Preparing talks and mathematical texts in English: general principles and practice

Sergey V. Smirnov

Moscow State University
Department of Mathematics and Mechanics

Spring semester, 2023

Disclaimer

The author is a *non-native* English speaker with a considerable experience in communicating Mathematics in English during his lectures, talks or in his papers. Hence all material given below is a collection of the author's thoughts on how mathematical ideas should be expressed in English and is *by no means* an indisputable truth.

What are we going to discuss?

- Basic mathematical clichés
- Specifics of English mathematical terminology
- Typical examples of mistakes coming from difference in the structure of English and Russian
- How to translate mathematical text from Russian to English?
- Structure of a paper: abstract, introduction and acknowledgements
- How to prepare a talk or a public lecture: main principles and standard problems
- How to read formulas in English?
- Various tools: whiteboard handwriting VS beamer presentation

Minimalistic approach by A. Sossinsky

- Writing good mathematical texts does not require fluent English.
- Good mathematical English comes from good mathematical understanding and use of correct mathematical terminology.
- Use limited number of standard clichés as building blocks for simple and short sentences.
- Never translate from Russian!

Example or terribly awful literal translation

Sossinsky gives a real-life example of a word-for-word Russian to English translation [S]:

Назовем допустимым узлом PL-вложение $f:\mathbb{R}^1 \hookrightarrow \mathbb{R}^3$, если образ $f(\mathbb{R}^1)$ асимптотически стремится к прямой $x=y=z$ при $t \to \pm\infty$, $t \in \mathbb{R}^1$. В дальнейшем рассматриваются только допустимые узлы. Каждому (допустимому) узлу мы ставим в соответствие элемент $I(f)$ группы когомологий $H^1(E)$ пространства E функций ограниченной вариации, которое определено ниже.

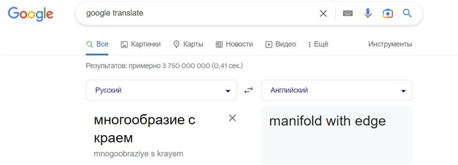
Translation looks completely inappropriate:

Let us call the admissible knot a PL-embedding $f:\mathbb{R}^1 \hookrightarrow \mathbb{R}^3$, if an image $f(\mathbb{R}^1)$ asymptotically tends to a line $x=y=z$ for $t \to \pm\infty$, $t \in \mathbb{R}^1$. In the further text considered only are the admissible knots. To every (admissible) knot we put in correspondence the element $I(f)$ of a cohomology group $H^1(E)$ of the space of the functions of bounded variation, which is defined below.

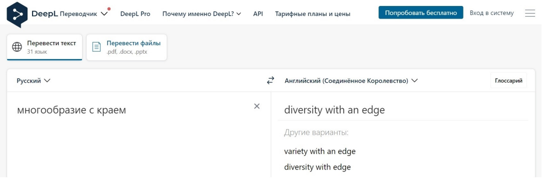
What's wrong with this translation?

- Literal translation of Russian clichés into English instead of using appropriate English clichés.
- Incorrect translation of mathematical terms.
- Order of words in English sentences is not as free as it is in Russian sentences.
- Lack of gender for nouns in English leads to ambiguity in the use of the word ‘which’.
- Lack of articles in Russian causes troubles with their use in translation.
- Due to simple case grammar in English direct translation often leads to an ugly chain of nouns linked together by the same preposition.

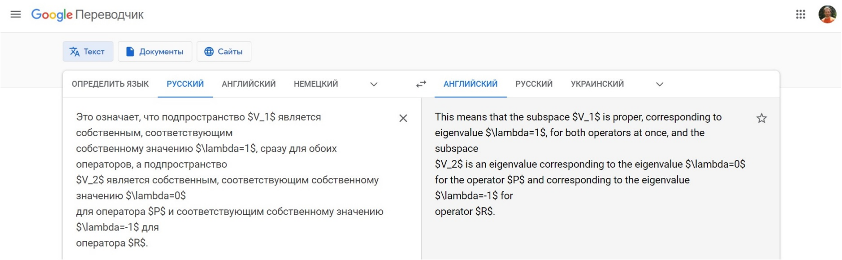
Is Google translate or DeppL an appropriate solution?



Oops. . . Manifolds can possibly have boundary, but not edge!

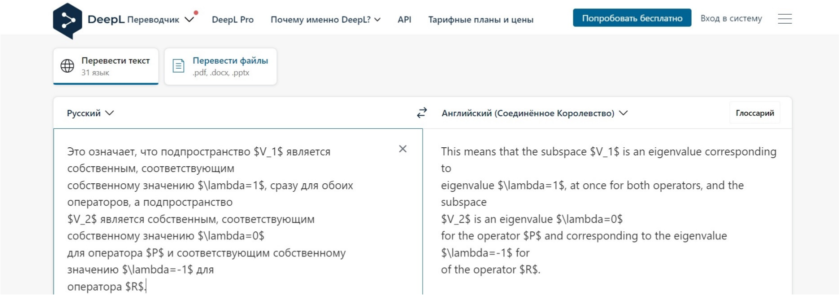


What happens when AI translates the whole sentence?



Result given by Google translate is completely inappropriate. Ok, but what if we use DeepL, which promotes itself as the best ever translating system?

What happens when AI translates the whole sentence?



We get the result that is inappropriate as well...

Well. . . But how to convert Maths from Russian to English?

That's why Sossinsky proclaims the principle "Never translate!". If you need to translate some Maths from Russian to English, the only possible way is to understand mathematically what is written and to say this in English using standard chichés and *not preserving* the structure of Russian sentences. We will practice in doing this later.

References

S А. Б. Сосинский. *Как написать математическую статью по-английски*. М., Факториал-Пресс, 2000.

M Д. В. Миллионщиков. *Методика подготовки докладов и ведения дискуссий на иностранном языке*. Мехмат, 2021.

Mi J. Milnor. *Morse Theory*. Princeton University Press, 1963.

Ha A. Hatcher. *Algebraic Topology*. Cambridge University Press, 2002.

He J. Heffron. *Linear Algebra*, 4th ed., Orthogonal Publishing, 2020.

Hu J. Humphreys. *Introduction to Lie Algebras and Representation Theory*, 3rd ed., Springer, 1980.

Ne B. O'Neill. *Elementary Differential Geometry*, 2nd ed., Academic Press, 2006.

Clichés

Cliché is a fixed text with variable entries (i.e. blank spaces to be filled in by words, expressions or formulas of appropriate type) [S].

Example

FOR ANY *var1* THERE EXISTS A *var2* SUCH THAT *var3*.

Alexander's Theorem: For any *knot* $\mathfrak{k}$ there exists a *braid* b such that $\mathfrak{k}$ *is represented by closure of b*.

Remark

Of course Alexander's Theorem can be formulated without using this cliché: *Any knot can be represented as a closure of some braid*. But clichés give a tool to express Maths in English almost automatically.

Types of variables

Sossinsky considers 5 types of variables:

- Objects: linear space X, Riccati equation, $\varepsilon > 0$, orientable manifold with non-trivial fundamental group, punctured Riemann surface...
- Openers: Then, Therefore, Hence, Without loss of generality we can assume that, This means that, Suppose that...
- References: Theorem 1, the previous lemma, formula (3), the Poincare conjecturé...
- Links: and, or, if, such that, since, when, which, whenever...
- Properties: continuity, completeness, Liouville integrability,...

Claims

Note that the whole construction is recursive: one cliché can possibly be a variable for another one:

EXAMPLE

FOR ANY $\boxed{\varepsilon > 0}$ THERE EXISTS A $\boxed{\delta > 0}$ SUCH THAT $\boxed{|x - a| < \delta \text{ IMPLIES } |f(x) - f(a)| < \varepsilon}$.

Sossinsky uses the word *claim* for formulas and combined statements that enter clichés. In our case,

$$|x - a| < \delta \text{ IMPLIES } |f(x) - f(a)| < \varepsilon$$

is a claim. Claims often appear on both sides of the links in clichés:

$\boxed{\textit{claim}}$ AND $\boxed{\textit{claim}}$, $\boxed{\textit{claim}}$ IF $\boxed{\textit{claim}}$, etc.

Examples of commonly used clichés

- *object* IS *object*
- *objects* ARE *objects*
- SUPPOSE *object*

Example (S)

Suppose $n \in \mathbb{N}$. An *n-dimensional manifold* is a pair $(M, \mathcal{A})$, where M is a topological space and $\mathcal{A}$ is an *atlas*. Here the atlas $\mathcal{A}$ is a set $\{(U_\alpha, \varphi_\alpha)\}$ such that for any α

i) $U_\alpha \subset M$ is an open set,

ii) $\varphi_\alpha : U_\alpha \to \mathbb{R}^n$ is a homeomorphism;

iii) $\cup_{\alpha \in J} = M$.

Examples of commonly used clichés, continued

- THERE EXISTS A *object* SUCH THAT *claim*
- LET *object* BE A *object*

Example (Mi)

Let P be a non-degenerate critical point for f. Then there exists a local coordinate system $(y^1, \ldots, y^n)$ in a neighbourhood U of P with $y^i = 0$ for all i and such that the identity

$$f = f(P) - (y^1)^2 - \cdots - (y^\lambda)^2 + (y^{\lambda+1})^2 + \cdots + (y^n)^2$$

holds throughout U, where λ is the index of f at P.

Examples of commonly used clichés, continued

- *object* IS CALLED *object* IF *claim*
- *claim* IF AND ONLY IF *claim*

Example (Hu)

Root system Φ is called *irreducible* if it cannot be partitioned into the union of two proper subsets such that each root in one set is orthogonal to each root in the other.

Example (Ha)

S^n has a continuous field of nonzero tangent vectors if and only if n is odd.

Introducing notation

If you need to introduce notation, you may use one of the following clichés (probably the list is not exhaustive):

- DENOTE *object* BY *symbol*
- LET US DENOTE *object* BY *symbol*
- BY *symbol* WE DENOTE *object*
- FOR *object* WE USE THE NOTATION *symbol*
- *symbol* STANDS FOR *object*

Example (Hu)

If V is a finite dimensional vector space over $\mathbb{F}$, denote by $\operatorname{End} V$ the set of all linear transformations $V \to V$.

Clichés used in definitions

- *object* IS CALLED *object* IF *claim*
- *object/symbol* IS DEFINED AS *object*
- LET US DEFINE *object* AS *object*
- WE CALL *object* IF *claim*
- *object* IS *object*

Probably this list is also not exhaustive.

Example (Hu)

A vector space L over a field $\mathbb{F}$, with an operation $L \times L \to L$, denoted $(x, y) = [x, y]$ and called the *bracket* or the *commutator* of x and y, is called a *Lie algebra* over $\mathbb{F}$ if the following axioms are satisfied:...

Starting clichés

Sometimes it is difficult to start a paragraph (a definition, a theorem). One of the following clichés may help:

- LET *object* BE A *object*
- CONSIDER *object*
- SUPPOSE *claim*
- ASSUME *claim*

Example (He)

Consider $\mathcal{P}_3 = \{a_0 + a_1x + a_2x^2 + a_3x^3 \mid a_0, a_1, a_2, a_3 \in \mathbb{R}\}$, the set of polynomials of degree three or less.

Example (Ha)

Suppose a space X is decomposed as the union of a collection of path-connected open subsets A_α, each of which contains the basepoint $x_0 \in X$.

Clichés used in theorems

- IF *claim*, THEN *claim*
- SUPPOSE THAT *claim*; THEN *object*
- LET *claim*, LET *claim*, AND LET *claim*; THEN *claim*
- *object* IS *object / property*
- IF *claim*, THEN *object* HAS THE FOLLOWING PROPERTIES: *claim*; *claim*;...
- *claim* IF AND ONLY IF *claim*
- *claim* A NECESSARY AND SUFFICIENT CONDITION FOR *object* TO BE *object / property*
- FOR *claim* IT IS NECESSARY AND SUFFICIENT THAT *claim*
- IF *claim*, THEN THE FOLLOWING CONDITIONS ARE EQUIVALENT: *claim*; *claim*;...

Theorems: examples

- Suppose X is path-connected, locally path-connected, and semilocally simply-connected. Then for every subgroup $H \subset \pi_1(X, x_0)$ there is a covering space $p : X_H \to X$ such that $p_*(\pi_1(X, \tilde{x}_0)) = H$ for a suitably chosen basepoint $\tilde{x}_0 \in X_H$. [Ha]
- A transformation t is diagonalizable if and only if there is a basis $B = \left\{ \overrightarrow{\beta}_1, \ldots, \overrightarrow{\beta}_n \right\}$ and scalars $\lambda_1, \ldots, \lambda_n$ such that $t(\overrightarrow{\beta}_i) = \lambda_i \overrightarrow{\beta}_i$ for each i. [He]
- If all elements of L are ad-nilpotent, then L is nilpotent. [Hu]

Clichés used in proofs

- IT FOLLOWS THAT *claim*
- IT FOLLOWS FROM *reference*, THAT *claim*
- IT IS EASILY VERIFIED THAT *claim*
- ONE CAN EASILY VERIFY THAT *claim*

Example (Mi)

It follows from the inverse function theorem that $v_1, \dots, v_n$ will serve as coordinate functions within some sufficiently small neighbourhood U_3 of 0. It is easily verified that f can be expressed as

$$f = \sum_{i \leqslant r} \pm v_i^2 + \sum_{i,j \leqslant r} v_i v_j H'_{ij}(v_1, \dots, v_n).$$

Clichés used in proofs, continued

- SINCE *claim*, IT FOLLOWS THAT *claim*
- *claim*, SINCE *claim*

Example (Hu)

1. Since $e_{ij}e_{kl} = \delta_{jk}e_{il}$, it follows that

$$[e_{ij}, e_{kl}] = \delta_{jk}e_{il} - \delta_{li}e_{kj}.$$

2. Since they are polynomials in x, x_s and x_n commute with each other, as well as with all endomorphisms that commute with x.
3. Evidently each of $p(T)$, $q(T)$ has zero constant term, since $p(T) \equiv 0 \pmod{T}$.

Clichés used in proofs, continued

- WE HAVE *claim*
- WE OBTAIN *claim*
- USING *reference* WE OBTAIN *claim*
- THIS MEANS THAT *claim*
- BY ASSUMPTION, *claim*

Example (Ha)

1. Denoting the latter map by $\bar{f}_t : X/A \to X/A$, we have $qf_t = \bar{f}_t q$ in the first of the two diagrams at the right.
2. So we obtain a well-defined lift $\tilde{F}$ on all of $Y \times I$.
3. This means there is a homotopy h_t of h_0 to a loop h_1 that lifts to a loop $\tilde{h}_1$ in $\tilde{X}$ based at $\tilde{x}_0$.

Clichés used in proofs, continued

- *reference* IMPLIES THAT *claim*
- THIS IMPLIES THAT *claim*
- *reference*/*property* IMPLIES *reference*/*property*
- *property* IS NECESSARY (SUFFICIENT) FOR *property*
- ACCORDING TO *reference*, *claim*

Example

1. This implies that the geodesic segment $t \mapsto \gamma(t) = \exp_P(tv)$, $0 \leqslant t \leqslant r$ is actually a minimal geodesic from p to q. [Mi]
2. According to Proposition 2.21, the inclusions $C_n(A+B) \hookrightarrow C_n(X)$ induce isomorphisms on homology groups. [Ha]
3. We saw in Section 2 that compactness implies completeness. [Ne]

Conclusions in proofs

- THEN *claim*
- THUS *claim*
- THEREFORE, *claim*
- HENCE *claim*

Example

1. Denote the A_α containing $f([s_{i-1}, s_i])$ by A_i, and let f_i be the path obtained by restricting f to $[s_{i-1}, s_i]$. Then f is the composition $f_1 \ldots f_m$ with f_i a path in A_i. [Ha]
2. Thus the only polynomials without roots in $\mathbb{C}$ are constants. [Ha]
3. We remark that two commuting semisimple endomorphisms can be simultaneously diagonalized; therefore, their sum or difference is again semisimple. [Hu]
4. Hence X generates a 1-parameter group of diffeomorphisms $\varphi_t : M \to M$. [Mi]

Ending a proof

- IT REMAINS TO PROVE THAT *claim*
- THIS CONCLUDES THE PROOF OF *reference*
- THIS PROVES *reference*
- THIS PROVES THAT *claim*
- THIS COMPLETES THE PROOF
- *reference* IS PROVED
- WE CONCLUDE THAT *claim*

Example (Ha)

1. Hence $H_i(U;R) = 0$, so z is a boundary in U and therefore in M, and we conclude that $H_i(M;R) = 0$.
2. This proves the result for homology, and the result for cohomology then follows by the universal coefficient theorem.

Proof by contradiction

- ASSUME THE CONVERSE.
- SUPPOSE ON THE CONTRARY THAT *claim* . . . THIS CONTRADICTS *claim*
- ASSUME *claim* . . . THIS CONTRADICTS *claim*
- THUS WE HAVE ARRIVED AT A CONTRADICTION.

Example (Ha)

Suppose on the contrary that $h(x) \neq x$ for all $x \in D^2$. Then we can define a map $r: D^2 \to S^1$ by letting $r(x)$ be the point of S^1 where the ray in R^2 starting at $h(x)$ and passing through x leaves D^2... But this contradicts the fact that $\pi_1(S^1)$ is nonzero.

Proof by induction

- So by induction on n we obtain a cell structure $\mathbb{C}P^n = e_0 \cup e_2 \cup \cdots \cup e_{2n}$ with cells only in even dimensions. [Ha]
- For any two interreducible matrices A and B there is some minimum number of row operations that will take one to the other. We proceed by induction on that number. [He]
- For that we use induction on the row number i. [He]
- This suggests doing the proof by induction since the dimension of the range space $\mathcal{R}(t-3)$ is less than that of the starting space. [He]
- It follows by induction on n that $\mathbb{R}P^n$ has a cell complex structure $e_0 \cup e_1 \cup \cdots \cup e_n$ with one cell e_i in each dimension $i \leqslant n$. [Ha]
- Because p is injective on $\tilde{U}_i$ and $p\tilde{F} = p\tilde{F}'$, it follows that $\tilde{F} = \tilde{F}'$ on $[t_i, t_{i+1}]$, and the induction step is finished. [Ha]

Links

- Define a real valued function f on $\mathbb{C}P^n$ by the identity

$$f(z_0 : z_1 : \cdots : z_n) = \frac{\sum_{j=0}^{n} c_j |z_j|^2}{\sum_{j=0}^{n} |z_j|^2},$$

  where $c_0, c_1, \ldots, c_n$ are distinct real constants. [Mi]
- This is obviously a group homomorphism, and *hence* an isomorphism. [Ha]
- Thus f restricts to a homeomorphism from $e^i \times e^j$ onto $e^j \times e^i$, and *therefore* $e^i e^j = e^j e^i$ in $H_*(SO(n); \mathbb{Z}_2)$. [Ha]
- For vertices $v_1, \ldots, v_n$ of a simplicial complex X, the intersection $\mathrm{st}\, v_1 \cap \cdots \cap \mathrm{st}\, v_n$ is empty *unless* $v_1, \ldots, v_n$ are the vertices of a simplex σ of X, in which case $\mathrm{st}\, v_1 \cap \cdots \cap \mathrm{st}\, v_n = \mathrm{st}\, \sigma$. [Ha]

Links, continued

- For each point $x \in X$ there is an open neighborhood U of x in X *such that* $p^{-1}(U)$ is a union of disjoint open sets each of which is mapped homeomorphically onto U by p. [Ha]
- The last step in the proof is to observe that since the $\tilde{F}$'s constructed above on sets of the form $N \times I$ are unique when restricted to each segment $\{y\} \times I$, they must agree *whenever* two such sets $N \times I$ overlap. [Ha]
- One can regard $S(X)$ as a Δ-complex model for X, *although* it is usually an extremely large object compared to X. [Ha]
- $\tilde{f}$ induces a quotient map $f : L \to L'$ for $L' = L_m(l'_1, \dots, l'_n)$ *provided* that $k_j l_j \equiv l'_j \mod m$ for each j. [Ha]

Openers

- *Similarly*, $\mathbb{C}P^\infty$ has a cell structure with one cell in each even dimension. [Ha]
- *In the same way*, we construct $\tilde{K}_{n+1}$ from $\tilde{K}_n$ for all n, and then we set $\tilde{K} = \bigcup_n \tilde{K}_n$. [Ha]
- *As above*, we have described the subspace as a collection of unrestricted linear combinations. [He]
- *In this case* Φ is not surjective. [Ha]
- *Moreover*, X_d can be chosen to have the simplest CW structure consistent with its cohomology, namely a single cell in each dimension a multiple of d. [Ha]
- *Nevertheless*, as in the real case there is a map $\mathbb{C}P^{2k-1} \to \mathbb{C}P^{2k-1}$ without fixed points. [Ha]

Openers, continued

- *However*, some things that hold for isomorphisms fail to hold for homomorphisms. [He]
- *Clearly* f^* depends only on the basepoint-preserving homotopy class of f, and it is obvious that f^* is a homomorphism if we replace K_n by ΩK_{n+1} and use the composition of loops to define the group structure. [Ha]
- *It is clear* that the covering space condition is satisfied for p. *Conversely*, every covering space of X is a graph that inherits a 2-orientation from X. [Ha]
- *Obviously* this system has the trivial solution $A_1 = 0$, $A_2 = 0$, for the case where the mass is at rest. [He]

Openers, continued

- *Furthermore*, $\bigcup_i X_i$ is closed since it is a union of closed edges and X has the weak topology. [Ha]
- *Evidently* there is some distinction being made between the two factors of $e^i \times e^j$. [Ha]
- *In addition*, because the rank of this matrix is two (we can spot this by eye or get it with Gauss's Method), any map represented by this matrix has a two-dimensional range space.
- *For example*, if L is an arbitrary finite dimensional vector space over $\mathbb{F}$, we can view L as a Lie algebra by setting $[x, y] = 0$ for all $x, y \in L$. [Hu]
- *Because* $\|t(v)\| = \|v\|$, the map t sends any v somewhere on the circle about the origin that has radius equal to the length of v. *In particular*, e_1 and e_2 map to the unit circle. [He]

Openers, continued

- *Indeed*, one can see from these examples that a covering space of $S^1 \vee S^1$ is nothing more than an efficient graphical representation of a pair of permutations of a given set. [Ha]
- *Besides*, if some vector appears more than once in an expression then we can always combine the coefficients. [He]
- *Consequently*, if matrices are row equivalent then they are also matrix equivalent since we can take Q to be the identity matrix. [He]
- *For the same reason*, any subset of D is closed, so D has the discrete topology. [Ha]
- *The same argument shows* $H^*(\mathbb{H}P^n;\mathbb{Z}) \approx \mathbb{Z}[\alpha]/(\alpha^{n+1})$ with $|\alpha| = 4$. [Ha]

Remarks

- *Note that* e^λ is contained in the "handle" H. [Mi]
- *Recall* from Chapter 0 the general definition of a homotopy as a family $\varphi_t : X \to Y$, $t \in I$, such that the associated map $\Phi : X \times I \to Y$, $\Phi(x,t) = \varphi_t(x)$, is continuous. [Ha]
- *Notice that* meridional circles in $S^1 \times S^1$ bound disks in the first solid torus, while it is longitudinal circles that bound disks in the second solid torus. [Ha]

Examples of other standard clishés

- *To avoid ambiguity*, the notation $\mathrm{ad}_L\, x$ or $\mathrm{ad}_K\, x$ will be used to indicate that x is acting in L (respectively, K). [Hu]
- To simplify the description, *we may assume without loss of generality* that the given f is an inclusion $A \hookrightarrow X$ by replacing X by the mapping cylinder of f . [Ha]
- *On the one hand*, this limit is exactly the stable operations by Proposition 4L.1 and the definition of a stable operation. *On the other hand*, the preceding theorem implies that this limit is $\mathcal{A}_p$ since it says that all elements of $H^*(K(\mathbb{Z}_p, n); \mathbb{Z}_p)$ below dimension $2n$ are uniquely expressible as sums of admissible monomials applied to i_n. [Ha]
- *In other words*, $g_1(U) \cap g_2(U) \neq \emptyset$ implies $g_1 = g_2$. [Ha]

Formulas and calculations

If a mathematical text contains formulas, then the following clichés are often being used:

- WE HAVE *formula*
- WE GET *formula*
- WE OBTAIN *formula*
- HENCE, *formula*
- THEREFORE, *formula*
- USING *reference* WE GET *formula*
- TAKING INTO ACCOUNT *reference*, WE OBTAIN *formula*
- ADDING *formula* TO BOTH SIDES OF *formula*, WE GET *formula*

Formulas and calculations, continued

- SUBSTITUTING *formula* INTO *formula* FOR *formula*, WE OBTAIN *formula*
- *formula* IMPLIES *formula*
- *formula* YIELDS *formula*
- COMBINING *formula* WITH *formula*, WE GET *formula*
- *object* SATISFIED THE FOLLOWING SYSTEM:
- REWRITE *formula* IN THE FORM *formula*

Example of a mathematical text with formulas

This fragment is taken from a paper by A. N. W. Hone:

Finally, we can construct the general solution of (1.9) by applying the discrete analogue of the method of variation of parameters to find an arbitrary element in the affine space of solutions of (3.10). In other words, we can write the solution as the sum of a particular integral plus an arbitrary linear combination of solutions of the homogeneous equation (3.13). For variation of parameters, the initial ansatz is to write the solution of (3.10) in the form

$$y_n = \sum_j f_n^{(j)} y_n^{(j)}, \tag{3.14}$$

where the index j ranges over the three lower case Roman numerals i, ii, iii, and then impose the constraints that

$$\sum_j (f_{n+1}^{(j)} - f_n^{(j)}) y_{n+1}^{(j)} = 0 = \sum_j (f_{n+1}^{(j)} - f_n^{(j)}) y_{n+2}^{(j)}, \tag{3.15}$$

which together imply that $y_{n+1} = \sum_j f_n^{(j)} y_{n+1}^{(j)}$, $y_{n+2} = \sum_j f_n^{(j)} y_{n+2}^{(j)}$. Putting all this into (3.10) gives

$$L_n(y_n) = C_n^{(3)} x_{n+3}^{-1} \sum_j (f_{n+1}^{(j)} - f_n^{(j)}) y_{n+3}^{(j)} + \sum_j f_n^{(j)} L_n(y_n^{(j)}) = C_n^{(3)} x_{n+3}^{-1} \sum_j (f_{n+1}^{(j)} - f_n^{(j)}) y_{n+3}^{(j)},$$

which must equal F_n. Combining the latter equality with the two constraints (3.15) gives a linear system for the three differences $f_{n+1}^{(j)} - f_n^{(j)}$, which is solved to yield

Theorem 3.7. *The general solution of (3.10) can be written in the form (3.14), where*

$$f_n^{(j)} = f_0^{(j)} + \sum_{k=0}^{n-1} v_k^{(j)}, \qquad j = i, ii, iii, \qquad \text{for } n \geq 0, \tag{3.16}$$

One more example of a mathematical text with formulas

This fragment is taken from a paper by C. Athorne:

Immediately we have

$$yy_1y_1' - \mathbf{x}(\partial_{u_1}h + x_1\partial_{u_2}h)\mathbf{x_1}^T - y_1\mathbf{x}h\mathbf{x}_1'^T = 0 \tag{5.21}$$

Replacing y_1y_1' by $\frac{1}{2}a'(x_1)$, taking the $x = \infty$ residue of

$$\frac{1}{2}ya'(x_1) - y_1\mathbf{x}h\mathbf{x}_1'^T = \mathbf{x}(\partial_{u_1}h + x_1\partial_{u_2}h)\mathbf{x_1}^T \tag{5.22}$$

and by elimination of y_1 as before, we find:

$$\frac{1}{2}\left((h\mathbf{x}_1^T)_4^2 - h_{44}a(x_1)\right)' = 2\sqrt{h}_{44}(\wp_{112} + 2\wp_{122}x_1 + \wp_{222}x_1^2)$$

prime denoting differentiation with respect to x_1: that is, we ignore the implicit x_1 dependence of the $\wp_{ij}$. In fact the right hand side of this equation is easily seen to be cubic in x_1 and not, as at first sight it appears, quintic. Exploiting the symmetry of h gives us,

$$\begin{aligned}\sqrt{h}_{44}(\wp_{112} + 2\wp_{122}x_1 + \wp_{222}x_1^2) \quad + \quad & \begin{vmatrix} h_{33} & h_{34} \\ h_{43} & h_{44} \end{vmatrix} x_1^3 + \frac{3}{2}\begin{vmatrix} h_{23} & h_{24} \\ h_{43} & h_{44} \end{vmatrix} x_1^2 \\ & + \left(\begin{vmatrix} h_{13} & h_{14} \\ h_{43} & h_{44} \end{vmatrix} + \frac{1}{2}\begin{vmatrix} h_{22} & h_{24} \\ h_{42} & h_{44} \end{vmatrix}\right) x_1 \\ & + \frac{1}{2}\begin{vmatrix} h_{12} & h_{14} \\ h_{42} & h_{44} \end{vmatrix} = 0\end{aligned} \tag{5.23}$$

It is straightforward to eliminate (5.13) from the above to leave a second quadratic identity,

$$2\sqrt{h}_{44}(\wp_{112} - \wp_{222}x_1^2) + \begin{vmatrix} h_{23} & h_{24} \\ h_{43} & h_{44} \end{vmatrix} x_1^2 + \begin{vmatrix} h_{22} & h_{24} \\ h_{42} & h_{44} \end{vmatrix} x_1 + \begin{vmatrix} h_{12} & h_{14} \\ h_{42} & h_{44} \end{vmatrix} = 0$$

Use of prepositions

- $f(x) \to f(a)$ as $x \to a$
- Graph of an even function is symmetric about the y-axis.
- Differentiate f with respect to x.
- Note that P is a polynomial in t.
- Recall that f is a function of x.
- Function f is continuous at the point a.
- Replace x by y in formula (1).
- Substitute y for x in (1).
- Restrict operator f to invariant subspace.
- This mapping depends on α but is independent of β.
- Let V be a vector space over $\mathbb{C}$.

Use of prepositions: OF

a function of x; a solution of equation (1); the set of all upper triangular matrices; one of the following manifolds; a subset of $\mathbb{C}^n$; closure of X; neighbourhood of x; system of equations; group of transformations; angle of rotation; consists of all elements; generalization of Theorem 1; transpose of the matrix; complex conjugate of x; the product of a and b; $Z(L)$ is the center of L; linear span of $e_1 \dots, e_n$

Use of prepositions: **TO**

x belongs to X; x is equal to y; x corresponds to y; a_n tends to 0; l_1 is parallel to l_2; assign a matrix to each linear operator; $T_P M$ is tangent to M; attach a handle to M; extend the action of G to M; relative to the topology τ; add a to both sides of equation (1); apply Theorem (1) to M; 15 is relatively prime to 32; x is a solution to equation (1)

Use of prepositions: BY

G is generated by g and h; set X is defined by condition (1); denote $T_{AB}SL(n, \mathbb{C})$ by V; function f is bounded by M; multiply both sides of equation (1) by A; solutions y_1 and y_2 differ by a constant; Lemma 1 is obtained by induction; f is approximated by $\{f_n\}$; general solution is given by (1); rotation by the angle $\frac{\pi}{2}$

Use of prepositions: **IN** and **ON**

x is contained in X; M lies in $\mathbb{R}^n$; p is a polynomial in x; A is everywhere dense in X; Y is compact in the weak topology; subspaces U and V intersect in a two-dimensional plane; convex domain in $\mathbb{R}^n$; expression (1) can be represented in the following form; curve γ in the xy-plane

points on a curve; X depends on α; topology on the set of continuous functions; graph on n vertices; our conjecture on cohomology of M

Use of prepositions: UNDER and OVER

under the action of G; under the condition; under the mapping; under the identification; under the equivalence relation; group under multiplication

module over the ring $\mathbb{Z}$, maximum over an interval; integrate ω over M; summation over $\mathbb{N}$; cone over X; bundle over B; vectors $v_1, \dots v_n$ are linearly independent over $\mathbb{R}$; as g ranges over G; remains constant over time

Use of prepositions: **FOR**, **FROM** and **WITH**

X_n is compact for all $n \in \mathbb{N}$; a basis for the subspace; excision for homotopy groups; solve equation (1) for$x$; substitute x for y in (1)

it follows from Theorem 1; f is bounded from above; subtract a from both sides of (1); projection from the point; homotopy from f to g; interval from 0 to 1; path from x_0 to x_1; map from $S^2 \setminus \{N\}$ to $\mathbb{R}^2$

coincides with; equipped with a metric; angle of $\frac{\pi}{3}$ with a plane; take the product with; M intersects with N; fibration with fiber F and base B; homology with coefficients; retraction is a map $r : X \to X$ with $r^2 = r$

Use of prepositions: **AS**, **AT** and **ALONG**

as $n \to \infty$; expressed as a commutator; as an example; regarded as subspaces of the xy-plane

at a point; at time t; at infinity; has at most two solutions; at least one element; arrive at a contradiction; at constant speed

P moves along a curve; derivation along ξ; pullback along the projection; cutting along a diagonal; attach $X \times I$ along $X \times \{1\}$; consists of a set V along with two operations

Use of prepositions: **ONTO** and **INTO**

a homeomorphism onto X; projection onto the x-axis; retraction onto a subspace

X is mapped into Y; decomposition into a product; divided into the two classes; partitioned into; taking into account; combine into a single term; this makes S^3 into a group; substitute into equation (1); transform one matrix into another; rearrange into a new order

Remark

When the preposition ONTO specifies the action of a certain mapping, it is always assumed that this mapping is surjective.

Use of ABOUT and WITH RESPECT TO

reflection about a plane; rotation about the origin

differentiation with respect to x; weak topology with respect to a subspace; group with respect to $\circ$; behaviour with respect to lifting; with respect to a basis

Warning: difference in naming theorems

There are examples when well-known mathematical results are called in English not in the same way as in Russian.

- In Calculus, Mean Value Theorem (Теорема Лагранжа) states that if f is a continuous function on $[a, b]$ and differentiable on (a, b), then there exists a point $c \in (a, b)$ such that the tangent at c is parallel to the secant line through the endpoints $(a, f(a))$ and $(b, f(b))$.
- The relation between the scalar product of two vectors and the product of their norms is usually called the *Cauchy–Schwarz inequality*, although in Russian it is "неравенство Коши–Буняковского".
- In Differential Geometry, structure equations for hypersurface and its fundamental forms are often called *Gauss–Codazzi* equations although the standard Russian name for these equations is "уравнения Гаусса–Петерсона–Кодацци".

Hypothesis VS Conjecture

In mathematics, the *Riemann hypothesis* is the conjecture that the Riemann zeta function has its zeros only at the negative even integers and complex numbers with real part $\frac{1}{2}$ (taken from Wikipedia).

- Pairs of spaces (X, A) satisfying the *hypothesis of the theorem* will be called good pairs. [Ha]
- The compactness *hypothesis* is essential, since a translation of R has $\tau = 1$ but no fixed points. [Ha]
- Use the inductive *hypothesis* to break up the k-term sum on the left. [He]
- We may *conjecture* that the wave speed v depends on the spacing d between the dominoes, the height h of each domino, and the acceleration due to gravity g. [He]

Hypothesis VS Conjecture, continued

- The word *conjecture* implies that one has some mathematical reason to suspect a claim of being true.
- The word *hypothesis* is used either when some assumptions are being recalled or when a claim is very famous and has great mathematical significance (Riemann Hypothesis, Continuum Hypothesis).
- There is one obvious exception: the *Poincaré Conjecture*.
- The word *hypothesis* could possibly sound a bit old-fashioned.

Related terminology

After obtaining mathematical results one prepares a *paper* (not an article), that is sent to a *journal* (not to a magazine). Those who attend conferences give *talks* there (not reports). Young participants are sometimes presenting a *poster* during the *poster session* (do not try to translate "стендовый доклад" literally). The most important part of almost every conference is the *conference dinner* (not the conference banquet).

Note that in English the word *paragraph* stands for the Russian "абзац" and the English analog of Russian word "параграф" is *section* (or *sub-section*).

Important abbreviations

Abbreviations *etc.*, *e.g.* and *i.e.* come from Latin

- etc. is *et cetera*=and so on. When it is used in the end of the list it should be preceded by a comma.
- e.g. is *exempli gratia*=for example. Comma should be used before and after *e.g.*
- i.e. is *id est*=that is. Comma should be used before and after *i.e.*

Example

1. More precisely, we will develop a formula for 1×1 matrices, one for 2×2 matrices, etc. [He]
2. The condition $[L, L] \neq 0$ (i.e., L is nonabelian) is imposed in order to avoid giving undue prominence to the one dimensional algebra. [Hu]

Important abbreviations, continued

- *Q.e.d.* is sometimes used instead of the black box in the end of a proof. Comes from Latin *Quod erat demonstradum*.
- *cf* stands for “compare to” and comes from Latin *conferre*. It is used without comma: (cf Theorem 1).
- *et al.* means “and others” and comes from Latin *et alii*. It is used in place of listing multiple authors and is never followed by a comma: A. Doliwa et al. *Integrable lattices and their sub-lattices...*
- *TBA* means “to be announced”. It is used when a prominent mathematician is going to give a talk somewhere but is too busy to give the title of his talk in advance.
- *RHS* (*LHS*) means “right-hand side” (“left-hand side”).

Camille Jordan in a prison cell

A matrix has *Jordan form* over $\mathbb{C}$. It consists of *Jordan blocks*, not of Jordan cells (or Jordan's cells)! Similarly, "камеры Вейля" are translated to English as *Weyl chambers*.

Example: Factors and Quotients

- *Quotient* is the result of dividing one number by another.
- *Factors* are numbers you multiply to get the original number: 3 and 4 are factors of 12.
- Both notions *quotient space* and *factor space* are used (the first one seems to be more frequent). The same with *quotient groups* and *factor groups*.

Example (Ha)

This is equivalent to saying that $\mathbb{R}P^n$ is the quotient space of a hemisphere D^n with antipodal points of ∂D^n identified.

Example: Factors and Multiples

- *Factors* are what we can multiply to get the number.
- *Multiples* are what we get after multiplying the number by an integer: $6, 9, 12, 15, \ldots$ are multiples of 3.
- The word *factor* can be used as a verb: factor out $\cos x$ to simplify the whole expression.

Example (Ha)

1. One can think of $X \wedge Y$ as a reduced version of $X \times Y$ obtained by collapsing away the parts that are not genuinely a product, the separate factors X and Y.
2. An example is the action of $\mathbb{Z}$ on S^1 in which a generator of $\mathbb{Z}$ acts by rotation through an angle α that is an irrational multiple of 2π.

Mistakes in terminology

- Coefficient of a polynomial can be *leading* (not elder).
- Natural number can be *prime*, not simple. But group that has no non-trivial normal subgroups is called *simple*.
- Topological space can be *pathwise connected* (or *arcwise connected*, or *path-connected*), but not linearly connected.
- *Compact* is not used as a noun in English: a set can be *compact*.
- Russian word *многообразие* is translated as *manifold* when we think about a topological space equipped with an atlas... But the solution set to a system of algebraic equations is called *algebraic variety*.
- *Обобщенные функции* are *distributions* in English (the term "generalized function" is not used).

Mistakes in terminology, continued

- In Topology, a handle is *attached* to a manifold (not glued).
- *Собственный вектор* is translated as *eigenvector*, not as proper vector.
- *Ступенчатый вид матрицы* is translated to English as *row-echelon form of a matrix*.
- Let p be a polynomial and let A be a square matrix. If $p(A) = 0$, then p is called an *annihilating polynomial* for A (although a marriage can be annulated, there is no word "annulating" in English).
- English analog of the word "ротор" is *curl*. Notation $\mathrm{rot}\,\xi$ looks awful because *rot* in English means "гниль".
- People study *moduli spaces* and not spaces of modules.

Mistakes in terminology, continued

- In Representation Theory, representation is called *faithful* if different elements are represented by different maps (not exact).
- *Представление Лакса* is translated to English as *Lax representation* (not presentation).
- Manifold can possibly have a *boundary*, but not an edge.
- A mapping can be *well-defined*, not correctly defined.
- Sylvester's criterion is dealing with *principal* minors of a matrix (not main minors)
- We can define the *exterior product* (or the *wedge product*) $\omega_1 \wedge \omega_2$ of two differential forms (not external product)
- The word "прямая" is translated to English as *straight line* (not just "line")

Mistakes in terminology, continued

- Minimal surface has zero *mean curvature* (not average curvature)
- In a graded Lie algebra $\bigoplus_{i=0}^{\infty} L_i$ each subspace L_i has a *grading* (not gradation)
- In the Theory of Integrable Systems, *algebraic-geometric* methods are often being used (not algebro-geometric)
- Signed minors are usually called *cofactors* in English (do not try to translate "алгебраическое дополнение" literally)
- Complex number $x - iy$ is *conjugate* to $x + iy$, but operator f^* in an Euclidean space such that $(f(u), v) = (u, f^*(v))$ for all u, v is called *adjoint*; the space of linear functions V^* on vector space V is called *dual* space.
- A word "notation" is not used in plural (even if we have "обозначения")

Other popular mistakes

- Both variants *a is equal to b* and *a equals b* are correct. But "*a* equals to *b*" (which is commonly used) is **not** correct.
- The word *let* is always followed by an infinitive: *Let M be a manifold* is correct, but "Let *M* is a manifold" is **not** correct.
- By the same reason *Let M have a singularity at the point P* is correct, but the variant using the form "has" is **not** correct.
- *Now we can prove Theorem 1* is correct, but "Now we can to prove Theorem 1" is **not** correct. The same with *must prove*.

Mistakes in naming theorems

Both *the Lagrange Theorem* and *Lagrange's Theorem* are correct. But "the Lagrange's Theorem" is not correct. Note that *a Lagrange theorem* could also be possible. Hatcher uses two ways to name a theorem within one paragraph:

The van Kampen Theorem

Suppose a space X is decomposed as the union of a collection of path-connected open subsets A_α, each of which contains the basepoint $x_0 \in X$. By the remarks in the preceding paragraph, the homomorphisms $j_\alpha : \pi_1(A_\alpha) \to \pi_1(X)$ induced by the inclusions $A_\alpha \hookrightarrow X$ extend to a homomorphism $\Phi : *_\alpha \pi_1(A_\alpha) \to \pi_1(X)$. The van Kampen theorem will say that Φ is very often surjective, but we can expect Φ to have a nontrivial kernel in general. For if $i_{\alpha\beta} : \pi_1(A_\alpha \cap A_\beta) \to \pi_1(A_\alpha)$ is the homomorphism induced by the inclusion $A_\alpha \cap A_\beta \hookrightarrow A_\alpha$ then $j_\alpha i_{\alpha\beta} = j_\beta i_{\beta\alpha}$, both these compositions being induced by the inclusion $A_\alpha \cap A_\beta \hookrightarrow X$, so the kernel of Φ contains all the elements of the form $i_{\alpha\beta}(\omega) i_{\beta\alpha}(\omega)^{-1}$ for $\omega \in \pi_1(A_\alpha \cap A_\beta)$. Van Kampen's theorem asserts that under fairly broad hypotheses this gives a full description of Φ:

Using capital letters

Words like *якобиан*, *лагранжиан*, *гамильтониан*, *риманова поверхность* are usually written with capital letter in English.

- quadratic Hamiltonian
- non-degenerate Poisson structure
- finitely generated Abelian group
- compact Riemann surface
- Sylow subgroup
- n-dimensional pseudo-Riemannian metric

Remark

Nevertheless writing *abelian group* with lower-case letters does not seem to be a mistake.

Riemann VS Reimannian

Don't ask me why

- metric on a manifold is called *Riemannian* although connected complex one-dimensional manifold is called *Riemann surface*;
- linear space equipped with a positive-definite symmetric bilinear function is called *Euclidean space*, although a topological space satisfying the T_2 separation axiom is called *Hausdorff space*;
- curvature in differential geometry is *Gaussian* although skew-symmetric bilinear mapping that satisfies the Jacobi identity is called *Poisson structure*.
- commutative group is called *Abelian* but people study *Sylow subgroups*.

Non-standard mathematical plurals

Regularly, plural form of a noun is made of singular form by attaching “-s” or “-es” in the end. But there are some exceptions that originate mostly from Latin words.

- axis (of symmetry) — axes
- basis — bases
- index — indices
- genus — genera
- formula — formulae
- phenomenon — phenomena
- spectrum — spectra
- locus — loci

Mistakes in the word order

In English the order of words cannot be changed in most cases. Typical simple mathematical sentence has the structure

Opener subject verb direct complement other complements.

Example

1. Therefore the association $g \mapsto L_g$ defines a homomorphism from G_α to the group $P(W)$ of all permutations of W. [Ha]
2. Hence a loop f in $X \times Y$ based at (x_0, y_0) is equivalent to a pair of loops g in X and h in Y based at x_0 and y_0 respectively. [Ha]

Mistakes in the word order, continued

Do not separate the subject and the verb in a sentence:

EXAMPLE (S)

"Consider now the variety V_2" does not look English. Reorganize the sentence as "Now consider the variety V_2".

Do not separate the verb and its direct complement:

EXAMPLE (S)

"The set $v_1, \dots, v_n$ generates in the complex case the required subalgebra" looks awful. Correct structure assumes the proper word order:

The set $v_1, \dots, v_n$ generates the required subalgebra in the complex case.

Use of adverbs

In English adverbs that are related to the verb are put in front of it. But with one exception: adverbs are used after the verb *is*. They help to diversify the text. In mathematical texts the following adverbs are used most frequently:

easily, obviously, clearly, immediately, always, sometimes, never, directly, entirely, naturally, sufficiently, necessarily, partially, additionally, evidently

Use of adverbs, continued

- These permutations obviously determine the covering spaces uniquely, up to isomorphism. [Ha]
- The right-hand vertical map is obviously an isomorphism. [Ha]
- We sometimes refer to Gauss's Method as Gaussian elimination. [He]
- The entries of a vector are sometimes called components. [He]
- Then from df we immediately get the partial derivatives of f, and, in fact, all its directional derivatives. [Ne]
- This will be entirely a matter of algebra. [Ha]
- Since a singular simplex always has path-connected image, $C_n(X)$ splits as the direct sum of its subgroups $C_n(X_\alpha)$. [Ha]
- Actually, $\mathbf{x}$ does not entirely cover Σ, but the missing semicircle can be covered by a patch like $\mathbf{x}$. [Ne]

Use of adverbs, continued

In these examples the word order is reverse but this does not break the rule since here adverb is related not to the verb:

- If $n > 1$ we deduce immediately from cellular cohomology that $H^*(J(S^n), \mathbb{Z})$ consists exactly of $\mathbb{Z}$'s in dimensions a multiple of n. [Ha]
- In fact, it follows immediately from (1) above that the function $dx_i(U_j)$ has the constant value δ_{ij}. [Ne]
- For example, suppose β is a unit-speed curve that lies entirely in the sphere Σ of radius a centered at the origin of $\mathbb{R}^3$. [Ne]

Use of adverbs, continued

It is also possible to use an adverb as an opener:

- Sometimes it is desirable, however, to contemplate Lie algebras abstractly. [Hu]
- Evidently L is an abelian algebra if and only if $[L, L] = 0$. [Hu]
- Clearly $\Gamma_{\mathbb{Q}}[\alpha]$ is the same as $\mathbb{Q}[\alpha]$. [Ha]
- Naturally we would like $f_t(x)$ to depend continuously on both t and x, and this will be true if we have. . . [Ha]
- Obviously φ is surjective, and its kernel consists of the words of even length. [Ha]
- Directly from the definitions, compute the simplicial cohomology groups of $S^1 \times S^1$ with $\mathbb{Z}$ and $\mathbb{Z}_2$ coefficients, using the Δ-complex structure given in 2.1. [Ha]

Use of articles: general principles

The concept of *article* has no analog in Russian and is therefore one of the most difficult things in English for Russian native speakers (and sometimes even leads to mistakes in mathematics). There are 3 articles: definite article the, indefinite article a (or an if the successive noun starts from a vowel) and no article, which will be denoted by □.

- The means "uniquely defined by the context" or "already mentioned".
- A means "some element", "certain element".
- □ is used with names, general concepts or when the object is unique in the whole world.
- □ is used when one of non-formal language quantifiers precedes the noun: *some, each, any, every, a certain, etc.*

Use of articles, simple examples

EXAMPLE (S)

1. The group G considered in the previous Section is not simple.
2. $(\mathbb{Z}_n, +)$ is a group.
3. The number $P = \max\{a_j\}$ is positive.
4. Choose a number $n \in \mathbb{N}$ such that $n > \pi$.
5. In topology □ continuity is the main notion.
6. □ Grothendieck defined the notion of □ scheme, which is now a fundamental concept in □ algebraic geometry.

Use of articles: consequences of general principles

- No article in needed before the name of a general theory (e.g., □ algebraic topology). Although when a theory is not so general, then the definite article could be used: *this is a standard theorem in the topology of smooth manifolds*.
- If an object has already been introduced, then it is named without an article: *we conclude that □ M is a compact manifold*. Note that if the name is used together with the characterization of the set to which the object belongs, then an article is required depending on the context:
  - Consider a manifold M; given a continuous function f, etc.
  - The function f that was defined earlier is continuous at the point a; consider the root space V_λ corresponding to the maximal eigenvalue λ, etc.

For more details see *axiomatic approach to articles* in [S].

Use of articles: more examples

Given an action of a group G on a space Y, we can form a space Y/G, the quotient space of □ Y in which □ each point y is identified with all its images $g(y)$ as □ g ranges over □ G. The points of □ Y/G are thus the orbits $G_y = \{g(y)\,|g \in G\}$ in □ Y, and □ Y/G is called the orbit space of the action. For example, for a normal covering space $\tilde{X} \to X$, the orbit space $\tilde{X}/G(\tilde{X})$ is just □ X. [Ha]

Use of articles: more examples, continued

First, the orthogonal complement $M^{\perp}$ is a subspace of $\mathbb{R}^n$ because it is a null space, namely the null space of the orthogonal projection map. To show that the space $\mathbb{R}^n$ is the direct sum of the two, start with □ any basis $\mathcal{B}_M = \{\mu_1, \ldots, \mu_k\}$ for □ M. Expand it to a basis for the entire space and then apply the Gram-Schmidt process to get an orthogonal basis $\mathcal{K} = \{\kappa_1, \ldots, \kappa_n\}$ for □ $\mathbb{R}^n$. This $\mathcal{K}$ is the concatenation of □ two bases: $\{\kappa_1, \ldots, \kappa_k\}$ with the same number of members, k, as □ $\mathcal{B}_M$, and □ $\mathcal{D} = \{\kappa_{k+1}, \ldots, \kappa_n\}$. The first is a basis for □ M so if we show that the second is a basis for □ $M^{\perp}$ then we will have that the entire space is the direct sum. [He]

Use of articles for plurals

- If the article a is used in the singular form, then no article is needed in the plural form.
- If the definite article is required in the singular form, then it is also required in the plural form.

Example (Ha)

1. Since τ and $p_\sharp$ commute with □ boundary maps, we have a short exact sequence of □ chain complexes, yielding the long exact sequence of □ homology groups.
2. This implies $n = 2$ since if $n > 2$ the spaces $\mathbb{R}P^{n-1}$ and S^{n-1} have different homology groups (or different fundamental groups).

Use of articles: special cases

- No article is used before attributes of a notion (e.g., radius of a circle, degree of a polynomial, dimension of a space etc.)
- An article is not needed in the reference for theorems and formulas that are numbered.
- The article a is used in the construction "There exists a unique..."

Example

1. A circle of ☐ radius r.
2. A manifold of ☐ dimension 3.
3. A function of ☐ bounded variation.
4. This proves ☐ Theorem 1.
5. It follows from (2) that ψ is a solution for ☐ equation (1).

Use of articles: special cases, continued

- Let σ be an n-simplex of K and let $\tau = f(\sigma)$, a simplex of L of □ dimension n or less. [Ha]
- Thus σ_i is a homogeneous polynomial of □ degree i. [Ha]
- In case Φ is irreducible there exists a unique highest root (relative to a fixed base Δ of Φ. [Hu]

More special cases are discussed in [S].

Example of inappropriate use of articles

How many mistakes in using articles do you see here? [S]

Let us call the admissible knot a PL*-embedding* $f : \mathbb{R}^1 \hookrightarrow \mathbb{R}^3$, *if an image* $f(\mathbb{R}^1)$ *asymptotically tends to a line* $x = y = z$ *for* $t \to \pm\infty$, $t \in \mathbb{R}^1$. *In the further text considered only are the admissible knots. To every (admissible) knot we put in correspondence the element* $I(f)$ *of a cohomology group* $H^1(E)$ *of the space of the functions of bounded variation, which is defined below.*

Use of commas

The good news is that in English there are almost now rules for using commas. The bad news is that in this situation one still has to decide whether to use or not to use them in every particular situation.

- Comma is not used before or in between the words "such that".
- Comma is *generally* not used before "that" or "than".

Example

1. We denote by M^a the set of all points $x \in M$ such that $f(x) \leqslant a$. [Mi]
2. Points that are already in X do not move. [Ha]
3. Exchange δ_{k+1} for β_i to get a new basis $\mathcal{B}_{k+1}$ with one δ more and one β fewer than the previous basis $\mathcal{B}_k$. [He]

In Russian in all these cases we *should* have put a comma.

Use of commas, continued

- We should use a comma before the last "and" in a list with multiple "and"'s (in Russian comma is *not* used in such situation).
- We should put a comma before "then" or "and" if it connects two independent clauses.

Example (Ha)

1. Most often these algebraic images are groups, but more elaborate structures such as rings, modules, and algebras also arise.
2. Such paths are called loops, and the common starting and ending point x_0 is referred to as the basepoint.
3. As explicit generators we can choose for each edge e_α of $X - T$ a loop f_α that starts at a basepoint in T, travels in T to one end of e_α, then across e_α, then back to the basepoint along a path in T.

Use of commas, continued

Use of comma in the construction "if. . . then" probably depends on the author, on the case, and on occasion:

Example

1. If a matrix has both a left inverse and a right inverse then the two are equal. [He]
2. If m or n is 1, then $G_{m,n}$ is infinite cyclic since in these cases the relation just expresses one generator as a power of the other. [Ha]
3. If $\tilde{f}_1(y) \neq \tilde{f}_2(y)$ then $\tilde{U}_1 \neq \tilde{U}_2$, hence $\tilde{U}_1$ and $\tilde{U}_2$ are disjoint and $\tilde{f}_1 \neq \tilde{f}_2$ throughout the neighborhood N. [Ha]
4. If $g_1(\tilde{U}) \cap g_2(\tilde{U}) \neq \emptyset$ for some $g_1, g_2 \in G(\tilde{X})$, then $g_1(\tilde{x}_1) = g_2(\tilde{x}_2)$ for some $\tilde{x}_1, \tilde{x}_2 \in \tilde{U}$. [Ha]
5. If there where a subspace of TM_P of dimension greater than λ on which f_{**} were negative definite then this subspace would intersect V, which is clearly impossible. [Mi]

General principles

Never translate the text word by word or even sentence by sentence. In some cases there is no one-to-one correspondence between Russian and English sentences. If you need to write a mathematical text in English, then write it in English from the beginning. Writing text in Russian and then translating it to English is much more difficult. Nevertheless sometimes we have to do the translation.

Algorithm of translation to English

1. Read a paragraph from the original Russian text.
2. Understand mathematics presented there.
3. Verbalize your understanding into English text using standard clishés (or otherwise). Of course it is better to remain as close to the original text as you can.

Do not keep the Russian word order

Example (S)

Russian phrase

Поэтому группа $\mathbb{Z}_6$ *содержит порожденную элементом* $\bar{2}$ *подгруппу* $\mathbb{Z}_3$

should be translated to English with permutation of words:

Therefore, the group $\mathbb{Z}_6$ contains the subgroup $\mathbb{Z}_3$ generated by the element $\bar{2}$.

Note that preserving the word order

Therefore the group $\mathbb{Z}_6$ *contains the generated by the element* $\bar{2}$ *subgroup* $\mathbb{Z}_3$.

is a mistake.

Use appropriate mathematical terminology

There are many cases when literal translation can possible lead to funny or even awkward situations like sending Jordan to prison.

- If you are not sure in the use of terminology try to google out the correct English analog. If you have a guess, then browse this expression in order to verify that it is really used in appropriate mathematical situation.
- If you do not know the English name of a certain method, procedure or theorem, it makes sense to look into an English textbook on the subject (there is no way to guess that the English analog of the notion "ступенчатый вид матрицы" is "row-echelon form of a matrix").

Problems with literal translation

Consider a fragment of a paper by E. Glukhov and O. Mokhov.

> Если операторы Вайнгартена невырожденны в некотором смысле, то базис в касательном пространстве подмногообразия, состоящий из общих собственных векторов операторов Вайнгартена, является голономным, т. е. подмногообразие можно параметризовать так, чтобы координатные линии являлись линиями кривизны подмногообразия (голономная сеть линий кривизны).

Literal translation with correct use of terminology and articles can possible lead to some ambiguity:

If the Weingarten operators are non-degenerate in a certain sense, then the basis in a tangent space to the submanifold that consists of common eigenvectors of the Weingarten operators is holonomic, i.e. the submanifold can be parameterized in such a way that curvature lines are coordinate lines of this submanifold (holonomic net of curvature lines).

Problems with literal translation, continued

In this example ambiguity comes from the fact that in Russian *basis* is a masculine noun and *submanifold* is a neuter noun and therefore the masculine form of the word “состоящий” indicates that the complement is related to the word basis. In English translation this information is lost. Possible solutions:

- reformulate using the “-ing” construction;
- divide the sentence into a number of short sentences.

The first option is not always possible. The second option requires thorough mathematical understanding.

Minimize the use of WHICH, WHOSE and THAT

The most probable literal translation of the phrase

Рассмотрим производную Ли векторного поля ξ, которое легко может быть вычислено явно.

is

Consider the Lie derivative of the vector field ξ, which can be easily calculated explicitly.

- The vector field or the Lie derivative can be calculated explicitly?

We can avoid this problem by not using the word WHICH:

The vector field ξ can be easily calculated explicitly. Consider its Lie derivative.

Minimize the use of WHICH, WHOSE and THAT, continued

In the Hatcher's book the word WHICH is used 650 times, i.e., once in a page in average. In case of literal translation from Russian the frequency would have been much higher. The word WHOSE is the analog of the Russian word "которого" and is even less frequent (once in 4 pages in the Hatcher's book).

Example (S)

The phrase below can be translated not using WHICH:

Всякая группа G, которая содержит свободное прямое слагаемое F, эпиморфно отображается на циклическую группу.

Suppose that the group G possesses a free direct summand F. Then there exists an epimorphism of G onto a cyclic group.

Use of WHICH and THAT

- that introduces a restrictive clause
- which introduces a nonrestrictive clause

Example

1. The van Kampen theorem gives a method for computing the fundamental groups of spaces that can be decomposed into simpler spaces whose fundamental groups are already known. [Ha]
2. More generally, the n-dimensional torus, which is the product of n circles, has fundamental group isomorphic to the product of n copies of $\mathbb{Z}$. [Ha]

In case when the word WHICH is used as an explanation, this explanation should be surrounded by commas.

Use of IT and THIS

Words IT and THIS can be used as nouns.

- IT substitutes a mathematical object.
- THIS substitutes a mathematical claim.

Example

1. A simply-connected covering space of X is therefore called a universal cover. It is unique up to isomorphism, so one is justified in calling it the universal cover. [Ha]
2. If we measure the linking of C with A and B by two integers, then the "forwards" and "backwards" cancel and both integers are zero. This reflects the fact that C is not linked with A or B individually. [Ha]

Ways to avoid multiple OF

Literal translation of the phrase "G есть группа преобразований пространства Фреше функций ограниченной вариации" is

G is the group of transformations of the space of Freéchet of functions of bounded variation.

The problem comes from the fact the case grammar in English is very simple. How to avoid use of multiple OF? [S]

- Transform nouns into adjectives:
  - *group of transformations* → *transformation group*
  - *space of Fréchet* → *Fréchet space*

G is the transformation group of the Fréchet space of functions of bounded variation.

Ways to avoid multiple OF, continued

Literal translation of the phrase “Воспользуемся (1) для построения группы преобразований пространства X” is

Let us use (1) for the construction of the group of transformations of the space X.

- Use the “-ing” ending to get rid of one OF: *Let us use (1) for constructing the transformation group of the space X.* [S]
- The phrase “The theory of differential equations of shallow waves of second order of the form...” can be improved by transforming OF to FOR:
  - *The theory of second order differential equations for shallow waves of the form...* [S]

Ways to avoid multiple OF, continued

- Transforming the phrase using 's also helps sometimes:
  - *the theorem of Cauchy* $\to$ *Cauchy's theorem*
- Literal translation of "Вычислим эйлерову характеристику множества нулей квадратичного отображения пространства функций класса C^∞" contains six OFs [S]. Radical transformations are required if none of the other options work:
  - *Suppose F is the space of C^∞ functions and Z is the zero set of the quadratic map $q : F \to \mathbb{R}$; let us compute the Euler characteristic of Z.*

DiversifyING the text

UsING of these clichés helps to diversify the text:

- USING reference/object, claim
- IDENTIFYING object WITH object, claim

Example

1. Using the less cumbersome notation $\gamma\tilde{x}_0$ for $L_\gamma(\tilde{x}_0)$, this relation can be written more concisely as $\gamma h(\tilde{x}_0) = h(\gamma\tilde{x}_0)$. [Ha]
2. Identifying e^k with $\mathbb{R}^k$, let $B_1, B_2 \subset e^k$ be the closed balls of radius 1 and 2 centered at the origin [Ha].

Use of GIVEN

- GIVEN *object*, *claim*

EXAMPLE

1. Given a path $f : I \to X$, let c be the constant path at $f(1)$, defined by $c(s) = f(1)$ for all $s \in I$. [Ha]
2. Given a loop $f : I \to X$ at the basepoint x_0, we claim there is a partition $0 = s_0 < s_1 < \cdots < s_m = 1$ of I such that each subinterval $[s_{i-1}, s_i]$ is mapped by f to a single A_α. [Ha]

Use of BOTH and NONE

The words BOTH and NONE are helpful to avoid lengthy and unnecessary constructions like SUCH THAT.

Example

1. For example, if X and Y are both closed intervals, then we are collapsing two opposite faces of a cube onto line segments so that the cube becomes a tetrahedron. [Ha]
2. If a matrix has both a left inverse and a right inverse then the two are equal. [He]
3. Projecting the four faces of the tetrahedron radially onto the sphere, we obtain a cover of S^2 by four closed sets, none of which contains a pair of antipodal points. [Ha]
4. If the echelon form matrix is nonsingular then none of its diagonal entries is zero. [He]

Whenever, Wherever...

Whenever, wherever
We're meant to be together...

Whenever, Wherever. . . , continued

Words like WHENEVER, WHEREVER, UNLESS, ELSEWHERE, OTHERWISE. . . have no exact Russian analogs. But they can be used in order to make the text more English. Not using them, Shakira would have to insert a couple of extra lines into her song.

Example

1. Notice that the definition of $\tilde{X}_\rho$ makes sense whenever we are given an action ρ of $\pi_1(X, x_0)$ on a set F regarded as a space with the discrete topology. [Ha]

2. Wherever forms appear, the exterior derivative applied twice always gives zero. [Ne]

3. For odd p we will show that $\theta_i = 0$ unless $i = 2k(p-1)$ or $2k(p-1)+1$. [Ha]

Whenever, Wherever. . . , continued

Example

4. An easy application of this, together with results proved elsewhere in the book, will be to compute $\pi_{n+1}(S^n)$ and $\pi_{n+2}(S^n)$ for all n. [Ha]

5. From this viewpoint the empty word is the common identity element of all the subgroups G_α, which are otherwise disjoint. [Ha]

6. By precomposing p with reflections of the summands of $\bigvee_k S^n$ if necessary, we can make each local degree either $+1$ or -1, whichever we wish. [Ha]

7. The Frenet frame field of β is full of information about β, whereas the natural frame field contains none at all. [Ne]

Use of EITHER and NEITHER

The words EITHER and NEITHER also do not have exact Russian analogs and help to diversify the text.

Example

1. Let C be a nontrivial cyclic subgroup of A, either finite or infinite. [Ha]
2. This requires that either X or K have some special structure. [Ha]
3. This is one step in a proof that neither of these graphs embeds in $\mathbb{R}^2$. [Ha]

Use of EITHER. . . OR and NEITHER. . . NOR

Example

1. If f is not cellular, we can either appeal to Theorem 4.8 which says that f is homotopic to a cellular map, or we can use the following argument. [Ha]

2. If neither w_j nor w_k is the barycenter b of $[v_0, \ldots, v_n]$, then these two points lie in a proper face of $[v_0, \ldots, v_n]$ and we are done by induction on n. [Ha]

Use of NOT ONLY...BUT ALSO

This is an example of one more construction that is frequently used in mathematical texts and that determines the logical structure of a sentence.

Example

One especially useful property of suspension is that not only spaces but also maps can be suspended. [Ha]

Use of tenses in mathematical texts

In general there is no need to use tenses other than *Present Simple* in mathematical papers. In most cases Russian phrases in past or in future can be translated into English phrases in the Present Simple:

- В статье [1] было доказано, что... → It is proved in [1] that...
- В параграфе 3 мы покажем, что... → In Section 3 we prove that...

Standard English construction *have been (has been)* is used in Hatcher's book only 30 times.

In most cases conditional mode also can be simplified:

- If we could prove that... it would then be possible... → If we prove that... then it is possible...

Correct spelling of names

It is important not to make a mistake in a mathematician's name in a paper or in a talk. Do you know these names?

- d'Alembert
- Clairaut
- Chasles
- Darboux
- Descartes
- Hadamard
- Lebesgue
- Maupertuis
- Pappus
- Tzitzeica

For French names use French spelling: Carnot, Gâteaux, Hermite, Lamé, Painlevé, Poincaré, Viète.

Contractions

Sometimes use of contractions in mathematical papers is considered as a *mauvais ton*.

- We've proved... $\to$ We have proved...
- It isn't proved... $\to$ It is not proved...
- We'll consider... $\to$ We will consider...

In the last case it is better to use the construction *Let us consider*...

Example

1. Thus if there is a vertex that is the endpoint of infinitely many edges, then the weak topology cannot be a metric topology. [Ha]
2. We will use Gauss's Method throughout the book. [He]

British and American spelling

There are some differences in the two main versions of English.

British	American
burnt, dreamt,. . .	burned, dreamed. . .
defence, licence. . .	defense, license. . .
generalise, emphasise. . .	generalize, emphasize. . .
factorisation, linearisation. . .	factorization, linearization. . .
diagonalisable. . .	diagonalizable. . .
metre, litre, centre. . .	meter, liter, center. . .
colour, behaviour. . .	color, behavior. . .
analogue, catalogue. . .	analog, catalog. . .

Writing an abstract

In the abstract to a paper we need

- to list the main results and/or to formulate the main purpose **very** briefly;
- to highlight the significance of the results by the reference to previous results or the commonly known mathematical facts;
- to list the methods used, if this is appropriate.

Example of an abstract by C. Athorne:

> **Abstract.** We present a new systematic method to construct Abelian functions on Jacobian varieties of plane, algebraic curves. The main tool used is a symmetric generalisation of the bilinear operator defined in the work of Baker and Hirota. We give explicit formulae for the multiple applications of the operators, use them to define infinite sequences of Abelian functions of a prescribed pole structure and deduce the key properties of these functions. We apply the theory on the two canonical curves of genus three, presenting new explicit examples of vector space bases of Abelian functions. These reveal previously unseen similarities between the theories of functions associated to curves of the same genus.

Writing an abstract, continued

- In this paper we *prove/study/construct/develop/consider/generalize*. . .
- This paper *is devoted to/deals with/is concerned with*. . .
- In this paper. . . is *defined/studied/considered/generalized*.
- Recently. . . Here we *prove/present/study/consider*. . .

Example of an abstract by A. N. W. Hone:

Abstract

We consider the growth rate of the Mahler measure in discrete dynamical systems with the Laurent property, and in cluster algebras, and compare this with other measures of growth. In particular, we formulate the conjecture that the growth rate of the logarithmic Mahler measure coincides with the algebraic entropy, which is defined in terms of degree growth. Evidence for this conjecture is provided by exact and numerical calculations of the Mahler measure for a family of Laurent polynomials generated by rank 2 cluster algebras, for a recurrence of third order related to the Markoff numbers, and for the Somos-4 recurrence. Also, for the sequence of Laurent polynomials associated with the Kronecker quiver (the cluster algebra of affine type $\tilde{A}_1$) we prove a precise formula for the leading order asymptotics of the logarithmic Mahler measure, which grows linearly.

Writing an introduction

In the introduction to a paper we need

- to describe the area and the problems related to the subject of the paper that were considered before;
- to explain the importance of the problem and the significance of the results obtained;
- to give a review of the main results related to the problem with the reference to the previous papers;
- to give the main results of the paper;
- to describe to structure of the paper.

Writing an introduction is the most difficult thing for those who are not fluent in English since Sossinsky's minimalistic approach does not work here.

Writing an introduction, continued

Example of an introduction by A. N. W. Hone:

1 Introduction

Elliptic curves are a fundamental tool in modern cryptography. The abelian group structure on an elliptic curve makes it suitable for versions of Diffie-Hellman key exchange and ElGamal key encryption, as well as providing techniques for primality testing and integer factorization, among many other applications relevant to network security [4,22,32,36]. In this paper we consider an approach to integer factorization using elliptic curves.

The elliptic curve method (ECM) due to Lenstra [24] is one of the most effective methods known for finding medium-sized prime factors of large integers, in contrast to trial division, Pollard's rho method, or the $p-1$ method, which quickly find small factors, or sieve methods, which are capable of finding very large prime factors. For factoring an integer N, the basic idea of the ECM is to pick (at random) an elliptic curve E and a point $\mathcal{P} \in E$, then compute the scalar multiple $s\mathcal{P} = \mathcal{P} + \cdots + \mathcal{P}$ (s times) in the group law of the curve, using arithmetic in the ring $\mathbb{Z}/N\mathbb{Z}$, take a rational function f on E with a pole at the point $\mathcal{O}$ corresponding to the identity in the group E, and evaluate $f(s\mathcal{P})$ for some s chosen as the largest prime power less than some fixed bound B_1, or as the product of all such prime powers. For certain choices of E and $\mathcal{P}$, this computation may lead to an attempt to divide by a non-unit in the ring, resulting in a factor of N being found.

Writing an introduction, continued

To be more precise, traditionally one starts with a Weierstrass cubic defined over $\mathbb{Q}$, which can be taken with integer coeffients as

$$y^2 = x^3 + Ax + B, \qquad A, B \in \mathbb{Z},$$

so that arithmetic mod N corresponds to working with the pseudocurve (or group scheme) $E(\mathbb{Z}/N\mathbb{Z})$ consisting of all $(x, y) \in (\mathbb{Z}/N\mathbb{Z})^2$ that satisfy the cubic equation together with $\mathcal{O}$, the point at infinity; but when N is composite the group addition $\mathcal{P}_1 + \mathcal{P}_2$ is not defined for all pairs of points $\mathcal{P}_1, \mathcal{P}_2 \in E(\mathbb{Z}/N\mathbb{Z})$. Typically f is taken to be the coordinate function x, and the method is successful if computing the scalar multiple $s\mathcal{P}$ leads to an x-coordinate with a denominator D which is not a unit in $\mathbb{Z}/N\mathbb{Z}$, such that $\gcd(D, N) > 1$ is a non-trivial factor of N. When this fortunate occurrence arises, it indicates that there is a prime factor $p|N$ for which $s\mathcal{P} = \mathcal{O}$ in the group law of the bona fide elliptic curve $E(\mathbb{F}_p)$, which is guaranteed if s is a multiple of the order $\#E(\mathbb{F}_p)$.

The original description of the ECM was based on computations with affine coordinates for a Weierstrass cubic; computing the scalar multiple $s\mathcal{P}$ is now known as "stage 1" of the ECM, and there is a further "stage 2", due to Brent, involving computing multiples $\ell s\mathcal{P}$ for small primes ℓ less than some bound $B_2 > B_1$, but here we only focus on stage 1. Improvements in efficiency can be made by using various types of projective coordinates and/or Montgomery curves (see chapter 7 in [4]). However, all of these approaches share an inconvenient feature of the addition law for $\mathcal{P}_1 + \mathcal{P}_2$ on a Weierstrass cubic, namely that the formulae for $\mathcal{P}_2 = \pm\mathcal{P}_1$ or $\mathcal{P}_2 = \mathcal{O}$ are different from the generic case.

Writing an introduction, continued

An important new development was the proposal of Bernstein and Lange [1] to consider a different model for E, namely the Edwards curve [6]

$$E_d: \qquad x^2 + y^2 = 1 + dx^2y^2 \tag{1}$$

(d is a parameter), for which the addition law

$$(x_1, y_1) + (x_2, y_2) = \left(\frac{x_1y_2 + y_1x_2}{1 + dx_1x_2y_1y_2}, \frac{y_1y_2 - x_1x_2}{1 - dx_1x_2y_1y_2}\right) \tag{2}$$

has the advantage that it is also valid for a generic pair of points $\mathcal{P}_1, \mathcal{P}_2 \in E_d$, even when $\mathcal{P}_1 = \mathcal{P}_2$, so it can be used for doubling (following [1], we have used a rescaled curve compared with the original version in [6]). The fact that the addition law (2) on E_d is unified in this sense is implicit in the classical addition formula for the Jacobi sine function (see chapter XXII in [35], or chapter 22 in [28]), for we have

$$\mathrm{sn}(z+w) = \frac{\mathrm{sn}(z)\mathrm{cd}(w) + \mathrm{cd}(z)\mathrm{sn}(w)}{1 + k^2\mathrm{sn}(z)\mathrm{sn}(w)\mathrm{cd}(z)\mathrm{cd}(w)},$$

$$\mathrm{cd}(z+w) = \frac{\mathrm{cd}(z)\mathrm{cd}(w) - \mathrm{sn}(z)\mathrm{sn}(w)}{1 - k^2\mathrm{sn}(z)\mathrm{sn}(w)\mathrm{cd}(z)\mathrm{cd}(w)},$$

Writing an introduction, continued

In this paper we explore implementations of the ECM using other models of elliptic curves, which arise in the context of QRT maps, an 18-parameter family of birational maps of the plane introduced by Quispel, Roberts and Thompson [30] to unify diverse examples of maps and functional relations appearing in dynamical systems, statistical mechanics and soliton theory. A QRT map is one of the simplest examples of a discrete integrable system, being a discrete avatar of a Hamiltonian system with one degree of freedom, with an invariant function (conserved quantity) and an invariant measure (symplectic form) [5].

Each orbit of a QRT map corresponds to a sequence of points $\mathcal{P}_0 + n\mathcal{P}$ on a curve of genus one, and in the special case $\mathcal{P}_0 = \mathcal{O}$ the orbit consists of the scalar multiples $n\mathcal{P}$, being closely related to an elliptic divisibility sequence (EDS) [34]. Thus we can implement the ECM by iterating a QRT map with a special choice of initial data, and performing all the arithmetic in $\mathbb{Z}/N\mathbb{Z}$.

A terse overview of QRT maps is provided in the next section; see [5,20,21,33] for further details. Section 3 briefly introduces Somos sequences and related EDS, showing how three particular examples of QRT maps arise in this context, namely the Somos-4 QRT map, the Somos-5 QRT map, and the Lyness map. Each of the subsequent sections 4-6 is devoted to one of these three types of QRT map, including the doubling map that sends any point $\mathcal{P}_1 \mapsto 2\mathcal{P}_1$, and a corresponding version of the ECM. In section 7 we analyse the complexity of scalar multiplication, concentrating on the Lyness case in projective coordinates, and the final section contains some conclusions.

Acknowledgements

- The author is grateful to *Prof. Smith* for valuable discussions.
- The author would like to thank *Prof. Smith* for useful comments and valuable remarks.
- The work was carried out when the author was visiting *TU Berlin*; he is grateful for hospitality and excellent working conditions.
- The author wishes to thank *the University of Glasgow* for hosting him during his visit and for additional funding.
- The work is partially supported by *RSCF grant ...*.

Acknowlwdgements, continued

Example of acknowledgements by A. N. W. Hone:

Acknowledgments

This research was supported by Fellowship EP/M004333/1 from the Engineering & Physical Sciences Research Council, UK. The author thanks the School of Mathematics and Statistics, University of New South Wales, for hosting him twice during 2017-2019 as a Visiting Professorial Fellow, with funding from the Distinguished Researcher Visitor Scheme. He is also grateful to John Roberts and Wolfgang Schief for providing additional support during his stay in Sydney, where the idea behind this work originated, and to Reinout Quispel for useful discussions and hospitality during his visit to Melbourne in May 2019.

Types of talks

- talk at a conference (15-50 minutes)
- talk at a seminar (60-90 minutes)
- popular lecture (60-90 minutes)

The aim, principles and methods used are completely different in these cases.

Different categories of conference talks

- *plenary talks* (40-50 minutes) are given by chief experts at major conferences with large number of participants;
- *invited talks* (30-40 minutes) are given either by well-known experts in the area or by young mathematicians who recently obtained significant results;
- *sectional talks* (15-30 minutes) are given at large conferences with a number of parallel sessions;
- *posters* are presented by young mathematicians and are not talks at all.

Mathematics VS Grammar

Mathematics matters, not Grammar!

- The best option is good mathematics expressed in good English.
- But good mathematics communicated in poor English is much better than poor mathematics communicated in good English: the audience consists not of grammar teachers and they are able to understand the talk even with serious grammar mistakes unless this affects the formulation of mathematical claims.

Main principles

- Even narrow specialists cannot understand the details after a short talk. Your talk is an invitation to read your paper.
- Regularly most of the people in the room are not familiar with the problem you are talking about. One of the aims is to explain the formulation of the problem to those who are not narrow specialists.
- Nice talk leaves a good impression even if one does not understand the main idea of your talk.

Before preparing a talk

- You need to understand *clearly* what exactly do you want to communicate to the audience during *this particular* talk: a talk is not a collection of all mathematics you know and not the list of all results you have ever obtained.

Before preparing a talk

- You need to understand *clearly* what exactly do you want to communicate to the audience during *this particular* talk: a talk is not a collection of all mathematics you know and not the list of all results you have ever obtained.
- You need to feel *confident* that the results you are planning to communicate are significant and interesting. Moreover, you need to *persuade* the audience of this.

Before preparing a talk

- You need to understand *clearly* what exactly do you want to communicate to the audience during *this particular* talk: a talk is not a collection of all mathematics you know and not the list of all results you have ever obtained.
- You need to feel *confident* that the results you are planning to communicate are significant and interesting. Moreover, you need to *persuade* the audience of this.
- Preparing a talk requires *much* time and effort: a theorem equipped with its proof and surrounded by a number of necessary definitions is not yet a presentation for your talk.

Before preparing a talk

- You need to understand *clearly* what exactly do you want to communicate to the audience during *this particular* talk: a talk is not a collection of all mathematics you know and not the list of all results you have ever obtained.
- You need to feel *confident* that the results you are planning to communicate are significant and interesting. Moreover, you need to *persuade* the audience of this.
- Preparing a talk requires *much* time and effort: a theorem equipped with its proof and surrounded by a number of necessary definitions is not yet a presentation for your talk.
- In general talks are more *informal* than papers.

How to prepare a conference talk?

Plan your talk very carefully:

- Choose *one* main idea (theorem, construction, etc.) that you want to communicate.

How to prepare a conference talk?

Plan your talk very carefully:

- Choose *one* main idea (theorem, construction, etc.) that you want to communicate.
- Decide what concepts or notions are really *absolutely* necessary for understanding of your results and introduce only them.

How to prepare a conference talk?

Plan your talk very carefully:

- Choose *one* main idea (theorem, construction, etc.) that you want to communicate.
- Decide what concepts or notions are really *absolutely* necessary for understanding of your results and introduce only them.
- Prepare a review of results obtained elsewhere in order to show that your results *naturally* arise within the frame of the development of the whole area.

How to prepare a conference talk?

Plan your talk very carefully:

- Choose *one* main idea (theorem, construction, etc.) that you want to communicate.
- Decide what concepts or notions are really *absolutely* necessary for understanding of your results and introduce only them.
- Prepare a review of results obtained elsewhere in order to show that your results *naturally* arise within the frame of the development of the whole area.
- Think of $1-2$ *simple* examples that clearly illustrate your results.

How to prepare a conference talk?

Plan your talk very carefully:

- Choose *one* main idea (theorem, construction, etc.) that you want to communicate.
- Decide what concepts or notions are really *absolutely* necessary for understanding of your results and introduce only them.
- Prepare a review of results obtained elsewhere in order to show that your results *naturally* arise within the frame of the development of the whole area.
- Think of $1 - 2$ *simple* examples that clearly illustrate your results.
- Think of a *good* explanation of importance and significance of your results.

Conference talks: precautions

- Never assume that “grown-up” audience *understands* and *knows* all that you know yourself.

Conference talks: precautions

- Never assume that "grown-up" audience *understands* and *knows* all that you know yourself.
- Do not go deep into the details: if the formulation of your main theorem requires many *technical* assumptions then in most cases there is no need to list them — better restrict yourself by writing "under a certain number of technical assumptions".

Conference talks: precautions

- Never assume that "grown-up" audience *understands* and *knows* all that you know yourself.
- Do not go deep into the details: if the formulation of your main theorem requires many *technical* assumptions then in most cases there is no need to list them — better restrict yourself by writing "under a certain number of technical assumptions".
- Never use a pdf of your paper instead of presentation. Prepare a *new* presentation for each talk regarding the audience and the venue.

Conference talks: precautions

- Never assume that "grown-up" audience *understands* and *knows* all that you know yourself.
- Do not go deep into the details: if the formulation of your main theorem requires many *technical* assumptions then in most cases there is no need to list them — better restrict yourself by writing "under a certain number of technical assumptions".
- Never use a pdf of your paper instead of presentation. Prepare a *new* presentation for each talk regarding the audience and the venue.
- In your review of the results obtained elsewhere do not forget to name the persons who are responsible for them.

Conference talks: precautions, continued

- Do not prepare too many slides: $6 - 7$ is the maximal number of slides for a 15 min talk.

Conference talks: precautions, continued

- Do not prepare too many slides: $6 - 7$ is the maximal number of slides for a 15 min talk.
- Do not overload your slides with information: a formula or a statement should be placed onto the slides only if it essential for your talk.

Conference talks: precautions, continued

- Do not prepare too many slides: $6 - 7$ is the maximal number of slides for a 15 min talk.
- Do not overload your slides with information: a formula or a statement should be placed onto the slides only if it essential for your talk.
- Do not try to prove your theorems.

Conference talks: precautions, continued

- Do not prepare too many slides: $6-7$ is the maximal number of slides for a 15 min talk.
- Do not overload your slides with information: a formula or a statement should be placed onto the slides only if it essential for your talk.
- Do not try to prove your theorems.
- Control the time of your talk and be prepared to skip something if you understand that there is not enough time to fully implement your plan.

Conference talks: precautions, continued

- Do not prepare too many slides: $6-7$ is the maximal number of slides for a 15 min talk.
- Do not overload your slides with information: a formula or a statement should be placed onto the slides only if it essential for your talk.
- Do not try to prove your theorems.
- Control the time of your talk and be prepared to skip something if you understand that there is not enough time to fully implement your plan.
- There is no sense in wasting time for announcing the plan in a short talk.

Conference talks: important tips

- The title is very important. Try to make is as attractive as possible.

Conference talks: important tips

- The title is very important. Try to make is as attractive as possible.
- The structure is very important. Try to organize your presentation along one general idea.

Conference talks: important tips

- The title is very important. Try to make is as attractive as possible.
- The structure is very important. Try to organize your presentation along one general idea.
- The talk should be logically developed.

Conference talks: important tips

- The title is very important. Try to make is as attractive as possible.
- The structure is very important. Try to organize your presentation along one general idea.
- The talk should be logically developed.
- If your knowledge of English allows, prepare a couple of jokes that support your presentation.

Conference talks: important tips

- The title is very important. Try to make is as attractive as possible.
- The structure is very important. Try to organize your presentation along one general idea.
- The talk should be logically developed.
- If your knowledge of English allows, prepare a couple of jokes that support your presentation.
- Using images or figures is helpful.

Conference talks: important tips

- The title is very important. Try to make is as attractive as possible.
- The structure is very important. Try to organize your presentation along one general idea.
- The talk should be logically developed.
- If your knowledge of English allows, prepare a couple of jokes that support your presentation.
- Using images or figures is helpful.
- Use large fonts: the text should be visible without additional effort from the end of the room.

Standard structure of a talk

- Introduction
- Brief historical review
- Formulation of the problem
- Examples
- Formulation of the main result
- Explanation of its significance
- Discussion (open problems, future results, etc.) — if there is enough time

REMARK

This it just a typical structure of a conference talk. The order of the above parts could be different if that's appropriate.

Conference talk: introduction

- Before you start, thank the organizers for giving the opportunity to present your results.

Conference talk: introduction

- Before you start, thank the organizers for giving the opportunity to present your results.
- If you are a beginner, introduce yourself while showing the title page: *I am Ivan Sidorov from MSU. I am the student of Prof. Taras Panov. I will talk about complex geometry of moment-angle manifolds*.

Conference talk: introduction

- Before you start, thank the organizers for giving the opportunity to present your results.
- If you are a beginner, introduce yourself while showing the title page: *I am Ivan Sidorov from MSU. I am the student of Prof. Taras Panov. I will talk about complex geometry of moment-angle manifolds*.
- In case you present joint results do not forget to mention this: *My talk is based on joint paper with Victor Buchstaber*.

During the talk

- All *key statements* should be written down.

During the talk

- All *key statements* should be written down.
- It's better to have all the *names* written because your English pronunciation could be incorrect.

During the talk

- All *key statements* should be written down.
- It's better to have all the *names* written because your English pronunciation could be incorrect.
- In case of a question from the audience during the talk generally it makes sense not to deviate from the original plan and not to start answering unless you know the answer and the answer is short. You may say "*Let's discuss this after the talk*".

Finishing a talk

In most presentations prepared for conference talks the last slide looks like this:

Thank you!

Title and abstract for a talk

Both *title* and *abstract* are very important since people often decide whether they will attend a talk after having read the title and the abstract. This is true for all kinds of talks. Sending the abstract of your talk to the organizers is not just a formality that everyone has to do. This is true for all kinds of talks. Here are some ideas from Dan Margalit (*Some Dos and Don'ts for Writing Abstracts*):

- Do give a *broad* title. A narrow title indicates a narrow talk.
- Do use one paragraph, roughly 5 – 6 sentences.
- Do start from scratch, *using words that everyone knows*. For a conference on a specialized topic, you can assume more.
- Do give context for your work — where does it fit into the *big picture*?
- Do mention your collaborators

Title and abstract for a talk, continued

- Do be *inviting*
- Do *de-symbolize* your abstract as much as possible, just for the sake of readability
- Do not cram *too many* technical words into the title or abstract
- Don't use *LaTeX* or *HTML* unless you really need to (or unless you know for sure which format is preferred).
- Don't try to explain the entire talk. Just explain the *main idea* in broad strokes

Talks at seminars

Main principles are the same as for conference talks. The only difference is that you have more time. So you have an opportunity

- to give more details;
- to formulate clearly all technical assumptions that could be omitted at a conference talk for the sake of brevity;
- to communicate proofs if that makes sense.

I, We or One

- In a talk (or in a paper) it is better to use the pronoun WE: *we prove that*, *we add A to both sides of (2)*.
- The pronoun I is normally used for personal sentiment in talks: *I am grateful to the organizers of this conference for inviting me to give a talk*. In most cases it is not used in papers.
- Third person reference ONE is sometimes used in papers: *By grouping similar terms ONE can easily understand that (3) is a direct consequence of (2)*. But the reference ONE may seem to be too formal to be used in talks.
- In the Acknowledgements section of a paper the word "the author" is usually used instead of the pronoun I.

Public lectures

- Find an attractive topic that does not require much presumed knowledge.

Public lectures

- Find an attractive topic that does not require much presumed knowledge.
- Talk about general results and not about particular ones obtained by yourself.

Public lectures

- Find an attractive topic that does not require much presumed knowledge.
- Talk about general results and not about particular ones obtained by yourself.
- Do not assume the audience knows anything beyond the standard curriculum.

Public lectures

- Find an attractive topic that does not require much presumed knowledge.
- Talk about general results and not about particular ones obtained by yourself.
- Do not assume the audience knows anything beyond the standard curriculum.
- Formulate the main goal of the lecture and generate a strategy and a plan that allow to reach this goal.

Public lectures

- Find an attractive topic that does not require much presumed knowledge.
- Talk about general results and not about particular ones obtained by yourself.
- Do not assume the audience knows anything beyond the standard curriculum.
- Formulate the main goal of the lecture and generate a strategy and a plan that allow to reach this goal.
- Talk only about the things that are necessary to reach the goal.

Public lectures

- Find an attractive topic that does not require much presumed knowledge.
- Talk about general results and not about particular ones obtained by yourself.
- Do not assume the audience knows anything beyond the standard curriculum.
- Formulate the main goal of the lecture and generate a strategy and a plan that allow to reach this goal.
- Talk only about the things that are necessary to reach the goal.
- Support your lecture by many simple examples. Examples are more important than general constructions: public lecture is not a part of a compulsory course.

Operations

multiplication ab	a times b
division $\frac{a}{b}$	a over b
power a^b	a to the power b
squares and cubes a^2 (a^3)	a squared (a cubed)
natural powers a^4	a to the fourth power
square root $\sqrt{a}$	square root of a
logarithm $\log_a x$	logarithm with the base a of x
scalar product $u \cdot v$	u dot v
wedge product $u \wedge v$	u wedge v
vector product $u \times v$	u times v
tensor product $u \otimes v$	u tensor v
transposition A^t	A transpose
adjoint operator A^*	A adjoint
complex conjugation $\bar{z}$	z conjugate

Relations

inequality $a < b$	a is less than b
inequality $a > b$	a is greater than b
inequality $a \leqslant b$	a is less or equal than b
inequality $a \geqslant b$	a is greater or equal than b
isomorphisms $U \cong V$	U is isomorphic to V

Differentiation, integration and summation

differential operator $\frac{d}{dx}$	d by dx
derivative f'	f prime
second derivative f''	f double prime
definite integral $\int_a^b f(x)dx$	integral of f over the interval from a to b
sum $\sum_{n=0}^{\infty}$ a(n)	summation from n equals zero to infinity of a of n

Various symbols

x_k	x sub k
a_0	a naught
capital A and a	A capital and a lowercase
$\tilde{a}$	a tilda
$\hat{a}$	a hat
$\bar{a}$	a bar
Greek letters π, φ, ξ...	"pai, fai, ksai..."
$(\ )$	parenthesis
$\{\ \}$	braces
$[\ ]$	brackets
$\langle\ \rangle$	angle brackets

Other typical situations

- If you need to talk about some quotient, you may call the numerator (the denominator) the *expression above (below) the fraction bar*.
- In order not to read a part of formula (3) you may say "the *expression in the right-hand side* of (3)".
- You may refer to $\tan x + \tan 2x$ in $a^{\tan x + \tan 2x}$ as *the expression in the exponent*.

Beamer presentation: pros and cons

- An opportunity to prepare a lecture in advance.

Beamer presentation: pros and cons

- An opportunity to prepare a lecture in advance.
- No need to remember formulas, etc.

Beamer presentation: pros and cons

- An opportunity to prepare a lecture in advance.
- No need to remember formulas, etc.
- Preparation takes much time.

Beamer presentation: pros and cons

- An opportunity to prepare a lecture in advance.
- No need to remember formulas, etc.
- Preparation takes much time.
- In most cases showing slides is too quick for the audience.

Online handwriting: pros and cons

"Professionals" in online lecturing say that MS OneNote is the best tool for handwriting.

- Handwriting establishes a proper speed for acquiring information.

Online handwriting: pros and cons

"Professionals" in online lecturing say that MS OneNote is the best tool for handwriting.

- Handwriting establishes a proper speed for acquiring information.
- Good handwriting creates a home atmosphere during the lecture.

Online handwriting: pros and cons

"Professionals" in online lecturing say that MS OneNote is the best tool for handwriting.

- Handwriting establishes a proper speed for acquiring information.
- Good handwriting creates a home atmosphere during the lecture.
- You need to think, write and speak English simultaneously.

Online handwriting: pros and cons

"Professionals" in online lecturing say that MS OneNote is the best tool for handwriting.

- Handwriting establishes a proper speed for acquiring information.
- Good handwriting creates a home atmosphere during the lecture.
- You need to think, write and speak English simultaneously.
- Good handwriting requires a good tablet and perfect skills in using it.

Online handwriting: pros and cons

"Professionals" in online lecturing say that MS OneNote is the best tool for handwriting.

- Handwriting establishes a proper speed for acquiring information.
- Good handwriting creates a home atmosphere during the lecture.
- You need to think, write and speak English simultaneously.
- Good handwriting requires a good tablet and perfect skills in using it.
- You need to establish your own style, colors, etc for theorems, definitions and other pieces of text and to follow these rules throughout the whole lecture.

Concluding remarks

- The list of clichés we considered is far from being complete. You are welcome to *adopt* constructions from books or papers by native English speakers. But you need to be careful in order not to misuse some construction.
- Use papers in your field for *specific* terminology.
- Develop *your own* style for beamer presentations: this helps to make a good impression for the audience.
- Probably it makes sense to write the *whole* text of your talk when you prepare the talk in English for the first time.
- Listen to the talks given by others from the point of view of Dos and Don'ts.
- Practice and develop your skills!

Further reading

For those who are interested in details:

1. P. Halmos. *How to write mathematics*. Enseign. Math. **16** (1970), 123-152.
2. D. Goss. *Some Hints on Mathematical Style*.
3. E. Swanson. *Mathematics into Type*. AMS, 1999.

(Can be googled.)