

Introduction to Kac-Moody Lie algebras

Nicolas Perrin

<http://www.hcm.uni-bonn.de/?id=961>

version of 2011

Errata and addenda by Darij Grinberg**0.1. Errata**

I am not an expert in Lie theory; hence, please approach the corrections below with a critical eye.

- **Definition 2.2.1:** Replace $x \otimes y - z \otimes x - [x, y]$ by $x \otimes y - y \otimes x - [x, y]$. (Could this be due to the switched "y" and "z" keys on the German keyboard layout?)
- **Theorem 2.2.4:** The product $x_1^{a_1} \cdots x_n^{a_n}$ should be $e_1^{a_1} \cdots e_n^{a_n}$ here.
- **Proposition 3.1.2:** Replace ℓ by n (in "size ℓ ").
- **Proof of Proposition 3.1.2:** "These is easy" should be "This is easy".
- **Proof of Proposition 3.1.2:** I don't understand the part of this proof that begins with "For this, we may assume that $A_3 = 0$ and $A_4 = 0$ " and ends with "and the matrix C is non degenerate". Why can we assume that $A_3 = 0$ and $A_4 = 0$ without changing things, and why do we have the $(\text{Vect}(\dots))^\perp = \text{Vect}(\dots)$ relations (particularly the second one)?

(Here is how I would show that the matrix C is nondegenerate: Since $\alpha_1, \alpha_2, \dots, \alpha_n$ are linearly independent and $\langle \cdot, \cdot \rangle$ is a nondegenerate bilinear form, the block matrix

$\begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \\ X_1 & X_2 \end{pmatrix}$ has rank n . But each row of the

"middle part" (by this I mean the $\begin{pmatrix} A_3 & A_4 \end{pmatrix}$ part) of this matrix is a linear combination of the rows of the "upper part" (the $\begin{pmatrix} A_1 & A_2 \end{pmatrix}$ part) (because

$\text{rank} \begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix} = \text{rank} A = \ell = \text{rank} A_1 \leq \text{rank} \begin{pmatrix} A_1 & A_2 \end{pmatrix}$). Hence, by

performing row operations to the matrix $\begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \\ X_1 & X_2 \end{pmatrix}$, we can replace the

$\begin{pmatrix} A_3 & A_4 \end{pmatrix}$ part by zeroes.¹ Since row operations don't change the rank,

¹Is this what you mean by "assume that $A_3 = 0$ and $A_4 = 0$ "?

this yields that $\text{rank} \begin{pmatrix} A_1 & A_2 \\ 0 & 0 \\ X_1 & X_2 \end{pmatrix} = \text{rank} \begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \\ X_1 & X_2 \end{pmatrix}$. Thus,

$$n = \text{rank} \begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \\ X_1 & X_2 \end{pmatrix} = \text{rank} \begin{pmatrix} A_1 & A_2 \\ 0 & 0 \\ X_1 & X_2 \end{pmatrix} = \text{rank} \begin{pmatrix} A_1 & A_2 \\ X_1 & X_2 \end{pmatrix}.$$

Now,

$$\begin{aligned} \text{rank } C &= \text{rank} \begin{pmatrix} A_1 & A_2 & 0 \\ A_3 & A_4 & I_{n-\ell} \\ X_1 & X_2 & 0 \end{pmatrix} = \text{rank} \begin{pmatrix} A_1 & A_2 & 0 \\ X_1 & X_2 & 0 \\ A_3 & A_4 & I_{n-\ell} \end{pmatrix} \\ &\quad (\text{since permutations of rows don't change the rank of a matrix}) \\ &= \underbrace{\text{rank} \begin{pmatrix} A_1 & A_2 \\ X_1 & X_2 \end{pmatrix}}_{=n} + \underbrace{\text{rank}(I_{n-\ell})}_{=n-\ell} \\ &\quad \left(\begin{array}{l} \text{since any block matrix of the form } \begin{pmatrix} U & 0 \\ V & I_m \end{pmatrix} \\ \text{satisfies } \text{rank} \begin{pmatrix} U & 0 \\ V & I_m \end{pmatrix} = \text{rank } U + m \end{array} \right) \\ &= n + n - \ell = 2n - \ell, \end{aligned}$$

so that C is nondegenerate, qed.)

- **Definition 3.1.4:** You say: "Any matrix can be decomposed as a direct sum of indecomposable matrix". Maybe you should add "(up to simultaneous permutation of rows and columns)" here.
- **Definition 3.1.5:** Maybe add "for all $h \in \mathfrak{h}$ and $h' \in \mathfrak{h}$ " to the defining relations.
- **Proof of Theorem 3.1.6:** It would be better to explicitly distinguish between the vector space \mathfrak{h} which belongs to the realization of A , and the subspace \mathfrak{h} of the Lie algebra $\tilde{\mathfrak{g}}(A)$. It is clear that there is a canonical surjection from the former space to the latter space, but it is not a priori clear that this surjection is a bijection (i. e., that the relations given in Definition 3.1.5 don't force some elements of \mathfrak{h} to become zero). This does not become clear until the following argument in your proof:

"Assume there is a relation $n_- + h + n_+ = 0$ with $n_- \in \tilde{\mathfrak{n}}_-$, $h \in \mathfrak{h}$ and $n_+ \in \tilde{\mathfrak{n}}_+$ [...] so that $h = 0$ "

The h in the beginning of this argument means an element of $\tilde{\mathfrak{g}}(A)$, whereas the h in the end of this argument means a corresponding element of the original vector space \mathfrak{h} . Hence, this argument actually shows that if some

element h of our original vector space \mathfrak{h} becomes 0 in $\tilde{\mathfrak{g}}(A)$, then it must have been 0 to begin with. This justifies the identification of \mathfrak{h} with the image of $\mathfrak{h} \rightarrow \tilde{\mathfrak{g}}(A)$. It would help a lot if you make this explicit. (Maybe even in the theorem itself, not just in the proof...)

- **Proof of Theorem 3.1.6:** In the formula

$$\begin{aligned} (he_i - e_i h)(a \otimes v_j) &= [\dots] \\ &= [\dots] \\ &= [\dots] = \langle \alpha_i, h \rangle e_i(a \otimes v_j) \end{aligned}$$

(the parts that I have omitted are correct), replace both occurrences of $a \otimes v_j$ by $v_j \otimes a$.

- **Proof of Theorem 3.1.6:** When you say "there is a surjective map $U(\mathfrak{n}_-) \rightarrow T(V)$ ", you might want to add "and this map is an algebra homomorphism" (since otherwise the next sentence is not clear).
- **Theorem 3.1.6 and its Proof:** It seems that you write \mathfrak{n}_+ for $\tilde{\mathfrak{n}}_+$ (and, similarly, \mathfrak{n}_- for $\tilde{\mathfrak{n}}_-$) several times here (e. g., in part (iv) of the theorem).
- **Proof of Theorem 3.1.6:** When you write $n_-(1) = \langle \alpha, h \rangle 1$, I think you mean $n_-(1) = -\langle \alpha, h \rangle 1$.
- **Proof of Theorem 3.1.6:** In the formula

$$\mathfrak{n}_\pm = \bigoplus_{\alpha \in Q_+, \alpha \neq 0} \tilde{\mathfrak{g}}_\pm,$$

the $\tilde{\mathfrak{g}}_\pm$ should be $\tilde{\mathfrak{g}}_{\pm\alpha}$.

- **Proof of Theorem 3.1.6:** When you say "we have the inequality $\dim \tilde{\mathfrak{g}}_\alpha \leq n^{|\text{ht } \alpha|}$ ", you might want to add "for $\alpha \neq 0$ ".
- **Definition 3.1.8:** Add a whitespace after "(i)".
- **Definition 3.1.8:** Replace "is it also contained" by "it is also contained".
- **Between Definition 3.1.8 and Definition 3.1.9:** When you say "We have the estimate $\dim \mathfrak{g}_\alpha < n^{|\text{ht } \alpha|}$ ", you might want to add "for $\alpha \neq 0$ ".
- **Remark 3.1.10:** In the relation $\alpha_{i_1} + \alpha_{i_2} = \dots + \alpha_{i_k} = \alpha$, the first = sign should be a + sign.
- **Remark 3.1.10:** Replace $\omega(\Delta_+) = \Delta_-$ by $\omega^*(\Delta_+) = \Delta_-$ (or do you denote by ω any map canonically induced by ω ?).
- **Proposition 3.1.11:** In part (i), you should put a comma after " $\mathfrak{g}(A) = \mathfrak{g}'(A) + \mathfrak{h}$ ", lest the reader think that this holds "if and only if $\det(A) \neq 0$ ".

- **Proposition 3.1.11:** I think that when you refer to α_i in this proposition, you really mean α_i^\vee (this error occurs twice). Also, when you say " $\mathfrak{g}'(A) \cap \mathfrak{g}_\alpha = \mathfrak{g}_\alpha$ ", you mean it only for $\alpha \neq 0$ of course.
- **Proof of Proposition 3.1.11:** Replace " $\det(A) = 0$ " by " $\det(A) \neq 0$ ".
- **Proposition 3.2.5:** In (iii), replace $j(k+1-j)v_{j-1}$ by $i(k+1-i)v_{i-1}$. Also, replace $v_{k+1} = v_0 = 0$ by $v_{k+1} = v_{-1} = 0$.
- **Proof of Proposition 3.2.5:** In the second formula in the proof of part (i), the minus sign before $f^{k-1} \otimes h$ should be a plus sign.
- **Proof of Proposition 3.2.5:** The proof of part (ii) has several typos, all of them in the end: $-i(i-1)f^j(v) + jf^j(h(v)) = j(a-j+1)f^j(v)$ should be $-j(j-1)f^{j-1}(v) + jf^{j-1}(h(v)) = j(a-j+1)f^{j-1}(v)$.
- **Proof of Proposition 3.2.5:** In the proof of part (iii), it is better to replace "as soon as they don't vanish" by "as long as they don't vanish", or at least so I believe. Also, "Let l be the smallest number" should be "Let l be the largest number".
- **Proof of Proposition 3.2.6:** Replace $g_{(i)}$ by $\mathfrak{g}_{(i)}$.
- **Proof of Lemma 3.2.7:** Maybe say at the beginning of the proof that you are going to consider $x \in \mathfrak{n}_+$ (not $x \in \mathfrak{n}_-$).
- **Proof of Lemma 3.2.7:** Replace "It it clear" by "It is clear".
- **Proof of Lemma 3.2.7:** The argument why $\text{ad}(f_i)$ sends \mathfrak{i} to \mathfrak{i} seems somewhat unclear to me. Instead I would say that we can rewrite any element of the form

$$\begin{aligned} & (\text{ad } f_i) \left([e_{i_1}, [\dots, [e_{i_s}, [h_1, [\dots, [h_k, x] \dots]]] \dots] \right) \\ & = ((\text{ad } f_i) (\text{ad } e_{i_1}) \dots (\text{ad } e_{i_s}) (\text{ad } h_1) \dots (\text{ad } h_k)) (x) \end{aligned}$$

as a linear combination of elements of the form

$$((\text{ad } e_{u_1}) \dots (\text{ad } e_{u_s}) (\text{ad } h_{w_1}) \dots (\text{ad } h_{w_r}) (\text{ad } f_{v_1}) \dots (\text{ad } f_{v_t})) (x)$$

(by the easy part of the Poincaré-Birkhoff-Witt theorem), and elements of the latter form vanish whenever $t > 0$ and lie in \mathfrak{i} whenever $t = 0$.

- **Proof of Lemma 3.2.7:** Replace " $a = 0$ " by " $x = 0$ ".
- **Proof of Proposition 3.2.6 (continued after proof of Lemma 3.2.7):** Replace "Lemme 3.2.5" by "Proposition 3.2.5 (ii)". In the formula directly below this, replace $f_i^{a_{i,j}}$ by $f_i^{-a_{i,j}}$.

- **Proof of Proposition 3.2.6 (continued after proof of Lemma 3.2.7):** I think that $-a_{i,j}(\operatorname{ad} f_i)^{-a_{i,j}}(f_i)$ should be $a_{j,i}(\operatorname{ad} f_i)^{-a_{i,j}}(f_i)$ here. Now you use the " $a_{i,j} = 0 \implies a_{j,i} = 0$ " condition from Definition 3.2.1 to see that this vanishes. (But I may very well be mistaken.)
- **Proof of Lemma 3.2.8:** By "block matrix" you mean "block-diagonal matrix".
- **Proposition 3.2.9:** In part (ii), replace $g'(A)$ by $\mathfrak{g}'(A)$.
- **Proposition 3.2.9:** The conclusion of part (iii) should not be " $g'(A)/\mathfrak{c} = 0$ " but it should be " $\mathfrak{g}'(A)/\mathfrak{c}$ is simple".
- **Proof of Proposition 3.2.9:** You write: "If for any α we have $\mathfrak{i} \cap \mathfrak{g}_\alpha = 0$ then $\mathfrak{i} \subset \mathfrak{h}$ ". Here, it would be better to replace "any" by "all", since "any" could also mean "some". Also, again you should say that you are only considering $\alpha \neq 0$.
- **Proof of Proposition 3.2.9:** You write: "We can therefore take α a root minimal". By "minimal" you mean "minimal among the roots in Q_+ " (not all of Q).
- **Proof of Proposition 3.2.9:** Replace "colinear" by "collinear".
- **Proof of Proposition 3.2.9:** It would be better not to speak of "the center" here, but just say \mathfrak{c} , because "the center" might also mean the center of $g'(A)$ (and I am not sure whether this is the same center).
- **Proof of Proposition 3.2.9:** Replace "where $n \in \mathfrak{n}_- \oplus \mathfrak{n}_+$ and $h \in \mathfrak{h}$ " by "where $n \in \mathfrak{n}_- \oplus \mathfrak{n}_+$ and $h \in \mathfrak{h}'$ ".
- **Proof of Proposition 3.2.9:** At the moment when you write "By minimality, this implies that $\gamma = \alpha_i$ ", I am losing track of what you are doing. However, it is not hard to complete the proof from here:

Since $[f_i, x] \in \mathfrak{i}$ and $[f_i, x]_{\gamma - \alpha_i} \neq 0$, we get a contradiction to the minimality of γ unless either $[f_i, x]_0 \neq 0$ or $[f_i, x] \in \mathfrak{c}$. So we conclude that either $[f_i, x]_0 \neq 0$ or $[f_i, x] \in \mathfrak{c}$. In the former case, we must have $x_{\alpha_i} \neq 0$ (since $[f_i, x_{\alpha_i}] = [f_i, x]_0 \neq 0$). In the latter case, we must have $x_{\alpha_i} \neq 0$ as well (since $[f_i, x] \in \mathfrak{c} \subseteq \mathfrak{h}$ and thus $[f_i, x] = [f_i, x]_0$, so that $[f_i, x_{\alpha_i}] = [f_i, x]_0 \neq 0$). Hence, in both cases, we have $x_{\alpha_i} \neq 0$. Thus, x_{α_i} is a nonzero scalar multiple of e_i (since $x_{\alpha_i} \in \mathfrak{g}_{\alpha_i}$). Hence, $[f_i, x_{\alpha_i}]$ is a nonzero scalar multiple of $[f_i, e_i] = -\alpha_i^\vee$, therefore a nonzero scalar multiple of α_i^\vee . Since $[f_i, x_{\alpha_i}] = [f_i, x]_0$, this shows that $[f_i, x]_0$ is a nonzero scalar multiple of α_i^\vee . Since $[f_i, x]_0 \in \mathfrak{i}$ (because $[f_i, x] \in \mathfrak{i}$ and by Lemma 3.1.7), this yields $\alpha_i^\vee \in \mathfrak{i}$. Since $\alpha_i^\vee \notin \mathfrak{c}$ (this is easy to prove using Proposition 3.1.12 and the fact that A is an indecomposable Cartan matrix), this yields that there exists an element

$h \in \mathfrak{i} \cap \mathfrak{h}$ not in \mathfrak{c} (namely, $h = \alpha_i^\vee$). As you already have shown above, this concludes the proof.

- **Lemma 4.1.2:** In part (i), replace “ x, y and z ” by “ x and y ”.
- **Proof of Lemma 4.1.2:** You write: “Applying it to the adjoint representation gives the result.” Why? If you apply the formula

$$(\operatorname{ad} x)^k [y, z] = \sum_{i=0}^k \binom{k}{i} [(\operatorname{ad} x)^i y, (\operatorname{ad} x)^{k-i} z] \quad \text{in } U(\mathfrak{g})$$

to the adjoint representation, you get

$$\operatorname{ad} \left((\operatorname{ad} x)^k [y, z] \right) = \operatorname{ad} \left(\sum_{i=0}^k \binom{k}{i} [(\operatorname{ad} x)^i y, (\operatorname{ad} x)^{k-i} z] \right) \quad \text{in } \mathfrak{g},$$

which does not immediately yield $(\operatorname{ad} x)^k [y, z] = \sum_{i=0}^k \binom{k}{i} [(\operatorname{ad} x)^i y, (\operatorname{ad} x)^{k-i} z]$ in \mathfrak{g} unless we know that \mathfrak{g} has trivial center. Maybe you wanted to use Corollary 2.2.5 (i), but then you wouldn't need the adjoint representation. Am I understanding something wrong?

- **Proof of Corollary 4.1.3:** You write: “In particular both parts of the equality are well defined.” Why is the left hand side well-defined?
- **Lemma 4.1.4:** In part (ii), replace “(resp. locally nilpotent element)” by “(resp. locally nilpotent) element”.
- **Proof of Lemma 4.1.5:** Replace “ $t \in C$ ” by “ $t \in \mathbb{C}$ ”.
- **Proof of Lemma 4.1.5:** There are some opening brackets missing and/or some closing brackets too much in certain equations in this proof. For example: $\exp(\operatorname{ad} y)(x)$.
- **Corollary 4.1.7:** Replace “and locally nilpotent” by “are locally nilpotent”.
- **Proposition 4.2.2:** Replace “integral” by “integrable”.
- **Proposition 4.2.2:** Replace “ $\mathfrak{g}_{(i)}$ ” by “ $\mathfrak{g}_{(i)}$ ”.
- **Proposition 6.1.2:** In part (i), replace “symmetrisable generalised Cartan matrix” by “a symmetrisable generalised Cartan matrix”.
- **Proposition 6.1.2:** In part (ii), replace “symmetric” by “symmetrisable”.
- **Proposition 6.1.2:** In part (iii), replace “symmetric indecomposable” by “symmetrisable indecomposable”.

- **Proof of Proposition 6.1.2:** Replace "These solutions" by "These equations".
- **Proof of Proposition 6.1.2:** Replace "Furthermore because all the $a_{i_j, i_{j+1}}$ are non negative" by "Furthermore because all the $a_{i_j, i_{j+1}}$ and a_{i_{j+1}, i_j} are negative".
- **Proof of Proposition 6.1.2:** Replace "me may assume" by "we may assume".
- **Proposition 6.1.3:** Replace "for all sequence $i_1 \cdots i_k$ " by "for all sequences (i_1, \cdots, i_k) ".
- Your use of American English vs. British English ("realization" vs. "realisation") is inconsistent.
- **Proposition 6.2.1:** The sentence "Let A be symmetrizable and indecomposable." could be better placed at the very beginning of this proposition, not inside part (i), because it concerns all three parts (i), (ii) and (iii).
- **Proposition 6.2.1:** In part (ii), replace "resctriction" by "restriction".
- **Proof of Proposition 6.2.1:** Replace "Let $D = \text{Diag}(\epsilon_i)$ be a diagonal matrix" by "Let $D = \text{Diag}(\epsilon_i)$ be a nondegenerate diagonal matrix".
- **Proof of Proposition 6.2.1:** When you write " $(\alpha_i^\vee, \alpha_j^\vee) = \langle \alpha_i, \alpha_j^\vee \rangle \epsilon_i = \langle \alpha_j^\vee, \alpha_i \rangle \epsilon_j = (\alpha_j^\vee, \alpha_i^\vee)$ ", you should replace $\langle \alpha_j^\vee, \alpha_i \rangle$ by $\langle \alpha_j, \alpha_i^\vee \rangle$.
- **Proof of Proposition 6.2.1:** Replace " $\langle \sum_i c_i \epsilon_i \alpha_i^\vee, h' \rangle = 0$ " by " $\langle \sum_i c_i \epsilon_i \alpha_i, h' \rangle = 0$ ".
- **Proof of Proposition 6.2.1:** In your proof of $(s_i(h), s_i(h')) = (h, h')$, you should replace $\langle \alpha_i, h' \rangle \langle \alpha_i, h' \rangle$ by $\langle \alpha_i, h \rangle \langle \alpha_i, h' \rangle$. (This typo appears twice.) Also, replace $\langle \alpha_i, h' \rangle \langle h', \alpha_i \rangle$ by $\langle \alpha_i, h' \rangle \langle h, \alpha_i \rangle$.
- **Proof of Proposition 6.2.1:** Replace "Let us set $\epsilon_i = ((\alpha_i, \alpha_i^\vee)) / 2$ " by "Let us set $\epsilon_i = ((\alpha_i^\vee, \alpha_i^\vee)) / 2$ ".
- **Remark 6.2.3:** I do not see why $(\alpha_i, \alpha_i) > 0$ should hold unless we choose the ϵ_i positive in the construction of the form (\cdot, \cdot) .
- **Proof of Theorem 6.2.5:** Replace "For $\alpha = \sum_i \alpha_i$ " by "For $\alpha = \sum_i k_i \alpha_i$ ".
- **Proof of Theorem 6.2.5:** I think what you call $|\alpha|$ here is what you called $ht \alpha$ in Chapter 3.

- **Proof of Theorem 6.2.5:** You say: "this proves the invariance since the other conditions all vanish". This is not exactly the case (for example, the condition $([e_i, h], f_j) = (e_i, [h, f_j])$ does not vanish, nor does the condition $([f_j, e_i], h) = (f_j, [e_i, h])$). Still it is probably fair to say that the other conditions are similarly proven.
- **Proof of Theorem 6.2.5:** You write: "where all the elements a, b, c and d as well as the brackets $[[a, b], c], [b, [c, d]], [[a, c], b], [a, [b, c]], [a, c], [b, d], [[b, c], d]$ and $[c, [b, d]]$ are in $\mathfrak{g}(N-1)$ ". This condition is not enough (for the proof at least); you also need $[b, c]$ to lie in $\mathfrak{g}(N-1)$.
- **Proof of Theorem 6.2.5:** Replace $([[s_j, t_j], u_i], v_j)$ by $([[s_j, t_j], u_i], v_i)$.
- **Proof of Theorem 6.2.5:** You write: "Then we have to define (x, y) and (y, x) for $x \in \mathfrak{g}_N$ and $y \in \mathfrak{g}_{-N}$ ". But you define only (x, y) . This, of course, is easy to fix: just define (y, x) to mean (x, y) . As a consequence of this definition, we see by induction that the form (\cdot, \cdot) on $\mathfrak{g}(N) \times \mathfrak{g}(N)$ is symmetric.
- **Proof of Theorem 6.2.5:** You write:

"For the invariance, we still need to prove that for $x \in \mathfrak{g}_N$, for $y \in \mathfrak{g}_{-N}$ and for all h we have the relations

$$(x, [h, y]) = ([x, h], y) \quad \text{and} \quad ([x, y], h) = (x, [y, h]).$$

"

This is not enough. First of all, I think you need also to prove the relation $(h, [x, y]) = ([h, x], y)$ (but that's easy: it follows from $([x, y], h) = (x, [y, h])$ using the symmetry of (\cdot, \cdot) and the antisymmetry of $[\cdot, \cdot]$). Secondly, you also need to show that $([x, y], z) = (x, [y, z])$ holds whenever one of the vectors x, y, z lies in either \mathfrak{g}_N or \mathfrak{g}_{-N} and the other two lie in $\mathfrak{g}(N-1)$. It seems to me that the latter part is easy, but I am not sure whether it immediately follows from the definition of (x, y) as $\sum_i ([x, u_i], v_i) = \sum_j (s_j, [t_j, y])$

(at least it does not follow without some rewriting using the symmetry of (\cdot, \cdot) and the antisymmetry of $[\cdot, \cdot]$; and even then there are a lot of cases to consider).

- **Proof of Theorem 6.2.5:** In your proof of $(x, [h, y]) = ([x, h], y)$ (for $x \in \mathfrak{g}_N$, for $y \in \mathfrak{g}_{-N}$ and for all h), you should replace all \sum_i signs by \sum_j signs.