Understanding plasma sources

A. Ganguli and R. D. Tarey[†]

*Centre for Energy Studies, and [†]Department of Physics, Indian Institute of Technology Delhi, New Delhi 110 016, India

The paper discusses the basic physical mechanisms behind different plasma discharges. An attempt is made to integrate the analysis by considering the salient changes brought about in discharge formation and plasma behaviour, as the frequency of the applied power is raised from dc to microwave frequencies. Following this, discharge systems commonly used in plasma sources are examined and discussed. These include the capacitively and inductively coupled rf discharges, helicon discharges, the microwave, electron cyclotron resonance and surface wave discharges.

Introduction

The last few decades have seen a rapid growth in the use of plasmas for various applications. The demands on the quality and properties of materials, and the increasing need for materials with novel properties are some factors that have given impetus to the induction of plasmas and plasma-based technologies into industry. Today, one has a whole gamut of industrial applications where the use of plasmas-based technologies offers distinct advantages over other more conventional technologies. Some of these applications may be listed as: etching and deposition (in semiconductor wafer processing), deposition and polymerization (in fabrication of optical fibres, treatment of fibrous materials, etc.), surface modification (by deposition of diamond films, corrosion resistance coatings, etc.), plasma immersed ion implantation (for plasma nitriding, etc.), flat plate displays, high-pressure arcs and jets (for ceramic coating and plasma pyrolysis applications, etc.)[1–7].

The transition to plasma-based technologies mentioned above has been aided to a great extent by the knowledge and expertise gained from our attempts to understand and control the high-temperature plasmas required for attaining thermonuclear fusion in the laboratory[8–10]. In particular, these attempts have given us novel methods for producing plasmas with widely different characteristics, some of which (as listed above) are already finding suitable applications. All this has led to the concept of dedicated *plasma sources* that are specially designed plasma discharge systems, for the production of plasmas with specific end-uses in the industry. Different types of discharges and discharge configurations are used in these sources. This paper analyses and discusses the important physical mechanisms that go to explain the discharges commonly used in plasma sources. Only low-pressure, non-equilibrium plasmas are considered. (In such plasmas, the electrons and ions do not interact too frequently and thermal equilibrium between these species is absent. Thus electrons and ions usually have very different temperatures in such plasmas.)

The paper has two sections. The first deals with breakdown and other related aspects of dc and rf/microwave discharges. Also discussed in detail are the qualitative changes that take place in the formation mechanisms and the characteristics of the plasmas, as the frequency of the applied power source is raised from dc to the microwave range. The second section discusses some discharge systems that find frequent application in plasma sources. These cover the different types of rf discharges (including the helicon discharge) together with the microwave, electron cyclotron resonance (ECR) and surface wave discharges.

Basic discharge mechanisms

DC discharge

Breakdown mechanism: A schematic of the set-up used for obtaining a dc discharge is shown in Figure 1. The operation of the discharge in steady state can be understood as follows[1,11]. Imagine an electron leaving the cathode surface. Let α be the probability that this electron will ionize a gas atom as it travels 1 cm along the discharge tube (α is the first Townsend coefficient). This results in an amplification factor of $\exp(\alpha d)$, i.e. $\exp(\alpha d)$ electrons are released as the primary electron reaches the anode at distance d. This also implies that

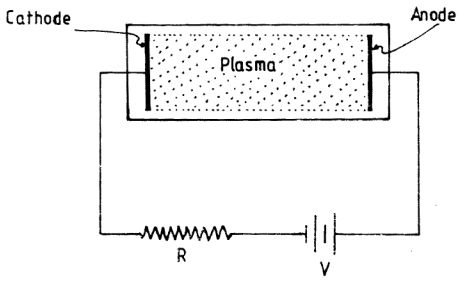

Figure 1. Schematic of a dc discharge system.

[†]For correspondence. (e-mail: rdtarey@physics.iitd.ernet.in)

$\exp(\alpha d)$ ions are produced by the primary electron. To preserve charge neutrality in steady state, these ions have to be given up at the cathode. On striking the cathode the ions produce secondary electrons with efficiency γ (γ is the second Townsend coefficient.). Thus $\gamma \exp(\alpha d)$ electrons are released by the bombarding ions. Now, in the previous pass we had assumed one electron leaving the cathode, which had eventually been lost by recombination at the anode. Thus, to compensate for the loss of this electron and to keep the discharge going, the ions bombarding the cathode must release at least one secondary electron. The condition for sustaining the discharge, can therefore be expressed as: $\gamma \exp(\alpha d) = 1$.

Breakdown voltage: Let l_m be the mean free path for collisions, i.e. the distance traversed by an electron between two collisions. As per the definition of α, the probability that an electron will be released by ionization as an electron travels a distance 1 cm, is equal to the number of collisions in 1 cm ($= 1/l_m$) times the probability P, that a collision will lead to ionization. P can be expressed as $\exp[-I_I/W_c]$, where I_I is the ionization energy of the gas and W_c is the energy gained by the electron between two collisions. If E_B is the electric field acting on the electrons due to the applied dc voltage, then one has $W_c = e\, l_m\, E_B$, where e is the electron charge. On using this in P gives $\alpha = P/l_m = \exp[-I_I/W_c]/l_m$. One may express this relation in a more conventional form[1] as

$$V_B = \frac{Apd}{\ln(pd)+B}. \quad (1)$$

Here we have used $V_B = E_B\, d$ for the breakdown voltage and have taken the mean free path l_m to be inversely proportional to the gas pressure p. A and B are constants (that also absorb the other constants). Equation 1 exhibits a minimum with respect to pd, called the Paschen minimum. At large values of pd (i.e. $d/l_m \gg 1$), the number of collisions within the length of the tube is large. Thus the electrons pick up very little energy between collisions so that $(I_I/e\, l_m\, E_B) \gg 1$, and the probability of ionization decreases. To compensate for this decrease, V_B rises. For small pd, V_B rises again because the collisions within the tube are too few. A plot of the Paschen curve is given in Figure 2. There are other important aspects of the dc discharge like its I–V characteristic, and the alternate dark and luminous bands that correspond to various regions like the cathode and anode sheaths, the positive column, the Faraday dark space, etc. These aspects are treated in detail in standard references[1,2,11], and will not be discussed here.

Remarks: It is to be noted that the continuous loss of electrons to the anode represents a rather severe loss to the system. It renders the dc discharge quite inefficient, since the lost electrons have to be constantly replenished by the release of secondary electrons into the system, i.e. by ion bombardment of the cathode. Thus, the plasma densities obtained in the dc discharge are rather modest.

RF/microwave discharges

Breakdown mechanism: In the presence of a dc electric field, an electron can keep gaining energy continuously till it has sufficient energy to ionize a gas atom. Collisions (elastic), therefore do not play any significant role in the ionization process in such discharges, except to retard the motion of the electrons. In rf discharges on the other hand, the electrons can only gain oscillatory energy from the ac fields. In the absence of collisions, the electron motion is coherent and no power can be absorbed from the rf fields once the oscillatory motions have acquired their steady-state values. For moderate field strengths, this oscillatory energy could be far short of the ionization energy, so that the ionization efficiency should be poor. This however is not the case, since even for moderate rf powers fairly modest plasma densities can be obtained. The latter feature is primarily due to the role played by the collisions in the ionization process. When an electron suffers a collision, its oscillatory motion is disturbed and its momentum randomized. In steady state, the average value of the random energy ($\approx T_e$, the electron temperature in energy units) is such that the energy imparted to the neutral gas atoms in a collision ($\approx T_e\, \delta$; $\delta = 2m_e/M_n$; m_e and M_n are the electron and gas atom masses) equals the energy gained by the electrons in the time between two collisions ($= P_{abs}/\nu_c$, where P_{abs} is average power absorbed from the ac fields and ν_c is the collisions frequency). Thus $T_e \approx P_{abs}/(\nu_c \times \delta)$, and because δ is very small, T_e can become comparable to the ionization energy I_I even for moderate values of P_{abs}.

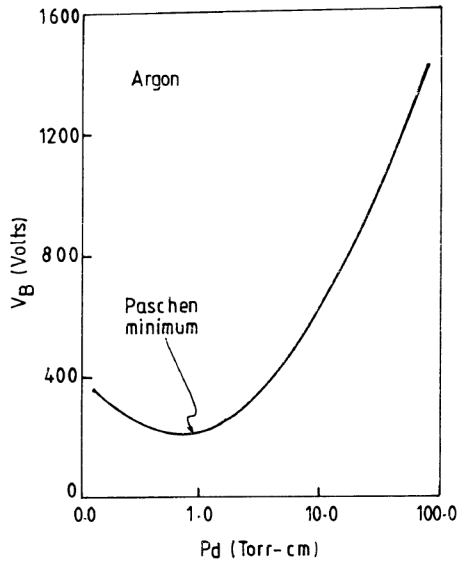

Figure 2. Plot of the breakdown voltage for argon as a function of the product of pressure and electrode spacing (from ref. 1).

The above argument assumes that the maximum excursion of the electrons during the oscillations is smaller than the chamber dimensions. In addition, the argument holds mainly for moderate pressures where $\nu_c \geq \omega$ (ω is the angular frequency of the applied rf). The rf power absorbed due to collisions is given by[12,13]

$$P_{abs} = \frac{e^2 \nu_c E_o^2}{2m_e(\nu_c^2 + \omega^2)}, \quad (2)$$

where E_o is the amplitude of the rf field. Using this in $P_{abs} \approx T_e \nu_c \delta$, it can be shown that in the limit of moderately high pressures ($\nu_c \gg \omega$) the rf breakdown field increases with increasing ν_c, and hence with increasing pressure (since p is proportional to ν_c). Thus in steady state, the absorbed energy goes primarily into heating the neutrals. In the low-pressure limit ($\nu_c \ll \omega$), on the other hand, the absorbed power is dumped mostly into ionizing collisions. To determine the dependence of the breakdown field strength on the pressure, we can regard the discharge as a diffusion-controlled one, i.e. a discharge in which diffusion loss is the primary loss mechanism. In steady state, this loss will be balanced by ionization. This balance is expressed by the relation[11]

$$\frac{\nu_I}{D} = \frac{1}{L_D^2}. \quad (3)$$

Here ν_I is the ionization frequency, $D = T_e/m_e\nu_c$ is the diffusion coefficient and L_D is the diffusion scale length. The above argument can be used to show that at low pressures, the breakdown field strength increases with decreasing pressure. We thus see that the breakdown field rises both in the high- and low-pressure regimes. This implies that there must be a region in between where the breakdown field is a minimum. It turns out that this minimum occurs at a pressure such that $\omega \approx \nu_c$ at the minimum. Physically one can see why this has to be so. For $\nu_c \gg \omega$ a large number of collisions take place within a wave period, so that the electrons are unable to build-up their steady-state oscillation energy. This reduction in the energy acquired by the electrons between collisions is compensated for by an increase in the electric field strength. In the other limit $\nu_c \ll \omega$, the collisions are few and far between, so that many cycles of the wave period go by before a collision occurs. Thus the power absorption is not optimized and the field strength again rises. It may be noted that the electrical power loss in a resistor is also due to collisions and so one may regard the condition $\omega \approx \nu_c$ as similar to the impedance matching condition for maximum power transfer in rf circuits.

Discharge as a function of frequency: A useful limit that characterizes the rf discharge is the 'oscillation amplitude limit' which occurs when the oscillation amplitude x_o of a charged particle

$$x_{oe,i} = \frac{e_{e,i} E_o}{m_{e,i}\omega\left(\nu_{ce,i}^2 + \omega^2\right)^{\frac{1}{2}}}, \quad (4)$$

becomes comparable to a relevant chamber dimension (L), i.e. when $x_o \approx L/2$. (The subscripts e and i in eq. (4) refer to electrons and ions) Beyond this limit (i.e. when $x_o > L/2$) the respective charged particles hit the chamber walls and hence present a major source of loss for the system. To see the role of the oscillation amplitude limit, we consider the changes in the discharge as the frequency is increased from near dc[13]. A plot of the measured breakdown field strength with respect to the rf wavelength λ near the oscillation amplitude limit is given in Figure 3 (ref. 11).

At very low frequencies, both electrons and ions are able to respond to the ac fields quickly and the oscillation amplitude for both species is well beyond the oscillation amplitude limit ($x_{oe,i} \gg L/2$). Both species therefore hit the electrodes at the ends. Thus, as in the dc discharge, the plasma in this case is also sustained by secondary electrons produced by ion bombardment of the cathode. The only difference is that the cathode and anode alternate in each half cycle of the wave, so that the electron–ion bombardment also switches round each half cycle.

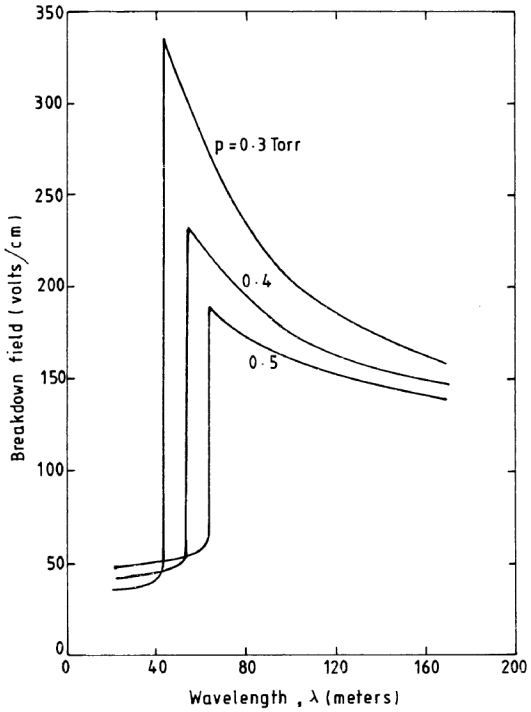

Figure 3. Plot of breakdown field strength versus rf wavelength λ (from ref. 11).

As the frequency is increased, a few things happen simultaneously. Firstly, the ion response to the ac field becomes more sluggish, although the electrons follow the ac fields without any difficulty. Thus, the ion current to the cathode decreases so that both ion bombardment and secondary electron production go down. To compensate for this reduced secondary electron production, the breakdown field rises. This increase in the breakdown field as the wavelength is reduced (at the high wavelengths) can be seen clearly in Figure 3. Finally, it may be noted that because the electron loss to the anode remains unchanged while the ion current to the cathode decreases, the plasma develops a net time-averaged positive potential to trap the electrons and reduce their loss. This potential also helps in enhancing the ambipolar diffusion of ions from the system (see Box 1).

The next major change as the frequency is increased further, occurs when the oscillation amplitude for the ions falls below the oscillation amplitude limit ($x_{oi} \ll L/2$). In this regime, there is a sharp reduction in (ion) bombardment of the cathode, along with a concomitant decrease in secondary electron production. (This situation occurs typically for rf discharges at frequencies around a few megahertz; at these frequencies, the ions cannot respond to the rf fields.) The electron response, however, is still almost instantaneous ($x_{oe} > L/2$) and the electron loss, quite severe. Due to the acute shortage of secondary electrons, the discharge is difficult to maintain and the breakdown field strength rises steeply. Simultaneously, the time-averaged plasma potential attains moderately high values in order to draw out the ions from the plasma, to preserve charge neutrality. The increased ion bombardment of the electrodes enhances secondary electron production that partially compensates, the electron loss from the system. It may be noted that unlike the low-frequency case where the ion bombardment takes place on the instantaneous cathode, the bombardment in the present case occurs symmetrically on both electrodes (for equal electrode areas). The plasma potential in such regimes can be as high as a few hundred volts.

Box 1. Plasma potential

When a finite-sized plasma is formed, the electrons being the more mobile species, tend to escape from the plasma much more rapidly than the ions. The plasma is therefore left with an excess of ions, which makes it slightly positively charged with respect to its original neutral state. In most cases, the plasma is in intimate contact with conducting surfaces like chamber walls or metallic electrodes. These conducting surfaces (unless left floating) usually have well-defined potentials assigned to them with respect to a common ground. By virtue of its intimate contact with these surfaces and due its own conducting nature, the plasma also acquires a definite potential called *the plasma or space potential*. This potential depends on the geometry and extent of contact with the different electrodes, the potentials on these electrodes and the amount of excess positive charge left behind with the plasma (by the escaped electrons). The potential adjusts itself so that it is more positive than the most positive electrode (that makes any significant contact with the plasma). If this were not so, there would be an excessive bleeding of electrons (to large electrodes) that could seriously disturb the plasma equilibrium. Moreover, in order to be able to trap the electrons effectively, the plasma potential (measured with respect to the largest and most positive electrode) must equal approximately the electron thermal energy ($\approx T_e$). It can thus be seen that the plasma potential generally serves two purposes. Firstly, it tends to retain the electrons leaving the plasma, by attracting them back into the plasma. Secondly, it enhances the escape rate of ions from the plasma, so that electrons and ions are lost from the plasma at approximately equal rates (ambipolar diffusion), so that overall charge neutrality (quasineutrality) can be maintained within the plasma.

The plasma potential would tend to remain constant within the plasma, if the electrons in the plasma were in thermal equilibrium. However, the electrons are usually at local equilibrium only, and gradients in electron temperature (and density) are generally present. These gradients give rise to potential variations within the plasma, from which the electric fields within the plasma can be determined (in principle). In general, large electric fields cannot be sustained within the plasma. Near the electrodes, however, the plasma potential has to change sharply to conform to the potentials on the electrode surfaces. This is achieved by the potential changing across a sheath region that acts as an interface between the plasma and the electrode. In contrast to the plasma, which is quasineutral, the sheath is non-neutral, comprising predominantly of species of like charge. The sheath shields any large electric fields that could arise within the plasma due to the electrodes. Exceptions to this situation occur when time-varying potentials are applied to the electrodes. In such cases, depending on the frequency of the applied signal (in relation to the plasma response time) large electric fields can penetrate the plasma. It may be noted that, in the absence of any reference electrodes in contact with the plasma (e.g. in inductively coupled discharges in *closed* quartz tubes), the plasma potential becomes undefined.

1. Chen, F. F., *Introduction to Plasma Physics and Controlled Fusion*, Plenum Press, NY, 1984, p. 290.
2. Chen, F. F., in *Plasma Diagnostic Techniques* (eds Huddlestone, R. H. and Leonard, S. L.), Academic Press, NY, 1965.

The above scenario continues as long as the electrons are beyond the oscillation amplitude limit. (Actually, the discharge behaviour is highly system-specific, since the oscillation amplitude limit is determined by the system dimensions.) As the frequency is raised further, the maximum excursion of electrons also goes below oscillation amplitude limit ($x_{oe} < L/2$). At this point, the heavy loss of electrons ceases and the breakdown field strength drops abruptly by a very large amount, as can be seen from Figure 3. In addition, the time-averaged plasma potential also reduces sharply, since the ion currents are no longer needed for maintaining the discharge. In microwave discharges, for example, where the electron oscillation amplitude is about 10 to 100 microns, the plasma potential is only a few tens of volts. It may be noted that the analysis in the previous sub-section is actually an accurate description of rf discharges in this regime.

Remarks: The dc and rf discharges used in practice usually operate on the principles discussed above. The exceptions are the wave-heated discharges like ECR, helicon and surface wave discharges that attain very high plasma-production efficiencies. In these discharges, suitable waves are excited in the plasma and it is the absorption of these waves by the plasma particles that is responsible for the deposition of the rf or microwave power into the plasma. In the ECR discharge the fundamental ECR condition is also exploited and that further enhances its efficiency.

In the next section, in addition to a few rf discharge schemes that are commonly used for applications such as reactive ion etching, sputtering and plasma processing, we shall also discuss the ECR, helicon and surface wave discharges.

Discharge schemes commonly used in plasma sources

RF discharges

Capacitively-coupled discharges (*CCDs*): Such discharges are primarily used for sputtering or reactive ion etching[1,12]. The basic scheme is shown in Figure 4. The frequency range is usually between 1 and 100 MHz. The corresponding wavelengths (300–3 m) are large compared to the dimensions of the plasma reactor. The pressure range is typically between 1 and 10 Pa. Electron densities obtained are in the range 10^9–10^{10} cm^{-3}. Power to one of the electrodes is fed through a capacitor. The other electrode is grounded. An important feature of CCDs is that a large, negative, self-bias voltage develops on the powered electrode with respect to the plasma (or the ground). This potential helps draw out the ions from the plasma and due to the large potential drop, these ions acquire sufficient energy (up to a few hundred electron volts) to cause effective sputtering or etching of the electrode. The latter are therefore used as target material in the process.

To see how the self-bias voltage develops[13,1], we present a simplified analysis based on the direct-coupled reactor shown in Figure 5. It may be recalled from the earlier discussion on rf plasmas, that at these frequencies the ions do not follow the rf field, while the electrons are still lost from the system. Consequently, the plasma develops a net time-averaged potential that helps expel the ions from the plasma by driving time-averaged ion currents to both electrodes. For ease in our discussion, we imagine rectangular pulses being applied at the electrodes. We let the ratio of the areas of the powered electrode A to that of the grounded electrode B, be β (< 1). The time-averaged ion current to each electrode will be in the ratio of their areas, since the dc potential on both electrodes is the same (i.e. $\approx$ ground potential). Thus, if the ion–particle current to the grounded electrode is I_o, the current to the powered electrode will be βI_o. In Figure 5, the direction of the arrow gives the positive direction of the electric current. The *net electric current due to the ions* (assuming unit ion charge) is therefore $(1 - \beta)I_o$, while the *total number of ions leaving the system* is $(1 + \beta)I_o$. For charge neutrality, the number of electrons leaving the system to the instantaneous anode (in any part of the cycle) must also be $(1 + \beta)I_o$. When the powered electrode has a positive voltage, all the electrons travel to this electrode. These electrons constitute a net electric current $(1 + \beta)I_o$, so that the total current in the circuit due to both ions and electrons becomes approximately, $(1 - \beta)I_o + (1 + \beta)I_o \approx 2I_o$. Similarly, in the negative half cycle all the electrons travel to the grounded electrode and it can be seen that the net electric current in the circuit will now be $(1 - \beta)I_o - (1 + \beta)I_o \approx -2\beta I_o$. Taking an average of the currents in the two cycles gives the time-averaged dc current in the circuit as $(1 - \beta)I_o$.

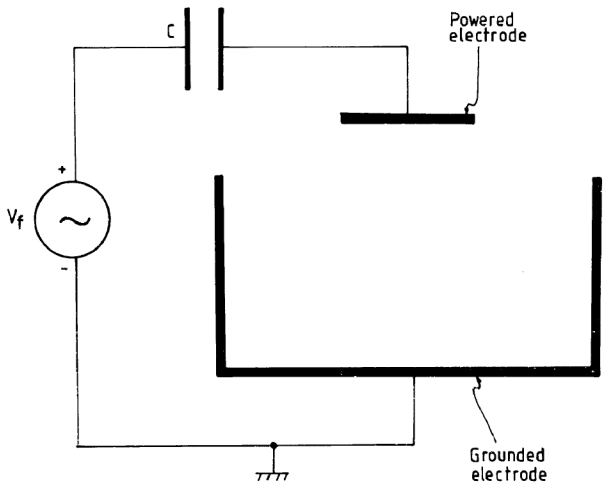

Figure 4. Schematic of capacitively coupled discharge system (from ref. 1).

It is thus seen that in the direct-coupled reactor, a net dc current flows in the circuit whenever the electrode areas are asymmetric. In the present situation, $\beta < 1$ and the dc current flows into the grounded electrode. If a capacitor is now introduced into the circuit as in Figure 4, this dc current will be blocked. This is attained by attracting more ions to the powered electrode, which it may be noticed, becomes floating after introduction of the capacitor. Thus in steady state, the plasma potential that would normally be acquired otherwise, is now manifested as a negative potential on the floating, powered electrode. This negative, self-bias potential on the powered electrode plays the role of the (positive) plasma potential by drawing out the ions, but in a way so that the net dc current is zero. (Basically sufficient electrons accumulate at this electrode, to reduce the net current to zero.)

It may be noted that the self-bias voltage decreases at higher pressures, so that the ion bombardment energy also reduces. This is due to the fact that for $\nu_c \gg \omega$, the electrons do not get to execute their full oscillation swing. As a result, the electron loss decreases and the need for the high plasma potentials for driving the ions out of the system also disappears. Thus, the self-bias voltage is low or nominal. For most applications related to reactive ion etching, etc. the large ion bombardment energy is desirable and so for such applications a reduction in the self-bias voltage may not be an attractive feature. Such applications can therefore, be undertaken in very specific pressure ranges only.

Inductively-coupled discharges (*ICDs*): In such discharges (Figure 6), the rf is fed to helical coils wound axially around a cylindrical, dielectric plasma vessel (usually of quartz)[12]. The power is coupled into the discharge by transformer action (coupling by radiation is poor for the small coil size to free space rf wavelength ratios used). The frequency range is typically between 1 and 100 MHz. At these frequencies transformer coupling is efficient, although a magnetic transformer core is absent. The electric field lines induced in ICDs form closed loops (in planes normal to the coil axis), so that the electron accelerations and oscillations occur in curved orbits. This reduces significantly the electron losses to the walls. This is in contrast to the CCDs, where these oscillations are linear and losses to the electrodes are inevitable for $x_0 > L/2$. (In CCDs and dc discharges, the electric field, while it provides the energy required for ionization on the one hand, also helps transport electrons out of the system on the other, giving rise to a rather heavy loss of electrons.) Due to the reduced loss of electrons in ICDs, the plasma densities attained in these are much higher than those in CCDs. Typically, the plasma densities in ICDs are as high as 10^{12} cm^{-3}. In addition, for discharges in contact with grounded electrodes (as in Figure 6), the plasma potentials are also low, since the ions no longer need to be driven to the walls in large numbers for release of secondary electrons. It may be noted that the breakdown mechanism and discharge operation in ICDs is similar to that described in the section on 'RF/microwave discharges'. ICDs have been used for applications such as thin-film deposition, plasma etching and ion sources for mass spectroscopic analysis[12].

Helicon discharges (*HDs*): Although the set-up used for HDs seems similar to that used for ICDs, the working principle of the former is quite different[3,15]. As mentioned earlier, the rf power in HDs is coupled via suitable waves that are excited and absorbed in the plasma at the applied rf frequency. This renders the process extremely efficient. For typical HDs, the frequency of operation is between 1 and 50 MHz. It may be noted that for a wave to be efficiently absorbed in the plasma, the length of the plasma column should be several absorption scale-lengths. In addition, the coupling efficiency to the waves, using simple antenna structures, should be moderately high. In the chosen range of frequencies, such waves are available in magnetized plasmas. In fact, the waves in question are the so-called helicon waves which are low frequency whistler waves that propagate parallel to the applied magnetic field much below the ECR frequency (i.e. $\omega \ll \omega_{ce}$; ω_{ce}: electron cyclotron frequency in radians). Since the applied frequencies are between 1 and 50 MHz, a magnetic field of 100–300 G usually suffices.

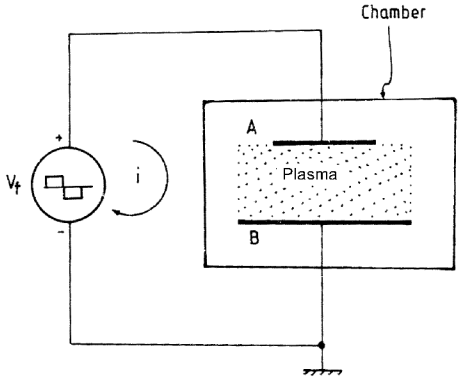

Figure 5. Schematic of direct-coupled rf discharge system.

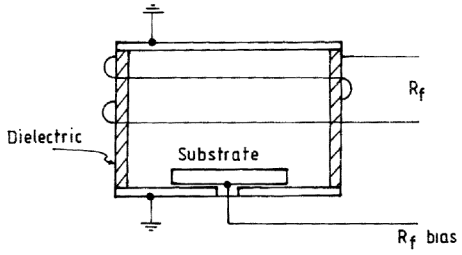

Figure 6. Schematic of inductively-coupled discharge system (from ref. 2).

In general, helicon waves are right-circularly polarized waves, though left-polarized helicons[14] are also possible in cylindrically bounded plasmas. The frequency of these waves is so low that the electron gyrations in the magnetic field can be neglected and only the guiding centre motions retained. In addition, at these low frequencies, the wave is not purely electromagnetic (em), but is partly electrostatic, as well. Helicon waves are excited in the discharge only beyond moderately high power levels ($\geq$ 300 W). At low powers, the coupling is non-resonant and the plasma densities obtained are low. A discontinuous jump in the density takes place as the power is raised beyond a few hundred watts and the helicon waves are excited.

Figure 7 shows some antenna structures that have been used for exciting helicon waves for 'plasma source' applications[12,15]. The antennas can be placed either outside or inside the plasma vessel. In the former case, since the coil size is negligible in comparison to the free-space rf wavelengths, the radiation efficiency of the structure is poor and coupling to the plasma wave takes place by transformer action. (Note that the currents in different arms of the coils are almost in phase.) To obtain the desired coupling to the waves in this case, one has to ensure that these currents are oriented (and placed) so that appropriately phased B-fields are induced inside the plasma and a travelling helicon wave is excited axially. When located inside the plasma vessel, the coil is immersed in the plasma medium where the rf has a wavelength of about a few tens of centimetres (i.e. the wavelength of the helicon wave). Thus, in this case the coil can be designed to *radiate* the rf power into the helicon waves. (In contrast to the earlier case, a *current wave* will be excited on the coil now.) Figure 8 gives plots of typical density profiles obtained using helicon wave sources[3].

Chen[14,15] has discussed various aspects of HDs in considerable detail. Using the helicon-wave dispersion relation and the boundary conditions, and matching the wave phase velocity to the speed of the electron population to be accelerated (primaries), one can arrive at useful design equations for the discharge system. One obtains the relations, $\lambda f/v_e \approx C_1$ and $B_o/(n_e\ a\ v_e) \approx C_2$ (C_1, C_2: constants; f, λ: frequency and wavelength of the helicon wave; v_e: electron speed to which wave phase velocity is matched; B_o: applied magnetic field; n_e: electron density; a: plasma column radius). In the above, v_e is fixed by letting it correspond to the electron speed (energy) at which the ionization cross-section (by electron impact) peaks. One may determine λ from the coupling efficiency G of the antenna, since $G \sim (0.61\lambda/a)^2$ (ref. 14). Typical values of G are about 60–100. f and B_o can then be determined from the formulae for given n_e and a. There is a restriction on the choice of the frequency, since rf sources are usually available at a frequency of 13.56 MHz or one of its harmonics. In any case G, and hence λ, can be adjusted to yield a suitable frequency. The wavelength λ fixes the antenna length, which is usually kept about one half-wavelength long. For argon, the energy corresponding to peak ionization is $\approx$ 50 eV and if one chooses $a \cong 10$ cm, $n_e \approx 10^{13}$ cm^{-3} one obtains a magnetic field of about of 1 kG, which is quite reasonable.

Using the above design procedure, one observes discrepancies among the plasma densities projected and those actually attained. This is particularly true of the larger diameter discharges where one observes the designed values only near the axis, with the average electron density $\langle n_e \rangle$ being lower across the column. Substitution of the values of B_o, a and $\langle n_e \rangle$ in the above formulae, yields values of electron energies (matched to the wave) that are too high. On the other hand, using the peak values of the plasma density yields reasonable values for the energies. In addition, the scale lengths for Landau damping are also much larger than the length of the plasma column. Nonetheless, the design conditions discus-

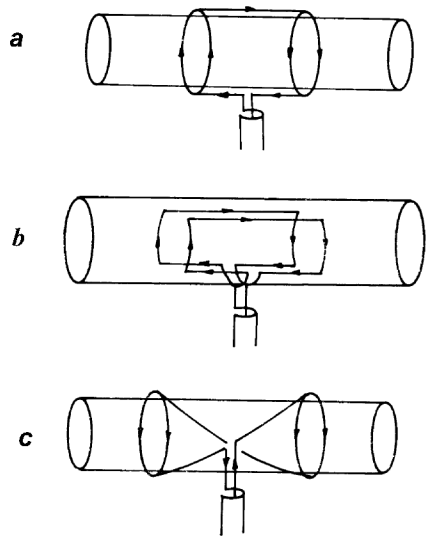

Figure 7. Half wavelength $m = 1$ antenna configurations for helicon discharges: ***a***, Nagayo type III; ***b***, Boswell type; and ***c***, Helical Shoji type (from ref. 12).

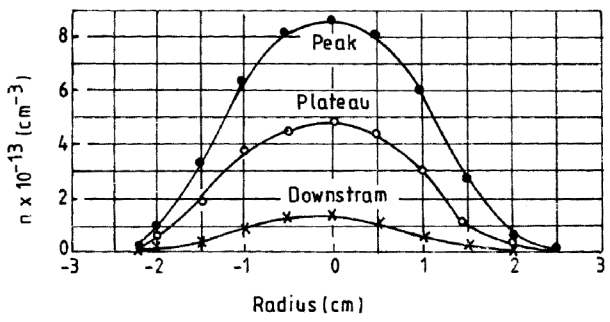

Figure 8. Plasma density profile in a helicon discharge system (from ref. 3).

sed above are quite useful since they provide basic guidelines for optimizing the discharge.

The above discussion is based on the assumption that Landau damping of the primaries is responsible for transfer of energy from the wave to the plasma. Recent work by Chen and Blackwell[16] however, casts doubts on this hypothesis, since non-Maxwellian electron populations are yet to be detected in helicon discharges. In fact, their findings suggest that the energy transfer occurs most likely, via the quasistatic Trivelpiece–Gould (TG) modes excited at the plasma surface and which, due to their short wavelengths, are absorbed quickly while propagating inwards.

Microwave/ECR and surface wave discharges

Microwave discharges (*MDs*): As in rf discharges, MDs can be operated both with and without external magnetic fields. When operated without external magnetic fields, MDs usually use frequencies between a few GHz and several GHz, the frequency 2.45 GHz being the most common. Due to the high frequencies involved, the power is almost invariably coupled through radiation, which bypasses all sheath losses. The antenna can be located either inside the vessel or outside. (In either case, the entire set-up has an electromagnetic (em) shield to avoid radiation hazard.) On penetration into the plasma, the microwave couples to a mode of the plasma and it is the absorption of this mode by the plasma particles that helps maintain the plasma. The power absorption depends on the pressure. At moderate pressures the absorption is through electron–neutral collisions. For chamber sizes much larger than the wavelength, the excited plasma mode is an em mode that experiences a cut-off near the critical density, given by the condition $\omega \approx \omega_{pe}$ ($\omega_{pe} = \{n_e e^2/\varepsilon_o m_e^2\}^{1/2}$ is the plasma frequency and ε_o is the vacuum permittivity). (Actually, the wave may also propagate along the plasma surface as a surface wave; see later in the article.) At 2.45 GHz this critical density is $\approx 7 \times 10^{10}$ cm^{-3}. When the chamber dimensions are comparable to, or smaller than the microwave wavelength, the modes in the system are guided-wave modes of the plasma column loaded inside the plasma chamber (with the latter acting as a waveguide). These guided modes may be either em or quasistatic TG modes for small-diameter plasma vessels. Somewhat higher densities are possible in the latter case, since the TG modes also propagate as surface waves and can be absorbed efficiently over several skin depths near the surface. It was pointed out earlier that the oscillation amplitude for electrons in MDs is about 10–100 microns, so that both electron loss and plasma potentials are much lower than in rf discharges. Electron losses are primarily by diffusion in these discharges.

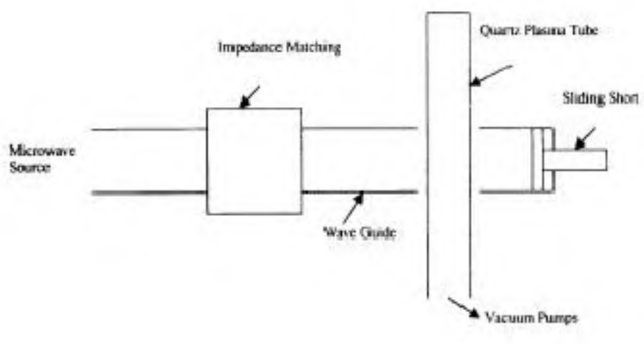

Figure 9. Schematic of microwave discharge system.

Antenna structures used for launching the microwaves can be varied depending on need and convenience[12]. One commonly used set-up is shown in Figure 9, where the discharge is initiated inside a quartz tube placed inside a conducting waveguide. Microwaves may also be launched directly from an open-ended waveguide into the plasma chamber.

ECR discharges (*ECRDs*): The frequencies used for ECR-based applications can vary from a few GHz to several tens of GHz. In tokamaks, ECR preionization and heating are usually undertaken at frequencies between 30 and 100 GHz, depending on the value of the toroidal magnetic field. In ECR ion-beam sources, the requirements are for high plasma densities and efficient electron heating. The frequencies used are lower (typically between 10 and 18 GHz), since lower frequency sources are cheaper and the magnetic fields required are lower. In plasma processing-related applications, the frequency is about a few GHz, the typical value chosen being $\approx$ 2.45 GHz. This is so since sources at these frequencies are the cheapest, and the magnetic field requirement, least stringent. It may be noted that all frequencies used in these devices have to comply with the International Scientific and Medical (ISM) band specifications. In what follows we shall discuss the role of ECR in processing-related applications and assume the frequency of operation to be $\approx$ 2.45 GHz.

In ECRDs, the plasma is initiated by exploiting the fundamental resonance condition $\omega \approx \omega_{ce}$. The applied axial magnetic field (along the chamber axis) is non-uniform, the typical profiles being either that of a mirror or a half-mirror (Figure 10). The microwave is usually launched axially from the high-field region, where $\omega_{ce} > \omega$. It is believed that the microwaves excite whistler waves in the plasma that propagate axially along the magnetic field for $\omega_{ce} > \omega$, to be eventually absorbed at the resonance layer where $\omega \approx \omega_{ce}$. It may be noted that for typical chamber sizes used in processing applications, the modes of the system are guided waves. Theoretical and experimental investigations reveal[17,18] that overdense plasmas in moderate-sized chambers can support multiple, long and short-wavelength guided modes. The excited long-wavelength mode is usually the dominant

wave of the system and is present all along the length of the plasma. (It is also a nonresonant wave and propagates through the resonance layer.) Typically it is found to be an $m = -1$ wave, even when an $m = +1$ mode is launched. Two reasons could be identified for this. Firstly, $m = +1$ long-wavelength guided modes cannot propagate in such plasmas. Secondly, the polarization pattern of the $m = -1$ wave exhibits polarization reversals within the plasma, i.e. it is left-circularly polarized on the axis, but turns into a right-hand polarized wave a little off the axis. Thus, the launched $m = +1$ mode, that is predominantly right-handed near the axis, can excite the $m = -1$ wave easily. The short-wavelength waves, on the other hand, can be both resonant (those that resonate at the $\omega \approx \omega_{ce}$ layer) or nonresonant. These are excited in numerous bursts along the length of the plasma column, and usually last only a few wavelengths. They can be identified as $m = \pm 1$ modes. Both modes exhibit polarization reversals within the plasma, as do the long waves. Thus the $m = +1$ wave that is a right-hand polarized wave on the axis, turns into a left or linearly polarized wave at higher radii, and vice versa. As a result, for the short waves, even small perturbations along the plasma column can lead to coupling between the different modes, so that both modes are usually found to be present together. It is therefore difficult to retain mode purity for these waves.

Since microwaves are coupled through radiation, the structure of the coupling antenna plays a key role in the final plasma densities obtained and the overall efficiency of plasma production. Standard antennas are structures like horns, open-ended waveguides or half-wave dipoles. One may also consider other types of structures like slow-wave antennas[19–22] that slow down the phase velocity of the wave along some direction, which in turn enhances the interaction of the wave with the plasma electrons along that direction. Such structures have very high plasma-production efficiencies. In some recent experiments, 15 cm diameter, 80 cm long, fairly uniform plasma columns with densities about 5×10^{12} cm^{-3} have been reported at pressures of about 10^{-4} Torr in argon, using ~ 2 kW of cw power at 2.45 GHz (Figure 11). These experiments indicate 100% ionization throughout the plasma volume. The antenna used in these experiments was an optimally designed, slotted, helical antenna energized by a conical horn (Figure 12).

It may be noted that considerable care is needed for designing slow-wave antennas. For instance, in free space, the helical antenna supports only one mode which is a slow-wave mode ($v_{ph} < c$; v_{ph}: phase velocity of the wave; c: speed of light in vacuum). However, when placed coaxially inside a conducting waveguide, it supports the usual slow mode (observed in free space) and additional fast wave modes with $v_{ph} > c$ (refs 23 and 24). Moreover, the latter tend to become excited preferentially, since the waveguide is intrinsically a fast-wave structure. To avoid the spontaneous excitation of the fast waves, one fixes the length of the helical antenna so that it acts like a resonant cavity for the slow-wave mode and not for the fast wave. This procedure allows the slow wave to be

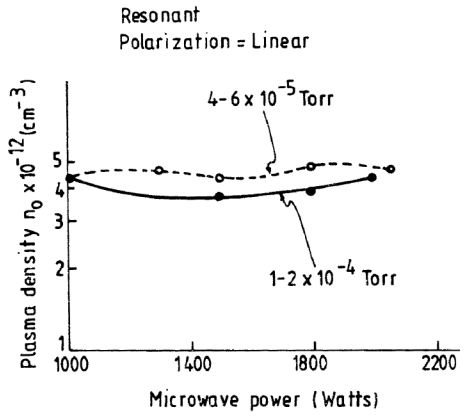

Figure 11. Plot of plasma density versus microwave power for ECR discharges obtained using slotted, helical antenna energized as in Figure 12 (from ref. 19).

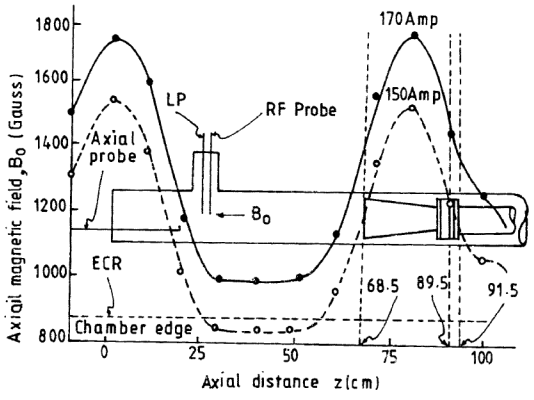

Figure 10. Plasma chamber and relative location of horn antenna with respect to axial magnetic field in an ECR discharge system (from ref. 18).

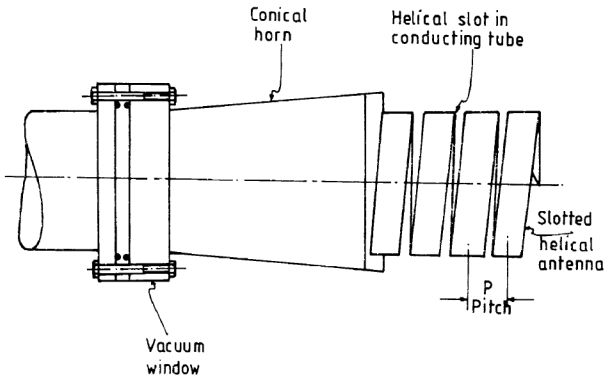

Figure 12. Slotted, helical antenna energized by a conical horn in ECR sources (from ref. 20).

excited in isolation, as is required for efficient plasma production[19–21].

In another recent work[25], moderately overdense plasmas were obtained by launching microwaves in the ordinary mode polarization from a side port (normal to the axial magnetic field). The magnetic field at the site of the microwave launch was kept much lower than the ECR value. According to the theory[17], the coupling to the plasma can take place via surface waves supported by the plasma at such low values of the magnetic field. In the experiments[25], both surface wave field profiles and electron heating were observed near the site of the launch. This method of plasma production has the advantage that no antenna structure needs to be placed within the plasma and because of the side-port launch, microwave window degradation can be avoided.

In applications like sputtering and ion-assisted plasma etching or RIE (reactive ion etching), the ions are required to bombard the target or substrate with moderately high energies. As discussed in detail earlier, such energies are naturally acquired by the ions in capacitively-coupled rf discharges due to the self-bias on the powered electrode. Although ECRDs have the advantages of higher densities and larger volumes and diameters, their plasma potentials are quite low. (The reason for this has already been analysed earlier.) Thus in spite of their inherent advantages, it is not possible to use ECRDs for RIE applications without suitable modifications. Figure 13 suggests such a modification where an additional rf (or dc) bias is provided on the substrate in the usual configuration for an ECRD[6,12]. The rf creates an auxiliary plasma close to the substrate that forms an integral part of the microwave plasma flowing downstream from the discharge chamber. At low-to-medium operating pressures (~ a few mtorr) a sufficiently large sheath voltage develops, that helps accelerate the ions and leads to the onset of ion-assisted plasma etching.

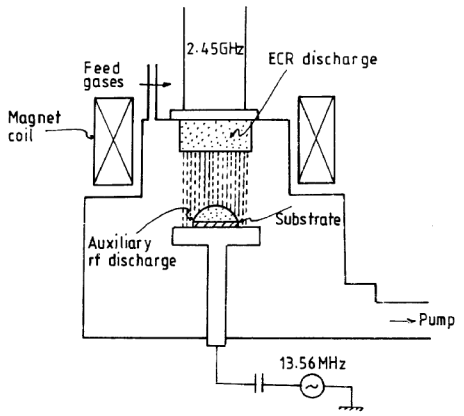

Figure 13. Schematic of an ECR-rf hybrid plasma etching system. The rf is capacitively coupled to the substrate to create an auxiliary discharge (from ref. 6).

In this context, a new concept that emerges is that of an ECR ion beam – microwave hybrid system[6]. This involves creation of a primary ECRD from which an ion beam is extracted. This ion beam can be characterized with respect to its current, energy and uniformity, so that low energy (50–200 eV) uniform ion beams can be directed onto the substrate for controlled 'damage' to the substrate, and its etching. A second ECR discharge just above the substrate provides the plasma environment required for the etching reactions. The ion beam from the primary ECRD enters the second discharge from above and strikes the substrate with sufficient energy to trigger ion-assisted plasma etching. A major advantage of this scheme is that the plasma potential of microwave discharges being low, the energy of the ion beam entering the second discharge (above the substrate) would not be altered significantly by this discharge. Consequently, the energy with which the ions bombard the substrate can still be controlled independently by the primary ECR discharge parameters and the voltages on the extraction grids. This adjustable control over the ion-beam energy can be used for controlled etching of the substrate, without any damage to it. This is to be contrasted with conventionally used rf-biased system discussed above, where one has very little control over the actual plasma potentials developed. The latter cause the ions to strike the substrate with very high energies, causing uncontrolled wafer damage.

Surface wave discharges (*SWDs*): It is well known that em radiation *cannot* propagate in unbounded, non-magnetized plasma, when the frequency of the radiation is lower than the plasma frequency. On the other hand, the plasmas encountered in the experiments are actually *bounded plasmas* that have *boundaries* with other dielectric (e.g. vacuum) or conducting media. Under such circumstances, while there is no difficulty for any radiation to propagate within the dielectric, the induced currents within the plasma prevent the radiation from penetrating the plasma, if the radiation frequency is below the plasma frequency. (This is a purely inductive effect, similar to that encountered in the penetration of em waves into conducting media.) Thus, radiation below the plasma frequency is able to penetrate the plasma only up to few 'skin depths' ($\approx c/\omega_{pe}$). In such cases, although the radiation cannot propagate *into* the plasma, it can propagate *along* the plasma–vacuum interface. Such waves are localized to the plasma surface and are called *surface waves*. The field amplitudes of these waves usually peak at the interface, and drop-off sharply on the plasma side. On the dielectric side, two different situations may arise[2]. If the adjoining dielectric or vacuum has infinite extent, then the wave amplitudes decay sharply on this side as well. However, if the dielectric is in contact with a second medium that is conducting, the fields no longer need to fall-off sharply on the dielectric

side. It may be noted that surface waves in unmagnetized plasmas are usually electromagnetic in nature. However, if the plasma radius is too small, the waves may become slow waves ($v_{ph}/c \ll 1$), and acquire a quasistatic nature (TG modes).

Surface waves gain their importance from the fact that in experiments, the em radiation is usually launched in vacuum, close to the plasma surface. Under these circumstances, if the conditions are right, the launched radiation, instead of penetrating the plasma, may propagate as a surface wave at the plasma–vacuum interface. This situation can arise easily when the plasma is overdense, surrounded by a dielectric layer that may or may not be in contact with another conducting surface.

It is possible for discharges to be sustained by such surface waves. Such discharges are then known as *surface wave discharges* (SWDs). Calculation[2] shows that the surface waves can actually propagate up to frequencies *slightly below* the plasma frequency, and not exactly the plasma frequency. For instance, in slab geometry, the waves propagate for $\omega < \omega_{pe}/(1 + \varepsilon_D)^{1/2}$ where, ε_D (≥ 1) is the dielectric constant of the adjoining dielectric. It turns out that the waves also encounter a resonance ($k \to \infty$; k: propagation constant along the plasma surface) at the frequency $\omega_{pe}/(1 + \varepsilon_D)^{1/2}$. In discharge applications, one usually tries to exploit this resonance, since the aim here is to produce the highest possible density. The existence of surface waves is also confirmed in cylindrical geometry, where the plasma column is surrounded by a thick dielectric[2]. Though the dispersion relation is more complicated in this case, it reduces to that for the planar case for large plasma radii.

In SWDs, both cylindrical and slab geometries have been used. In cylindrical geometry, plasma columns up to about 15 cm diameter have been produced, although 3–10 cm diameters are more common[26]. Frequencies used range typically from 1 MHz to 10 GHz, although for high densities microwaves in the range 1–10 GHz are preferred. The power in the wave is absorbed by electrons near the surface, by collisional heating. The heated electrons diffuse into the bulk of the plasma, where they help maintain the plasma through impact ionization of the gas atoms. The absorption length for surface waves is large, so that efficient absorption of the wave demands large system lengths. Thus, SWD systems in cylindrical geometry are operated in large aspect ratios (i.e. large length to diameter ratios). Since material processing usually requires low aspect ratios, the cylindrical configuration is not too useful for such applications. However, large-area, planar (rectangular) configurations have been developed that are suitable for processing applications[27].

An issue that is of considerable practical interest, is the existence of surface waves in magnetized, high-density (overdense) plasma columns loaded inside conducting waveguides. This is so, since this configuration is used in a variety of applications with the plasma being produced by rf or microwaves. A theory was undertaken for such a configuration recently[17], where the free space wavelength of the em radiation was taken to be comparable to the diameter of the waveguide. Under such circumstances, the modes of the plasma-loaded guide are the guided wave modes of the system. Numerical computations reveal that at frequencies much below the ECR frequency, the waves propagating along the plasma guide are surface waves that tend to become body waves at higher values of the magnetic field. It is possible to exploit this situation for plasma production in a magnetic mirror configuration, by keeping the mid-plane field much below the ECR, and the field at the throats close to the ECR value. The microwaves coupled from a side-port near the mid-plane could easily couple to the surface waves in that region. However, as the waves propagate outwards axially, they could transform into body waves carrying the power into the interior of the plasma, only to be absorbed efficiently near the resonance region at the throats.

As mentioned earlier, an experiment based on the above ideas was undertaken recently[25], in which the microwaves were launched from a side-port at the mirror mid-plane in the ordinary-mode polarization. This polarization was used since the dominant electric field component of the surface waves is along the applied magnetic field. In these experiments, one obtained typically overdense plasmas with densities about 10^{11} cm^{-3} at pressures of about 10^{-4} Torr in argon, without any significant electron heating (bulk $T_e \approx 5$ eV). The plasmas produced in such cases had all the basic features of ECR-produced plasmas[18–22] and these results are in accordance with the theoretical ideas presented above. In addition to the above results, cases were also obtained[25], where significantly warm bulk electrons with $T_e \approx 25$ eV could be observed near the mirror mid-plane. These results indicate a direct absorption of the surface waves near the mirror mid-plane itself. The plasma potential in this region was about 125 V, indicating that confinement of these warm electrons is electrostatic. The plasma characteristics obtained in these latter cases are significantly different from those of ECR discharges, and present a new regime of plasma parameters that can be exploited for plasma-source applications in the future.

Conclusions

In this article, we have discussed some common plasma-discharge schemes used in plasma sources. Only non-equilibrium, low-to-moderate pressure discharges have been treated. The discharges are classified with respect to the frequency of the power source and the mode of power coupling. Using the latter classification as a basis, the physical mechanism behind each type of discharge is discussed in detail. The frequency range covered is from dc to microwaves, both with and without magnetic fields.

The qualitative changes brought about in the discharges with the frequency of the power source are discussed, along with a simplified treatment on the origin of the self-bias in capacitively coupled rf discharges. Finally a new concept using ECR discharges is presented for (ion-assisted) plasma etching without damage to wafers.

1. Cecchi Joseph, C., in *Handbook of Plasma Processing Technology* (eds Rossnagel Stephen, Cuomo Jerome, J. and Westwood William, W.), Noyes Publications, Park Ridge, NJ, 1990.
2. Lieberman Michael, A. and Lichtenberg Allan, J., *Principles of Plasma Discharges and Materials Processing*, John Wiley, New York, 1994.
3. Chen Francis, F., *IEEE Trans. Plasma Sci.*, 1995, **23**, 20.
4. Graves David, B., *ibid*, 1994, **22**, 31.
5. Manos Dennis, M. and Flamm Daniel, L. (eds), *Plasma Etching – An Introduction*, Academic Press, CA, 1989.
6. Ganguli, A., Akhtar, M. K., Tewari, D. P. and Taery, R. D., in *Reviews in Contemporary Plasma Physics* (eds Raj Kamal, Maheshwari, K. P. and Sawhney, R. L.), Wiley Eastern Ltd, New Delhi, 1992.
7. Tarey, R. D. and Ganguli, A., *Phys. News*, 1998, **29**, 46.
8. *Phys. Today*, 1979, **32**, Special issue on magnetically confined fusion.
9. Teller, E. (ed.), *Magnetic Confinement*, Academic Press, NY, 1981, vol. I.
10. Hazeltine, R. D. and Meiss, J. D., *Plasma Confinement*, Addison-Wesley, CA, 1992.
11. Brown Sanborn, C., *Introduction to Electrical Discharges in Gases*, John Wiley, NY, 1966.
12. Conrads, H. and Schmidt, M., *Plasma Sources Sci. Technol.*, 2000, **9**, 441.
13. Ganguli, A., Workshop on 'Plasma Surface Interaction and Plasma Processing', Pune Univ. Pune, 24–29 May 1993; Unpublished.
14. Chen Francis, F., *Plasma Phys. Controlled Fusion*, 1991, **33**, 339.
15. Chen Francis, F., *J. Vac. Sci. Technol.*, 1992, **A10**, 1389.
16. Chen Francis, F. and Blackwell David, D., *Phys. Rev. Lett.*, 1999, **80**, 2677.
17. Ganguli, A., Akhtar, M. K. and Tarey, R. D., *Phys. Plasmas*, 1998, **5**, 1178.
18. Ganguli, A., Akhtar, M. K., Tarey, R. D. and Jarwal, R. K., *Phys. Lett. A*, 1998, **250**, 137.
19. Tarey, R. D., Jarwal, R. K., Ganguli, A. and Akhtar, M. K., *Plasma Sources Sci. Technol.*, 1997, **6**, 189.
20. Ganguli, A., Jarwal, R. K., Tarey, R. D. and Akhtar, M. K., *IEEE Trans. Plasma Sci.*, 1997, **25**, 1086.
21. Ganguli, A., Appla Naidu, P, and Tewari, D. P., *ibid*, 1991, **19**, 433.
22. Ganguli, A., Baskaran, R. and Pandey, H. D., *ibid*, 1990, **18**, 134.
23. Ganguli, A. and Appala Naidu, P., *J. Phys. D. Appl. Phys.*, 1986, **19**, 2265.
24. Ganguli, A. and Appala Naidu, P., *J. Appl. Phys.*, 1990, **68**, 3679.
25. Ganguli, A., Akhtar, M. K. and Tarey, R. D., *Plasma Sources Sci. Technol.*, 1999, **8**, 519.
26. Moisan, M. and Zakrzewski, Z., *J. Phys. D. Appl. Phys.*, 1991, **24**, 1025.
27. Komachi, K., *J. Vac. Sci. Technol.*, 1992, **A11**, 164.