Critical potentials of mercury with a Franck–Hertz tube

Peter Nicoletopoulos

Faculté des Sciences, Université Libre de Bruxelles, Brussels, Belgium

E-mail: pnicolet@skynet.be and pnicolet@ulb.ac.be

Received 17 April 2002, in final form 16 July 2002
Published 6 September 2002
Online at stacks.iop.org/EJP/23/533

Abstract

A classic experiment, where a four-electrode Franck–Hertz tube is used to show various energy states (critical potentials) to which a neutral mercury atom can be raised by electron impact, is upgraded and elucidated thanks to a novel arrangement of applied inter-electrode potentials. By that means, a suitably located plasma region, the working field-free scattering chamber, can be created dynamically and maintained thereafter under appropriate conditions of cathode emission current and mercury vapour pressure. Concurrently, a more efficient mechanism for displaying excitation peaks becomes operant in the adjoining electron collection cell, by means of the window for slow-electron transmission that is available in mercury below the resonance peak in the cross section at about 0.4 eV. The resultant enhanced resolution in current–voltage curves permits the identification and measurement within 0.05 eV of the leading features in the known electron impact excitation spectrum of mercury.

1. Introduction

This paper is about an experiment that can lift itself by its own bootstraps: the attempted setup, inefficient at the outset, is spontaneously redesigned during operation by a conspiracy of dynamical phenomena whose actuator is the very element that is being subjected to investigation.

Suppose that you want to demonstrate—indeed, measure—atomic critical potentials. Take a vacuum diode. Interpose two grids distant from each other near the electrodes, say grid 1 close to the cathode and grid 2 close to the anode.

(a) Short the grids so that their interspace will be equipotential. Insert a pure gas of atoms, say mercury.

(b) Adjust the pressure so that the electronic excitation free path is longer than the gap between the cathode and grid 1.

Thus, when a voltage is applied across this gap, preaccelerated electrons traversing the vapour between the grids will conserve their energy unless they suffer inelastic collisions.

1 Address for correspondence: Rue Joseph Cuylits 16, 1180 Brussels, Belgium.
(c) Provide a small retarding potential between grid 2 and the anode in order to prevent slow electrons from reaching the anode.

Now vary the voltage of grid 1 and simultaneously monitor the electron current between grid 2 and the anode. You expect to see downturns in this current whenever the accelerating voltage passes through a critical potential of the gas.

At first sight this looks simple enough for an undergraduate laboratory. The question is, how well does it work? Usually not so well, as attested by the literature: the data commonly obtained are so elusive that textbooks of laboratory practice proposing this method have been loath to show sample results [1, 2].

In 1979 I introduced this experiment in our undergraduate laboratory because it is an immediately available extension of the standard Franck–Hertz experiment, thanks to a Franck–Hertz tube of the type described in [2] which incorporates grid 1 as a luxury (see section 2). It took years to grope a way toward a reasonably consistent interpretation of the electron excitation spectrum of mercury obtained in that way, and some time more to find a theoretical basis for the gaseous discharge phenomena that are involved. The main outcome of this search is presented here. The technical elaboration of the transport-theoretical and plasma-physical aspects will be presented elsewhere [3, 4].

Although this experiment dates from the days of the early quantum theory, only scarce data exist in the literature with which to compare the results on mercury. Three papers contain current–voltage curves in the range between zero and about 25 V. The authors were:

(1) Mohler et al [5] in 1920 at NBS,
(2) Einsporn [6], a student of James Franck, in 1921 in Berlin, and

These plots were explained by matching the peaks with suitable combinations of two or three of the lowest critical potentials of mercury, multiple inelastic scattering being favoured at the pressure employed. Einsporn interpreted his or her plot in terms of the $3P_1$ and $1P_1$ states. Mohler et al, who presented curves at more than one value of pressure, concluded likewise but noted that one or two additional critical potentials might be involved. Liu, whose features are sharper and more numerous, was forced to invoke a third critical potential, the $3P_2$ state, and yet four features remained outside this scheme.

Curves resembling one or another of the ones published are readily obtained with a Franck–Hertz tube of the type described in [2]. At times, however, additional peaks show up among the ones usually present and continually frustrate all attempts to interpret the spectrum. The problem is that the pressure required for producing strong peaks and valleys in the anode current is so large that positive ions are produced very early by repeated electron impact on atoms excited to the metastable $3P_0$ and $3P_2$ states. Thus, the electrostatic potential distribution between the grids evoked by condition (a) is strongly modified by the dynamics of the interaction of positive and negative space charge. In addition to that, the effectiveness of condition (c) becomes questionable because the anode current is bipolar. It was realized that a maximum of structure is obtainable if condition (c) is abandoned or reversed and a sufficiently large decelerating voltage is applied across the grids. This structure is identifiable as the superposition of two spectra (see section 3). The inference is that the space potential between the grids has been distorted into a steplike configuration of two field-free plasma regions at different potentials joined by a narrow space charge layer. This phenomenon is well known in plasma physics.

Varying the intergrid potential can move the narrow zone dividing the two drift spaces inwards or outwards. In this way one can eliminate the intermixing of spectra. Particularly good results are obtained under conditions that privilege the display of the spectrum excited at the lower potential, namely in the field-free plasma adjacent to the outer grid.

The reason is that we are then enabled to shift from the intended beam-depletion experiment to its complementary version in which critical potentials are displayed as peaks in the diffusion current of slow electrons crossing the ‘detection chamber’ between grid 2 and the anode.
Critical potentials of mercury with a Franck–Hertz tube

Table 1. Observed critical potentials and their interpretation.

<table>
<thead>
<tr>
<th>No</th>
<th>Measured value (eV)</th>
<th>Attribution</th>
<th>Accepted value (eV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.4</td>
<td>(5d^{10}6s^26p)^3P state of Hg^-</td>
<td>0.42°</td>
</tr>
<tr>
<td>2</td>
<td>4.9</td>
<td>(5d^{10}6s6p)^3P_1</td>
<td>4.89</td>
</tr>
<tr>
<td>3</td>
<td>5.5</td>
<td>(5d^{10}6s6p)^3P_2</td>
<td>5.47</td>
</tr>
<tr>
<td>4</td>
<td>6.7</td>
<td>(5d^{10}6s6p)^3P_1</td>
<td>6.70</td>
</tr>
<tr>
<td>5</td>
<td>7.8</td>
<td>Unresolved cluster of (5d^{10}6s7s)^3S_1</td>
<td>7.73</td>
</tr>
<tr>
<td>6</td>
<td>8.8</td>
<td>(5d^{6}6s^26p')^3D_1</td>
<td>8.80</td>
</tr>
<tr>
<td>7</td>
<td>9.8</td>
<td>Twofold 6^3P_1</td>
<td>9.78</td>
</tr>
<tr>
<td></td>
<td></td>
<td>or (5d^{6}6s^26p')^1P</td>
<td>9.77</td>
</tr>
<tr>
<td>8</td>
<td>10.4</td>
<td>6^3P_1 + 6^1P_2</td>
<td>10.36</td>
</tr>
<tr>
<td></td>
<td></td>
<td>or (5d^{6}6s^26p')^3(3/2 1/2)</td>
<td>10.35</td>
</tr>
<tr>
<td>9</td>
<td>11.0</td>
<td>Twofold 6^3P_2</td>
<td>10.94</td>
</tr>
<tr>
<td></td>
<td></td>
<td>or (5d^{6}6s^26p')^3P</td>
<td>11.00</td>
</tr>
</tbody>
</table>

° Peak in the elastic cross section.

Such electrons, of the very kind that we had initially planned to hold back, are subject to preferential collection near their thresholds of production due to an effect intrinsic to elastic e^-–Hg scattering: a low-energy shape resonance (see sections 4 and 5.1).

The features displayed by this contrivance are stronger and much sharper than those obtained ordinarily and can be measured within 0.05 eV. Thus it becomes possible to interpret the structure in terms of the known electron impact excitation spectrum of mercury (see table 1).

The coarser curves of [5–7] are spectra obtained under conditions that privilege the display of the spectrum excited in the high-potential plasma. The consistency of these data with the series in table 1 is established in sections 5.2 and 6. A detailed consideration is necessary because a further difficulty caused by space charge must be faced.

The phenomena underlying the changes of the space potential are quite complicated and their details are outside the scope of this paper. The main points are outlined in section 7.

The dependence of optimal operation on the pressure of mercury vapour and on the emission current density and velocity distribution of the electron beam is given in section 8.

2. Apparatus

The tube used in this experiment is a commercially available Franck–Hertz tube (Leybold–Heraeus 55580) an earlier version of which (55080) was described in [2]. Although a triode is sufficient for the standard Franck–Hertz experiment, this tube contains an extra grid g_1 close to the oxide-coated cathode K. This grid, designed to act merely as a ‘current control’ electrode, is promoted here to accelerating grid.

All electrodes are cylindrical except for g_1 which is a flat, 0.2 cm wide helix evenly coiled around two upright supports siding the 0.05 cm diameter cathode. A sort of shield, close to and shorted with g_1, obstructs the vertical ends of this assembly. The open central part, 0.7 cm tall, contains about 40 turns of thin (0.005 cm?) wire.

The outer grid g_2 is a cylindrical net made of 0.02 cm wire (0.1 cm square mesh); it is 3 cm tall. The height of the anode is 2 cm. The distances d_1, d_2, d_3, between K and g_1, g_1 and g_2, g_2 and A (the anode), respectively, are: d_1 = 0.02, d_2 = 0.8, d_3 = 0.2 cm.

The experimental arrangement is shown in figure 1. The tube lies vertically inside a small homemade electric oven, essentially a 50 Hz ac coil. The temperature was measured with an ordinary mercury thermometer inserted between the coil’s support and the tube through the
Figure 1. Schematic diagram of tube and circuit.

lid. Temperature readings $T$ ($^\circ$C) vary somewhat along the axis of the oven. The readings of $T$ given in the text were taken with the tip of the thermometer positioned about midway between the drop of mercury at the bottom of the tube and the electrodes.

M1 and M2 are ordinary vacuum tube meters. The role of M1 is essentially to provide a standard, called $I_c$, of the emission capability of the cathode. The quantity $I_c$ is defined arbitrarily as the reading of M1 (operated in current mode) at $T =$ room temperature, $V_a = 8.6$ V, $V_f = 3.5$ V, where $V_a$ is the voltage between the cathode and $g_1$, and $V_f$ is the effective value of the ac filament voltage. As will be discussed in section 8, the value of $I_c$ depends on the state of activation of the cathode, which is an important factor.

The output of M2 is fed to the Y terminals of an X–Y recorder whose X-axis is driven by the accelerating voltage $V_a$. The scale of X was set at 1 V cm$^{-1}$. $V_a$ is varied very slowly by hand via a 5 kΩ helipot. Smoothness of recording was enhanced by adding the condenser $C$ (640 µF). The curves shown in this paper are representative of hundreds of similar plots obtained in the period 1985–7.

Preparations must also be made for small constant voltages of either sign to be applicable between $g_1$ and $g_2$ and between $g_2$ and A. The algebraic values (in volts) of these potentials are called $\Delta g$ and $\Delta a$, respectively; positive values accelerate outward-bound electrons.

The magnitude of $\Delta g$, accurately controlled by a finely adjustable constant-voltage generator, is a crucial element in this experiment.

3. Untangling the energy scale: the main step

Figure 2 shows a plot of anode current versus $V_a$ obtained with $\Delta g = 0$ and $\Delta a = 0$. A discussion of how different portions of such curves are influenced by mercury vapour pressure, cathode state of activation and filament voltage is deferred until section 8; meanwhile, it will be implicit that these factors have been adjusted for optimal visibility of a maximum of structure throughout the energy range.

The excitation curve of figure 2 is considerably more complex than the original curves of Einsporn [6] and Mohler et al [5]. It differs, too, from the modern curve of Liu [7]. The central purpose of this section is to show that the increased quantity of structure obtained here is due to the fortuitous presence of a property intrinsic to well-seasoned samples of our type of tube.

Before we can measure critical potentials, we need a reference for calibrating the energy scale. To all appearances, the first large peak, feature $a$, should correspond to a poorly
resolved clump of the lowest multiplet of excited states of mercury, the $6^3\text{P}$ triplet around 5 V, which is the state involved in the standard Franck–Hertz experiment. Despite the differences between figure 2 and the excitation curves of previous authors, a similar peak is displayed in all occasions, and is commonly placed at 4.9 V to calibrate the energy scale.

On the other hand, it is clear that there is a point $V_a = V_0 \approx 2.5 \text{ V}$ on the voltage axis, below which there is no anode current. This must mean that the energy scale has been shifted to the right by a retarding (negative) contact potential difference (cpd) of 2.5 V between K and the grids, a fact that would be consistent with a low-work-function cathode surface.

If the abrupt rise of current is the zero of absolute energy, then feature 1 (also shown in amplified form) is compatible with the peak at 0.4 eV in the elastic cross section due to a shape resonance (see section 4). Feature a, however, which we would like to attribute to the $6^3\text{P}$ triplet, appears at about 3 V on that scale instead of 5 V. We are therefore confronted with a serious inconsistency.

A clue for the solution of this problem is obtainable from a more elaborate exploration of the origin of the retarding cpd: let $U_{g_1}$ and $U_{g_2}$ be the contact potentials between K and $g_1$ and between K and $g_2$. If $W_c$, $W_{g_1}$, $W_{g_2}$ are the work functions of these three electrodes, then $eU_{g_1} = W_c - W_{g_1}$ and $eU_{g_2} = W_c - W_{g_2}$. The true electrostatic voltages $T_a$ and $T_{g_2}$ felt by cathode electrons traversing $g_1$ and $g_2$, respectively, are $T_a = V_a + U_{g_1}$ and $T_{g_2} = V_a + U_{g_2}$. Therefore $T_{g_2} - T_a = U_{g_2} - U_{g_1} = (W_{g_1} - W_{g_2})/e$. This means that the electrostatic potential drop between shorted grids is zero only when $U_{g_2} = U_{g_1}$ or $W_{g_1} = W_{g_2}$.

Now $V_0 = V_a - T_a = V_a - T_a - \Delta U$, where $\Delta U = U_{g_2} - U_{g_1}$. But the location of peak a on the $V_a$ axis indicates that $T_a$ is close to $V_a$. And if the current to $g_1$ is plotted against the voltage $V_{g_1}$, its onset potential is found to be very close to $V_a = 0$, which means that the cpd between K and $g_1$ must be negligible so that indeed $T_a = V_a$. Thus $V_0 = -\Delta U = (W_{g_1} - W_{g_2})/e$. In other words the 2.5 V delay of the onset of anode current must be due to a negative contact potential between $g_1$ and $g_2$.
Figure 3. Mixed excitation curve with $\Delta g = -0.5$; feature b has come into view, f and g have been overtaken by 6 and 7.

If this is true, then upon addition of an external negative voltage $\Delta g$, $V_0$ should advance to $V_0 + |\Delta g|$. Curves obtained for $\Delta g = -0.5, -0.8, -1.6$, are displayed in figures 3–5. It is seen that in each case the onset of anode current has indeed been delayed by an additional $|\Delta g|$ volts. The threshold of $g_1$ current is not affected. This confirms that the cpd is effective between $g_1$ and $g_2$.

Unexpectedly, however, the changes of intergrid potential have also affected the overall aspect of the excitation curves, certain portions of which have been significantly modified.

The key to solving the puzzle of the energy scale is recognizing that the source of the alterations in the excitation curves is dual. Firstly, certain peaks (those labelled by numbers) are shifted to the right by $|\Delta g|$ volts concurrently with the shift of the threshold of anode current. Secondly, there is a class of peaks (labelled by low-case letters) whose location is independent of the value of $\Delta g$. Thus, as $\Delta g$ is made more negative, a standing peak may be overtaken and partly or totally masked by a prominent moving one. For instance, in figure 4, feature 4 has crept forward to almost overtake peak d. The unmasking of a previously invisible standing feature can also occur, as is the case for peak b in figures 3 and 4.

The standing structure (except of course for feature a) is overtaken entirely at $\Delta g \approx -1.2$.

For $\Delta g < -1.2$, namely after this mergence, the standing peaks do not re-emerge. The moving spectrum—now unblemished—improves toward its optimal form, attained at $\Delta g \approx -1.5$. Curves recorded in the range of $\Delta g$ between $-1.5$ and $-1.8$ offer the strongest display of structure. Beyond $\Delta g \approx -1.8$, the peaks become narrower and their amplitude decreases.

Because the location of moving peaks is governed by the potential $T_{g_2}$ of $g_2$ with respect to the cathode, it is natural to suspect that these features show excitations effected by electrons with energy equivalent to this potential, these electrons somehow reaching the vicinity of $g_2$ (decelerated by $\Delta V = \Delta g + \Delta U$) without suffering previous inelastic collisions. Impressive support for this conjecture is obtained if the pure moving spectrum of figure 5 is measured from $T_{g_2} = 0$, namely with respect to the onset of anode current. As discussed in section 4, the
Critical potentials of mercury with a Franck–Hertz tube

Figure 4. Mixed excitation curve with $\Delta g = -0.8$; feature 4 has almost overtaken peak $d$.

Figure 5. Excitation curve with $\Delta g = -1.6$, showing a pure ‘moving’ spectrum plus feature $a$. 
values obtained by this procedure correspond to all the prominent features commonly observed in the electron impact spectrum of mercury.

What about the origin of the standing peaks? By their very name they must be due to excitations occurring in the vicinity of the first grid, by electrons with energy equivalent to the true potential $T_a (\equiv V_a)$ of $g_1$ with respect to the cathode. We can now explain the inconsistency encountered with regard to the position of feature α: since this peak is stationary, it must be measured on the $g_1$ voltage scale and on this basis its location is not incompatible with its interpretation in terms of the $6^3P$ state.

Needless to say, the existence of two scales set off from each other, means that different peaks in the same curve can show the same energy level excited in different regions in the tube. For instance, features 2 and 3 correspond to the two highest members of the $6^3P$ triplet excited at $g_2$ potential. Clearly, much better resolution is attainable amongst the moving structure.

But the very existence of the moving structure is critically dependent on the presence of a retarding potential of sufficient magnitude between the two grids. It is easily verified that far less satisfactory results would have been achieved had we used a tube with no cpd between $g_1$ and $g_2$ ($\Delta U = 0$) in order to really accomplish the arrangement desired originally: indeed, the effect of compensating the negative value of $\Delta U$ by making $\Delta g$ increasingly positive is to blunt the exposed structure by favouring the standing peaks. Quite soon, at $\Delta g \cong +0.4$, namely for $\Delta V$ still in the retarding direction but inferior in magnitude to $|\Delta g + \Delta U| \cong 2.1$, all the moving peaks are washed out. Thereafter, the remaining structure becomes broader until ultimately a stable form is attained at $\Delta g \cong +1.5$. This type of curve, shown in figure 6, is quite similar to the plots obtained in [5, 6].

The fact that even for $\Delta g = 0$ a retarding value of $\Delta V$ is operant in the present tube is doubtless due to deposition of oxide material from the cathode coating onto the surface of the close-lying first grid. As a result, the cpd initially present between the cathode and a clean $g_1$ surface has been neutralized, but a large retarding potential has become effective between
Critical potentials of mercury with a Franck–Hertz tube

Figure 7. Intergrid space potential, established for $\Delta g < 0.4$ and a strong beam, showing two plasma regions and a space charge double layer (DL) with the high-potential plasma at (a) applied value of grid 1 voltage and (b) raised value operant beyond feature $A$. Observed values of $\delta$ range from 1.6 to 2.7.

Let us take stock: our initial intention of creating field-free drift space within the tube by shorting the grids was thwarted by the presence of a retarding potential between these electrodes. Far from being a handicap, this mishap is a bonus, and enhancing its effect upgrades the experiment. Thus, it is found that in the most effective arrangement for showing critical potentials of mercury with the present four-electrode tube, the second grid must be maintained negative with respect to the accelerating grid by 4–4.3 V, including the cpd. The structure in excitation curves recorded with retarding values of $\Delta g$ is interpretable in terms of inelastic events occurring in regions of the tube where the electron beam energy is equivalent to either the potential of $g_1$, or that of $g_2$, with respect to the cathode.

If it is correct, that interpretation implies that only an insignificant fraction of the inelastic events can take place at intermediate energies. Therefore, it would require the presence, in the interspace between grids, of a stepwise decreasing potential configuration consisting of two equipotential regions joined by a precipitous fall where the beam is abruptly decelerated from $g_1$ to $g_2$ potential, as shown in figure 7(a). The natural unit here is the electronic excitation free path $\lambda_e$. So, expressed in this scale, the widths $p_1$ and $p_2$ of the high- and low-potential regions and the width $s$ of the sheath joining them should be in accordance with the relations: $p_1 \gg 1$, $p_2 \gg 1$, $s \approx 1$.

Potential configurations of this type can indeed arise from the global interaction of positive and negative space charges. In fact, we have come across an instance of a nonlinear phenomenon which is of current interest in plasma physics: the formation of a narrow space charge sheath between two weakly ionized plasma regions at different potentials (see section 7).

The foregoing discussion indicates that for values of $\Delta g$ exceeding +0.4, the sheath advances outward through the intergrid volume so that $p_2 \approx 1$, since only the standing spectrum is observed. For $\Delta g < -1.2$ the converse phenomenon takes place.

A closer look at the structure of the stationary spectrum beyond feature $a$ (see section 6) will bring to light a further complication: the shift of scale $|\Delta V|$, namely the potential difference between the two field-free regions, jumps from $|\Delta g + \Delta U|$ to a higher value $|\Delta g + \Delta U| + \delta$ in
the course of operation. Thus, the final shape of the high-potential region is the one shown in figure 7(b). The formation of a hump in the space potential is another topic of current research.

4. Measuring critical potentials

A striking trait of figure 5 is that the pattern comprising peaks 2–9 is met repeatedly, with somewhat reduced amplitude but no appreciable change in shape. Two complete cycles (unprimed and primed numbers) and part of a third (doubleprimed numbers) are included in our energy range.

The energy interval between similar peaks of adjacent periods is 6.7 eV. This value agrees precisely with the energy of the $^1P_1$ singlet state of mercury. Primed and doubleprimed peaks are therefore interpreted as multiple step excitations involving the corresponding unprimed level plus one and two $^1P_1$ transitions, respectively. The exclusive involvement of the $^1P_1$ state in these periodic multiple-impact transitions is in accordance with the fact that the excitation cross section for this state does not drop significantly after reaching a peak but extends virtually undiminished well beyond its maximum, so that above 12 eV it exceeds the excitation functions of other prominent levels by several orders of magnitude [8, 9].

Peak 2 is located nearly 5 V beyond the onset of anode current. Therefore, the eight peaks within the first period should reflect the structure of the $e^−$–Hg excitation spectrum in a range of 6.7 eV past the first excitation threshold. Paired excitations involving the lowest lying levels (the 6$^3P$ triplet) are energetically possible only beyond feature 6 so that we must be observing at least five single-impact critical transitions.

Feature 1 is special and will be discussed individually in the later part of this section. The energy scale whose zero is the steep onset of anode current is calibrated by placing peak 4′, the second occurrence of the strongest feature, at 13.4 V. This procedure was found to give better overall agreement than placing peak 4 at 6.7 V, because peak 4′ always appears at the same distance from the dip at feature 1, whereas peak 4, and the two preceding it, are sometimes shifted (by as much as 0.2 V) farther from feature 1 and closer to their primed counterparts. This effect is more noticeable for less negative values of $\Delta g$. Apparently a slight scale distortion can exist during an interval of about 2 V following the creation of the low-potential drift space. Beyond feature 4, complete linearity is established since the energies obtained for peaks 5–9, 2′ and 3′ by measuring from peak 4′, are precisely 6.7 eV below those of their primed counterparts 5′–9′, 2′′ and 3′′.

The values found for the nine basic features of figure 5 are given in table 1. Remarkably, this is an exhaustive list of the electron impact spectrum of mercury obtainable with a resolution of 0.2 eV or so in the range $0 < E \leq 12$ eV (see the spectra in [10, 11]).

Peaks 2–5 correspond respectively to the 6$^3P_1$, 6$^1P_2$, 6$^1P_1$, and to a cluster of the 7$^3S_1$ and 7$^1S_0$ states of the 5d$^{10}6s$ configuration of mercury converging to the well known $^3S_{1/2}$ ionization potential at 10.4 V. The lowest member of the 6$^3P$ triplet, 6$^3P_0$ at 4.7 eV, is not observed.

Prominent structures commonly displayed in $e^−$–Hg excitation spectra above 8 eV are attributed nowadays to states of the 5d$^96s^2$ configuration converging to the $^2D_{5/2}$ and $^2D_{3/2}$ ionization potentials at 14.8 and 16.7 V respectively. Therefore, feature 6 at 8.8 V is assigned to the 6p$^3D_{3/2}$ level in that configuration, a metastable state as discussed in [12].

Peaks 7–9 correspond to the three pairwise combinations of features 2 and 3 and would surely be identified with those were it not that single-impact events at the same energies are seen in $e^−$–Hg spectra obtained by one method or another.

Single-impact candidates for features 7 and 9 are the 6p$^3P_1$ and 6p$^3P_3$ states at 9.77 and 11.00 eV, both of which appear strongly in the energy loss spectra of [9]. As to feature 8, a case has been made of late [13] for a new (metastable) state close to 10.5 eV which is claimed to be related to a pronounced resonance structure observed [13, 14] at about 10.35 eV. This neutral state is best described in $j−j$ coupling notation (see table 1).
Next we turn to feature 1 which is of a different nature. As mentioned earlier, it is natural to attribute this structure to the 'shape' resonance identified in [15] and confirmed elsewhere [16, 17]. Remarkably, this feature had already been discovered 50 years earlier with mercury and cadmium in precisely our type of apparatus (with \( \Delta a > 0 \)) and correctly recognized as being due simply to a variation with energy of the elastic mean free path [18].

A shape resonance of this type is an unstable compound state (negative ion) formed when an impacting electron is temporarily trapped in a neutral atom by a combination of polarization and centrifugal barrier potentials [19]. This effect produces an anomaly—in this case a pronounced asymmetric peak of considerable width—in the total elastic scattering cross section \( \sigma_T(E) \), near the energy of the compound state. Comprehensive reviews of experimental and theoretical studies of negative ion resonances are provided in [20, 21].

How is the shape resonance observed here? Suppose that electrons of energy \( E \) are injected into a field-free region of gas of width \( x \). It is known [22, 23] that at moderate pressure and in the absence of inelastic events, the current traversing this region is proportional to the total elastic free path \( \lambda_T(E) \) and thus inversely proportional to \( \sigma_T(E) \). Therefore, in a plot of current versus \( E \), the resonance will appear as a dip and the peak in the cross section should be near the minimum.

With our calibration of the energy scale, the dip at feature 1 is found to lie near 0.4 eV, in agreement with the total cross section results of [17]. Thus, arbitrary as it may have seemed, the procedure of identifying peaks at features 2–9 with thresholds of excited states is consistent with the determination of the energy of feature 1 at the dip. The absence of feature 1 in the plots of [5–7] is understandable since a retarding potential (\( \Delta a < 0 \)) was applied.

5. Two complementary detection mechanisms

5.1. The low-potential spectrum

Obviously, the mechanism responsible for observing feature 1 should also operate at thresholds of excited states whereupon electrons of very low energy, stemming from inelastic collisions in the low-potential drift space, diffuse through the vapour in the adjacent detection chamber: a sharp peak should appear near each threshold due to the narrow window for slow-electron transmission formed below the energy of the resonance. Thus, the entire low-potential spectrum is formed by pre-eminent collection rather than rejection of the slowest electrons.

The simple formulae of [22, 23] for the diffusion-dominated anode current are only applicable to a field-free \( g_2-A \) region. A generalization of these results for \( \Delta a \neq 0 \) was presented by McMahon [24]. This extended treatment is used in [3] in order to model the effect of the elastic cross section on the formation of our moving peaks at various pressures and to demonstrate that \( \Delta a \) must be zero or slightly positive for optimal operation.

5.2. The stationary spectrum

Regardless of the secondary current of slow electrons, the primary beam (what is left of it) is not prevented from reaching the anode and its proper variations compose added structure amongst the slow-electron peaks, as in figures 2–4. But this structure is due to inelastic collisions occurring in the high-potential drift space as evidenced by their ‘standing’ nature.

A second experiment is being performed simultaneously, but a retarding potential is now interposed between the scattering chamber and the detector. As a result, critical potentials are displayed by the complementary detection mechanism as conceived originally, namely the removal of the slower electrons.

A pure standing spectrum is shown in figure 6. Although the structure is much coarser than in figure 5, it is clear that the pattern formed by \( d, e, f, g, \ldots \) is a loose sketch of the sequence \( 4', 5', 6', 7', \ldots \) of figure 5, with successive periods differing by 6.7 V. Finer standing spectra are obtainable with lower values of the filament voltage \( V_f \).
We also see that a dip labelled $z$, $z'$, $z''$, better displayed in figure 2, shows up in the standing spectrum. Its position at 0.5 eV or so below $d$, $d'$, $d''$ suggests that $z$ may be due to the structure at $E \approx 12.8–13$ eV exhibited in [9, 13, 14]. Analogues of $z$ and $z'$ in the pure moving spectrum would be situated, inconveniently, inside the pronounced pits preceding peaks $4'$ and $4''$. On occasion, they have indeed been seen to stand out in curves like figure 5.

6. Untangling the energy scale: the final step

There is still a serious difficulty concerning the absolute location of the pattern of structure in figure 6 following feature $a$. If peak $d$ is the analogue of feature $4'$, then it should be close to 13.4 V on the scale of the $g_1$ potential. Instead, we find that it occurs at $V_a = 10.8$ V in figures 2–4 and slightly higher (11 V) in figure 6.

We are forced to infer that a sudden leap of 2.4–2.6 V in the $V_a$ scale has taken place between $a$ and $d$ so that the space potential has taken the form shown at (b) in figure 7. A similar instance of abrupt vitiation of the energy scale due to space charge, also with mercury, has been observed elsewhere with an entirely different apparatus [25]. Remarkably, the jump of potential reported there (2.7 V) is very close to our values.

The signature for the occurrence of a leap in $g_1$ potential is structure $A$, the clear discontinuity in anode current at $V_a \approx 9.5$ V of figures 2 and 3, which is weaker but still discernible in figures 4–6. The higher the cathode emission, the stronger is feature $A$ and the closer to 2.7 V is the jump in accelerating potential. Conversely, it was observed that if the cathode is insufficiently activated, feature $A$ is smoothed out, and the entire pattern comprising $b$, $c$, $d$, ..., moves to the right. Thus $d$ advances past 10.8 V. Even so, this peak never reaches 13.4 V, which shows that a significant rise of potential always takes place. Observed values of $\delta$ range from 1.6 to 2.7 V.

Evidently, the uphill region is subject to large potential variations between $a$ and $A$ so that no standing peaks are seen in this interval. The sudden reappearance of standing peaks after feature $A$ indicates that a sudden transformation takes place at that point which creates the field-free plateau (b) of figure 7. A detailed scenario describing this phenomenon is given in [4].

The essence of this is that the distance between peak $a$ and the subsequent structure of the high-potential spectrum is affected by the magnitude of the emission current. No wonder then that the identification of features has been haphazard in the past: of all the peaks that could have been chosen by earlier authors as the reference for reading the spectrum, peak $a$ is actually the most unsuitable!

Consider, for instance, the excitation curve of Liu [7]. Starting from peak number ‘2’, the entire spectrum of [7], including the four unidentified peaks, is found to agree perfectly with the structure of figure 5, if Liu’s feature ‘4’ is taken as the counterpart of our peak $4'$ (13.4 eV). The location of the first maximum (peak ‘1’) in Liu’s curve indicates that a negative cpd of 1.65 V exists between the cathode and $g_1$. It follows that Liu’s peak ‘4’ (located at 13.25 V) is really at 11.6 V (13.25–1.65) on the $g_1$-potential scale. This shows that a jump of potential of 1.8 V has taken place somewhere between the first and the second peak.

Summing up, it seems that we can safely infer that:

1. onwards from peak $b$, the sequence of maxima—or should we say downturns?—of the stationary spectrum is identical to the sequence of peaks of the moving spectrum, and
2. peak $a$ is unsuitable for calibrating the $g_1$ energy scale because a jump of space potential inevitably occurs before the emergence of the remaining standing structure.

7. Space charge phenomena

Potential configurations resembling feature (a) of figure 7 are well known in plasma physics. A narrow zone encompassing most of the potential drop between wide field-free regions at
different potentials is created by the strife among the various components of space charge to attain stress balance. Such zones are composed of two adjacent sheaths of opposite space charge excess and are thus called ‘double layers’ or simply DLs. Several kinds of DLs have been studied with various devices in laboratory experiments, computer simulations and in the formation of auroras [26, 27].

A bulging high-potential region similar to that of (b) in figure 7 is often seen to accompany DL formation in gaseous discharge experiments (see for instance [27], p 65). At the pressure used in this experiment, this phenomenon always takes place. As a result, the high-potential plasma is inevitably above the potential of grid 2, regardless of the value of $\Delta g$. Hence the drift space producing the standing spectrum is always followed by a decelerating space charge sheath.

In contrast, no significant jump analogous to $\delta$ (namely a uniform depression below $g_2$ potential followed by an ascending sheath) takes place on the low-potential side. This asymmetry is due to a fundamental difference in the dynamics of formation of the high- and low-potential plasmas [4].

The superior detection mechanism involving the low-energy shape resonance can only operate if the drift space preceding the detection chamber is very close to $g_2$ potential. Therefore, the success of this experiment hinges upon the creation and subsistence of the low-potential plasma. The presence of this plasma requires (a) an adequately large beam current density and (b) a sufficiently negative intergrid potential.

The current density requirement is understandable: the cathode beam produces plasma at the exit of the electron gun and must remain strong enough after traversing this region to create a second plasma close to the outer grid. The initial formation of this plasma occurs just below feature 2. At that point the beam energy at $g_2$ exceeds 4.7 eV and thus positive ions become available in that neighbourhood by cumulative ionization thanks to the metastable nature of the $6^3P_0$ state.

The problem is that shortly afterwards, at feature A, a strong transformation takes place on the high-potential side which produces the field-free plateau of figure 7(b). The charge density in this region increases considerably because single-impact ionization becomes energetically possible. As a result, the DL is pushed outwards and the useful low-potential plasma tends to disappear. This is where the requirement on $\Delta g$ comes in.

One can estimate the lowest value of the DL amplitude $\psi = |\Delta U + \Delta g| + \delta$ that will sustain the display of moving peaks. As we saw in section 3, for both plasmas to be present we need $\Delta g < 0.4$. Assuming that $\delta \cong 2.6$ and $\Delta U \cong -2.5$ we find that, past feature A, $\psi > 4.7$.

The remarkable closeness of this lower bound to the excitation potential of the $6^3P_0$ state suggests that the values of $|\Delta V|$ sustaining the presence of a DL inside the intergrid volume should exceed the threshold of production of metastable atoms. In that case an additional internal discharge is initiated: a fraction of the slow electrons produced by inelastic scattering in the low-potential region are accelerated backwards towards the high-potential plasma by the DL. When the amplitude of the DL exceeds 4.7 V, the threshold for cumulative ionization, the backflowing beam can confine the high-potential plasma by modifying the composition of space charge at the expanding front.

8. The influence of gas pressure and electron emission current

Irrespective of the value of $\Delta g$, the intensity and definition of the structure displayed in our curves are conditioned by the pressure of mercury vapour and by the state of activation and temperature of the cathode. This section is a description of the changes induced by varying each of these three parameters assuming that a value of $\Delta g$ in the optimal range ($-1.5$ to $-1.8$) is operant.
8.1. Mercury vapour pressure

At room temperature, at most some frail structure corresponding to peaks 2–4 is discernible in the anode current, feebly modulating a smoothly rising curve.

As the oven temperature is raised, the average anode current is attenuated, but the overall form of the plots is not affected. An end to this ‘smooth’ regime takes place abruptly at about 110° when the sudden drop at feature A appears in the curve. Structure is now clearly observable in the portion of the curve past the break A. Near 120° the features become narrower and more numerous. At 130° the peaks are clear enough for the curve to look like a weak replica of figure 5 but the lowest excitation peak displayed is peak 4. The curve then evolves into its optimal form, attained, as in figure 5, around 160°. Above this temperature no significant changes are observed until, past 180°, the peaks become narrower and their amplitude diminishes perceptibly.

Feature 1 is brought to view as a peak followed by a clear dip at about 120°. Below this temperature it looks more like an inflection and is very sensitive to $\Delta a$.

8.2. Cathode state of activation

At any value of filament voltage, the emission characteristic of the cathode is dependent on its state of activation; a measure for the latter, denoted $I_c$, was defined in section 2. If a tube remains idle for several weeks, $I_c$ dwindles, sometimes by as much as 50-fold. Such ‘poisoning’ [28] of the cathode prevents the occurrence of the moving spectrum despite the application of a negative value of $\Delta g$.

Fortunately, a process for reactivating a poisoned cathode was discovered in the course of the early tryouts. In an attempt to determine the effect of higher values of filament voltage $V_f$, it was observed that for $V_f \cong 7$ V, as the voltage $V_{hel}$ of the accelerating helipot is cranked upwards, a critical value, say $V_{osc}$, of the voltage $V_a$ across $g_1$ is reached whereupon an oscillatory behaviour sets in: instead of growing past this value on the graph, $V_a$ drops suddenly by a few tenths of volt, rises less rapidly back to $V_{osc}$, drops again abruptly and so on. When $V_f$ is subsequently lowered, stable behaviour is restored, but upon cooling the tube, $I_c$ is found to have grown at least tenfold, an obvious indication that substantial activation of the cathode has taken place. Of course, raising $V_f$ can increase the current by orders of magnitude and might thus appear to be hazardous for the life of the tube; all the same, this procedure has been employed time and again with no visible damage, whenever reactivation was desirable.

No useful activation has ever been achieved unless oscillations actually set in. The procedure is most effective when $T$ and $V_f$ exceed certain limits, which seem to depend on the degree of poisoning; usually $T > 170°$ works best and $V_f \cong 7$ V will suffice, but occasionally higher values of $V_f$ (up to 9 V) have been necessary.

The minimal value of $V_a$ at which a poisoned tube will oscillate is usually the voltage of feature A (about 9.5 V) but it can be as high as 11–12 V in cases of severe poisoning. To reactivate, $V_{hel}$ should be raised until oscillations set in and then be held fixed when the rate of oscillations has reached about 1 Hz. If reactivation is becoming effective, then upon bringing $V_{hel}$ back to zero and gradually increasing it again, it will be found that the value of $V_{osc}$ has retracted to 9.5 V. When good activation is finally attained, the oscillating regime may even be triggered at peak a.

For our purposes, activation is optimal when $I_c$ is between 1000 and 1500 $\mu$A, but values as low as 600 $\mu$A have been seen to be acceptable.

Activation is not always a rapid procedure. In one of our two model-55580 tubes, called tube I, a dozen or so oscillations at feature A have commonly been sufficient, even when poisoning is severe. Tube II, however, usually had to be left oscillating for minutes or even hours for $I_c$ to recover adequately. Two older-model tubes (Leybold–Heraeus 55080) which had remained inactive for several years could not be activated beyond $I_c \cong 200$ $\mu$A, which was inadequate for producing the moving spectrum with any tube at any value of $\Delta g$. 
Once a tube’s cathode has been sufficiently activated, $I_c$ stays remarkably constant for hours on end. And although $I_c$ diminishes somewhat, usually no reactivation is required for several months if the experiment functions frequently. When a tube was left inactive for months, poisoning invariably took place.

An oxide cathode is a metal core, usually nickel, coated with a mixture of the alkaline earth oxides CaO, SrO, BaO. Activation consists in the formation in the coating, by ion impact [28], of excess atoms of Ca, Sr and Ba. The sequence of events underlying the phenomenon of relaxation oscillations probably breaks up space charge limitation of the ion current flowing to the cathode and produces puffs of increased ion bombardment which induce activation in the coating [4].

8.3. Filament voltage

When electrons escape from a thermionic cathode, their velocity distribution is half-Maxwellian. Hence its width is proportional to the temperature of the cathode and we would expect the energy resolution to improve with decreasing filament voltage $V_f$. This effect is indeed perceptible: it was verified that the peaks in the curve of figure 5 are conspicuously sharper than they would have been, had we used a higher value of $V_f$. On that basis it would appear that lowering $V_f$ would be profitable. In reality, this is only partly true. As $V_f$ is reduced below 3.5 V, the peaks past feature 4 in figure 5 are indeed sharpened. On the other hand, however, in the portion of the curve below peak 4, the resolution declines.

Insofar as it is desirable to display a maximum of structure throughout the energy range with good resolution, the setting of $V_f$ should be the best compromise between the narrowing of the energy distribution and the smearing of the early moving peaks. In both our tubes the best range was between 3.3 and 3.6 V.

9. Conclusion

This mode of medium-pressure electron impact spectroscopy was introduced in the literature more than 80 years ago. It is surprising, therefore, that despite the progress of physics, modern accounts [2, 7] of the experiment do not differ substantially from those of the original authors. Perhaps the reason for this is that while the scope of the experiment is in atomic physics and quantum mechanics, its conception really belongs to the plasma field of gaseous discharges, a complex and less well known subject.

As a result, it is not easy to make this a routine experiment for undergraduate university laboratories, unless an apparatus is developed that provides completely straightforward access to plasma formation at the second grid. Among the types of Leybold tubes, only the 55580, duly conditioned, is guaranteed to produce the moving spectrum. Older tubes did not, as I pointed out. A newer-model tube (55585) acquired in 1987 was briefly tried with no success. Since then the author has not carried out experimental work on this matter.

Readers whose tube is only able to produce the standing excitation spectrum are at least supplied, for the first time, with the interpretation of the observed critical potentials, and with the correct way to calibrate the energy scale. In addition, direct observation of the resonance at feature 1 should always be possible with non-negative $\Delta a$ and negative $\Delta g$.

Acknowledgments

I am grateful to Paul Burrow for several motivating overseas conversations in the early stages of this venture and for sending me an unpublished electron transmission spectrum of mercury. As to plasma double layers, I am indebted to Jean Wallenborn for introducing me to the literature and to Hans Schamel for teaching me about them and for his constant concern. Of course without Martine this work would never have been completed.
References

[3] Nicoletopoulos P The 0.4 eV shape resonance of electron scattering from mercury in a Franck–Hertz tube, to be submitted
[22] Langmuir I 1931 Phys. Rev. 38 1656
[28] Herrmann G and Wages S 1951 The Oxide-Coated Cathode vol 2 (London: Chapman and Hall) ch 6