Understanding and modelling plasma–electrode interaction in high-pressure arc discharges: a review

To cite this article: M S Benilov 2008 J. Phys. D: Appl. Phys. 41 144001

View the article online for updates and enhancements.

You may also like

- On the occurrence of gauge-dependent secularities in nonlinear gravitational waves
  Fabio Briscese and Paolo Maria Santini
- Quantum tasks require islands on the brane
  Alex May and David Wakeham
- Ground state energy of the -Bose and Fermi gas at weak coupling from double extrapolation
  Sylvain Prolhac
Understanding and modelling plasma–electrode interaction in high-pressure arc discharges: a review

M S Benilov

Departamento de Física, Universidade da Madeira, 9000 Funchal, Portugal

Received 26 December 2007, in final form 18 April 2008
Published 4 July 2008
Online at stacks.iop.org/JPhysD/41/144001

Abstract
Considerable advances have been attained during the last decade in the theoretical and experimental investigation of electrode phenomena in high-pressure arc discharges, in particular, in low-current arcs that occur in high-intensity discharge lamps. The aim of this paper is to deliver a concise review of the understanding achieved and modelling methods developed.

(Some figures in this article are in colour only in the electronic version)

Nomenclature

\begin{align*}
A & \quad \text{area of the cathode cross section} \\
A_f & \quad \text{work function of the cathode material} \\
A_i & \quad \text{ionization energy of the plasma-producing gas} \\
c_p & \quad \text{specific heat of the cathode material} \\
D_{ei}, D_{ia} & \quad \text{coefficients of diffusion of the electrons} \\
& \quad \text{and the ions} \\
d & \quad \text{ionization length} \\
E & \quad \text{electric field strength} \\
e & \quad \text{electron charge} \\
h & \quad \text{cathode height} \\
I & \quad \text{arc current} \\
I_1 & \quad \text{arc current corresponding to the limit} \\
& \quad \text{of stability of the diffuse mode} \\
I_t & \quad \text{arc current corresponding to a turning} \\
& \quad \text{point of a spot mode (the upper limit of its} \\
& \quad \text{existence range)} \\
j & \quad \text{local density of electric current} \\
j_e & \quad \text{density of electron current to the electrode} \\
& \quad \text{surface} \\
& \quad \text{(0)} \\
j_{e(0)} & \quad \text{density of electron current to the electrode} \\
& \quad \text{surface that occurs at zero sheath} \\
& \quad \text{voltage} \\
& \quad \text{(sat)} \\
j_{e(sat)} & \quad \text{density of electron saturation current to the} \\
& \quad \text{electrode surface} \\
j_{em} & \quad \text{density of electron current emitted by the} \\
& \quad \text{electrode surface} \\
& \quad \text{surface} \\
& \quad \text{surface that occurs at zero sheath} \\
& \quad \text{volt} \\
& \quad \text{(sat)} \\
k & \quad \text{Boltzmann constant} \\
k_i & \quad \text{rate coefficient of ionization of an atom} \\
& \quad \text{by electron impact} \\
L & \quad \text{local length scale} \\
L_{lm} & \quad \text{length scale on which heat conduction} \\
& \quad \text{by heavy particles is comparable to the} \\
& \quad \text{rate of energy exchange between} \\
& \quad \text{electrons and heavy particles} \\
L_{tp} & \quad \text{length scale on which heat conduction in} \\
& \quad \text{the plasma is comparable to radiative} \\
& \quad \text{losses} \\
m_e, m_i & \quad \text{electron and ion masses} \\
n & \quad \text{direction locally orthogonal to the} \\
& \quad \text{cathode surface and directed outside the} \\
& \quad \text{cathode} \\
n_a, n_e, n_i & \quad \text{number densities of atoms, electrons} \\
& \quad \text{and ions} \\
n_c & \quad \text{distribution of the charged-particle} \\
& \quad \text{number density in the ionization layer} \\
n_{c(sh)} & \quad \text{characteristic value in the sheath of the} \\
& \quad \text{density of the charged particles that are} \\
& \quad \text{attracted by the sheath electric field}
\end{align*}
\[ r_{\text{Saha}} \] charged-particle density evaluated under the assumption of ionization equilibrium
\[ p \] plasma pressure
\[ P_e \] Pécel number
\[ Q_c \] power removed by heat conduction from the cathode body to the cooling fluid
\[ Q_p \] total power received by the cathode from the plasma
\[ Q_t \] power irradiated by the cathode surface
\[ q \] density of heat flux removed from the cathode surface into the bulk by heat conduction
\[ q_e \] density of energy flux delivered to the cathode surface by fast plasma electrons
\[ q_{\text{em}} \] density of losses of energy by the cathode surface due to thermionic emission
\[ q_i \] density of energy flux delivered to the cathode surface by ions
\[ q_p \] density of the plasma-related energy flux to the cathode surface
\[ q_r \] density of losses of energy by the cathode surface through radiation
\[ R \] radius of a rod cathode
\[ R_{\text{arc}}, R_{\text{at}} \] characteristic radii of the arc column and the arc attachment
\[ r \] radial coordinate; see figure 7
\[ T \] temperature distribution in the cathode body
\[ T_c \] temperature of the base of the cathode
\[ T_e \] electron temperature in the plasma
\[ T_e^* \] average electron temperature in the near-cathode layer
\[ T_h \] heavy-particle temperature in the plasma
\[ T_{\text{pl}} \] temperature of an LTE plasma
\[ T^{(\text{sh})} \] characteristic value in the sheath of the temperature of the particles that are repelled by the sheath electric field
\[ T_{\text{tip}} \] average temperature of the cathode tip
\[ T_w \] temperature of the surface of the electrode time
\[ U \] near-cathode voltage drop (voltage drop in the near-cathode layer)
\[ U_{\text{con}} \] combined voltage drop in the region of constricted plasma and the constriction zone
\[ U_{\text{sh}} \] sheath voltage
\[ U_{\text{fns}}, U_{\text{fp}} \] voltages in the layer of thermal non-equilibrium and the layer of thermal perturbation
\[ v \] characteristic velocity of plasma flow
\[ v_{\text{is}} \] ion speed at the edge of the space-charge sheath
\[ w_{\text{rad}} \] rate per unit volume at which energy is lost by the plasma as a result of radiative processes
\[ x \] coordinate counted from the electrode surface into the plasma
\[ z \] axial coordinate; see figure 7
\[ \alpha \] inverse of the Knudsen number for the ionization layer
\[ \Gamma_c \] base of the cathode
\[ \Gamma_h \] part of the cathode surface that is in contact with the plasma and the cold gas
\[ \epsilon \] emissivity of the cathode material
\[ \epsilon_0 \] permittivity of free space
\[ \kappa, \kappa_\epsilon, \kappa_h, \kappa_{\text{pl}} \] thermal conductivities of the cathode material, the electron gas, the heavy-particle gas and the plasma on the whole
\[ \lambda_D \] Debye length
\[ \lambda_e, \lambda_i \] electron and ion mean free paths
\[ \bar{\nu}_{\text{ch}} \] average frequency of collisions of an electron with heavy particles
\[ \rho, \rho_c \] mass density and electrical resistivity of the cathode material
\[ \sigma \] Stefan–Boltzmann constant
\[ \sigma_{\text{pl}} \] electrical conductivity of the plasma
\[ \varphi \] azimuthal angle; see figure 7
\[ \chi_{\text{pl}} \] thermal diffusivity of the plasma
\[ \omega \] ionization degree of the plasma

### 1. Introduction

An adequate organization of current transfer from the plasma to electrodes is critical for the performance and lifetime of any arc discharge device. As far as high-intensity discharge (HID) lamps are concerned, an inadequate organization of current transfer to electrodes can cause highly undesirable effects such as degradation of the electrodes, blackening of the walls of the lamp, enhanced heat conduction losses and electromagnetic interference (EMI) emissions.

Electrodes of arc discharges have been under intensive investigation for many decades; a review on electrodes of arc discharges in ambient gas published in 1961 by Ecker [1] already included more than 600 references and numerous further references were cited in subsequent reviews [2–7]. However, the experimental investigation has been hindered by the fact that phenomena governing plasma–electrode interaction take place in thin layers adjacent to electrode surfaces which are extremely difficult to resolve due to their very small dimensions as well as by extreme conditions typical for arc discharges. Even basic quantities such as near-cathode and near-anode voltage drops have been reliably measured relatively recently. The lack of reliable experimental data was detrimental to theory, as was a late realization of the fact that different modes of current transfer, typical for electrodes of arc discharges in general and, in particular, for electrodes of HID lamps, represent self-organization phenomena and must be described as such. As a consequence, reliable experimental data and self-consistent theoretical models started to emerge only in the 1990s and mostly for cathodes.

The aim of this paper is to summarize the present understanding of cathodes and anodes of high-pressure arc discharges. The focus is on theoretical models and methods of simulation and on low-current discharges. Most of
the concepts under review are applicable not only to low-current high-pressure arc electrodes but also to electrodes of arc discharges in ambient gas in general, in particular, to cathodes of low-pressure dc discharges in fluorescent lamps (so-called thermionic arcs), and some concepts also apply to arc discharges in vacuum. On the other hand, some physical mechanisms that are unimportant under conditions of low-current high-pressure arcs, although they may come into play under other conditions, are beyond the scope of this paper.

The paper is organized as follows. In section 2, the basic physics of near-electrode plasmas in high-pressure low-current arc discharges is analysed with the aim of identifying mechanisms governing current transfer to cathodes and anodes. A review of recent papers on theory and modelling of plasma–cathode interaction and of the most important results is given in section 3. In section 4, some papers on the plasma–anode interaction are reviewed and the main factors hindering or facilitating investigation of the near-anode physics are discussed. Works concerned with modelling electrodes of ac arcs and integration of models of plasma–electrode interaction into codes describing the whole system arc-electrodes are discussed in section 5. Conclusions are summarized in section 6.

2. Basic physics of near-electrode plasmas: why are the cathode and the anode so different?

Near-cathode and near-anode phenomena in high-pressure arc discharges are manifestly different: the near-anode voltage drop is usually smaller than the near-cathode voltage drop and can be both positive and negative, the temperature of the anode tip is below the temperature of the cathode tip at low arc currents and above the cathode tip temperature at high currents, energy fluxes from the plasma received by cathodes are substantially different from those received by anodes, spectacular mode changes observed on cathodes at low currents are not normally observed on anodes etc. In order to find the root reason for these differences, one needs to consider the basic physics of near-electrode regions.

2.1. Structure of the near-electrode perturbation region

Different kinds of perturbations introduced into the arc plasma by electrodes manifest on different length scales. This allows one to divide the near-electrode perturbation region into a number of sub-regions with different physics.

2.1.1. Constriction zone and region of constricted plasma.

The attachment of an arc to an electrode is usually considerably narrower than the arc column as schematically shown in figure 1: \( R_{\text{att}} \ll R_{\text{arc}} \), where \( R_{\text{att}} \) and \( R_{\text{arc}} \) are characteristic radii of the arc attachment and, respectively, the arc column. Therefore, the most far-reaching perturbation introduced into the arc plasma by an electrode consists of a decrease in the radius of the arc channel in the vicinity of the electrode. This decrease occurs at distances from the electrode of the order of \( R_{\text{arc}} \). The region in which this decrease occurs is marked in figure 1 as the constriction zone.

In the region positioned ‘at the bottom’ of the constriction zone at distances of the order of \( R_{\text{att}} \) from the electrode, the density of electric current is of the same order of magnitude as that at the electrode surface. This region is marked in figure 1 as the constricted plasma. Let us estimate typical values of plasma parameters and length scales characterizing different physical processes in this region.

Let us consider a mercury plasma under pressure \( p = 30 \) bar as a representative example of HID lamp plasmas and \( 10^7 \) A m\(^{-2} \) as a characteristic value of the electric current density at the electrode surface. (The latter value approximately corresponds to an attachment of radius \( R_{\text{att}} = 250 \) \( \mu \)m for arc current \( I = 2 \) A.) Assuming that the constricted plasma is close to LTE and that its energy balance is dominated by radiation, i.e. that the Joule heating is locally balanced by the energy lost by the plasma as a result of radiative processes, one finds that the local plasma temperature \( T_{\text{pl}} \), electron number density \( n_e \) and ionization degree are approximately 8800 K, \( 4.5 \times 10^{23} \) m\(^{-3} \) and 1.9\%. (Note that transport, kinetic and radiative properties of the mercury plasma are evaluated in this work along the lines of [8]. The assumption of local thermodynamic equilibrium, or LTE, implies, in particular, that the radiation escape does not significantly affect the kinetics of excited states. For the mercury plasma, the latter assumption is justified if the electron density substantially exceeds \( 10^{19} \) m\(^{-3} \), which is clearly the case under the conditions considered.)

The assumption of energy balance of the constricted plasma being dominated by radiation, \( j E = u_{\text{rad}} \), implies that Joule heating and radiative losses of energy dominate over all the other mechanisms of energy transport, which are the convective transport of energy, enthalpy transport by diffusion fluxes of species and heat conduction. (Here \( j \) is the local density of electric current, \( E \) is the electric field and \( u_{\text{rad}} \) is the rate per unit volume at which energy is lost by the plasma as a result of radiative processes.) The contribution of convective transport of energy compared with that of heat conduction is of the order of the Péclet number \( Pe = L v / \chi_{\text{pl}} \), where \( L \) is a local length scale, \( v \) is a characteristic flow velocity and \( \chi_{\text{pl}} \) is a characteristic thermal diffusivity of the plasma. Alternatively, the Péclet number may be expressed as \( Pe = \rho L v / T_{\text{pl}} \kappa_{\text{pl}} \), where \( \kappa_{\text{pl}} \) is a characteristic thermal conductivity of the plasma. According to the above, one can set \( p = 30 \) bar, \( L = R_{\text{att}} = 250 \) \( \mu \)m, \( T_{\text{pl}} = 8800 \) K while

![Figure 1. Structure of the near-electrode perturbation region.](image-url)
analysing the constricted plasma. The thermal conductivity of an LTE mercury plasma at these values of temperature and pressure is \( \kappa_e = 0.22 \text{ W m}^{-1} \text{ K}^{-1} \). Values of the maximum gas velocity in HID lamps calculated by most authors are of the order 0.1–0.2 m s\(^{-1}\) (e.g. [9–14]), although substantially higher calculated values of 6–8 m s\(^{-1}\) have also been reported [15,16]. Since the gas velocity in the immediate vicinity of the electrode is smaller than the maximum gas velocity in the lamp, one can set \( v = 0.1 \text{ m s}^{-1} \). Then one finds \( \text{Pe} = 0.04 \) and the contribution of convective transport of energy is much smaller than that of heat conduction.

The main component of enthalpy transport by diffusion fluxes of species is enthalpy transport by electric current. The density of the corresponding energy flux is of the order of \( jkT/\kappa \), where \( k \) is the Boltzmann constant and \( \kappa \) is the electron charge. The ratio of this quantity to a characteristic density of the corresponding energy flux is of the order of magnitude as heat conduction and need not be estimated separately.

It follows that it is sufficient to check whether radiative losses in the constricted plasma substantially exceed heat conduction; if the answer is positive, then the energy balance of the plasma is dominated by radiation. Let us define \( L_{\text{ip}} \) as a length scale on which heat conduction in the plasma is comparable to radiative losses. This definition amounts to the order-of-magnitude equation \( \kappa_{\text{pl}} T_{\text{pl}}/L_{\text{ip}}^2 \sim u_{\text{rad}} \), so one can set

\[
L_{\text{ip}} = \sqrt{\kappa_{\text{pl}} T_{\text{pl}}/u_{\text{rad}}}. \tag{1}
\]

Note that a somewhat different formula for this length scale can be obtained by taking into account that the product \( \sigma_{\text{pl}} u_{\text{rad}} \) is a rapidly varying function of \( T_{\text{pl}} \) (here \( \sigma_{\text{pl}} \) is the electrical conductivity of the plasma): such a formula differs from (1) by the multiplier \( [d \ln(\sigma_{\text{pl}} u_{\text{rad}})]/d \ln T_{\text{pl}} \) on the right-hand side [17]. The latter multiplier equals 0.27 for the conditions considered and does not provoke order-of-magnitude changes. Therefore, this multiplier is not introduced in this work.

On a length scale \( L \neq L_{\text{ip}} \), the ratio of contributions of heat conduction and radiative losses is of the order of \( (L_{\text{ip}}/L)^2 \). \( L_{\text{ip}} = 130 \mu\text{m} \) and \( L = R_{\text{att}} = 250 \mu\text{m} \) for the conditions considered, so the ratio of contributions of heat conduction and radiative losses is in the constricted plasma of the order of 0.25. It follows that the assumption of energy balance of the constricted plasma being dominated by radiation, while not being very accurate, is acceptable in the framework of the qualitative analysis that we are engaged in.

Significant deviations from local LTE can occur in the column of high-pressure arc discharges at low arc currents. For example, deviations from LTE on the axis of a wall-stabilized arc discharge in atmospheric-pressure argon become appreciable when the current is decreased to several tens of amperes [18]. The current density in the region of constricted plasma is considerably higher than the current density in the arc column, so there are more chances of encountering LTE here than in the column. Nevertheless, the existence of LTE even in the region of constricted plasma in a low-current high-pressure arc is far from certain and must be analysed separately for each case.

Three kinds of deviations from LTE are most important in the near-electrode physics: a violation of thermal equilibrium, i.e. a divergence between electron and heavy-particle temperatures \( T_e \) and \( T_h \); a violation of the ionization equilibrium, i.e. a deviation of the electron density \( n_e \) from the electron density \( n_{\text{Saha}} \) predicted by the Saha equation; and a violation of quasi-neutrality, i.e. a divergence between the electron and ion densities \( n_e \) and \( n_i \). The manner in which deviations from the thermal equilibrium are analysed depends on the relation between the thermal conductivities of the heavy-particle gas, \( \kappa_h \), and the electron gas, \( \kappa_e \). If the ionization degree is low enough and \( \kappa_h \gg \kappa_e \), then \( T_h \) is decoupled from, and may be determined independently of, \( T_e \). One should find out whether the electron temperature \( T_e \) follows, or not, the heavy-particle temperature \( T_h \). To this end, terms of the electron energy equation should be analysed. Such an analysis would involve the length of relaxation of electron energy in a weakly ionized plasma, which coincides with the length of relaxation of the isotropic part of the electron distribution function in a weakly ionized plasma (e.g. [19]) and equals \( M_e \sqrt{m_e/m_i} \) (here \( M_e \) is the mean free path of electrons and \( m_e \) and \( m_i \) are the electron and ion masses).

If \( \kappa_e \ll \kappa_h \), then \( T_e \) is decoupled from \( T_h \) and the question is whether the heavy-particle temperature \( T_h \) follows, or not, the electron temperature \( T_e \); the terms of the heavy-particle energy equation should be analysed. In this case, terms where \( \kappa_h \) and \( \kappa_e \) are comparable, the same result is obtained by analysing either energy equation.

Given that \( \kappa_h = 0.08 \text{ W m}^{-1} \text{ K}^{-1} \) and \( \kappa_e = 0.14 \text{ W m}^{-1} \text{ K}^{-1} \), we can set \( \text{Pe} = 0.002 \), i.e. negligible.

Ionization equilibrium is perturbed by ambipolar diffusion of the charged particles and by convective transport of the charged particles. The ratio of the contributions of convective transport and ambipolar diffusion is of the order of the diffusion \( \text{Péclet number} \) \( Pe_1 = L_v/D_a \) defined in terms of the coefficient of ambipolar diffusion \( D_a \). \( D_a = 1.2 \times 10^{-5} \text{ m s}^{-1} \) under the conditions considered and one finds that \( Pe_1 \) is approximately 2.1, i.e. of order unity. Hence, convective transport of charged particles produces an effect of the same order of magnitude as ambipolar diffusion and need not be estimated separately. The dominating mechanism of ionization of neutral atoms in arc plasmas is ionization by electron impact and its rate per unit
volume is \( k_i n_e n_a \), where \( k_i \) is the rate constant of ionization of an atom by electron impact and \( n_a \) is the atomic density. The rate of transport of charged particles due to ambipolar diffusion is comparable to the ionization rate on a length scale of the order of the so-called ionization length:

\[
d = \frac{1}{C_2} \sqrt{\frac{D_i k T_i}{k_i e}}
\]

Here \( C_2 \) is a dimensionless coefficient which is defined by equation (14) of [21] and varies between approximately 0.67 and 1. \( d = 0.15 \mu m \) in the conditions considered. The relative deviation of \( n_e \) from \( n_{Sat} \) in the constricted plasma is of the order of \( (d/R_{at})^2 = 3.6 \times 10^{-5} \), i.e. also negligible.

Violations of quasi-neutrality are characterized by the Debye length \( \lambda_D = \sqrt{\varepsilon_0 k T_i / n_e e^2} \). \( \lambda_D = 0.01 \mu m \) in the conditions considered. The relative divergence between \( n_e \) and \( n_i \) in the constricted plasma is of the order of \( (\lambda_D/R_{at})^2 = 1.6 \times 10^{-9} \), i.e. also negligible.

Thus, one comes to the conclusion that the constricted plasma is indeed close to LTE under conditions typical for HID lamps.

### 2.1.2. Near-electrode non-equilibrium layer.

‘At the bottom’ of the constricted plasma, the role of a local length scale is played not by the attachment radius \( R_{at} \) but rather by the distance to the electrode surface: \( L = x \), where \( x \) is a coordinate counted from the electrode surface in the direction into the plasma. As \( x \) decreases, quantities \( (L_p/L)^2 \), \( (L_m/L)^2 \), \( (d/L)^2 \), and \( (\lambda_D/L)^2 \) increase proportionally to \( 1/x^2 \) and eventually become comparable to or exceed unity. In other words, there is a thin layer ‘at the bottom’ of the constricted plasma where the energy balance of the plasma is no longer dominated by radiation and the plasma is no longer in LTE. This layer is marked in figure 1 as the near-electrode non-equilibrium layer.

According to the above estimates, the typical values of length scales \( L_p \), \( L_m \), \( d \) and \( \lambda_D \) differ significantly, being \( \lambda_D \ll d \ll L_m \ll L_p \). This hierarchy suggests that the deviation of the plasma energy balance from being dominated by radiation and different kinds of deviations from LTE come into play not simultaneously but rather sequentially. Therefore, the near-electrode non-equilibrium layer may be divided into a number of sub-layers with different physics. This fine structure is illustrated by table 1 and may be explained as follows.

In the region of constricted plasma, i.e. at \( x \) of the order of \( R_{at} \), the plasma is quasi-neutral, ionization and thermal equilibria hold and the plasma energy balance is dominated by radiation: \( n_e = n_i, n_e = n_{Sat}, T_e = T_i, j E = w_{rad} \). When \( x \) is decreased, i.e. when one moves from the region of constricted plasma to the electrode surface, the first deviation to come into play is the one characterized by the biggest length scale, which is \( L_p \). Hence, one can introduce a layer which is adjacent to the region of constricted plasma, has a thickness of the order of \( L_p \) and in which the balance between Joule heating and irradiated energy is violated due to losses of energy caused by heat conduction in the direction of the electrode. The plasma in this layer is still in LTE. One can say that this is a layer in which cooling of the plasma by the electrode comes into play. This layer is marked in table 1 as the layer of thermal perturbation.

When one approaches the electrode surface closer, the next deviation to come into play is the violation of thermal equilibrium. Hence, one can introduce a layer which is adjacent to the layer of thermal perturbation and in which the electron and heavy-particle temperatures are no longer equal. The plasma is still quasi-neutral in this layer and the ionization equilibrium still holds. This is the layer marked in table 1 as the layer of thermal non-equilibrium. The scale of its thickness may be evaluated in terms of the characteristic values of plasma parameters in this layer by means of equation (2) if \( \kappa_b \ll \kappa_e \) in this layer; otherwise the scale of thickness of the layer of thermal non-equilibrium should be deduced from the equation of electron energy.

Still closer to the electrode surface, the ionization equilibrium breaks down as well, although the plasma is still quasi-neutral: \( n_e = n_i \neq n_{Sat} \). This is the layer marked in table 1 as the ionization layer. The scale of its thickness may be evaluated in terms of the characteristic values of the plasma parameters in this layer by means of equation (3).

Quasi-neutrality breaks down in the layer immediately adjacent to the electrode surface, a space-charge sheath. Let us designate by \( n_{Sat}^{(sh)} \) and \( T^{(sh)} \) the characteristic values in the sheath of the density of the charged particles that are attracted by the sheath electric field and, respectively, of the temperature of the particles that are repelled. If the sheath voltage \( U_{sh} \) is moderate, \( |U_{sh}| \ll T^{(sh)}/e \), then the scale of thickness of the sheath is given by the local Debye length:

\[
\lambda_D = \sqrt{\varepsilon_0 k T^{(sh)}/n_{Sat}^{(sh)} e^2}
\]

If \( U_{sh} \) is high, \( |U_{sh}| \gg k T^{(sh)}/e \), then the scale of thickness is given by the Debye length in which thermal energy \( k T^{(sh)} \) is replaced by the electrostatic energy \( e U_{sh} \):

\[
\lambda_D = \sqrt{\varepsilon_0 U_{sh}/n_{Sat}^{(sh) e}}
\]

Thus, a structure of the near-electrode perturbation region in low-current high-pressure arc discharges typical for HID lamps has been established: it is shown in figure 1 and table 1. Note that the sub-structure of the near-electrode non-equilibrium layer shown in table 1 is similar to the structure of the near-anode perturbation region in high-current arcs proposed in [17, 22]. Somewhat more involved structures have been considered in [2, 4].

As an example, distributions of parameters in the near-anode non-equilibrium layer under the above-specified conditions are shown in figure 2. Here and further \( E \) is the projection of the electric field over the \( x \)-axis and \( T_w \) is the

<table>
<thead>
<tr>
<th>Layer</th>
<th>Constricted plasma</th>
<th>Layer of thermal perturbation</th>
<th>Layer of thermal non-equilibrium</th>
<th>Ionization layer</th>
<th>Space-charge sheath</th>
</tr>
</thead>
<tbody>
<tr>
<td>( n_e = n_i )</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>( T_e = T_i )</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>( j E = w_{rad} )</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>( d )</td>
<td>( L )</td>
<td>( d )</td>
<td>( d )</td>
<td>( d )</td>
<td>( d )</td>
</tr>
<tr>
<td>( d/L )</td>
<td>( 1 )</td>
<td>( 1 )</td>
<td>( 1 )</td>
<td>( 1 )</td>
<td>( 1 )</td>
</tr>
<tr>
<td>( \lambda_D )</td>
<td>( 1 )</td>
<td>( 1 )</td>
<td>( 1 )</td>
<td>( 1 )</td>
<td>( 1 )</td>
</tr>
<tr>
<td>( (d/L)^2 )</td>
<td>( 1 )</td>
<td>( 1 )</td>
<td>( 1 )</td>
<td>( 1 )</td>
<td>( 1 )</td>
</tr>
<tr>
<td>( (\lambda_D/L)^2 )</td>
<td>( 1 )</td>
<td>( 1 )</td>
<td>( 1 )</td>
<td>( 1 )</td>
<td>( 1 )</td>
</tr>
</tbody>
</table>
temperature of the electrode (anode, in this case) surface which in this modelling was set equal to 3000 K. These distributions have been calculated by means of a unified modelling [8] in which equations written without simplifying assumptions such as thermal equilibrium, ionization equilibrium and quasi-neutrality are solved in one dimension from the electrode surface right up to the constricted plasma. Boundary conditions at the electrode surface are formulated in [8] taking into account thermionic emission of electrons.

Although the data plotted in figure 2 have been obtained by modelling the whole of the non-equilibrium layer by means of a single set of equations, i.e. without dividing it into different sub-layers, the modelling results clearly reveal the structure shown in table 1. CP, TP, TN, IL and SH in figure 2 stand for, respectively, the region of constricted plasma, the layer of thermal perturbation, the layer of thermal non-equilibrium, the ionization layer and the space-charge sheath; vertical dashed lines are used to separate adjacent sub-layers (these separation lines are shown for illustrative purposes only; of course, no distinct boundaries between different sub-layers exist in reality).

The calculations being considered have been performed in the planar geometry; hence, the current density is constant and so are all parameters in the region of constricted plasma, as one can see in figure 2. Then one can extrapolate the (linear) potential distribution in the region of constricted plasma to the electrode surface. The difference between this extrapolated value and the actual potential of the electrode surface has the meaning of the near-electrode voltage drop. The near-anode voltage drop calculated in this way for conditions of figure 2 is −0.4 V.

2.2. Ion and electron currents to the electrodes

In most cases, the thickness of the space-charge sheath is much smaller than the ionization length, so the generation of charged particles in the sheath is negligible compared with the particle fluxes coming to the sheath edge from the ionization layer. Therefore, densities of ion and electron currents to the electrode surface, \( j_i \) and \( j_e \), are determined by particle fluxes from the ionization layer to the sheath edge.

Let us designate by \( j_i^{(0)} \) and \( j_e^{(0)} \) the densities of the ion and electron currents to the sheath edge that occur in the case where the role of the electric field in the ionization layer is limited to maintain quasi-neutrality, without accelerating or suppressing diffusion of charged particles to the electrode. (Note that the sheath voltage is zero in this case.) With the use of appendix B of [21], one can derive

\[
j_i^{(0)} = e D_{ia} \frac{n_{Saha}}{d}, \quad j_e^{(0)} = e D_{ea} \frac{n_{Saha}}{d}, \tag{4}
\]

where \( n_{Saha} \) is the local charged-particle density evaluated under the assumption of ionization equilibrium; \( D_{ia} \) and \( D_{ea} \) are coefficients of diffusion of the ions and the electrons in a weakly ionized gas, the pressure of the gas being equal to the pressure in the arc and the gas temperature being equal to the local temperature of heavy particles. Note that the simplicity of these formulae is due to the presence of coefficient \( C_2 \) in the definition of the ionization length (3).

Let us consider the case where there is a strong electric field in the ionization layer suppressing the electron flux to the electrode surface: \( j_e \ll j_i^{(0)} \). (Note that the electrode is under a high negative potential relative to the sheath edge in this case, i.e. the sheath voltage is positive and high, \( U_{sh} \gg kT_e/e \).) The ion current density tends to a value \( j_i^{(sat)} \) given by the formula (see appendix B of [21])

\[
j_i^{(sat)} = e D_{ia} \left( 1 + \frac{T_e}{T_n} \right) \frac{n_{Saha}}{d}. \tag{5}
\]

Note that \( j_i^{(sat)} \) exceeds \( j_i^{(0)} \) by the factor \( (1 + T_e/T_n) \), which accounts for an enhancement of diffusion of the ions in the ionization layer due to an accelerating electric field. Since \( j_i^{(sat)} \) is independent of the sheath voltage (at \( T_e, T_n \) and \( p \)), it can be called the ion saturation current density.

Now let us consider the case where there is a strong electric field in the ionization layer suppressing the ion flux to the electrode surface: \( j_i \ll j_e^{(0)} \). (Note that the sheath voltage is negative and high in this case, \(-U_{sh} \gg kT_i/e\).) The electron current density tends to a value \( j_e^{(sat)} \) given by the formula

\[
j_e^{(sat)} = e D_{ea} \left( 1 + \frac{T_i}{T_n} \right) \frac{n_{Saha}}{d}. \tag{6}
\]

Again, this formula can be derived with the use of appendix B of [21].

Under conditions of the near-anode layer shown in figure 2, \( T_n \approx 3000 \) K and \( T_e \approx 6000 \) K at the edge of the ionization layer. One finds \( d = 0.4 \mu m \) and \( j_e^{(0)} = 4 \times 10^1 \) A m\(^{-2}\), \( j_i^{(sat)} = 3 \times 10^2 \) A m\(^{-2}\).

These estimates manifest the root cause of the difference between the operation of the anode and the cathode. The ratio \( j_e^{(0)}/j_i^{(0)} \) exceeds unity; hence, electron heat conduction from the layer of thermal non-equilibrium to the ionization layer is more than sufficient to sustain an electron density in the ionization layer needed to transport current to the anode. In fact, three
quarters of the flux of electrons that would come to the sheath edge without an accelerating or repelling electric field in the ionization layer should be suppressed. This explains why the electric field in the ionization layer (and in the space-charge sheath) in figure 2 is negative, i.e. directed from the plasma to the anode surface. One can conclude that the sheath plays a minor role in the near-anode phenomena.

Most of the electric current on the cathode is transported by electrons emitted by the cathode surface. Schottky-amplified thermionic emission is a dominating mechanism of electron emission under conditions of high-pressure arcs (e.g. [23]), so each emitted electron takes with it an energy equal to the work function on leaving the cathode. Thus, electron emission produces a cooling effect over the cathode, and there should be a mechanism responsible for overcoming this cooling and heating the cathode surface up to the temperatures necessary for thermionic emission. The role of such a mechanism can be played by the heating of the cathode by ions coming from the plasma and accelerated in the space-charge sheath. Therefore, the ion current at the cathode surface must not be too low.

Let us recast this reasoning into a quantitative form, with the aim of obtaining an order-of-magnitude estimate for \( j_i \) the ion current density at the cathode surface. An ion, on coming to the cathode surface after having been accelerated by the sheath electric field, brings to the surface the kinetic energy \( eU_{sh} \). The energy released at the cathode surface as a result of neutralization of an ion equals \( A_i - A_e \), where \( A_i \) is the ionization energy and \( A_e \) is the work function. Then the power of ion heating per unit area of the cathode surface equals \( (eU_{sh} + A_i - A_e) j_i/e \). This power is partially taken away by electrons emitted from the cathode surface and partially removed from the surface inside the cathode body by heat conduction. Thus, the power balance of the cathode surface may be written as

\[
\frac{j_i}{e}(eU_{sh} + A_i - A_e) = \frac{j_{em}}{e} A_i + q, \tag{7}
\]

where \( j_{em} \) is the electron emission current density and \( q \) is the density of the heat conduction flux. Clearly, this equation is written in a very simplified form. In particular, it bears no account of heating of the cathode surface by fast plasma electrons capable of overcoming the repelling electric field in the sheath, of terms of the order of thermal energies of electrons and heavy particles, of cooling of the cathode surface by radiation, of energy lost by ions in the space-charge sheath in collisions with neutral atoms. However, this equation is sufficiently accurate for order-of-magnitude estimates.

If the cathode is undoped and there are no additives in the plasma which could contribute to a reduction in the work function through formation of monolayers on the cathode surface, then the temperature of the cathode tip necessary for sufficiently strong thermionic emission is around 3000 K. Assuming 1000 K for the temperature of the base of the cathode, 10 mm for the cathode height and 100 W mK\(^{-1}\) for thermal conductivity of the cathode material (thermal conductivity of tungsten at 2000 K), one can estimate \( q \) to be \( 2 \times 10^7 \text{ W m}^{-2} \) and the so-called volt equivalent of heat losses, \( q/j_i \), to be 2 V. Substituting \( j_{em} \) in equation (7) by \( j - j_i \), setting \( U_{sh} = 10 \text{ V}, A_i = 10.4 \text{ eV} \) (value for mercury), \( A_e = 4.5 \text{ eV} \) (a value typical for most metals of interest) and solving this equation, one finds \( j_i = 0.32 j \). Thus, a fraction of the ion current at the cathode surface is of the order of tens of per cent.

The ratio \( j_i^{(sat)}/j \) estimated for conditions of figure 2 equals \( 3 \times 10^{-3} \), i.e. it is far lower than that. It follows that the charged-particle density in the ionization layer, which is sustained by the electron heat conduction from the layer of thermal non-equilibrium under conditions of figure 2, is insufficient for transporting current to the cathode. A considerably higher ionization level and, consequently, additional heating of electrons are needed, and this heating may come only from acceleration in the space-charge sheath of the electrons emitted by the cathode. In other words, considerable electrical power must be deposited in the space-charge sheath in order to sustain current on a cathode. Thus, the space-charge sheath plays an extremely active role in a cathode, in contrast to what happens on an anode.

These conclusions are illustrated by figure 3, in which distributions of parameters in the near-cathode non-equilibrium layer obtained by means of unified modelling [8] are shown for the same conditions as those of figure 2. Again, the modelling results clearly reveal the structure shown in table 1. The plasma temperature in the layer of thermal perturbation decreases with decreasing \( x \) similarly to what happens in the layer of thermal perturbation in the near-anode layer shown in figure 2, although this decrease is barely visible on the scale of figure 3. On the other hand, the behaviour of \( T_e \) in the layer of thermal non-equilibrium, in the ionization layer and in the space-charge sheath shown in figure 3 is very different from that seen in figure 2: as \( x \) decreases, \( T_e \) first slightly decreases, then slightly increases and then dramatically increases up to very high values of around \( 7 \times 10^4 \text{ K} \), after which it rapidly decreases down to values of around \( 3 \times 10^5 \text{ K} \). This very high maximum of \( T_e \) occurs inside the space-charge sheath and represents a consequence of the heating of electron gas through acceleration.
by the sheath electric field of the electrons emitted by the cathode. The charged-particle density in the ionization layer under conditions of figure 3 is about two orders of magnitude higher than that under conditions of figure 2. The electric field in the space-charge sheath is also very high (about two orders of magnitude higher than that in the near-anode layer), as is the power input into the near-cathode layer (the near-cathode voltage is 38 V).

In summary, figure 3 combined with figure 2 confirms the above reasoning of the root cause of the difference between the operation of electrodes of high-pressure arc discharges: electron heat conduction from the layer of thermal non-equilibrium to the ionization layer is sufficient to sustain a charged-particle density in the ionization layer needed to transport current to the anode but not to the cathode; considerable additional electrical power must therefore be deposited in the space-charge sheath on the cathode; the space-charge sheath, while playing a minor role in the anode, is of central importance on the cathode.

### 3. Cathodes of low-current high-pressure arc discharges

Current transfer from high-pressure arc plasmas to thermionic cathodes may occur in a diffuse mode, where the current is distributed over the front surface of the cathode in a more or less uniform way, or in a spot mode, where most of the current is localized in one or more small areas (cathode spots). In a certain arc current range, either of the modes can occur.

In a number of aspects, cathode spots are damaging; for example, transient spots that can appear on electrodes of ac HID lamps at the beginning of the cathodic phase cause rapid blackening of the walls of the burner. On the other hand, the spots result in a lower average temperature of the cathode and, consequently, in lower heat conduction losses; cathode spots enhance the brightness of the arc which is especially important for compact light sources such as projection or automotive lamps. But in any case, stable operation of the cathode in the desired mode is of extreme importance.

It is a more or less common understanding nowadays that the existence of multiple modes of current transfer to high-pressure arc cathodes is not necessarily a manifestation of essentially different physical mechanisms, such as Schottky-amplified thermionic emission versus thermo-field or field emission. Rather, it is a manifestation of the non-uniqueness of multidimensional thermal balance of a finite body heated by an external energy flux depending in a nonlinear way on the local surface temperature. These and other aspects of theory and modelling of interaction of thermionic cathodes with high-pressure low-current arc discharges are reviewed in this section.

#### 3.1. The model of nonlinear surface heating

As shown in section 2.2, two sections of the near-cathode perturbation region play a particular role in the plasma–cathode interaction, the ionization layer and the space-charge sheath. The ion flux to the cathode surface is generated in the ionization layer. The ions, after having been accelerated by the sheath electric field, bring their energy to the cathode surface. This is the main mechanism of heating of the cathode, although some additional heating may be provided by fast plasma electrons that can overcome the decelerating electric field in the sheath and reach the cathode surface. The energy necessary for ionization in the ionization layer comes from acceleration by the sheath electric field of electrons emitted by the cathode.

In the following, the ionization layer and the space-charge sheath are jointly termed the near-cathode layer; see figure 4.

One can expect that the near-cathode layer gives a dominating contribution to the voltage drop in the whole near-cathode perturbation region, i.e. that \( U \gg U_{\text{in}}, U_{\text{pl}}, U_{\text{con}} \), where \( U \) is the combined voltage drop in the space-charge sheath and in the ionization layer, \( U_{\text{in}} \) and \( U_{\text{pl}} \) are voltage drops in, respectively, the layer of thermal non-equilibrium and the layer of thermal perturbation and \( U_{\text{con}} \) is the combined voltage drop in the region of constricted plasma and the constriction zone. Of particular importance is the last inequality, \( U \gg U_{\text{con}} \); it ensures that the electrical resistance of the near-cathode layer considerably exceeds the combined resistance of the constricted plasma and the constriction zone; therefore, the distribution of current over the cathode surface is governed by the near-cathode layer rather than the constriction zone and/or the constricted plasma. Let us check the validity of this inequality. An order-of-magnitude estimate for \( U_{\text{con}} \) may be obtained by assuming that the current expansion is semi-spherical, then \( U_{\text{con}} = I/2\pi \sigma_{\text{pl}} R_{\text{att}} \). Setting \( \sigma_{\text{pl}} = 10^3 \text{ S m}^{-1} \) (a value for an LTE mercury plasma at \( p = 30 \text{ bar} \) and \( T = 8800 \text{ K} \), \( I = 2 \text{ A} \) and \( R_{\text{att}} = 250 \mu\text{m} \), one finds \( U_{\text{con}} = 1.3 \text{ V} \). \( U \) is of the order of 10 V; hence the inequality \( U \gg U_{\text{con}} \) is justified, i.e. the resistance of the near-cathode layer is indeed much higher than the combined resistance of the constricted plasma and the constriction zone. It follows that the voltage drop \( U \) in the near-cathode layer is approximately the same for all points of the arc attachment.

The above means that the plasma–cathode interaction in high-pressure low-current arc discharges is governed by a thin near-cathode plasma layer comprising the space-charge sheath and the ionization layer, and this near-cathode plasma layer is to a first approximation unaffected by the outside plasma. Therefore, there is in principle no need to calculate the whole system arc–cathode simultaneously: one can first find a solution for the near-cathode layer, then a solution describing
the cathode and finally a solution for the arc column. More specifically, the procedure is as follows.

As a first step, one calculates the characteristics of the near-cathode layer. Since this layer is much thinner than a characteristic dimension of the arc attachment, it may be treated as locally one-dimensional (1D). Therefore, the characteristics of the near-cathode layer are found as functions of the local surface temperature $T_w$ and the near-cathode voltage drop $U$. (It should be stressed that in the framework of such an approach current transfer across the near-cathode layer to any point of the arc attachment depends only on the temperature of the cathode surface at this point but not at other points; the near-cathode voltage drop $U$ represents a combined voltage drop in the space-charge sheath and in the ionization layer and is the same for all points of the arc attachment.)

In particular, densities of the net energy flux and the electric current from the plasma to the cathode surface are found: $q = q(T_w, U)$, $j = j(T_w, U)$. These characteristics do not depend on the total arc current or the shape of the cathode and it is sufficient to calculate them only once for every given combination of the cathode material, the plasma gas and its pressure.

At the second step, one calculates the temperature distribution inside the cathode body and on the surface. This amounts to solving inside the cathode the (multidimensional) heat conduction and current continuity equations with the boundary conditions supplied by the dependences $q(T_w, U)$ and $j(T_w, U)$. If the Joule heat production in the body of the cathode is unessential, which is usually the case for low-current arcs, the current continuity equation may be excluded from consideration. After this problem has been solved, one can substitute the surface temperature distribution obtained as a part of the solution into the dependences found at the previous step, thus determining distributions of parameters of the near-cathode layer along the cathode surface. Integrating the current density distribution, one finds the arc current corresponding to the considered value of the near-cathode voltage $U$.

If the principal aim of a study consists of a modelling of a cathodic part of the discharge, which is the case, for example, in investigations dealing with cathode erosion, the procedure of solution terminates here. Otherwise, one can calculate the arc plasma outside the near-cathode layer, taking necessary boundary conditions from the solution obtained at the second step; the third step.

The above-described approach may be called the model of non-linear surface heating. The most important feature of this model is the existence under certain conditions of more than one solution at the same value of the arc current, with different solutions describing different modes of current transfer. This feature allows for a self-consistent calculation of different modes, thus eliminating the necessity of switching different mechanisms (such as thermionic electron emission versus thermo-field or field emission) ‘by hand’ in order to calculate different modes, which is the usual way of simulating different modes in other models.

The model of nonlinear surface heating was apparently first suggested in 1963 by Bade and Yos [24], who not only gave a mathematical formulation of the problem but also forwarded an assumption of the existence of two or more solutions for any given set of input conditions, corresponding to different modes of cathode operation. However, such multiple solutions were not found in [24], neither was their existence proved, and the importance of this work was not properly appreciated at the time. Besides, the work of Bade and Yos has apparently never been published in a research journal, while the technical report [24] was not widely accessible.

As a consequence, the model of nonlinear surface heating was virtually forgotten for several decades and has apparently been re-discovered, with some or other variations, more than once; see, e.g., works [25, 26] and their discussion in [23], or the recent work [27]. An important role in promoting this model was played by the book [3] in which a detailed description of the model with reference to the work [24] was given, although—somewhat ironically—without mentioning the possibility of the existence of multiple solutions describing different modes.

A revival of the model of nonlinear surface heating was started in the 1990s, after it was proved that the model indeed gives multiple solutions describing different modes of current transfer. It is interesting to note that the first such proof [28] was given by means of the bifurcation analysis, which is a usual tool in the studies of self-organization in nonlinear dissipative systems. Understanding of the fact that different modes of current transfer to cathodes of high-pressure arc discharges represent self-organization phenomena and must be described as such was also beneficial for the subsequent development of the theory. By now, the model of nonlinear surface heating has become virtually universally accepted.

As described above, finding a solution in the framework of the model of nonlinear surface heating comprises three steps:

1. Finding a solution for the near-cathode layer,
2. finding a solution inside the cathode body and, if necessary, 
3. joining a solution for the cathodic part of the discharge and a solution for the arc plasma outside the near-cathode layer. Modelling of the near-cathode layer is considered in the next section. Finding the temperature distribution in the cathode is treated in section 3.3. Results on the plasma–cathode interaction on the whole are reviewed in section 3.4. Works in which a solution for the cathodic part of the discharge is joined to a solution for the bulk plasma are reviewed in section 5.

### 3.2. Finding solution on the plasma side

#### 3.2.1. Ionization layer—space-charge sheath models

According to the above, a model of a near-cathode layer in a high-pressure arc discharge must include descriptions of the ionization layer and the near-cathode space-charge sheath, with some or other procedure of the matching of solutions describing these two zones. A simple model may be obtained as follows. An equation governing distribution of the density $n_c$ of the ions and electrons in the ionization layer is written taking into account ionization, recombination and ambipolar diffusion. The ambipolar diffusion term is written with the use of Fick’s law or the Stefan–Maxwell equations [29–31], depending on whether the ionization degree on the plasma side of the ionization layer is small or of order unity. The Saha
equation serves as a boundary condition for $n_e$ on the plasma side of the ionization layer. Since the charged-particle density in the sheath is much smaller than inside the ionization layer, $n_e$ is set equal to zero on the cathode side of the ionization layer. Variations of the heavy-particle and electron temperatures $T_h$ and $T_e$ in the ionization layer are neglected. The above-described nonlinear diffusion equation governing $n_e$ admits an analytic first integral, resulting in formula (5) for the density of ion current to the cathode surface. $T_h$ is assumed to be equal to the temperature of the cathode surface. $T_e$ is governed by the equation of conservation of the electron energy in the ionization layer written in the form of integral balance. Sources in this equation are energy brought in the ionization layer by electrons emitted by the cathode and accelerated in the space-charge sheath and the work of the electric field over the electrons inside the ionization layer. Sinks are energy carried away by fast electrons leaving the ionization layer for the sheath, energy carried away by electrons leaving the ionization layer for the bulk plasma and losses of electron energy for ionization inside the layer. The space-charge sheath is described by a model of electron-free sheath formed by ions that enter the sheath with a negligible velocity and cross it without collisions, the Child–Langmuir model [32, 33].

Let us consider some of the modifications introduced into the above-described simple model by different workers. There is an elaborate theory of collisionless space-charge sheaths in the literature (see, e.g., review [34]) developed mostly in connection with problems of plasma–surface interaction in low-pressure gas discharges and of plasma edge in fusion devices. A commonplace in this theory is a model of a collisionless ion sheath taking into account the presence of Boltzmann-distributed plasma electrons with the Bohm criterion [35, 36] serving as a boundary condition on the plasma side (edge) of the sheath. This model was introduced into the theory of near-cathode layers in high-pressure arc discharges in [37]. Also introduced in [37] was an account of a finite temperature of the ions in the sheath which amounts to a simple integration of the ion Vlasov equation with an assumed ion velocity distribution on the sheath edge, another commonplace in sheath theory. An enhanced-accuracy model of a collision-dominated ion sheath taking into account the presence of Boltzmann-distributed plasma electrons was developed in [38].

Let us proceed to papers dealing with different aspects of the description of the ionization layer. Numerical treatments of the ionization layer accounting for variability of heavy-particle and electron temperatures were given in [39, 40].

If the cathode operates in the spot mode, the electron temperature $T_e$ in the near-cathode layer may reach very high values, several electronvolts and higher. Multiply charged positive ions can appear at $T_e$ that high. The effect of multiply charged ions on the ionization layer was studied numerically in [41]. It was found that the ion current to the cathode is formed in an inner section of the layer where the singly charged ions are dominant; hence, the ion current can be calculated accurately enough neglecting the presence of multiply charged ions. One can expect that this conclusion is not restricted to atmospheric-pressure argon arcs treated in [41]: this is a consequence of a successive decrease in the rate constants of each subsequent ionization, which is a general trend rather than a specific property of argon.

A scale of thickness of the ionization layer is represented by the ionization length $d$ in the framework of the above-described diffusion analysis. At $T_e$ not very high, $d$ considerably exceeds the ion mean free path $\lambda_i$. It follows from equation (3) that $d$ varies approximately proportionally to $\sqrt{(T_e + T_h)/k_e}$, with variation of $T_e$ at fixed $p$ and $T_h$. As $T_e$ increases, $\lambda_i$ increases faster than $T_e$ in the range of $T_e$ of interest, including at high $T_e$. Hence, the ionization length $d$ decreases with the increase in $T_e$ and the ratio $\lambda_i/d$, which has the meaning of a local Knudsen number, increases. At $T_e$ sufficiently high, $\lambda_i/d$ becomes comparable to or greater than unity. This means that coupling between the ions and the atoms in the ionization layer becomes not strong enough and a diffusion description of the ion–atom relative motion in the layer is no longer valid. A theory appropriate for such conditions can be developed on the basis of the so-called (multi) fluid description of the plasma, which amounts to treating the ions and the atoms as separate fluids coexisting with each other, e.g. [42–44]. Equations of motion of the ion and atomic fluids in such an approach involve terms accounting for inertia, the pressure gradient, the electric field force affecting the ions, the momentum exchange (friction force) between the ions and the atoms due to elastic collisions, the momentum exchange due to volume ionization and recombination. It should be stressed that the difference from diffusion theory resides in the terms accounting for inertia and the momentum exchange due to volume reactions: the Stefan–Maxwell equations comprise all the above-mentioned terms except these two. Multifluid theory of the ionization layer in high-pressure arc discharges was pursued in [45–49]. As shown in [47], results given by multifluid theory conform to the available experimental data.

As an example, the ion current generated in the ionization layer of the atmospheric-pressure argon plasma is shown in figure 5 (taken from [23]). The solid and dashed lines represent the results given by the multifluid model [46] and, respectively, by diffusion formula (5). Also shown are the (equilibrium) ionization degree $\omega$ on the plasma side of the ionization layer and the ratio $\alpha = d/\lambda_i$, which has the meaning of the inverse of the Knudsen number. As expected, diffusion theory provides a good approximation at $T_e$ not too high, where $\alpha \gg 1$. As $T_e$ grows and $\alpha$ becomes comparable to and then much smaller than unity, the growth of the ion current becomes slower than that predicted by diffusion theory and then gives way to a decrease. Thus, while the ion current evaluated in the diffusion approximation monotonically increases in the whole range of $T_e$ considered, multifluid theory predicts a non-monotonic dependence, with the maximum value under conditions of figure 5 being equal to $2.3 \times 10^7$ A m$^{-2}$ and attained at $T_e = 3.7$ eV. It should be emphasized that the latter value is considerably higher than values at which full ionization is attained on the plasma side of the ionization layer. In other words, there is no direct relation between the limitation of the ion current and the full ionization of the plasma. A further discussion of figure 5 can be found in [23].

Note that although the multifluid approach produces results that are sufficient for practical purposes and agree
with the experiment, it is still incomplete theoretically: the question what are boundary conditions governing charged-particle distribution in the ionization layer for \( \alpha \) below a certain value remains open and challenging.

Let us now consider recent models of the near-cathode layer on the whole. There are three models with a collisionless space-charge sheath: the model formulated in [37] and modified in [23, 30] (a summary of equations of this model can be found in [51]), the model [52] and the model [53]. There is also a model [16], which is similar to [37] except that the ion motion across the space-charge sheath is treated as collision-dominated and the sheath is described by means of the solution [38].

Let us compare the models [23, 30, 50], [52] and [53] in some detail. The model [23, 30, 50] relies on the multifluid solution for the ionization layer [46], which accounts for inertia and the momentum exchange due to volume reactions and is supposed to apply for all \( \alpha \). The model [52] has been formulated for the case \( d \gg \lambda_i \) and, consequently, employs a diffusion solution, which accounts for neither inertia nor the momentum exchange due to volume reactions. The model [53] has been formulated for the case \( d \gg \lambda_i \); however, in this work an attempt has been made to expand the range of applicability of the solution for the ionization layer by means of taking into account the momentum exchange due to volume reactions, although not inertia.

In the case \( d \lesssim \lambda_i \), the multifluid solution for the ionization layer can in a natural way be matched [37] through the Bohm criterion with a solution for the collisionless sheath. The situation is different in the case \( d \gg \lambda_i \): a collision-dominated ionization layer and a collision-free sheath are not adjacent but rather separated by an intermediate layer. (In mathematical terms: there is no overlapping between the region of validity of a diffusion solution for the ionization layer and the region of validity of a solution neglecting collisions in the sheath; hence, no unambiguous way exists to formulate matching conditions.)

Any attempt to directly match a diffusion solution for the ionization layer with a collisionless solution for the sheath, such as in the works [39, 40], where the ions were assumed to leave the ionization layer and enter the sheath with a velocity equal to their mean thermal speed \( \sqrt{3kT_i/2m_i} \) or, respectively, the Bohm velocity \( \sqrt{kT_e/m_i} \), is unjustified.

The difficulty may be overcome with the use of a solution for the above-mentioned intermediate layer separating a collision-dominated ionization layer and a collision-free sheath: a layer with a scale of thickness equal to \( \lambda_i \), in which collisions between the ions and the atoms are essential but not dominating, the so-called Knudsen layer. The models [52] and [53] employ to this end a kinetic solution for the Knudsen layer derived in [54] on the basis of a one-dimensional Boltzmann equation for ions under the assumption of symmetric charge exchange ion-atom collisions with constant collision frequency.

There is no necessity to specially introduce a solution describing the Knudsen layer in the framework of the model [37]. The multifluid solution [46] for the ionization layer is employed in [37] not only in the case \( d \lesssim \lambda_i \) but also in the case \( d \gg \lambda_i \). In the latter case, no separate solution describing the Knudsen layer is required, since the multifluid solution describes both the ionization layer itself (which is dominated by diffusion) and the Knudsen layer; see the work [46] where these layers are described by, respectively, the outer and inner asymptotic expansions in powers of the small parameter \( \alpha^{-1} \) in section 4. Note that the capability of the approach [37] to treat by means of the same formulae both the case \( d \lesssim \lambda_i \), in which an account of inertia of the atoms and the ions and the momentum exchange due to volume reactions is of primary importance, and the case \( d \gg \lambda_i \), is useful in practice.

Thus, the matching of solutions for the ionization layer and for the space-charge sheath in the model [37] in the case \( d \gg \lambda_i \) and in the models [52, 53] occurs through the intermediate Knudsen layer; however, descriptions of the Knudsen layer differ, the one based on multifluid equations versus the kinetic one. A rigorous kinetic derivation of multifluid equations can be given only in a special case [55]; therefore, the multifluid model cannot be considered as rigorous as the Boltzmann equation. On the other hand, it was shown in [43] in connection with the investigation of a low-pressure glow discharge column that the results of the (multi) fluid model give a smooth transition from a good agreement with the free-fall model to identically the collision-dominated model; thus, the multifluid model spans the whole range of conditions from collision-free to collision-dominated. The multifluid model is understood to ensure sufficient accuracy in applications and is widely used in modelling of low-pressure gas discharges, e.g. [44] and references therein.

In practical terms, the only result of the difference in descriptions of the Knudsen layer which is passed on to the overall scheme of evaluation of the near-cathode layer is the difference in \( v_i \), the ion speed at the edge of the space-charge layer.

Figure 5. The ion current density to the cathode surface, ratio \( \alpha \) of the ionization length to the mean free path for collisions of an ion with neutral atoms and the ionization degree \( \omega \) on the plasma side of the ionization layer versus the electron temperature. Dashed: the ion current density evaluated in the diffusion approximation. Ar plasma, \( p = 1 \) bar, \( T_i = 4000 \) K.
sheath. In [37], \( v_{is} \) is given by the formula
\[
v_{is} = \sqrt{\frac{k(T_e + T_0)}{m_i}}.
\]
(8)

Similar formulae from [52, 53] may be written as
\[
v_{is} = \sqrt{\frac{kT_e}{m_i}} \left( c_1 + \frac{T_0}{T_e} \right),
\]
(9)

where \( c_1 = 0.9107 \) and \( c_2 = 0.2363 \) [52] or \( c_1 = 0.9107\sqrt{2} = 1.288 \) and \( c_2 = 0.2363\sqrt{2} = 0.3342 \) [53]. If \( T_0 \ll T_e \), formula (8) assumes the form \( v_{is} = \sqrt{kT_e/m_i} \), which is the well-known exact result for the limiting case of cold ions; cold ions are accelerated in a pre-sheath up to the Bohm velocity \( \sqrt{kT_e/m_i} \) [35, 36]. In our view, this feature of formula (8) is important since in the Knudsen layer the electron temperature \( T_e \) is indeed considerably higher than the temperature \( T_0 \) of heavy particles, which is close to the temperature of the cathode surface. On the other hand, differences in the numerical values of \( v_{is} \) given by the above-described formulas are not huge.

There are differences also in the models of the near-cathode space-charge sheath employed in [37, 52, 53]. In [37], a model is used with an ion velocity distribution of a finite width and electrons coming from the plasma. The ion velocity distribution at the edge of the space-charge sheath satisfies with the equality sign the Bohm criterion for this particular sheath model as it should. It should be stressed that the account of a finite width of the ion velocity distribution in the sheath is consistent with the account of finite ion temperature in the ionization layer. Besides, it is indispensable for the Bohm criterion being satisfied: the Bohm velocity for the sheath model with monoenergetic (cold) ions equals \( \sqrt{kT_e/m_i} \) [35, 36], i.e. different from the values given by either of equations (8) or (9).

The Child–Langmuir model of an electron-free sheath formed by ions entering the sheath with a negligible velocity is used in [52]. A model of a sheath with monoenergetic ions and electrons both coming from the plasma and emitted by the cathode surface is used in [53].

Differences in the models of the space-charge sheath, used in [37, 52, 53], cause differences in the expression for the electric field at the cathode surface, which is needed for the evaluation of the electron emission current and is the only equation contributed by the sheath model to the overall evaluation scheme, cf equations (10) of [37], (20) of [52] and (37) of [53]. However, these differences are not very considerable for conditions of practical interest and can hardly cause dramatic differences in the evaluation results.

A comparison of results of calculations of near-cathode layers by means of the model [23, 37, 50] and the model [53] can be found in [53]. The results are quantitatively similar and numerical differences are not dramatic, especially at not too high near-cathode voltages.

From the computational point of view, a calculation of the near-cathode layer in the framework of the above approaches amounts to a solution of transcendental equations. A numerical solution of these equations causes no difficulties except for the following: in some cases, these equations have more than one solution for the same \( T_w \) and \( U \). (For example, for \( U < 10 \) V two roots are sometimes encountered, one with \( T_0 \) below \( 5 \times 10^4 \) K and the other with \( T_0 \) between \( 5 \times 10^4 \) and \( 10^5 \) K.) A solution associated with lower \( T_0 \) is chosen in such cases. It remains unclear whether other solutions are physically meaningful.

The most important result of the calculation of the near-cathode layer is the dependence of the density \( q \) of the net energy flux from the plasma to the cathode surface on the local surface temperature \( T_w \) and the near-cathode voltage drop \( U \), which is evaluated by means of the formulae
\[
q = q_p - q_r, \quad q_p = q_i + q_e - q_{em}, \quad q_r = \varepsilon \sigma T_w^4.
\]
(10)

Here \( q_p \) is the density of the plasma-related energy flux to the cathode surface, \( q_i \) is the density of losses of energy by the cathode surface through radiation, \( q_i \) and \( q_r \) are the densities of energy fluxes delivered to the cathode surface by incident ions and fast plasma electrons, \( q_{em} \) is the density of losses of energy by the cathode surface due to thermionic emission, \( \varepsilon \) is the emissivity of the cathode material and \( \sigma \) is the Stefan–Boltzmann constant. The second formula in equation (10) can also be written in an equivalent form [23, 50]:
\[
q_p = jU - j(A_1 + 3.2kT_c)/e.
\]
(11)

Equation (11) has a distinct physical sense: the plasma-related energy flux to the cathode surface represents the difference between \( jU \) the electrical power deposited per unit area of the near-cathode layer and \( j(A_1 + 3.2kT_e)/e \) the energy transported from the layer into the bulk plasma by the (electron) current, calculated taking into account the energy necessary for extracting electrons from the cathode. The factor 3.2 represents the sum of the coefficient 5/2 accounting for enthalpy transport due to the electric current and of a thermal-diffusion coefficient equal to 0.703. Note that the latter value corresponds to the limiting case of a strongly to fully ionized plasma (e.g. [20, p 410]); this limiting case is appropriate since the corresponding term of equation (11) is the most important at high \( T_w \), where the plasma is fully ionized.

An example of the dependence \( q(T_w, U) \) and its components can be seen in figures 6(a) and (b), respectively (calculations by means of the model [23, 37, 50]). Also shown in figure 6(b) are the electron temperature in the ionization layer, the ionization degree of the plasma at the edge of the ionization layer and parameter \( \alpha \). A detailed discussion of this dependence can be found in [23]; here we only note the following. One can see from figure 6(b) that the dominating mechanisms of energy exchange between the plasma and the cathode in the case \( U = 15 \) V are ion heating and cooling by thermionic emission. At relatively low \( T_w \) the ion heating grows with the increase in \( T_w \) faster than thermionic cooling, so the net energy flux to the cathode surface increases. As \( T_w \) increases and the ionization degree on the plasma side of the ionization layer approaches unity, the increase in the ion current slows down; the increase in thermionic cooling overcomes the increase in ion heating and the net energy flux starts to decrease. Another interpretation is suggested by
exceeds unity in the whole range of $T_w$ equation (11): the electron temperature increases in the range of plasma, its components versus the local surface temperature. W cathode, Ar Figure 6. J. Phys. D: Appl. Phys. 41 voltages. The second maximum is due to a rapid increase in the diffusion approximation, in contrast to the maximum at low temperatures. Eventually the second term on the right-side of equation (11) starts to grow faster than the first term and the dependence of $q$ on $T_w$ falls.

It should be emphasized that the parameter $\omega$ at $U = 15$ V exceeds unity in the whole range of $T_w$ considered. Hence, the non-monotony of the dependence of $q$ on $T_w$ at low voltages may be understood in the framework of the diffusion approximation and is caused by the slowing down of the increase in the ion current that occurs when the plasma becomes strongly and then fully ionized, which can be seen in figure 5.

The dependence of $q$ on $T_w$ in the case $U = 50$ V possesses two maxima. The first maximum originates in the above-discussed deviation of the ion current to the cathode surface from the diffusion value, which can be seen in figure 5, and therefore cannot be understood in the framework of the diffusion approximation, in contrast to the maximum at low voltages. The second maximum is due to a rapid increase in the heating by plasma electrons, which is subsequently overcome by thermionic cooling.

The non-monotony of the dependence of $q$ on $T_w$, which is seen in figure 6(a), is the root reason for the existence of multiple modes of current transfer to thermionic cathodes. The growing section of this dependence, which occurs at relatively low $T_w$ around 3000 K, is potentially unstable: a local increase in the surface temperature will result in an increase in the local energy flux from the plasma; the latter will cause a new increase in the local temperature, etc, i.e. a thermal instability may develop. Note that a growing dependence $q(T_w)$ is unusual from the point of view of conventional heat exchange: if the temperature of a heated surface increases, the net external heat flux to the surface will normally decrease and not increase.

Integrating equation (11) over the arc attachment, one arrives at an equation of integral balance of energy of the near-cathode layer which has been well known since the work [56]:

$$IU = Q_p + I (A_f + 3.2kT_e^*)/e, \quad (12)$$

where $Q_p$ is the total power received by the cathode from the plasma and $T_e^*$ is a weighted average value of the electron temperature in the near-cathode layer. Note that the power $Q_p$ may also be interpreted as the total cathode heat losses and represented as $Q_p = Q_c + Q_t$, where $Q_c$ is the integral power removed by heat conduction from the cathode body to the cooling fluid and $Q_t$ is the integral power irradiated by the cathode surface.

3.2.2. Other models. There are works in which the treatment of the ionization layer is replaced by assumptions which in most cases involve some or other variant of the formula $j_i = 1/4 \sqrt{n_{saha}(8kT_e/\pi m_e)^{1/2}}$, e.g. [3,24,26, 57–60]. In essence, this formula gives the ion chaotic flux evaluated on the plasma side of the ionization layer. Since the ion current to the cathode is generated just in the ionization layer, such approaches do not describe the physics involved. Neither do they offer simplifications in programming. In summary, one should be discouraged from using such approaches, even if in certain cases [59, 60] they can give results not very different from those given by more adequate models.

There are models of high-pressure arc discharges in which the near-cathode space-charge sheath is discarded, e.g. [61–66] and references therein. Obviously, such models can hardly be used for the investigation of the plasma–cathode interaction, and most of the authors do not aim at such an objective. A question to what extent results for the whole system arc–electrodes are affected by disregarding the near-cathode space-charge sheath is still open. Additionally, models which employ the assumption of thermal equilibrium in the whole near-cathode layer do not describe transport of energy from the near-cathode layer to the arc column, which plays an essential role in the energy balance of the near-cathode region as discussed in the next section.

3.2.3. Unified numerical modelling of the near-cathode non-equilibrium layer. The structure shown in table 1 reflects the fact that different physical mechanisms in many cases,
although not always, come into play on different length scales. However, the usage of this structure as a basis for a calculation model inevitably involves quite a bit of intuitive consideration, and this is indeed what we have seen above.

An alternative to approaches relying on an a priori introduction of different sub-layers is to model the whole of the near-cathode non-equilibrium layer in the framework of a single set of equations without simplifying assumptions such as thermal equilibrium, ionization (Saha) equilibrium and quasi-neutrality. After such unified modelling has been completed, one will be able to identify physical mechanisms dominating different regions and thus to pin down appropriate sub-layers. While being highly desirable, a unified numerical modelling of near-cathode non-equilibrium layers in high-pressure arc discharges is quite complex computationally. Note that a two-dimensional modelling of high-pressure arc plasmas without assumptions of thermal or ionization equilibrium has already been reported (e.g. [62,67]); however, the assumption of quasi-neutrality is more difficult to relax, given a high density of charged particles and, consequently, a high degree of quasi-neutrality in the bulk plasma. First results of a unified 1D modelling of near-cathode non-equilibrium plasma layers in high-pressure arc discharges in Ar on the basis of diffusion equations have been reported in [8]; an example of the modelling for Hg is seen in figure 3 above. The results confirm the physical picture discussed in section 2, in particular, a critically important role played by the near-cathode space-charge sheath. It is found, in agreement with conclusions of [68], that the electric power deposited into the near-cathode layer is transported not only to the cathode but also to the arc column. The latter effect is caused by the electron enthalpy transport (see the discussion following equation (11)), which substantially exceeds heat conduction and cannot be described by models employing the assumption of thermal equilibrium, $T_e = T_h$. Modelling results for an atmospheric-pressure Ar arc support results given by the simplified model [23,37,50]; however there is a significant difference for a very high-pressure Hg arc.

### 3.3. Finding temperature distribution in the cathode

Suppose now that the near-cathode plasma layer has been calculated, in particular, the functions $q = q(T_w, U)$, $j = j(T_w, U)$ found, and let us proceed to find the temperature distribution inside the cathode. In this section, a mathematical problem governing the temperature distribution inside the cathodes is discussed. Results of its solution, which provide a complete description of the plasma–cathode interaction, are discussed in the next section.

#### 3.3.1. Formulating the boundary-value problem

First, let us estimate the role of Joule heat production in the cathode body or, in other words, let us compare the electric power dissipated inside the cathode with the power supplied to the cathode surface from the plasma. Assuming that the latter is of the order of $IU$, one can compare the voltage drop inside the cathode with $U$ the voltage drop in the near-cathode plasma. If the cathode operates in the diffuse mode, then the voltage drop inside the cathode may be estimated by order of magnitude as $I\rho_h h/A$, where $\rho_h$ is the electrical resistivity of the cathode material, $h$ is the cathode height and $A$ is the area of the cathode cross section. Electrical resistivity of tungsten increases from $5.3 \times 10^{-8} \text{\,}\Omega\text{\,m}$ at room temperature to $94 \times 10^{-8} \text{\,}\Omega\text{\,m}$ at 3000 K [69]; let us assume $\rho_h = 10^{-6} \text{\,}\Omega\text{\,m}$ for the estimate. Assuming that $I = 2A$, $h = 10$ mm and $A = 1$ mm$^2$, one finds $I\rho_h h/A = 20 \text{\,mV}$ which is much smaller than $U$, so the Joule heating in the cathode body may be safely neglected.

If the cathode operates in a spot mode, there are two components of the voltage drop inside the cathode: the voltage drop in the bulk of the cathode body and the voltage drop in a domain inside the cathode adjacent to the spot in which the current density decreases from values typical for the spot to those of the order of $I/A$. The first component is of the order of $I\rho_h h/A$ and is small as estimated above. The second component may be estimated similarly to how its analogue on the plasma side, $U_{con}$, was estimated in section 3.1 and is of the order of $I\rho_h/2\pi R_m$. One finds that this component does not exceed approximately 30 mV even for $R_m$ as small as 10 $\mu$m, so the Joule heating in the cathode body may be neglected in the spot mode as well.

The temperature distribution in the cathode is governed by the heat conduction equation:

$$\rho c_p \frac{\partial T}{\partial t} = \nabla \cdot (\kappa \nabla T), \quad (13)$$

where $t$ is time and $\rho$, $c_p$ and $\kappa$ are, respectively, mass density, specific heat and thermal conductivity of the cathode material, which are treated as known functions of the local temperature: $\kappa = \kappa(T)$, $\rho = \rho(T)$, $c_p = c_p(T)$. This equation bears no account of heat generation inside the cathode due to the Joule effect as justified above. An account of heat generation inside the cathode due to melting (latent heat) can be introduced into this equation by means of methods employed in simulations of gas–metal arc welding or metal casting (e.g. [70]); for example, the enthalpy method has been used in [71] and the effective specific heat method has been used in [72].

Equation (13) is solved inside the cathode with the following boundary conditions. Let us designate by $T_c$ the base of the cathode and by $T_h$ the part of the cathode surface that is in contact with the arc and the cold gas and may exchange energy with it; see figure 7. Then the boundary conditions read

$$\Gamma_c : \ T = T_c; \quad \Gamma_h : \ \kappa \frac{\partial T}{\partial n} = q(T_w, U), \quad (14)$$

where $T_c$ is the temperature of the base, which is governed by the external cooling arrangement and is treated here as a control parameter, $q(T_w, U)$ is the dependence found by means of one of the methods described in section 3.2 (for example, with the use of equation (10)) and $n$ is the direction locally orthogonal to the cathode surface and directed outside the cathode.

Temporal variations of the thermal regime of the cathode develop on a time scale which is of the order of time of heat propagation in the cathode. Taking $10^{-3}$ $\text{\,m}^2\text{\,s}^{-1}$ as a typical value of thermal diffusivity of electrode materials and 0.1 mm as a lower estimate of an electrode dimension, one finds that this time is $10^{-3}$ s or larger. This time is much greater than the...
in certain cases of multiple solutions, one of these solutions is extremely important feature of this problem is the existence of one solution of iterations. (In mathematical terms: each 3D steady-state simulation defines to which one of the multiple solutions the numerical solution will converge and is therefore of extreme importance.)

Most electrodes of high-pressure arc discharges are shaped in such a way that the arc current is related to the temperature distribution inside the cathode body, and the arc voltage drop must be treated as a control parameter while calculating a steady-state mode in the vicinity of an extreme point of its CVC or a turning point. Thus, it is desirable to impose an additional condition in order to specify the azimuthal boundary distribution on an axially symmetric cathode, one must consider the multiplicity of solutions and should be discussed.

The choice of initial approximation in steady-state simulations determines to which one of the multiple solutions the iterations will converge and is therefore of extreme importance.

Several authors reported computed steady-state spot modes that break off at a certain current. This behaviour normally does not occur in nonlinear dissipative systems: a solution can turn back or join another solution but cannot just disappear. One should suspect therefore that this break-off represents a failure of the numerical method rather than a physical effect, and indeed this follows from the results of subsequent works.

Some researchers believe that the failure of iterations of a steady-state solver to converge indicates the physical instability of the steady state being sought. This idea is in general unjustified: iterative processes in steady-state solvers need not be equivalent to relaxation in time.

Variations of $U$ and $I$ along a given steady-state mode of current transfer may be non-monotonic. In other words, a mode may pass through extreme points of its CVC $U(I)$ and through turning points; examples are shown below. It is helpful to use $I$ or, respectively, $U$ as a control parameter while calculating a steady-state mode in the vicinity of an extreme point of its CVC or a turning point. Thus, it is desirable to provide for the possibility of switching the control parameter while developing a code.

While numerically calculating a steady-state 3D temperature distribution on an axially symmetric cathode, one must impose an additional condition in order to specify the azimuthal position of the 3D spot system and thus to ensure convergence of iterations. (In mathematical terms: each 3D steady-state solution on an axially symmetric cathode represents an element of a continuous set of 3D solutions that exist for the same arc current and are identical to the accuracy of a rotation describing the diffuse mode of current transfer and the others describing different spot modes.)

3.3.2. Hints for numerical solution. From the computational point of view, the above-formulated problem is not very difficult nowadays. It has been solved by different authors and in different ways: by means of home-made codes [23, 58, 73], with the use of commercial finite-element platforms ANSYS [57, 60, 71, 74] and COMSOL Multiphysics (e.g. [75–77]). However, there are delicate points in the numerical solution that originate in the multiplicity of solutions and should be discussed.

The choice of initial approximation in steady-state simulations determines to which one of the multiple solutions the iterations will converge and is therefore of extreme importance.

Several authors reported computed steady-state spot modes that break off at a certain current. This behaviour normally does not occur in nonlinear dissipative systems: a solution can turn back or join another solution but cannot just disappear. One should suspect therefore that this break-off represents a failure of the numerical method rather than a physical effect, and indeed this follows from the results of subsequent works.

Some researchers believe that the failure of iterations of a steady-state solver to converge indicates the physical instability of the steady state being sought. This idea is in general unjustified: iterative processes in steady-state solvers need not be equivalent to relaxation in time.

Variations of $U$ and $I$ along a given steady-state mode of current transfer may be non-monotonic. In other words, a mode may pass through extreme points of its CVC $U(I)$ and through turning points; examples are shown below. It is helpful to use $I$ or, respectively, $U$ as a control parameter while calculating a steady-state mode in the vicinity of an extreme point of its CVC or a turning point. Thus, it is desirable to provide for the possibility of switching the control parameter while developing a code.

While numerically calculating a steady-state 3D temperature distribution on an axially symmetric cathode, one must impose an additional condition in order to specify the azimuthal position of the 3D spot system and thus to ensure convergence of iterations. (In mathematical terms: each 3D steady-state solution on an axially symmetric cathode represents an element of a continuous set of 3D solutions that exist for the same arc current and are identical to the accuracy of a rotation describing the diffuse mode of current transfer and the others describing different spot modes.)

Most electrodes of high-pressure arc discharges are shaped more or less rotationally symmetric; therefore, let us restrict the consideration to axially symmetric cathodes. Let us introduce cylindrical coordinates $(r, \varphi, z)$ with the origin at the centre of the front surface of the cathode and the $z$-axis directed along the axis of symmetry inside the cathode, as shown in figure 7. Then the boundary distributions associated with the diffuse mode are axially symmetric, $T = T(r, z, t)$. Temperature distributions associated with spot modes can be both axially symmetric (a spot at the centre of the front surface of the cathode, a ring spot, etc) or 3D (three-dimensional; an off-centre spot or a system of two or more spots), $T = T(z, \varphi, z, t)$.

3.3.2. Hints for numerical solution. From the computational point of view, the above-formulated problem is not very difficult nowadays. It has been solved by different authors and in different ways: by means of home-made codes [23, 58, 73], with the use of commercial finite-element platforms ANSYS [57, 60, 71, 74] and COMSOL Multiphysics (e.g. [75–77]). However, there are delicate points in the numerical solution that originate in the multiplicity of solutions and should be discussed.

The choice of initial approximation in steady-state simulations determines to which one of the multiple solutions the iterations will converge and is therefore of extreme importance.

Several authors reported computed steady-state spot modes that break off at a certain current. This behaviour normally does not occur in nonlinear dissipative systems: a solution can turn back or join another solution but cannot just disappear. One should suspect therefore that this break-off represents a failure of the numerical method rather than a physical effect, and indeed this follows from the results of subsequent works.

Some researchers believe that the failure of iterations of a steady-state solver to converge indicates the physical instability of the steady state being sought. This idea is in general unjustified: iterative processes in steady-state solvers need not be equivalent to relaxation in time.

Variations of $U$ and $I$ along a given steady-state mode of current transfer may be non-monotonic. In other words, a mode may pass through extreme points of its CVC $U(I)$ and through turning points; examples are shown below. It is helpful to use $I$ or, respectively, $U$ as a control parameter while calculating a steady-state mode in the vicinity of an extreme point of its CVC or a turning point. Thus, it is desirable to provide for the possibility of switching the control parameter while developing a code.

While numerically calculating a steady-state 3D temperature distribution on an axially symmetric cathode, one must impose an additional condition in order to specify the azimuthal position of the 3D spot system and thus to ensure convergence of iterations. (In mathematical terms: each 3D steady-state solution on an axially symmetric cathode represents an element of a continuous set of 3D solutions that exist for the same arc current and are identical to the accuracy of a rotation describing the diffuse mode of current transfer and the others describing different spot modes.)

Most electrodes of high-pressure arc discharges are shaped more or less rotationally symmetric; therefore, let us restrict the consideration to axially symmetric cathodes. Let us introduce cylindrical coordinates $(r, \varphi, z)$ with the origin at the centre of the front surface of the cathode and the $z$-axis directed along the axis of symmetry inside the cathode, as shown in figure 7. Then temperature distributions associated with the diffuse mode are axially symmetric, $T = T(r, z, t)$. Temperature distributions associated with spot modes can be both axially symmetric (a spot at the centre of the front surface of the cathode, a ring spot, etc) or 3D (three-dimensional; an off-centre spot or a system of two or more spots), $T = T(z, \varphi, z, t)$.
about the cathode axis. Hence, one must impose an additional condition in order to single out one solution from this set, otherwise the problem will be ill-stated and iterations will not converge.) Assuming that the 3D temperature distribution \( T(r, \phi, z) \) being sought possesses planar symmetry, one can do so by means of restricting the calculation domain to half the cathode, say, to \( 0 < \phi < \pi \), and imposing the symmetry condition, \( \partial T / \partial \phi = 0 \), at the plane \( \phi = 0, \phi = \pi \). If, for example, the solution being sought describes a mode with one spot, then the hottest point of the spot belongs either to the half-plane \( \phi = 0 \) or to \( \phi = \pi \), thus the azimuthal position of the spot is fixed. Of course, the solution is still not unique: there are two solutions for the same arc current, one with the hottest point in the half-plane \( \phi = 0 \) and the other with the hottest point in the half-plane \( \phi = \pi \). However, these two solutions are not close to each other; therefore, a code may be forced to converge to a desired solution by means of an appropriate choice of initial approximation. This approach has been used in [75, 76]. Of course, this approach allows one not only to fix the azimuthal position of a 3D spot system but also to save RAM and CPU time.

The same question arises in the numerical calculation of a non-stationary 3D temperature distribution on an axially symmetric cathode if the temperature distribution at the initial moment (the initial condition) is axially symmetric, as is the case in the modelling of evolution of the diffuse mode on a rod cathode into a mode with a non-stationary spot [57, 60, 72]. In [57, 60, 72], this question was resolved in the same way as described above, i.e. by restricting the calculation domain to half the cathode. Applicability of this approach in the non-stationary case is limited by the assumption that the 3D temperature distribution \( T(r, \phi, z, t) \) being sought possesses planar symmetry at each \( t \) and the plane of symmetry is the same for all \( t \).

It remains unclear whether problem (13)–(15) for axially symmetric cathodes admits 3D solutions that cannot be found by means of this approach, such as solutions describing steady-state spots that do not possess planar symmetry or solutions describing evolution of the diffuse mode into a mode with a rotating spot. Note that a bifurcation of steady-state spot modes that breaks planar symmetry has been found in [78].

3.3.3. Quasi-1D analytical solution. In the particular case of a thin rod cathode operating in the dc mode, an approximate analytical solution to problem (13), (14) may be found by means of the quasi-1D approximation. The temperature distribution inside the cathode in the framework of this approximation is assumed to be a function of \( z \), \( T = T(z) \) and is governed by the ordinary differential equation

\[
\frac{\pi R^2}{2} \frac{d}{dz} \left( \kappa \frac{dT}{dz} \right) + 2\pi R q(T, U) = 0, \tag{16}
\]

where \( R \) is the cathode radius. Equation (16) has a clear physical sense and its direct derivation is straightforward. Alternatively, this equation may be derived by integrating equation (13) over the cross section of the cathode, invoking the divergence theorem in two dimensions, substituting equation (14) and neglecting variations of the temperature in the cross section. As before, \( q(T, U) \) in equation (16) represents the dependence of the density of the net energy flux from the plasma to the cathode surface on the local surface temperature and the near-cathode voltage drop; the difference is that the local surface temperature \( T_w \) is replaced in the quasi-1D approximation with the local value of the function \( T(z) \).

Equation (16) does not involve \( z \); hence its order may be reduced. This can be done by multiplying each term by \( \kappa (dT/dz) \) and integrating over \( z \):

\[
\frac{\pi R^2}{2} \left( \kappa \frac{dT}{dz} \right)^2 + 2\pi R \int_{L_z}^T \kappa q(T, U) \, dT = \frac{Q_c^2}{2\pi R^2}. \tag{17}
\]

Given that \( \kappa = \kappa(T) \), the integral on the left-hand side may be evaluated as a function of \( T \), after which equation (17) may be solved for \( dT/dz \). The resulting equation again does not involve \( z \) and may be integrated once more:

\[
\int_{L_z}^T \frac{\kappa}{\sqrt{F}} \, dT = h - z, \tag{18}
\]

where

\[
F = \left( \frac{Q_c}{\pi R^2} \right)^2 - \frac{4}{R} \int_{L_z}^T \kappa q(T, U) \, dT. \tag{19}
\]

Equation (18) represents a desired solution describing temperature variation along a thin rod cathode. Note that the assumption \( T = T(z) \), being a good approximation in the bulk of a thin cathode, does not hold in the vicinity of the tip since heat transfer there is multidimensional, especially if the cathode operates in the spot mode. Therefore, the value of the function \( T(z) \) at \( z = 0 \) should be considered as an extrapolated value of the average temperature of the front surface of the cathode.

A further simplification can be achieved by assuming that the power input from the plasma to the cathode is localized in the vicinity of the tip while the rest of the cathode surface loses energy through radiation. This amounts to setting \( q = -\dot{q}_1 = -\varepsilon \sigma T^4 \) on the cathode surface except in the vicinity of the tip. Assuming further that the product \( \kappa \varepsilon \) does not vary much in the temperature range being considered, one can evaluate the integral on the left-hand side of equation (17) analytically and obtain

\[
\frac{\pi R^2}{2} \left( \kappa \frac{dT}{dz} \right)^2 = \frac{Q_c^2}{2\pi R^2} + \frac{2\pi}{5} R \kappa \varepsilon \sigma (T^5 - T^5_w). \tag{20}
\]

Applying this equation at \( z = 0 \), one arrives at a relationship between the average temperature of the front surface of the cathode, \( T(0) \), and the total power \( Q_p \) received by the cathode from the plasma:

\[
Q_p^2 = Q_c^2 + \frac{4\pi^2}{5} R^3 \kappa \varepsilon \sigma [T^5(0) - T^5_w]. \tag{21}
\]

Such quasi-1D solutions have been known since the work [24]. Nowadays multidimensional codes are widely available, also free of charge [79], and a theoretical value of such solutions is limited. On the other hand, such solutions are useful in the analysis of measured cathode temperature distributions, e.g. [80, 81].
3.4. Results on plasma–cathode interaction

3.4.1. Steady-state modes of current transfer. The assumption of the existence of multiple solutions to the nonlinear boundary-value problem governing the temperature distribution in the body of a thermionic arc cathode, with different solutions describing different modes of current transfer, was formulated in [24]. The existence of such multiple solutions was proved in [28] by means of bifurcation analysis for an idealized model of the cathode, a right cylinder with a thermally and electrically insulated lateral surface. The diffuse mode of current transfer is described by a 1D analytical solution in the framework of this model, which is why this model is convenient for qualitative analysis. Also in the work [28] a general pattern of various steady-state modes and their stability was suggested on the basis of bifurcation analysis and general trends characteristic for nonlinear dissipative systems.

The first actually calculated multiple solutions were axially symmetric and described the diffuse mode and a mode with a spot at the centre of the front surface of the cathode [74,82,83]. A detailed numerical investigation of these modes was reported in [23] and [73], respectively. In [75], simulations were reported of a 3D steady-state spot attached to a rounded edge of a cylindrical cathode. A whole ‘zoo’ of very diverse modes was revealed by a 3D modelling [76].

A pattern of different modes on a rod cathode is illustrated by figure 8 (data taken from [73, 76]). This figure refers to a cylindrical cathode of radius $R = 2\,\text{mm}$ and height $h = 10\,\text{mm}$, a geometry convenient for illustrative purposes. Two of the modes shown in this figure are axially symmetric, one of these modes being the diffuse mode and the other being the first axially symmetric spot mode. The diffuse mode exists at all currents and its CVC has two branches, one falling and the other rising, separated by a point of minimum [73]; only the falling branch is seen in figure 8. The first axially symmetric spot mode embraces steady states with a spot at the centre of the front surface of the cathode and exists in a limited current range, $I < I_t = 145\,\text{A}$. The upper limit of the existence range, $I = I_t$, may be called the turning point of this mode. One can say that the first axially symmetric spot mode has two branches, a low-voltage one and a high-voltage one, separated by the turning point. The 3D modes shown in figure 8 are modes with one, two, three or four spots at the edge of the front surface, and a mode with a spot at the centre and two spots at the edge. The 3D modes also exist in limited current ranges.

Under conditions of figure 8, a value of the arc current exists for each 3D mode such that at this current the cathode temperature distribution associated with this mode turns axially symmetric and the corresponding state belongs to one of the axially symmetric modes. This phenomenon, called bifurcation, or branching, of solutions, is well known in mathematical physics and frequently occurs in nonlinear systems admitting states of different symmetries. The 3D modes with spots at the edge branch off from the diffuse mode, and the mode with a spot at the centre and two spots at the edge branches off from the first axially symmetric spot mode. Points at which this branching occurs coincide with bifurcation points predicted by bifurcation theory [50], which are shown in figure 8 by circles.

It should be stressed that an intersection of CVCs does not necessarily mean a bifurcation: not only must near-cathode voltages of a 3D mode and an axially symmetric mode be the same at a bifurcation point but also temperature distributions. For example, the CVC of the low-voltage branch of the first axially symmetric spot mode in figure 8 is very close to the CVC of the diffuse mode; however, thermal regimes of these two modes are different, hence no bifurcation occurs.

At large distances from the bifurcation points, all spot modes tend to the region of small currents and high voltages,
i.e. approach the axis of voltages. General considerations explaining this behaviour are given in [28]. In fact, the whole pattern of different steady-state modes seen in figure 8 may be understood on the basis of general trends characteristic for nonlinear dissipative systems [28].

In the following, the 3D mode with a spot at the edge of the front surface of the cathode is termed the first 3D spot mode. The point at which this mode branches off from the diffuse mode is termed the first bifurcation point and the corresponding arc current is designated $I_1$. The section of the first 3D spot mode comprised between the first bifurcation point and the turning point is termed the low-voltage branch, and the section beyond the turning point is termed the high-voltage branch. In the case of a wide cathode shown in figure 8, the first bifurcation point and the turning point are not distant and the low-voltage branch is not well pronounced, i.e. the current range $I_1 \leq I \leq I_2$ in which this branch exists is rather narrow. As the cathode becomes thinner, the current range $[I_1, I_2]$ expands and the low-voltage branch becomes better pronounced. An example can be seen in figure 9: the first bifurcation point moved into the range of very high voltages, well in excess of 100 V, and both branches of the first 3D spot mode exist in a wide range of conditions. Note that the CVCs of the two branches of the first 3D spot mode in figure 9 intersect, so the terms ‘low-voltage’ and ‘high-voltage’ in such cases refer to the branches of the spot mode which manifest lower or, respectively, higher values of $U$ in the vicinity of the turning point.

The CVC shown in figure 9 by the solid and dashed lines has been calculated by means of the code [76]. (There is not enough information on the cathode geometry in [58], so it is simulated as a rod with a fixed temperature of 1000 K at 10 mm from its tip.) $T_{tip}$ is evaluated by averaging the calculated surface temperature distribution over the flat section of the front surface and the rounded rim of the cathode (table 2) or over the flat front surface of the cathode (tables 3 and 4); $T_{c}^{*}$ the average electron temperature is evaluated in terms of the calculated $U$ and $Q_p$ by means of equation (12).

Although conditions to which tables 2–4 refer are quite different, the average temperature of the tip in the diffuse mode does not vary too much and is not very different from 3000 K, a feature which is characteristic of the diffuse discharge on a tungsten cathode. If the discharge under conditions of table 2 switches from the diffuse mode to the low-voltage branch of the first 3D spot mode, the tip temperature decreases by 62 K and the total cathode heat losses by 8%. The decrease is much stronger if the discharge switches from the diffuse mode to the high-voltage branch: $T_{tip}$ decreases by 457 K and $Q_p$ by 33%. A similar trend is seen in tables 3 and 4: if the discharge switches to the low-voltage branch, $T_{tip}$ decreases by 77 K and $Q_p$ by 11% under conditions of table 3 and by 46 K and 9% under conditions of table 4; if the discharge switches to the high-voltage branch, $T_{tip}$ decreases by 261 K and $Q_p$ by 26% under conditions of table 3 and by 310 K and 32% under conditions of table 4. The near-cathode voltage drop on the low-voltage branch under conditions of table 2 is virtually the same as that in the diffuse mode, the voltage drop on the high-voltage branch is a bit higher. Variations of $U$ from one mode to the others in table 3 are smaller than those in table 2 and variations in table 4 are still smaller. The average electron temperature on the high-voltage branch is considerably higher than that in the diffuse mode and on the low-voltage branch.

Figure 9. Current–voltage characteristics. Rod W cathode, $R = 0.75$ mm, $h = 20$ mm, rounding 100 μm, $T_c = 300$ K, Ar plasma, $p = 2.6$ bar.

The non-uniformity of the cathode surface temperature distribution on the high-voltage branch is considerably stronger than that at the same arc current in the diffuse mode and on the low-voltage branch. Since the current density is a very strong function of the local surface temperature, the same arc current in the case of a less uniform temperature distribution corresponds to a lower average temperature. One should expect therefore that the average temperature $T_{tip}$ of the cathode tip and, consequently, the total cathode heat losses $Q_p$ are considerably lower on the high-voltage branch than the corresponding values at the same arc current in the diffuse mode and on the low-voltage branch.

This is indeed the case, which can be seen from tables 2–4. Results of numerical modelling for conditions of figure 9 at $I = 4$ A are shown in table 2 and results for conditions reproducing experiment [58] are shown in tables 3 and 4. The non-uniformity of the cathode surface temperature distribution on the high-voltage branch is considerably higher than that in the diffuse mode and on the low-voltage branch.

...
A linear analysis of stability of steady-state modes of current transfer to thermionic cathodes of high-pressure arc discharges on the basis of the model of nonlinear surface heating was given in [78, 85]. In the framework of this analysis, instability is a result of competition between a positive feedback, which is present on the growing section of the dependence of the density \( q \) of energy flux from the plasma on \( T_{\text{wall}} \), the local surface temperature (see the discussion of figure 6(a)), and heat conduction in the cathode body, which tends to smooth out perturbations, i.e. produces a stabilizing effect.

A pattern of stability that has been established for a current-controlled arc on a rod cathode on the basis of an analytical treatment [85] and numerical results [78] is as follows. In all cases, the spectrum of perturbations is real, which means a monotonic development or decay of perturbations. All modes other than the diffuse mode and the first 3D mode are unstable. (This includes modes with more than one spot and the axially symmetric mode with a spot at the centre.) The pattern of stability of the diffuse mode and the first 3D mode is shown in figure 10 (taken from [85]). The diffuse mode is stable beyond the first bifurcation point, i.e. at \( I > I_1 \), and unstable at lower currents. The low-voltage branch of the first 3D spot mode is unstable; the high-voltage branch is stable. The transition between the diffuse mode and the high-voltage branch of the first 3D spot mode is non-stationary and accompanied by hysteresis as shown by arrows. The first 3D spot mode on an axially symmetric cathode is neutrally stable against infinitesimal rotations.

Similarly to the pattern of different steady-state modes shown in figure 8, the above-described pattern of their stability has been predicted and may be understood on the basis of general trends characteristic for nonlinear dissipative systems [28].

The arc current corresponding to the first bifurcation point, \( I = I_1 \), represents the limit of stability of the diffuse mode. A method and results of its calculation are described in [50]. The stability limit is much more sensitive to variations of control parameters than characteristics of the diffuse mode, the strongest effect being produced by variations of cathode dimensions and the work function of the cathode material.

In addition to theoretical CVCs of the first spot mode, also shown in figure 9 are the results of careful measurements of the near-cathode voltage in the spot mode reported shown in figure 8 are stable and which are not. Unstable solutions must be discarded and stable solutions must be compared with the experiment.

A pattern of stability that has been established for a current-controlled arc on a rod cathode on the basis of the model of nonlinear surface heating was given in [78, 85]. In the framework of this analysis, instability is a result of competition between a positive feedback, which is present on the growing section of the dependence of the density \( q \) of energy flux from the plasma on \( T_{\text{wall}} \), the local surface temperature (see the discussion of figure 6(a)), and heat conduction in the cathode body, which tends to smooth out perturbations, i.e. produces a stabilizing effect.

A pattern of stability that has been established for a current-controlled arc on a rod cathode on the basis of an analytical treatment [85] and numerical results [78] is as follows. In all cases, the spectrum of perturbations is real, which means a monotonic development or decay of perturbations. All modes other than the diffuse mode and the first 3D mode are unstable. (This includes modes with more than one spot and the axially symmetric mode with a spot at the centre.) The pattern of stability of the diffuse mode and the first 3D mode is shown in figure 10 (taken from [85]). The diffuse mode is stable beyond the first bifurcation point, i.e. at \( I > I_1 \), and unstable at lower currents. The low-voltage branch of the first 3D spot mode is unstable; the high-voltage branch is stable. The transition between the diffuse mode and the high-voltage branch of the first 3D spot mode is non-stationary and accompanied by hysteresis as shown by arrows. The first 3D spot mode on an axially symmetric cathode is neutrally stable against infinitesimal rotations.

Similarly to the pattern of different steady-state modes shown in figure 8, the above-described pattern of their stability has been predicted and may be understood on the
in [75]. Without comparing at the moment the theory and the experiment, we note the following. The experiments are performed in a limited current range, \( I \leq 6 \, \text{A} \). Both the diffuse mode and the high-voltage branch of the first spot mode are stable in this range. Hence, no reproducible transition between the diffuse and spot modes can occur if the experiment is well-controlled and quasi-stationary: a mode which occurs immediately after the ignition of the discharge will be maintained during the whole run; if a mode change is systematically observed in such an experiment under quasi-stationary conditions, it means that the experiment is not well controlled. This conclusion conforms to the general trend that the transition between the diffuse and spot modes is difficult to reproduce in the experiment.

The authors [75] assumed, on the basis of arguments stemming from Steenbeck’s principle of minimum voltage for discharges with a fixed current, that a mode with the lowest near-cathode voltage drop is the preferred one. The above-described results of the accurate linear stability theory [78,85] have not confirmed this assumption: stability of different modes is unrelated to which mode operates at a lower voltage drop; e.g. figure 10. Note that the occurrence of hysteresis in experiments on transition between the spot and diffuse modes is by itself an unambiguous indication of incorrectness of this assumption.

3.4.3. Comparison with the experiment. There is a variety of methods of experimental investigation of plasma–cathode interaction in high-pressure arc discharges (e.g. [86,87]), such as spectroscopic measurements of plasma parameters in the near-cathode region, determination of the near-cathode voltage drop by means of electrostatic probe measurements, study of changes in modes of current transfer to the cathode by means of visual observation and/or measurement of intensity of light emitted by the cathodic part of the arc, correlated with variations of the arc voltage, and pyrometric measurements of the temperature of different points of the cathode surface. Radiation losses \( Q_r \) can be derived from pyrometrically measured temperature distribution of the cathode surface with the use of a tabulated emissivity of the cathode material. Calorimetric measurements (e.g. [81]) represent a direct way of determining the heat conduction losses \( Q_c \); unfortunately, this method is not simple, so in most cases \( Q_c \) is derived, with the use of the quasi-1D solution discussed in section 3.3.3, from the measured cathode surface temperature distribution. Electrostatic probe measurements [88] represent a direct way of determining the near-cathode voltage drop; unfortunately, this method is also not simple, so in most cases \( U \) is determined in terms of the derived total cathode heat losses \( Q_p \) by means of equation (12) with \( T^*_e \) somehow estimated or the term involving \( T^*_e \) dropped altogether.

According to the modelling results shown in table 2, switching from the diffuse mode to (the high-voltage branch of) the spot mode in these conditions, while being accompanied by an 8% increase in the near-cathode voltage, is accompanied also by a 33% decrease in the total cathode heat losses. In other words, variations of \( U \) and \( Q_p \) in this case are unrelated. Similar trends can be seen in tables 3 and 4. These examples demonstrate limitations of the above-mentioned method of determination of the near-cathode voltage drop from the cathode heat losses by means of equation (12): this method is acceptable in the case of the diffuse mode, where \( T^*_e \) is usually not dramatically different from \( 10^4 \, \text{K} \); however, it becomes problematic in the case of the spot mode because of unknown and high values of \( T^*_e \).

A considerable amount of information on plasma parameters in the near-cathode region has been obtained by spectroscopic measurements performed by different research groups, e.g. [89–91] and references therein. Unfortunately, available realizations of the model of nonlinear surface heating are insufficient for a quantitative comparison since they treat the near-cathode non-equilibrium layer in too simplified a way (by means of separate treatments of different sub-layers and with the use of an equation of integral balance for the electron temperature, as described in section 3.2.1). However, the conclusion of the work [91] is of particular interest for theory already at the present stage. Using different methods, the authors [91] found that there is a region in front of the cathode approximately 500 \( \mu \text{m} \) in dimension which shows noticeable deviations from LTE in spite of the high electron densities of about \( 2 \times 10^{23} \, \text{m}^{-3} \). The authors pointed out that this deviation cannot be caused by losses of energy by radiation covered by Griem’s criterion. Neither can it be explained by the inward drift of cold atoms, which is the supposed source for deviations from LTE in high-current arcs. The authors concluded that the physical processes which drive the plasma away from LTE in the vicinity of the cathode remain to be clarified.

In our view, the physical process which drives the plasma away from LTE in the vicinity of the cathode is clear: this is heating of the electron gas through acceleration in the space-charge sheath of the electrons emitted by the cathode. This process represents a key element of the physical picture on which the model of nonlinear heating is based, as discussed in detail in section 2.2. Thus, the measurements [91] confirm, on a qualitative level, the basis of the model of nonlinear heating.

Detailed electrical and thermal measurements in arcs operating in inert gases were performed during the last decade by Mentel and co-workers in Bochum. An extensive comparison has convincingly validated the model of nonlinear surface heating for the steady-state diffuse mode, e.g. [23,75,76,92]. One should also mention an elegant verification of the model by means of an experiment with an additional heating of the cathode by laser light [59]. There is good qualitative agreement between the theory and the experiment on the steady-state spot mode. However, quantitative agreement is not quite satisfactory in some cases, which is not surprising given that the physics of the spot mode is considerably more complex than that of the diffuse mode. There is good agreement in figure 9 between the measured CVC of the spot mode and both sets of theoretical data. However, the authors [75] found that the value of electron temperature predicted theoretically for the high-voltage branch of the spot mode is about twice the temperature obtained by spectroscopic measurements [93] in the spot mode at \( I = 4 \, \text{A} \), which is about \( 2.5 \times 10^4 \, \text{K} \). On the other hand, the temperature of \( 2.5 \times 10^4 \, \text{K} \) conforms to \( T^*_e \) predicted for...
the low-voltage branch. This, along with the fact that the cathode surface temperature calculated for the high-voltage branch is far beyond the melting temperature of tungsten, led the authors [75] to assume that it is the low-voltage branch of the spot mode that is observed in the experiment. However, this assumption was not confirmed by the above-discussed stability theory [78, 85]. Note that values of $T_e$ on the high- and low-voltage branches calculated for the same conditions by means of the code [76] and given in table 2 are substantially lower than the above-mentioned values predicted by the code [75]. On the other hand, a significantly higher electron temperature of about $3.6 \times 10^5\,\text{K}$ has been measured in [94] for conditions that apparently are not very different. Clearly, this matter requires further work.

As discussed at the end of section 3.4.1, the average electron temperature on the high-voltage branch of the spot mode is considerably higher than that in the diffuse mode. This means that there is a strong flux of energy from the near-cathode layer into the bulk plasma in the spot mode transported by the electron current, cf equation (12). This is a likely reason for the appearance in the experiment (e.g. [95]) of an area of intense plasma radiation in front of the cathode attached to the spot.

The theoretical near-cathode voltage drop on the high-voltage branch of the first spot mode is usually slightly higher than that in the diffuse mode, as discussed at the end of section 3.4.1. In the experiment, the arc voltage in the spot mode is slightly lower than that in the diffuse mode (e.g. figure 12 of [95]). This may be explained by a decrease in the resistance of the part of the bulk plasma adjacent to the cathode due to the above-mentioned strong flux of energy from the near-cathode layer into the bulk plasma that occurs in the spot mode.

It is indicated in [75] that under conditions of figure 9 the global cathode temperature in the spot mode is about $150\,\text{K}$ lower than that in the diffuse mode, an appreciably smaller variation than the theoretically predicted difference between values of $T_{tip}$ in the diffuse mode and the (high-voltage branch of the first 3D) spot mode in table 2. Experimental data for conditions of tables 3 and 4 taken from figure 2 of [58] are as follows: at $I = 0.7\,\text{A}$, the discharge operated in the diffuse mode with $T_{tip} = 3192\,\text{K}$ and $Q_p = 8.2\,\text{W}$; at $I = 0.5\,\text{A}$, the discharge operated in the spot mode with $T_{tip} = 2677\,\text{K}$ and $Q_p = 5.0\,\text{W}$. This time there is reasonably good agreement between the experiment and the theory not only on the diffuse mode but also on the spot mode.

The theoretical pattern of stability of stationary modes discussed in section 3.4.2 conforms to trends observed in the experiment: the diffuse mode is observed at high currents and a mode with a spot at the edge of the cathode at low currents; the transition between the diffuse mode and the mode with a spot at the edge cannot be realized in a quasi-stationary way and represents a non-stationary process; this transition is monotonic (for example, the luminosity of the cathode surface during this transition varies in time monotonically) and accompanied by hysteresis; patterns with more than one spot are not normally observed; a stationary arc spot attached to the centre of the (flat) front surface of a rod cathode is not observed in the experiment (e.g. [75]) unless stabilized by an axial gas flow; spots sometimes change their angular position, however no systematic rotation is observed; the effect produced on the limit of stability of the diffuse mode by variations of control parameters (cathode dimensions, work function of the cathode material, plasma-producing gas and its pressure) conforms to the trends observed experimentally [50]. However, a transition between stationary diffuse and spot modes observed in [75] does not conform to the stability theory discussed in the preceding section. On the other hand, there is good agreement between the theory and the experiment on a transition from a non-stationary diffuse mode to a non-stationary spot mode; see the beginning of the next section.

In summary, the model of nonlinear surface heating has gone through a detailed validation by the experiment as far as the diffuse mode is concerned; however, further work is required on the spot mode, both in the experiment and in the theory. One direction is to obtain more detailed information on the electron temperature in the near-cathode region by means of spectroscopic measurements and to compare the results with those predicted by the model of nonlinear surface heating. It should be stressed that there is not much sense in comparing results of spatially resolved measurements with theoretical values of $T_e$ obtained from an equation of integral balance; hence, the comparison will be meaningful only after the simplified treatment of the near-cathode non-equilibrium layer in the model of nonlinear surface heating has been replaced by a unified modelling described in section 3.2.3. Another direction is to perform more detailed pyrometric measurements of temperature distributions over the cathode surface at the same or close values of arc current in the diffuse and spot modes and to compare the results with those predicted by the model of nonlinear surface heating. Note that there is not much sense in first deriving $Q_r, Q_c$ and $T_{tip}$ from a measured temperature distribution and then comparing these quantities with the calculated values; a direct comparison of measured and calculated temperature distributions would be simpler and methodologically correct.

The most reliable experimental data have been obtained in experiments with arcs in inert gases at the pressure of a few bars. A question arises whether these data are representative of cathodes of real HID lamps, most of which operate under the pressure of tens of bars with mercury being a dominant species of the filling. In this connection, let us compare distributions of parameters in the near-cathode non-equilibrium layer shown in figure 3, which refers to the mercury plasma under the pressure of 30 bar, with the distributions shown in figure 11, which have been obtained for the argon plasma under the pressure of 1 bar by means of unified modelling [8]. Both figures refer to the same current density of $10^7\,\text{A}\,\text{m}^{-2}$. The value of $3300\,\text{K}$ was chosen for the temperature of the cathode surface in the calculations shown in figure 11, in order that the near-cathode voltage drop in the two cases be approximately the same (around $38\,\text{V}$). While the calculations shown in figure 3 have been performed for a planar cathode under the assumption of plane-parallel current transfer through the non-equilibrium layer, the calculations shown in figure 11 have been performed for a spherical cathode of radius $1\,\text{mm}$ under the assumption
of spherically symmetric current transfer through the non-equilibrium layer. One can conclude that the physics of the near-cathode layer in the two cases is similar. Hence, the answer to the above-stated question is positive: yes, experimental data obtained in experiments with arcs in inert gases at the pressure of a few bars are representative of cathodes of HID lamps.

3.4.4. Other topics. Works [57, 60, 71, 72, 96] concern modelling of non-stationary spots induced on arc cathodes by rapid variations of the arc current; a topic important for practice in connection with occurrence of spots on cathodes of ac HID lamps after the change in polarity. (One can also mention the experimental work [97], in which measurements are reported of the electrode sheath voltage in arcs operated in noble gases with sinusoidal and switched dc currents.) Spots induced by a rectified sine current wave were investigated by means of a 2D numerical modelling in [71] and experimentally in [96]. In [57, 60, 72], results were reported of a 3D modelling and an experimental investigation of spots induced by a current step (after a few seconds at a constant arc current, the current is instantly increased to a higher value which is maintained constant until the transition process terminates). In a general case, temporal evolution of a stable diffuse mode perturbed by a rapid increase in the arc current can follow one of the three scenarios [72]: (a) a transition between two stable diffuse states via intermediate diffuse states, (b) a transition from a stable diffuse state to a stable state with a stationary spot and (c) a transition between two stable diffuse states via intermediate states with a transient spot. In the experiments [57, 60, 72] the first and third scenarios have been observed. The breakdown of a non-stationary diffuse mode and the formation of transient spots were found to be easier to reproduce in the experiment than transitions between stationary diffuse and spot modes. Good agreement between the theory and the experiment was found. In [72], a possibility of the prevention of the appearance of transient spots by means of a brief reduction in the arc current shortly after the initial current increase was demonstrated both numerically and experimentally and it was shown that blackening of burners of HID lamps that accompanies the appearance of transient cathode spots is due to evaporation of the cathode material rather than sputtering.

The appearance of cathode spots was studied in [57, 60, 71, 72] by means of a straightforward numerical simulation of the transition process, without attempting to theoretically describe an instability behind the breakdown of the non-stationary axially symmetric diffuse mode and the appearance of a 3D spot. In [77], the stability of a non-stationary current transfer is studied by means of applying a dc stability criterion to intermediate states of a non-stationary process neglecting their dependence on time, which is justified if the time of development of the instability is much smaller than a characteristic time of variation of the diffuse mode. Such an approach is frequently used for the investigation of the stability of non-stationary processes in systems of various types and contributes to a better understanding of the relevant physics.

A study of evolution of a spot on a hot refractory cathode of a high-pressure arc was reported in [98]. The modelling in this work, as well as in the other works cited in this section up to now, has been based on the non-stationary model of nonlinear surface heating. Such an approach can also be useful in the investigation of the glow-to-arc transition; a problem which has attracted considerable attention (e.g. [99–107]) but still requires further work.

The effect produced on the diffuse and spot modes and the limit of stability of the diffuse mode by variations of dimensions of the cathode is studied in [50, 76]. A reduction in cathode dimensions enhances heat conduction in the cathode body, which is a stabilizing effect. Therefore, a reduction in cathode dimensions is favourable for the diffuse mode. The same idea may be expressed in other words which are familiar to designers of arc devices: if a cathode is small, it is easier to heat it up to temperatures necessary for the diffuse operation. The sensitivity of the spot mode and the current range in which it exists with respect to variations of cathode dimensions is higher (much higher in certain cases) than that of the diffuse mode.

The effect produced on the spot mode by a rounding of the edge of the front surface of the cathode may be understood in the same terms: better cooling conditions on a rounded edge reduce the current range in which the spot mode exists, cf the CVC shown in figure 9 by the solid and dashed lines and the corresponding CVC shown in figure 14 of [76]. (Note that a similar effect is produced by an enhanced thermal conductivity of the cathode material [108].) A transient spot on a rounded cathode is formed considerably later, is considerably cooler and lives much less than the spot on a cathode with the flat front surface [72].

A protrusion on a thermionic arc cathode dramatically changes the pattern of steady-state modes of current transfer in certain cases, as shown by simulations [109]: it may happen that there is only one steady-state mode in a wide current range and this mode is spot-like. This finding may be related to the so-called super-spot mode observed in the experiment [110].
and to observations [111] indicating that a protrusion on the front surface facilitates stable operation of the cathode in a mode with a spot.

An upper limit of the cathode temperature and a lower limit of the temperature inside low-voltage well-developed spots can be estimated without actually solving the heat conduction equation inside the cathode [73].

A model of a current-collecting spot surrounded by a current-free region is useful in cases where the cathode is large enough and the spot can be treated as solitary. A self-consistent formulation of this model can be found in [112]; see also [113].

A simple approximate way to generalize a model of the near-cathode plasma layer in order that it is applicable to a plasma composed of neutral particles of several species, positive ions of several species, and electrons is given in [51].

The presence of metal halides affects current transfer to cathodes of ceramic metal halide lamps, including mercury-free lamps, in two ways: through a variation of properties of the near-cathode plasma layer due to the presence of metal halides in the gas phase and through a variation of the work function of the cathode surface due to formation on the surface of an alkali metal monolayer, the so-called gas-phase emitter effect. Both effects were studied theoretically in [114]. It was found that the presence of metal halides in the gas phase results in a small decrease in the cathode surface temperature and the near-cathode voltage drop and in an expansion of the range of stability of the diffuse mode. The gas-phase emitter effect was studied in [114] for Na–Hg and Cs–Hg plasmas. The work function of the surface of a tungsten cathode covered by a monolayer of Na or Cs was calculated as a function of the surface temperature with the use of semi-empirical formulae given in [115] and [116], respectively. The formation of the Na monolayer affects the diffuse mode of current transfer in the same direction as the presence of metal atoms in the gas phase does: the cathode surface temperature and the near-cathode voltage drop decrease weakly while the range of stability of the diffuse mode expands. In general, formation of the Na monolayer affects the plasma–cathode interaction only moderately. In contrast, formation of the Cs monolayer changes the plasma–cathode interaction dramatically: the temperature of the cathode surface decreases very strongly; the CVC of the diffuse mode becomes \(N \sim S\)-shaped.

As far as Na is concerned, theoretical results [114] conform to the experimental observation [58] indicating that Na does not affect the electrode behaviour compared with that in Hg-only lamps. Another conclusion of [58] was that Dy produces a significant gas-phase emitter effect: the average cathode tip temperature is drastically reduced and the cathode arc attachment changes from diffuse to constricted. This conclusion has been confirmed by subsequent experiments [117, 118]; in particular, a correlation between the density of Dy in front of the electrode and the emitter effect was observed [117]. Although the gas-phase emitter effect produced by Dy was not studied in [114], it is interesting to note that the drastic reduction in the average cathode tip temperature observed for Dy is similar to the one predicted in [114] for Cs.

In [119], equations describing the near-cathode plasma were solved jointly with an equation of electron emission from the cathode surface. Two solutions have been found for each surface temperature and near-cathode voltage drop, one of them with low values of the current density and the electric field at the cathode surface and another with high values. The cathode operates via Schottky-amplified thermionic emission in the framework of the first solution and via thermo-field or field emission in the framework of the second solution.

The explanation of the multiplicity of modes of current transfer to thermionic arc cathodes in the framework of the model of non-linear surface heating is based on the non-uniqueness of multidimensional thermal balance of a finite body heated by an external energy flux which depends in a nonlinear way on the local surface temperature: the equation of heat conduction in the cathode body may have multiple solutions for the same arc current even if the function \(q(T_w, U)\) is single-valued. However, the results [119] indicate that plasma states in front of a cathode with a given temperature of the surface and a given near-cathode voltage drop may be non-unique, and it is irrelevant in this context whether these non-unique states are, or are not, governed by essentially different physical mechanisms. It was mentioned in section 3.2.1 that transcendental equations describing the near-cathode layer in the framework of the model [23, 37, 50] admit multiple roots in some cases; possibly another indication in the same direction. It follows that the function \(q(T_w, U)\) may be multi-valued, and this may be an additional reason for non-uniqueness of temperature distributions in the cathode: one may think of solutions with different branches of the function \(q(T_w, U)\). This question remains to be clarified.

The above-discussed diffuse and spot modes are not the only modes of current transfer to thermionic cathodes of high-pressure arc discharges observed in the experiment. For example, the authors [120] observed, dependent on pressure, arc current and cathode material, four to five modes, which were not equally reproducible. The identification of different modes is always somewhat difficult, and the authors [121] rightly pointed out that different working groups tend to create their own names for effects they observe in their discharges. One should keep in mind therefore that the term ‘spot mode’ as used in this work, as well as in most recent papers on near-cathode phenomena in low-current high-pressure arc discharges, has quite a definite meaning; for example, it does not apply to modes with concentrations of current caused by non-uniformities on the cathode surface (such as areas with a reduced work function) or geometrical factors (such as conical geometry of the cathode).

An example of another (neither diffuse nor spot) mode of current transfer to cathodes of high-pressure arc discharges is represented by the above-mentioned super-spot mode [110], which occurs on a tiny tip on the cathode surface. Another example is represented by the so-called blue-core mode which occurs on cathodes of low-current free-burning arcs [91, 120, 121]. As compared with the diffuse mode, the cathode in the blue-core mode is very much colder, while the emission of the plasma above the cathode is much stronger [120]. The arc voltage in the blue-core mode is slightly higher than that in the diffuse mode [121], in contrast to the arc voltage in the spot mode being slightly lower than that in the diffuse.
mode. A transition from the blue-core mode to the diffuse mode occurs when the arc current decreases [120], in contrast to how the spot–diffuse transition occurs. The blue-core mode is observed mostly on cathodes made of thoriated tungsten; with pure tungsten cathodes it is more difficult to obtain [120]. The authors [121] believe that a convective motion of the gas is necessary for the formation of the blue-core mode (a gas was fed into the arc chamber in the experiments [91, 120, 121] apparently through a pipe in the cathode block).

Presumably, some of these other modes can be described by means of the same model of nonlinear surface heating that has been successfully applied to the diffuse and spot modes. For example, this is presumably the case of the super-spot mode: it has already been mentioned above that this mode may be related to the change in the pattern of steady-state modes of current transfer caused by a protrusion, detected by simulations [109] performed in the framework of the model of nonlinear surface heating. In connection with the blue-core mode, one could think of the above-mentioned solutions with different branches of the function \( q(T_w, U) \): a branch with low values of \( T_w \) would describe the diffuse and spot modes and a branch with high \( T_w \) would describe the blue-core mode. Data in the papers [91, 120, 121] are insufficient to judge whether the convective effects substantially change the steady-state current transfer or merely alter the stability of certain modes without strongly affecting the modes itself. If the latter is the case, it is sufficient to introduce an account of different branches of the function \( q(T_w, U) \) (as well as of the function \( j(T_w, U) \)) into the model of nonlinear surface heating. If the former is the case, the model of nonlinear surface heating for the cathode and the near-cathode layer will have to be supplemented with a model of the bulk plasma taking into account convective effects.

Steenbeck’s principle of minimum power, or voltage, for discharges with fixed current, which was proposed more than seven decades ago for a better illustration of observations (see, e.g., the textbook [122, p 184]), is still in use in physics of arc discharges, e.g. [52, 62, 75, 104, 123–128], including in works on plasma–cathode interaction in high-pressure arc discharges [52, 62, 75, 123]. Most of the authors, starting apparently from Peters [129], seem to believe that Steenbeck’s principle is a corollary of laws of thermodynamics. However, this is not the case: Steenbeck’s principle bears no relation to fundamental physical laws, being in this respect fundamentally different from well-known variational principles such as the principle of least action in mechanics or the Fermat principle in geometrical optics. Equally, Steenbeck’s minimum principle bears no relation to the stability theory, which has been demonstrated in section 3.4.2. The present state of the theory allows one to describe plasma–cathode interaction in high-pressure arc discharges in a self-consistent way on the basis of first principles, thus leaving no place for Steenbeck’s principle.

Concluding this section, it is worth saying a few words on the application of the model of nonlinear surface heating to plasma–cathode interaction in low-pressure and vacuum arc discharges. In [27], the model of nonlinear surface heating in the quasi-1D approximation was successfully used for a description of the diffuse and spot modes on cathodes of a low-pressure dc discharge in argon under conditions of fluorescent lamps. Note that such discharges are frequently termed thermionic arcs. This is misleading as far as the discharge column is concerned, since the neutral gas remains cold there; however, this term seems to be justified as far as the cathodic part of the discharge is concerned, since the physics of plasma–cathode interaction and its mathematical description are quite similar to those in the case of high-pressure arcs. As far as plasma–cathode interaction in vacuum arcs is concerned, the best developed at present are the theoretical models based on considering this interaction as a collective phenomenon: see, e.g., reviews [5, 130].

In other models, this interaction is treated as a sequence of individual events, termed microexplosions or microspots or ‘ectons’ [5, 131]. These models are not fundamentally different from the theoretical models of plasma–cathode interaction in high-pressure arcs. Therefore, the model of nonlinear surface heating may be applied in order to also describe plasma–cathode interaction in vacuum arc discharges, its applicability being limited only by the assumption of a collective phenomenon; see in this connection papers [5, 132–136].

3.5. On-line tool for simulation of axially symmetric current transfer to rod cathodes

The 2D simulation technique of plasma–cathode interaction has reached a point where it can be automated. A free on-line tool for simulation of axially symmetric steady-state current transfer to rod cathodes is available on the internet [79]. The database of plasma-producing gases includes, but is not limited to, He, Ne, Na, Ar, Kr, Xe, Cs, Hg, air, mixtures Na–Hg and Cs–Hg, plasmas of mercury or xenon with addition of metal halides. The database of cathode materials includes W, Mo, Fe, Nb, Zr.

At present, the tool simulates the diffuse mode. Future versions will calculate bifurcation points positioned on the diffuse mode, including the first bifurcation point that represents the limit of stability of the diffuse mode, and simulate the first axially symmetric spot mode.

It should be stressed that the tool is destined not only for researchers but for engineers as well, so there is no need to study theoretical papers in order to be able to use the tool.

4. What do we know about anodes of high-pressure arcs?

Over the years, near-anode phenomena in high-pressure arc discharges have been assumed to be less critical for performance of arc devices than near-cathode phenomena and have therefore attracted considerably less attention. A comprehensive and relatively recent review on the subject can be found in [4]; another review is currently under preparation [137]. Therefore, we mention here only a few works, mostly published during the last decade.

The work [138] reported an experimental investigation of multiple anode constriction in free-burning arcs with water-cooled electrodes in atmospheric-pressure argon and xenon in the intermediate current range. The multiple constriction
was shown to be a result of an instability in the near-anode region rather than in the bulk plasma. Further experimental and theoretical work on multiple anode constriction was reported in [139, 140].

The 1D modelling of the near-anode layers in high-pressure arc discharges was reported in [141–143]. The models [141, 142] comprised an (ambipolar) diffusion equation and separate energy equations for the electrons and heavy particles, i.e. did not rely on assumptions of ionization and thermal equilibrium. On the other hand, it was assumed that quasi-neutrality prevails throughout the whole plasma, i.e. the near-anode space-charge sheath was discarded. Calculation results were given for an atmospheric-pressure argon arc at the current density of \(0.9 \times 10^7\) A m\(^{-2}\). In [142], a careful analysis of the physics of the near-anode layer was performed by means of a comparison of different terms of the governing equations. An account of the near-anode sheath has been introduced in [143]. Besides, in [143] an attempt was made to get an idea of the effect of radial diffusion of the charged particles, radial transport of the electron and heavy-particle energies and the radial electric field. These effects have been estimated with the use of a prescribed variation of the radius of the arc channel; additionally, the approximation of quasi-neutrality has been employed in order to estimate the radial electric field.

In [67], the effect of the gas flow in high-intensity arcs on the negative anode fall and the heat flux to the anode was studied by means of a 2D numerical modelling which did not rely on assumptions of ionization and thermal equilibrium but still employed the quasi-neutrality assumption. Calculation results were given for the atmospheric-pressure argon plasma at the current of 200 A. It was found that the anode boundary layer strongly depended on the mass-flow rate if the arc attachment was diffuse. In contrast, the anode boundary layer appeared not to depend on the mass-flow rate if the attachment was constricted.

In [17], a theoretical picture of a near-anode layer in high-current arcs in atomic plasmas established in [22] was modified in order to be applicable to a near-anode layer in atmospheric-pressure nitrogen.

A detailed experimental investigation of anode phenomena in a model lamp filled with inert gases (Ar, Xe, Kr) was reported in [144]. The arc current varied in the range from 0.5 to 10 A and the plasma pressure from 1 to 10 bar. The arc attachment occurred in the constricted mode, which is characterized by a plasma flow from the anode to the bulk plasma. A negative voltage drop is present in front of the anode, similarly to what happens in high-current arcs. It was found that the power input into the anode varies approximately linearly with the arc current and the proportionality constant, the so-called anode heating voltage, is around 7–8 V. Note that the measured power input into the anode is in agreement with the modelling [143]. On the other hand, other authors speak of somewhat lower values of the anode heating voltage, see, e.g., [60]. The anode heating voltage is mainly governed by the work function of the anode material and seems to be unrelated to the electrically measured anode fall and scarcely dependent on the anode dimensions.

The measurements [144] have shown that the arc attachments on cold and hot anodes are quite different and may represent different modes. The constriction is more pronounced on cold anodes than on hot anodes. Maxima of \(T_e\) and \(n_e\) in front of the anode surface have been observed. With increasing anode temperature distances separating the \(T_e\) and \(n_e\) maxima from the anode surface increase and \(T_e\) in front of the anode is reduced.

An interpretation of the measurements suggested in [144] fits the physical picture discussed in section 2: the sheath in front of the anode delivers electrical power to the circuit; hence, there is a power supply into the sheath from the adjacent plasma. In the case of a constricted mode the power requirement of the sheath is covered by an enhanced electrical power input into the constriction zone. In the diffuse mode of anodic arc attachment, often realized at high currents, the power requirement is covered by a plasma flow from the bulk plasma to the anode.

Xe lamps operated at very high pressures develop under certain conditions voltage oscillations accompanied by electromagnetic interference (EMI) emissions, e.g. [145]. This effect has been pinned down to phenomena in near-anode layers [146, 147].

Let us now briefly discuss main factors hindering or facilitating investigation of anode physics. In accord with what has been said in section 2, the energy flux to the anode surface is formed in the layers of thermal perturbation and thermal non-equilibrium, i.e. at much larger distances from the electrode surface than distances at which the energy flux to the cathode is formed. Furthermore, it was shown in section 3.1 that the voltage drop \(U\) in the near-cathode layer considerably exceeds \(U_{\text{con}}\) the combined voltage drop in the constriction zone and the region of constricted plasma, meaning that the resistance of the near-cathode layer considerably exceeds the combined resistance of the region of constricted plasma and the constriction zone. Therefore, the distribution of current over the cathode surface is governed by the near-cathode layer rather than the constriction zone and/or the constricted plasma. The voltage drop in the near-anode layer is considerably smaller than that in the near-cathode layer and may even be negative, while \(U_{\text{con}}\) remains more or less the same; thus \(U \lesssim U_{\text{con}}\) in the case of the anode. It follows that the resistance of the near-anode layer is of the same order of magnitude or smaller than the combined resistance of the constriction zone and the region of constricted plasma; hence, a distribution of current over the anode surface is partially or wholly governed by the constriction zone and/or the region of constricted plasma.

One can conclude that a decoupling of the near-electrode layer from the bulk plasma, while being a reasonable approximation in the case of the cathode, would be unjustified in the case of the anode. It follows that simulation of the arc plasma–anode interaction requires modelling not only of the near-anode non-equilibrium layer but also of the adjacent bulk plasma. This, of course, complicates the modelling.

As shown in section 2.1, \(L_p\), the scale of thickness of the layer of thermal perturbation in mercury plasma at \(p = 30\) bar and \(j = 10^7\) A m\(^{-2}\), being about 0.13 mm, is substantially smaller than the typical radius \(R_{\text{att}}\) of the arc attachment, so the energy balance of the constricted plasma in mercury-based HID lamps is to a first approximation dominated by
radiation. An estimate for the Ar plasma at $p = 1–10$ bar and the same current density gives $L_{\text{pp}} = 1–0.9$ mm, which is comparable to or higher than $R_{\text{at}}$. Hence, the energy balance of the constricted plasma in model lamps is not dominated by radiation. Furthermore, there is no sense in talking about the layer of thermal perturbation under conditions of model lamps. Thus, the physics of current transfer in model lamps filled with inert gases and in mercury-based HID lamps is in a general case different, and this definitely hinders experimental investigation of anode phenomena, more bad news.

On the other hand, near-anode non-equilibrium layers are described by the same equations that describe near-cathode layers and may be simulated by means of the same code, cf. figures 2 and 3. Methods of simulation of the bulk of arc discharges are well developed. Hence, there should be, at least in principle, no major problems in the modelling of the arc plasma–anode interaction.

5. Modelling electrodes of ac arcs and the whole system arc–electrodes

Modelling of electrodes of ac HID lamps was reported in [58, 83, 148]. Different variants of the model of nonlinear surface heating have been used for a description of the cathodic phase. The anodic phase was described in a rather crude way; for example, in the modelling [58] a fixed value taken from the experiment has been assumed for the anode heating voltage and the current density was assumed to be uniformly distributed over a prescribed part of the anode surface.

Cathodic arc attachment after a steady-state anodic phase was numerically investigated in [60]. A fixed value taken from the experiment has been assumed for the anode heating voltage and the anode temperature distribution was replaced by one that corresponds to the cathode regime with an equal integral heating power.

The model of nonlinear surface heating has also been used in simulations of the whole system arc–electrodes. In [149], the near-cathode space-charge sheath was assumed to be collision-free and the plasma–cathode interaction was described by the model [37]. The arc column was described by a two-temperature model, i.e. the assumption of thermal equilibrium was not used. A similar approach was adopted in [68], with the difference that the modelling was extended to the anode. However, the near-anode region was not considered in [68] and its treatment was replaced by an assumption. It was found that the electric power deposited into the near-cathode layer is transported not only to the cathode but also to the arc column; see the discussion in section 3.2.3.

In [16], the model of a collision-dominated near-cathode space-charge sheath [38] was integrated into a numerical model of an HID lamp as the whole. Another model of a collision-dominated sheath was implemented in a numerical model of an HID lamp in [150]. The arc column was described by an LTE model in both works.

6. Conclusions

The plasma–cathode interaction in low-current high-pressure arc discharges is governed by a thin near-cathode plasma layer. A self-consistent and universally accepted theory and simulation methods of the plasma–cathode interaction have been developed and validated experimentally, although further work is still needed on the spot mode.

A 2D simulation technique of plasma–cathode interaction has reached a point where it can be automated. A free on-line engineering tool for simulation of axially symmetric steady-state current transfer to rod cathodes is available at [79].

There is a possibility of simulation of complex 3D geometries, including cathodes of complex shapes consisting of several pieces made of different materials, of non-stationary processes and of plasmas of complex chemical composition (e.g. ceramic metal halide lamps, including mercury-free lamps).

The plasma–anode interaction in high-pressure arc discharges is governed not only by a thin near-anode layer but also by the adjacent part of the bulk plasma. In other words, the decoupling of the near-electrode layer from the bulk plasma, while being a reasonable approximation in the case of cathode, would be unjustified in the case of anode. The physics of current transfer in model lamps filled with inert gases and in mercury-based HID lamps is in a general case different, and this hinders investigation of anode phenomena. Due to advances achieved in numerical modelling of near-cathode non-equilibrium layers and the bulk of arc discharges, a numerical simulation of near-anode phenomena is, at least in principle, straightforward.

Simulation of the plasma–electrode interaction in ac arcs and of the whole system arc–electrodes with the use of modern methods of modelling of plasma–cathode interaction has already started. Further progress in this field will follow after understanding of the plasma–anode interaction improves.

Acknowledgments

The work on this paper was supported by projects POCI/FIS/60526/2004 and PPCDT/FIS/60526/2004 Modes of current transfer to cathodes of high-pressure arc discharges and their stability of FCT, POCI 2010 and FEDER and by the project Centro de Ciências Matemáticas of FCT, POCTI-219 and FEDER. The author is grateful to Professor J Mentel and Dr G M J F Luijks for critically reading the manuscript and many useful comments.

References


[61] Fischer E 1987 Modelling of low-power high-pressure discharge lamps Philips J. Res. 42 58–85


[79] http://www.arc_cathode.uma.pt


[94] Mitrofanov N K and Shkol’nik S M 2007 Two forms of attachment of an atmospheric-pressure direct-current arc in argon to a thermionic cathode Tech. Phys. 52 711–20
[122] Raizer Y P 1991 Gas Discharge Physics (Berlin: Springer)
[123] Zhou X and Heberlein J 1994 Analysis of the arc-cathode structures and determination of radius of arc spots on cathodes Phys. Scr. 58 383–6


[131] Vasenin Y 1999 Parameters of plasma layer within the vacuum arc cathode spot IEEE Trans. Plasma Sci. 27 858–3


