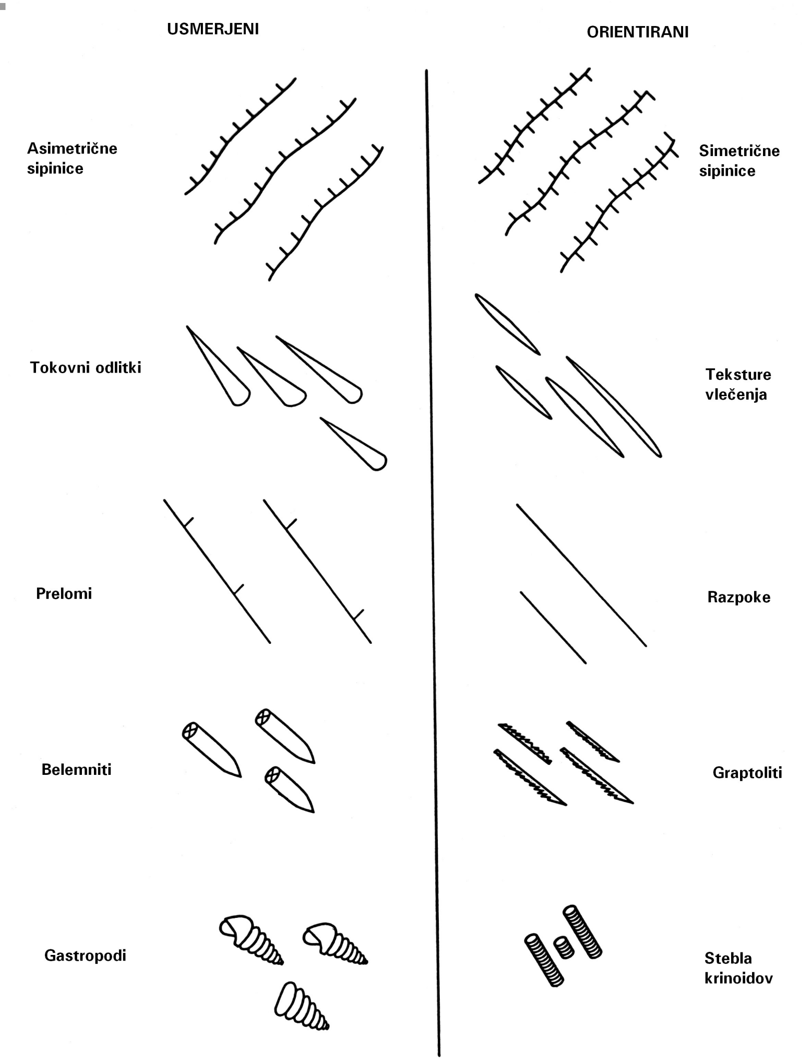
# 5. USMERJENI IN ORIENTIRANI PODATKI

## 5.1. UVOD

### 5.1.1. Definicije in vrste podatkov

- Krožni podatki so dveh vrst:
  - Usmerjeni
  - Orientirani

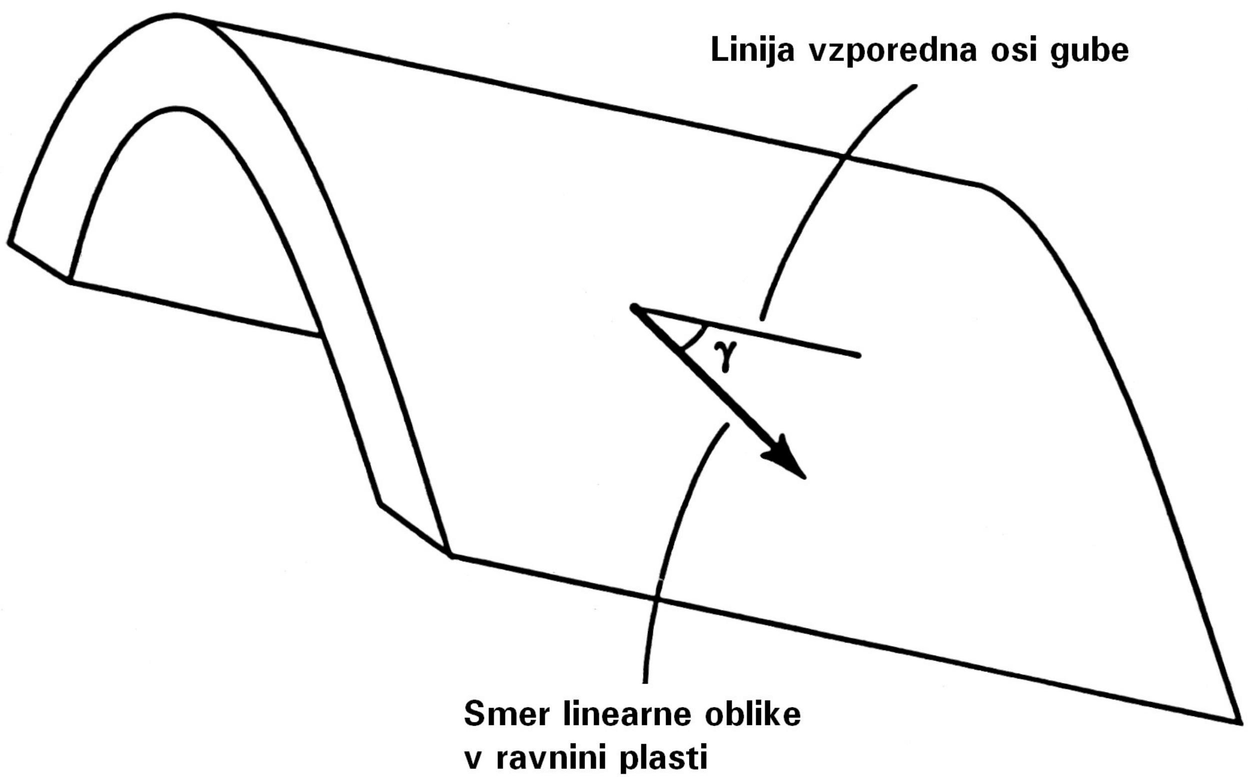


# 5.1.1. Definicije in vrste podatkov

- Preden pričnemo analizirati krožne smeri, je pogosto bistveno, da odstranimo vpliv tektonskega nagiba.
- Prvotni položaj neke linearne oblike lahko ugotovimo, če poznamo kot in smer nagiba sistema gub.
- Nagib celotne gube zavrtimo po navpični ravnini v vodoravno lego.
- Ravnino plasti potem zavritmo okrog osi gube v vodoraven položaj.
- Kot med linearno obliko in prvotno smerjo vpada se ohrani, nagib pa v narisu ohrani smer.

# 5.1.1. Definicije in vrste podatkov

- Prvoten vpad linearne oblike je enak smeri nagiba in kotu v ravnini plasti med linearno obliko in nagibom.



# 5.1.1. Definicije in vrste podatkov

- Če je:

  $\Theta$ smer nagiba linearne oblike (merjene v vodoravni ravnini)

  $\alpha$ kot nagiba linearne oblike (merjene v napični ravnini)

  $\Phi$ smer nagiba gube (merjene v vodoravni ravnini)

  $\beta$ kot nagiba gube (merjene v napični ravnini)

- Poiščemo enotni vektor ($x_1$, $y_1$, $z_1$) linearne oblike iz:

$x_1 = \cos\alpha \cdot \sin\Theta$

$y_1 = \cos\alpha \cdot \cos\Theta$

$z_1 = \sin\alpha$

- In enotni vektor ($x_p$, $y_p$, $z_p$) nagnjene gube iz:

$x_p = \cos\beta \cdot \sin\Phi$

$y_p = \cos\beta \cdot \cos\Phi$

# 5.1.1. Definicije in vrste podatkov

- Če je kot med linearno obliko in nagnjeno gubo v ravnini plasti $\gamma$, lahko načrtamo pravokotni trikotnik s hipotenuzo z dolžino ene enote.

  $\cos\gamma = a$

  kjer je a dolžina pravokotne projekcije enega vektorja na drugega.

- Zmnožek vektorjev je dolžina projekcije:

$$\cos\gamma = \cos\alpha \cdot \sin\Theta \cdot \cos\beta \cdot \sin\Phi + \cos\alpha \cdot \cos\Theta \cdot \cos\beta \cdot \cos\Phi + \sin\alpha \cdot \sin\beta$$

# 5.1.1. Definicije in vrste podatkov

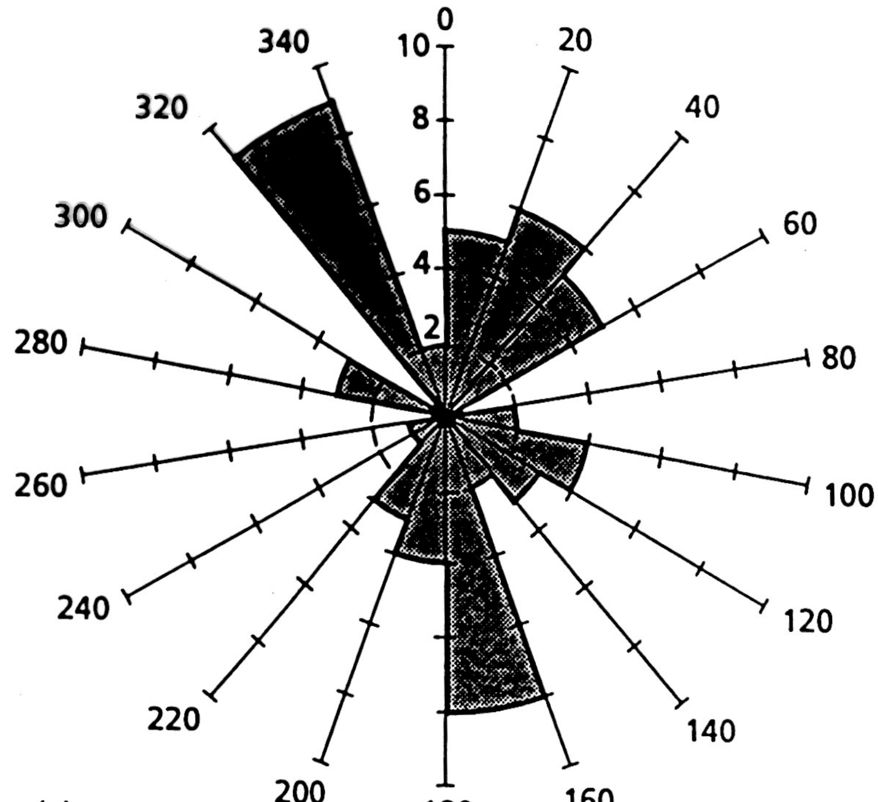
- Originalen vpad je torej:

$$\Phi+\cos^{1}(\cos\alpha\cdot\sin\Theta\cdot\cos\beta\cdot\sin\Phi+\cos\alpha\cdot\cos\Theta\cdot\cos\beta\cdot\cos\Phi+\sin\alpha\cdot\sin\beta)$$

- Izračun lahko uporabimo le za linearne oblike, ki ležijo v ravnini plasti ter kadar je med gubanjem prišlo le do enostavnih rotacij.
- Kadar gre za deformacijo ob strigu, se koti v ravnini plasti ne ohranijo.
- Že najmanjša deformacija onemogoči kakršnokoli oceno originalnega vpada.

# 5.1.2. Grafična predstavitev

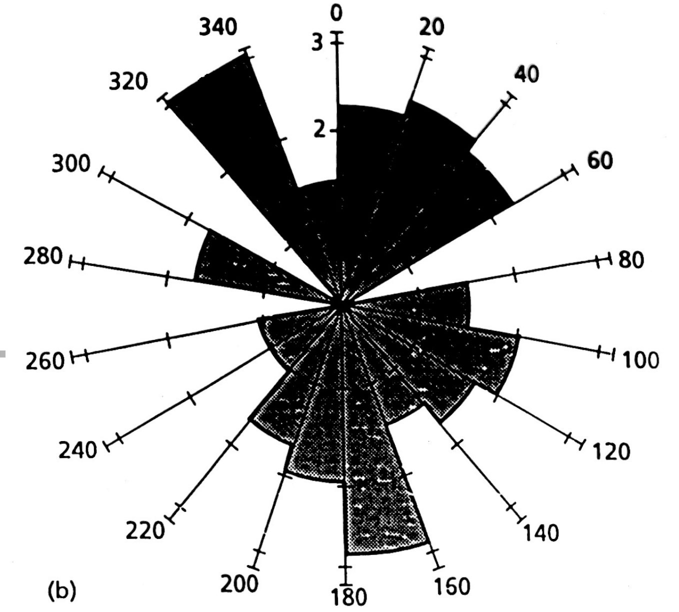
- Najpreprostejša grafična ponazoritev usmerjenih podatkov je s točkami, nanešenimi na obod kroga.
- Univariatnemu histogramu je pri krožnih podatkih enak veterni (rose) diagram, sestavljen iz delov kroga.
- Polmeri delov so sorazmerni frekvenci.

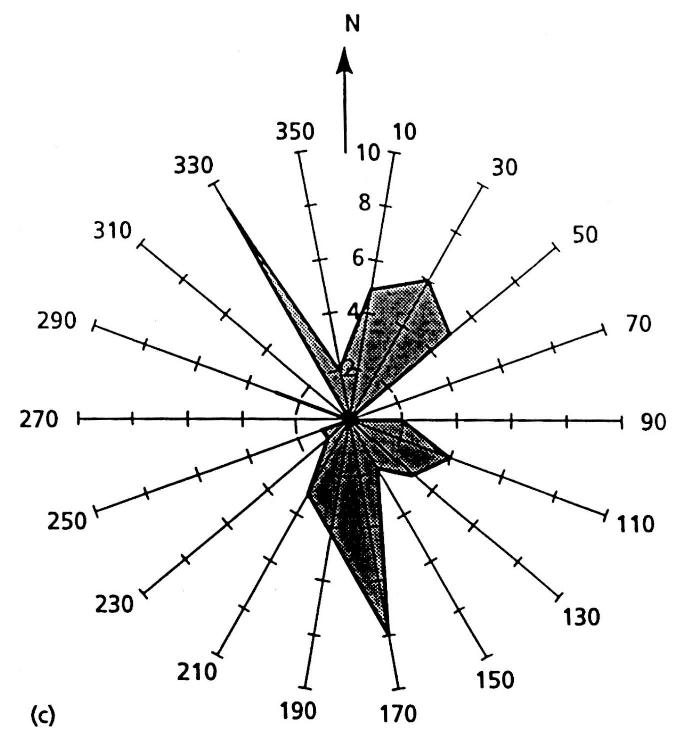


(a)

# 5.1.2. Grafična predstavitev

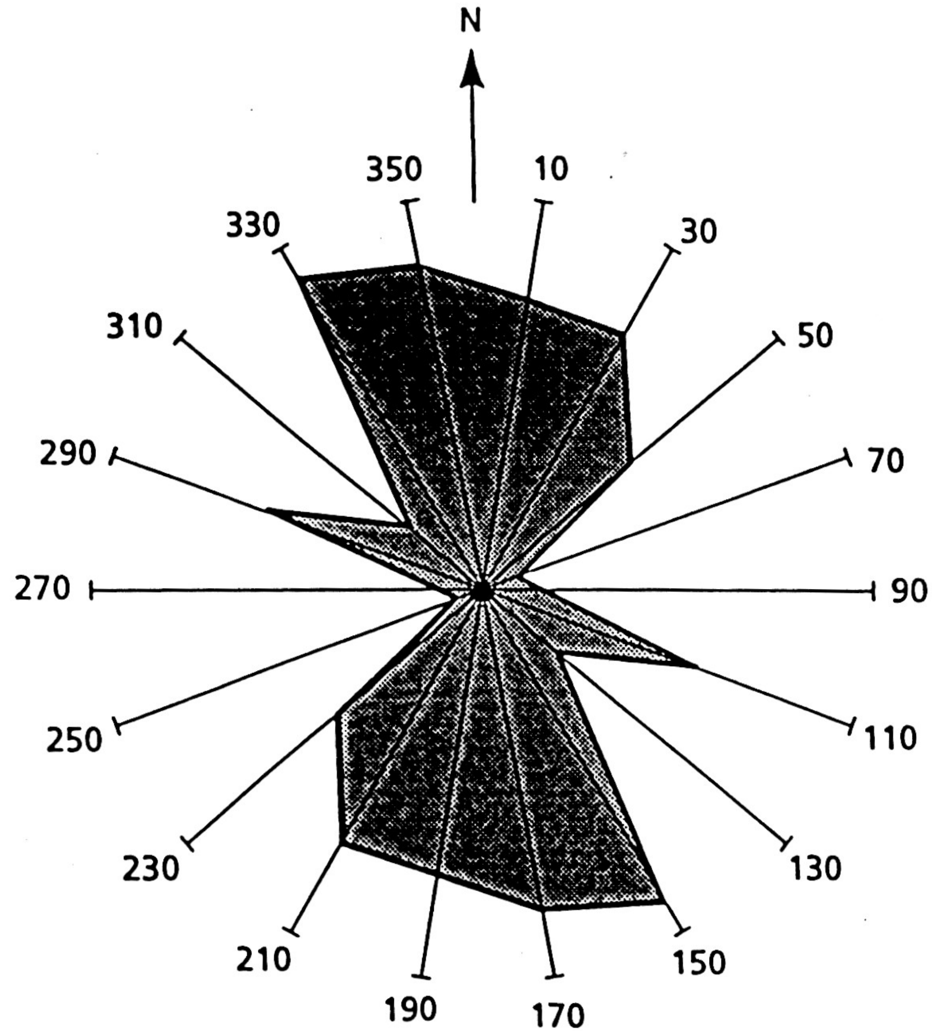
- Vizualni vtis dela je sorazmeren njegovi površini, ta pa r$^2$.
- Visoke frekvence so zato vizualno preveč, nizke premalo povdarjene, kar lahko vodi do napačnega vtisa prednostne smeri popolnoma naključnih podatkov.
- Polmere zato rošemo sorazmerne kvadratnemu korenu ali uporabimo zmajasti (kite) diagram.



# 5.1.2. Grafična predstavitev

- Pri orientiranih podatkih so si za $180^{0}$ nasprotne strani enakovredne, zato grafični prikaz omejimo na polkrog.
- Rotacijsko simetrijo lahko ohranimo tako, da imajo nasproti ležeče si smeri, enake frekvence.

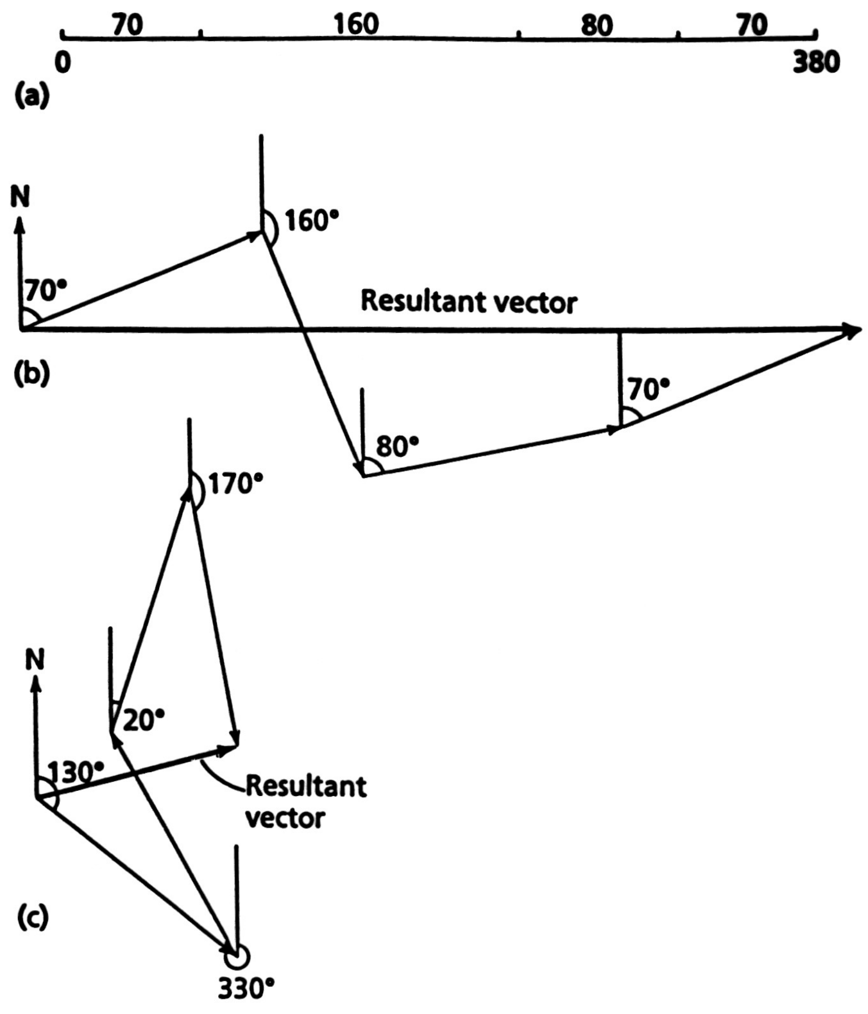


# 5.2. STATISTIKE Z USMERJENIMI PODATKI

- Potrebujemo postopek, ki obravnava 1 in 395 kot podobni vrednosti ter 0 in 360 kot enaki.
- Rešitev je v uporabi trigonometričnih funkcij – sinus, cosinus in tangens, ki se ponovijo vsakih $360^0$, kjer je $\sin(0^0) = \sin(360^0)$, ter enako cosinus in tanges.

# 5.2.1. Geometrijski koncept seštevanja smeri

- Usmerjene podatke predstavimo tako, da vsaka meritev določa smer vektorja in ne dolžine, ki je stalna.
- Vsoti ustreza premica, ki povezuje začetno in končno točko – rezultanta.
- Njena smer je srednja vrednost smeri.
- Dolžina rezultante pa odraža varianco.



# 5.2.2. Izračun povprečne smeri in mere sipanja

- Ponazorimo podatke s $\Theta_1$, $\Theta_2$, ..., $\Theta_n$.
- Komponenti vsakega enotnega vektorja sta:

  $x_i = \sin\Theta_i$

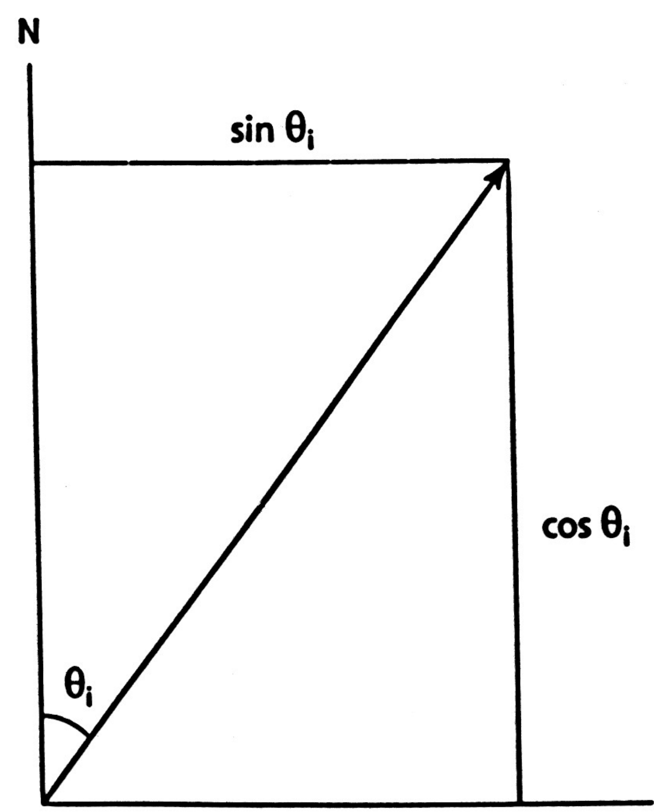
  $y_i = \cos\Theta_i$
- Komponenti vektorja rezultante:

  $x_r = \Sigma\sin\Theta_i$

  $y_r = \Sigma\cos\Theta_i$
- Srednja vrednost smeri je smer hipotenuze:

$\overline{\Theta}$

$$\overline{\Theta} = \tan^{-1}\left(\frac{x_r}{y_r}\right)$$



# 5.2.2. Izračun povprečne smeri in mere sipanja

- Opozorila:
  - Upoštevajte, da večina računalnikov normalno delijo v radijanih in ne stopinjah (1 rad = 57,2956$^0$).
  - Ker se funkcija tangens ponovi vsakih 180$^0$, daje tan$^{-1}$ pogosto zavajujoče rezultate, če računamo z osebnim ali žepnim računalnikom – vrednost bo izračunana v obsegu -90$^0$ do 90$^0$, dejanski rezultat pa bi moral biti v razponu 0$^0$ do 360$^0$.

# 5.2.2. Izračun povprečne smeri in mere sipanja

- Pravo vrednost poiščemo tako, da dodamo $0^0$, $180^0$ ali $360^0$ obrnjeni vrednosti, glede na predznaka komponent x in y.

| x | y | dodaj |
|---|---|---|
| + | + | $0^0$ |
| + | - | $180^0$ |
| - | - | $180^0$ |
| - | + | $360^0$ |

# 5.2.2. Izračun povprečne smeri in mere sipanja

- Dolžina rezultante je:

$$R = \sqrt{x_r^2 + y_r^2}$$

- Sipanje podatkov je odvisno od velikosti vzorca n, zato uporabljamo srednjo vrednost dolžine rezultante:

$$\overline{R} = \frac{R}{n}$$

- Neugodna lastnost je, da višji $\overline{R}$ pomeni nižjo varianco, zato raje uporabljamo krožno varianco:

$$\sigma_0 = 1 - \overline{R}$$

# 5.2.3. Frekvenčne porazdelitve

- Ničelna naključna frekvenčna porazdelitev usmerjenih podatkov je enotna porazdelitev, brez modalne najvišje vrednosti.
- Večina geoloških razmer daje usmerjene podatke s sipanjem vrednosti okrog prednostne smeri, ki morda nakazuje prevladujočo smer toka.
- Zato lahko pri statističnih izračunih uporabimo kot pomoč normalno porazdelitev.

# 5.2.3. Frekvenčne porazdelitve

- Če z usmerjenimi podatki uporabimo normalno porazdelitev, se bosta skrajna konca porazdelitve nujno ovijala po krožnici, tako da se bodo skrajne vrednosti začele približevati srednji vrednosti iz nasprotne smeri!
- Včasih zato uporabljamo “ovito normalno”, pogosteje pa von Misesovo porazdelitev.

# 5.2.3. Frekvenčne porazdelitve

- Von Mises ima dva parametra:
  - Srednjo vrednost smeri $\overline{\Theta}$
  - Parameter koncentracije $\kappa$
- Parameter koncentracije $\kappa$ ocenimo iz srednje vrednosti dolžine rezultante.
- Njegove vrednosti naraščajo s padajočim sipanjem, tako da $\kappa = 0$ pomeni enotno porazdelitev.

# 5.2.4. Test značilnosti povprečne smeri

## 5.2.4.1. Test povprečne dolžine rezultante

- Rayleighov test značilnosti srednje vrednosti smeri izračunamo:

$$\bar{R} = \frac{1}{n}\sqrt{\left(\sum_{i=1}^{n}\sin\Theta_i\right)^2 + \left(\sum_{i=1}^{n}\cos\Theta_i\right)^2}$$

- $H_0$, da ni značilne prevladujoče smerizavrnemo, kadar izračunana vrednost presega kritično.

| | α (%) | | | |
|---|---|---|---|---|
| *n* | 10 | 5 | 2.5 | 1 |
| 3 | 0.86 | 0.96 | 1.03 | 1.13 |
| 4 | 0.75 | 0.84 | 0.91 | 1.03 |
| 5 | 0.67 | 0.75 | 0.82 | 0.94 |
| 6 | 0.62 | 0.69 | 0.76 | 0.87 |
| 7 | 0.57 | 0.64 | 0.70 | 0.82 |
| 8 | 0.53 | 0.60 | 0.66 | 0.77 |
| 9 | 0.50 | 0.57 | 0.63 | 0.73 |
| 10 | 0.48 | 0.54 | 0.59 | 0.70 |
| 11 | 0.46 | 0.52 | 0.57 | 0.67 |
| 12 | 0.44 | 0.49 | 0.54 | 0.64 |
| 13 | 0.42 | 0.48 | 0.52 | 0.62 |
| 14 | 0.40 | 0.46 | 0.51 | 0.60 |
| 15 | 0.39 | 0.44 | 0.49 | 0.58 |
| 16 | 0.38 | 0.43 | 0.47 | 0.56 |
| 17 | 0.37 | 0.42 | 0.46 | 0.54 |
| 18 | 0.36 | 0.41 | 0.45 | 0.53 |
| 19 | 0.35 | 0.39 | 0.44 | 0.52 |
| 20 | 0.34 | 0.38 | 0.42 | 0.50 |
| 22 | 0.32 | 0.37 | 0.41 | 0.48 |
| 24 | 0.31 | 0.35 | 0.39 | 0.46 |

# 5.2.4.2. Test enotnosti

- Pri naključni porazdelitvi, ki se ne razlikuje bistveno od enotne, je pričakovana frekvenca v vsakem razredu enaka, in sicer preprosto n/k, kjer je k število razredov.
- Testiramo jo s $X^2$ testom, kjer $H_0$ enotne porazdelitve zavrenemo, če izračunana vrednost presega kritično.

# 5.2.4.2. Test enotnosti

- $H_0$: podatki izhajajo iz enotno porazdeljene populacije

  $H_1$: podatki izhajajo iz neenotno porazdeljene populacije

$$X^2 = \frac{\sum (O_j - E_j)^2}{E_j}$$

- Ker je pričakovani model enoten, je $E_j$ enak za vse razrede, zato poenostavimo:

$$X^2 = \frac{1}{E_j} \sum (O_j - E_j)$$

# 5.2.4.3. Bimodalne in polimodalne porazdelitve

- Rayleighov test ne bo ovrgel ničelne hipoteze, kadar so podatki bi- ali polimodalni, zlasti če si ležita modusa nasproti.
- Pri dodajanju členov k dolžini rezultante se nasprotni si smeri izničita.
- Bi- ali polimodalni podatki se kljub temu značilno razlikujejo od enotnega modela v $X^2$ testu.

# 5.2.5. Testiranje hipotetične povprečne smeri
## 5.2.5.1. En vzorec

- Izračunamo interval zaupanja srednje vrednosti smeri in preverimo, če predpostavljena smer znotraj tega intervala.
- Pri $\alpha = 0{,}05$ bo dejanska srednja vrednost populacije v razponu od:

$$\overline{\Theta} \pm 1{,}96 \cdot s_e$$

- kjer je $s_e$ standardna napaka srednje vrednosti, n velikost vzorca, $\kappa$ koncentracijski parameter (tabele) in $\bar{R}$ srednja vrednost dolžine rezultante:

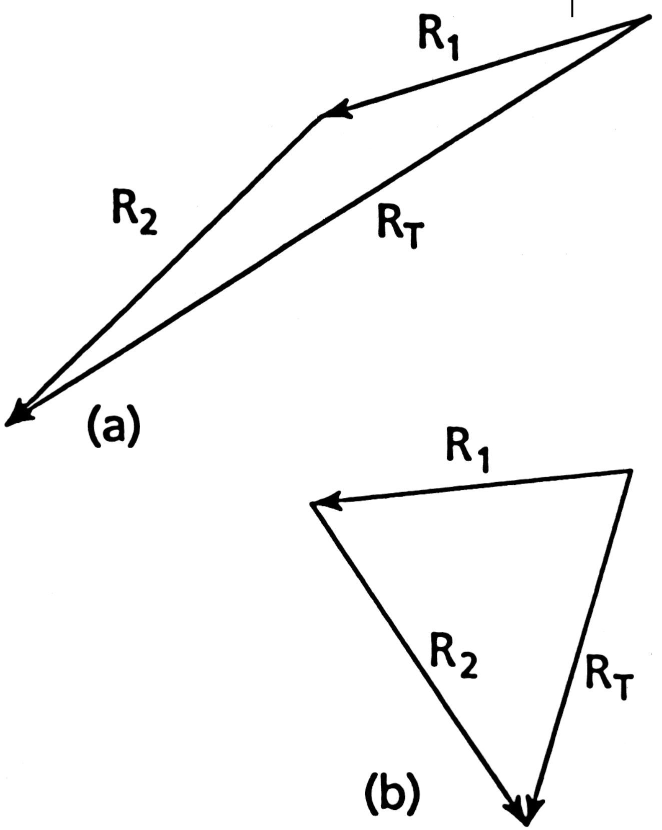
$$s_e = \frac{1}{\sqrt{n\bar{R}\kappa}}$$

# 5.2.5.1. En vzorec

- Opozorila:
  - Enote $s_e$ so radiani in ne stopinje, pretvorimo jih z množenjem z 57,2956.
  - Podatki se morajo vsaj približno ujemati s von Misesovo porazdelitvijo.
  - Najprej naredimo Rayleighov test prednostne smeri. Če je ni, bo standardna napaka velika in predpostavljena srednja vrednost bo zato verjetno padla znotraj intervala zaupanja, a bo rezultat nesmiselen.
  - Ker ne poznamo verjetnosti zavrnitve $H_0$, ne moremo zaupati ujemanju podatkov z določeno srednjo vrednostjo. To je ustrezno, ker je tudi pri podatkih, ki jih Rayleighov test potrdi, $s_e$ pogosto velika in daje interval zaupanja, ki zajema velik del kroga.

# 5.2.5.2. Dva vzorca

- Za test enakost dveh srednjih vrednosti smeri primerjamo dolžini rezultant $R_1$ in $R_2$, z rezultanto združenih dolžin obeh vzorcev $R_T$.
- Če sta si srednji vrednosti vzorcev blizu, bo vsota dolžin $R_1 + R_2$ podobna vrednosti $R_T$ (a), če pa sta si zelo različni, bo združeni rezultirajoči vektor precej krajši (b).



# 5.2.5.2. Dva vzorca

- $H_0$: vzorca izhajata iz populacij z enakima srednjima vrednostima smeri
- $H_1$: vzorca izhajata iz populacij z različnima srednjima vrednostima smeri

$$F = \left(1 + \frac{3}{8\kappa}\right)\frac{(n-2)(R_1 + R_2 - R_T)}{(n - R_1 - R_2)}$$

- κ poiščemo v tabelah na osnovi $\overline{R}_T$.
- Izračunani F primerjamo s kritičnim $F_{\alpha,1,n-2}$ in $H_0$ zavrnemo, če izračunana vrednost presega tabelirano.

| $\bar{R}$ | $\kappa$ | $\bar{R}$ | $\kappa$ |
|---|---|---|---|
| 0.00 | 0.00000 | 0.59 | 1.47543 |
| 0.02 | 0.04001 | 0.60 | 1.51574 |
| 0.04 | 0.08006 | 0.61 | 1.55738 |
| 0.06 | 0.12022 | 0.62 | 1.60044 |
| 0.08 | 0.16051 | 0.63 | 1.64506 |
| 0.10 | 0.20101 | 0.64 | 1.69134 |
| 0.12 | 0.24175 | 0.65 | 1.73945 |
| 0.14 | 0.28279 | 0.66 | 1.78953 |
| 0.16 | 0.32419 | 0.67 | 1.84177 |
| 0.18 | 0.36599 | 0.68 | 1.89637 |
| 0.20 | 0.40828 | 0.69 | 1.95357 |
| 0.22 | 0.45110 | 0.70 | 2.01363 |
| 0.24 | 0.49453 | 0.71 | 2.07685 |
| 0.26 | 0.53863 | 0.72 | 2.14359 |
| 0.28 | 0.58350 | 0.73 | 2.21425 |
| 0.30 | 0.62922 | 0.74 | 2.28930 |
| 0.31 | 0.65242 | 0.75 | 2.36930 |
| 0.32 | 0.67587 | 0.76 | 2.45490 |
| 0.33 | 0.69958 | 0.77 | 2.54686 |
| 0.34 | 0.72356 | 0.78 | 2.64613 |
| 0.35 | 0.74783 | 0.79 | 2.75382 |
| 0.36 | 0.77241 | 0.80 | 2.87129 |

# 5.2.5.2. Dva vzorca

- Opozorila:
  - Oba vzorca morata izhajati iz von Misesove porazdelitve.
  - Zgornja enačba velja zgolj za $10>\kappa>2$. $\kappa>10$ je zelo malo verjeten, in če je $\kappa<2$, je zelo verjetno, da ni prednostne smeri, ali da sta si srednji vrednosti smeri očitno različni.

# 5.2.6. Nekateri hitri neparametrični testi

- Prednost neparametričnih usmerjenih metod je, da ne zahtevajo von Misesove porazdelitve, ter da jih hitro in preprosto izračunamo.
- Taki sta npr.:
  - Hodges – Ajnejev test prednostne smeri
  - Test nizov enakosti srednjih vrednosti smeri

# 5.2.6.1. Hodges – Ajnejev test prednostne smeri

- Podatke narišemo na obod kroga.
- Kadar obstaja prednostna smer, bodo točke zgoščene na enem delu krožnice.
- Potujemo po krožnici, dokler ne dosežemo največje razlike med številom točk na vsaki strani.
- Hodges – Ajnejeva statistika $M_{min}$ je najmanjše število točk na eni strani.
- Test ni zelo občutljiv.

# 5.2.6.2. Test nizov enakosti srednjih vrednosti smeri

- Podatke dveh vzorcev združimo in jih razvrstimo v naraščajočem vrstnem redu.
- Niz predstavlja vrsto zaporednih točk istega vzorca.
- Kadar dva vzorca izhajata iz populacij z enakima smerema pričakujemo, da bosta vzorca med seboj pomešana in bo število nizov visoko.
- Če se smeri močno razlikujeta, bodo opazovanja razvrščena le v dva niza.
- Razprševanje podatkov mora biti v obeh vzorcih približno enako, ker lahko pride do anomalno majhnega števila nizov, tudi ob enakih srednjih vrednostih smeri, če so opazovanja enega izmed vzorcev tesno skupaj.

# 5.2.6.2. Test nizov enakosti srednjih vrednosti smeri

- Testna statistika je :

$$Z = \frac{U - \overline{U}}{\sigma_{\overline{U}}}$$

- Pričakovano število nizov je:

$$\overline{U} = 1 + \frac{2n_1 n_2}{n_1 + n_2}$$

- In pričakovana varianca:

$$\sigma_{\overline{U}}^2 = \frac{2n_1 n_2 (2n_1 n_2 - n_1 - n_2)}{(n_1 + n_2)^2 (n_1 + n_2 - 1)}$$

# 5.3. Statistike orientiranih podatkov

- Izračune za orientirane podatke izvajamo enako kot za usmerjene, le da so vsi koti že v naprej podvojeni.
- Vse kotne rezultate, zato razpolovimo.
- Kote podvajamo zato, da postanejo smeri, ki se razlikujejo za $180^0$ (enaka orientacija), $360^0$ narazen in s tem identične v trigonometričnih izračunih.
- Vsi ostali testi so veljavni s podvojenimi podatki.