① Design of Experiments

time = money

experiments = money

Inputs ⇉ [System] ⇉ Outputs

Complex → DOE statistical method to get max data in min # of experiments.

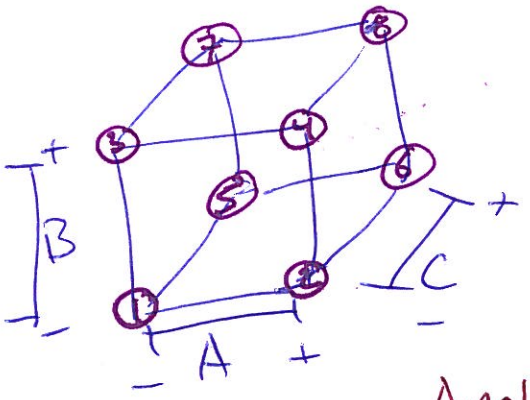
→ not taught in std. physics/chem curricula!

---

② Factorial 2^n experimental Design (two level)

1. consider inputs A, B, C
2. inputs can be hi(+) or lo(−)
3. can be mapped onto 3D space

Simplest design → more complex designs via jump needed for more variables



→ can "map" response space w/ 8 experiments.

Analyze 3D space.

An Example: Pilot Plant

① Identify Factors

Three factors

	(−)	(+)
T = Temperature (T):	160	180
C = Cat Concentration (%):	20	40
K = Catalyst Type:	A	B

② Create Run Table of coded units

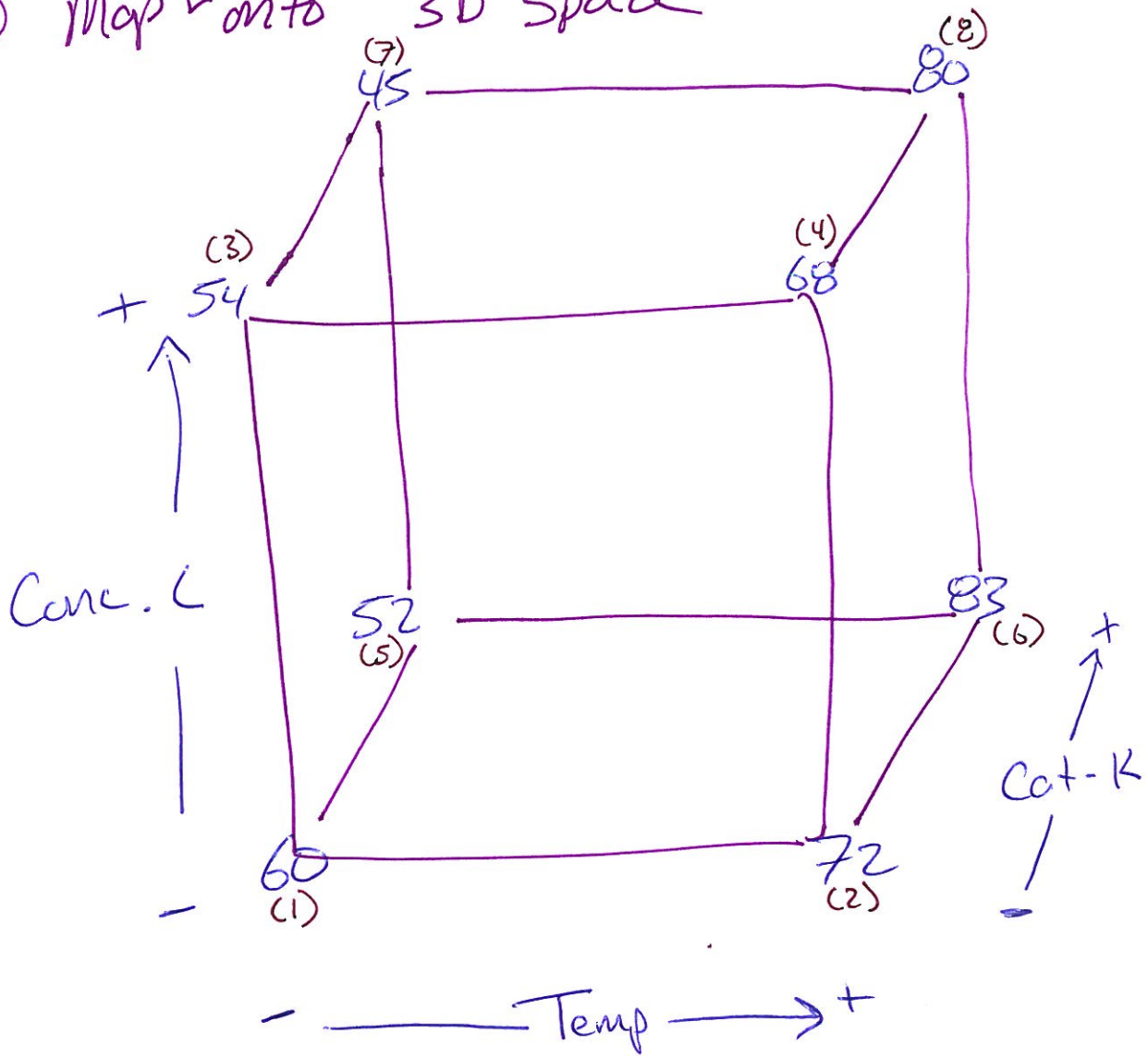
Run	T	C	K	Yield (Y)
1	−	−	−	60
2	+	−	−	72
3	−	+	−	54
4	+	+	−	68
5	−	−	+	52
6	+	−	+	83
7	−	+	+	45
8	+	+	+	80

(Runs arranged in standard order)

Note that these vectors are completely orthogonal.

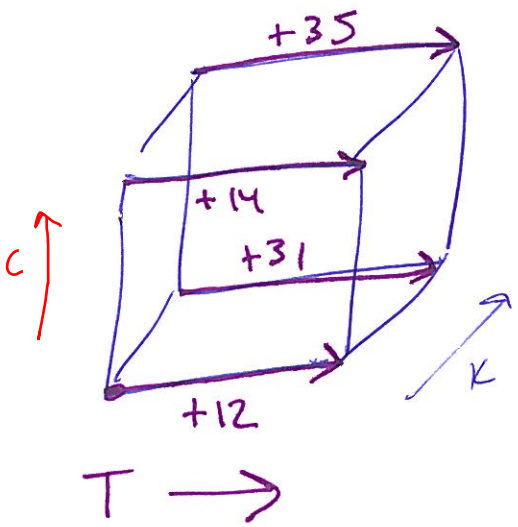
(represent degrees of freedom)

③ Map Response onto 3D space



④ Calculate Main Effects

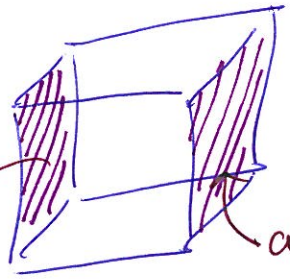
compare single values where one variable goes from ~~lo~~ Lo to hi



Average effect of T $= \frac{12+14+31+35}{4}$

$= \underline{23}$

another way to visualize



average response on this face

$T \rightarrow$

average response on this face

$$T = \frac{\sum Y_+}{n_+} - \frac{\sum Y_-}{n_-}$$

where Y is a response (i.e. data.

- all three main effects can be calculated this way.
- notice that all 8 observations are being used to estimate each of the main effects
- precision of a four-fold replicated difference; same precision in one-at-a-time experiment would require 8 runs for each factor!

Advantage #1

(5) Calculate interaction effects.

Notice T has a larger average effect $\left(\frac{+35 + 31}{2} = +33\right)$ when K is hi (+) than when K is lo (-) $\left(\frac{14+12}{2} = 13\right)$

→ The T and K parameters are coupled and interact

→ 1/2 of difference between two (average effects) sides of T- → + is interaction parameter

$$TK = \frac{33 + 13}{2} = \frac{20}{2} = \boxed{10}$$

work out effects that contribut

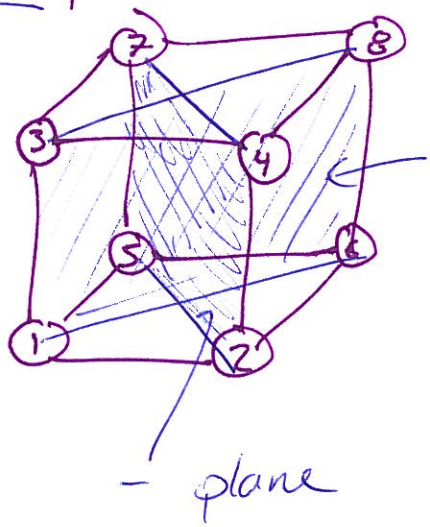
$$TK = \left[\frac{(Y_8 - Y_7) + (Y_6 - Y_5)}{2}\right] - \left[\frac{(Y_4 - Y_3) + (Y_2 - Y_1)}{2}\right]$$

re-arrange

$$TK = \frac{(Y_1 + Y_3 + Y_6 + Y_8)}{4} - \frac{(Y_2 + Y_4 + Y_5 + Y_7)}{4}$$

again notice TK is calculated by using all eight measurements as a sum of +'s and −'s.

Visually

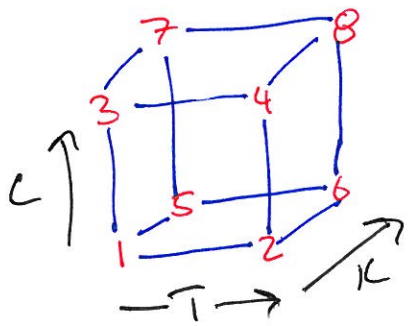


Two factor interactions all averages between two diagonal planes,

→ same calculations can be performed for TC and CK interactions.

Three - Factor interactions : TCK

Consider TC — two measures of TC are available, one at high K and one at low K



hi-K

$$TC_1 = \frac{(Y_8 - Y_7) - (Y_6 - Y_5)}{2} = 2$$

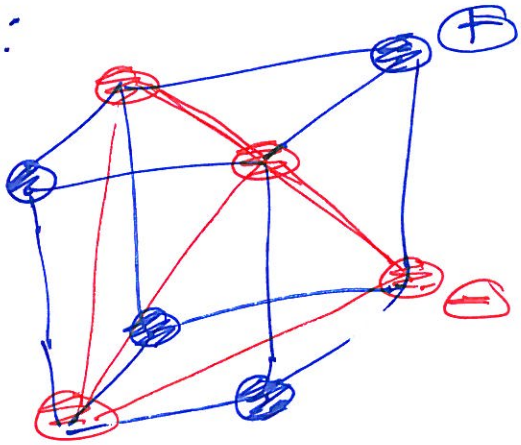
lo-K

$$TC_2 = \frac{(Y_4 - Y_3) - (Y_2 - Y_1)}{2} = 1$$

$$TCK = \frac{2-1}{2} = 0.5$$

Can rearrange terms again to be the sum of all (+)'s and all (−)'s.

graphically → comparing responses from the corners of the cube:



Tedious → design matrix method: simple linear algebra

Mean	T	C	K	TC	TK	CK	TCK	response
+	−	−	−	+	+	+	−	Y_1
+	+	−	−	−	−	+	+	Y_2
+	−	+	−	−	+	−	+	Y_3
+	+	+	−	+	−	−	−	Y_4
+	−	−	+	+	−	−	+	Y_5
+	+	−	+	−	+	−	−	Y_6
+	−	+	+	−	−	+	−	Y_7
+	+	+	+	+	+	+	+	Y_8

product of signs (i.e. dot product of the vectors) yields the correct signs for the interaction vector! Simply add interactions with the signs indicated... show excell example.

How to estimate error of measurement?

a. run replicates (good idea to look for instrument drift)

or

b. assume 3rd-order interactions and higher are negligible (usually very good assumption)

(show excell example)

~~Standard~~ Mean Error of an Effect

$$\left(\text{error effect}\right)^2 = \frac{1}{n} \sum_{n'=1}^{n} \left(XYZ_{n'}\right)^2$$

i.e. sum of squares of all negligible terms! divided by number of negligible terms

DOE Calculation Example.xls

Trial	A	B	C	AB	AC	CB	ABC	Effect
1	-1	-1	-1	1	1	1	-1	60
2	1	-1	-1	-1	-1	1	1	72
3	-1	1	-1	-1	1	-1	1	54
4	1	1	-1	1	-1	-1	-1	68
5	-1	-1	1	1	-1	-1	1	52
6	1	-1	1	-1	1	-1	-1	83
7	-1	1	1	-1	-1	1	-1	45
8	1	1	1	1	1	1	1	80

	A	B	C	AB	AC	CB	ABC
	-60	-60	-60	60	60	60	-60
	72	-72	-72	-72	-72	72	72
	-54	54	-54	-54	54	-54	54
	68	68	-68	68	-68	-68	-68
	-52	-52	52	52	-52	-52	52
	83	-83	83	-83	83	-83	-83
	-45	45	45	-45	-45	45	-45
	80	80	80	80	80	80	80
Sum	92	-20	6	6	40	0	2
divide by n+(or n-)	23	-5	1.5	1.5	10	0	0.5

blue = residual = no effect

red/orange = effect

→ a "normal probability plot" is statistical way to determine if observed "effect" is explainable by random chance. Here the effects are obvious.

other

→ → statistics can be used to determine to what confidence interval you believe your results

→ to build a linear model.

$$\text{Effect} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_{12} X_1 X_2$$

here for a two-level two-effect system (more complex above)