

Reviewer comment:

1) Relation to strain/stress a) A "shear strain" is calculated by summing up the ratios of displacement increments and thickness. As it appear the samples are thinning with increasing displacement, this "shear strain" is neglecting the thinning component, the derived "shear strain" cannot be used to calculate a strain ellipse (no functional relation) and overestimates strain in the sample. A more correct procedure would be one where progressive simple and pure shear are treated to occur concurrently, and reporting a unique measure of strain.

Response:

By "measured layer thickness" we mean the layer thickness measured at the same record number at which the shear displacement increments were measured, i.e. it is the instantaneous layer thickness. Consequently, we take into account the sample thinning with increasing displacement. Perhaps we need to better explain this in the manuscript.

However, we recognize that you have recently recalculated strains in ~~simple~~ general shear experiments that take the layer thinning into account (Kilian, R., & Heilbronner, R. 2017, April. Texture transition in experimentally deformed quartzite. In EGU General Assembly Conference Abstracts, Vol. 19, p. 6966). We hope you can advise us in more detail how we could most appropriately make use of such calculations in our current study.

Reply to Response:

Summing up individual steps (delta displacement/current thickness) does not produce a shear strain which accounts for the deformation imposed on the sample. This measure could be termed "apparent shear strain" but it overestimates the true (shear) strain. This apparent shear strain cannot be used to derive a strain ellipse (and accordingly strain related measures i.e. LINFs, ISA, ... any measure of strain). Any correlations of e.g. microstructural features or any other supposedly strain dependent material property should greatly benefit from a correct measure of strain. For a simple procedure to derive bulk strain under the assumption of simultaneous thinning and shearing see e.g. Fossen&Tikoff (1993).

Reviewer comment:

b) The authors report (and obtain their measurements) from the shiny surface where they assume along which most of the displacement is realised. Assuming that this surface is actually a thin layer, the strain within that layer must always be larger than the strain derived from the entire sample. Hence one might speculate that a functional relation between the determined "shear strain" and R2 is at best a rough approximation. In case, as indicated by some of the comments on the microstructure, the compaction (sample thinning?) is localised as well and not homogeneous, a bulk sample, "strain estimate" is even more likely to be unrelated to the state of deformation in the analysed layer.

Response:

We agree with your statement that the measured bulk shear strain is most probably significantly lower than the shear strain accumulated only within the thin shear surfaces. However, we expect the shear strain variations within these surfaces to be linearly correlated with the measured bulk shear strain within a sample. Thus, we believe that a correlation between shear strain and R2 is in fact possible. Nevertheless, we will relate to this relationship as a 'rough approximation'.

Additionally, brittle deformation in most fault gouges is observed to occur in a localized way, with displacement focused on a series of through-going and anastomosing shear surfaces (Craw and Upton, 2014). In polyphase gouges, graphite is quite commonly focused into these anastomosing shear surfaces (e.g. Nakamura et al., 2015; Kirilova et al., in press). So, in any natural fault that may have experienced a shear strain that we measure from its total displacement and the thickness of the deforming zone, it is also likely that individual graphite-bearing layers actually accommodated much higher shear strains, probably similar to the ones in our experiments. Therefore, correlation between bulk shear strain and R2 may actually be quite a good approximation because it is applicable to complex natural systems.

Reply to Response:

(a) It is unclear to me, on what base the relationship between bulk (sample) strain and strain within the thin, deforming layer(s) should be linear over the range of all experiments. Is the number and the thickness of these layers known and if so, are they always constant? Otherwise, to my opinion a linear relation appears to be speculative.

(b) The complexity of a naturally deformed, polymineralic rock where graphite may form layers (which may tend to localize strain, which may be arranged in an anastomosing network and which may differ in thickness and hence experience a large variation in strain) is not necessarily identical to the situation in the experiments. Simply because an experiment is a complex, its complexity may not automatically balance the complexities found in nature. Certainly there's a good chance that deformation localizes into graphite layers in nature, but it remains unclear to me why this should occur at the same (linear? as proposed above) ratio with respect to bulk strain as in the pure graphite experiments. It remains unclear to me why it should be likely that in a polymineralic rock, strain partition would between graphite and e.g. quartz in a similar way as in a pure graphite system.

Reviewer comment:

c) The authors use a mixture of surface related measures (friction coefficients, normal stress, slip rate) and volume measures (strain) which might be confusing in places, e.g. it might not be directly evident that the experiments with the 25 MPa normal stress should actually be stronger than the 5 MPa experiments.

Response:

With our mechanical data, presented in Figure 1, we are using standard measurements, i.e. double direct shear experiments on powdered material, to test the frictional properties of graphite. The vast literature (Brace & Byerlee, 1966; Byerlee 1978; Blampied et al., 1995; Marone, 1998; etc.) that put the base for fault strength is based on experiments similar to those presented in our work. Our experiments are performed at different values of normal stress (5 and 25 MPa) and at different sliding velocities (1-10-100 microns/s), and represent the deformation of a shear zone that has an initial thickness of 3 mm. In all the experiments with increasing displacement, i.e. with increasing strain, we observe an initial phase of strengthening until a peak stress is reached, then we observe strain weakening until a steady state friction is achieved. Mechanical data clearly show that: 1) Friction coefficient is lower at high normal stress and this means that the experimental fault at 25 MPa is weaker than the one at 5 MPa; 2) The sliding velocity does not influence too much the frictional properties of the experimental fault. During deformation, a typical fabric develops within the entire shear zone (e.g. Marone, 1998, fig 11) and the fault strength is strongly related to fabric evolution. In other words, frictional strength, fabric and strain are strongly connected.

Reply to Response:

The answer of the authors reflects that a mixture of surface properties and strain is indeed confusing. Friction coefficients give a ratio of shear to normal stress but are not a measure of the actual strength of a volume of rock. Expressing sample "strength" in terms of friction coefficient may be more appropriate if it is assumed that the sample is actually deforming only along a surface, however in that case the concept of strain is not existent.

Besides all, for the main and important part of the manuscript, I think the mechanical data (displacement-friction coefficient or strain-shear stress) are only relevant to make any point that the samples are deforming by a brittle, semi-brittle or ductile mechanism. The measures of strain (and in this case, estimates of strain inside the deforming layers of the sample) are those relevant to the points addressed in (2), the relation between Raman spectra and strain.

Reviewer comment:

2) R2 correlation with shear strain a) Following the number in formula 1 and as shown in Figure 3., shear strain is evaluated as a function of R2, the inverse would be logical and a fit is numerically not equivalent. Additionally, formula 1 is wrong.

Response:

We plotted R2 as a function of shear strain and fitted the curve one more time. On the attached figure (Fig. 1) you can see the resulting figure together with the new formula, which was automatically calculated by Excel. Could you please advise us if the revised version of the figure is correct?

Reviewer comment:

b) What is the physical basis that R2 and strain should have a power law relationship?

Response:

If we take the R2 value associated with fitting of a curve in Excel as a measure, then a power law function provides the best fit to the relationship between shear strain and R2. It may be possible to fit the curve using other types of equations, and we are open to your suggestions on what other functions we should try.

Reply to Response:

The way data is evaluated (R2 as a function of "strain" and not the other way around) is now correct. However, why should it have now a quadratic relationship. At high shear strain, R2 will decrease again (also visible in your new plot)! So the way to evaluate the data ($R2=f(\text{strain})$) is correct, a quadratic fit however does not really make sense to me.

Reviewer comment:

3) Origin of D1,2 bands Obtaining Raman spectra at surface ledges on (001) or generally grain boundary regions, the appearance of the D bands has been observed (e.g. Tuinstra & Koenig, 1970; Pimenta et al. 2007). For small grain sizes, the ratio of G and D bands is actually used to establish a grain size determination. The authors mention that the increasing area of D1, D2 peaks is related to a decreasing crystallinity of graphite, is it possible that the crystallinity does not change but rather with a smaller grain size more grain boundary area with a disturbed lattice is measured? The authors estimate the minimum resolution of their optical system at 0.4 μm , however, from the text it becomes not clear whether this relates to the analysed point and/or if the analysed point could actually be identified and if so, whether measurements with a large R2 come from areas with a smaller grain size? If crystallinity is defined in relation to intragranular defect density/lattice perfection, it should be verified that only intragranular measurements are evaluated. If the grain size is so small that most likely grain aggregates are measured, maybe the grain size effect could be corrected. The authors mention that they actually examine the structural disordering of graphite, so depending how this is defined, it needs to be considered separately to grain boundary effects.

Response:

We acknowledge that the measured increase in D bands can result both from (1) decrease in crystallinity and (2) spectra obtained on grain boundaries. Furthermore, more surfaces are likely to be created as a result of brittle deformation. In addition, in natural fault zones graphite commonly appears with significantly smaller grain size (e.g. <1 micron in the Alpine Fault cataclasites; Kirilova et al, in press) than in our experimental samples. Thus, the calibrated Raman thermometer could yield temperatures that significantly underestimate the peak metamorphic temperatures experienced by the host rocks. Nevertheless, our experimental study proves that the calibrated Raman thermometers are unreliable in active tectonic settings.

However, in our study: (1) We attempted to avoid grain boundaries as much as possible, and thus most (if not all) of our measurements were obtained from intragranular areas; (2) our SEM data (fig. 4b) shows that the accumulated shear surfaces are compiled of grains > 10 microns in size, which is significantly larger than the laser spot size (approximately 412 nm). Therefore, we believe that the detected increase of D bands in our experimental data in fact reflects disorder of the internal structure of graphite rather than grain size reduction.

Reply to Response:

Both, a strongly reduced grain size as well as a reduced crystallinity will certainly lead to increased D bands and an underestimation of a RSCM temperature; this is the important message in the manuscript. However, based on the presented data, I do not clearly see how the two possible effects could be separated. It was mentioned in the manuscript that the intent was to avoid grain boundaries, however, to me it is not clear how this was done. Inspecting e.g. Figures 4a, a BSE image, I find it difficult to detect grains or grain boundaries. Are grain boundaries so much more obvious in the LM images of the Raman system? Is the assumed grain size of 10-50 μm really present in the thin deforming layers which were actually measured. A nice micrograph showing these features would be quite helpful.

Reviewer comment:

4) Applicability of a strain corrected RSCM thermometer a) in real rocks, it might be difficult to estimate strain and it might not be clear in which way deformation partitions between graphite and other minerals and so I'd encourage the authors to share some thoughts on how a strain correction should be applicable in a real-rock situation.

Response:

We agree that estimating shear strain in real rocks could be a challenging task. Therefore, the suggested calibration could be used only in the case when shear strain can be undoubtedly identified. Furthermore, in the initial version of the manuscript we acknowledge the importance of sliding velocities, and thus suggest that shear strain calibration may not be sufficient for reliable temperature estimates in active tectonic settings (L272-273).

Reply to Response:

a) It is often challenging to derive a reliable strain measure from a natural rock (even when neglecting e.g. a complex deformation-uplift history). Following the earlier arguments, I think it is even more challenging to estimate the strain that an individual graphite seam inside such a rock may have experienced. Assuming that the authors would establish their "strain-R²"-relation based on an estimate of the strain in the thin deforming layer, it would still be very difficult in natural cases to "undoubtedly identify" strain accumulated in the graphite seams. I think this is a sufficiently complex problem without actually having to start thinking about any effect of deformation rates (not sliding rates) or other parameters.

Reviewer comment:

b) A correction for the thermometer might depend on the relation of temperature and strain: e.g. during exhumation of a rock deformation takes place at increasingly lower temperature. Using a simple "strain" correction, would imply that the deformation temperatures during exhumation would not have to be considered. Given that most minerals show different deformation mechanisms at different temperatures, it might be reasonable to assume that this is the case as well for graphite. So strictly speaking, a correction should only consider lattice defects introduced by the identical process which is occurring in the calibration experiments. I'd encourage the authors to comment on this complication of such an effort.

Response:

As you have mentioned during exhumation of a rock, deformation takes place at increasingly lower temperature. However, graphite structure is sensitive only to increase in temperature i.e. graphite crystallinity increases with increasing temperature. Retrograde metamorphism is known not to affect the degree of graphite crystallinity that has been previously achieved. Thus, in the suggested scenario the accumulated shear strain will be the main parameter affecting graphite structural order. On the contrary, we expect that shear strain in high temperature conditions would yield significantly different results than the ones presented in our study.

Reply to Response:

b) This comment was not about the effect of retrograde metamorphism but rather on the effect of exhumation at increasingly lower temperatures, where it is not well known when brittle behaviour of graphite starts to dominant. The presented experiments nicely demonstrate that graphite deformed at room temperature shows a reduced RSCM temperature. Despite graphite deformation at higher temperature might indeed be a different story, it is primarily very challenging to identify the part of deformation accumulated in a graphite layer which was established under brittle conditions in order to apply the "strain-correction" derived by the presented, supposedly brittle experiments. I can perfectly imagine that higher temperature deformation does a similar RSCM temperature resetting to graphite, but there might be a different functional relation compared to the one which is intended in the present study.

Reviewer comment:

c) Using any thermometer to determine peak temperatures, often the measurements yielding the highest results are considered as representing peak conditions to overcome the problem of a partial lower temperature overprint. Given that some measurements in deformed samples still yield a low R₂, it would be helpful to see where those measurements are actually determined. Are those from within grains while those with large D and D' bands contain areas with a high grain boundary density?

Response:

We attempted to obtain all our measurements within grains. Therefore, we do not expect to be observing an effect from a variable grain boundary density.

Reply to Response:

Similarly to my comments above, a convincing micrograph may enhance the manuscript. The Raman microscope seems to have a normal optical system, so a light micrograph where grains/boundaries in an analysed area would be recognisable could be a good addition to the manuscript. The authors state (L200ff) that it might have been possible that some undeformed grains might have been accidentally measured through fractures in the surface. A nice micrograph might be able to explain why fractures could accidentally be overlooked but grain boundaries be clearly identified.

I think the discussion about the applicability of the RSCM thermometer correction and the actual strain dependency should be conducted very carefully. An estimate on the strain within the deforming layer (e.g. via the thickness of the deforming layer, providing an upper estimate if it is assumed that all/most deformation is realised within those layers) in comparison to the minimum strain estimate derived from the bulk sample (if strain is correctly derived) may provide a more extensive explanation under which assumptions a "strain-correction" could realistically be applied.

All the best, Rüdiger