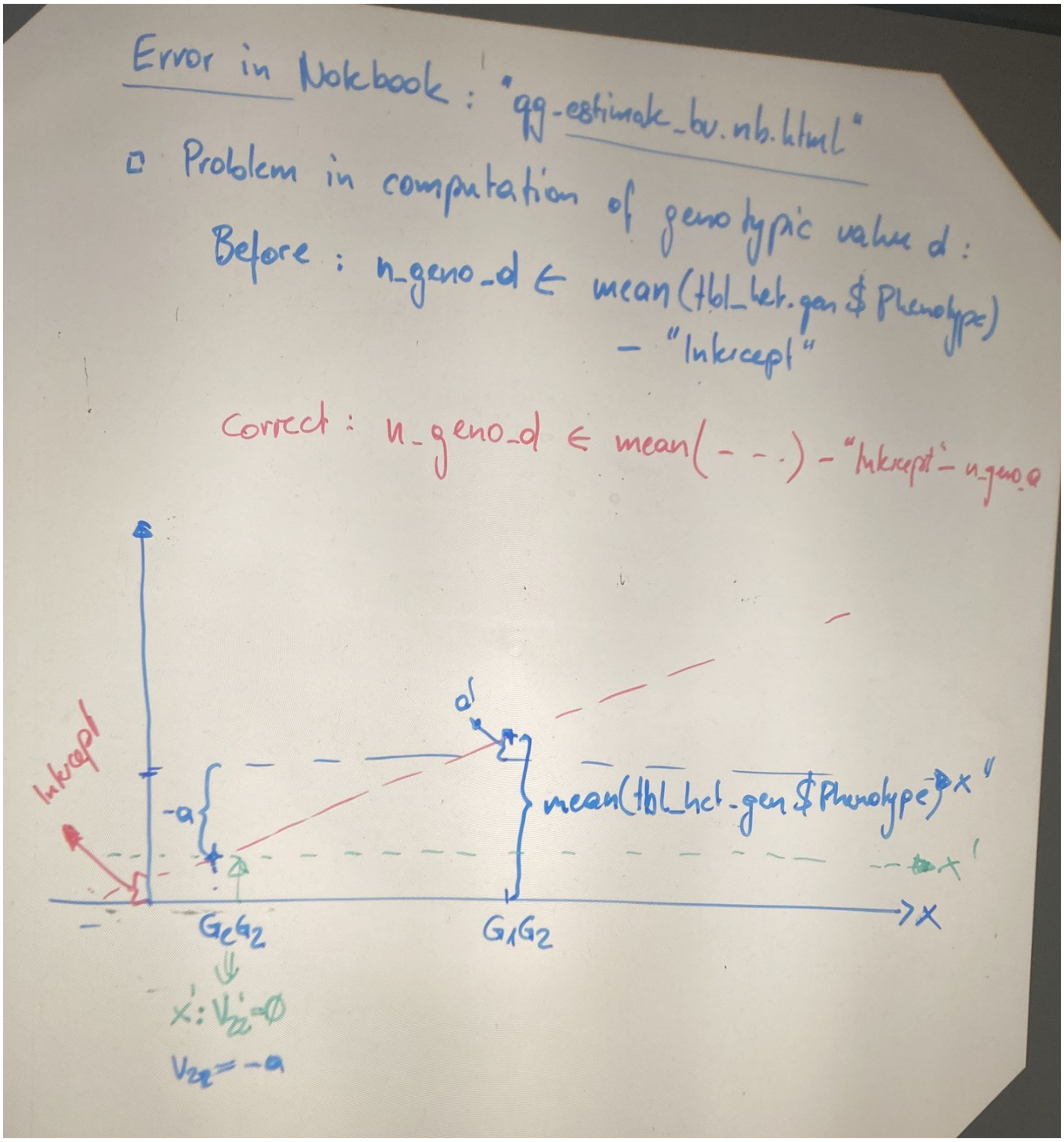
OHP Picture 1



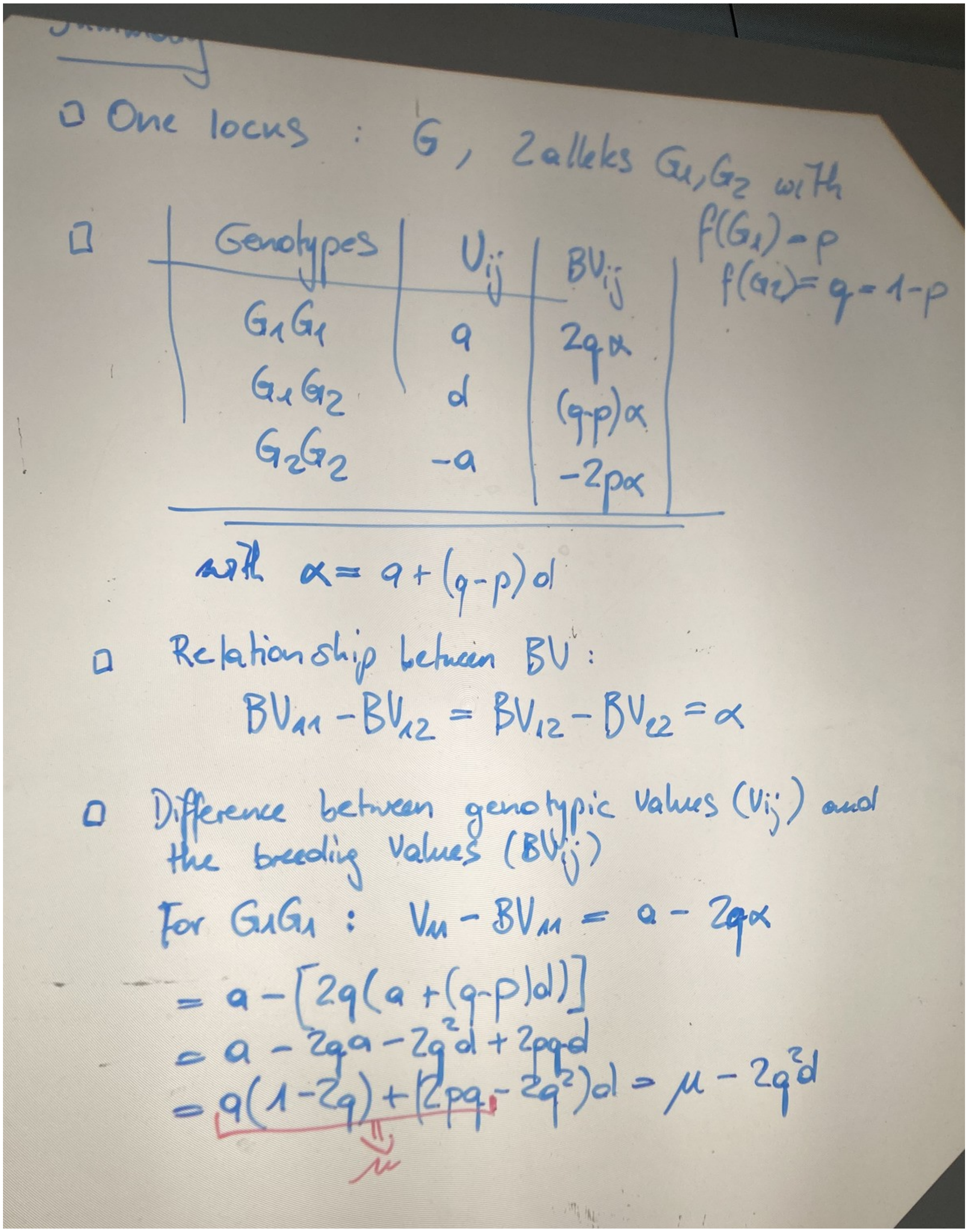
Error in Notebook: "qg-estimate_bv.nb.html"

☐ Problem in computation of genotypic value d:

Before: n_geno_d ← mean(tbl_het.gen$Phenotype) - "Intercept"

Correct: n_geno_d ← mean(- - -) - "Intercept" - n_geno_a

OHP Picture 2



☐ One locus : G, 2 alleles G_1, G_2 with

$f(G_1) = p$

$f(G_2) = q = 1 - p$

☐

Genotypes	V_{ij}	BV_{ij}
G_1G_1	a	$2q\alpha$
G_1G_2	d	$(q-p)\alpha$
G_2G_2	$-a$	$-2p\alpha$

with $\alpha = a + (q-p)d$

☐ Relationship between BV :

$$BV_{11} - BV_{12} = BV_{12} - BV_{22} = \alpha$$

☐ Difference between genotypic values (V_{ij}) and the breeding values (BV_{ij})

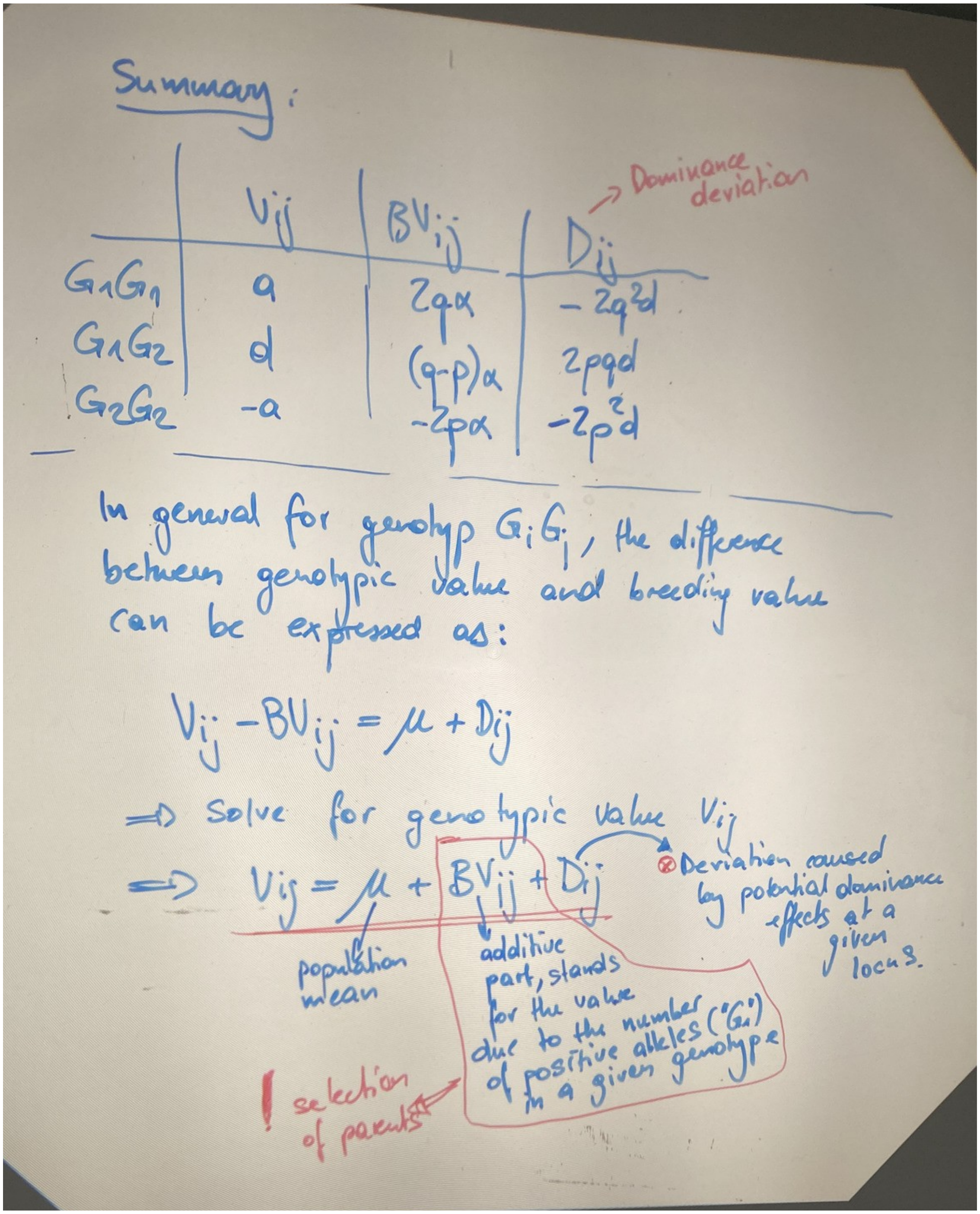
For G_1G_1 : $V_{11} - BV_{11} = a - 2q\alpha$

$$= a - [2q(a + (q-p)d)]$$

$$= a - 2qa - 2q^2d + 2pqd$$

$$= \underbrace{a(1-2q) + (2pq}_{\mu} - 2q^2)d = \mu - 2q^2d$$

OHP Picture 3



Summary:

Dominance deviation

	V_{ij}	BV_{ij}	D_{ij}
G_1G_1	a	$2q\alpha$	$-2q^2d$
G_1G_2	d	$(q-p)\alpha$	$2pqd$
G_2G_2	$-a$	$-2p\alpha$	$-2p^2d$

In general for genotyp G_iG_j, the difference between genotypic value and breeding value can be expressed as:

$$V_{ij} - BV_{ij} = \mu + D_{ij}$$

⇒ Solve for genotypic value V_{ij}

$$\Rightarrow \quad V_{ij} = \mu + \boxed{BV_{ij}} + D_{ij}$$

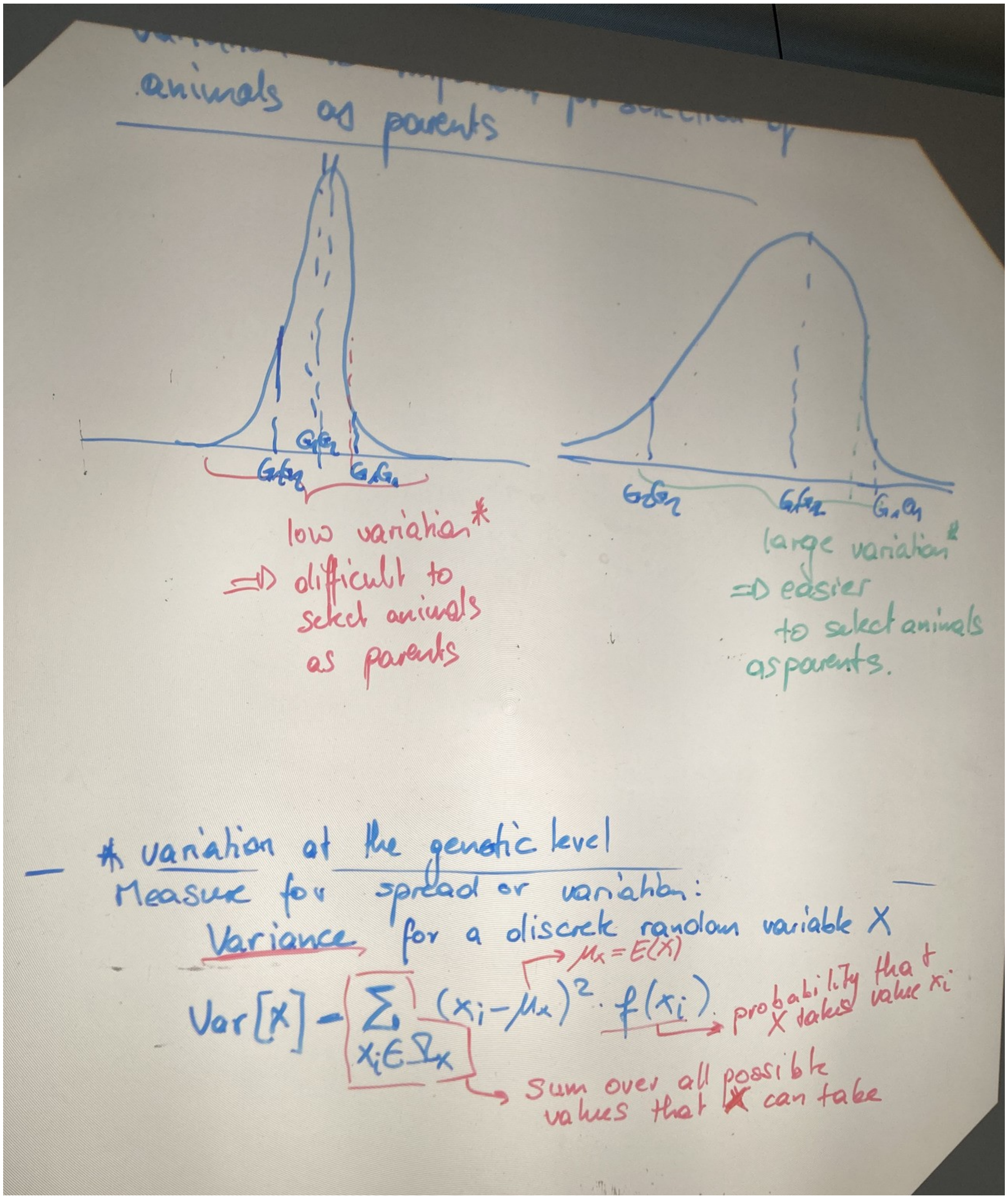
population mean

additive part, stands for the value due to the number of positive alleles ("G_i") in a given genotype

⊗ Deviation caused by potential dominance effects at a given locus.

! selection of parents

OHP Picture 4



animals as parents

low variation*

⇒ difficult to select animals as parents

large variation*

⇒ easier to select animals as parents.

* variation at the genetic level

Measure for spread or variation:

Variance for a discrete random variable X

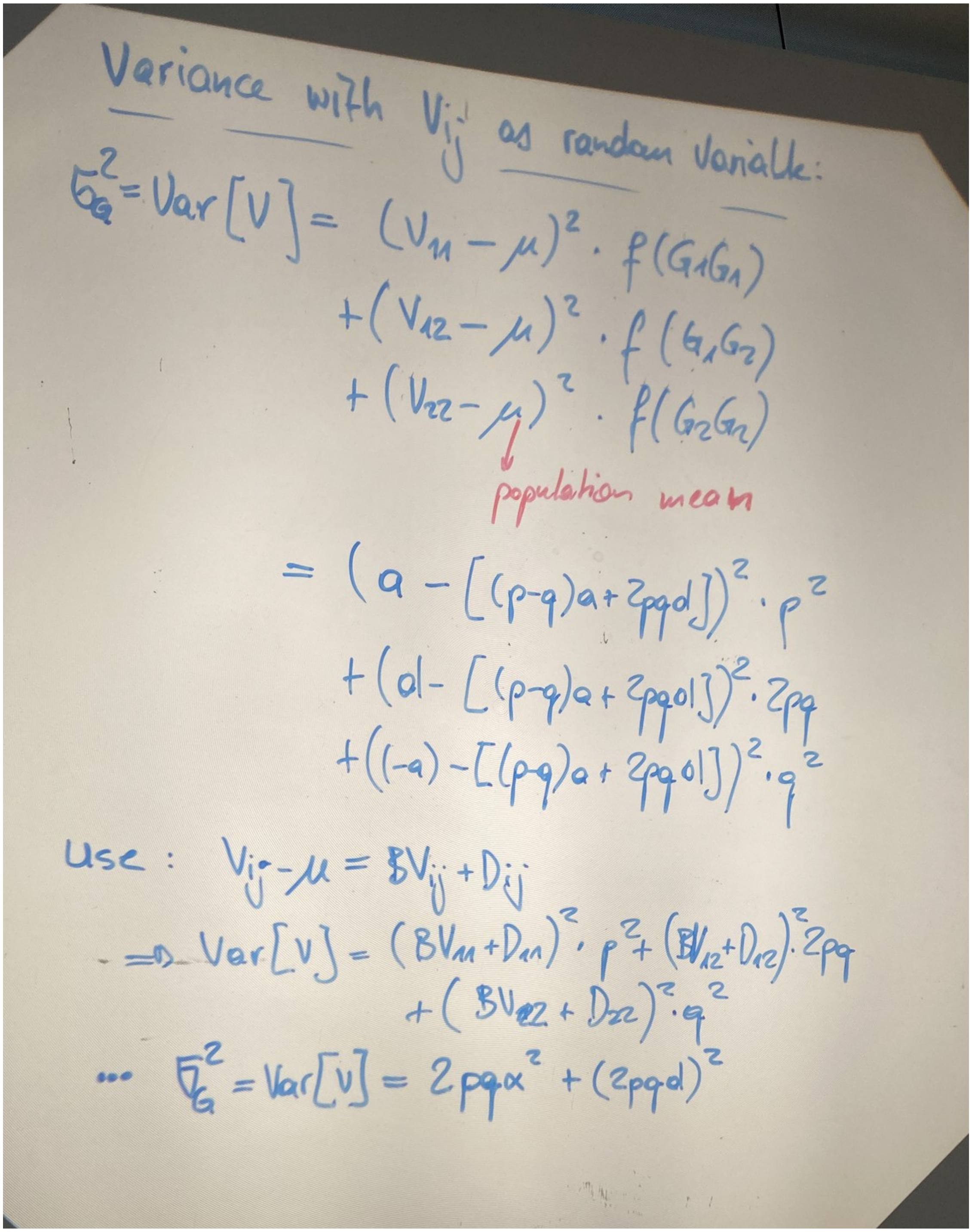
$$\mathrm{Var}[X] = \boxed{\sum_{x_i \in \Omega_X}} (x_i - \mu_X)^2 \cdot f(x_i)$$

$\mu_X = E(X)$

$f(x_i)$: probability that X takes value x_i

sum over all possible values that X can take

OHP Picture 5



Variance with V_{ij} as random variable:

$$\sigma_G^2 = Var[V] = (V_{11}-\mu)^2 \cdot f(G_1G_1) + (V_{12}-\mu)^2 \cdot f(G_1G_2) + (V_{22}-\mu)^2 \cdot f(G_2G_2)$$

population mean

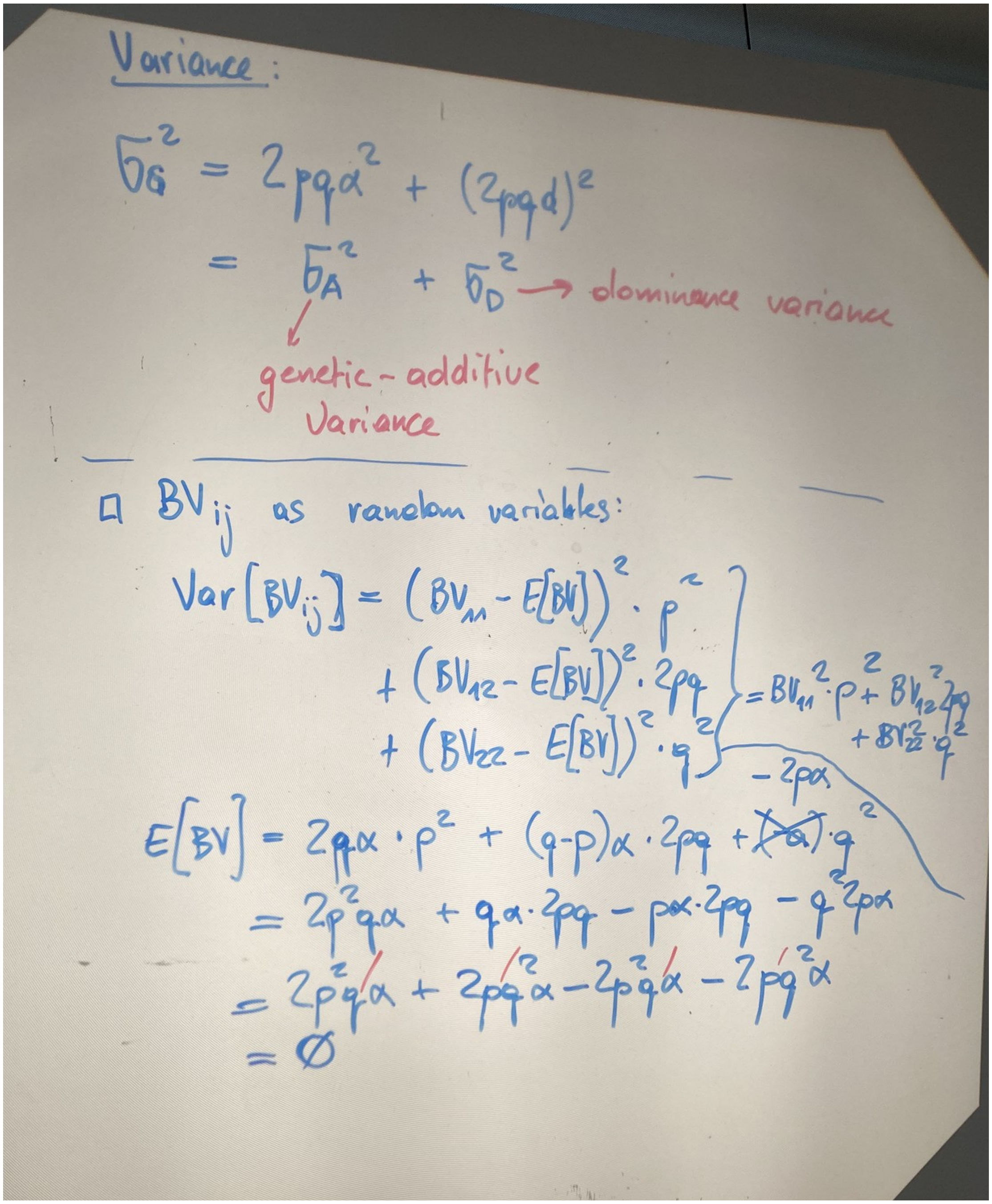
$$= (a - [(p-q)a+2pqd])^2 \cdot p^2 + (d - [(p-q)a+2pqd])^2 \cdot 2pq + ((-a) - [(p-q)a+2pqd])^2 \cdot q^2$$

Use: $V_{ij}-\mu = BV_{ij} + D_{ij}$

$$\Rightarrow Var[V] = (BV_{11}+D_{11})^2 \cdot p^2 + (BV_{12}+D_{12})^2 \cdot 2pq + (BV_{22}+D_{22})^2 \cdot q^2$$

$$\ldots \quad \sigma_G^2 = Var[V] = 2pq\alpha^2 + (2pqd)^2$$

OHP Picture 6



Variance:

$$\sigma_G^2 = 2pq\alpha^2 + (2pqd)^2$$

$$= \underbrace{\sigma_A^2}_{\text{genetic-additive Variance}} + \sigma_D^2 \rightarrow \text{dominance variance}$$

☐ BV_{ij} as random variables:

$$\left.\begin{aligned} Var[BV_{ij}] = {} & (BV_{11} - E[BV])^2 \cdot p^2 \\ & + (BV_{12} - E[BV])^2 \cdot 2pq \\ & + (BV_{22} - E[BV])^2 \cdot q^2 \end{aligned}\right\} = BV_{11}^2 \cdot p^2 + BV_{12}^2\, 2pq + BV_{22}^2 \cdot q^2$$

$$-2p\alpha$$

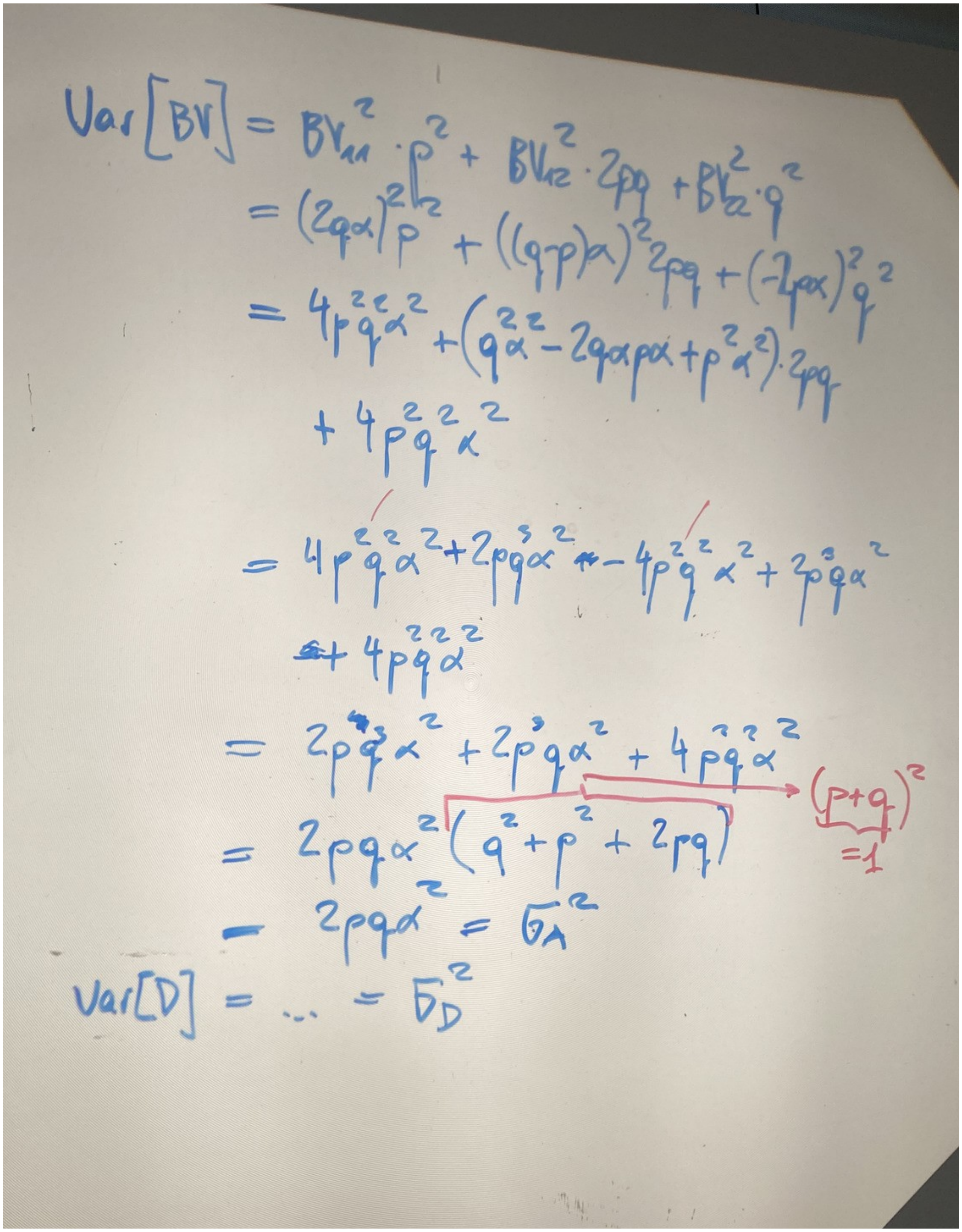
$$E[BV] = 2q\alpha \cdot p^2 + (q-p)\alpha \cdot 2pq + \cancel{(-\alpha)} \cdot q^2$$

$$= 2p^2q\alpha + q\alpha \cdot 2pq - p\alpha \cdot 2pq - q^2 2p\alpha$$

$$= 2p^2q\alpha + 2pq^2\alpha - 2p^2q\alpha - 2pq^2\alpha$$

$$= 0$$

OHP Picture 7



$$\mathrm{Var}[BV] = BV_{11}^2 \cdot p^2 + BV_{12}^2 \cdot 2pq + BV_{22}^2 \cdot q^2$$

$$= (2q\alpha)^2 p^2 + ((q-p)\alpha)^2 2pq + (-2p\alpha)^2 q^2$$

$$= 4p^2q^2\alpha^2 + (q^2\alpha^2 - 2q\alpha p\alpha + p^2\alpha^2) \cdot 2pq$$

$$+ 4p^2q^2\alpha^2$$

$$= \cancel{4p^2q^2\alpha^2} + 2pq^3\alpha^2 - \cancel{4p^2q^2\alpha^2} + 2p^3q\alpha^2$$

$$+ 4p^2q^2\alpha^2$$

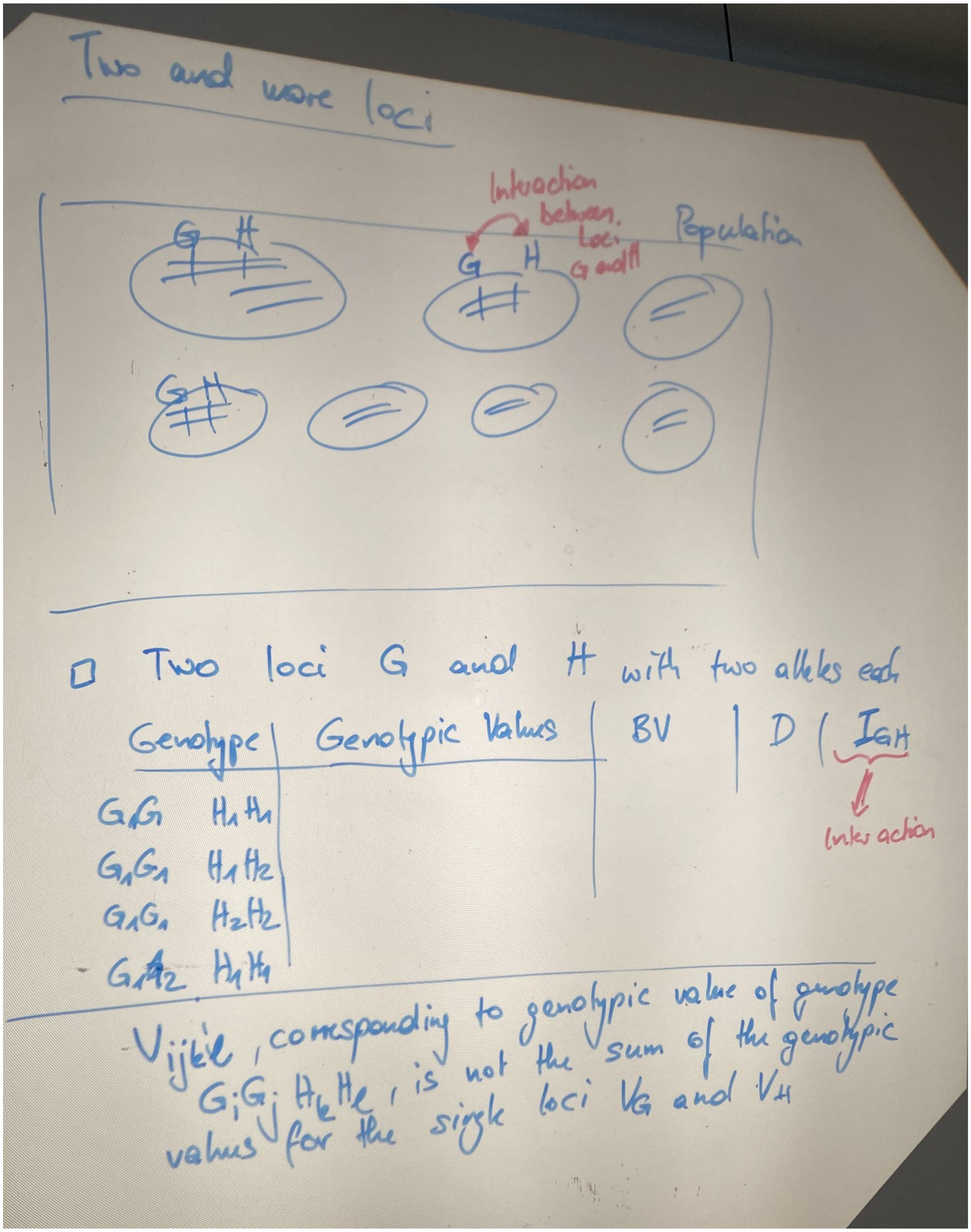
$$= 2pq^3\alpha^2 + 2p^3q\alpha^2 + 4p^2q^2\alpha^2$$

$$= 2pq\alpha^2 (q^2 + p^2 + 2pq) \rightarrow (\underbrace{p+q}_{=1})^2$$

$$= 2pq\alpha^2 = \sigma_A^2$$

$$\mathrm{Var}[D] = \ldots = \sigma_D^2$$

OHP Picture 8



Two and more loci

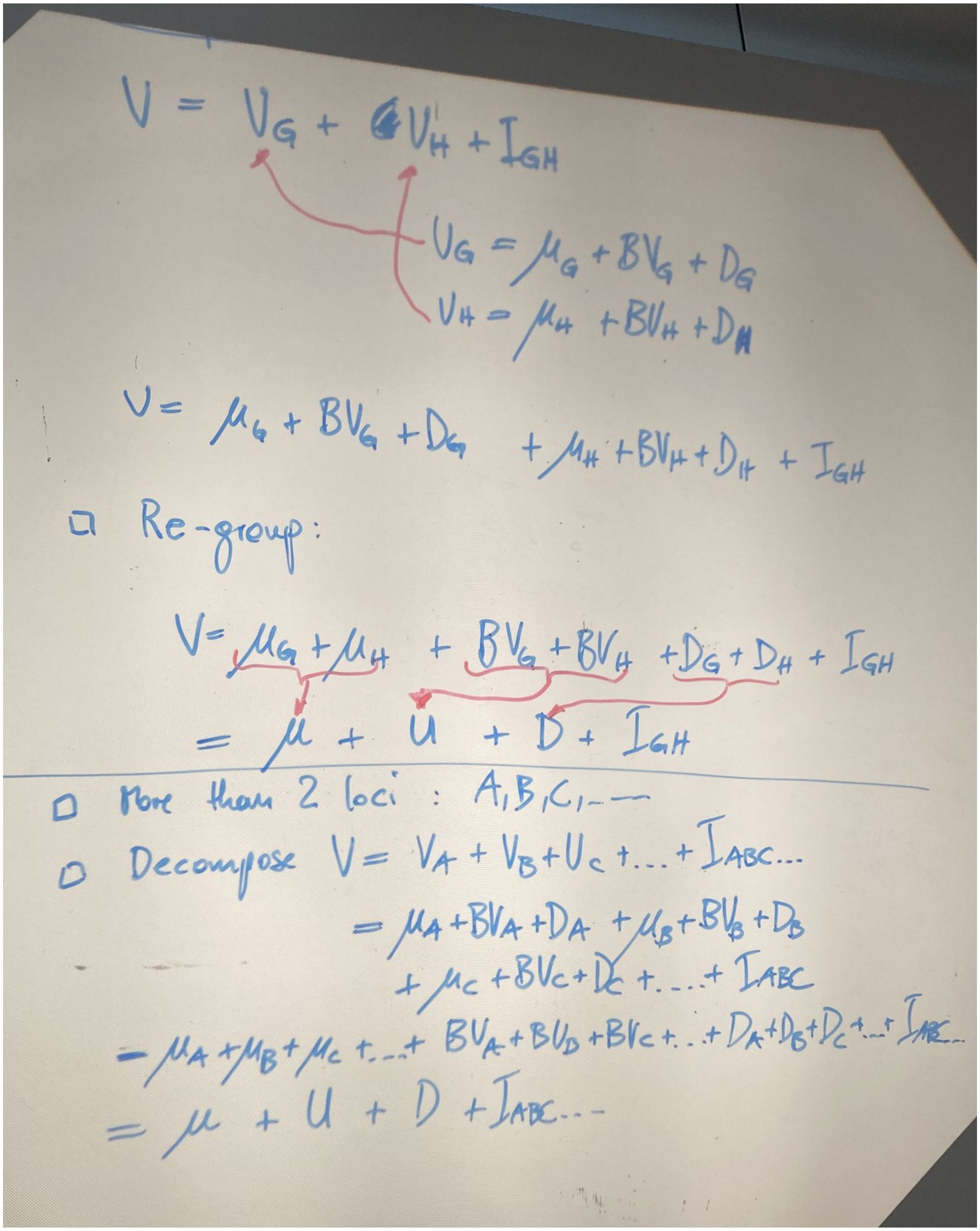
☐ Two loci G and H with two alleles each

Genotype	Genotypic Values	BV	D	I_{GH}
G_1G_1 H_1H_1				
G_1G_1 H_1H_2				
G_1G_1 H_2H_2				
G_1G_2 H_1H_1				

Interaction

V_{ijkl}, corresponding to genotypic value of genotype $G_iG_j H_kH_l$, is not the sum of the genotypic values for the single loci V_G and V_H

OHP Picture 9



$$V = V_G + V_H + I_{GH}$$

$$V_G = \mu_G + BV_G + D_G$$

$$V_H = \mu_H + BV_H + D_H$$

$$V = \mu_G + BV_G + D_G + \mu_H + BV_H + D_H + I_{GH}$$

☐ Re-group:

$$V = \underbrace{\mu_G + \mu_H}_{} + \underbrace{BV_G + BV_H}_{} + \underbrace{D_G + D_H}_{} + I_{GH}$$

$$= \mu + U + D + I_{GH}$$

☐ More than 2 loci : A, B, C, ...

☐ Decompose $V = V_A + V_B + V_C + \ldots + I_{ABC\ldots}$

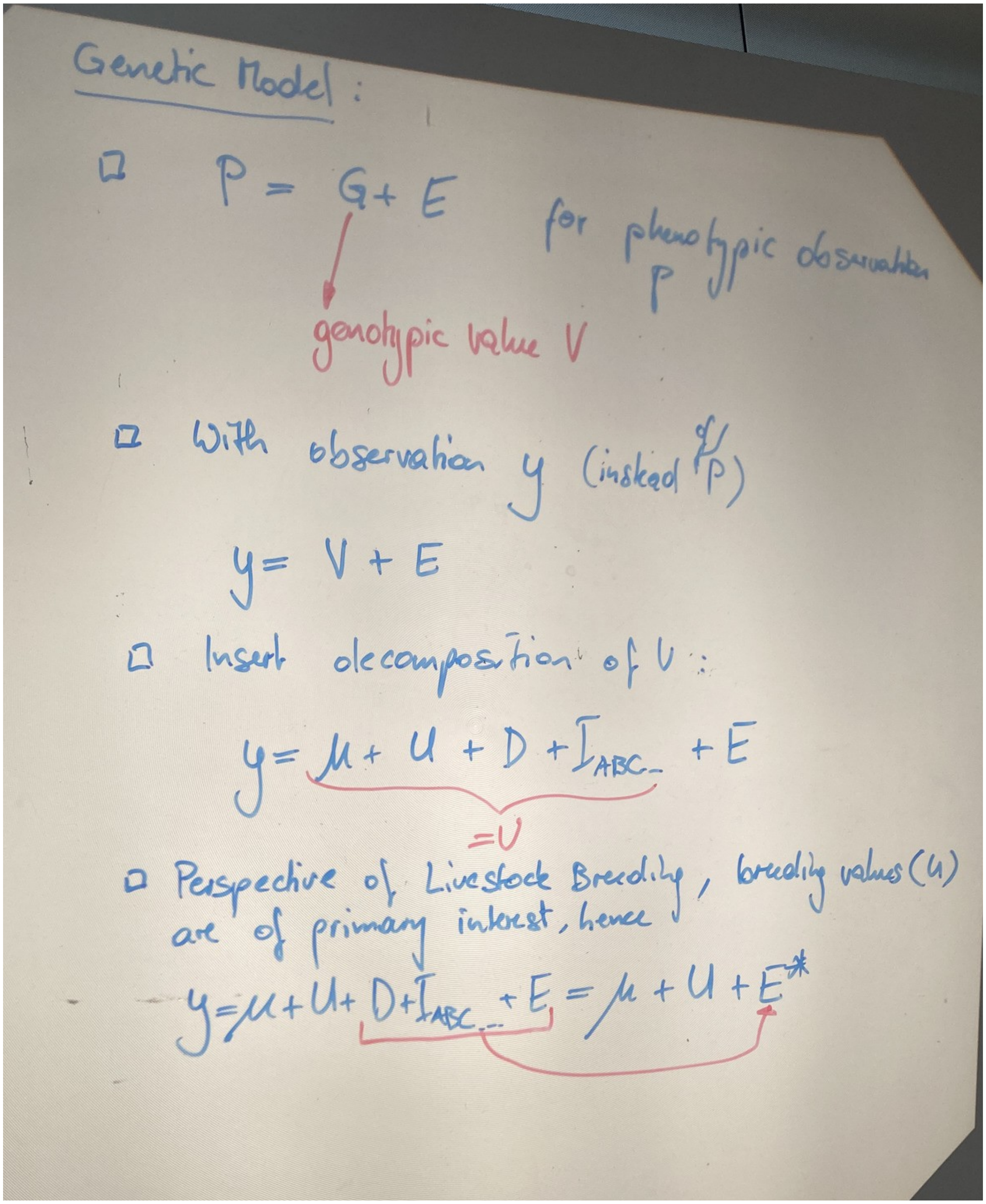
$$= \mu_A + BV_A + D_A + \mu_B + BV_B + D_B$$

$$+ \mu_C + BV_C + D_C + \ldots + I_{ABC}$$

$$= \mu_A + \mu_B + \mu_C + \ldots + BV_A + BV_B + BV_C + \ldots + D_A + D_B + D_C + \ldots + I_{ABC\ldots}$$

$$= \mu + U + D + I_{ABC\ldots}$$

OHP Picture 10



Genetic Model:

☐ $P = \underset{\downarrow \text{ genotypic value } V}{G} + E$ for phenotypic observation P

☐ With observation y (instead of P)

$$y = V + E$$

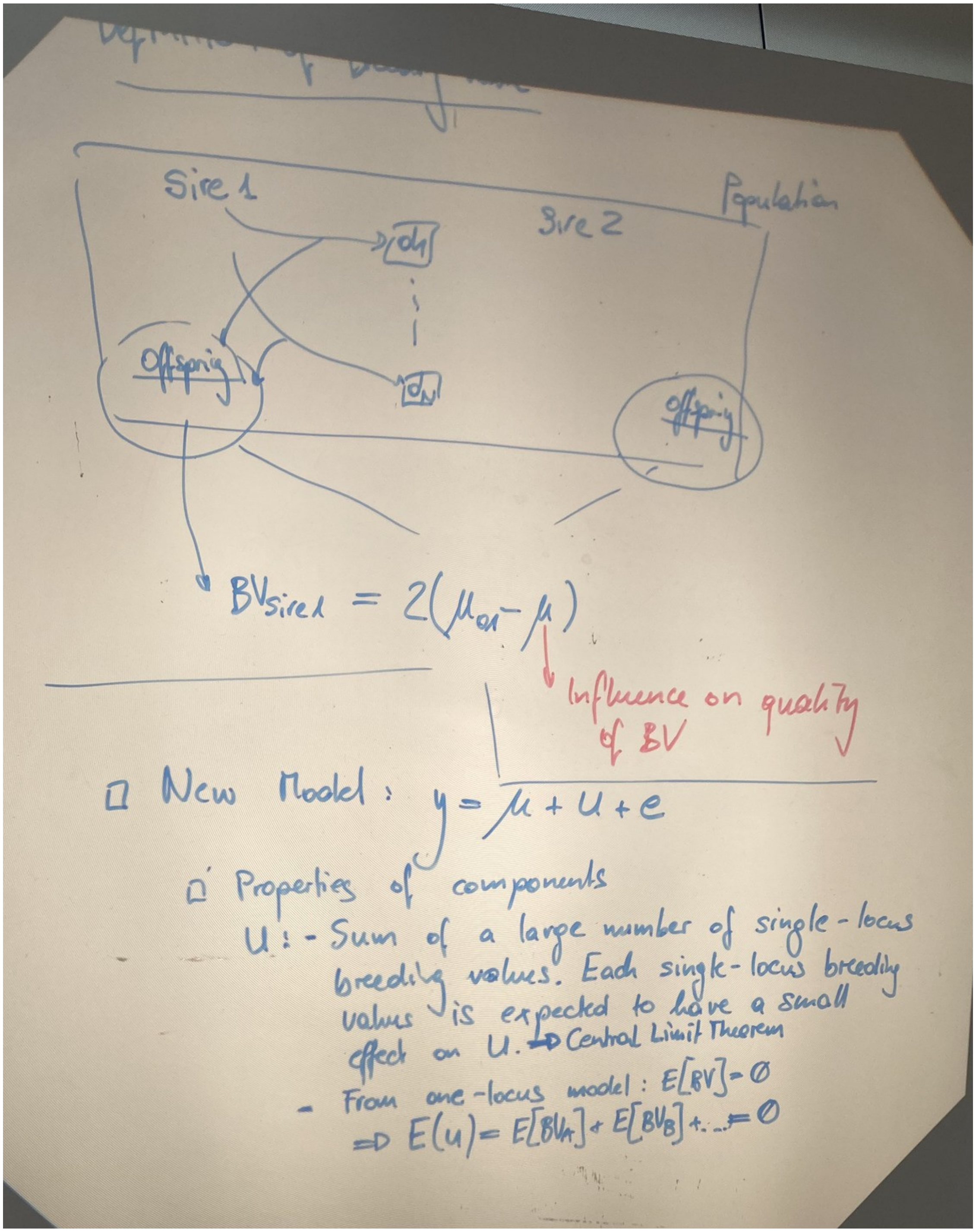
☐ Insert decomposition of V:

$$y = \underbrace{\mu + U + D + I_{ABC\ldots}}_{=V} + E$$

☐ Perspective of Livestock Breeding, breeding values (U) are of primary interest, hence

$$y = \mu + U + \underbrace{D + I_{ABC\ldots} + E}_{\to E^*} = \mu + U + E^*$$

OHP Picture 11



$BV_{sire1} = 2(\mu_{o1} - \mu)$

Influence on quality of BV

☐ New Model: $y = \mu + u + e$

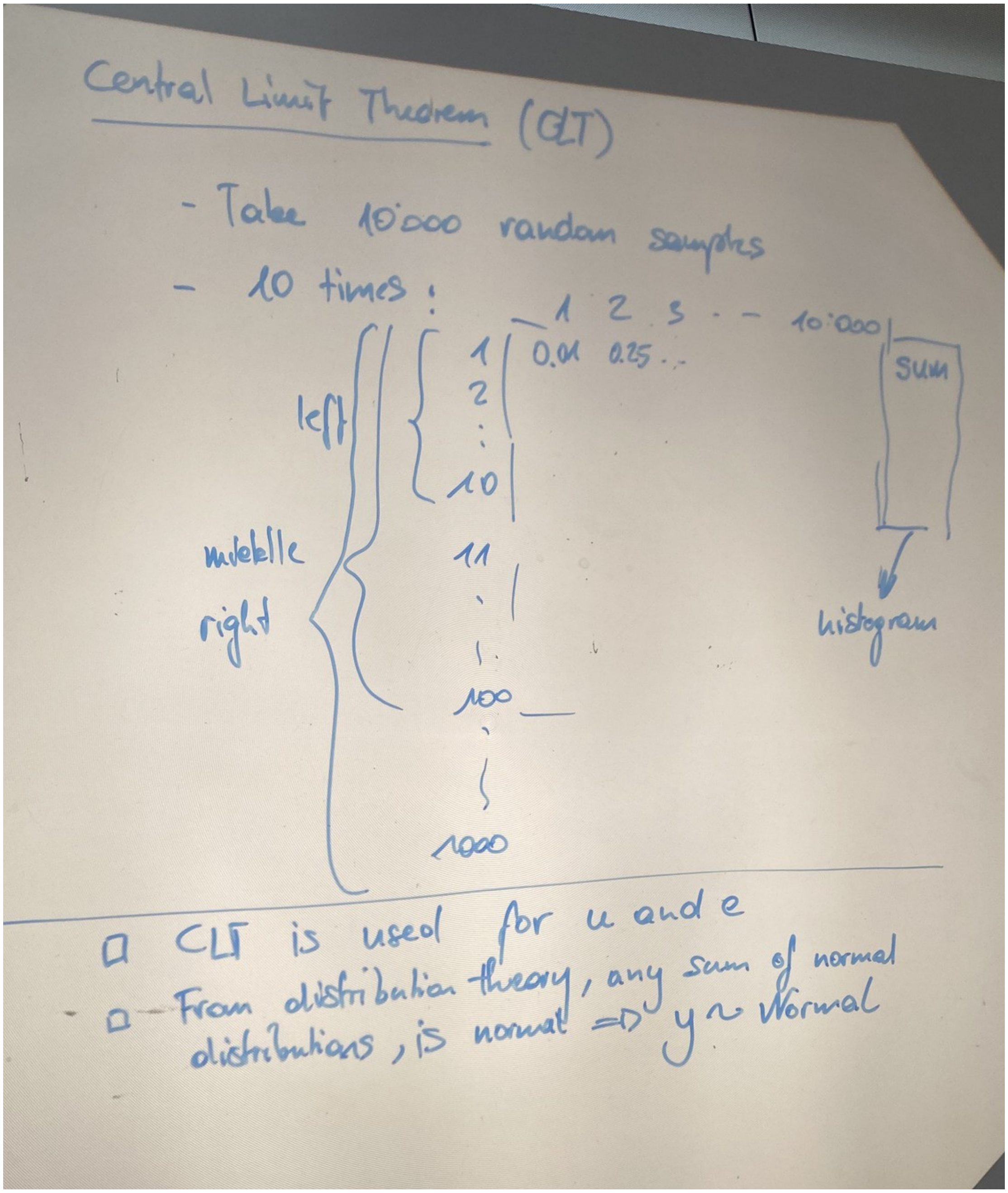
☐ Properties of components

u: - Sum of a large number of single-locus breeding values. Each single-locus breeding values is expected to have a small effect on u. → Central Limit Theorem

- From one-locus model: $E[BV] = 0$

$\Rightarrow E(u) = E[BV_A] + E[BV_B] + \ldots = 0$

OHP Picture 12



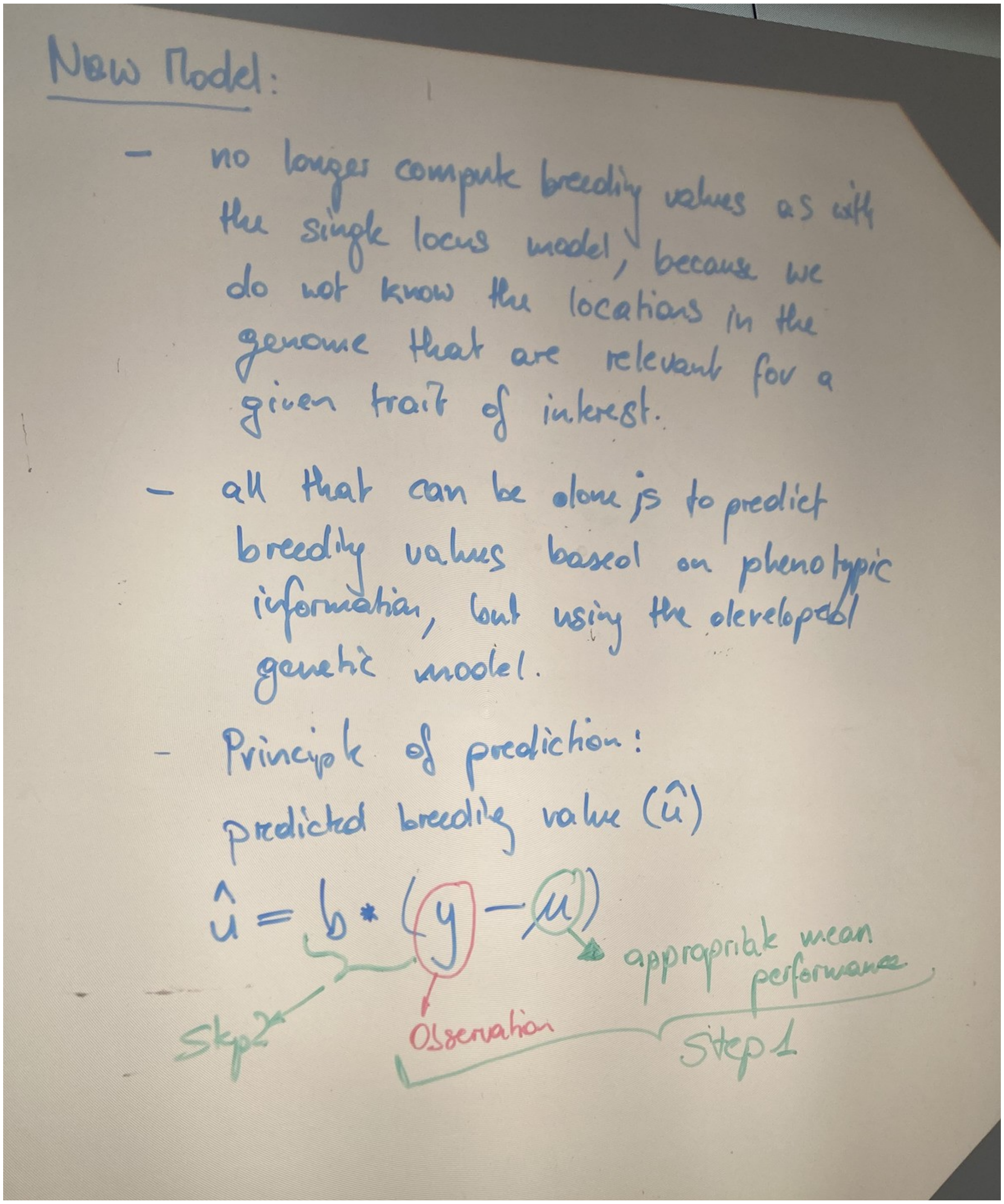
Central Limit Theorem (CLT)

- Take 10'000 random samples
- 10 times :

☐ CLT is used for u and e

☐ From distribution theory, any sum of normal distributions, is normal ⇒ y ~ Normal

OHP Picture 13



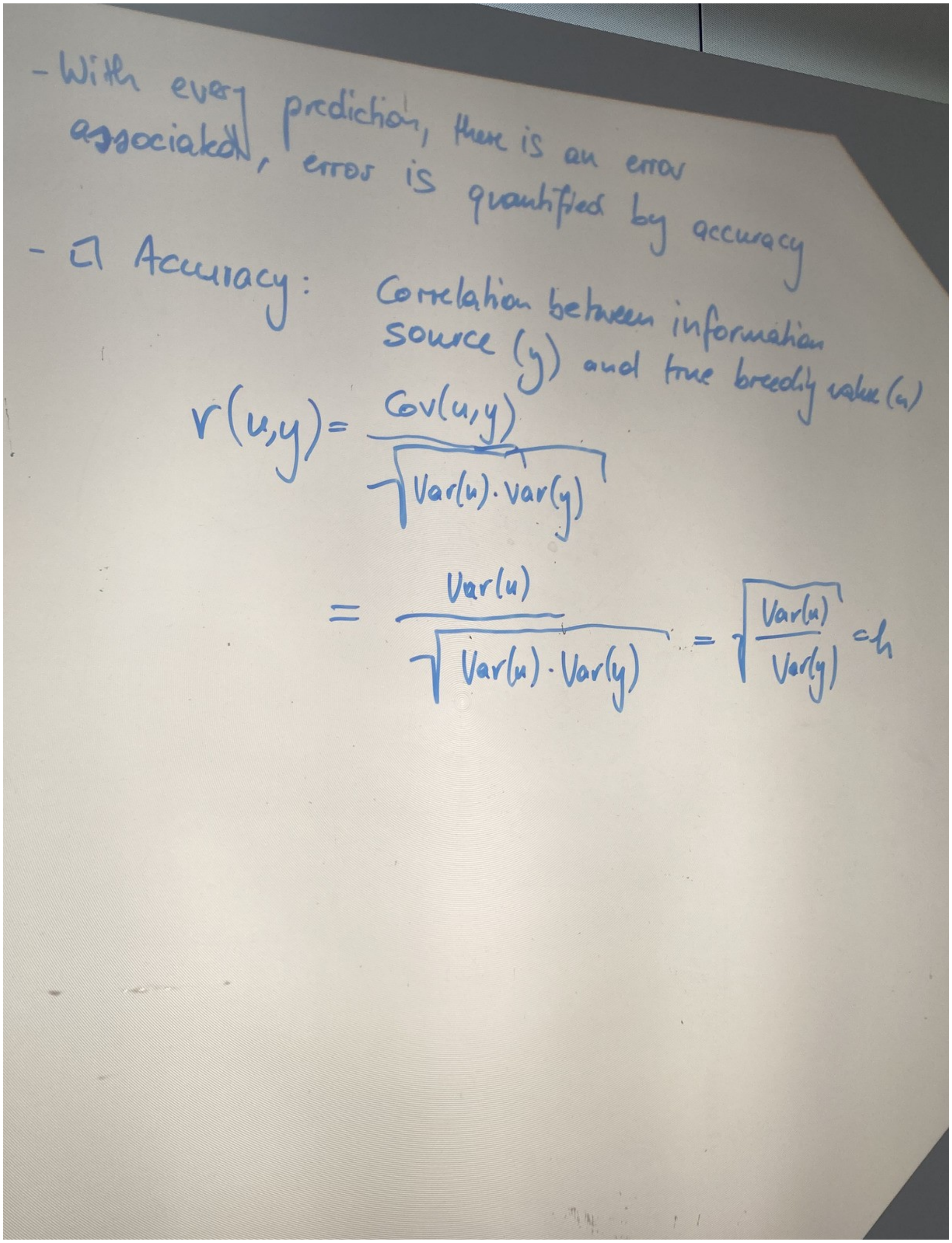
New Model:

- no longer compute breeding values as with the single locus model, because we do not know the locations in the genome that are relevant for a given trait of interest.
- all that can be done is to predict breeding values based on phenotypic information, but using the developed genetic model.
- Principle of prediction:

predicted breeding value ($\hat{u}$)

$$\hat{u} = \underbrace{b}_{\text{Step 2}} * (\underbrace{\underbrace{y}_{\text{Observation}} - \underbrace{\mu}_{\text{appropriate mean performance}}}_{\text{Step 1}})$$

OHP Picture 14



- With every prediction, there is an error associated, errors is quantified by accuracy
- ☐ Accuracy: Correlation between information source (y) and true breeding value (u)

$$r(u,y) = \frac{Cov(u,y)}{\sqrt{Var(u) \cdot var(y)}}$$

$$= \frac{Var(u)}{\sqrt{Var(u) \cdot Var(y)}} = \sqrt{\frac{Var(u)}{Var(y)}} = h$$