

# A new approach to optimize pH buffers

Marc Blétry\*

University of Tours, PCM2E, France

**Abstract** Buffer solutions are pervasive in chemistry, biochemistry, analytical chemistry, etc. A better understanding of buffer properties and what controls them is susceptible to be of interest in many scientific and technological fields. For instance, linear pH gradients are commonly used in electrophoresis and their optimization rests on numerical optimization of the concentrations of various weak species. It is probably generally assumed that no basic progress could be made on optimization approaches. We introduce here a new strategy to buffer optimization, based on a parametric study of the roots of the first derivative of the buffer index. In this way, it is possible to find mathematically optimal sets of parameters (pKa and concentrations). The method is applied to mixtures of 2, 3 and 4 monovalent species, which represent simple cases that do not call for overly elaborate numerical optimization techniques, but are nevertheless of practical interest in various branches of analytical chemistry.

**Keywords** Buffer solutions ; Optimization ; Constant buffer index ; Monovalent species

## 1 Introduction

Buffers are of primary importance in many fields, from biology [1] to geochemistry [2] as a scientific notion, but they also are used in technological applications—among many, one can cite water treatment [3], catalytic reduction [4, 5], drug solubility [6], control of enzyme activity [7], polarography [8, 9, 10] and, in analytical chemistry, atomic absorption spectrophotometry [11], chromatography [12, 13], electrophoresis [14, 15] – in particular the use of immobilized pH gradients in electrophoresis [16, 17], binary buffers for cell electrophoresis [18] or as an important parameter influencing peak spreading [14]. In all these practical uses, it is important to obtain as good a control as possible over the buffer index, sometimes across several pH units.

In particular, the search for an even buffer capacity across large pH ranges is primordial for the production of linear pH gradients. Mathematical optimization of such mixtures rests mostly on objective functions minimization, typically through least square methods, by adjusting the concentrations of known buffer substances [19, 20, 21]. On the other hand, theoretical studies of pH mixtures that lead to real values prediction has mostly been conducted analytically, greatly limiting the situations that could be dealt with as the complexity of the system grows extremely rapidly with the composition [22, 23, 24, 25, 26, 27]. Such studies are not primarily interested with actual buffers but rather with the mathematical conditions that drive a buffer system. In this way, it is possible, for instance, to find the conditions in terms of dissociation constants (or  $pK_a$ ) and concentrations that guarantee the largest possible pH range of even buffer index for a two weak acid mixture, provided that one can neglect the contribution of  $H_3O^+$  and  $OH^-$ . One can, analytically, hardly go much further given the degree of the master polynomials that drive the more complicated systems—this is the case also for complexation equilibria where overlapping involving two steps already leads to intricate calculations [28], even though they do not involve formally an equilibrium with the solvent. Hence, it is hopeless to expect analytical solutions that could predict the best chemical properties to produce constant buffer index over large pH ranges—i.e. more than 1 pH unit, roughly, as going further necessarily involves using more than two chemical components. For that matter, it is even hard to guarantee finding the best mathematical conditions to reach the flattest possible index buffer curve for extreme pH, as then the contribution of  $H_3O^+$  or  $OH^-$  cannot be neglected anymore, which renders the system impossible to solve analytically for two weak acid mixtures.

General formulation of buffer capacity  $\beta$  have been proposed (see, for instance, [22, 29, 30]) and Urbansky and Schock have suggested the use of the derivative of the buffer index  $\beta'$  to locate critical points [31]. However, to the best of our knowledge, this suggestion has not been the object of extended

---

\*mbl@univ-tours.fr, Bâtiment J, UFR Sciences et Techniques, Parc de Grandmont, 37200 Tours, FRANCE

study, even though it turns out to bring interesting results for buffer optimization in the framework of a parametric study.

The object of the present paper is to introduce a new approach, resting on a numerical analysis of the roots of the  $\beta'$  function, which allows finding the optimal mathematical conditions to produce flat buffer index curves. It has been applied to solutions of 2, 3 and 4 weak acids, which are simple enough to be analyzed without numerical optimization techniques.

## 2 Preliminary considerations and notations

The buffer index,  $\beta$ , was first proposed by Van Slyke [32, 33] as:

$$\beta = \frac{dB}{dpH}$$

where  $B$ , in Van Slyke's original paper, was the concentration of strong base in gram equivalent per L, whereas here we will instead use molarity to define  $\beta$ , as is generally done.

In what follows,  $pK_a$ 's will be numbered  $pK_{a_1}$ ,  $pK_{a_2}$ ,  $pK_{a_3}$  and  $pK_{a_4}$  and we will always consider  $pK_{a_1} < pK_{a_2} < pK_{a_3} < pK_{a_4}$ . The associated concentrations may also be numbered  $C_1$ ,  $C_2$ ,  $C_3$  and  $C_4$ . We will note  $\Delta pK_a$  the difference between the highest  $pK_a$  values and  $pK_{a_1}$  (as it is the smallest one by definition here). We will neglect ionic strength effects, hence activities reduce to concentrations in the equilibrium constants, and we will use the shorthand notation  $[H_3O^+] = h$ .

## 3 Buffers of two weak acids

This section deals with solutions of two weak acids with their acidity constant noted  $pK_{a_1}$  and  $pK_{a_2}$  and concentrations  $C_1$  and  $C_2$ . We will note the average  $pK_a$ ,  $\bar{pK}_a = (pK_{a_1} + pK_{a_2})/2$ . Such mixtures have already been studied at great length. To quote two particularly relevant contributions, one should note that Ricci has studied the presence of a minimum of  $\beta$  in the pH interval between  $pK_{a_1}$  and  $pK_{a_2}$ , and provided conditions for its existence when neglecting the contribution of  $H_3O^+$  and  $OH^-$  (see pp. 185-191 in [22]). Rilbe ([26]) has shown that this minimum disappears when  $pK_{a_2} - pK_{a_1} \leq 1.14$  for equimolar mixtures, also when neglecting the contributions of  $H_3O^+$  and  $OH^-$ .

In the present section, we will first study the equimolar case without taking  $H_3O^+$  and  $OH^-$  into consideration in the relationship for  $\beta$ , and find the same result as Rilbe. Then, we study the mixtures of two weak acids of different concentrations and establish for which value of  $\Delta pK_a = pK_{a_2} - pK_{a_1}$  the minimum exists or not and, finally, we study the effect of the contribution of  $H_3O^+$  and  $OH^-$  and show how to determine the optimal values of concentrations and  $pK_a$ 's in order to obtain a flat buffer index curve. All this will be done using numerical methods, as they can be extended to more complicated mixtures and to the inclusion of  $H_3O^+$  and  $OH^-$ , which is already difficult analytically for a mixture of two weak monovalent species.

### 3.1 Neglecting the contribution of $H_3O^+$ and $OH^-$

#### 3.1.1 Two weak acids with the same concentration

In the case of two weak acids with the same concentration  $C$ , the buffer index is given by:

$$\beta = \ln(10)C \left( \frac{K_{a_1}h}{(K_{a_1} + h)^2} + \frac{K_{a_2}h}{(K_{a_2} + h)^2} \right)$$

and its derivative  $\beta'$ :

$$\beta' = \frac{d\beta}{dpH} = \ln^2(10)Ch \left( \frac{K_{a_1}(h - K_{a_1})}{P_1^3} + \frac{K_{a_2}(h - K_{a_2})}{P_2^3} \right)$$

with  $P_1 = (K_{a_1} + h)$  and  $P_2 = (K_{a_2} + h)$ . This derivative is zero when

$$P(h) = K_{a_1}(h - K_{a_1})(K_{a_2} + h)^3 + K_{a_2}(h - K_{a_2})(K_{a_1} + h)^3 = 0$$

which does not depend on the concentration of the conjugated species as it could be factored out. By finding the roots of this polynomial in, say, the interval  $pH \in [0 : 14]$ , one can determine the number of

maxima and minima of the buffer index and the values of the corresponding roots. Figure 1.a represents the  $pH$  values for which  $\beta' = 0$  in the case where  $\overline{pK_a} = 4$  as a function of  $\Delta pK_a = pK_{a2} - pK_{a1}$ <sup>i</sup>. The first obvious point is that a single root is always present, that for  $pH = \overline{pK_a}$ . A point at  $\Delta pK_a^c = 1.1439$ <sup>ii</sup> corresponds to the transition from the one to the three roots regime of the buffer index curve and matches the flattest buffer index that can be obtained at  $pH = 4$  using two weak acids [26, 25], in this case with  $pK_{a1} = 3.428$  and  $pK_{a2} = 4.5720$  (see figure 1.b). For  $\Delta pK_a < 1.1439$  the single root in the derivative of  $\beta$  corresponds to a maximum, whereas it becomes a minimum surrounded by two maxima for  $\Delta pK_a > 1.1439$ .

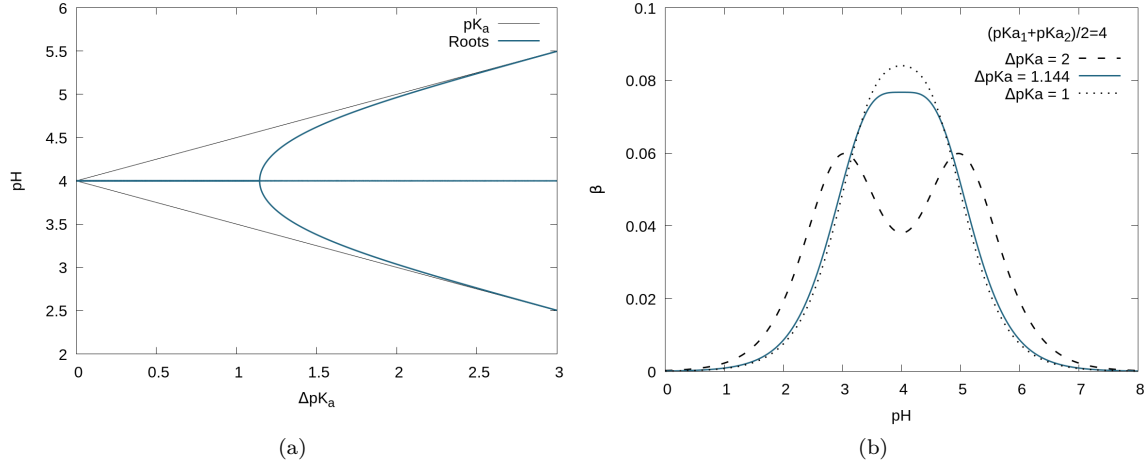


Figure 1: a) Values of  $pH$  for which the derivative of the buffer index cancels. The number of roots goes from 1 to 3 at  $\Delta pK_a = 1.1439$ . b) Buffer index for three sets of  $pK_a$  with average  $pK_a = 4$ .

### 3.1.2 Two weak acids with different concentrations

In the case of two weak acids with two different concentrations, the buffer index is given by:

$$\beta = \ln(10)C_1 \left( \frac{K_{a1}h}{(K_{a1} + h)^2} + \frac{rK_{a2}h}{(K_{a2} + h)^2} \right)$$

with  $r = C_2/C_1$ . Its derivative writes:

$$\beta' = \frac{d\beta}{dpH} = \ln^2(10)C_1h \left( \frac{K_{a1}(h - K_{a1})}{P_1^3} + \frac{rK_{a2}(h - K_{a2})}{P_2^3} \right)$$

with  $P_1 = (K_{a1} + h)$  and  $P_2 = (K_{a2} + h)$ .

We observe a significant change in the behaviour of the roots of  $\beta'$  when  $C_1 \neq C_2$ . Indeed, as shown in figure 2.a, the triple point disappears and is split in two, even if there still exists a 1- and a 3- roots regime, with a transition point for a given value of  $\Delta pK_a^c$  that appears dependent upon the ratio  $r = C_2/C_1$  (see figure 2.b) but not on the magnitude of the concentrations. Here, as  $r < 1$ ,  $C_2$  varies below  $C_1$ . In that case, the single root regime corresponds to a maximum that is closer to  $pK_{a1}$  than to  $pK_{a2}$  and not anymore equal to  $\overline{pK_a}$ . However, it logically goes to 4 when  $\Delta pK_a \rightarrow 0$ . In the three roots regime, the "middle root" is above  $\overline{pK_a}$  and goes to  $\overline{pK_a}$  as  $\Delta pK_a$  increases beyond  $\Delta pK_a^c$ . The situation is perfectly symmetrical when  $C_1$  is below  $C_2$ .

As shown in figure 2.b, the value of the split triple point,  $\Delta pK_a^c$ , increases as the difference in concentrations increases. When this split occurs (i.e. for  $\Delta pK_a = \Delta pK_a^c$ ), the buffer index of the solution presents a rather flat regime around  $pH = pK_{a2}$ , as can be seen in figure 2.c, but  $pH = pK_{a1}$  appears as a clear maximum of the buffer index. Below the critical value, the buffer index presents a single maximum and above that value, two maxima and a minimum are present, roughly and respectively for  $pH = pK_{a1}$ ,

<sup>i</sup>We chose that value of  $\overline{pK_a}$  for no particular reason as, in the considered situation, this curve is independent from this parameter. At this stage, we could have used reduced variables as proposed by Rilbe [23], however when we will consider the contribution of  $H_3O^+$  and  $OH^-$ , it will not be possible to use such variables anymore so we prefer this homogeneous presentation of all equations studied here.

<sup>ii</sup>A similar case can be made considering a solution of a weak diacid, in which case  $\Delta pK_a^c = 1.204$ , see [23, 27].

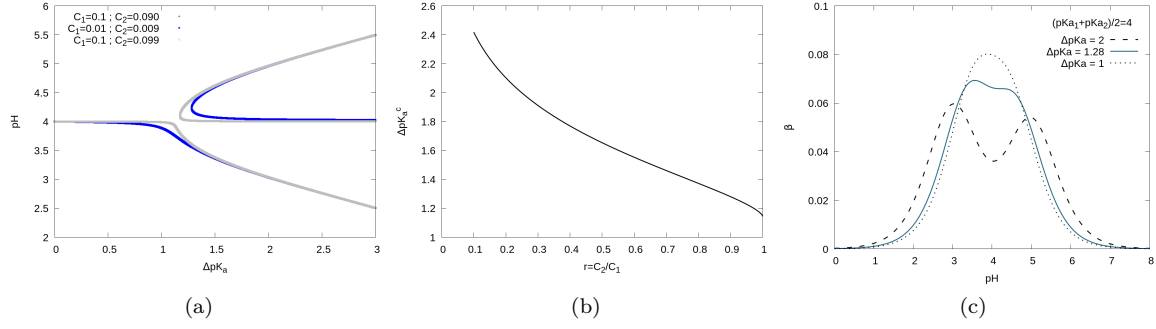


Figure 2: a) Values of  $pH$  for which the derivative of the buffer index cancels for different values of  $C_1$  and  $C_2$ , the curves with the same  $r$ -value are perfectly superimposed. b) Dependency of the triple point upon the ratio  $r = C_2/C_1$ . c) Buffer index of solutions of weak species with  $(pK_{a1} + pK_{a2})/2 = 4$ ,  $C_1 = 0.1 \text{ mol.L}^{-1}$  and  $C_2 = 0.09 \text{ mol.L}^{-1}$  – in that case,  $\Delta pK_a^c = 1.28$ .

$pK_{a2}$  and  $\overline{pK_a}$  (the exact values of these extrema must be computed as a function of  $r$  and present a dependency similar to the ones presented in figure 2.a).

### 3.2 Including the contribution of $H_3O^+$ and $OH^-$

Now, we include the contribution of  $H_3O^+$  and  $OH^-$ . In that case, the buffer index writes:

$$\beta = \ln(10) \left( \frac{K_w}{h} + h + \frac{C_1 K_{a1} h}{(K_{a1} + h)^2} + \frac{C_2 K_{a2} h}{(K_{a2} + h)^2} \right)$$

and its derivative is:

$$\beta' = \frac{d\beta}{dpH} = \ln^2(10) \left( \frac{K_w P_2^3 P_1^3 - h^2 P_2^3 P_1^3 + h^2 (C_1 K_{a1} (h - K_{a1}) P_2^3 + C_2 K_{a2} (h - K_{a2}) P_1^3)}{h P_1^3 P_2^3} \right)$$

which cancels out if the numerator cancels, obviously, as above.

#### 3.2.1 Two weak acids with the same concentration

We first consider the case where  $C_1 = C_2$ . As is shown by figure 3, the triple point is broken when the contributions of  $H_3O^+$  and  $OH^-$  are taken into account even if  $C_1 = C_2$ , as long as  $\overline{pK_a} \neq 7$ . This behaviour is symmetrical whether this parameter is above or below 7, and these curves appear very similar to the ones obtained for  $C_1 \neq C_2$  when neglecting  $H_3O^+$  and  $OH^-$  (see figure 2). The further  $\overline{pK_a}$  is from 7, the stronger the divergence from the  $\Delta pK_a^c$  curve established in the absence of  $H_3O^+$  and  $OH^-$  contributions. Moreover, these curves become dependent upon the concentration of the buffer solutions, as it cannot be factored out: the higher the dilution, the higher  $\Delta pK_a^c$ , as in the case of solutions with higher differences in concentrations.

Figure 3.c illustrates how the buffer index evolves for values of  $\Delta pK_a$  above, equal or below  $\Delta pK_a^c$  in the case where  $C = 0.1 \text{ mol.L}^{-1}$  and  $\overline{pK_a} = 3$  (here,  $\Delta pK_a^c = 1.249$ ).

#### 3.2.2 Two weak acids with different concentrations

Now, it becomes possible to compensate for the asymmetry due to the  $H^+$  or  $OH^-$  contributions by changing the concentration of either species in order to improve the buffer index in the target  $pH$ -area, which is obtained by converging towards a set of concentrations  $C_1$  and  $C_2$  such that the single triple root is restored. In practice, here, we reached that goal by setting the value of  $C_1$  and then optimizing  $C_2$ .

As an example, we treat the case of the solution seen above, with  $C_1 = 0.1 \text{ mol.L}^{-1}$  and as target  $pH = 3$ . In fig. 4.a are represented the roots in function of  $\Delta pK_a$  for  $C_1 = C_2 = 0.1 \text{ mol.L}^{-1}$  – the starting point – and  $C_1 = 0.1 \text{ mol.L}^{-1}$ ,  $C_2 = 0.1104 \text{ mol.L}^{-1}$ , when the convergence was satisfying. In that situation,  $\Delta pK_a^c = 1.099 \approx 1.1$ , value that differs significantly from 1.144 found in the absence of the  $H_3O^+$  contribution (which is the important one at  $pH = 3$ , where  $OH^-$  can safely be neglected) or the one found for  $C_1 = C_2 = 0.1 \text{ mol.L}^{-1}$  ( $\Delta pK_a^c = 1.249$ ).

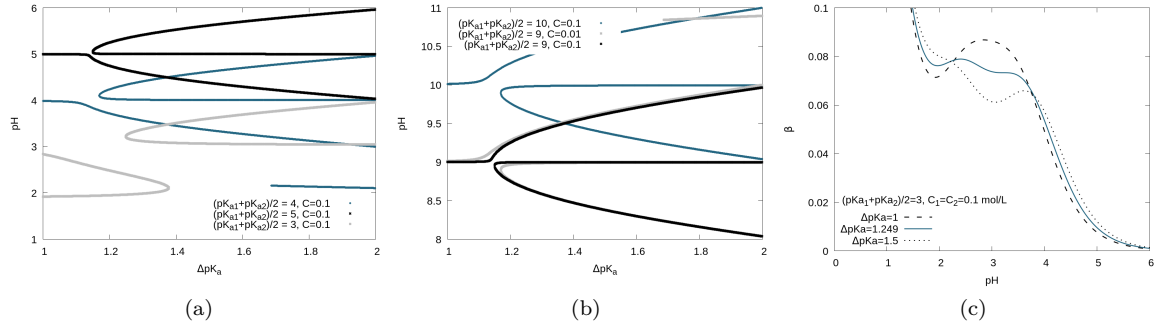


Figure 3: Value of the roots as a function of  $\Delta pK_a$  for a) low and b) high values of  $pK_a$ . c) Buffer index for 3 solutions with the same concentrations ( $C = 0.1$  mol.L<sup>-1</sup>) and varying  $\Delta pK_a$ ; for these concentrations and  $\overline{pK_a}$ ,  $\Delta pK_a^c = 1.249$ .

Figure 4.b shows the dependency of the buffer index with  $pH$  for various sets of  $C_1$ ,  $C_2$  and  $\Delta pK_a$ . The benefit of selecting the right value of  $C_2$  and  $\Delta pK_a = \Delta pK_a^c$  with  $C_1$  set is obvious, as in that case  $\beta$  turns out to be perfectly flat over a range of almost 1 pH unit, centered around 3.

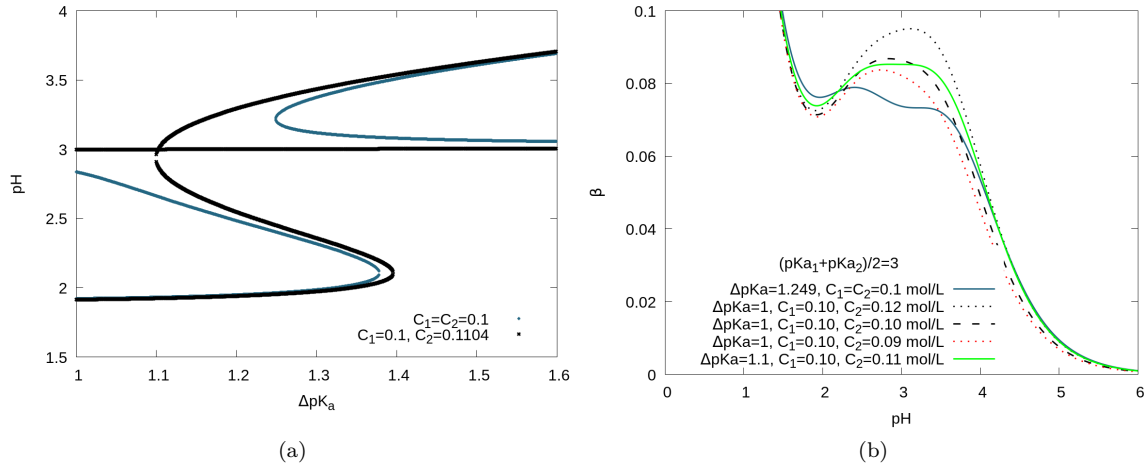


Figure 4: a) Value of the roots as a function of  $\Delta pK_a$  for optimized and non optimized values of  $C_2 = 0.11$  mol.L<sup>-1</sup>,  $C_1$  is set at  $0.1$  mol.L<sup>-1</sup>. b) Buffer index for various solutions with different concentrations and  $\Delta pK_a$  ( $C_1 = 0.1$  mol.L<sup>-1</sup>).

## 4 Three weak acids

We do not know of any detailed study dealing with mixtures of three weak acids in detail. The closer situation we found was that of trivalent species. In his study of the buffer capacity of trivalent protolytes, Rilbe found no value of  $\Delta pK_a$  that allowed a flat dependency of the buffer index [24]. As a matter of fact, this is only possible if the three  $pK_a$  correspond to monovalent species of different concentrations, as we are about to see, and this is obviously impossible for trivalent species.

We consider here the situation without taking into account the contribution of  $H_3O^+$  and  $OH^-$ , as we want to keep to situations that allow simple optimization. Otherwise, more advanced and fully automated numerical developments would be needed, which could be the object of future studies. Hence, the results established here will be of interest in pH and concentration ranges where these contributions can be neglected.

#### 4.1 Isoconcentration of extreme $pK_a$ and centered $pK_a$ intervals

Here, we take  $C_1 = C_3 = C \neq C_2 = C_m$  and  $r = C/C_m$ . We also note  $\Delta pK_a = pK_{a3} - pK_{a1}$  and  $\overline{pK_a} = pK_{a2} = (pK_{a1} + pK_{a3})/2$ . In that case, it comes:

$$\beta = \ln(10)C_m \left( \frac{rK_{a1}h}{(K_{a1} + h)^2} + \frac{K_{a2}h}{(K_{a2} + h)^2} + \frac{rK_{a3}h}{(K_{a3} + h)^2} \right)$$

and its derivative  $\beta'$ :

$$\beta' = \frac{d\beta}{dpH} = \ln^2(10)C_m h \left( \frac{rK_{a1}(h - K_{a1})}{P_1^3} + \frac{K_{a2}(h - K_{a2})}{P_2^3} + \frac{rK_{a3}(h - K_{a3})}{P_3^3} \right)$$

where  $P_1$ ,  $P_2$  and  $P_3$  are defined as in the previous section.

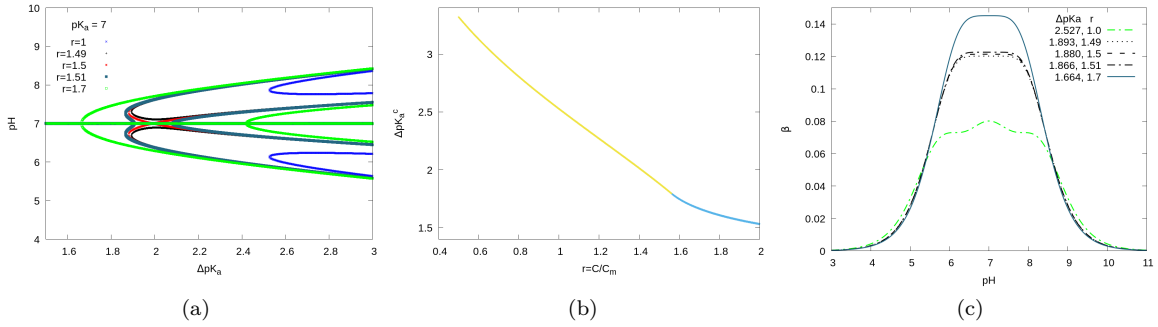


Figure 5: a) Roots for  $\overline{pK_a} = 7$  and various values of  $r$ . b) Dependency of  $\Delta pK_a^c$  with  $r$  – colours correspond to the number of roots for  $\Delta pK_a = \Delta pK_a^c + \varepsilon$ . c) Buffer index of various solutions at  $\Delta pK_a = \Delta pK_a^c$  as determined in figure a).

Figure 5.a shows the dependency of the roots as a function of  $\Delta pK_a$  for various  $r$ . First of all, for low values of  $\Delta pK_a$ , there always exists a single maximum and for  $\Delta pK_a > \Delta pK_a^c$ , multiple roots appear. The value of  $\Delta pK_a^c$  depends on  $r$  (see figure 5.b). Additionally, one can note that the multiplicity of the root evolves also with  $r$ . For large  $r$  values, as  $\Delta pK_a$  increases, there actually exists a domain with a single root, then with 3 roots and finally with 5 roots. In contrast, for low values of  $r$ , the transition goes straight from a single to a 5 roots situation. The value of  $r$  found to be the transition between these two types of behaviour is  $r \approx 1.56$ . Another value of interest turns out to be the one for which the intermediate branches come to touch the  $pH = 7$  solution, observed at  $r = 1.5$ . Indeed, figure 5.c shows the evolution of the buffer index for solutions of various values of  $r$  and  $\Delta pK_a = \Delta pK_a^c$ . The value of  $r = 1.5$  appears as the transition to a very flat behaviour, when mixtures parametrized with lower values of  $r$  present a non flat behaviour even if with a single maximum<sup>iii</sup>.

#### 4.2 Varying concentration and non centered $pK_a$ intervals

We now take  $pK_{a1} < pK_{a2} < pK_{a3}$  and  $C_l \neq C_h \neq C_m$  and  $r_l = C_l/C_m$ ,  $r_h = C_h/C_m$ . We also note  $\Delta pK_a = pK_{a3} - pK_{a1}$ . In that case, it comes:

$$\beta = \ln(10)C_m \left( \frac{r_l K_{a1} h}{(K_{a1} + h)^2} + \frac{K_{a2} h}{(K_{a2} + h)^2} + \frac{r_h K_{a3} h}{(K_{a3} + h)^2} \right)$$

and its derivative  $\beta'$ :

$$\beta' = \frac{d\beta}{dpH} = \ln^2(10)C_m h \left( \frac{r_l K_{a1}(h - K_{a1})}{P_1^3} + \frac{K_{a2}(h - K_{a2})}{P_2^3} + \frac{r_h K_{a3}(h - K_{a3})}{P_3^3} \right)$$

where  $P_1$ ,  $P_2$  and  $P_3$  are defined as in the previous sections.

<sup>iii</sup>Here, it is illustrated for the value  $r = 1$  and  $\Delta pK_a^c = 2.527$ , more or less corresponding to trivalent species – Rilbe found  $\Delta pK_a^c = 2.58$  for trivalent species, which is rather close [24]. Obviously, these two situations are different and it is no more surprising to find different values of  $\Delta pK_a^c$  than it was when comparing mixtures of two monovalent acids with divalent species in section 3.2.1.

We do not present here a general study of this situation as too many parameters are involved for a simple visualisation, however it is still possible to optimize solutions "by hand" as we exemplify below. In these asymmetrical situations, it becomes handy to work also with  $r_{\Delta pK_a} = \frac{\Delta pK_{a_h}}{\Delta pK_{a_l}}$ , where  $\Delta pK_{a_h} = pK_{a_3} - pK_{a_2}$  and  $\Delta pK_{a_l} = pK_{a_2} - pK_{a_1}$ .

**Optimizing the parameter of a buffer solution** Now, it is possible to optimize a solution of three weak acids with asymmetrical  $\Delta pK_a$  by varying the concentrations of one of the weak acid with the extreme  $pK_a$ , and, in this case, without the contribution of  $H_3O^+$  and  $OH^-$ . This is interesting to obtain as flat a dependency of  $\beta$  with  $pH$  as possible, when using weak acids with irregular  $pK_a$  intervals.

The same procedure that was used in the case of two weak acids (taking into account the contribution of  $H_3O^+$  and  $OH^-$ , see section 3.2.2) allows performing just such a task. To exemplify that point, figure 6.a displays the roots landscape in the case where  $pK_{a_2} = 7$  and for  $r_{\Delta pK_a} = 1$  and 1.1, with or without optimization of  $r_l$  and  $r_h$ . A first situation is simply the one obtained for  $\overline{pK_a} = 7$  and  $r = 1.5$  (and, here,  $r_{\Delta pK_a} = 1$  and  $r_l = r_h$ ) in section 4.1. Then the roots evolve in an asymmetrical manner when  $r_{\Delta pK_a} = 1.1$  is imposed with  $r_l = r_h$ , leading to a skewed plateau of the buffer index. Finally, by optimizing  $r_h$ , it is possible to recover a single multiple roots and the corresponding  $\beta$  curve appears as flat as the first one. The difference in height is due to the fact that the optimization was obtained by increasing  $C_h$ , leaving the other two unchanged, hence the overall buffer index increases.

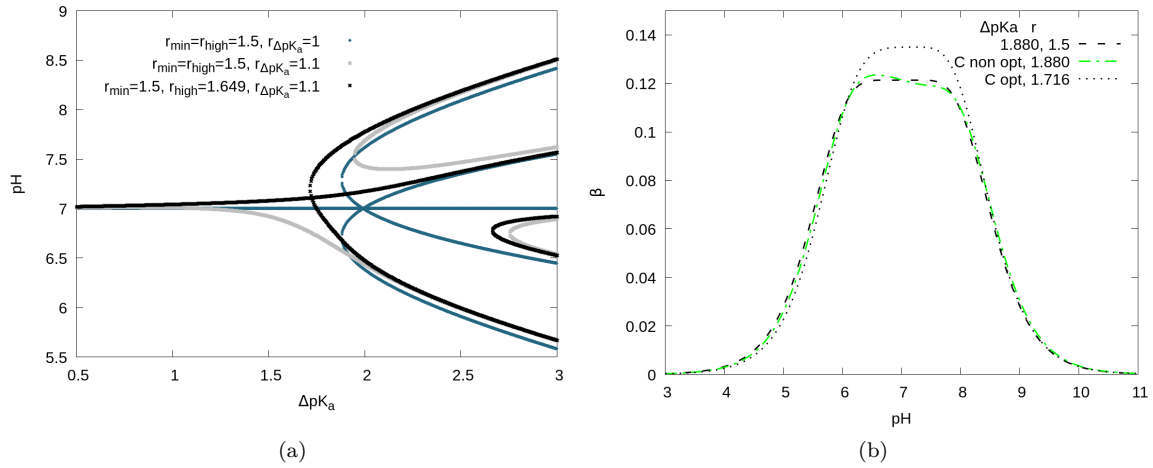


Figure 6: a) Root spaces for non optimized and optimized solutions of three weak acids. b) Buffer index for optimized and non optimized buffer solutions at  $\Delta pK_a = \Delta pK_a^c$ .

## 5 Four weak acids

We could not find any detailed study of mixtures of four weak species in the literature and it turns out that the tools proposed here allow studying this case, provided that we restrict ourselves to "nice" combinations of parameters. Indeed, the four acids mixture represents an interesting and still somewhat simple situation under the following conditions: first, the contribution of  $H_3O^+$  and  $OH^-$  is neglected and, second, we consider a highly symmetrical set of concentrations and  $pK_a$ 's in the sense defined hereafter.

In accordance with what was said before, we take  $pK_{a_1} < pK_{a_2} < pK_{a_3} < pK_{a_4}$ . Additionally, we take  $C_1 = C_4$ ,  $C_2 = C_3 = C$  and  $r_C = C_1/C_2$ , hence  $C_1 = r_C C$ . We also note  $\Delta pK_a = (pK_{a_4} - pK_{a_1})/2$  and we consider that  $\overline{pK_a} = \frac{pK_{a_1} + pK_{a_4}}{2} = \frac{pK_{a_2} + pK_{a_3}}{2}$ . It then comes  $pK_{a_1} = \overline{pK_a} - \Delta pK_a$  and  $pK_{a_4} = \overline{pK_a} + \Delta pK_a$ . Moreover, we have  $r_{\Delta pK_a} = \frac{pK_{a_3} - pK_{a_2}}{pK_{a_4} - pK_{a_1}}$ , which gives  $pK_{a_2} = \overline{pK_a} - r_{\Delta pK_a} \Delta pK_a$  and  $pK_{a_3} = \overline{pK_a} + r_{\Delta pK_a} \Delta pK_a$ . In that case, it comes:

$$\beta = \ln(10)C \left( \frac{r_C K_{a_1} h}{(K_{a_1} + h)^2} + \frac{K_{a_2} h}{(K_{a_2} + h)^2} + \frac{K_{a_3} h}{(K_{a_3} + h)^2} + \frac{r_C K_{a_4} h}{(K_{a_4} + h)^2} \right)$$

its derivative  $\beta'$ :

$$\beta' = \frac{d\beta}{dpH} = \ln^2(10)Ch \left( \frac{r_C K_{a_1} (h - K_{a_1})}{P_1^3} + \frac{K_{a_2} (h - K_{a_2})}{P_2^3} + \frac{K_{a_3} (h - K_{a_3})}{P_3^3} + \frac{r_C K_{a_4} (h - K_{a_4})}{P_4^3} \right)$$

where  $P_1, P_2, P_3$  and  $P_4$  are defined as in the previous section. It will also be interesting to study the second derivative, which writes:

$$\beta'' = \frac{d^2\beta}{dpH^2} = 2\ln^3(10)Ch^2 \left( r_C K_{a_1} \frac{2K_{a_1} - h}{P_1^4} + K_{a_2} \frac{2K_{a_2} - h}{P_2^4} + K_{a_3} \frac{2K_{a_3} - h}{P_3^4} + r_C K_{a_4} \frac{2K_{a_4} - h}{P_4^4} \right)$$

Then, by varying  $r_C$  and  $r_{\Delta pK_a}$ , it becomes possible to map the properties of all mixtures of four weak acids with the same average  $pK_a$ . Figure 7 presents the study of the first critical point for these systems by plotting (a) the value of  $\Delta pK_a^c$  as a function of  $r_C$  and  $r_{\Delta pK_a}$ , (b) the dependency of the number of roots of  $\beta'$  for  $\Delta pK_a > \Delta pK_a^c$  and (c) the corresponding number of roots of  $\beta''$ .

The domain plotted was restricted to an "interesting" area, as for larger values of  $r_C$  and/or  $r_{\Delta pK_a}$ , the behaviour becomes monotonous. Indeed, we observe the existence of a domain on the border of which  $\Delta pK_a^c$  values undergo an inflexion, and this border corresponds to a zone of transition of the number of roots from 5 to 3 for  $\Delta pK_a$  immediately above  $\Delta pK_a^c$ . The maximum of that area is observed at  $r_C = 1.37$  and  $r_{\Delta pK_a} = 0.293$  and it disappears for  $r_C > 3.12$ .

Figures 7.c and d display the behaviour of the buffer index for two values of  $r_C$ . It is clear that exiting the 5 roots domain leads to buffer index curves with a single maximum while at the same time the values of  $\Delta pK_a^c$  decreases as  $r_{\Delta pK_a}$  increases in the 3 roots domain, leading to narrower flat region in the buffer index curve. Hence, the optimal solution must be found on the 5-3 boundary. Figure 7.e and f display respectively the buffer index and its derivative along the critical line for  $\Delta pK_a = \Delta pK_a^c$ . At low values of  $r_C$ ,  $\beta$  actually displays several shoulder that tend to disappear for  $r_C > 1.5$ . In order to optimize more precisely, it becomes necessary to proceed as above, by examining the root space in detail.

Hence, by iteratively searching over the boundary between the 5 and 3 roots domains, we can find an optimal combination, which is obtained for  $r_C = 1.6217$ ,  $r_{\Delta pK_a} = 0.2862$ ,  $\Delta pK_a^c = 2.572$ . It corresponds to a situation where branches start to touch in the root space, as in the 3 weak acids case of section 4.1. Hence, for 4 weak acids mixture the optimal combination found here for an interval centered around  $pH = 7$  is  $pK_{a_1} = 5.714$ ,  $pK_{a_2} = 6.632$ ,  $pK_{a_3} = 7.368$ ,  $pK_{a_4} = 8.286$  and  $r_C = 1.6217$ , as depicted in figure 8.a (root space) and 8.b (corresponding buffer index curves at  $\Delta pK_a = \Delta pK_a^c$ ).

## 6 Discussion and conclusion

In this paper, we have presented the interest that lies in studying buffer solutions in the root space of the derivative of the buffer index. It has allowed us to easily find optimal values of concentrations and  $pK_a$  in the case of 2 and 3 weak acids mixtures, and, with greater difficulty, an optimal point in a symmetrical 4 weak acids case.

As we restricted ourselves to simple optimization techniques, only the case with 2 weak acids could be dealt with while including the contribution of  $H_3O^+$  and  $OH^-$ . Going beyond – for instance finding the optimal combination of parameters for mixtures of 3 weak acids at low or high pH target values – would involve numerical multiparameter optimization techniques, as is done in other approaches by combining stochastic and steepest descent methods [19], which goes beyond the scope of the present study. The question that will face such an enterprise will be that of the proper optimization technique as well as the proper target function(s). This will be the object of upcoming work. But it should be underlined that the flatness of buffer index curves obtained here appears encouraging for optimization methods based on the approach proposed here, as it finds mathematically well defined optima. Another route of investigation would be to optimize the dynamics buffer capacity [33], however in that case an additional parameter would have to be taken into account as the dilution effects depend on the concentration of the added solution.

In a more practical light, the results and techniques presented in this study allow defining optimal buffer solutions which might or might not be feasible when considering the actual  $pK_a$  values of monovalent species. However, finding combinations of real weak acids that come as close as possible to such optima will lead to solutions with improved buffer-index stability over a wider range of pH and can serve as a guide for buffer preparation that do not aim at pH ranges wider than 3 pH units.

Moreover, this method could obviously be extended to situations including the contribution of polyprotic species, for instance in order to find specific optimal points in the presence of a given polyacid whose  $pK_a$ 's are known and set, while trying to flatten the dependency of  $\beta$  with pH by including one or more monovalent species.



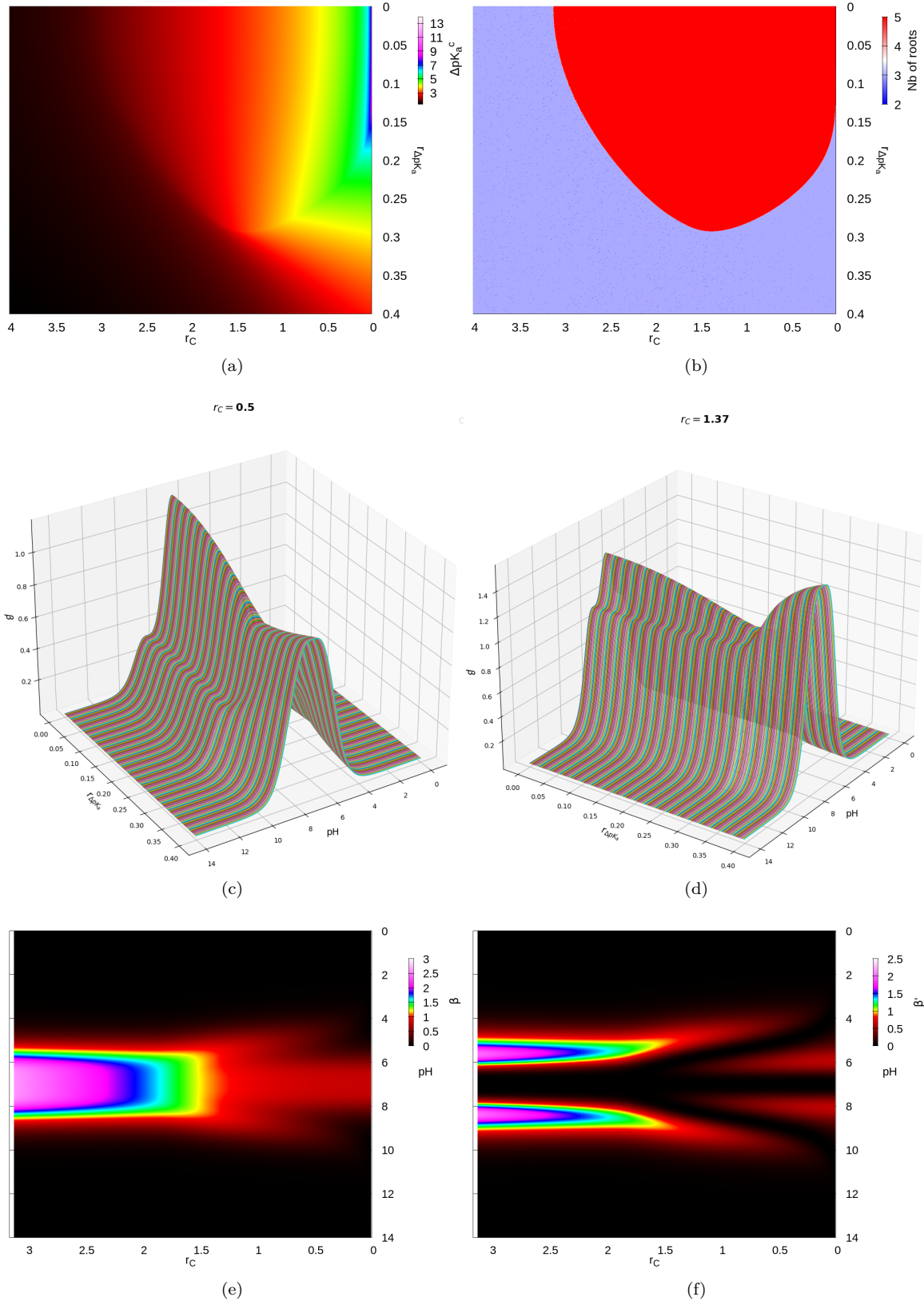


Figure 7: a) Values of  $\Delta pK_a^c$  as a function of  $r_C$  and  $r_{\Delta pK_a}$ . b) Number of roots of the derivative  $\beta'$  for  $\Delta pK_a$  values just above  $\Delta pK_a^c$ . c) Buffer indexes at  $\Delta pK_a = \Delta pK_a^c$  with varying  $r_{\Delta pK_a}$  for c)  $r_C = 0.5$  d)  $r_C = 1.37$ . e) Buffer index and f) its derivative along the 5-3 boundary.

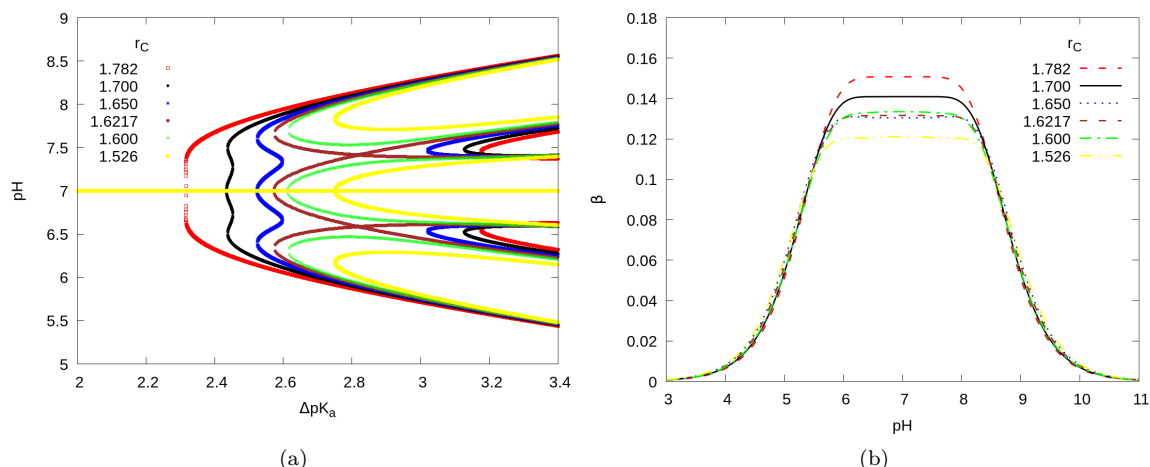


Figure 8: a) Roots of  $\beta'$  for parameters laying on the 5-3 border. b) Buffer indexes obtained for  $\Delta pK_a = \Delta pK_a^c$  as calculated in figure c.

## Conflict of interest

There are no conflict of interest to declare.

## References

- [1] E.P. Widmaier, H. Raff, and K.T. Strang. *Human physiology*. McGraw Hill Education, 2019.
- [2] C.M. Bethke. *Geochemical and Biogeochemical Reaction Modeling*. Cambridge University Press, 2008.
- [3] Hejie Qin, Yuankui Sun, Hongyi Yang, Peng Fan, Junlian Qiao, and Xiaohong Guan. Unexpected effect of buffer solution on removal of selenite and selenate by zerovalent iron. *Chemical Engineering Journal*, 334:296–304, 2018.
- [4] Sungjun Bae, Junyoung Jung, and Woojin Lee. The effect of ph and zwitterionic buffers on catalytic nitrate reduction by tio2-supported bimetallic catalyst. *Chemical Engineering Journal*, 232:327–337, 2013.
- [5] Cetin Kantar and Muhammet Samet Bulbul. Effect of ph-buffering on cr(vi) reduction with pyrite in the presence of various organic acids: Continuous-flow experiments. *Chemical Engineering Journal*, 287:173–180, 2016.
- [6] K.A. Levis, M.E. Lane, and O.I. Corrigan. Effect of buffer media composition on the solubility and effective permeability coefficient of ibuprofen. *International Journal of Pharmaceutics*, 253(1):49–59, 2003.
- [7] V.S. Stoll and J.S. Blanchard. Chapter 6 buffers: Principles and practice. In Richard R. Burgess and Murray P. Deutscher, editors, *Guide to Protein Purification, 2nd Edition*, volume 463 of *Methods in Enzymology*, pages 43–56. Academic Press, 2009.
- [8] Crawford Jordan. Ionic strength and buffer capacity of wide-range buffers for polarography. *Microchemical Journal*, 25(4):492–499, 1980.
- [9] M. V. Damle, R. Kaushal, M. Tiwari, and P. T. Malshe. Polarography of aldoximes in buffer solutions. *Bulletin des Sociétés Chimiques Belges*, 89(11):969–976, 1980.
- [10] Eva Horakova, Jiri Barek, and Vlastimil Vyskocil. Determination of methyl violet 2b using polarographic and voltammetric methods at mercury electrodes. *Analytical Letters*, 49(1):56–65, 2016.

- [11] P. Chattopadhyay and M. Mistry. Direct flame atomic absorption spectrophotometric determination of major oxides in rock samples of diverse composition using matrix buffer. *Microchemical Journal*, 50(1):78–87, 1994.
- [12] S. Moldoveanu and V. David. *Selection of the HPLC Method in Chemical Analysis*. Elsevier, 2016.
- [13] Petr Kozlik, Jana Vaclova, and Kveta Kalikova. Mixed-mode hydrophilic interaction/ion-exchange liquid chromatography – separation potential in peptide analysis. *Microchemical Journal*, 165:106158, 2021.
- [14] S. Blanco, M. J. Clifton, and J.-L. Joly. Sensitivity models in simulation of capillary zone electrophoresis: The effect of buffer power and initial buffer concentration. *Electrophoresis*, 19(10):1662–1673, 1998.
- [15] Tianfeng Xue, Zhongqi Xu, and Tao He. Study on a novel buffer system to separate and quantify glycated hemoglobin a1c (hba1c) and the mechanism between its key constituents and hba1c. *Microchemical Journal*, 163:105916, 2021.
- [16] P.G. Righetti. *Immobilized pH Gradients: Theory and Methodology*. Elsevier Science, 1990.
- [17] Görg, A. and Obermaier, C. and Boguth, G. and Harder, A. and Scheibe, B. and Wildgruber, R. and Weiss, W. The current state of two-dimensional electrophoresis with immobilized pH gradients. *Electrophoresis*, 21:1037–1053, 2000.
- [18] G. Weber, D. Grimm, and J. Bauer. Application of binary buffer systems to free flow cell electrophoresis. *Electrophoresis*, 21(2):325–328, 2000.
- [19] F. Celentano, E. Gianazza, G. Dossi, and P.G. Righetti. Buffer systems and pH gradient simulation. *Chemometrics and Intelligent Laboratory Systems*, 1(4):349–358, 1987.
- [20] C. Tonani and P.G. Righetti. Immobilized pH gradients (ipg) simulator—an additional step in pH gradient engineering: I. linear pH gradients. *Electrophoresis*, 12(12):1011–1021, 1991.
- [21] F. Kröner and J. Hubbuch. Systematic generation of buffer systems for pH gradient ion exchange chromatography and their application. *Journal of Chromatography A*, 1285:78–87, 2013.
- [22] J.E. Ricci. *Hydrogen ion concentration*. Princeton University Press, 1952.
- [23] Harry Rilbe. On the use of dimension-less parameters in acid-base theory. i. the buffer capacity of simple ampholyte solutions. *Electrophoresis*, 13(1):811–816, 1992.
- [24] Harry Rilbe. On the use of dimensionless parameters in acid-base theory. iii. the molar buffer capacity of trivalent protolytes. *Electrophoresis*, 14(1):591–596, 1993.
- [25] Harry Rilbe. On the use of dimensionless parameters in acid-base theory. v. buffers composed of binary mixtures of monovalent weak acids and bases. *Electrophoresis*, 14(1):1271–1277, 1993.
- [26] Harry Rilbe. On the use of dimensionless parameters in acid-base theory: Vi. the buffer capacities of equimolar binary mixtures of monovalent weak protolytes as compared to that of bivalent protolytes. *Electrophoresis*, 15(1):580–583, 1994.
- [27] Agustin G. Asuero. Buffer capacity of a polyprotic acid: First derivative of the buffer capacity and pk a values of single and overlapping equilibria. *Critical Reviews in Analytical Chemistry*, 37(4):269–301, 2007.
- [28] A.G. Asuero. Inflection points of the formation curve for 2-step overlapping equilibria. *Microchemical Journal*, 47(3):386–398, 1993.
- [29] J. N. Butler and D. R. Cogley. *Ionic equilibrium: Solubility and pH calculations*. NewYork,NY: John Wiley and Sons, 1998.
- [30] L. Van Vooren, M. Van De Steene, J.P. Ottoy, and P.A. Vanrolleghem. Automatic buffer capacity model building for the purposeof water quality monitoring. *Water Science and Technology*, 43(7):105–113, 2001.

- [31] E.T. Urbansky and M.R. Schock. Understanding, deriving, and computing buffer capacity. *J. Chem. Educ.* 2000, 77, 12, 1640, 77(12):1640, 2000.
- [32] D.D. Van Slyke. On the measurement of buffer values and on the relationship of buffer value to the dissociation constant of the buffer and the concentration and reaction of the buffer solution. *J. Biol. Chem.*, 52:525, 1922.
- [33] Michałowska-Kaczmarczyk, A.M., Michałowski, T. Dynamic buffer capacity in acid–base systems. *J Solution Chem*, 44:1256–1266, 2015.