside surface depletion is the position of the maxima in the column density in Plate 3. If we allow no depletion of the sodium concentration in the surface layers, the maximum should occur at the dayside limb because the ejection rates for photon stimulated desorption and thermal desorption are maximum at the subsolar point (just behind Mercury apparent surface in panels Plates 2 and 3). In Plate 3 the band of maximal column density moves from the limb at TAA $= 181°$ to a band close to the sunrise terminator at TAA $= 305°$. This change in position is linked to the surface temperature which defines the region where the full depletion of the dayside surface starts (see Plate 1). The case of TAA $= 16°$ is due to the fact that the Sun rotation around Mercury did change sign (at TAA $= 335°$) with respect to other Mercury's positions displayed in Plates 2 and 3. As a consequence the maximum in the column density is formed at the sunset terminator. However this effect remains rather small since at TAA $= 16°$ the Sun's rotation around Mercury is almost null.

At a phase angle greater than 180°, an observer would see the afternoon side of Mercury. At such phase angles, the shape of the sodium exosphere appears significantly different from that at smaller phase angles. Figure 7 provides two images of the exosphere seen from its afternoon side. On

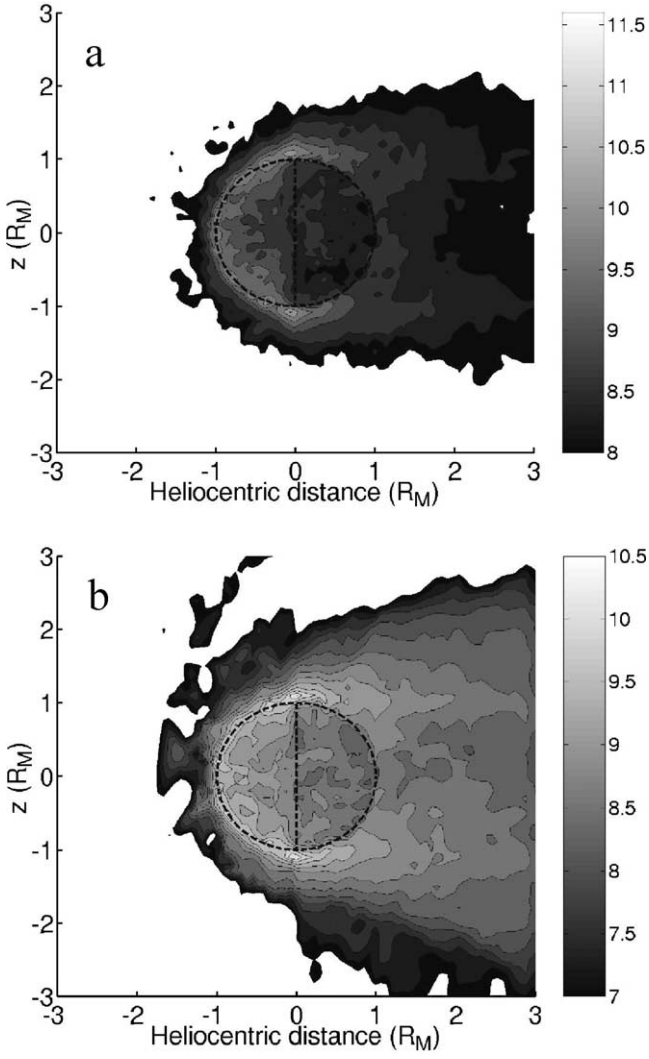

Fig. 7. Distribution of the Na column density around Mercury (dashed circle) as seen from the Earth with phase angles corresponding to the afternoon side of Mercury. The dashed line inside the dashed circle represents the dusk terminator as seen by the observer. $+Z$ values correspond to the North pole. The Sun is placed at negative heliocentric distance (left side of the figure). The scale is logarithmic in Na/cm^2. (a) Phase angle of 270° and TAA $= 125°$ (same image as in Plate 2 but at a different phase angle). (b) Phase angle of 270° and TAA $= 289°$.

both images, taken at two different TAA for a phase angle of 270°, strong maxima in the column density appear close to both North and South poles. These maxima are not the effect of the solar wind sputtering but the extension above the limb of the high column density band which is shown in Plate 3 as illustrated by comparing Fig. 7a with Plate 2, TAA $= 125°$ for which the same scale bar has been used.

3.3. Mercury morning/afternoon asymmetry

Sprague (1990a) and Sprague et al. (1997) described a large set of observations of the sodium exosphere made at several heliocentric Mercury's positions for variable phase angles. These authors found an average morning/afternoon

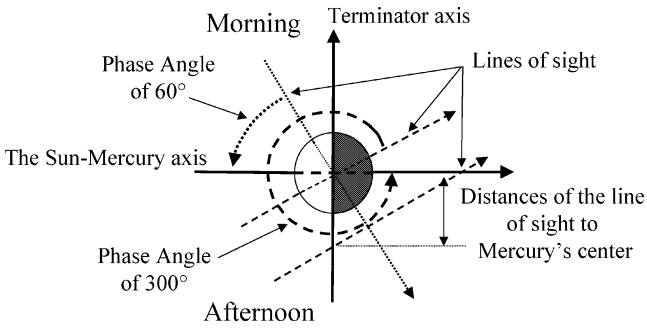

Fig. 8. Sketch of the method used to calculate the morning/afternoon column densities ratio. Represented is Mercury projected onto its orbital plane. The axis of Mercury's rotation is oriented perpendicularly to the figure towards the reader.

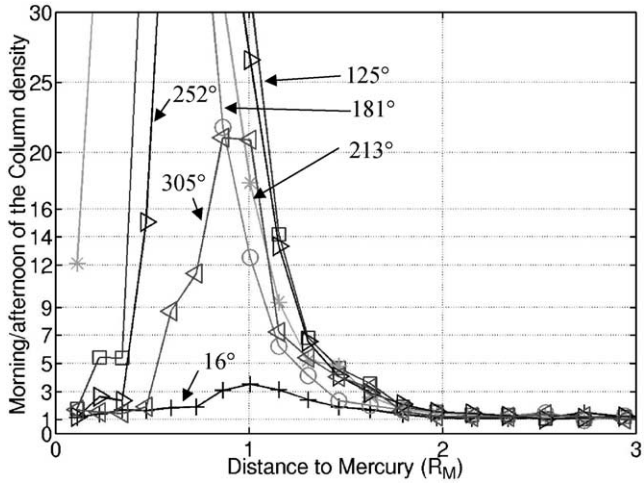

Fig. 9. Ratio between sodium column density on the morning side of Mercury and sodium column density on the afternoon side. Cross line: TAA = 16°. Square line: TAA = 125°. Circle line: TAA = 181°. Star line: TAA = 213°. Triangle dark line: TAA = 252°. Triangle light line: TAA = 305°. The axis of the ratio has been limited to a value of 30 in order to detail the profiles close to zero.

ratio of the sodium column density greater than 3. This implies a sodium exosphere where most of the sodium atoms are ejected from the surface in the morning, probably by photon stimulated and thermal desorptions. The simulation displayed in Plates 1, 2, and 3 can provide information about the possibility of observing such an asymmetry with the method used by Sprague et al. (1990a, 1997). For each of the positions in Plates 2 and 3, we consider two different phase angles, 60° and 300° the first with a line of sight perpendicular to Mercury's morning surface and the other to Mercury's afternoon surface (see Fig. 8).

For each phase angle, we calculate the column density along lines of sight with the same phase angle and placed inside Mercury's orbital plane at different distances from Mercury. For each of these, we calculate the distance to the center of the intersection of the line of sight with the terminator axis (Fig. 8). For a phase angle of 60° we only consider those lines of sight on the morning side of Mercury that intersect the terminator above the sun axis in Fig. 8 and for a phase angle 300° on the afternoon side below the sun axis. From the column densities versus such distance to Mercury's center for both phase angles we calculate the ratio of these common densities vs. distance to Mercury's center. This corresponds to the ratio between morning/afternoon column densities defined by Sprague et al. (1997).

In Fig. 9 the morning/afternoon column density ratio with respect to distance to Mercury's center is shown at different TAA. This plot has been limited to ratios below 30 in order to show detail for small ratios. Actually, the curve at TAA = 181° (star line) peaks at a value close to 100, but depends on the spatial resolution used. A lower spatial resolution gives a smaller peak value. The maximum value of the morning/afternoon ratio is obtained for a TAA = 181° at the aphelion and the minimum value at perihelion.

Close to perihelion thermal desorption dominates even close to sunrise. Therefore, the migration towards the nightside is so quick that any column density increase is very localized in longitude and is difficult to observe (Killen and Morgan, 1993b). On the contrary, at the aphelion, the depletion of the surface is less affected by thermal desorption so that photon stimulated desorption and solar wind sputtering contribute significantly to the flux of sodium atoms ejected from Mercury's morning surface over a larger range of longitude. This leads to a morning enhancement in the exosphere density over a large area. Moreover because photon stimulated desorption and solar wind sputtering are more energetic ejection processes than thermal desorption, this enhancement is increased by the larger residence time in the exosphere of an ejected sodium atom. Averaged between 0 to 1 R_M from Mercury center, we found that between TAA = 320° and TAA = 40° the ratio between morning and afternoon column densities is close to one, whereas between TAA = 40° and TAA = 140° and between TAA = 230° and TAA = 320° it is larger than 5 and a maximum between TAA = 140° and TAA = 230° (larger than 10).

The asymmetry introduced by micro-meteoroid vaporization (Section 2.4) does not play a significant role. Indeed the same simulation but with a uniform spatial distribution of the ejecta by micro-meteoroid vaporization at Mercury's surface gave the same results. Another potential source of the morning/afternoon asymmetry is the energetic solar particles which can impact Mercury's surface with a larger flux on the morning side than on the afternoon side (Leblanc et al., 2003). However Leblanc et al. (2003) have shown that such energetic particles are not efficient sources of sodium atoms except at the arrival of the shock associated with an energetic particle event. A third possible source of asymmetry in the ejecta distribution is magnetospheric ion sputtering. Following Delcourt et al. (2003) and Ip (1993), the impacting flux of magnetospheric ions at the surface could have a significant morning/afternoon asymmetry because of the electric field of convection associated with the solar wind (see also Killen and Sarantos, 2003).

4. Comparison with observations

Detailed observations of Mercury's sodium exosphere have been recently published by Potter et al. (2002). These authors provide sodium column densities and velocity profiles in Mercury's orbital plane at two Mercury positions: TAA = 125° on the 05/26/2001 and TAA = 114° on the 06/05/2000. In this paper, we limit our analysis to their observation at TAA = 125° which they observe with a phase angle of ∼ 120°. A simulation starting at TAA = 16°, with the initial surface density distribution displayed in Plate 1 and initial exosphere density distribution displayed in Plate 2, TAA = 16°, is carried out until Mercury reaches TAA = 125° (∼ 500 hours of real time). From TAA = 16° to TAA = 125° a few million test-particles were ejected and followed around Mercury. At TAA = 125°, we integrated the 3D distribution of the density around Mercury as described in Section 2. Micro-meteoroid vaporization was assumed to be the only supply of new sodium atoms.

Potter et al. (2002) provided an image of the sodium exosphere in their Fig. 2. The corresponding result from our model is shown in Plates 2 and 3 for TAA = 125°. The two images are roughly in agreement with the same enhancement in the exosphere above the northern and southern latitudes and the same decrease of the column density with increasing distance from Mercury.

The density profile in Mercury's orbital plane is integrated along lines of sight placed at different distances to Mercury and for a phase angle equal to 120°. The resulting column densities vs. heliocentric distance from the center of Mercury are shown in Fig. 10a. In this calculation we assumed the slits covered 1 R_M square as in Potter et al. (2002). In Fig. 10a, the heliocentric distance to Mercury has been corrected for the difference in phase angle by dividing the line of sight distance from Mercury's center by the sine of the phase angle. Negative distances correspond to the direction towards the Sun. In Fig. 10a, the dark crosses are the observations of Potter et al. (2002). On the dayside of Mercury and above the limb (indicated by the vertical dashed line at −3000 km heliocentric distance from Mercury), the comparison between simulation and observation is good. The maximum of the simulated column density and the maximum of Potter et al. (2002) observations are of the same order. However the maximum of the simulated column density occurs close to the terminator (the vertical dashed line at −1500 km) whereas the maximum of the observations of Potter et al. (2002) is slightly in Mercury's nightside which is probably due to smearing effect. The good agreement between both maximum values indicates that in our model the total quantity of sodium in Mercury exosphere/surface layer is of the right order. Moreover, the position of the observed maximum is closer to the Sunrise terminator than to the limb. This indicates that the ejection rate of sodium atoms is larger close to the terminator than close to the subsolar region in agreement with our suggestion that a strong surface depletion occurs at Mercury.

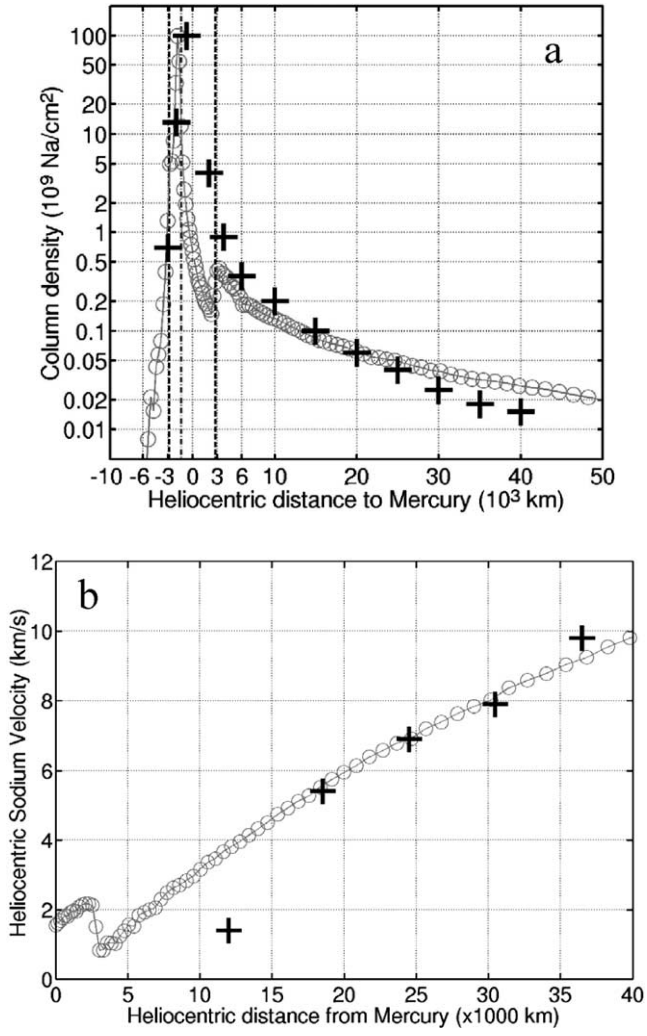

Fig. 10. Column density and velocity profiles of Mercury sodium exosphere at TAA = 125° for a phase angle of 120°. Negative heliocentric distances to Mercury correspond to the Sun direction. (a) Sodium column density. Circle line: our simulation. Dark crosses: Potter et al. (2002) observations. Mercury' surface for such a phase angle is between the two dark vertical dashed lines at ±3000 km. The vertical dash-dot line at −1500 km represents the terminator as seen by the observer. (b) Heliocentric velocity of the sodium atoms relative to Mercury. Dark crosses: Potter et al. (2002) observations. Circle line: our simulation.

On the nightside, between 3000 and 20,000 km from the center of Mercury's disc, there is also a good agreement between the simulation and observations. This shows that the amount of sodium ejected from Mercury' surface and able to reach the nightside is also correctly estimated in our model. These are the energetic sodium atoms ejected at the dayside by photo stimulated desorption, solar wind sputtering and micro-meteoroid vaporization. Beyond 20,000 km, the decrease of the simulated column density with increasing heliocentric distance is much slower than that observed (Potter et al., 2002). However these authors suggested that they may have failed to properly follow the tail centerline with increasing heliocentric distance leading to a more rapid decrease of the column density. In a recent report, Potter, Morgan

and Killen ("The sodium tail of Mercury," DPS conference, 2002) showed new observations having a smaller decrease of the observed column density with increasing heliocentric distance and observed a column density at 40,000 km from Mercury, of the order of the one we obtained.

Figure 10b provides the profile of the average heliocentric radial velocity of sodium atoms relative to Mercury which is plotted vs. heliocentric distance. These values are calculated along lines of sight with phase angle equal to 120° using the same resolution as for Fig. 10a (1 R_M squared slits placed perpendicularly to the Sun–Mercury axis). Figure 10b displays the velocity profile at TAA = 125° as calculated by our simulation (circles) and the measured velocity profile (Potter et al., 2002) (crosses). The observations below 15,000 km from Mercury are "probably dominated by unshifted emissions from the bright planet scattered into the field of view" (Potter et al., 2002). Therefore, in Fig. 10b below 15,000 km no comparison between simulation and observation should be made. Above this distance, there is a very good agreement between our simulation and the observations of Potter et al. (2002). We find that the radial velocity profile shows a significant dependence on the TAA due to the significant change in the solar radiation pressure with heliocentric position (Section 3.2).

Potter and Morgan (1990, 1997b) published images of Mercury's Na exosphere displaying significant high latitude emission brightness on the 02/16–18/1989 at TAA = 130° and on the 07/07/1989 at TAA = 170° for phase angle of 70° and 80°, respectively (Potter and Morgan, 1990) and on the 12/07/1990 at TAA = 275° and phase angle of 80° (Potter and Morgan, 1997b). Their observations can be compared to Plate 3 for panels with close TAA that is TAA = 125° for the 02/16–18/1989, at TAA = 181° for the 07/07/1989 and at TAA = 252° for the 12/07/1990. In all three panels, maxima in the column density of Na exospheric atoms appear at high latitudes. Their positions are at slightly higher latitude than the ones reported in these observations. This can be due to the resolution of the observations of Potter and Morgan or to smearing effects. It can be also due to the fact that position and nature of the maxima of the surface Na concentration displayed in Plate 1 are probably different from those obtained using the idealized conditions in our model. As an example, North/South asymmetries of the observations which do not appear on Plate 3 can be explained by surface Na concentration asymmetries explained in the next paragraph and in Section 5.2. In any case, the observed high latitude emissions reported can be roughly reproduced without solar wind sputtering but with the formation of maxima of the surface Na concentration at high latitude just before the dawn terminator.

Potter and Morgan (1997a) reported observations of North/South high latitude emissions when Mercury was at TAA ≈ 289° with a phase angle of ∼ 270°. These observations can be compared with the result displayed in Fig. 7b. As explained in Section 3.2 and illustrated in Fig. 7b, because Potter and Morgan (1997a) observed the afternoon side of Mercury, the high latitudes emissions they reported should be partly due to the extension above the limb of the band of maximal column density close to sunrise. Actually, these maxima would have occurred between Mercury's disc (below the limb) and the observer if they were due to the solar wind sputtering of the surface. The three-day variability that these authors observed can be understood if we remember that in our model the surface density distribution is idealized. That is, it is symmetric with respect to the equator. Moreover, significant variations in the surface density distribution can occur on time scales shorter than the time scale of the surface density variation described here. There are a number of processes which can induced such variations and asymmetries among which the highly variable and non-symmetric solar wind sputtering has been widely described. The North/South motion of the peak in the emission brightness observed by Potter and Morgan (1997a) is potentially due to a North/South asymmetry in the surface density distribution.

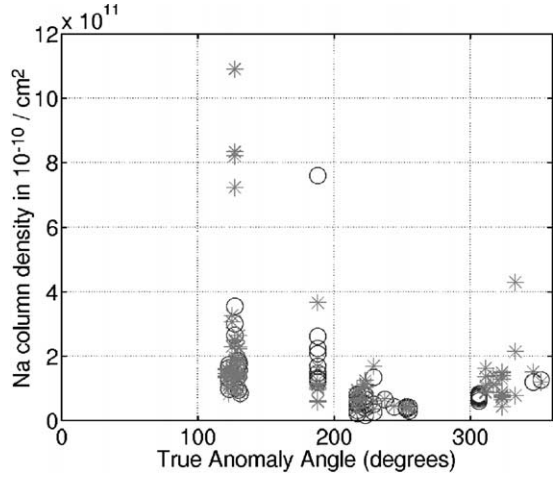

Fig. 11. D1 and D2 column densities with respect to TAA issued from Table II of Sprague et al. (1997). Only column density obtained by observations with morning side viewing (positive phase angle in Sprague et al. (1997)) and close to the equatorial plane ($Y \leqslant 0.3$ Mercury radius in Table II) have been plotted. This figure was kindly suggested by A.L. Sprague.

Sprague et al. (1997) published the largest ground based data set available on Mercury's Na exosphere. We selected in their Table II D1 and D2 column densities obtained from morning side (positive phase angle value in Table I) and with the Y value in Table II lower than 0.3 of Mercury's radius. This selection is done in order to suppress any morning/afternoon asymmetry effect and any high latitude emissions potentially associated with inhomogeneities in the surface concentration. The result is plotted in Fig. 11 vs TAA. Figure 11 shows that the selected column densities display a clear maximum close to aphelion in agreement with our calculation of the evolution of the total content of Mercury's Na exosphere (Fig. 5).

Potter et al. (1999) reported more than 7 Earth days of observations of Mercury sodium exosphere from November 13 to 20th 1997. These authors observed a variation by a

factor three of the total amount of sodium around Mercury with maximal emission brightness moving from northern to southern hemispheres. During this time Mercury moved from TAA = 213° to TAA = 239° with a quasi constant phase angle of 50°. Killen et al. (2001) improved the calibration of these observations and deduced that the total amount of sodium in Mercury exosphere changed from 1.1×10^{28} sodium atoms the first day up to 3.1×10^{28} sodium atoms the last day. Our simulation of the same period, predicts a total amount of sodium in Mercury exosphere decreasing from 1.1 to 0.9×10^{28} sodium atoms between TAA = 213° and TAA = 239° (see Fig. 5). Therefore the increase of the total exosphere content reported by Potter et al. (1999) is clearly in contradiction with our model, except for the first day of their observations which shows a surprisingly close agreement with our model. It suggests that some of the parameters in the simulation changed between November 13 to 20th, as it will be discussed in Section 5.2.

5. Discussion

5.1. Supply and loss rates

Loss and supply rates of sodium atoms at Mercury are linked and should be equal over long time scales, of the order of a Mercury year. In Sections 3 and 4, the only supply considered was micro-meteoroid vaporization. Other sources have been suggested in Section 2. In order to address this we perform simulations using a second model in which the supply of sodium atoms into Mercury's surface layer is large enough to maintain a constant surface layer density (e.g., a large diffusion rate inside the surface layer grains and the exposure of fresh grains by meteoroid mixing of the regolith). The concentration of sodium atoms in Mercury's surface is set equal to 0.0053 and the surface density to 7.5×10^{14} atoms/cm^2. The same type of simulation as described in Section 4 of more than 500 hours of real Mercury time is carried out. Since the surface density is now assumed to be uniform, there is no need to simulate several years of Mercury. Therefore, the simulation is stopped at TAA = 125°. Figure 12a compares the result of this simulation (circles) with the observations of Potter et al. (2002) (crosses). In Fig. 12a, the maximum in the column density is up to 500 times higher than the observed peak. Moreover, this maximum occurs at the limb, −3000 km from Mercury center, as expected since the maximum ejection rate occurs at the subsolar point when there is no surface depletion. Although the agreement between simulation and observation is rather good at 3000 km on the nightside, global agreement cannot be obtained due to the peak difference. This suggests that the observations cannot be explained without a strong dayside depletion of Mercury's surface.

As a second test case, we assume that the supply of new sodium is equal to the meteoroid vaporization supply plus additional sources, such that the total supply rate is increased by a factor ∼ 2 with respect to the meteoroid supply. This could be due to enhanced diffusion in grains and meteoroid mixing of the regolith. For this simulation, we use the method described in Section 2. A first simulation with low statistics for more than 6 Mercury's years is carried out in order to obtain the surface and exosphere density distributions with respect to Mercury's heliocentric position. A second simulation using the result of this first simulation is then carried out with better statistics in order to get a good description of Mercury's sodium exosphere. The result of this second simulation at TAA = 125° and for a phase angle of 120° is given in Fig. 12b. The maximum of the simulated column density (line with circles) is larger by a factor two than the maximum of the observations (dark crosses). Both maxima in this case are at the same spatial position (−1500 km). Contrary to the result in Fig. 12a, reducing the simulated column density by a factor two would significantly improve the agreement between simulation and observation. However, this would imply reducing the supply rates and sources rates by the same factor. The total supply rate of new sodium atoms into Mercury exosphere/surface layer cannot then be much larger than the average supply rate of 3.5×10^{23} Na/s that has been considered in Sections 3 and 4.

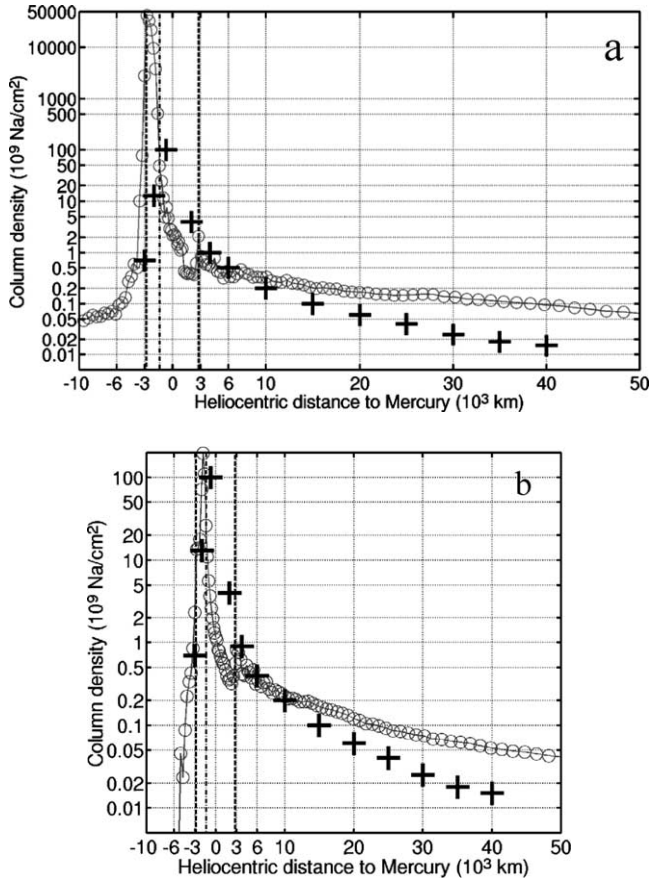

Fig. 12. Same legend as Fig. 10a. (a) In the case of no depletion in the surface. (b) In the case of a supply rate of new sodium atoms four times the supply rate due to micro-meteoroid vaporization only.

The loss of sodium atoms from Mercury is composed of neutral particles which reach large distances from Mercury (50 R_M) and have a negligible probability of returning to Mercury's surface, and of neutral atoms ionized by solar photons and accelerated away from Mercury's magnetosphere by the solar wind electric field of convection. Most of the ions created outside Mercury's magnetosphere ($\sim 12\%$ of the total amount of new ions) are lost. But most of the pick-up ions are created inside Mercury's magnetosphere some of which impact the surface. The percentage impacting the surface depends strongly on the IMF orientation which also defined the proportion of open magnetic field lines of Mercury's magnetosphere (Luhmann et al., 1998; Killen et al., 2001; Ip and Kopp, 2002). Ip (1990) estimated that 30% of the ionized sodium atoms escape Mercury's magnetosphere whereas Hunten and Sprague (1997) estimated this value to be equal to 60%. The probability that a newly ionized particle will impact the surface is highly dependent on its position and velocity vector when ionized (Delcourt et al., 2003). Therefore it is difficult to accurately calculate this proportion without an accurate model of the neutral exosphere. To obtain an estimate, we assumed that a new ion created at an altitude smaller than 10 km is most probably reabsorbed by Mercury's surface. These ions represent $\sim 8\%$ of the total with the rest considered as lost. With these assumptions, the loss averaged over one Mercury year due to neutral particles reaching large distances from Mercury is equal to $\sim 3 \times 10^{22}$ Na/s. The rate of creation of new ions outside Mercury's magnetosphere is of $\sim 4 \times 10^{22}$ Na/s and the loss due to those created inside Mercury's magnetosphere of $\sim 2.8 \times 10^{23}$ Na/s for a net average loss $\sim 3.5 \times 10^{23}$ Na/s equal to the average supply rate of new sodium atoms to the exosphere/surface layer. The part of the loss outside of the magnetosphere is $\sim 7 \times 10^{22}$ Na/s. This becomes $\sim 2.4 \times 10^{23}$ Na/s if 60% of the ions created inside Mercury's magnetosphere are lost (Hunten and Sprague, 1997) and up to $\sim 1.5 \times 10^{23}$ Na/s if only 30% are lost (Ip, 1990). A larger or smaller supply would be associated with a larger or smaller sodium loss, respectively.

5.2. Temporal variations and spatial inhomogeneities

In the results shown in the previous sections, we observe significant North and South maxima:

- At any TAA with morning side viewing, north and south maxima in the line-of-sight column densities can be seen in Plate 2 and even better in Plate 3. These maxima at high latitude are not associated with solar wind sputtering which produces on average less than 0.5% of the total ejecta. Rather, they are due to maxima in the surface Na concentration at high latitudes on the nightside just before the dawn terminator. In a more realist model asymmetries between north and south maxima, as well as maxima at lower latitudes, can be produced by latitudinal inhomogeneities in the Na concentration.

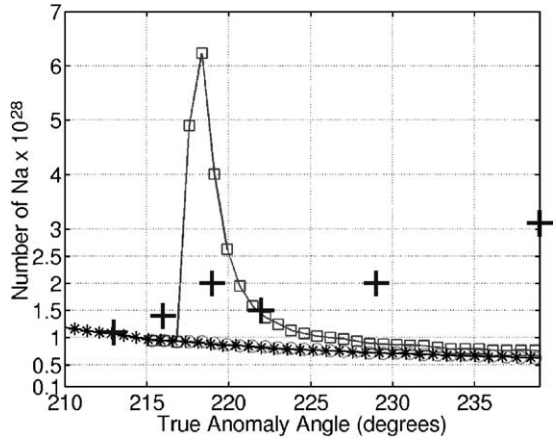

Fig. 13. Total number of sodium atoms in Mercury's exosphere with respect to TAA. Solid star line: for quiet solar wind conditions (same as dashed line in Fig. 5). Solid circle line: for quiet solar wind conditions up to TAA = 217° followed by a brutal increase of the solar wind density by a factor 3 maintained until TAA = 239°. Solid square line: for quiet solar wind conditions up to TAA = 217° followed by a brutal shift of the sputtered surface by 50° in longitude towards the dawn terminator maintained until TAA = 239°. Crosses: Potter et al. (1999) measurements.

- The second case for which we observe strong north and south local maxima in the column density occurs when Mercury is seen from its afternoon side (phase angle larger than 180°), as shown in Fig. 7. As explained in Section 3.2, such maxima correspond to the extension above the limb of the band of column density maxima produced on the morning side. These maxima could also reflect any North/South inhomogeneities in the surface Na concentration.

Different scenarios for the large temporal variations in the total content of the sodium exosphere observed by Potter et al. (1999) will be discussed. The solid line with stars in Fig. 13 between TAA = 210° to TAA = 239° provides the total content of the sodium exosphere obtained by the simulation defined in Section 2 for quiet solar wind conditions as illustrated in Sections 3 and 4 (this is a zoom of Fig. 5). In Fig. 13, we also plot in crosses the observations of Potter et al. (1999) reanalyzed by Killen et al. (2001). As can be seen, there is a good agreement at TAA = 213° between the observations of Potter et al. (1999) and our simulation.

In order to explain the large variation observed from TAA = 213° to TAA = 239°, Killen et al. (2001) suggest that the solar wind density changes from 8.8 particles/cm^3 November 13 up to 168.9 particles/cm^3 November 20. In Fig. 13, we plot the total content of the sodium exosphere with respect to TAA (solid circle line) assuming that the solar wind density changes from a density of ~ 60 particles/cm^3 for quiet solar wind conditions (Section 2.3), up to a density of ~ 180 particles/cm^3, seven Earth days before Mercury reaches TAA = 239° (corresponding to the period from November 13, TAA = 217° to November 20th, TAA = 239° in Potter et al. (1999) observations). To model

this case, we use the same maximum surface area impacted by solar wind particles (Fig. 4) with the same half hour periodic variation. The maximum area corresponds roughly to the area invoked by Killen et al. (2001) for November 20th. As shown by the solid line with circles in Fig. 13, the total amount of sodium in Mercury's exosphere does not change with increasing solar wind density with respect to permanent quiet solar wind conditions (solid line with stars). This result can be explained by the depletion of the surface in the region bombarded by solar wind particle and, therefore, the larger flux of solar wind ions does not increase the ejection rate.

The second scenario we considered is a shift of the bombarded surface towards the dawn terminator. This is towards a region not strongly depleted in sodium and colder so that it more efficiently reabsorbs sodium. For quiet solar wind conditions we shift the impact region towards the dawn terminator by 50° in longitude, in order that the maximal bombarded surface reaches the dawn terminator seven Earth days (at TAA = 217°) before Mercury reaches TAA = 239°. Such a shift can occur due to a variation of the IMF leading to a significant change in the open magnetic field line configuration at Mercury (Sarantos et al., 2001; Ip and Kopp, 2002). The result of this simulation is shown in Fig. 13 (solid square line). In contrast to our previous result, we now observe a significant increase of the total amount of sodium in Mercury's exosphere after TAA = 217°. This increase is even larger than the increase observed by Potter et al. (1999). It is immediately followed at TAA = 218° by a rapid decrease of the total content due to the depletion of the bombarded surface. Therefore, the observations of Potter et al. (1999) could be due to a slow shift of the bombarded surface towards the terminator leading to the slow increase in the total content of the exosphere, as observed by Potter et al. (1999). A shift of the bombarded surface is a more efficient way to generate a significant change in Mercury's exosphere than an increase of the solar wind density. Therefore, in some situations solar wind sputtering can be a significant contributor to the sodium exosphere. However, this likely only lasts for short periods of time (less than one Earth day, the time for sodium in the upper layers of the local surface to be partially depleted by sputtering as shown Fig. 13).

Finally, another explanation is suggested by the earlier observations of Potter and Morgan (1997a) showing emission brightness peaks above the limb that move from the northern to the southern hemisphere. These observations suggest that the model for the surface density distribution used in this paper is too simple, as explained in Section 4. If the surface density of the early morning region, which is the principal source of the sodium exosphere, changes significantly on shorter time scales than that found in this paper, it could lead to shorter time variations in the exosphere content, potentially of the size observed by Potter et al. (1999). Changes in the surface density distribution on shorter time scales can be due to variations in the solar wind electric field of convection which results in variations in ion recycling which is non-uniform at the surface (Delcourt et al., 2003). It can also be due to solar wind sputtering variations due to IMF changes, to coronal mass ejection encounters which can induce diffusion from depth in the grains of new sodium atoms, large impacts at the surface, to local spatial inhomogeneity at the surface like the Caloris basin (Sprague et al., 1990b), etc. If there are significant spatial inhomogeneities or time variations, the results provided in this paper have to be understood as the description of the average state of the exosphere/surface layer for one Mercury year. This can then be significantly perturbed on short spatial or time scales.

6. Conclusion

The modeling of Mercury's exosphere has lead to considerable disagreement on the morphology and content of the exosphere and on the dominant desorption processes. Because Mercury rotates very slowly while it orbits the sun and its orbit is eccentric, a comprehensive simulation is required to account for the changes in the sodium content in the surface layers and the transport and re-deposition of sodium. In particular, thermal desorption has been neglected in most studies of Mercury's sodium exosphere (Ip, 1990; Smyth and Marconi, 1995a; Killen et al., 2001) despite its importance as suggested by Sprague et al. (Sprague, 1990a, 1992a, 1992b; Hunten and Sprague, 1997, 2002). In recent laboratory experiments, Yakshinskiy et al. (2000) showed that thermal desorption is a very efficient process for ejecting sodium atoms from a surface at temperatures larger than 400 K. Since Mercury's subsolar temperature varies from 575 to 700 K in going from aphelion to perihelion, thermal desorption can be a very efficient ejection process. This can deplete the sodium content in the surface layers of exposed grains. Although more laboratory data is needed on sodium diffusion in grains, particularly on the role of grain charging, we have shown here that neglecting this process can lead to incorrect estimates of the importance of photon stimulated desorption and solar wind sputtering in the formation of Mercury's sodium exosphere.

In this paper, we carried out the first comprehensive simulation in which we describe self consistently the distribution of sodium atoms in Mercury's exosphere/surface layer by considering the effect and variation of thermal desorption, photon stimulated desorption, solar wind sputtering and micro-meteoroid vaporization as a function of Mercury's heliocentric position. We also examined the effect of suggested sodium supply rates. The calculations here are direct simulations based on reasonable descriptions of the various desorption processes and is not a fit to the observational data.

We showed that significant variations in the surface content of sodium, that is, the sodium available for desorption, can occur in going from night to dayside, from low to high latitudes and from morning to afternoon terminators. The distribution in the surface content of sodium changes with respect to Mercury's heliocentric position and has a signif-

icant influence on the morphology of the exosphere. The importance of the different ejection processes is, therefore, significantly affected by this distribution. We also showed that the depletion of available sodium can be the origin of a strong morning/afternoon asymmetry as reported by Sprague et al. (1997), and that the considerable disagreement in the literature on the principal desorption mechanism is likely due to the temporal and spatial variations in the sodium available for desorption.

Potter et al. (2002) recently published measurements of the sodium exosphere which describe the dependence on heliocentric distance of the sodium column density and average velocity. Our model provides surprisingly good agreement with their measurements especially considering the number of a priori choices made in describing the details of the various desorption processes and the pick-up ion loss rate. This comparison highlights the role of the surface depletion on the observations. If the ejection mechanisms and pick-up ion loss rates are described correctly, then the additional supply rate needed to produce the sodium exosphere observed by Potter et al. (2002) is comparable to the supply rate due to micro-meteoroid vaporization only. A large change in the pick-up ion loss rate could change this balance.

We also suggested an explanation of the origin of the strong enhancement in the emission brightness at northern and southern latitudes observed by Potter and Morgan (1997a) on Mercury's afternoon side. Moreover, when viewing Mercury's morning side, significant North/South maxima are correlated with North/South maxima in the surface Na concentration and do not require solar wind sputtering. The calculated maxima are similar to the observations reported by Potter and Morgan (1990, 1997b) and by Sprague et al. (1998). The strong variation in the total amount of sodium in Mercury's exosphere (Potter et al., 1999) is probably not simply related to an increase of the solar wind density, as suggested by Killen et al. (2001). We suggest that it is more likely related to a shift in the region impacted by solar wind particle towards the dawn terminator or, possibly, to surface density variation beyond that described in this paper.

These simulations of one Mercury year have shown that a number of seemingly contradictory observations can be roughly understood using a comprehensive model which includes a surface reservoir of sodium available for desorption. Although there is considerable work to do in improving the descriptions of the various desorption processes and in determining the rate of supply of fresh sodium to the surface, these simulations also suggest that there is no compelling reason for a fresh sodium supply rate significantly in excess of that believed to be delivered by micro-meteoroid vaporization. Of course, there are additional sources. Fresh grains are exposed by meteoroid bombardment, radiation enhanced diffusion delivers sodium to the grain surface, charging can affect solid state transport of sodium, there are a variety of surface binding sites, and fresh regolith is occasionally formed. Therefore, additional modeling is required. In addition to describing the solid state effects more carefully, a better understanding of the relationship between the solar wind magnetic fields and the morphology of the flux to Mercury's surface is required. However, the results obtained here show that any improvements in the desorption, loss, or source processes has to be evaluated using a comprehensive model of the type described here. Of critical importance in such a model is the role of thermal desorption in depleting the available sodium in the surface layers of exposed grains.

Acknowledgments

The work at Virginia is supported by NASA's Planetary Atmospheres Program and Planetary Geology and Geophysics Program. We thank R.M. Killen, T.E. Madey, and B.V. Yakshinskiy for helpful comments. Figure 11 has been kindly suggested by A.L. Sprague.

References

Brinkman, R.T., 1970. Departures from Jeans' escape rate for H and He in the Earth's atmosphere. Planet. Space Sci. 18, 449–478.

Butler, B.J., 1997. The migration of volatiles on the surface of Mercury and the Moon. J. Geophys. Res. 102, 19283–19291.

Cheng, A.F., Johnson, R.E., Krimigis, S.M., Lanzeroti, L.J., 1987. Magnetosphere, exosphere and surface of Mercury. Icarus 71, 430–440.

Cintala, M.J., 1992. Impact-induced thermal effects in the lunar and mercurian regolith. J. Geophys. Res. 97, 947–973.

Cremonese, G., Boehnhardt, H., Crovisier, J., Rauer, H., Fitzsimmons, A., Fulle, M., Licandro, J., Pollacco, D., Tozzi, G.P., West, R.M., 1997. Neutral sodium from Comet Hale–Bopp: a third type of tail. Astrophys. J. 490, L199–L202.

Delcourt, D.C., Grimald, S., Leblanc, F., Bertherlier, J.-J., Millilo, A., Mura, A., 2003. A quantitative model of planetary Na^+ contribution to Mercury's magnetosphere. Ann. Geophys. In press.

Eichhorn, G., 1978. Heating and vaporization during hypervelocity particle impact. Planet. Space Sci. 26, 463–467.

Flavill, R.P., Carey, W.C., McDonnell, M., Ashworth, D.G., Allison, R.J., 1980. Progress in defining the solar wind sputter rate on protoplanets and interplanetary matter. Planet. Space Sci. 28, 511–524.

Hale, A.S., Hapke, B., 2002. A time-dependent model of radiative and conductive thermal energy transport in planetary regoliths with application to the Moon and Mercury. Icarus 156, 318–334.

Hapke, B., Cassidy, W., 1978. Is the Moon really as smooth as a billiard ball? Remarks concerning recent models of sputter-fractionation on the lunar surface. Geophys. Res. Let. 5, 297–300.

Hunten, D.M., Morgan, T.H., Shemansky, D.E., 1988. The Mercury atmosphere. In: Vilas, F., Chapman, C.R., Matthews, M.S. (Eds.), Mercury. Univ. of Arizona Press, Tucson, pp. 562–612.

Hunten, D.M., Sprague, A.L., 1997. Origin and character of the lunar and mercurian atmospheres. Adv. Space Res. 19, 1551–1560.

Hunten, D.M., Sprague, A.L., 2002. Diurnal variation of sodium and potassium at Mercury. Meteor. Planet. Sci. 37, 1191–1196.

Ip, W.-H., 1987. Dynamics of electrons and heavy ions in Mercury's magnetosphere. Icarus 71, 441–447.

Ip, W.-H., 1990. On solar radiation-driven surface transport of sodium atoms at Mercury. Astrophys. J. 356, 675–681.

Ip, W.-H., 1993. On the surface sputtering effects of magnetospheric charged particles at Mercury. Astrophys. J. 418, 451–456.

Ip, W.-H., Kopp, A., 2002. MHD simulations of the solar wind interaction with Mercury. J. Geophys. Res. 107, 1348. 10.1029/2001JA009171.

Johnson, R.E., 1989. Application of laboratory data to the sputtering of a planetary regolith. Icarus 78, 206–210.

Johnson, R.E., 1990. Energetic Charged-Particle Interactions with Atmospheres and Surfaces. Springer-Verlag, Berlin.

Johnson, R.E., Baragiola, R.A., 1991. Lunar Surface: sputtering and secondary ion mass spectrometry. Geophys. Res. Let. 18, 2169–2178.

Johnson, R.E., Leblanc, F., Yakshinskiy, B.V., Madey, T.E., 2002a. Energy distribution for desorption of sodium and potassium from ice: the Na/K ratio at Europa. Icarus 156, 136–142.

Johnson, R.E., 2002b. Surface boundary layer atmosphere. In: Mendillo, M., Nagy, A., Waite, J.H. (Eds.), Atmosphere in the Solar System: Comparative Aeronomy. In: Geophys. Monograph Ser., Vol. 130, pp. 203–219.

Jurac, S., Baragiola, R.A., Johnson, R.E., Sittler, E.C., 1995. Charging of ice grains by low-energy plasmas: application to Saturn's E-ring. J. Geophys. Res. 100, 14821–14831.

Kabin, K., Gombosi, T.I., DeZeeuw, D.L., Powell, K.G., 2000. Interaction of Mercury with the solar wind. Icarus 143, 397–406.

Killen, R.M., Potter, A.E., Morgan, T.H., 1990. Spatial distribution of sodium vapor in the atmosphere of Mercury. Icarus 85, 145–167.

Killen, R.M., Morgan, T.H., 1993a. Maintaining the Na atmosphere at Mercury. Icarus 101, 293–312.

Killen, R.M., Morgan, T.H., 1993b. Diffusion of Na and K in the uppermost regolith of Mercury. J. Geophys. Res. 98, 23589–23601.

Killen, R.M., Ip, W.-H., 1999. The surface bounded atmospheres of Mercury and the Moon. Rev. Geophys. 37, 361–406.

Killen, R.M., Potter, A.E., Reiff, P., Sarantos, M., Jackson, B.V., Hick, P., Giles, B., 2001. Evidence for space weather at Mercury. J. Geophys. Res. 106, 20509–20525.

Killen, R.M., Sarantos, M., 2003. Space weather at Mercury. Adv. Space Res. In press.

Leblanc, F., Johnson, R.E., Brown, M.E., 2002. Europa's sodium atmosphere: an ocean source? Icarus 159, 132–144.

Leblanc, F., Luhmann, J.G., Johnson, R.E., Liu, M., 2003. Solar energetic particle event at Mercury. Planet. Space Sci. 51, 339–352.

Luhmann, J.G., Russell, C.T., Tsyganenko, N.A., 1998. Disturbances in Mercury's magnetosphere: are the Mariner 10 "substorms" simply driven? J. Geophys. Res. 103, 9113–9119.

Lundin, R., Barabash, S., Brandt, P., Eliasson, L., Naim, C.M.C., Norberg, O., Sandahl, I., 1997. Ion acceleration processes in the hermean and terrestrial magnetospheres. Adv. Space Res. 19, 1593–1607.

Madey, T.E., Yakshinskiy, B.V., Ageev, V.N., Johnson, R.E., 1998. Desorption of alkali atoms and ions from oxide surfaces: relevance to origins of Na and K in atmospheres of Mercury and the Moon. J. Geophys. Res. 103, 5873–5887.

Madey, T.E., Johnson, R.E., Orlando, T.M., 2002. Far-out surface science: radiation-induced surface processes in the Solar System. Surf. Sci. 500, 838–858.

McGrath, M.A., Johnson, R.E., Lanzerotti, L.J., 1986. Sputtering of sodium on the planet Mercury. Nature 323, 694–696.

Morgan, T.H., Zook, H.A., Potter, A.E., 1988. Impact-driven supply of sodium and potassium to the atmosphere of Mercury. Icarus 75, 156–170.

Morgan, T.H., Killen, R.M., 1997. A non-stoichiometric model of the composition of the atmospheres of Mercury and the Moon. Planet. Space Sci. 45, 81–94.

OKeefe, J.D., Ahrens, T.J., 1977. Impact-induced energy partitioning, melting, and vaporization on terrestrial planets. Lunar Planet. Sci. 8, 3357–3374.

Potter, A.E., Morgan, T.H., 1985. Discovery of sodium in the atmosphere of Mercury. Science 229, 651–653.

Potter, A.E., Morgan, T.H., 1990. Evidence for magnetospheric effects on the sodium atmosphere of Mercury. Science 248, 835–838.

Potter, A.E., 1995. Chemical sputtering could produce sodium vapor and ice on Mercury. Geophys. Res. Let. 22, 3289–3292.

Potter, A.E., Morgan, T.H., 1997a. Evidence for suprathermal sodium on Mercury. Adv. Space Res. 19, 1571–1576.

Potter, A.E., Morgan, T.H., 1997b. Sodium and potassium atmospheres of Mercury. Planet. Space Sci. 45, 95–100.

Potter, A.E., Morgan, T.H., Killen, R.M., 1999. Rapid changes in the sodium exosphere of Mercury. Planet. Space Sci. 47, 1441–1449.

Potter, A.E., Killen, R.M., Morgan, T.H., 2000. Variation of lunar sodium during passage of the Moon through the Earth's magnetotail. J. Geophys. Res. 105, 15073–15084.

Potter, A.E., Killen, R.M., Morgan, T.H., 2002. The sodium tail of Mercury. Meteor. Planet. Sci. 37, 1165–1172.

Sarantos, M., Reiff, P.H., Hill, T.W., Killen, R.M., Urquhart, A.L., 2001. A bx-interconnected magnetosphere model for Mercury. Planet. Space Sci. 49, 1629–1635.

Shemansky, D.E., Broadfoot, A.L., 1977. Interaction of the surfaces of the Moon and Mercury with their exospheric atmospheres. Rev. Geophys. 15, 491–499.

Schwenn, R., Marsch, E., 1991. Physic of Inner Heliosphere. Springer-Verlag, Berlin.

Smith, G.R., Shemansky, D.E., Broadfoot, A.L., Wallace, L., 1978. Monte Carlo modeling of exospheric bodies: Mercury. J. Geophys. Res. 83, 3783–3790.

Smyth, W.H., 1986. Nature and variability of Mercury's sodium atmosphere. Nature 323, 696–699.

Smyth, W.H., Marconi, M.L., 1995a. Theoretical overview and modeling of the sodium and potassium atmospheres of Mercury. Astrophys. J. 441, 839–864.

Smyth, W.H., Marconi, M.L., 1995b. Theoretical overview and modeling of the sodium and potassium atmospheres of the Moon. Astrophys. J. 443, 371–392.

Sprague, A.L., 1990a. A diffusion source for sodium and potassium in the atmospheres of Mercury and the Moon. Icarus 84, 93–105.

Sprague, A.L., Kozlowski, R.W.H., Hunten, D.M., 1990b. Caloris Basin: an enhanced source for potassium in Mercury's atmosphere. Science 249, 1140–1143.

Sprague, A.L., 1992a. Mercury's atmospheric sodium bright spots and potassium variations: a possible cause. J. Geophys. Res. 97, 18257–18264.

Sprague, A.L., 1992b. A correction to Mercury's atmospheric sodium bright spots and potassium variations: a possible cause. J. Geophys. Res. 98, 1231.

Sprague, A.L., Kozlowski, R.W.H., Hunten, D.M., 1997. Distribution and abundance of sodium in Mercury's atmosphere, 1985–1988. Icarus 129, 506–527.

Sprague, A.L., Schmitt, W.J., Hill, R.E., 1998. Mercury: sodium atmospheric enhancements, radar bright spots, and visible surface features. Icarus 135, 60–68.

Stern, S.A., 1999. The lunar atmosphere: history, status, current problems and context. Rev. Geophys. 37, 453–491.

Wiens, R.C., Burnett, D.S., Calaway, W.F., Hansen, C.S., Lykkem, K.R., Pellin, M.L., 1997. Sputtering products of sodium sulfate: implications for Io's surface and for sodium bearing molecules in the Io torus. Icarus 128, 386–397.

Wilson, J.K., Smith, M.S., Baumgardner, J., Mendillo, M., 1999. Modeling and enhancement of the lunar sodium tail during the Leonid meteor shower of 1998. Geophys. Res. Let. 26, 1645–1648.

Yakshinskiy, B.V., Madey, T.E., 1999. Photon-stimulated desorption as a substantial source of sodium in the lunar atmosphere. Nature 400, 642–644.

Yakshinskiy, B.V., Madey, T.E., Ageev, V.N., 2000. Thermal desorption of sodium atoms from thin SiO_2 films. Surface Rev. Lett. 7, 75–87.

ANNEXE 4: JOURNAL OF GEOPHYSICAL RESEARCH, 113, A08311, 2008.

LEBLANC F., WITASSE O., LILENSTEN J., R. FRAHM, ALI SAFAENILI, D. BRAIN, J. MOUGINOT, H. NILSSON, Y. FUTAANA, J. HALEKAS, M. HOLMSTROM, J.L. BERTAUX, D. WINNINGHAM, W. KOFMANN AND R. LUNDIN, , OBSERVATIONS OF AURORAE BY SPICAM UVS ON BOARD MARS EXPRESS: SIMULTANEOUS ASPERA-3 AND MARSIS MEASUREMENTS, JOURNAL OF GEOPHYSICAL RESEARCH, 113, A08311, DOI:10.1029/2008JA013033, 2008.

JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 113, A08311, doi:10.1029/2008JA013033, 2008

Observations of aurorae by SPICAM ultraviolet spectrograph on board Mars Express: Simultaneous ASPERA-3 and MARSIS measurements

F. Leblanc,[1,2] O. Witasse,[3] J. Lilensten,[4] R. A. Frahm,[5] Ali Safaenili,[6] D. A. Brain,[7] J. Mouginot,[4] H. Nilsson,[8] Y. Futaana,[8] J. Halekas,[7] M. Holmström,[8] J. L. Bertaux,[1] J. D. Winningham,[5] W. Kofman,[4] and R. Lundin[8]

Received 16 January 2008; revised 12 May 2008; accepted 4 June 2008; published 22 August 2008.

[1] We present a new set of observations of Martian aurorae obtained by Spectroscopy for the Investigation of the Characteristics of the Atmosphere of Mars (SPICAM) on board Mars Express (MEX). Using nadir viewing, several auroral events have been identified on the Martian nightside, all near regions of crustal magnetic fields. For most of these events, two to three consecutive events with variable intensities and separated by a few seconds to several tens of seconds have been observed, whereas simultaneous observations with Mars Advanced Radar for Subsurface and Ionosphere Sounding (MARSIS) and Analyzer of Space Plasma and Energetic Atoms (ASPERA-3) have been possible. In this paper, we present the data set for these events and discuss the possible correlation between the measured UV emission by SPICAM, the measured downward and/or upward flux of electrons by ASPERA-3 and the total electron content recorded by MARSIS. Despite the limited coverage of SPICAM ultraviolet spectrograph (UVS) on the Martian nightside (essentially in regions of high crustal magnetic fields), there is however a very good correlation between the regions with the locally smallest probability to be on closed crustal magnetic field lines, as derived from Mars Global Surveyor/Electron Reflectometer (MGS/MAG-ER), and the position of an aurora event. This suggests that the crustal magnetic fields, when organized into cusp-like structure, can trigger the few aurorae identified by SPICAM UVS. It confirms also the good probability, in the cases where SPICAM UVS measured UV emissions, that the increase in the measured total electron content by MARSIS and the simultaneous measured precipitating electron flux by the ASPERA-3/Electron Spectrometer may be related to each other.

Citation: Leblanc, F., et al. (2008), Observations of aurorae by SPICAM ultraviolet spectrograph on board Mars Express: Simultaneous ASPERA-3 and MARSIS measurements, *J. Geophys. Res.*, *113*, A08311, doi:10.1029/2008JA013033.

1. Introduction

[2] After 1 decade of observations of magnetic fields and electrons near Mars, the Magnetometer and Electron Reflectometer (MAG-ER) on board Mars Global Surveyor (MGS) has completely changed our view of Mars' plasma environment [*Nagy et al.*, 2004]. The discovery of strong crustal magnetic field structures revealed the existence of a past Martian dynamo [*Acuña et al.*, 2001] and strongly affected our understanding of the past history of Mars' atmosphere [*Chassefière et al.*, 2006]. The long-term observation of these magnetic structures has also demonstrated their present influence on Mars' interaction with the solar wind [*Nagy et al.*, 2004]. In particular, electron spike events have been observed in relation to maxima of the crustal magnetic field radial component [*Mitchell et al.*, 2001]. At 400 km, electron flux spikes have been identified as an electron energy distribution of magnetosheath type, measured during short periods (down to 8 s), and usually observed between two plasma voids (defined as regions in which the measured electron flux is below the noise level of MAG-ER/MGS). *Mitchell et al.* [2001] interpreted these measurements as the evidence of a past or present reconnection of the crustal magnetic field lines to the interplanetary magnetic field lines. *Mitchell et al.* [2001] also pointed out the analogy between such a magnetic field structure and Earth's magnetospheric cusp. In the same way as in the Earth's auroral cusps, in region of strong crustal magnetic fields, particles may precipitate into Mars' atmosphere and/or escape from it. This

[1]Service d'Aéronomie, CNRS, IPSL, Verrières-le-Buisson, France.
[2]Currently at Osservatorio Astronomico di Trieste, Trieste, Italy.
[3]Research and Scientific Support Department, ESA, ESTEC, Noordwijk, Netherlands.
[4]Laboratoire de Planétologie de Grenoble, Grenoble, France.
[5]Southwest Research Institute, San Antonio, Texas, USA.
[6]Jet Propulsion Laboratory, Pasadena, California, USA.
[7]Space Sciences Laboratory, University of California, Berkeley, California, USA.
[8]Swedish Institute of Space Physics, Kiruna, Sweden.

0148-0227/08/2008JA013033$09.00

conclusion has been recently confirmed by Analyzer of Space Plasma and Energetic Atoms (ASPERA-3) on board Mars Express (MEX) by *Lundin et al.* [2006] as well as by MAG-ER/MGS [*Brain et al.*, 2006]. In particular, *Lundin et al.* [2006] observed downward energetic electron fluxes with typical terrestrial auroral-type energy distributions in association with energetic upflowing ion fluxes. This points to possible processes of acceleration or heating between the Martian ionosphere and the Martian magnetosheath and magnetotail [*Dubinin et al.*, 2008a].

[3] A downward flux of energetic particles can also lead to auroral emissions of the type observed by Spectroscopy for the Investigation of the Characteristics of the Atmosphere of Mars ultraviolet spectrograph (SPICAM UVS) on Mars Express [*Bertaux et al.*, 2006]. The only observation published so far of such an emission was clearly associated with the strongest crustal magnetic field structure [*Bertaux et al.*, 2005]. This observation has also been recently reanalyzed and has been shown to be most probably produced by primary electrons of few tens eV energy [*Leblanc et al.*, 2006b]. Therefore, either these electrons have been transported from the Martian dayside ionosphere [*Dubinin et al.*, 2008b] or they are nonaccelerated magnetosheath electrons as observed in nightside crustal magnetic field regions by *Mitchell et al.* [2001] and *Liehmon et al.* [2003]. Actually, a recent analysis based on a 3-D magnetohydrodynamic approach of the magnetic field draping around Mars suggested that the most likely electrons which produced the observed UV aurora emission were nonaccelerated magnetosheath electrons rather than photoelectrons [*Liehmon et al.*, 2007].

[4] A downward flux of particles may also locally ionize the Martian atmosphere and may lead to a local increase of the ionospheric electron content. This has been recently observed by *Safaeinili et al.* [2007] using Mars Advanced Radar for Subsurface and Ionosphere Sounding (MARSIS) on board Mars Express. These authors found a significant correlation between the maxima of the total electron content (that is the total number of electrons between Mars Express and the surface in a column density) as measured by MARSIS on the nightside and the presence of a strong radial crustal magnetic field component.

[5] In this paper, we present nine new auroral-type emission events identified within the SPICAM UV observations on the Martian nightside (section 2). Eight of these events have been measured when two other Mars Express instruments were operating: ASPERA-3 and the MARSIS radar sounder. The possible correlation between SPICAM UV observation, ASPERA-3 particle measurements and MARSIS total electron content is therefore discussed in section 3.

2. SPICAM UVS Observations

[6] SPICAM UV spectrograph is part of the SPICAM instrument (also composed of an IR spectrograph covering 1.1 and 1.7 μm) which is dedicated to solar occultation and nadir viewing observations. SPICAM UV spectrograph has a 118–305 nm wavelength range with a 1.5 nm spectral resolution for the light passing through the narrow part of SPICAM slit and a 6 nm spectral resolution with a $\sim$8 times larger sensitivity for the light passing through the large part of the slit [*Bertaux et al.*, 2006]. During each observation by SPICAM UVS, five spatially adjacent parts (that we will call spatial bin) of the CCD are recorded leading to five individual spectra each second (the time for one integration lasting $\sim$0.6 s and a complete reading of these five lines of the CCD, $\sim$0.4 s). Typically, one to two thousands times five individual spectra are recorded during the 20 to 30 min of an observation. Of these five spatial bins, two have been obtained with the best spectral resolution (spatial bins numbered 0 and 1) and two with the worst spectral resolution (spatial bins numbered 3 and 4). The spatial bin numbered 2 cannot be used because it measured light with both spectral resolutions and cannot be properly calibrated. The procedure to subtract offset and nonuniform dark current and to integrate the spectral emission have been described in detail by *Leblanc et al.* [2006a] and will not be repeated here. The geometry of each observation, its UT time, and the solar conditions are given in Table 1.

[7] Following the first observation of an auroral-type emission event in the Martian atmosphere by the SPICAM UVS [*Bertaux et al.*, 2006; *Leblanc et al.*, 2006b], several dedicated observations have been made during periods when the periapsis of Mars Express was on the nightside. Nadir mode viewing has been preferentially selected in order to map the position where these aurorae occur and to obtain coordinated observations with other instruments on board Mars Express. We added two limb viewing observations in Table 1, one during the first identified event recorded during orbit 716 [*Bertaux et al.*, 2005; *Leblanc et al.*, 2006b] and the second during orbit 2800. However, we will not discuss them in this paper. An auroral event is identified when we observe a significant increase of the signal (by more than two standard deviations with respect to the mean signal) simultaneously on (1) the two spatial bins facing the large slit (highest sensitivity) and (2) on the total measured signal (summed over the five spatial bins). In order to avoid the parts of the spectra with low signal/noise ratio we have only considered a spectral region between 150 and 280 nm excluding the edges of the spectral range. The increase must last more than two consecutive integrations in order to be validated. We have tested observations from 66 nightside passes of MEX and found only six orbits with auroral events. It is difficult to deduce from this set of observations any significant statistics since from these 66 cases, only nine limb viewing orbits were obtained when the spacecraft was below 320 km and 21 nadir viewing orbits had a portion below 1000 km. This limit in altitude is imposed by the sensitivity of SPICAM UVS which constrains its capability to detect faint aurora emissions occurring below 200 km in altitude.

[8] Because these observations have been performed with nadir viewing, the typical intensity observed for the main auroral emissions is significantly smaller than during limb viewing as a result of the different path lengths through the region of the emission. This difference is typically of the order of a factor 13 for dayglow observations following *Leblanc et al.* [2006a]. The method used to extract the auroral emission intensities (subtraction of a background associated with the nightglow emission and integration of the emission lines) has been described by *Bertaux et al.* [2005] and *Leblanc et al.* [2006b]. For most observations it has been possible to identify the Cameron band system emission [*Leblanc et al.*, 2006a, b]. The Cameron band

Table 1. Set of Auroral Emission Events[a]

Date MEX Orbit	Ls (°)	F10.7	SW Conditions	SPICAM Viewing	Time Period (UT)	S/C Altitude (km)	Aurora Longitude (°)	Aurora Latitude (°)	Cameron Band Emission (R)	Emission at 289 nm (R)
2004 07 07 591	57.1	54.1	SEP event high SW pressure southward IMF	nadir	0539:52–0539.55	670	214	34S	241 ± 50	N/A
									105 ± 10	N/A
					0541:17–0541:36	586	214	39S	218 ± 34	N/A
									195 ± 10	13 ± 8
2004 08 11 716	72.4	39.4	dawnward - southward IMF	limb	0601:21–0601:28	266	178	50S	2040 ± 132	126 ± 56
									1546 ± 41	157 ± 26
2005 12 27 2515	347.3	38.3	no SEP event moderate to high SW pressure southward IMF	nadir	2147:31–2147:41	954	158	41S	341 ± 74	146 ± 70
									175 ± 16	50 ± 26
					2151:08–2151:16	700	158	52S	588 ± 90	N/A.
									662 ± 24	120 ± 30
2006 01 26 2621	2.4	32.7	no SEP event moderate SW pressure dawnward IMF	nadir	1404:01–1404:15	564	182	40S	414 ± 61	N/A
									274 ± 14	34 ± 15
2006 02 19 2705	13.9	30.0	no SEP event moderate to high SW pressure dawnward–southward IMF	nadir	0237:56–0238:11	387	206	40S	187 ± 36	N/A.
									105 ± 9	27 ± 11
					0238:34–0238:48	365	206	42S	663 ± 51	57 ± 37
									547 ± 16	94 ± 15
					0240:52–0241:02	309	206	52S	273 ± 46	31 ± 30
									214 ± 11	12 ± 8
2006 03 17 2800	25.5	29.3	southward IMF	limb	1709:46–1710:16	314	237	31S	597 ± 41	35 ± 16
									826 ± 17	164 ± 13

[a]Two values of the Cameron band emission and of the 289 nm emission are given for each event. The first value has been derived from the average spectrum of the two spatial bins with the best spectral resolution and the second value using the spatial bins with the worst spectral resolution. Each spatial bin covered 32 pixel rows of the CCD. F10.7 (10^{-22} W/m^2/Hz) values are calculated from daily average National Geophysics Data Center taking into account the relative positions of the Earth and Mars, the Sun rotation, and Mars' heliocentric distance. N/A indicates that the emission intensity cannot be estimated better than one sigma uncertainty. SW conditions: Solar wind conditions as inferred from low altitudes dayside measurement of the magnetic pressure by MGS/MAG-ER [*Crider et al.*, 2003]. SEP is for solar energetic particles [*Brain et al.*, 2006]. IMF direction is accurate at ~90°. S/C is for MEX spacecraft.

emission ($a^3\Pi - X^1\Sigma^+$) from 180 to 260 nm is the most intense emission measured in Mars' dayglow [*Barth et al.*, 1971; *Leblanc et al.*, 2006a] and is thought to be produced by the photo and electron-impact dissociation of the CO_2 atmospheric molecules leading to CO molecule in $a^3\Pi$ excited state. On the Martian dayside, a typical intensity of a few tens of kilorayleigh (kR) has been measured by SPICAM UVS [*Leblanc et al.*, 2006a]. In some cases, it has been possible to extract the 289 nm emission which corresponds to the CO_2^+ ultraviolet doublet ($B^2\Sigma^+ - X^2\Pi$) emission. Typically, the CO_2^+ emission's intensity is 10 kR on the dayside. It is produced by photon and electron-impact ionization of the atmospheric CO_2 molecule leading to the excited state CO_2^+ ($B^2\Sigma^+$). Figure 1 displays an example of auroral spectra measured during orbit 2621. These spectra are noisy with respect to dayglow measurements [*Leblanc et al.*, 2006a] because they are the average of only 28 individual spectra with respect to the 1000 individual spectra used on the dayside. However, in both Figures 1a and 1b, it is possible to clearly identify the main features of the Cameron band system using both types of spectral resolution and also to identify the 289 nm emission in Figure 1b which corresponds to the spatial bin with the lowest spectral resolution but the highest sensitivity. In the aurora spectra, no emissions like the bright dayglow O 130.4 nm emission, Lyman α H 121.6 nm emission, or the CO fourth positive emission [*Leblanc et al.*, 2006a] can be identified. However, the dayside O 130.4 nm and Lyman α H 121.6 nm emissions are essentially produced by resonant scattering of solar photons. Moreover, the CO fourth positive $A^1\Pi - X^1\ \Sigma^+$ emission was present in the aurora spectrum published by *Bertaux et al.* [2005] but because it has a weak spectral signature, its signal/noise ratio is probably too small when measured with SPICAM's nadir field-of-view orientation.

[9] Another difficulty in the identification of the auroral emission has been introduced by the nonnominal behavior of SPICAM UVS during the extended mission of Mars Express. Such behavior is characterized by the sudden increase or decrease during a few seconds of the measured signal on the used and masked pixels (that are pixels which are physically masked from the light and measured only the dark current of the CCD [see *Bertaux et al.*, 2006]). Figure 2 displays an example of such a behavior. As shown in Figure 2a, the signal measured on the masked pixel during the whole sequence of observation of orbit 2621 evolved smoothly showing the typical and nominal variation of the dark current. In contrast, Figure 2b displays the masked signal measured during orbit 2705. As shown, frequent short time variations of the masked signal are observed most of the time before 0234:04 and around 0235:33 and 0237:01. After 0237:20 up to the end of the sequence of observation, the masked signal is almost clear of any event. In the present work, we have considered the period after 0237:20 as nominal using the masked signal as an indicator of any nonnominal behavior. Three aurora events have been identified during orbit 2705 and are all clearly not associated with short-term intense variations of the masked signal.

[10] Figure 3 displays the trajectories of Mars Express during all the nadir viewing observations made when the spacecraft was below 1000 km altitude. These trajectories have been plotted on a map of the radial component of the crustal magnetic field as calculated at 200 km altitude by *Purucker et al.* [2000]. The white crosses represent the positions where the observations of auroral-type event occurred. As shown in Figure 3, all auroral events occurred

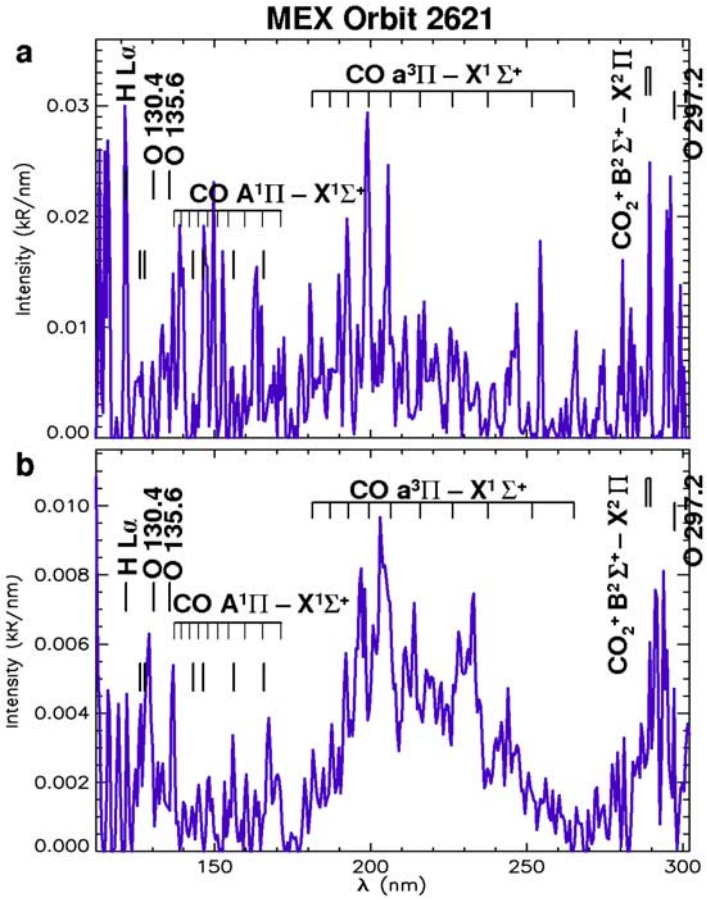

Figure 1. SPICAM ultraviolet spectrograph (UVS) spectra of the aurora event during orbit 2621. (a) As measured with the spatial bins facing the narrow part of the slit (best spectral resolution) during 14 s. (b) As measured with the spatial bins facing the large part of the slit (lowest spectral resolution). Each spectrum is the average of the 28 individual spectra measured during this event on the two spatial bins (14 each) either with the worst spectral resolution or the best one.

in the region of strongest crustal magnetic field (in the Southern Hemisphere centered on a region at 180° in longitude). Actually, during several trajectories over these magnetic structures no observation of auroral events occurred. No auroral events were observed southward of 60° in latitude, even though the spacecraft was below 1000 km altitude at high southern latitudes for some of these passes. There is also no obvious correlation between emission intensity and local time of the auroral event. All observations performed during these 21 orbits have been obtained between 18 h local time and 22 h local time, and the auroral events are uniformly distributed within this local time range. There is no obvious correlation between these observations and the occurrence of solar energetic particle events as inferred from the MGS/MAG-ER data set (Table 1).

3. Simultaneous Measurements by Other Instruments

3.1. ASPERA-3/MEX, MARSIS/MEX, and MAG-ER/MGS Instruments

[11] The main advantages of nadir viewing observations are the accurate location of the aurora on a magnetic field map (Figure 3) and the opportunity to obtain coordinated observations with other instruments on board Mars Express, in particular with ASPERA-3 and MARSIS. ASPERA-3 is composed of two energetic neutral analyzer detectors (Neutral Particle Imager, NPI, and Neutral Particle Detector, NPD), an electron spectrometer (ELS) and an ion mass analyzer (IMA) [*Lundin and Barabash*, 2004]. During the set of observations described in the previous section, only ELS and IMA were operating. The ELS is made up of a collimator followed by a standard spherical top hat electrostatic analyzer with a microchannel plate and anode ring to count electrons. ELS can measure the electron energy distribution within a 360° × 4° field of view (with 16 sectors each of 22.5° × 4°). The energy range is from 0.9 eV to 20 keV with an energy resolution of 8%. A complete energy sweep is performed every 4 s [*Winningham et al.*, 2006]. The IMA energy range is between 0.01 and 30 keV/q with a 5.6° × 360° field of view and a capability of electrostatic sweeping to cover the elevation angle of ±45° in 16 sectors. IMA is a spherical electrostatic analyzer followed by a circular magnetic separation section and a circular microchannel plate imager providing the mass and azimuth of the incident particle. A full 3-D spectrum accumulation sweep of 16 polar angles × 16 azimuthal sectors × 32 mass rings × 96 energies requires 192 s to complete, whereas a 16 azimuthal sectors × 32 mass rings × 96 energies requires 12 s [*Carlsson et al.*, 2006]. ELS is located on a scanning platform which is placed

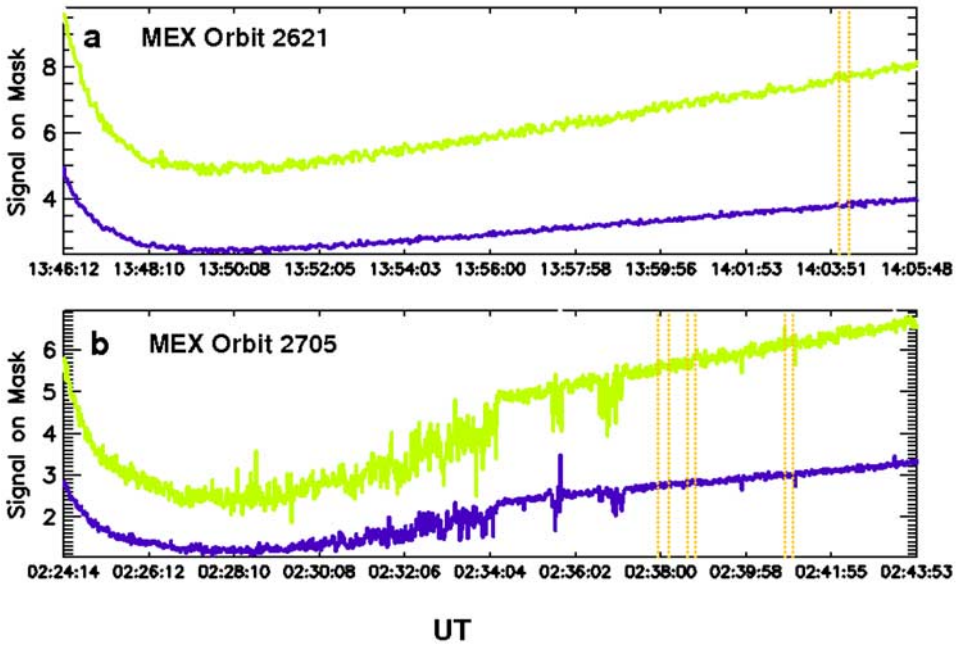

Figure 2. SPICAM UVS measurements of the masked signal during (a) orbit 2621 and (b) orbit 2705. The gray solid line corresponds to the signal measured on the spatial bin numbered 4 (worst spectral resolution), whereas the dark solid line corresponds to the masked signal measured on the spatial bin numbered 0 (best spectral resolution). The dark solid line was shifted by a factor 0.5 for clarity. The dotted vertical lines correspond to the period(s) during which an aurora event has been spectrally identified.

at the edge of a side of the spacecraft where the SPICAM UVS has been mounted (that is the nadir facing side of the spacecraft during the observations discussed here). IMA has been placed at the edge of the opposite side of the spacecraft. As a consequence, in the particular case of SPICAM UVS nadir pointing, part of the field of view of ELS will be covered by the spacecraft when looking for precipitating electrons and IMA will have its field of view covered partly by the spacecraft when looking for outflowing flux.

[12] MARSIS was deployed on MEX in May–June 2005. It is a multifrequency synthetic aperture orbital sounding radar which operates below 800 km and can be used either to sound the ionospheric plasma (Active Ionosphere Sounder mode, AIS) or to sound the subsurface with frequency

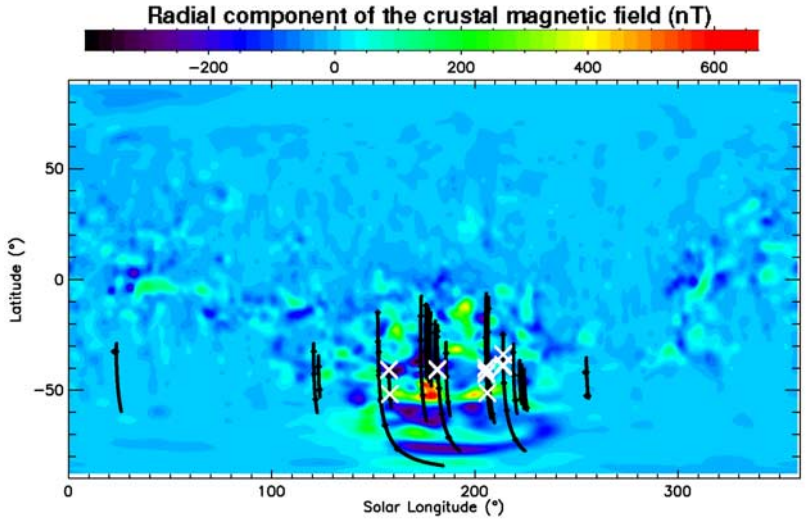

Figure 3. Trajectories of Mars Express during nadir viewing observations by SPICAM UVS. Only the part of each orbit with the spacecraft below 1000 km in altitude has been plotted. These trajectories are plotted on a 2-D map of the radial component of the Martian magnetic crustal field as calculated at 200 km in altitude [*Purucker et al.*, 2000]. The positions of the identified aurora during these nadir viewing orbits are also plotted with white crosses.

between 1.3 and 5.5 Mhz [*Picardi et al.*, 2004]. To sound the subsurface, the signal frequency must be larger than the maximum plasma frequency. In that case, the ionosphere will still disperse the signal phase. The ionospheric phase distortion may then be analyzed to retrieve the total electron content (TEC) of Mars' atmosphere between the altitude spacecraft and the surface [*Mouginot et al.*, 2008]. Such distortion has been intensively analyzed for 750 MEX orbits during their nightside operation and successfully correlated with Mars' crustal magnetic fields [*Safaeinili et al.*, 2007]. *Safaeinili et al.* [2007] found a high concentration of high value of the TEC in regions of high intensity of the crustal radial magnetic field as measured by MGS/MAG-ER. The lateral spatial resolution of MARSIS in subsurface sounding mode is 10 to 30 km in the cross-track direction, the along-track footprint resolution being between 5 and 10 km. A TEC is calculated every 2 s [*Mouginot et al.*, 2008].

[13] On board Mars Global Surveyor, the Electron Reflectometer (ER) and Magnetometer (MAG) instruments were also operating [*Mitchell et al.*, 2001]. The ER instrument consists of a hemispherical imaging electrostatic analyzer followed by a microchannel plate and a resistive imaging anode. An integration lasts between 2 to 48 s. The field of view is $360^\circ \times 14^\circ$ with $22.5^\circ \times 14^\circ$ angular resolution. The energy range is between 10 eV to 20 keV with an energy resolution of 25%. The count rate is accurate within a 10% uncertainty. The MAG instrument provides fast vector measurement (up to 32 samples/s) of the Martian magnetic field and is composed of two redundant triaxial magnetometers. MGS was on a circular orbit at 405 ± 36 km at a fixed local time of 0200/1400 (period of rotation of 1.96 h) during its orbital mapping phase at the time of the present discussed measurements.

3.2. ASPERA-3/MEX, MARSIS/MEX, and MAG-ER/MGS Measurements

[14] Figure 4 displays the different measurements obtained by SPICAM UVS (Figure 4g), MGS/ER (Figure 4f), ASPERA-3/ELS (Figures 4d and 4e), MARSIS (Figure 4c), and the Mars Express altitude (Figure 4a) and latitude (Figure 4b). The track of Mars Express during this orbit was at a local time of 2030 at a longitude of 180° that is above the most intense crustal magnetic field recorded by MGS/MAG. Between 1404:01 and 1404:15 (between the two dotted vertical lines of Figure 4), SPICAM UVS observed a significant increase of the total observed light (Figure 4g). The corresponding spectra measured during this period are displayed (Figure 1) and are clearly composed of the main Cameron band emission features highlighting the occurrence of an auroral event. During the same period (within the 4-s time resolution of ELS), ASPERA-3/ELS measured an increase of the total energy flux (Figure 4e), by one order of magnitude followed by a decrease of this flux by two orders of magnitude. The electron energy distribution is displayed in Figure 5. We only display the signal measured in sectors 1 to 11 of ASPERA-3/ELS which cover a total field of view of $247.5^\circ \times 4^\circ$, because sectors 0, 12, 13, 14, and 15 are partially facing the spacecraft. The energy distribution of Figures 4 and 5 is an isotropic electron distribution exhibiting an atmospheric loss cone with a peak in flux at an energy of 80 eV for all sectors numbered 1 to 11; the exception being sectors 6 and 7. During this orbit, the nadir direction lies between sectors 7 and 8. Therefore sectors 1, 2, and 3 measured downward electron fluxes along the radial direction. Sectors 4 and 5 measured electrons perpendicularly to the radial direction, whereas sectors 6 and 7 and sectors 8 and 9 cover angles between 0° and 45° from the nadir direction. Since ASPERA-3/ELS observations have been made above a strong region of radial crustal magnetic field, it is possible to use the *Cain et al.* [2003] model of the crustal magnetic field to infer the electron pitch angle distribution with respect to the sectors of ASPERA-3/ELS. This model suggests that sectors 1, 2, and 3 measured electrons with pitch angle between 40° and 96°, sectors 4 and 5 between 114° and 131°, sectors 6 and 7 to 155°, and sectors 8 and 9 between 120° and 130°. The angular distribution displayed in Figure 5 may therefore be interpreted as an isotropic electron distribution which is channeled into the atmosphere by the crustal magnetic field lines, interacting with the atmosphere by loosing energy which is going into the generation of the auroral signatures, that is a loss cone electron distribution already reported as associated to Martian aurorae [*Brain et al.*, 2006].

[15] ASPERA-3/IMA did not measure any ion flux during this event (and actually during none of the auroral events identified by SPICAM UVS). It is most probably due to the fact that IMA is mounted on the opposite side of the spacecraft facing the zenith, so that an ion beam coming from below the spacecraft might be partially blocked by the spacecraft. When IMA is in electrostatic scanning mode, the field of view partially covers the nadir direction but only 1/16 of the time, the scanning lasting 192 s (case of orbits 591, 2621, and 2705). For the case with no electrostatic scanning with a 12-s resolution (orbit 2515), IMA should be able to observe ions passing through the spacecraft, unless the path is blocked by the solar panel. Unfortunately, the nadir direction is between two sectors where the sensitivity of IMA is much lower. Therefore, a beam along the nadir direction must have a certain angular width to be observed, at least more than about 10°. According to *Lundin et al.* [2006], the typical ionospheric ions accelerated by a field-aligned electric field observed by ASPERA-3/IMA can have a very narrow angular width so that IMA is rather unlikely to observe these ion beams when SPICAM UVS is pointed toward the nadir direction.

[16] Figure 4c provides the total electron content as deduced from MARSIS subsurface sounding mode during orbit 2621. There is a clear peak of the TEC during the time of SPICAM UVS measurement of an auroral event. It is also interesting to notice that the TEC variation is much broader than in the case of ASPERA-3/ELS or SPICAM/UVS measurements. This broadening is larger than MARSIS spatial/temporal resolution so that it suggests a much larger region of increased electron density along the track of MARSIS. A second peak of the TEC just before the one corresponding to the auroral event can also be seen in Figure 4c (just before 1403:43). Examination of the ASPERA-3 measurement at that time shows a spike of electron flux. Moreover, a small increase of the signal has also been measured by SPICAM UVS during that period (Figure 4g) but is not significant enough to be identified as an auroral event according to our criteria.

[17] During orbit 2621, MGS was on the nightside at 240° longitude in the same range of latitude as MEX. The

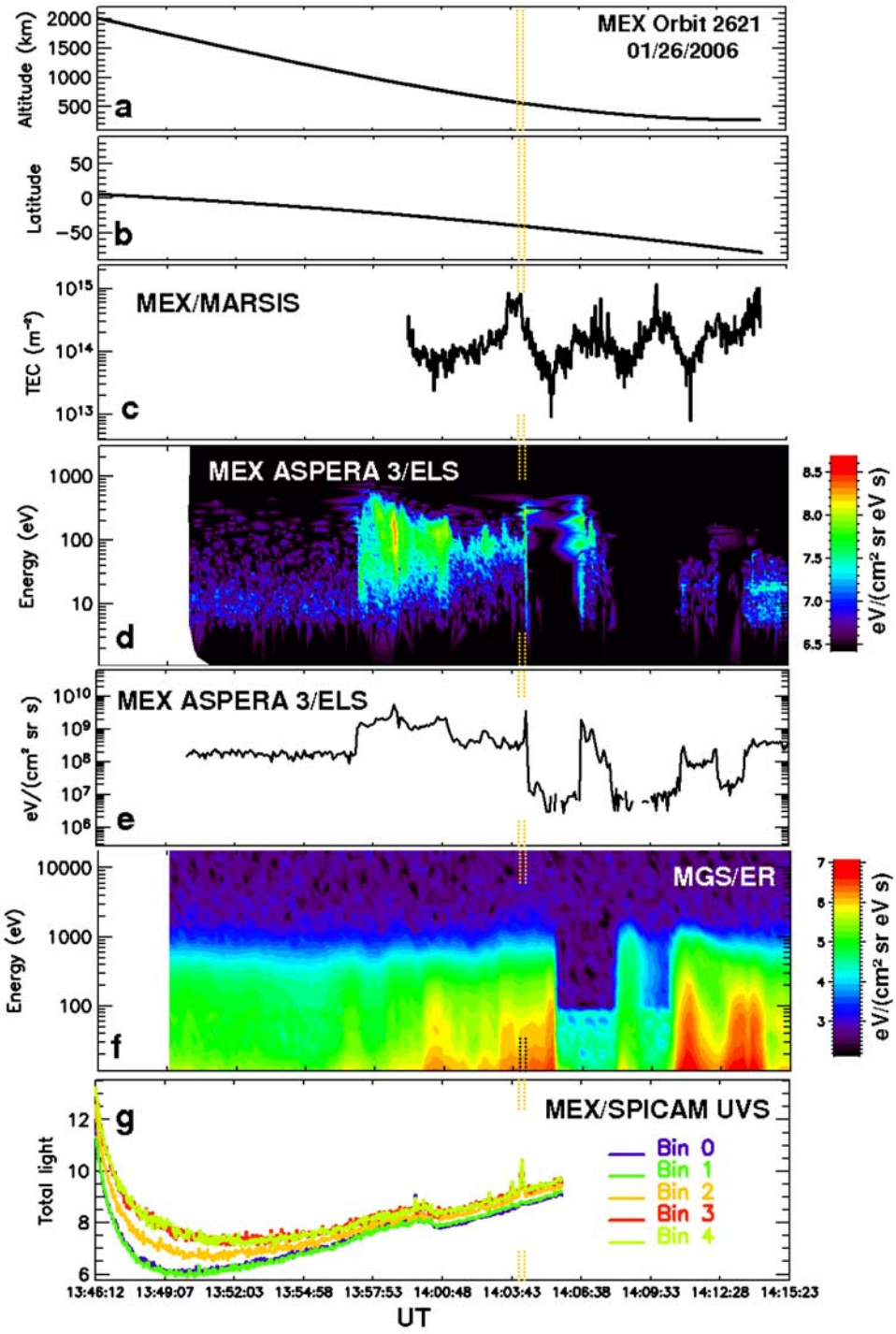

Figure 4. Time series of the measurements during orbit 2621. Shown is the MEX (a) altitude and (b) latitude, (c) MEX/MARSIS Total Electron Content, (d) MEX/ASPERA-3/ELS sector 1 (which is the sector facing the zenith) electron measurement in eV/(cm^2 ster eV s), (e) MEX/ASPERA-3/ELS energy flux of sector 1, (f) MGS/MAG-ER electron energy flux averaged over the full set of sectors in eV/(cm^2 ster eV s), and (g) MEX/SPICAM UVS measurements summed over the full spectral range for each spatial bin in ADU. Indicated on each by vertical dotted lines are the periods during which an aurora event has been identified in SPICAM UV observations.

electron energy spectrum is tail-like, and there may be some hints of photoelectrons at 40 eV and near 500 eV. MGS was passing from a region of field lines unconnected to crustal magnetic fields into a region of open and closed crustal field lines. At the moment of the auroral event, the electron energy distribution measured by MGS/ER is isotropic over the range of pitch angles sampled by MGS/ER. It usually indicates closed magnetic field lines [*Brain et al.*, 2007].

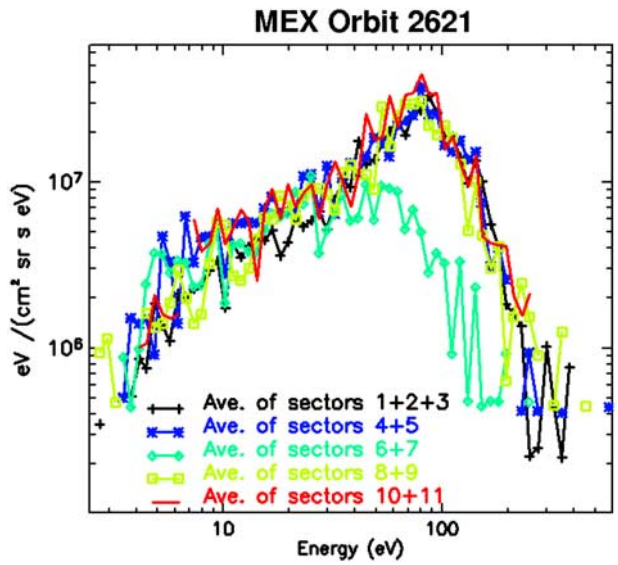

Figure 5. Electron differential energy flux measured by MEX/ASPERA-3/ELS during orbit 2621 between 1404:01 and 1404:15. In order to improve the figure, measurements in adjacent ELS sectors displaying a similar energy distribution have been averaged. Sectors 1 and 9 are 180° apart. No correction of the spacecraft potential has been done. Typical potentials between −4.5 and −5.5 V are estimated from the energy position of the photoelectron peaks during nearby time periods.

[18] In Figure 6, we display the same information as in Figure 4 but during orbit 2705. From the beginning of SPICAM measurements up to 0236 UT a typical non-nominal behavior of the instrument is seen (also seen in Figure 2b). It is followed by a nominal period that follows our criteria (see section 3.1). The dotted vertical lines indicate the times of auroral emissions observed by SPICAM UVS based on spectral identification. Mars Express was at less than 400 km in altitude during the auroral emissions. Figure 6 clearly shows a good temporal correlation (considering the time resolution of each instrument) between SPICAM aurora emissions (Figure 6f), downward electrons measured by ASPERA-3 (Figure 6d, sector 1 facing the zenith direction) and a significant increase of the Total Electron Content of the Martian ionosphere (by a factor 5, 6, and 3 with respect to the background signal in cases 2705a, 2705b, and 2705c, respectively, as displayed in Figure 6c). During this period, MGS was on the dayside passing over the fringes of a crustal magnetic field region near the North Pole. MGS did not show any high instrument background (no solar event) but recorded a dawnward IMF and a normal solar pressure. The measurements during orbit 2705 are particularly interesting since the SPICAM UV spectrograph observed two events, 2705a and 2705b, separated by 40 s (that is by approximately 200 km apart) whereas ASPERA-3 (in the 4-s time resolution mode) observed at 387 and 362 km an almost continuous downward flux of electron with similar energy distributions (see Figures 6d and 6e). At the time of the SPICAM UV observations (2705a and 2705b cases), Mars Express was above a crustal magnetic field anomaly with a 200–400 nT radial component at 200 km in altitude. Following the *Cain et al.* [2003] model of the crustal magnetic field, the sector 1 displayed Figure 6d measured electrons with pitch angle around 60° during events 2705a and 2705b, whereas this sector 1 measured electrons with pitch angle around 150° during event 2705c. In another way, from events 2705a and 2705b to event 2705c Mars Express crossed radial crustal magnetic field lines upward for the two first events and downward for event 2705c. The TEC derived from MARSIS measurements displays a somewhat similar configuration than that of orbit 2621. Two peaks in the TEC are observed almost simultaneously to SPICAM UVS auroral emission observations. These two peaks are seen in the MARSIS data at the top of a much broader peak centered between the two SPICAM UV aurora emissions, suggesting that MARSIS observed an extended region of large electron content.

3.3. Potential Correlation Between SPICAM UVS, ASPERA-3/ELS, and MARSIS

[19] In Table 2, we summarized the different parameters measured during each auroral event identified by SPICAM UVS with a nadir viewing. There is no obvious relation between impacting energy flux as measured by ASPERA-3/ELS and the observed Cameron band and 289 nm doublet emission intensities. However, the brightest Cameron band emission is associated with the largest incident energy flux (case 2515b). The other cases, namely 591a and b, 2515a, 2621, and 2705a, b, and c, do not show any particular relation between intensity of the incident electron flux and intensity of the emissions. This can be due to the variable altitude of the spacecraft (as an example in the case of 2515a the spacecraft was at an altitude of 940 km). The spacecraft potential may also have changed significantly the incident electron flux, particularly for the electron distribution peaking at low energy (591b, 2621a, and 2705c). Moreover, the total zenith observed emission is in a large part produced by secondary electrons of lower energy. This is particularly true for the Cameron band emission since its cross section peaks around 80 eV [*Avakyan et al.*, 1998]. The production of this secondary electron as well as their altitude of production depends significantly on the initial distribution in energy of the incident electrons. It is therefore not surprising to find a significant difference between cases 2621a and 2705a as an example and to a lesser extent between cases 2705a and 2705b where electron energy distributions as measured by ASPERA-3/ELS significantly differ from one another. Moreover, the transport of primary and secondary electrons within the Martian atmosphere, key parameter in constraining the integrated emission, will depend significantly on the magnetic field's configuration [*Liehmon et al.*, 2003]. In another way, this lack of relation between intensity of the emission and intensity of the incident energy flux may be simply due to fact that the measured ASPERA-3/ELS electron population is not representative of the precipitating population at the origin of the auroral emission. Several explanations are possible. For example, this is true if ASPERA-3/ELS' field of view did not cover the downward part of the electron distribution, if the electrons did not move radially between the spacecraft altitude and the Martian atmosphere, or if the electron distribution changed significantly from the spacecraft altitude to the Martian atmosphere (acceleration, backscattering...).

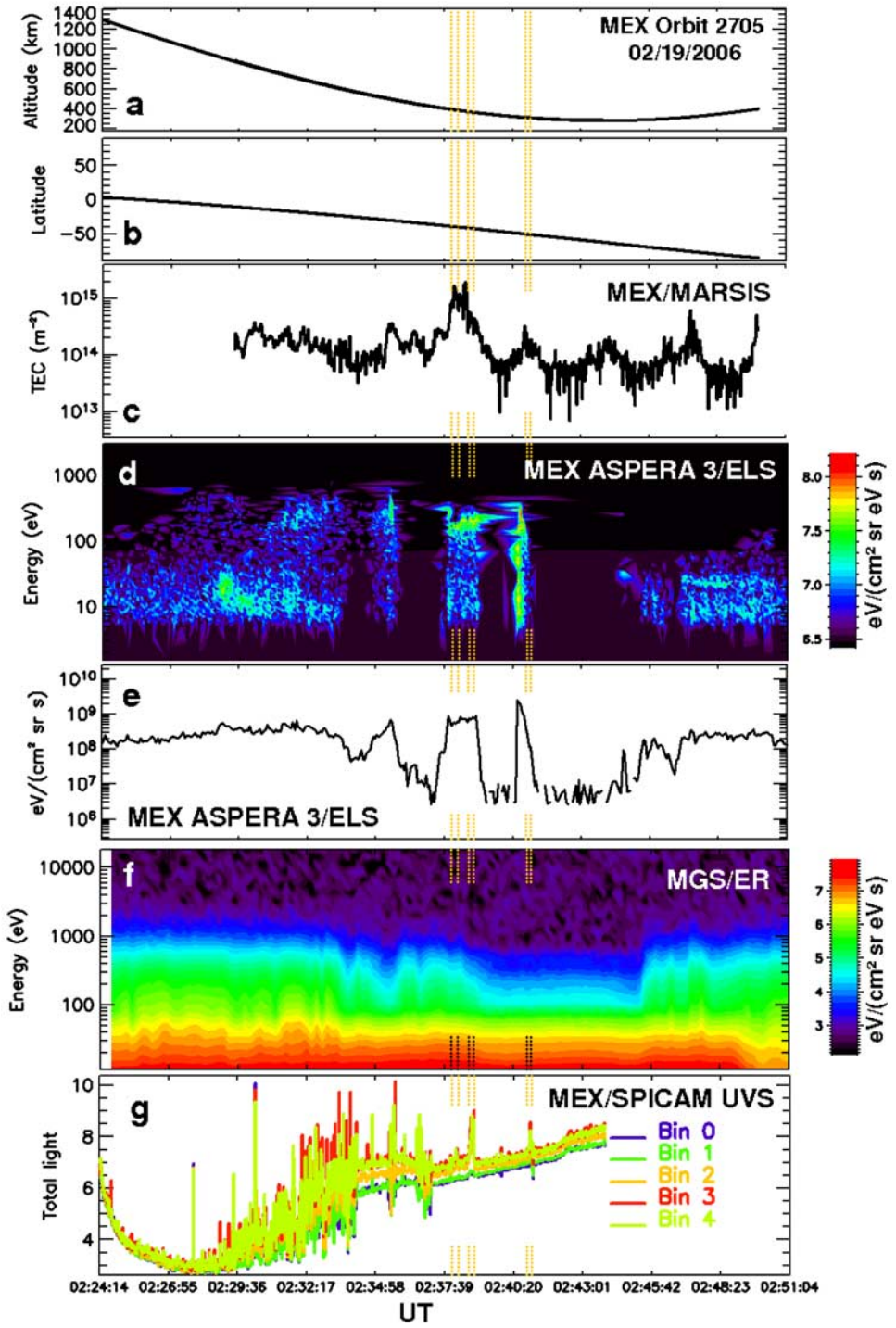

Figure 6. Time series of the measurements during orbit 2705. Shown is the MEX (a) altitude and (b) latitude, (c) MEX/MARSIS Total Electron Content, (d) MEX/ASPERA-3/ELS sector 1 (which is the sector facing the zenith) electron measurement in eV/(cm^2 ster eV s), (e) MEX/ASPERA-3/ELS energy flux of sector 1, (f) MGS/MAG-ER electron energy flux averaged over the full set of sectors in eV/(cm^2 ster eV s), and (g) MEX/SPICAM UVS measurements summed over the full spectral range for each spatial bin in ADU. Indicated on each panel by vertical dotted lines are the periods during which an aurora event has been identified in SPICAM UV observations.

[20] There is clearly, as shown in Figures 4 and 6, a high temporal correlation between the increase of the TEC as measured by MARSIS and the occurrence of an auroral event detected by SPICAM UVS. *Safaeinili et al.* [2007] have suggested that on the nightside, the observed peaks in the TECs with a high correlation with crustal magnetic field lines might indicate the local ionization of the Martian atmosphere where particles precipitate along these field lines. The convolution of

Table 2. Nadir Viewing Measured Aurora Events by SPICAM UVS/MEX[a]

Date MEX Orbit	Label	S/C Altitude (km)	Cameron Band Emission (R)	Emission at 289 nm (R)	Cameron Band/289 nm emissions	Electron Energy Flux 10^9 eV/(cm^2 sr s)	Peak in Energy (eV)	TEC Value (10^{15} m^{-2})
2004 07 07 591	591a	670	105 ± 10	N/A.	N/A.	4.2	∼350	N/A
	591b	586	195 ± 10	13 ± 8	15 ± 10	1.3	∼50	N/A
2005 12 27 2515	2515a	954	175 ± 16	50 ± 26	3.5 ± 2	4.1	∼190	N/A
	2515b	700	662 ± 24	120 ± 30	5.5 ± 1.6	7.8	∼150	0.27
2006 01 26 2621	2621a	564	274 ± 14	34 ± 15	8 ± 4	1.4	∼80	4.5
2006 02 19 2705	2705a	387	105 ± 9	27 ± 11	4 ± 2	2.2	∼160	11.6
	2705b	365	547 ± 16	94 ± 15	5.8 ± 1	2.3	∼200	3.8
	2705c	309	214 ± 11	12 ± 8	18 ± 13	0.4	∼40	0.9

[a]The Cameron band and 289 nm emissions are the values given in Table 1 for the lowest spectral resolution that is for the highest sensitivity of SPICAM UVS. The electron energy flux values are the integrated energy flux measured in sector 1 of ASPERA-3/ELS during the period of the aurora event. The peak of energy is the energy at which the electron flux distribution reached its maximum in sector 1. The TEC is the integrated total electron content during the period of the aurora event as measured by MARSIS. N/A indicates either that the emission was not estimated at better than one sigma uncertainty or that the TEC was not available because MARSIS was not deployed at that time or because MEX was too high in altitude.

the ionization cross sections of the main neutral species of the Martian atmosphere (essentially CO_2, the other species being minor) with the electron distribution measured by ASPERA-3/ELS therefore provide a first-order estimate of the total ionization rate induced by such an impacting flux. However, we did not find any evidence of a relation between measured TEC values and measured energy electron flux during the auroral events. As an example the largest TEC value from the 2705a event corresponds to a value of the electron energy flux close to the 2705b or 2621a events, whereas the measured TEC value is 3 times smaller during these two latter events. In contrast, the smallest TEC value (event 2515b) corresponds to the largest electron energy flux measured by ASPERA-3/ELS for this set of aurora events. This apparent lack of correlation between measured TEC and impacting electron fluxes may have three explanations. The first one is that the electrons measured by ASPERA-3/ELS did not precipitate directly below the S/C where the TEC is measured. The second explanation could be that there is no real direct relation between the ionization rate associated with the precipitating energy electron flux and the value of the atmospheric electron content. The third explanation is related to the difficulty in retrieving TEC values from the distortion in the MARSIS signal generated when the variation of the TEC is relatively short in time as in the cases of auroral events. Thus, *Mouginot et al.* [2008] estimated the nighttime mean sensitivity on a TEC determination as being equal to 1.5×10^{-14} m^{-2}, which is significantly large when compared to the values of the TEC given in Table 2, column 9.

[21] Another way to look for a potential link between ASPERA-3/ELS measurements and SPICAM UVS observations is to calculate the ratio, when available, between the measured intensity of the Cameron band emission and the intensity of the doublet at 289 nm. As explained by *Leblanc et al.* [2006b], this ratio is highly dependent on the energy of the precipitating electrons that may produce the observed emissions. Indeed the cross section of the production of the Cameron band emission peaks at an electron energy of around 80 eV, whereas the cross section for the production of the 289 nm doublet peaks at 150 eV [*LeClair and McConkey*, 1994]. In Table 2, column 7, we have calculated the ratios between the Cameron band emission and the 289 nm doublet emission. There is an evident relation between the large value of this ratio (between 8 and 18) and the electron distribution peaking below 80 eV (column 8), and between the small value of this ratio (between 3 and 6) and the electron distribution peaking above 150 eV. This result suggests that the particles at the origin of the measured emission probably have an energy distribution close to the energy distribution measured at higher altitudes, even if the energy flux intensity has probably changed from the altitude of ASPERA-3/ELS down to the regions of production of the observed glow.

3.4. Localization and Occurrence of UV Auroral Emissions

[22] Figure 7 is a zoom of the region where the aurora events have been observed by SPICAM UVS plotted on a map of the probability to be on a region of closed field lines at 400 km altitude as calculated by *Brain et al.* [2007]. This map has been built using 6 years of MGS/ER data set during the mapping orbit (between 1 July 1999 and 1 July 2005) on a nearly circular orbit of 405 ± 36 km at 0200/1400 local time. More than 20 million pitch angle electron distributions measured by MGS/ER on the nightside and in the energy channel 95–148 eV have been grouped into 27 different types of pitch angle distributions. In particular, regions with two-side loss cone electron distributions or no detection of electron flux (plasma voids region) have been considered regions of closed crustal magnetic field lines. The map displayed in Figure 7 is the probability to find such a type of distribution in MGS/ER data set with respect to planetary longitude and latitude. Similar maps for the open field line regions have also been built but are not used here because as explained by *Brain et al.* [2007], the association between electron pitch angle distributions and open field lines is more ambiguous. As shown in Figure 7, most of the auroral emissions identified by SPICAM UVS were detected in regions of local minimum of the probability to be on close field line regions, that is, in magnetic field regions that have cusp-like structure.

[23] *Lundin et al.* [2006] observed that the auroral-type electron energy distributions are usually observed at the boundary between open and closed field lines in good analogy with the Earth's auroral zone, so that auroral emission should preferentially occur at such a boundary. Auroral emissions should therefore be observed with a higher probability when the pass of the spacecraft crosses a boundary from closed to open field lines and vice versa. In some cases, we observed two consecutive and spatially close aurora events that may be associated to the passing of SPICAM UVS FOV over an open field line region bounded

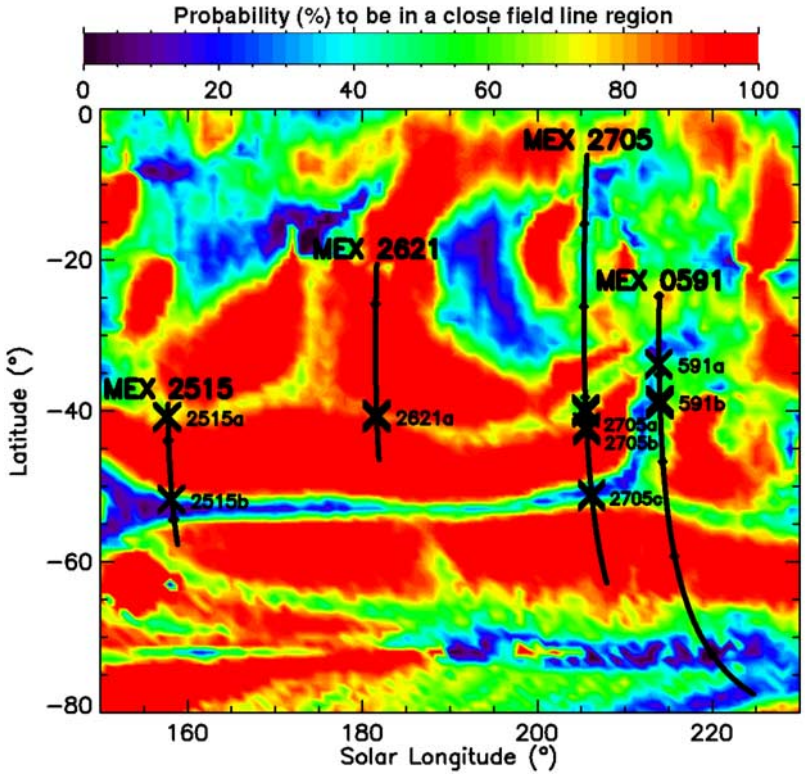

Figure 7. Map of the probability (expressed in percentage) to be in a closed field line region at 400 km in altitude on the Martian nightside as calculated from the Electron Reflectometer on board Mars Global Surveyor (MGS) during its mapping phase orbit. Electron pitch angle distributions have been recorded by the MGS magnetometer/electron reflectometer every 2 to 8 s during the spacecraft mapping orbit phase at around 400 km. Pitch angle distributions recorded in a single instrument energy channel (115 eV, a channel typically uncontaminated by photoelectrons) have been classified according to their shape. Also plotted are the trajectories of Mars Express (black line) below 1000 km altitude for the four orbits with aurora events. The crosses indicate the position of the aurora events identified in SPICAM UVS data. The spacecraft was moving from Northern to Southern Hemispheres.

by two closed crustal magnetic field line regions (e.g., passing into and out of a crustal cusp-like field structure). This may be what SPICAM UVS has observed in the case of orbit 2705 for events 2705a and 2705b (as also suggested by MARSIS measurements, Figure 6), and probably for the case of orbit 2621 as discussed previously (see Figure 4, MARSIS measurements) and the case of orbit 591 as shown in Figure 7. In cases 2515a, 2515b, and 2705c, no other auroral emission within ±20 s has been recorded even if Mars Express was above a large region of open field lines and relatively low in altitude (Figure 7). It may suggest that the current system was differently organized in these aurora regions. It may be also, as underlined by *Dubinin et al.* [2008b], that the map inferred from MGS/ER data obtained at a local time of 0200/1400 might be significantly different at a local time of 2030 (MEX local time during the considered orbits). As an example, *Dubinin et al.* [2008b] underlined the lack of any statistically significant enhancement of electron precipitation between latitude −50° and −55° in the same longitudinal range as measured by ASPERA-3/MEX, suggesting that the extended region of open field lines of Figure 7 might be significantly different at such a local time.

[24] On the other hand, as shown in Figure 3, several orbits occurred over the crustal magnetic field regions with no aurora event signature identified in SPICAM UVS data. Moreover, some of these orbits passed over regions where we positively identified an aurora event during previous or following orbits. As an example during orbit 2694, around 0031 UT, three days before orbit 2705, no aurora event was detected. Actually, precipitation of electrons into the Martian atmosphere has been shown by E. M. Dubinin et al. (Access of solar wind electrons into the Martian magnetosphere, submitted to *Annales Geophysicae*, 2008) and *Brain et al.* [2006] to depend strongly on the solar wind conditions. In particular, the interplanetary magnetic field orientation inducing a southward electric field direction should strongly favor the Southern Hemisphere access to the solar wind electrons (in good agreement with the dawnward direction of the IMF as inferred during orbit 2621 and 2705, see Table 1, column 4). During orbit 2694, the IMF, as determined half an hour before and 1 h after was close to a duskward direction (see http://sprg.ssl.berkeley.edu/∼brain/proxies/drapingdirxn.html). Such a dependency may explain the lack of auroral events observed during a MEX pass over the crustal magnetic field region but not the lack of consecutive events as discussed previously. Dubinin et al. (submitted manuscript, 2008) have built a map of the electron precipitation with energy spanning between 40 and 80 eV as measured by Mars Express/ASPERA-3/ELS between 250 and 600 km in altitude (Dubinin et al., submitted

manuscript, 2008, Figure 13c) in the highest crustal magnetic field regions. These authors have found that the largest flux of electrons occurred below −60° in latitude, whereas SPICAM UVS did not observe any auroral event below such latitude. Such an apparent discrepancy might be simply due to the poor statistic of SPICAM UVS detection.

4. Conclusions

[25] SPICAM UVS have identified several new auroral events for which it has been possible to obtain coordinated measurements with MARSIS and ASPERA-3/ELS. In order to avoid any ambiguity on the positions of the measurements, we used nadir configurations. This new set of observations shows quite strong coincidences between the occurrence of energetic precipitating electrons into the Martian atmosphere, the presence of crustal magnetic field anomalies and auroral-type glow. In some of these cases, there is a good probability that simultaneous measurements of the precipitating electron flux, of the induced UV aurora emission and of the produced total electronic content have been obtained leading to unprecedented information on the Martian aurora. Following the definition of *Brain et al.* [2007] of open/closed magnetic field lines, we observe that the few aurora detected by SPICAM UVS occur essentially in the region where the probability to be on closed field lines is small. Because the regions covered by SPICAM UVS are essentially where the crustal magnetic fields are present, we cannot firmly conclude on the role of crustal magnetic fields in triggering the Martian aurora. However, the positions of the identified aurorae in regions with the locally smallest probability to be on closed field lines suggests a significant relation between aurora events at Mars and the presence of cusp-like magnetic field line structures in good analogy with the Earth's auroral regions [*Lundin et al.*, 2006].

[26] **Acknowledgments.** Wolfgang Baumjohann thanks the reviewers for their assistance in evaluating this paper.

References

Acuña, M. H., et al. (2001), Magnetic field of Mars: Summary of results from the aerobraking and mapping orbits, *J. Geophys. Res.*, *106*, 23,403–23,417, doi:10.1029/2000JE001404.

Avakyan, S. V., et al. (1998), *Collision Processes and Excitation of UV Emission From Planetary Atmospheric Gases: A Handbook of Cross Sections*, Gordon and Breach, Amsterdam.

Barth, C. A., C. W. Hord, J. B. Pearce, K. K. Kelly, G. P. Anderson, and A. I. Stewart (1971), Mariner 6 and 7 Ultraviolet Spectrometer Experiment: Upper atmosphere data, *J. Geophys. Res.*, *76*, 2213–2227, doi:10.1029/JA076i010p02213.

Bertaux, J.-L., F. Leblanc, O. Witasse, E. Quémerais, J. Lilensten, A. S. Stern, B. Sandel, and O. Korablev (2005), Discovery of aurora on Mars, *Nature*, *435*, 790–794, doi:10.1038/nature03603.

Bertaux, J. L., et al. (2006), SPICAM on Mars Express: Observing modes and overview of UV spectrometer data and scientific results, *J. Geophys. Res.*, *111*, E10S90, doi:10.1029/2006JE002690.

Brain, D. A., J. S. Halekas, L. M. Peticolas, R. P. Lin, J. G. Luhmann, D. L. Mitchell, G. T. Delory, S. W. Bougher, M. H. Acuña, and H. Rème (2006), On the origin of aurorae on Mars, *Geophys. Res. Lett.*, *33*, L01201, doi:10.1029/2005GL024782.

Brain, D. A., R. J. Lillis, D. L. Mitchell, J. S. Halekas, and R. P. Lin (2007), Electron pitch angle distributions as indicators of magnetic field topology near Mars, *J. Geophys. Res.*, *112*, A09201, doi:10.1029/2007JA012435.

Cain, J. C., B. R. Ferguson, and D. Mozzoni (2003), An $n = 90$ internal potential function of the Martian crustal magnetic field, *J. Geophys. Res.*, *108*(E2), 5008, doi:10.1029/2000JE001487.

Carlsson, E., et al. (2006), Mass composition of the escape plasma at Mars, *Icarus*, *182*, 320–328, doi:10.1016/j.icarus.2005.09.020.

Chassefière, E., F. Leblanc, and B. Langlais (2006), The combined effects of escape and magnetic field histories at Mars, *Planet. Space Sci.*, *55*, 343–357, doi:10.1016/j.pss.2006.02.003.

Crider, D. H., D. Vignes, A. M. Krymskii, T. K. Breus, N. F. Ness, D. L. Mitchell, J. A. Slavin, and M. Acuña (2003), A proxy for determining solar wind dynamic pressure at Mars using Mars Global Surveyor data, *J. Geophys. Res.*, *108*(A12), 1461, doi:10.1029/2003JA009875.

Dubinin, E., G. Chanteur, M. Fraenz, and J. Woch (2008a), Field-aligned currents and parallel electric field potential drops at Mars: Scaling from the Earth's aurora, *Planet. Space Sci.*, *56*, 868–872.

Dubinin, E., M. Fraenz, J. Woch, J. D. Winningham, R. Frahm, R. Lundin, and S. Barabash (2008b), Suprathermal electron fluxes on the nightside of mars. ASPERA-3 observations, *Planet. Space Sci.*, *56*, 846–851.

Leblanc, F., J. Y. Chaufray, O. Witasse, J. Lilensten, and J.-L. Bertaux (2006a), The Martian dayglow as seen by SPICAM UV spectrometer on Mars Express, *J. Geophys. Res.*, *111*, E09S11, doi:10.1029/2005JE002664.

Leblanc, F., O. Witasse, J. Winningham, D. Brain, J. Lilensten, P.-L. Blelly, R. A. Frahm, J. S. Halekas, and J. L. Bertaux (2006b), Origins of the Martian aurora observed by Spectroscopy for Investigation of Characteristics of the Atmosphere of Mars (SPICAM) on board Mars Express, *J. Geophys. Res.*, *111*, A09313, doi:10.1029/2006JA011763.

LeClair, L. R., and J. W. McConkey (1994), (^{1}S) and CO($a^3\Pi$) production from electron impact dissociation of CO_2, *J. Phys. B At. Mol. Phys.*, *27*, 4039–4055.

Liehmon, M. W., D. L. Mitchell, A. F. Nagy, J. L. Fox, T. W. Reimer, and Y. Ma (2003), Comparisons of electron fluxes measured in the crustal fields at Mars by the MGS magnetometer/electron reflectometer instrument with a B field-dependent transport code, *J. Geophys. Res.*, *108*(E12), 5134, doi:10.1029/2003JE002158.

Liehmon, M. W., Y. Ma, A. F. Nagy, J. U. Kozyra, J. D. Winningham, R. A. Frahm, J. R. Sharber, S. Barabash, and R. Lundin (2007), Numerical modelling of the magnetic topology near Mars aurora observations, *Geophys. Res. Lett.*, *34*, L24202, doi:10.1029/2007GL031806.

Lundin, R., and S. Barabash (2004), Evolution of the Martian atmosphere and hydrosphere: Solar wind erosion studied by ASPERA-3 on Mars Express, *Planet. Space Sci.*, *52*, 1059–1071, doi:10.1016/j.pss.2004.07.020.

Lundin, R., et al. (2006), Plasma acceleration above Martian magnetic anomalies, *Science*, *311*, 980–983, doi:10.1126/science.1122071.

Mitchell, D. L., R. P. Lin, C. Mazelle, H. Réme, P. A. Cloutier, J. E. P. Connerney, M. H. Acuña, and N. F. Ness (2001), Probing Mars' crustal magnetic field and ionosphere with the MGS Electron Reflectometer, *J. Geophys. Res.*, *106*, 23,419–23,427, doi:10.1029/2000JE001435.

Mouginot, J., W. Kofman, A. Safaeinili, and A. Herique (2008), Correction of the ionospheric distortion on the MARSIS surface sounding echoes, *Planet. Space Sci.*, *56*, 917–926.

Nagy, A. F., et al. (2004), The plasma environment of Mars, *Space Sci. Rev.*, *111*, 33–114, doi:10.1023/B:SPAC.0000032718.47512.92.

Picardi, G., et al. (2004), Mars Express: A European mission to the red planet, *Eur. Space Agency Spec. Publ.*, *ESA SP 1240*, 51–70.

Purucker, M., D. Ravat, H. Frey, C. Voorhies, T. Sabaka, and M. Acuña (2000), An altitude-normalized magnetic map of Mars and its interpretation, *Geophys. Res. Lett.*, *27*, 2449–2452, doi:10.1029/2000GL000072.

Safaeinili, A., W. Kofman, J. Mouginot, Y. Gim, A. Herique, A. B. Ivanov, J. J. Plaut, and G. Picardi (2007), Estimation of the total electron content of the Martian ionosphere using radar sounder surface echoes, *Geophys. Res. Lett.*, *34*, L23204, doi:10.1029/2007GL032154.

Winningham, J. D., et al. (2006), Electron oscillations in the induced Martian magnetosphere, *Icarus*, *182*, 360–370.

J. L. Bertaux, Service d'Aéronomie, CNRS, IPSL, F-91371 Verrières-le-Buisson, France.

D. A. Brain and J. Halekas, Space Sciences Laboratory, University of California, Berkeley, CA 94720, USA.

R. A. Frahm and J. D. Winningham, Southwest Research Institute, San Antonio, TX 78228-0510, USA.

Y. Futaana, M. Holmström, R. Lundin, and H. Nilsson, Swedish Institute of Space Physics, Box 812, S-98 128, Kiruna, Sweden.

W. Kofman, J. Lilensten, and J. Mouginot, Laboratoire de Planétologie de Grenoble, F-38041 Grenoble CEDEX 9, France.

F. Leblanc, Osservatorio Astronomico di Trieste, Via Tieppolo 11, I-34131, Trieste, Italy. (francois.leblanc@aerov.jussieu.fr)

A. Safaenili, Jet Propulsion Laboratory, Pasadena, CA 91109, USA.

O. Witasse, Research and Scientific Support Department, ESA, ESTEC, NL-2200 AG Noordwijk, Netherlands.

ANNEXE 5: GEOPHYSICAL RESEARCH LETTERS, 36, L07201, 2009.

LEBLANC F., A. DORESSOUNDIRAM, N. SCHNEIDER, S. MASSETTI, M. WEDLUND, A. LÓPEZ ARISTE, C. BARBIERI, V. MANGANO, AND G. CREMONESE, SHORT TERM VARIATIONS OF MERCURY'S NA EXOSPHERE OBSERVED WITH VERY HIGH SPECTRAL RESOLUTION, GEOPHYS. RES. LETT., 36, L07201, DOI:10.1029/2009GL038089, 2009.

GEOPHYSICAL RESEARCH LETTERS, VOL. 36, L07201, doi:10.1029/2009GL038089, 2009

Short-term variations of Mercury's Na exosphere observed with very high spectral resolution

F. Leblanc,[1] A. Doressoundiram,[2] N. Schneider,[3] S. Massetti,[4] M. Wedlund,[1] A. López Ariste,[5] C. Barbieri,[6] V. Mangano,[4] and G. Cremonese[7]

Received 9 March 2009; revised 9 March 2009; accepted 20 March 2009; published 15 April 2009.

[1] Short time variations of Mercury's exosphere cannot be tracked easily from ground based observatories because of the difficulty of distinguishing them from Earth atmospheric effects. On July 13th 2008, using THEMIS solar telescope, we were able to simultaneously measure brightness, Doppler shift and width of the exospheric sodium D_2 emission line during half a day with a resolving power of ~370,000. Mercury's exosphere displayed an emission brightness peak in the Northern hemisphere which vanished in few hours and a more persistent Southern Hemispheric peak. The bulk Doppler shift of the exosphere suggests a period of strong escape from Mercury. The global changes of the Doppler shift and of the Doppler width suggest that a cloud of sodium atoms ejected before or at the beginning of our sequence of observations passed through THEMIS field of view moving anti-sunward. A preferentially southern ejection of sodium atoms leading to the observed persistent southern emission peak is consistent with the orientation of the Interplanetary Magnetic Field during that period. **Citation:** Leblanc, F., A. Doressoundiram, N. Schneider, S. Massetti, M. Wedlund, A. López Ariste, C. Barbieri, V. Mangano, and G. Cremonese (2009), Short-term variations of Mercury's Na exosphere observed with very high spectral resolution, *Geophys. Res. Lett.*, *36*, L07201, doi:10.1029/2009GL038089.

1. Introduction

[2] Mercury's sodium exosphere was observed by *Potter and Morgan* [1990] to be variable on timescales of an Earth day (that is ~10 minutes of Mercury solar time) with high latitude peaks in emission brightness appearing and disappearing in few Earth days. *Potter et al.* [2006] suggested that the solar wind might play a key role in producing such observed localized peaks and their short time variability. The reconnection between the interplanetary magnetic field (IMF) lines and Mercury's magnetic field lines would allow the solar wind to penetrate through Mercury's magnetospheric cusps [*Massetti et al.*, 2007]. The solar wind particles would then impact Mercury's surface inducing the sputtering of surface Na atoms leading to localized increases of the sodium exospheric density at high latitudes.

[3] Following *Massetti et al.* [2007], the solar wind density and velocity control the efficiency of the solar wind sputtering whereas the IMF orientation controls the size and position of the region where the solar wind particles impact Mercury's surface. When the radial component of the IMF is sunward and is the dominant component, the solar wind particles are predicted to preferentially impact the Southern hemisphere in a region whose size and position are driven by the two other components of the IMF. An opposite situation (anti-sunward IMF radial component) was encountered during MESSENGER first flyby with a more intense northern sodium emission than southern [*McClintock et al.*, 2008].

[4] The solar wind and IMF conditions changing on time scales much shorter than one Earth day, the exospheric signatures of the solar wind sputtering at Mercury should also change within few Earth hours. One of the problems with tracking short time variations from Earth is that, in few hours, Mercury is observed through various atmospheric conditions which are a significant potential source of misinterpretation.

[5] In this paper, we discuss the first simultaneous observations of Mercury's sodium brightness, Doppler shift and width during more than 10 continuous hours thanks to the very high resolution of THEMIS solar telescope and its capability to image Mercury's exosphere throughout daylight. Section 2 provides the information on these observations and the analysis performed, and sections 3 and 4 report discussion and conclusions, respectively.

2. Observations and Analysis

[6] THEMIS [*López Ariste et al.*, 2000] is a French-Italian solar telescope on the Canary Island of Tenerife with a 0.9 m primary mirror (with a central obscuration of 0.4 m) and a 15.04 m focal length. The slit size was 0.25" × 69.6" and the spectral resolution of 15.9 mÅ provided ~370,000 resolving power. One camera is used to measure the D_2 at 5890 Å Na emission line covering a spectral range of ~4 Å and is composed of 512 by 512 pixels at 7.8 mÅ/pixel spectral dispersion. The observation was obtained on the 13th July 2008 between 06:19 and 18:22 UT (08:19–20:22 local time). Mercury's true anomaly angle (TAA) was between 308.8 and 311.5°, the phase angle was between 67.3° and 65.4° and Mercury's radius (R_M) was 3.05". We were facing the dusk side of Mercury and were seeing 70% of Mercury's illuminated disk. Mercury's heliocentric distance was 0.33 AU and its heliocentric radial velocity was between −7.8 and −7.5 km/s (towards the Sun). The slit was oriented along

[1]LATMOS/IPSL, Université Versailles Saint Quentin, CNRS, Verrières-le-Buisson, France.
[2]LESIA, Observatoire de Paris, Meudon, France.
[3]LASP, University of Colorado, Boulder, Colorado, USA.
[4]IFSI, INAF, Rome, Italy.
[5]THEMIS, UPS853, CNRS, La Laguna, Spain.
[6]Department of Astronomy, University of Padova, Padova, Italy.
[7]Osservatorio Astronomico di Padova, INAF, Padova, Italy.

0094-8276/09/2009GL038089$05.00

Table 1. Subsequent Scans of July 13th 2008[a]

Scan n°	Time (UT, LT+2h)	Type of Scan	Seeing (″)	Average Emission (kR)	Average Doppler Shift (km/s)	Average Doppler Width (km/s)	North/South Hemispheric Emissions Ratio
3	06:19–06:50	HR, 9 positions separated by 0.5″	3.50 ± 0.28	N/A	−1.07 ± 0.02	2.90 ± 0.31	0.93
4	06:53–08:05	HR, 21 positions separated by 0.25″	1.50 ± 0.58	2.81 ± 0.28	−0.98 ± 0.02	2.79 ± 0.28	0.94
8	08:16–09:28	HR, 21 positions separated by 0.25″	1.33 ± 0.38	2.74 ± 0.15	−1.18 ± 0.01	2.84 ± 0.03	0.90
10	09:33–10:45	HR, 21 positions separated by 0.25″	1.50 ± 0.36	2.61 ± 0.15	−1.04 ± 0.01	2.87 ± 0.25	0.91
12	10:50–12:02	HR, 21 positions separated by 0.25″	1.67 ± 0.34	3.08 ± 0.13	−0.85 ± 0.01	2.78 ± 0.12	0.90
13	12:05–13:31	HR, 25 positions separated by 0.25″	1.83 ± 0.63	2.61 ± 0.32	−0.89 ± 0.01	2.74 ± 0.12	0.93
16	13:38–15:03	HR, 25 positions separated by 0.25″	1.67 ± 0.38	<2.60 ± 0.15	−0.89 ± 0.01	2.84 ± 0.27	0.78
17	15:05–16:31	HR, 25 positions separated by 0.25″	2.33 ± 0.47	2.73 ± 0.25	−0.97 ± 0.01	2.87 ± 0.27	0.81
21	16:55–17:38	LR, 12 positions separated by 0.5″	1.50 ± 0.42	3.28 ± 0.19	−0.98 ± 0.04	2.38 ± 0.45	0.82
22	17:39–18:22	LR, 12 positions separated by 0.5″	2.16 ± 0.53	3.11 ± 0.30	−1.05 ± 0.04	2.51 ± 0.45	0.75

[a]The average emission has been calculated by summing all the ADU measured within all the pixels of the scan, then by multiplying by the angular size of a pixel and then by dividing by the angular size of Mercury apparent disk. The average Doppler shift and width are calculated by fitting by a Gaussian distribution the sum of all spectra measured within all pixels with signal/noise ratio larger than 150 (Figure 3). Uncertainty on the Doppler shift and width (one standard deviation from Gaussian fit) should be considered as relative uncertainty. HR and LR are for high and low resolution respectively. Scan 16 did not cover the whole exosphere so that the average emission should be considered as a lower limit. For scan 3, coverage and seeing precluded a good estimate of the calibration. The method to calculate uncertainties on emission brightness and seeing is described by *Leblanc et al.* [2008].

Mercury's North/South axis and Mercury's exosphere was scanned with the slit automatically moved between each position in a direction perpendicular to the slit. Because the sky conditions were deteriorating, the two last scans were performed with lower resolution (220,000 and 27 mÅ spectral resolution) with a slit of 0.5" × 118" size. Details on each scan are provided in Table 1.

[7] The telescope provided tip-tilt corrections at ∼1 kiloHertz. At each slit position, ten individual exposures of twenty seconds were taken with negligible overhead for CCD readout and slit motions. The data were bias corrected and flat fielded. The flat field was obtained by observing the Sun using a special mode avoiding solar bright or dark spots. Spectral calibration was carefully performed, during the whole sequence of observation, using solar spectra obtained at four different times. The sky background was calculated from two segments at each end of the slit interpolated over the whole slit by fitting these two parts with a second degree polynomial. In order to subtract the reflected solar spectrum from Mercury's surface, we used the solar spectrum obtained for the closest atmospheric terrestrial conditions (similar air mass and zenith angle), shifted in wavelength, and scaled to the measurements. The exospheric emission line is then integrated subtracting an average background level calculated outside the emission line. We then fitted the emission line with a Gaussian function and derived the Doppler shift and the spectral full width at half maximum (FWHM) of the emission line after correction by the effect of the point spread function of the instrument. The Doppler shift of Mercury's exospheric line in Mercury's frame was determined using JPL/Horizon ephemerides. We estimated the uncertainty on the Doppler shift as ∼0.25 km/s. The brightness calibration is based on the Hapke theoretical model of the reflected solar light from Mercury's surface, which has the advantage of avoiding any uncertainty due to Earth's atmospheric absorption [*Sprague et al.*, 1997]. We have also developed a method to evaluate the uncertainty on the absolute calibration due to the uncertainty of position of the slit on Mercury's illuminated disk [*Leblanc et al.*, 2008].

3. Results

[8] Figure 1 shows the measured emission brightness during six of the best scans described in Table 1. The quality of the images is validated by our estimate of the seeing value and by the image of the continuum [*Leblanc et al.*, 2008]. As shown in Table 1, 4th column, these scans were obtained for comparable atmospheric conditions. As displayed in Figure 1, there are distinct Northern and Southern exospheric peaks structure during scans 4 and 8, followed by a progressive disappearance of the northern peak in less than five hours, whereas the southern peak is still evident up to scan 22 (more than 11 hours later, see also Table 1 last column). The average emission brightness intensity (Table 1: fifth column) displays minima during scans 10 and 13–16, a significant maximum during scan 12 and an increase from scan 17 up to the end of the observation.

[9] In Figure 2, the spatial distribution of the Doppler shift associated to each plot of Figure 1 is shown. In Table 1 (6th column) we also provide the Doppler shift of the sum

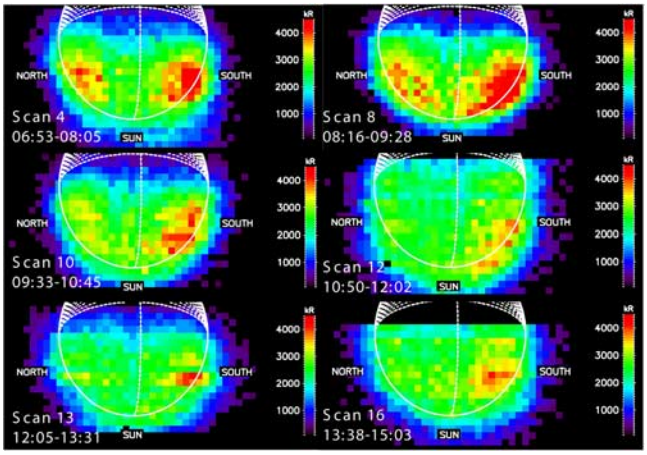

Figure 1. Emission brightness during the six consecutive scans on the 2008/07/13. Mercury's disk is plotted. The white dashed lines are longitudes on the nightside region. The scan of Mercury is done from bottom to top of each panel. Only pixels where the signal/noise ratio was greater than 150 are plotted.

of all pixels for each scan (Figure 3). This Doppler shift is always negative and displays minima at scan 8 and at the end of the observation and a maximum at scan 12. A positive Doppler shift corresponds to atoms moving away from the observer and mainly anti-sunward. The Doppler shift as seen from Earth depends on the angle between the line of sight and the vector normal to the surface, the sodium atoms being ejected preferentially perpendicularly to the surface. Since sodium atoms are also pushed anti-sunward by the solar radiation pressure, such a Doppler shift will also depend on the angle between the line of sight and the solar zenith angle. For quiet conditions, we should observe a globally negative Doppler shift of the atoms along the line of sight with maxima at the solar limb and in the anti-solar direction (bottom and top parts of each panel) and

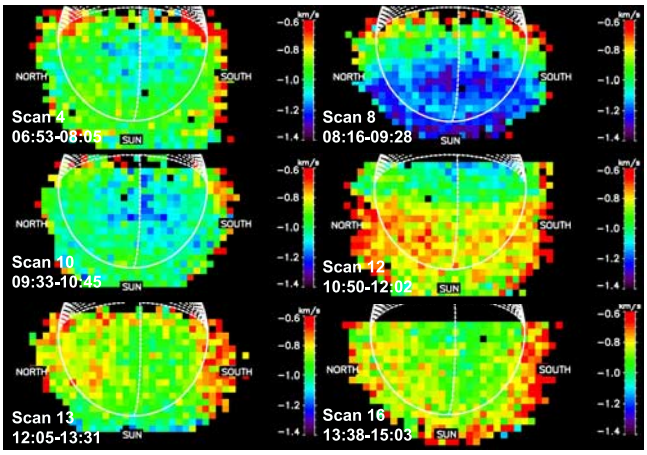

Figure 2. Doppler shift of the Na D_2 emission line in Mercury's frame. A positive Doppler shift means that the Na atoms move on average away from the observer whereas a negative Doppler corresponds to Na atoms moving towards the observer. Only pixels where the signal/noise ratio was greater than 150 are plotted.

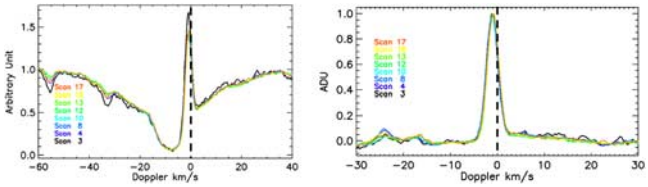

Figure 3. Spectra of the measured signal obtained by summing all pixels with signal/noise ratio larger than 150. (left) The sky contribution to the measured spectra was subtracted but not the solar reflected light on Mercury's surface. (right) Same as Figure 3 (left) but with the solar reflected light on Mercury's surface subtracted. Each spectrum is plotted in Mercury's exospheric frame (zero Doppler shift, vertical dashed line).

a minimum at the sub-Earth point. Such a minimum of the Doppler shift at the center of the apparent disk is observed in Figure 2, for scans 4, 8 and 10 as well as for scans 13 to 17. A globally negative Doppler shift (Figure 3) implies also that a significant part of the observed exosphere was not dominantly gravitationally bounded to Mercury. In the contrary, we should have observed a globally null Doppler shift in Mercury's frame.

[10] *Leblanc et al.* [2008] observed an enlargement of the Doppler spectral line in association with peak of emission brightness and interpreted this as the signature of an energetic process of ejection. The Doppler width distribution during the 6 scans displayed Figures 1 and 2 is shown in Figure 4. We observe a slightly hotter southern hemisphere than the northern during the whole sequence of observation. Two episodes of increase of the Doppler width appear also to occur during scans 8 and 10 and during scans 16 and 17 in Figure 4 (also evident in the average Doppler width displayed 7th column of Table 1). Therefore, during these two periods of maximum Doppler width, it is probable that solar wind sputtering increased leading to a local increase of the Doppler width and to the Doppler shift minima during scan 8 and less evidently during scans 16–17. The association between Doppler width and peak of emission is not as clear for these observations as for the one reported by *Leblanc et al.* [2008]. This can be explained by worse seeing conditions, Mercury being also closer to the Earth in 2007, but also as a signature of the weakness of the energetic processes during our observation. Moreover, most of the atoms producing the southern and northern brightness peaks were probably ejected before the start of our observation, so that they either already moved away from the region of ejection for the most energetic particles as suggested by Figure 4 or partially thermalized by reimpacting the surface.

[11] Scan 12 is the only scan which does not display a clear minimum of the Doppler shift at the center of the apparent disk. We interpret it as the path of another cloud of sodium moving anti-sunward, with positive Doppler shift as suggested by the peak of the average brightness, Figure 1, during scan 12 but without brighter high latitude peaks. Such atoms may have been ejected before scan 10, as suggested by the increase of the Doppler width. The rate of ejection towards the observer between scans 4 and 10, should have peaked during scan 8, as suggested by the contrast between the emission brightness of the high latitude

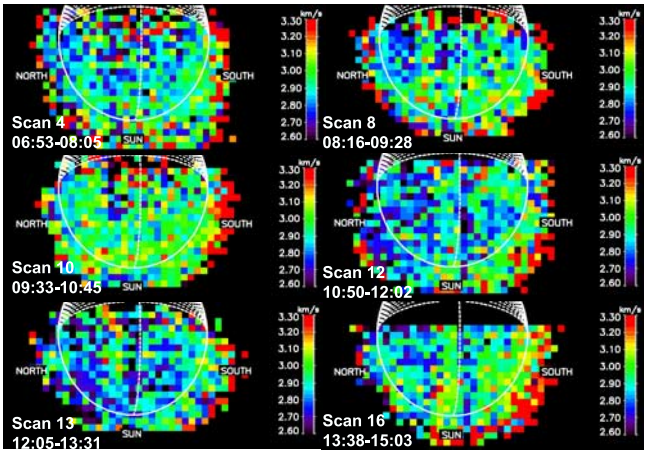

Figure 4. Doppler Width of the spectral Na D_2 emission line. Only pixels where the signal/noise ratio was greater than 150 are plotted.

peaks with the rest of the exosphere. An episode of strong ejection during scan 8 is also suggested by the minimum of the Doppler shift close to the subsolar region. The minimum of average emission brightness during scan 10 (Table 1, 5th column) could have been due to those atoms ejected before scan 10 being slowed down by solar radiation pressure (around -180 cm s^{-2} at TAA = 310° [*Smyth and Marconi*, 1995]) down to a null velocity with respect to the Sun (and therefore passing through a minimum of the number of solar photons a sodium atom could scatter) before moving anti-sunward. The minimum of the Doppler width during scan 13 would be then the signature of a quiet period during which energetic ejection decreased. If f_1 represents the total number of atoms of the bulk exosphere with a Doppler shift V_1 and f_2 the atoms in the cloud moving with a Doppler shift of V_2, then the average measured Doppler shift observed during scan 12 would be equal to $V = (V_1 \times f_1 + V_2 \times f_2)/(f_1 + f_2)$ if the exosphere is assumed to be optically thin. V_1 is the average Doppler shift measured when no second population is present, that is $V_1 \sim -0.9$ km/s (scan 13). The cloud population when ejected from the surface should have a velocity of the order of -1.2 km/s (scan 8) and should reimpact the surface with a similar velocity, that is, $V_2 \sim +1.2$ km/s, implying, $f_1 \sim 41 \times f_2$ ($V = -0.85$ km/s for scan 12). Therefore, apart from any consideration of scattering efficiency, the increase of exospheric atoms needed to produce the Doppler shift signature observed during scan 12 (f_2) should represent only few percent of the total exospheric sodium atomic population (f_1) in good agreement with the total emission brightness variation (Table 1, 5th column). A similar episode of ejection apparently occurred after scan 16 (from the Doppler width increase) but will be not discussed here because the conditions of observation worsened after scan 16.

[12] In summary, the presence of a persistent peak in emission brightness in the southern hemisphere during more than 11 hours seems to be due to at least two consecutive events of solar wind sputtering and preferentially in the Southern hemisphere. The Northern peak remained visible from Earth for a few hours which suggests that after being initiated by a dramatic increase of the ejection rate, probably before the start of our observation (if not we would have seen a signature in the Doppler width during scan 4), the peak in exospheric density progressively fades away during several hours being partially maintained by recycling and induced diffusion in the surface. In the contrary, the apparently more persistent southern peak might be explained by one or more subsequent episodes of preferentially southern ejection as indicated by the Doppler width signatures. These episodes of ejection were probably weak as suggested by the lack of significant variation of the total emission brightness.

[13] It is possible to infer the variation and orientation of the IMF during the period of our observations by using Advanced Composition Explorer measurements (ACE, which was 20 Earth radii above the equatorial plane) and the Wilcox Solar Observatory solar data (WSO). At that time, WSO indicates that both Earth and Mercury were above the equatorial plane by 4° and 2.5° respectively. The correction due to the Earth–Sun–Mercury angle as well as to the propagation time of the solar wind between the Earth and Mercury was evaluated using the interplanetary discontinuity linked to the equatorial coronal hole recorded by the SoHO EIT instrument as a marker. Mercury on the July 13th appears to have encountered a period of variable sign of the radial magnetic field component but with long period of sunward and strong values. As stated in the introduction, a strong sunward radial component should induce a preferentially bombardment of the Southern hemisphere as supported by our observation.

4. Conclusions

[14] Eight consecutive images of the brightness, Doppler shift and width of Mercury's exospheric D_2 emission line have been obtained during almost eleven hours by THEMIS solar telescope using very high resolving power (~370,000). Two high latitude peaks in emission brightness were visible at the beginning of our sequence of observation. The Northern hemispheric peak vanished during the first five hours of observation whereas the Southern peak lasted during the whole sequence. In the same time, the Doppler shift changed significantly as well as the Doppler width. These observations suggest that a strong ejection event occurred before or at the beginning of our sequence of observations producing in particular the Northern hemispheric peak. Such an event leads to the formation of a cloud of sodium atoms probably initially ejected towards the Sun, then slowed down and accelerated in the anti-sunward direction by the solar radiation pressure. This first event of ejection was then followed probably by a second episode of increased ejection rate preferentially from the Southern hemisphere.

[15] *Baumgardner et al.* [2008] and S. Okano et al. (personal communication, 2008) have recently observed Mercury's sodium tail and reported the presence of localized peak of density along the tail that they interpreted as potentially short time variation in Mercury's exosphere. We here show that indeed Mercury's exosphere significantly change in less than a few Earth hours.

References

Baumgardner, J., J. Wilson, and M. Mendillo (2008), Imaging the sources and full extent of the sodium tail of the planet Mercury, *Geophys. Res. Lett.*, *35*, L03201, doi:10.1029/2007GL032337.

Leblanc, F., A. Doressoundiram, N. Schneider, V. Mangano, A. López Ariste, C. Lemen, B. Gelly, C. Barbieri, and G. Cremonese (2008), High latitude peaks in Mercury's sodium exosphere: Spectral signature using THEMIS solar telescope, *Geophys. Res. Lett.*, *35*, L18204, doi:10.1029/2008GL035322.

López Ariste, A., J. Rayrole, and M. Semel (2000), First results from THEMIS spectropolarimetric mode, *Astron. Astrophys.*, *142*, 137–148.

Massetti, S., S. Orsini, A. Milillo, and A. Mura (2007), Modelling Mercury's magnetosphere and plasma entry through the dayside magnetopause, *Planet. Space Sci.*, *55*, 1557–1568.

McClintock, W. E., et al. (2008), Mercury's exosphere: Observations during MESSENGER's first Mercury flyby, *Science*, *321*, 92–94.

Potter, A. E., and T. H. Morgan (1990), Evidence for magnetospheric effects on the sodium atmosphere of Mercury, *Science*, *248*, 835–838.

Potter, A. E., R. M. Killen, and M. Sarantos (2006), Spatial distribution of sodium on Mercury, *Icarus*, *181*, 1–12.

Smyth, W. H., and M. L. Marconi (1995), Theoretical overview and modeling of the sodium and potassium atmospheres of Mercury, *Astrophys. J.*, *441*, 839–864.

Sprague, A. L., R. W. H. Kozlowski, D. M. Hunten, N. M. Schneider, D. L. Domingue, W. K. Wells, W. Schmitt, and U. Fink (1997), Distribution and abundance of sodium in Mercury's atmosphere, 1985–1988, *Icarus*, *129*, 506–527.

C. Barbieri, Department of Astronomy, University of Padova, Vicolo Osservatorio 2, I-35122 Padova, Italy.

G. Cremonese, Osservatorio Astronomico di Padova, INAF, Vicolo Dell'Osservatorio 5, I-35122 Padova, Italy.

A. Doressoundiram, LESIA, Observatoire de Paris, 5, place Jules Janssen, F-92195 Meudon CEDEX, France.

F. Leblanc and M. Wedlund, LATMOS/IPSL, Université Versailles Saint Quentin, CNRS, Réduit de Verrières BP 3, F-91371 Verrières-le-Buisson CEDEX, France. (francois.leblanc@latmos.ipsl.fr)

A. López Ariste, THEMIS, UPS853, CNRS, C/Via Lactea s/n, E-38200 La Laguna, Spain.

V. Mangano and S. Massetti, IFSI, INAF, Via del Fosso del Cavaliere, 100, I-00133 Roma, Italy.

N. Schneider, LASP, University of Colorado, Campus Box 392, Boulder, CO 80309-0392, USA.