

# Winding journey down (in Gibbs energy)

Jeffrey M. Dick

December 13, 2017

## 1 Introduction

`wjd()` appeared in CHNOSZ version 0.9-9. It implements the steepest descent algorithm for Gibbs energy minimization described by White, Johnson and Dantzig, 1958. This vignette shows some examples of its usage. The help page for `wjd()` in the CHNOSZ documentation has details of the implementation that are not covered in this vignette. The functions demonstrated here are not intended for aqueous solutions or heterogeneous phase equilibria, which are covered in some reviews (e.g. Nordstrom et al., 1979).

The arguments of `wjd()` have default values to simulate the example problem given by White et al., 1958. In their words, “The example taken is the determination of the composition of the gases arising from the combustion of a stoichiometric mixture of hydrazine and oxygen at 3500 °K and a pressure of 750 psi ( $\ln P_{\text{atm}} = 3.932$ ).” There are ten species included in this example. The value of temperature does not appear explicitly in the algorithm, or in the arguments to `wjd()`. Instead, the standard Gibbs energies of the species, at the given temperature, are provided in dimensionless form, i.e.  $\Delta G^\circ / RT$ . Note that White et al., 1958 use the terminology “free energy” and the notation  $F^\circ$ . The term “Gibbs energy” and corresponding notation are used here.

## 2 Is it winding?

The “winding” in the title refers to the observation that the abundances of species during a Gibbs energy minimization often do not all change in a monotonic fashion. This is because at each iteration a new direction of steepest descent is calculated; this direction is a vector of changes of mole numbers of all species in the system and is subject to mass balance constraints as well as the steepest-descent criterion.

Let’s run `wjd()` with its default settings and save the output.

```
> w <- wjd()
```

What are the most abundant species, and how many iterations were taken?

```
> # the order of species abundance
> oX <- order(w$X, decreasing=TRUE)
> # the stoichiometries of the two most abundant species
> w$A[head(oX,2),]
```

```
      H N O
[1,] 2 0 1
[2,] 0 2 0
```

```
> # the number of iterations
> niter <- length(w$lambda)
> niter
```

```
[1] 7
```

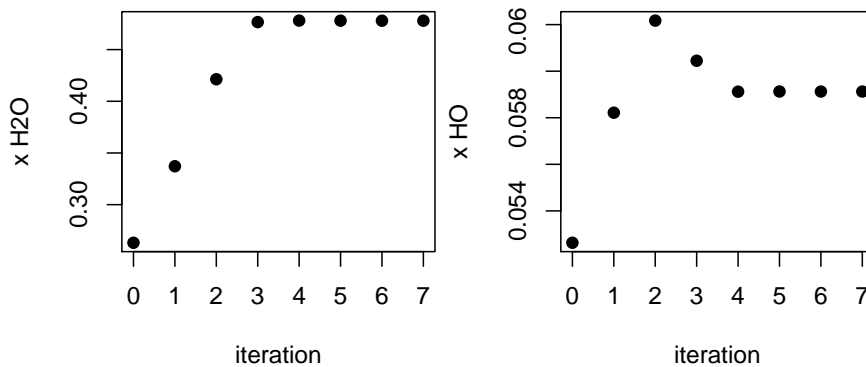
We find that H<sub>2</sub>O and N<sub>2</sub> are the most abundant species, after 7 iterations. The fourth most-abundant species, HO, is discussed below.

Let's track their mole fractions through the iterations. Write a function that returns the mole fractions of a species, after a specified number of iterations.

```
> iterfun <- function(imax, i) {
+   w <- wjd(imax=imax)
+   x <- w$X[i]/sum(w$X)
+   return(x)
+ }
```

Then apply the function over the different numbers of iterations, from 0 (initial conditions) to 7, and plot the values.

```
> par(mfrow = c(1, 2), mar = c(4.2, 4.2, .1, .1))
> sa <- sapply(0:niter, iterfun, i = oX[1])
> plot(0:niter, sa, xlab = "iteration", ylab = paste("x", f1), pch = 19)
> sa <- sapply(0:niter, iterfun, i = oX[4])
> plot(0:niter, sa, xlab = "iteration", ylab = paste("x", f4), pch = 19)
```



A bit of winding: the mole fraction of H<sub>2</sub>O generally increases (there is a very slight decrease at the fifth iteration), but the mole fraction of HO increases in the first two iterations, then decreases in following ones. This behavior is consistent with a decrease in Gibbs energy at *every* iteration; that can be verified by inspecting the values in the 7-iteration result:

```
> all(diff(w$F.Y) < 0)
```

```
[1] TRUE
```

The decrease in Gibbs energy becomes smaller with every iteration, as an equilibrium distribution of species is approached:

```
> identical(diff(w$F.Y), sort(diff(w$F.Y)))
```

```
[1] TRUE
```

### 3 Is it near the bottom? Testing for equilibrium

#### 3.1 Change in Gibbs energy of the system becomes small

Equilibrium is synonymous with a global Gibbs energy minimum. Becoming convinced that the output from wjd() represents a near-equilibrium condition can be difficult. One type of observation that can be helpful is the amount of change during the iterations of the algorithm. As equilibrium is approached, it makes sense that the fractional decreases in Gibbs energy would become smaller and smaller. This is used

as a stopping condition in the current implementation – if the fractional change, relative to the total energy of the system, reaches the value of `Gfrac` given in the arguments the iterations stop. Here are the fractional changes with each iteration:

```
> diff(w$F.Y)/w$F.Y[1:7]
```

```
[1] 1.9290e-02 3.5552e-03 5.8413e-04 9.7265e-06 2.4656e-06 1.1559e-07 2.2645e-10
```

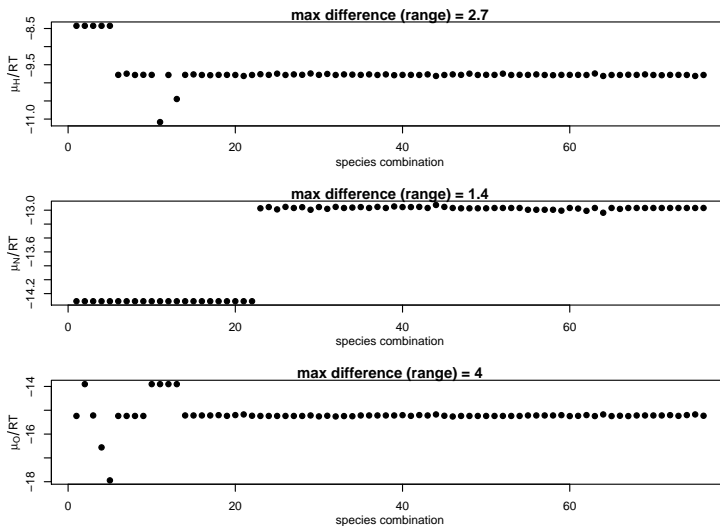
Only the last value is below the default value of `Gfrac` in `wjd()`.

### 3.2 Chemical potentials of the elements (from different species representing them) become more uniform

Equilibrium is also synonymous with uniformity of chemical potentials ( $\mu$ ) of the thermodynamic components throughout the system. For the default system in `wjd()`, and perhaps other systems of interest, the elements themselves are a conceivable set of the components. Often, the number and compositions of the species are such that multiple combinations of species satisfy the stoichiometric conditions necessary to be a basis set, and therefore can be used to compute the chemical potentials of the elements. At equilibrium, these different combinations of species would all yield the same chemical potentials of the elements.

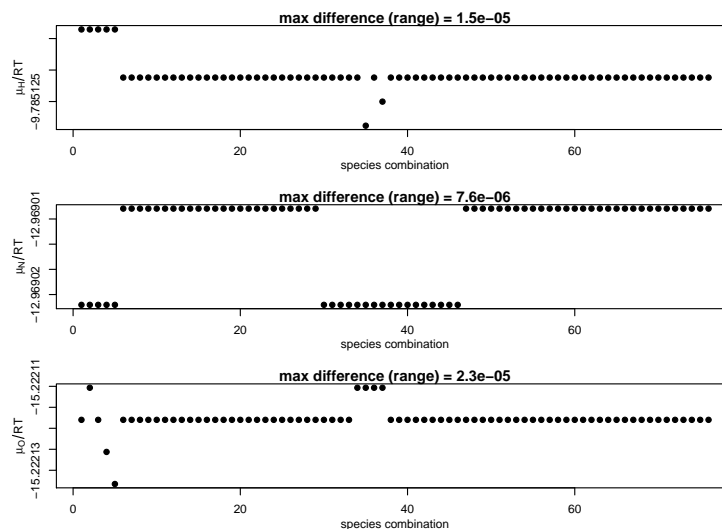
There is a supporting function, `element.potentials()`, that computes the chemical potentials of the elements using different eligible combinations of species. It has an option to plot the results. Let's first look at a plot showing the results after 3 iterations.

```
> par(mar = c(4.2, 4.2, 1.1, 1.1))
> par(mgp = c(2, 1, 0))
> w3 <- wjd(imax=3)
> ep3 <- element.potentials(w3, plot.it=TRUE)
```



Here's the plot for the default settings of `wjd()` which takes 6 iterations:

```
> par(mar = c(4.2, 4.2, 1.1, 1.1))
> par(mgp = c(2, 1, 0))
> ep <- element.potentials(w, plot.it=TRUE)
```



That one shows considerably less deviation than the first plot. The differences don't become zero, but perhaps they are small enough to accept as an operational solution. There is a function, `is.near.equil()`, that provides a logical type result based on the range of chemical potentials of the elements calculated from the different eligible species combinations. It returns TRUE if the maximum difference for any element is below  $\mu/RT = 0.01$ , but this limit can be changed through the `tol` argument to the function. With the default setting, `is.near.equil()` does pass the 7-iteration result, but by reducing the tolerance, the tests can be made to fail:

```
> is.near.equil(w3) # 3 iterations
[1] FALSE
> is.near.equil(w) # 7 iterations
[1] TRUE
> is.near.equil(w, tol=0.00001)
[1] FALSE
```

### 3.3 Different initial solutions yield similar results

**Different guesses for an underdetermined system** Although it seems to be the case in the examples shown so far, increasing the number of iterations does not necessarily bring one closer to true equilibrium, since being trapped in a local energy minimum is a possibility. Sometimes local minima can be shown to exist by starting with different mole numbers of species, that still yield the same bulk composition (mole numbers of elements). `guess()` exists to facilitate this sort of investigation. Its input the stoichiometric matrix of species in the system (the default refers to the same system as in `wjd()`) and the bulk composition; it outputs one set or a series of sets of mole numbers of species that satisfy the bulk composition constraints. In general, this is an underdetermined problem, so there are usually an infinite number of possible solutions.

Using the default method ("stoich") a single guess or multiple guesses are returned based on stoichiometric constraints (see the help page for details as well as the comments in the function code itself). If the "central" method is chosen instead, the single guess returned is the central solution from linear optimization using the package `limSolve`. Let's compare the guesses made by `guess()` with the default (from [White et al., 1958](#)) initial composition in `wjd()`.

```
> as.list(args(wjd))$Y
c(0.1, 0.35, 0.5, 0.1, 0.35, 0.1, 0.1, 0.1, 0.1, 0.1)
```

```

> Y1 <- guess(method="stoich")
> Y1

[1] 0.012 0.002 0.990 0.992 0.002 0.002 0.002 0.002 0.002 0.002

> #Y2 <- guess(method="central")
> #Y2

```

Now let's run the minimization using the different guesses, count the numbers of iterations, and test them for equilibrium.

```

> wY1 <- wjd(Y=Y1)
> niterY1 <- length(wY1$lambda)
> niterY1

[1] 9

> is.near.equil(wY1, tol=0.0001)

[1] FALSE

```

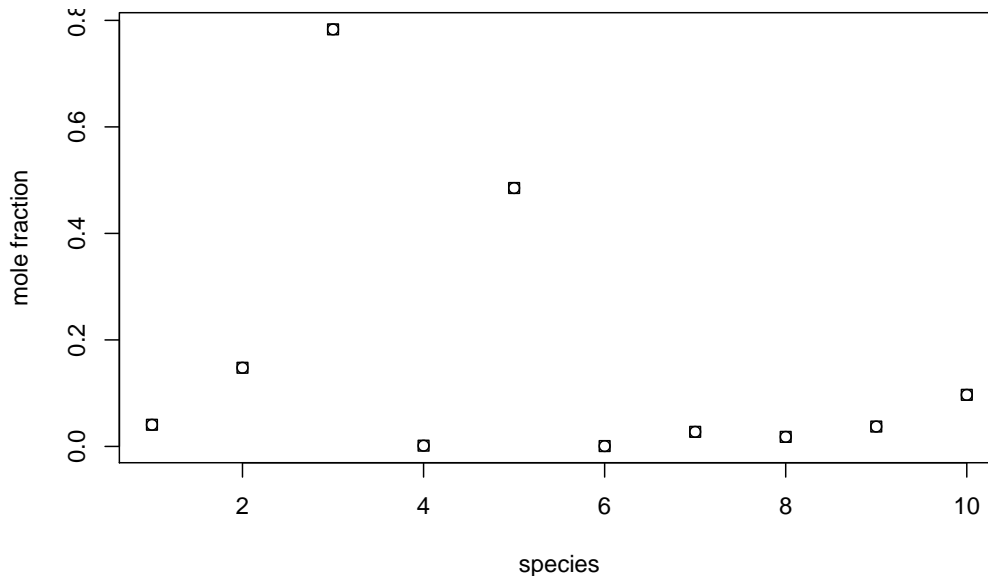
Using the first guess generated using the "stoich" method there were 2 more iterations than using the default initial solution in wjd().

Do the different initial guesses actually give similar results?

```

> par(mar=c(4.2, 4.2, 0.1, 0.1))
> plot(1:10, w$X, xlab="species", ylab="mole fraction")
> points(1:10, wY1$X, pch=0)
> #points(1:10, wY2$X, pch=2)

```



At that scale it's fairly difficult to tell them apart.

**Comparing many different guesses** Let's test for closeness to equilibrium of all of the guesses available from guess() for the default set of species and bulk composition. First generate all of the guesses. Note that guess() returns NA for compositionally eligible combinations of species that have some negative mole numbers, so we filter those out.

```

> Ys <- guess(iguess=NULL, method="stoich")
> # total number of species combinations
> length(Ys)

```

```
[1] 76
```

```
> # species combinations that have all-positive mole numbers
> iYs <- which(!is.na(Ys))
> nguess <- length(iYs)
> nguess
```

```
[1] 39
```

Now let's run the minimization using each of those guesses, and test if each one `is.near.equil()`:

```
> sapply(iYs,function(i) is.near.equil(wjd(Y=Ys[[i]])))

[1] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
[16] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
[31] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
```

Looks like they are. What if we're really picky though and want to make sure the chemical potentials of the elements are very well-defined? We can decrease the tolerance of `is.near.equil()`:

```
> sapply(iYs,function(i) is.near.equil(wjd(Y=Ys[[i]]),tol=0.0001))

[1] FALSE TRUE FALSE TRUE TRUE TRUE TRUE TRUE FALSE FALSE TRUE FALSE
[13] FALSE TRUE TRUE TRUE FALSE TRUE FALSE FALSE TRUE FALSE FALSE TRUE
[25] TRUE TRUE TRUE FALSE TRUE FALSE TRUE TRUE FALSE TRUE TRUE FALSE
[37] TRUE FALSE FALSE
```

Well some of the tests failed. But maybe this is because we have stopped iterating too soon. Let's iterate until a we have smaller change in Gibbs energy of the system:

```
> sapply(iYs, function(i) {
+   is.near.equil(wjd(Y=Ys[[i]], Gfrac=1e-9), tol=0.0001)
+ })

[1] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
[13] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
[25] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE FALSE
[37] TRUE TRUE TRUE
```

We're back to passing all but one of the equilibrium tests based on uniformity of chemical potentials of the elements. What are the standard deviations of the resulting species abundances?

```
> Xs <- sapply(iYs, function(i) wjd(Y=Ys[[i]], Gfrac=1e-9)$X)
> apply(Xs, 1, sd)

[1] 1.1573e-08 1.5358e-08 1.4431e-08 2.3977e-08 1.2291e-08 1.6666e-09
[7] 8.1648e-09 2.5680e-09 8.8796e-09 5.2164e-09
```

Based on this exercise, starting from different initial species abundances for the same bulk composition, and arriving at a similar near-equilibrium solution, it seems likely that we are near the global minimum.

(In fact, for ideal reactions in closed chemical systems there is only a unique solution (Othmer, 1976). Therefore, the algorithm should converge on the same result for any initial vector of species mole numbers with the same bulk composition.)

## 4 An example from the literature: prebiological atmospheres

We will try to reproduce a subset of the calculations of equilibrium between carbon-containing compounds in model prebiological atmospheres reported by [Dayhoff et al. \(1964\)](#), using standard Gibbs energies from [Dayhoff et al. \(1967\)](#). The standard Gibbs energies shown in Figure 2 of [Dayhoff et al. \(1967\)](#) for the selected compounds listed in Table 1 of [Dayhoff et al. \(1964\)](#) are provided with CHNOSZ. First let's have a look at what species there are:

```
> # read formulas and Gibbs energies
> file <- system.file("extdata/thermo/DLEN67.csv", package="CHNOSZ")
> dlen <- read.csv(file, as.is=TRUE, row.names=1)
> t(dlen[, 1, drop=FALSE])

      water carbon dioxide nitrogen methane ammonia carbon monoxide ethane
formula "H2O" "CO2"           "N2"      "CH4"  "NH3"  "CO"           "C2H6"
      formic acid propane hydrogen cyanide methylamine butane glycine
formula "CH2O2" "C3H8" "HCN"           "CH5N"  "C4H10" "C2H5NO2"
      benzene hexane naphthalene anthracene octane
formula "C6H6" "C6H14" "C10H8"      "C14H10" "C8H18"
```

Now let's write some code to define the system. The bulk composition in C:H:O space has 40 percent oxygen, a percentage of carbon given by  $x_C$ , and hydrogen as the remainder; the C:N ratio is taken to be 1:1 ([Dayhoff et al., 1964](#)). The temperature is 500 K and the pressure is 1 atmosphere.

```
> # turn formulas into a stoichiometric matrix
> A <- i2A(dlen$formula)
> # assemble Gibbs energies/RT at 500 K
> T <- 500 # K
> R <- 1.9872 # gas constant, cal K-1 mol-1
> GO.RT <- 1000 * dlen$G500 / R / T
> # a function to minimize Gibbs energy for system with
> # given mole fraction of carbon (xC)
> min.atmos <- function(xC) {
+ # the bulk composition C:H:N:O
+ B <- c(xC, 100-40-xC, xC, 40)
+ # guess the initial composition
+ Y <- guess(A, B)
+ w <- wjd(A=A, GO.RT=GO.RT, Y=Y, P=1, imax=90, Gfrac=1e-14)
+ if(!is.near.equil(w)) cat(paste("not near equilibrium for xC=", xC, "\n"))
+ return(w)
+ }
```

Notice the increase in  $imax$  and the decrease in  $Gfrac$  from the defaults! These changes are needed to get closer to equilibrium (and we'll know we're not there if a message is shown). What does the prebiological atmosphere look like with 15% (not counting nitrogen) carbon?

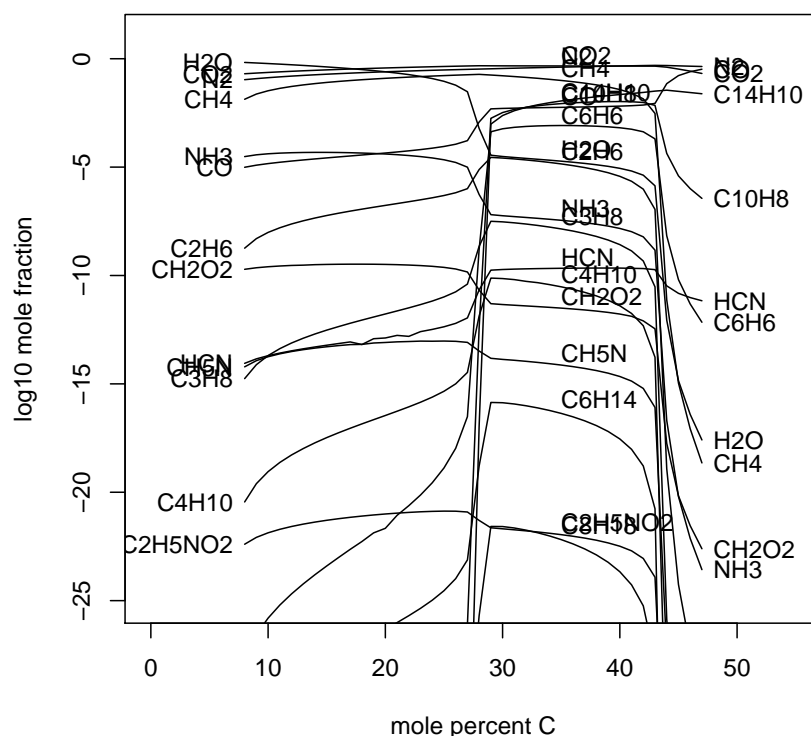
```
> sort(min.atmos(15)$X, decreasing=TRUE)

      H2O      CO2      N2      CH4      NH3      CO      C2H6
1.6249e+01 1.1875e+01 7.4991e+00 3.1240e+00 1.8522e-03 1.0304e-03 2.2665e-06
      CH2O2      C3H8      HCN      CH5N      C4H10      C2H5NO2      C6H6
1.2545e-08 1.1703e-11 2.3502e-12 1.9968e-12 1.2856e-16 1.9721e-20 1.5394e-22
      C6H14      C10H8      C8H18      C14H10
4.9248e-27 2.8495e-37 1.9641e-37 6.8381e-53
```

Well that's exciting! The order of abundances of H<sub>2</sub>O, CO<sub>2</sub>, CH<sub>4</sub>, NH<sub>3</sub>, CO, ethane (C<sub>2</sub>H<sub>6</sub>), formic acid (CH<sub>2</sub>O<sub>2</sub>) and so on matches the order on the left side of Figure 2 in Dayhoff et al. (1964). How about changing the carbon content? (Because it takes a while to run, the code is commented below; it is present in *demo(wjd)*, which was used to make the plot.)

```
> # xCs <- seq(8, 47, 1)
> # Xs <- sapply(xCs, function(xC) min.atmos(xC)$X)
> # # normalize the mole numbers to mole fractions
> # Xs <- t(t(Xs)/colSums(Xs))
> # plot(-10, 0, xlim=c(0, 55), ylim=c(-25, 1), xlab="mole percent C", ylab="log10 mole fraction")
> # for(i in 1:nrow(Xs)) lines(xCs, log10(Xs[i, ]))
> # text(48, log10(Xs[, length(xCs)]), dlen$formula, adj=0)
> # text(35, log10(Xs[, 27]) + 0.5, dlen$formula, adj=0)
> # text(7, log10(Xs[, 1]), dlen$formula, adj=1)
```

Prebiological atmospheres (Dayhoff et al., 1964)



We triggered some “not near equilibrium” messages (*not shown; run the demo to see them*), but overall it seems well behaved (note that using the “central” method in `guess()` is one way that leads to fewer of these messages). And, as expected, it is similar to Fig. 2 of Dayhoff et al. (1964), with a major crossing of curves at about 28% carbon, together with an increase in aromatic compounds (e.g. naphthalene, anthracene) going toward higher carbon content. Unlike the figure in Dayhoff et al. (1964), there appears to be a second major crossing of curves at about 43% carbon, corresponding to a rise in CO.

## 5 Running Down: Using a Thermodynamic Database

So far the examples shown here have been based on the chemical system defined by the default settings for the arguments to the functions. What if we're interested in a different system? It can be rather tedious



to manually construct the stoichiometric matrices and vectors of standard Gibbs energies of the species. The function `run.wjd()` pulls the compositional and thermodynamic data for requested species from the thermodynamic database and then calls `wjd()`.

## 5.1 At a single temperature

Let's start by finding the indices in the thermodynamic database of some liquid alkanes and aromatic compounds:

```
> alkanes <- c("n-hexane", "n-heptane", "n-octane", "n-nonane")
> ialk <- info(alkanes, "liq")
> aromatics <- c("naphthalene", "anthracene", "phenanthrene", "pyrene")
> iaro <- info(aromatics, "liq")
```

Let's find the equilibrium distribution of species for a bulk composition corresponding to a single mole of each species.

```
> waa <- run.wjd(c(ialk, iaro), Y=rep(1, 8))
> waa$elements

  [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10]
C   84   84   84   84   84   84   84   84   84   84
H  106  106  106  106  106  106  106  106  106  106

> is.near.equil(waa)

[1] FALSE
```

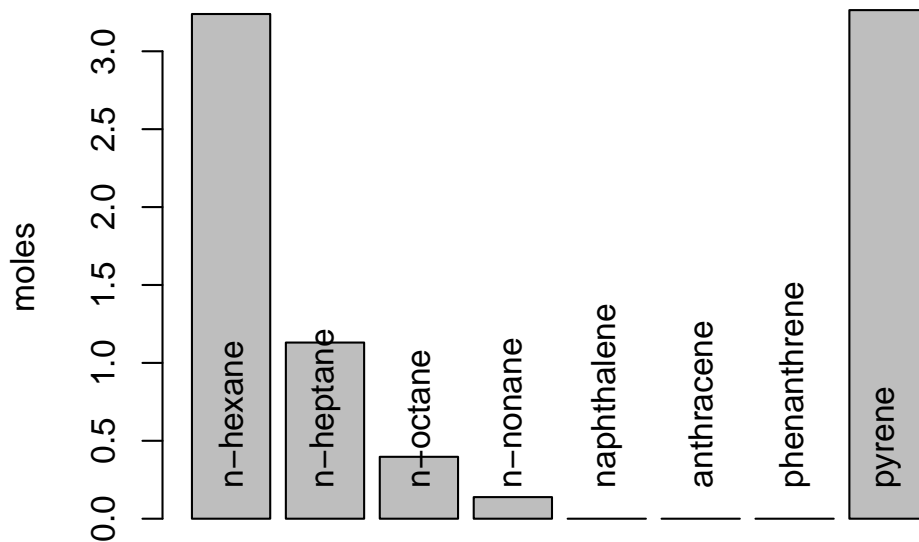
The `waa$elements` shows that the bulk chemical composition in the Gibbs energy minimization was  $C_{84}H_{106}$  ( $H/C = 1.261905$ ). That `FALSE` means the chemical potentials of the elements calculated from different combinations of species differ by more than 0.01, the default setting of `tol` in `is.near.equil()`. We can make it `TRUE` by running more iterations and increasing the number of intervals tested for fractional distance change at each step (`nlambda`):

```
> waa <- run.wjd(c(ialk, iaro), Y=rep(1, 8), imax=20, Gfrac=1e-14, nlambda=501)
> is.near.equil(waa)

[1] TRUE
```

Let's plot the abundances of the species:

```
> par(mar=c(1.1, 4.1, 1.1, 1.1))
> bp <- barplot(waa$X, ylab="moles")
> text(bp, rep(0.2,8), c(alkanes, aromatics), srt=90, adj=0)
```



Now let's calculate the chemical potentials of the elements (dimensionless values,  $\mu/RT$ ) ...

```
> ep <- equil.potentials(waa)
> print(ep)
```

```
      C      H
5863.4 -2610.6
```

... and the corresponding logarithms of chemical activities of a set of basis species:

```
> basis(c("graphite", "H2"), c("cr", "gas"))
```

```
      C H ispecies logact state
C    1 0    2037      0    cr
H2   0 2    3309      0    gas
```

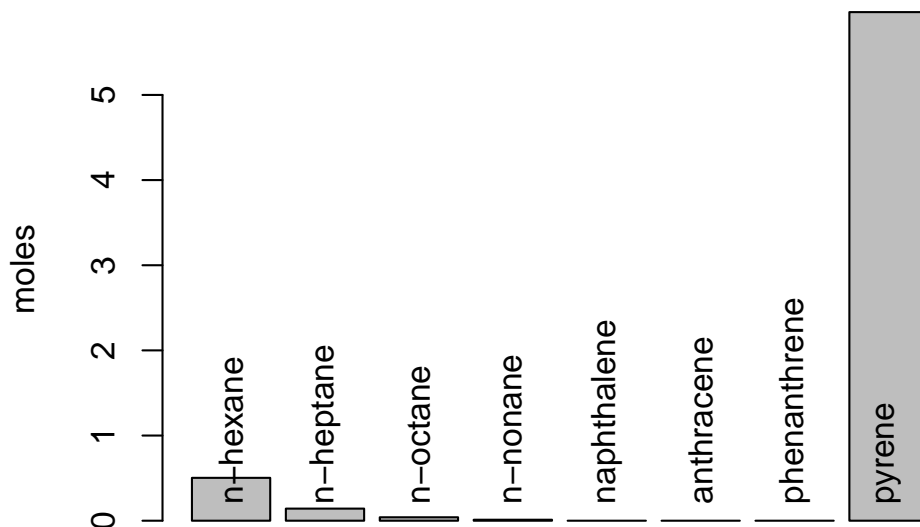
```
> basis.logact(ep)
```

```
      C      H2
4.2979 -3.8271
```

## 5.2 Bulk composition instead of moles of species

`run.wjd()` can be given `B` (bulk chemical composition), instead of `Y` (initial numbers of moles of species). When this happens, the function tests the different initial compositions generated by `guess()` until one is found that is within the specified tolerance of chemical potentials of elements. Here's an example:

```
> waa <- run.wjd(c(ialk, iaro), B="C100H70")
> par(mar=c(1.1, 4.1, 1.1, 1.1))
> bp <- barplot(waa$X, ylab="moles")
> text(bp, rep(0.2,8), c(alkanes, aromatics), srt=90, adj=0)
```



Compared with the previous example, here the lower  $H/C = 0.7$  defines a more oxidized system, so the increase in aromatic content is expected. As before, we can calculate the chemical potentials of the elements and of a corresponding set of basis species:

```
> print(ep <- equil.potentials(waa))
```

```
      C      H
5964.2 -2723.9
```

```
> basis.logact(ep)
```

```
      C      H2
4.3718 -3.9933
```

## 6 Document History

- 2012-01-01 Initial version
- 2012-06-16 Running Down using alkanes and aromatics
- 2012-09-20 Add example from Dayhoff et al.

## References

- M. O. Dayhoff, E. R. Lippincott, and R. V. Eck. Thermodynamic equilibria in prebiological atmospheres. *Science*, 146(3650):1461–1464, 1964. doi: [10.1126/science.146.3650.1461](https://doi.org/10.1126/science.146.3650.1461).
- M. O. Dayhoff, E. R. Lippincott, R. V. Eck, and G. Nagarajan. Thermodynamic equilibrium in prebiological atmospheres of C, H, O, N, P, S, and Cl. Report SP-3040, National Aeronautics and Space Administration, Washington, D. C., 1967.
- D. K. Nordstrom, L. N. Plummer, T. M. L. Wigley, T. J. Wolery, J. W. Ball, E. A. Jenne, R. L. Bassett, D. A. Crerar, T. M. Florence, B. Fritz, M. Hoffman, Jr. Holdren, G. R., G. M. Lafon, S. V. Mattigod, R. E. McDuff, F. Morel, M. M. Reddy, G. Sposito, and J. Thraikill. A comparison of computerized chemical models for equilibrium calculations in aqueous systems. In E. A. Jenne, editor, *Chemical Modeling in Aqueous Systems: Speciation, Sorption, Solubility, and Kinetics*, Series 93, pages 857–892. American Chemical Society, 1979. URL <http://www.worldcat.org/oclc/4591313>.

H. G. Othmer. Nonuniqueness of equilibria in closed reacting systems. *Chemical Engineering Science*, 31(11): 993–1003, 1976. doi: [10.1016/0009-2509\(76\)87020-0](https://doi.org/10.1016/0009-2509(76)87020-0).

W. B. White, S. M. Johnson, and G. B. Dantzig. Chemical equilibrium in complex mixtures. *Journal of Chemical Physics*, 28(5):751–755, 1958. ISSN 0021-9606. doi: [10.1063/1.1744264](https://doi.org/10.1063/1.1744264).